GANDHINAGAR UNIVERSITY

Gandhinagar University-Journal of Engineering and Technology

ISSN 2249–6157

Special Edition (15th Volume - I)

"2nd International Conference on Data Science and Intelligent Applications (ICDSIA-2023)"

Published By:

Gandhinagar University,
Khatraj - Kalol Road, Moti Bhoyan,
Tal. Kalol, Dist. Gandhinagar-382721
Phone: +91 9904405900 / 01
E-mail: director@git.org.in,
gu-jet@gandhinagaruni.ac.in
Website: www.gandhinagaruni.ac.in

GANDHINAGAR UNIVERSITY

ICDSIA
GANDHINAGAR UNIVERSITY

AICTE &GUJCOST Sponsored

2nd International Conference on Data Science and Intelligent Applications (ICDSIA-2023)

April 28-29, 2023

Organized by
Gandhinagar Institute of Technology, Gandhinagar University

Gandhinagar University

Gandhinagar University established under the Gujarat Private Universities Act, 2009 is, committed to excellence in education with high quality pedagogy and dedicated to promoting the all-faceted development of the students.

Gandhinagar University is empowering young minds by imparting quality education, research, and training. Its special emphasis on a multidisciplinary and skill-based approach as outlined in the New Education Policy NEP 2020 sets it apart as a pioneer in the field of higher education. This is University in its correct perspective, integrated ecosystem of education – by offering programmes under one roof. The University achieved many laurels of excellence in its short time of establishment and is striving hard to accomplish many more. The infrastructure of our university is huge, with equipped laboratories with highly qualified, dedicated and committed faculty.

Providing more than ten diverse disciplines and ninety courses of study along with lush green campus, sports pavilion, modern resource center is conducive to continuous learning to become all round competent individual for nation building.

About Gandhinagar University - Journal of Engineering and Technology (Formerly known as GIT- Journal of Engineering and Technology ISSN 2249-6157)

"Gandhinagar University - Journal of Engineering and Technology" is a Peer Reviewed Journal which aims to provide an opportunity to research scholars, faculties, industry experts, working professionals and UG/PG students with faculty guide as a co-author to interact and share their knowledge and experience in the field of Computer Engineering, Information Technology, Mechanical Engineering, Civil Engineering, Electrical Engineering, Electronics and Communication Engineering and Engineering Mathematics. GU-JET's main objective is to strengthen the boundaries of the journal, the reviewer database, and to motivate potential authors to provide quality research with minimum plagiarism. It includes both fundamental and advanced research-oriented ideas with an expectation that all those who wish to contribute diverse papers such as research articles, review papers.

Engineering & Technology is a field where a lot of research and study is required. **Gandhinagar University - Journal of Engineering and Technology** primarily focuses on providing a platform to the researchers of various engineering streams where a high level of teaching and learning is encouraged. Self-created theories, research; practical/theoretical work, application-based reports/studies are welcomed to add contribution to the better understanding of technical challenges.

Its scope encompasses relevant topics under the broad areas of Computer Engineering, Information Technology, Civil Engineering, Mechanical Engineering, Electronics and Communication Engineering, Electrical Engineering and Engineering Mathematics. Articles of interdisciplinary nature are welcome. The scope of GU-JET is not limited to above mentioned subfields, instead it encourages a broad spectrum of contribution in Engineering & Technology.

GU-JET Editors

Vice Chancellor's Message

In pursuance of the vision of continuing herculean task of nation building, Gandhinagar University has been making sincere efforts to promote excellence in education. There is constant endeavor to bring holistic approach through multidisciplinary education.

The perspective of Gandhinagar University is "high quality pedagogy plus choice-based credit system" with continuous performance management of the students to transform them to 'ready for industry' in accordance with the New Education Policy (NEP) 2020. Most of the programs of this University cover internship with local industry and businesses with research and innovation focus. The students at this University are determined to accept the challenges globally. In this context, I must state that it is a substantial fact that the Innovation Cell of this University has been rated by Ministry of Education, Government of India at 3.5 out of scale of 4.

This is the University that promote the students all faceted development of wellness- not only physical fitness but also psycho-social well- being with ethical mind-set to achieve specified learning outcomes as all our programs are outcome based. This young Gandhinagar University has been awarded as 'Prominent University for Futuristic Education' by Divya Bhaskar this year. The University achieved many laurels of excellence in its short time of establishment and is striving hard to accomplish many more.

I would like to congratulate the ICDSIA Team for their Commitment and superb drive in organizing this conference. Such a huge conference cannot be organized without the whole-hearted commitment and involvement of many people, be it faculty or students or sponsors. I admire their commitment and congratulate them on the success of the conference. I also profusely thank all the sponsors for their effort to encourage academic research by way of liberal sponsorships. I am very certain that this occasion will be able to provide a platform towards strengthening our relationships in knowledge sharing while at the same time providing the necessary thrust in joint research collaborations within the research society. It is my aspiration that this conference is a foundation for the growth of new ideas towards a better tomorrow. I congratulate all the researchers whose research papers are published in this Issue and express my sincere thanks to their mentors and referees. I congratulate the GU-JET committee who have worked determinedly to conceptualize and compile this publication.

Last but not least, I would also like to thank our Respected Trustees. With your continued support and interest in us, I am sure that the quest to make Gandhinagar University a top-class university is not going to be impossible to achieve.

Thank you.

Prof. Dr. K. N. Sheth
Vice Chancellor, Gandhinagar University

About ICDSIA Conference

The Gandhinagar Institute of Technology (GIT), a constituent institute of Gandhinagar University (GU) held the ***2nd International Conference on Data Science and Intelligent Applications (ICDSIA-2023) on 28th-29th April, 2023***. It was partially sponsored by All India Council for Technical Education (AICTE), New Delhi and Gujarat Council on Science and Technology (GUJCOST), Gandhinagar.

It emphasized novel, substantial, and exceptional research projects that may advance ideas and methods for artificial intelligence, communication, and data science technologies. The conference's objectives were to gather in an ample number of skilled submissions, encourage conversations about cutting-edge research among several academic trailblazers, scientists, industrial engineers, and students from around the world, and give researchers a forum. The conference, which brought together experts and participants from all around the world, also concentrated on cutting-edge concerns on an international stage. The demands and effects of Data Science and AI on society inspired the idea for the conference. It significantly and favorably affects society and the economy. AI and robots will assist people in performing their activities more effectively than they do now with the support of communicational aspects. Man, and machine working together will be unstoppable. Additionally, AI and data science dramatically lowers the likelihood of human error and uses historical data analysis to save costs.

A total of 367 research papers submission were received, out of which 77 research papers were selected and registered for presentation, which gave benefits to more than 231 researchers. The submission of research papers were contributed from 6 countries: Algeria, China, Taiwan, Morocco, USA, and Bangladesh, and 21 different states of India. Day1- 28th April 2023: First day started with registration of the participants. The Inauguration Ceremony followed by keynote address and the second half of the day had track wise paper presentations of the participants. At the end of each tracks certificates were distributed to the research groups/presenters. ***Day2- 29th April 2023:*** Online keynote address followed by track wise presentation. At the end of each tracks certificates were distributed to the research groups.

The proceedings of ICDSIA-2023 published in GU - JET (Journal of Engineering and Technology), https://www.gandhinagaruni.ac.in/publications/, ISSN 2249 – 6157 and selected papers among all the presented papers of the conference has been published in ITM WoC: https://www.itm-conferences.org/articles/itmconf/abs/2023/03/contents/contents.html, eISSN: 2271 – 2097. The ITM Web of Conference will submit this conference proceedings for the indexing of WoS (Web of Science) as per the fulfilment of the CPCI web of science indexing criteria.

ICDSIA-2023, Proceedings Editors:
Hetalkumar Shah, Rajan Patel, Nimisha Patel, Rajkumar Buyya and Indranath Chatterjee

Chief Patron(s)

Shri. Hareshbhai Rohera (President)
Ms. Vinita Rohera (Vice President)
Shri. Deepakbhai Ravani
Shri. Pravinbhai Shah
Smt. Varshaben Pandhi

Trustee & Governing Body, Gandhinagar University, India

Prof. Dr. K. N. Sheth

Vice Chancellor, Gandhinagar University, India

Patron & Convener

Prof. (Dr.) Hetalkumar N. Shah

Director, Gandhinagar University, India

Honorary Chair

Prof. (Dr.) Rajkumar Buyya

Redmond Barry Distinguished Professor and Director of CLOUDS Lab, University of Melbourne, Australia

Honorary Technical Program Committee Chair

Prof. (Dr.) Joel Rodrigues

AAIA & IEEE FELLOW

College of Computer Science and Technology, China University of Petroleum, Qingdao, China

Instituto de Telecomunicações, Portugal

Conference Coordinators

Prof. (Dr.) Rajan Patel

Dean (Engg. & Tech.), Professor, Gandhinagar University, India

Prof. (Dr.) Nimisha Patel

Professor, Gandhinagar University, India

Conference Co-coordinators

Prof. Madhuri Chopade

Assistant Professor, Gandhinagar University, India

Prof. Raxit Jani

Assistant Professor, Gandhinagar University, India

International Advisory Committee

International Keynote Speakers

Dr. Rajkumar Buyya, University of Melbourne, Australia delivered the keynote address. Title of his presentation was "Neoteric Frontiers in Cloud, Edge, and Quantum Computing". The presentation smoothly moved from understanding the basics of cloud computing with the help of examples to describing opportunities of Cloud in real life. The session was recorded and available on https://www.youtube.com/watch?v=aJl-yD-lhzo

Dr. Indranath Chatterjee, Professor, Tongmyong University, South Korea delivered his presentation online on the title "Intelligent Business: How NLP and AI are Revolutionizing the Future of Industry 4.0". Major focus of his presentation was on, how the AI & NLP are taking over the world and his examples and theories somewhere predicted that one day it might happen that machines may rule over human. So, there should be proper usage of technology in terms of applying intelligence to the machines. He briefed all the above terms with examples and also stated the scope of research available in these areas. The session had a great impact on the audience.

Technical Session-Chairs

Dr. Narendra M. Patel *BVM*	Dr. S. D. Panchal *GTU PG-SET*
Dr. Satyajit Pangaonkar *AU*	Dr. Rajan Patel *GU*
Dr. Mehul Shah *GCET*	Dr. Srinivas Padmanabhuni
Dr. Hiren B. Patel *VSITR*	Dr. Mehul Shah *GCET*
Dr. Nimisha Patel *GIT*	

List of Reviewers

Abhigya Mahajan *SMVDU*	Akanksha Soni *University Institute of Technology, RGPV*	Dr Amit Thakkar *Charusat*
Dr. Amit Barve *Parul University*	Ankita Gandhi *National Forensic Sciences University*	Ashwin Makwana *Charotar University of Science and Technology*
Dr. Bhavesh Borisaniya *Shantilal Shah Engineering College, Bhavnagar*	Dr. Dilip Motwani *Vidyalankar Institute of Technology, Mumbai*	Prof. Dinesh Kumar *DAV Institute of Engineering & Technology*
Dr Ankur Saxena *The Assam kaziranga University*	Dr Chintan Thacker *Parul Institute of Engineering and Technology PIET*	Dr Gautam Makwana *GTU - Graduate School of Engineering and Technology*
Dr Madhavi Dave *Adani Institute of Digital Technology Management*	Dr. Gaurang Raval *Nirma University*	Dr. Gaurang Lakhani *Government Polytechnic Himatnagar*
Dr. Geetali Saha *G H Patel College of Engineering and Technology, Gujarat*	Dr. Krual Vaghela *Marwadi Uiversity*	Dr. P.M.K. Prasad *GVP College of Engineering for Women, Visakhapatnam*
Dr. Parita Shah *Vidush Somany Institute of Technology and Research*	Dr. Purvi Ramanuj *LDCE*	Dr. Seema Joshi *Gujarat Technological University, Graduate School of Engineering and Technology, Ahmedabad*
Dr. Udesang Jaliya *BVM Engineering College*	Dr. Uttam Chauhan *VGEC*	Dr. Vijay Ukani *NIRMA University*
Dr. Kajal Patel *VGEC*	Dr. Kamini Solanki *Parul University*	Dr. Rajan Patel *Gandhinagar University*
Dr. Ramesh Prajapati *Shree Swaminarayan Institute of Technology, Bhat*	Dr. Gaurav Ameta *Parul Institute of Technology, Parul University (PU), Vadodara, Gujarat*	Indu Bala *Lovely Professional University, Punjab, India*

Index

Reconfigurable Ternary Computing for Convolutional code

Mayuri Chetan Rathi , Dr C.N.Deshmukh*

Research Scholor,PRMIT& R Badnera,Amravati,444607,India
**Professor,PRMIT& R Badnera,Amravati,444607,India*

Abstract

The need for trustworthy and efficient digital data transmission and storage systems has increased in recent years. The demand to handle and store digital information has grown in the business sector, the military, and the government as a direct result of the widespread availability of high-speed, large-scale data networks. This requirement must be met in order for the design of these systems to keep up with the rapid pace set by communication and computer technologies. When trying to overcome the variable degradation in real time, one of the most important factors to consider is the dependability of the broadband communication channel. Therefore, the use of convolutional codes and other channel-coding strategies is an essential component of any broadband communication system. DSL, WLAN, and 3G standards all require different configurations of convolutional coding, each of which must achieve a specific level of coding performance despite operating at a unique data rate (constraint length and code rate) Therefore, from the perspective of channel-coding techniques, hardware implementations for the development of an encoder are essential. This encoder should be able to support multiple networks using a reconfigurability approach and should be able to function across a variety of standards. Additionally, flexibility and hardware performance should both be prioritised. This calls for forward error control coding with reconfigurable logic, which provides high-speed, low-power dynamically dedicated hardware architectures under a number of speed/power performance constraints at different time intervals that can function within a variety of channel conditions. The ternary computation system has a number of benefits, the most notable of which are the availability of a high data rate, improved spectral efficiency, enhanced coverage, and lower latency. As a result, there is a growing demand for ternary-based systems that use convolutional encoders as a result of developments in technology
Keywords: Ternary computing ,Reconfigurable logic ,Ternary communication .

1. Introduction

Within the parameters of wireless communication standards The goal of reconfigurable computing is to satiate the need for both fast data rates and low power usage. In the present day, we may utilise digital mobile phones, the internet, and CD/DVD players to transmit digital data from one location to another or retrieve data from a storage device. Sometimes it's necessary to ensure that data is sent without any interruptions. In order to keep these mechanisms running smoothly. Efficient channel coding methods are utilised for this purpose. The error-correcting abilities of a system are determined by the coding rate, generator polynomial, and constraint length. In order to ensure the integrity of data transfer in modern wireless standards and satellite communication, convolutional codes are utilised as a fundamental building component. This includes the transmission of digital video, radio, and mobile communication. The message is encoded into symbols by the transmitter and sent along the transmission channel. The data is encoded in many stages using a combination of hardware and software for processing.

1.1 Digital computations operations on the digits 0, 1, and 2 (the "ternary" numbers) are carried out. creating high-density, large-scale chips High-bandwidth data transmission in parallel and serial communication, as well as enhanced manufacturing density that presents hurdles for interconnections and pin-out issues, are two of binary logic's drawbacks. Multi valued logic is an appealing and a key field of study because of the benefits it offers, chief among them decreased complexity in the design, which leads to greater manufacturing density and high-bandwidth data connection enabling parallel and serial data transfer on a smaller on-chip. As a special case of Multivalued logic, Ternary logic distinguishes between three distinct values: true, false, and intermediate.

*1.2 **Description of invention***

To begin modelling a convolutional encoder, one must first generate Ternary Random Data. After ternary random data was generated, the convolutional encoder's coding rate, constraint length, and number of shift registers were set to get the encoder's final state count. The zero-value starting point was used as the first setting for the encoder. The output's link to the shift registers is crucial to the Convolutional code encoder's design. The input bits are temporarily stored in the fixed length shift register. Each

* Dr C.N.Deshmukh
E-mail address:2014mmsoni@gmail.com

shift register's output is sent into a modulo 3 adder, which combines all of the results. In a convolutional code encoder, the most significant bit of a shift register is fed the message bits. It takes K=m+1 shifts for a message bit to go from the input to the output of an encoder with an m-stage shift register, and the encoder's memory is equal to the number of message bits. If you think of the Convolution encoder's shift register as a finite state machine, the set of delay elements inside has d potential states, where d is the number of delay elements in the register. based on the current state of affairs and the information at hand. In this case, a codeword is created from bits. These calculations are only the sums of predetermined shift register tap sequences. The history of the message bits and the current message word are stored in a register, the status of which affects each codeword. Consequently, the meanings of consecutive codewords rely on the meanings of the preceding ones. The efficiency of the convolution code is determined by the connections between the shift registers and the mod-3 adder. In this work, the convolutional encoder structure is shown for various code rates, taking into account the coupling of shift registers with output through a mod-3 adder. Convolutional codes, in contrast to block codes, don't have a fixed word length and may instead have their properties altered by changing the connections between nodes. In convolution, the multiplication action is realised by shifts and sums. When encoding, the addition operation requires a lot of dynamic energy. And so, while implementing using reconfigurable hardware, which is more difficult than shift operations. optimising adder use has paramount importance.

2. 2.Illustrations of Ternary Convolutional Encoder

Convolutional codes are often used as the coding method of choice in practical communication systems. Ones who can compress a large amount of information into a single word. The decoding of a symbol depends on both the current symbol and the one that came before it in the input stream. Different factors including coding rate R, constraint or memory length m, and free distance dfree are taken into account. We choose the best convolutional code for each application.

Figure 1 shows a convolutional encoder with a coding rate of 2/3 and a constraint length of K=3. The convolutional encoder for K=3 consists of three shift registers (Reg 1, Reg 2, and Reg 3) and two modulo-3 adders, with the generator polynomial in the upper path given by (1 0 2 1) and (1 2 0 1), and two shift registers and one modulo-3 adder, with the generator polynomial in the lower path given by (2 0 0 1). (101). The output sequence is calculated by using the message sequence and the generator polynomials.

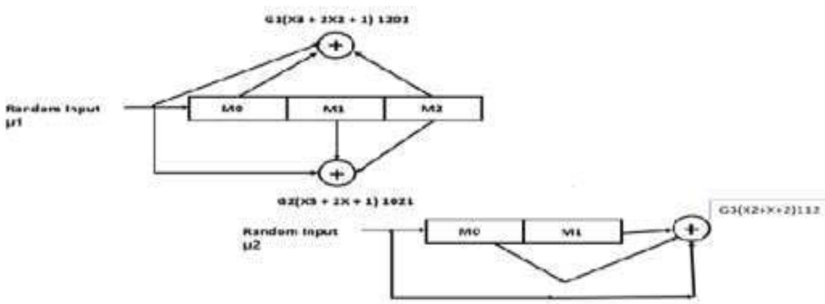


Fig. 1. Convolutional Encoder for code rate 2/3

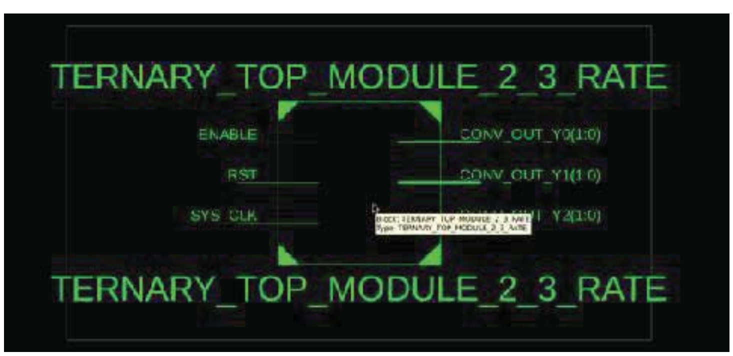


Fig. 2. RTL Schematic Of Convolutional Encoder for code rate 2/3

The generator polynomials are given by *gi(1)* = { $x1^3$+2$x1^2$+1} and *gi(2)* = { $x1^3$+2x1+1}and *gi(3)*={ $x2^2$+x2+2} using the expression of generator polynomial .

It is assumed that the message sequence m = (00101101) is completely at random, and the resulting sequence is calculated as follows. determined by the initial polynomial of the generator The calculated result is as follows: The XOR operation is accomplished by multiplying the first bit of the generator polynomial gi (1) by 1, the second bit by 1, and the third bit by 2. The first result from the higher modulo-3 adder When calculating the second output of a modulo-3 adder, the first bit of the message sequence (012012) is multiplied by 1, the second bit of the generator polynomial gi (1) is multiplied by 2, and the third bit is multiplied by 1. The output is found at the lower modulo-3 adder once the following computation is made. Upper modulo-3 adder O/P1 yields the following sequence of outputs: (012201). Upper modulo-3 adder O/P2 yields the following output sequence: (012012). Generating lower modulo-3 adder outputs with gi(3). The output of the lower modulo-3 adder is calculated by multiplying the first bit of the generating polynomial gi(3) by the first bit of the message, and then multiplying the second bit by factor 2. Sequence of results from the lower modulo-3 adder O/P2 is (00100110). V1 and V2 are updated to reflect the current conditions. m0,m1,m2 . The bits m0 and m2 hold the prior two messages. The left-to-right ', movement of the message bit. There is a current bit in m0.

The output switch first samples the input V1, and then it samples the input V2. Through the shift register, we may transfer the data from memory location m1 to memory location m2 and from memory location m0 to memory location m1. Then, we take the following input and store it in the variable m0. Based on this most recent combination of initial circumstances, m1 and m2 are generated once again. At that point, V1 and V2 are measured by the output switch. Each bit of the input message is represented by two encoded bits, V1 and V2. To encode a message using a convolutional encoder, you'll need three bits for a code word and two for the message.

For the top route, k=1 denotes the number of message bits and n=2 denotes the number of encoded output bits for a single message bit. For the lowest possible route, the number of message bits is k=1, and the number of encoded output bits for a single message bit is n=1.

3. DISCUSSION OF TERNARY CONVOLUTIONAL ENCODER

Table1 displays the stages of a 2/3 coding rate convolutional encoder. The current observed state, the encoded data generated by the convolutional encoder, and the value moved into the register to represent the next observed state are all recorded here. The following table displays the input bits (U1 and U2), the current states (m0, m1, and m2) of the encoder, and the related output. Given that there are three trit streams heading out (v1, v2, and v3) and only two leading in (u1, u2), the coding rate is 2/3. S0, S1, S2, and S3 are only three of the twenty-seven potential states. Upto S26 . The current state will transition to the next stage if the input value changes from 0 to 1 or 2.

State table for 2/3 code rate convolutional Encoder

Present state(m0m1m2)	Input u1	Next state (m0+m1+m2+)	Output(G1 G2)	Present state (m0m1)	Input u2	Next state (m0+m1+)	Output G3
001	0	000	201	00	2	20	1
010	2	201	202	10	2	21	0
110	1	111	020	11	0	01	2
020	0	002	120	21	0	02	2
120	1	112	200	22	1	12	1

4.RECONFIGURABLE TERNARY CONVOLUTIONAL ENCODER

A convolutional encoder is a finite state machine that, given an input sequence u of information blocks uj, u =... u2u1u0u1u2, performs a linear mapping, where each block uj contains k symbols, that is, uj = (u(1) j, u(2j,..., u(k)j), where u(i) The convolutional encoder receives this information stream in blocks. The encoder is given the k-ary block uj at time j, and it spits out the n-ary block vj. If uj is the "information block delivered at time j," then j is the "time index" of uj. The jth data block uj has a matching code block denoted by the letter vj. To express the number of symbols in each code block, vj, we write vj = (v(1) j, v(2) j,..., v(n) j), where v(i) j 2 R, for I = 1,..., n. To do this, it generates the output sequence v =... v2v1v0v1v2... which is treated as a single codeword by the encoder. Both u and v must begin at a certain moment and might potentially conclude at the same point. The most notable feature of a convolutional encoder is that the j-th code block vj is not only dependent on the current uj but also on the previous, say, fixed number of information blocks uj1, uj2,. This may be written as vj = ujG0 + uj1G1 +... + ujG. The inputs and outputs produced by this flexible top block are governed by the coding rate of the convolutional encoder. The logic block seen in Fig. 2 may be set up in many ways to accommodate any desired convolutional encoder code rate.

.

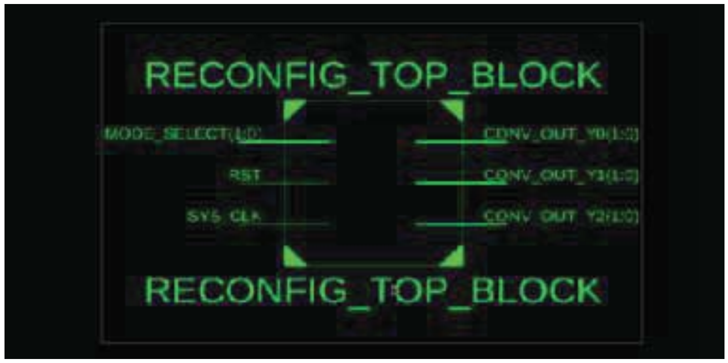


Fig. 2. Reconfigurable Convolutional Encoder for code rate ½ ,1/3,2/3

Since the data rate, bandwidth, and performance characteristics of the physical layer are defined by emerging wireless communication standards, the physical layer must dynamically adapt to accommodate the frequent changes in these standards. One of the main stumbling blocks to a dependable and high-throughput wireless network is channel faults. The physical layer's crucial block, consisting of convolutional encoders and Viterbi decoders, must be built to accommodate fluctuations in data rate and channel noise by running at high frequency on a low power supply and with the capacity to be reconfigured.

5. Simulation Result

Figure is a timing diagram for a Ternary Randomizer. In this case, ISE was used for every step of the Convolutional Encoder's design and implementation. Xilinx ISE 14.7 Simulator is used for simulation and testing, and it requires a virtual machine setup and the Verilog programming language. When creating a random ternary bit stream, it is common practise to first employ a pseudo random ternary sequence as a template. For the generation of this ternary bit stream, a linear feedback shift register might be employed (LFSR). Ternary convolution encoding receives data from the Ternary Randomzier. Convolution relies heavily on multiplication, thus we employ shift and addition to do the necessary multiplication operations. Therefore, it is crucial to optimise adder utilisation throughout the construction of reconfigurable hardware. At the outset of encoding, the convolutional encoder's linear feedback shift register is reset, so initialising all of its registers. As a generalisation, you may think of shift registers as "flip flops on steroids." These flip flops are chained together in sequential order to perform the updating and switching triggered by the clock pulse. When drawing a convolutional encoder diagram, each square represents a memory element, and the ringed numbers are the coefficients used to multiply the information in those memory cells. The circular sum represents a modulo-3 adder. The encoder's output may take on any value in the ternary number system, and the reset input makes it happen.

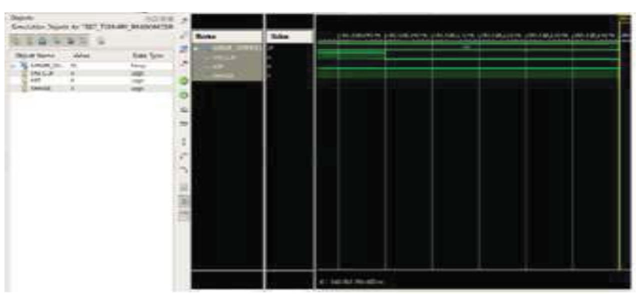


Fig. 3. Simulation result for Ternary Randomizer

This encoding scheme is a ternary convolutional encoder. Ternary randomizer outputs V1, V2, and V3 depending on clk, rst, and data d, as seen in the figure above; V1's timing wave form is 110001, V2's is 100110, and V3's is 011011; the Reset signal begins at the falling edge of the first clock signal. Convolution Encoder and Viterbi Decoder for xc3s250e-4-pq208 Spartan 3E board with constraint length 3 and coding rate 2/3 are shown here. The Hardware Description Language is used to document the design. The FPGA is independently functional as a Convolutional encoder over a range of coding rates, thanks to simulation and synthesis in Xilinx ISE 14. By setting it up in this fashion. Therefore, it is advantageous to employ ternary logic since it allows for the use of the same hardware.

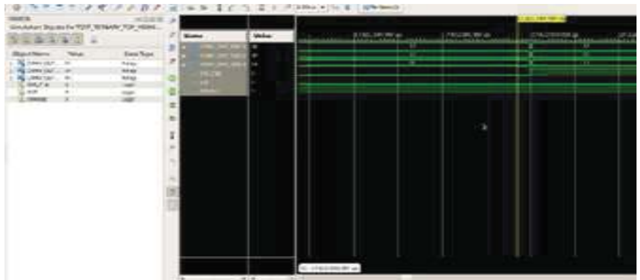


Fig. 4. Simulation result for Ternary Convolutional Encoder for code rate 2/3

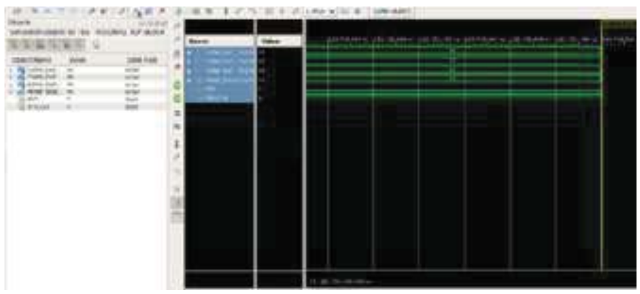


Fig. 5. Simulation result for Reconfigurable Convolutional Encoder

With regards to processing speed, wiring complexity, energy consumption, and wireless communication networks, ternary systems outperform their binary counterparts hands down. where. There are several benefits to using a ternary system instead of a binary one. Among the many benefits of the Ternary system are: As the amount of connections used to implement logic functions is decreased, a lot of space is freed up within the ship. Because more data may be sent across the same number of lines, less storage space is required for the same amount of data. This permits serial and certain serial-parallel operations to be carried out more quickly. On the other hand, the adaptive convolutional encoding technique constantly adapts to minimise computation time, iterations, and resources. Convolution codes, which are optimised for greater free distance, are preferred because of their error correcting capabilities, which are connected to polynomial strength. Typically, shift register (SR) is used to implement the convolutional encoder, which consists of delay components and modulo-3 adders (XOR gates). Convolution relies on multiplication, which is realised by shifts and adds. A lot more dynamic power is used up during encoding due to the addition operation, which is also the most difficult compared to the shift operations. Implementing with reconfigurable hardware necessitates, thus, careful consideration of adder use optimization. Reconfiguration is possible with little effort by simply altering a polynomial with the same constraint length and coding rate in a hardware implementation using ROM. Overcoming the dynamic decline in dependability of a broadband communication connection in real time is a crucial challenge. Consequently, channel-coding methods, such as convolutional codes, are crucial components of any broadband communication system.

A common place reconfigurable convolution encoder consists mostly of a shift register with N stages and v modulo-3 adders. A wide variety of modulo-3 adders may be constructed using XOR gates alone. A convolution encoder is defined by the parameters (N, k, v). In this notation, v represents the total number of encoder outputs, k represents the total number of encoder inputs, N1 represents the total number of memory components (Flip- Flops), and N represents the total length of the constraints. The k/v notation represents the encoding speed of the corresponding encoder. Into the encoder's N-stage shift register goes the source data sequence represented by Input(n) = (input(0), input(1), input(2),...). Using modulo 3 adders, we transform each input into a distinct series of output values (V1, V2). Because of the importance of the constraint length to convolution encoder performance, a better understanding of it is necessary. Keeping the constraint length and code rate the same while changing the generator polynomial via architectural reconfigurability might streamline the design.

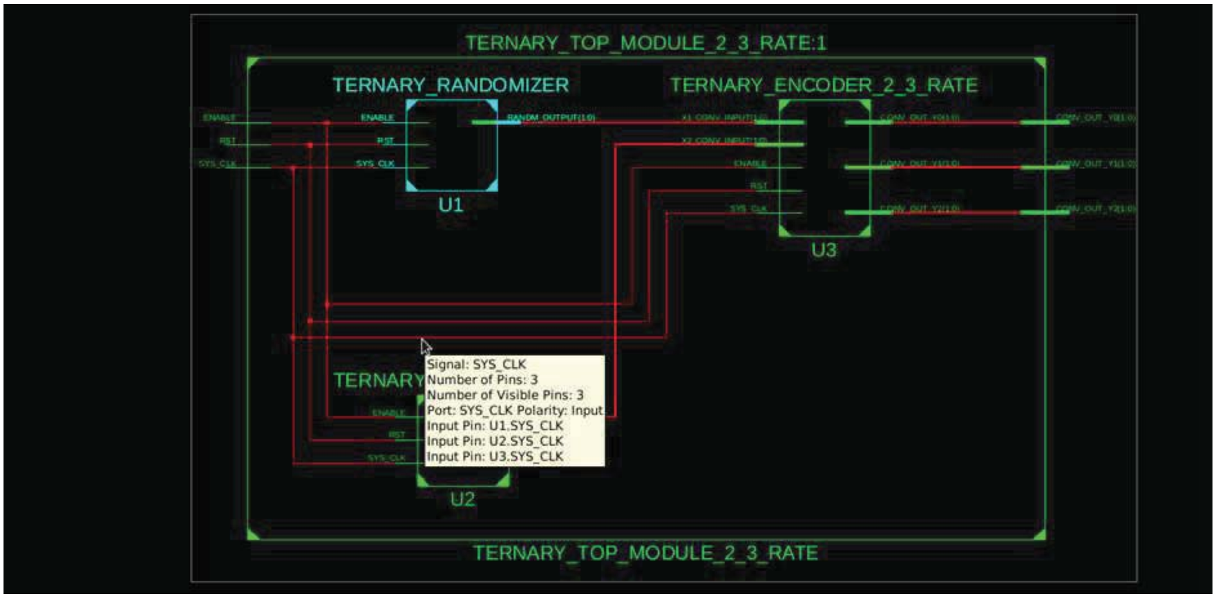


Fig. 7. Ternary Top module for1/2, 1/3 and 2/3 code rate With Reconfigurable Architecture

6. CONCLUSION

The capacity to reconfigure is crucial for satisfying the rising physical layer requirements of wireless communications, which call for faster data rates with lower power consumption. For long constraint lengths, the design and implementation of such structures presents a significant challenge. However, applications needing codes with enormous constraint lengths, such as 3G, generally render these designs impractical due to the hardware cost of reconfigurable logic of the convolutional codes increasing exponentially with the constraint length. Today, wireless connections are an integral part of most communication networks. Data services, such as websites and multimedia files, often use such connections to send and receive data, but they have varying needs in terms of bandwidth, latency, and reliability. The increasing demand for these offerings motivates researchers to look for ways to improve efficiency without increasing complexity. One way to do this is by use ternary symbols. Because of their potential to reduce the amount of hardware needed, systems based on ternary logic are viewed as a promising future technology. New requirements for channel coding arise with the advent of broadband Internet access. Using parallel and pipeline characteristics of the hardware resources, the architecture is built on programmable logic arrays. Through the use of HDL language, simulation, synthesis, and implementation are performed on the existing algorithm with the help of EDA tools based on FPGAs. The system as a whole now operates more efficiently in terms of both space and time.

Acknowledgements

Author 1: Mayuri Rathi is currently working as an Assistant Professor in Modern Education Society College of Engineering, Wadia campus, Pune. She is MTech in VLSI design and having teaching experience of 8 years. She is a pursuing PhD under the guidance of Dr C.N. Deshmukh. Her area of interest is Radio frequency communication, VLSI Design, Antenna Design.
Author 2: Dr C.N.Deshmukh is currently working as a Associate Professor in Prof Ram Megha Institute Of Technology and Research ,Badnera ,Amravati .He is having teaching experience of 30 years. He has guided many students for their research work.currently five research scholars are pursuing their PhD under his guidance. His area of interest are wireless communication, Digital Image Processing ,Ternary Computing .

References

1. S. Madumitha and G. Suchitra, “performance analysis of a convolutional encoder using digital modulation technique,” ictact journal on communication technology, march 2020, volume: 11, issue: 01
2. Fazal Noorbasha, G.Jhansi, K.Deepthi, K Hari Kishore, “ASIC Implementation of Convolution Encoder and Viterbi Decoder Based Cryptography System” (IJITEE) ISSN: 2278-3075, Volume-8 Issue-6S, April 2019.

3. Zhang Taotao, Zhang JingKe, Zhou Zhiwen, Yang Zhifei, Liu Wanhong, "FPGA-Based Large Constraint Length Convolution Code Encoder Verification, ICSP 2019
4. Dr. S Aruna, Mr. A.V. Adiseshu, Dr.K.Srinivasa Naik, "Design of Viterbi Decoder for Speech to Text Conversion Application using ACS Architecture,IOSR journal of VLSI and signal processing Jan 21, 2019
5. Gabriele Meoni, "Design Optimization for High Throughput Recursive Systematic Convolutional Encoders" ICSTCC 2018
6. Prof. Vijaya Bharathi M, Sneha.H., Mahesh.M., Shwetha.N., Sowmya.S, "Forward Error Correcting Implementation Using Convolutional Encoders and Viterbi decoding" International Journal of Electrical, Electronics and Computer Systems (IJEECS) Volume -6, Issue-3, 2018
7. Mr.J.Anuj Sai, Mr.P.Kiran Kumar, "FPGA Design and Implementation of Convolution Encoder and Viterbi Decoder IJSER,Vol 9,issue 3, March 2018
8. Girish D. Korde1, Sanjay L. Haridas, "Design of Asynchronous Viterbi Decoder Using Pipeline Architecture" IJRASET Volume 6 Issue I, January 2018
9. Mr Sanket Kadu, "Design and Implementation of Viterbi Encoder and Decoder on FPGA" IJIRST –International Journal for Innovative Research in Science & Technology| Volume 3 | Issue 10 | March 2017
10. Akash Thakur and Manju K Chattopadhyay, "Design and Implementation of Viterbi Decoder Using VHDL, 3rd International Conference on Communication Systems (ICCS-2017)
11. Moussa Hamdan1 and Abdulati Abdullah2, "Analysis and Performance Evaluation of Convolutional Codes over Binary Symmetric Channel Using MATLAB" Proceedings of The Second International Conference on Electrical and Electronic Engineering, Telecommunication Engineering, and Mechatronics, Philippines 2016
12. Mahmoud Abdelaziz and T. Aaron Gulliver,"Ternary Convolutional Codes for Ternary Phase Shift Keying" IEEE COMMUNICATIONS LETTERS, VOL. 20, NO. 9, SEPTEMBER 2016
13. Juganpreet Kaur Brar, R K Bansal, Savina Bansal, "Goodness Analysis of Generator Polynomial for Convolution Code with Varying Constraint Length", IJARCCE, Vol 5, issue 11, Nov 2016
14. Gaurav Purohit,1 Kota Solomon Raju,2 and Vinod Kumar Chaubey1, "XOR-FREE Implementation of Convolutional Encoder for Reconfigurable Hardware, Hindawi Publishing Corporation International Journal of Reconfigurable Computing Volume 2016
15. Deepa Kumari*, Madan Lal Saini "Design and Performance Analysis of Convolutional Encoder and Viterbi Decoder for Various Generator Polynomials" IJERA Vol. 6, Issue 5, (Part - 2) May 2016
16. Do Duy Tana and Yeon-Mo Yangb, "Design and Simulation of Rate One-Third Convolutional Codes with Viterbi Algorithm based Hidden Markov Model for Digital Communications" Journal of Multidisciplinary Engineering Science and Technology (JMEST) ISSN: 2458-9403 Vol. 3 Issue 10, October – 2016
17. Arpitha K H1, Dr. P A Vijaya2, "Design of Low Power Efficient Viterbi Decoder" (IJRSEEE) Volume 2, Issue 2, 2016, PP 1-7
18. Khloud Mostafa, "High performance reconfigurable Viterbi Decoder design for multi-standard receiver", 33rd NRSC,Feb 2016
19. Rakhi B. Menon, Dr. Gnana Sheela K, "Synthesis of Convolution Encoder and Viterbi decoder of rate 2/3 using Xilinx ISE tool, IJARECE Vol 5, issue 2, Feb 2016
20. Hao Peng, Rongke Liu, Yi Hou and Ling Zhao, "A Gb/s Parallel Block-based Viterbi Decoder for Convolutional Codes on GPU", JULY 2016
21. Bhaskar Nandy, "Analysis of Convolutional Encoder System",IJCET (April 2016)
22. Rakhi B. Menon, 2 Dr. Gnana Sheela, "Design and Implementation of Convolution Encoder and Viterbi Decoder,IJCSN,volume 4 ,issue 6, December 2015
23. Ranjitha S, Divya Preethi K, Megha K,, "An Efficient FPGA Implementation of Convolutional Encoder and Viterbi Decoder for DSP Applications", International Journal of Engineering Research & Technology (IJERT Vol. 4 Issue 10, October-2015
24. Gaurav Vijay, R.P Gupta, "Performance Analysis of Convolutional Codes with Bit Error Rate V/S SNR over Channel for Viterbi Algorithm"IJIRCCE ,Vol 3 ,issue 6, June 15
25. D. Vaithiyanathan a,, J. Nargis b, R. Seshasayanan n, "High performance ACS for Viterbi decoder using pipeline T-Algorithm, ELSEVIER 15 April 2015
26. NEHA, GH.MOHAMMAD RATHER, "CONVOLUTION ERROR CONTROL CODING -A REVIEW", International Journal of Electrical, Electronics and Data Communication, ISSN: 2320-2084 Volume-3, Issue-3, March-2015

Fake News Detection in Social Media: A Systematic Research Review

Mr. Parshwa R. Shah[a]*, Mr. Devam M. Sheth[b]

[a, b]*Student, Arizona State University, Tempe,AZ*

Abstract

With the expanding fame of social media and electronic discussions, the dissemination of phony news has become a significant danger to different areas and offices. This has reduced trust in the media, leaving users in a condition of perplexity. The spread of this kind of information represents a genuine threat to union and social prosperity since it cultivates political polarization and doubt of individuals regarding their politicians. The large measure of information that is spread through social media makes manual confirmation unworkable, which has advanced the plan and execution of different techniques for fake news detection. The makers of phony news utilize different complex stunts to advance the accomplishment of their manifestations, with one of them being triggering people's thought processes. Due to this sentiment analysis is accountable for deciding the extremity and strength of opinions commuted in content. So, Sentiment analysis can be used either as a premise of the framework or as a feature inside the database. In this article, we study the unique employments of detection of fake news using sentiment analysis with a conversation of the most pertinent benefits and shortcomings, and the necessities that should be met sooner rather than later, for example, multilingualism, fairness, the balance of inclinations, or treatment of negative media components.

Keywords: Phony News, Social Media, Linguistics, Semantics, Machine learning

---

1. Introduction

In the past, if anybody required any news, the individual would hang tight for the following day's paper [9]. In any case, with the development of online papers and the measure of time spent via social media stages individuals have discovered a superior and quicker approach to being educated regarding the matter of his/her advantage. It has become a main source for news reading, for instance, according to the Pew Research Centre's Journalism Project [2], in 2020, 53% of US grown-ups say they acquired news from web-based media "frequently" or "now and again", with 59% of Twitter clients and 54% of Facebook clients burning-through news on the website routinely. Bogus data can be proliferated by bots, criminal/fear-monger associations, extremist or political associations, governments, covered-up paid banners, state-supported savages, writers, helpful numbskulls, scheme scholars, people that profit from bogus data, and savages [3]. The aim of these entertainers can be to harm or hurt, to get monetary benefit by expanding site sees, to control general assessment, to make problems and disarray, to advance philosophical inclinations, or even as individual amusement [4].

1.1 Need and Motivation: -

The motivation behind fake news detection is to combat the negative impact of fake news on society. Fake news can have serious consequences, including spreading misinformation, influencing public opinion, and damaging trust in traditional news sources. By detecting and flagging fake news, researchers and journalists can help prevent the spread of misinformation, reduce the impact of propaganda and biased information, and increase the credibility of news sources. Additionally, developing effective techniques for fake news detection can also help in the fight against disinformation campaigns, online scams, and other malicious activities that exploit the vulnerabilities of online platforms. Ultimately, the motivation for fake news detection research is to promote the integrity and accuracy of the information in the digital age.

1.2 Challenges: -

- Volume of information: The vast amount of information available on the internet makes it challenging to identify fake news from authentic news sources. Moreover, fake news can spread rapidly through social media, making it even more difficult to track and monitor.
- Lack of labeled data: To train machine learning models to detect fake news, large amounts of labeled data are required. However, such data is scarce, making it difficult to train and evaluate models accurately.
- Diverse forms of fake news: Fake news can take many different forms, including misleading headlines, manipulated images or videos, and completely fabricated stories. As a result, it is challenging to develop a single model or approach that can detect all forms of fake news effectively.

---

* Mr. Parshwa R. Shah
E-mail address: parshwas09@gmail.com

- Human biases: Humans have inherent biases that can affect their ability to detect fake news. For example, people may be more likely to believe the fake news that aligns with their preconceived beliefs or biases.
- Evolving tactics: As technology and social media continue to evolve, the tactics used to spread fake news are also changing. Attackers are becoming more sophisticated and are constantly developing new ways to spread fake news and evade detection.
- Legal and ethical challenges: There may be legal and ethical challenges in detecting and flagging fake news, such as concerns around censorship and freedom of speech.

An intentionally tricky story is "fake news" Any way of late babbling on the web media's discussion is changing its definition. Some of them at present use the term to pardon the real factors counter to their supported perspectives [5]. An exceptional trait of information via web-based media is that anybody can enroll as a news distributor with no forthright expense (e.g., anybody can make a Facebook page professing to be a paper or news media association). Thusly, not just customary news, companies are progressively moving to web-based media. (https://www.comscore.com/Insights/Blog/TraditionalNews-Publishers-Take-Non-Traditional-Path-toDigital-Growth). Alongside this progress, as anyone might expect, there are developing worries about counterfeit news distributors posting "counterfeit" reports, and regularly spreading those broadly utilizing "counterfeit" supporters [6].

Here is a list of some fake news that had a very huge (negative) impact:

- Before the finish of the 2016 official political decision, a report assessed that more than 1 million tweets were found connected to the phony news story "Pizza Gate" [7]. With the most recent three months of 2016's U.S. official political decision, numerous individuals accepted that the phony news favored the result in either of the two competitors was more than 37 million times shared on Facebook and Twitter [7] Thus, the spread of phony news might cause enormous scope adverse consequences, frequently affecting or in any event, abusing public occasions.
- A picture on Facebook shows a man's skull hacked open that was seen more than multiple times. The Facebook clients who posted the pictures asserted they showed a slaughter in progress in the Gashish area of Plateau State, Nigeria by Fulani Muslims who were killing Christians from the locale Berom ethnic minority. As an outcome, a slaughter occurred in Gashish that weekend, and somewhere close to 86 and 238 Berom individuals were killed, as per estimates made by the police and by nearby local area pioneers. However, the images and videos of a man's skull were not even from Gashish. The video showing a man's head was cut, didn't occur in Nigeria and it was recorded in Congo, in 2012. [8]
- An investigation by Silverman [9] shows that for the main 20 political race stories in 2016, the best 20 fake reports had 8,711,000 shares, comments, and likes on Facebook. These client commitment numbers were essentially higher than those for the main 20 genuine stories, with 7,367,000 shares, comments, and likes on Facebook during a similar time frame. These worries rouse evaluating news validity and recognizing counterfeit word before it gets out on the web.
- In the financial field, a piece of phony news guaranteeing that Barack Obama was injured in a blast cleared out $130 billion in stock value [10].
- In the social field, many guiltless individuals were pounded into the ground by local people in India give a piece of fake news about child trafficking that was broadly spread on social media [11].

Since individuals are frequently incapable to invest sufficient energy to cross-check references and make certain of the validity of information, automatic detection of fake news is crucial. Along these lines, it is getting incredible consideration from the research community. A few techniques have been pointed out in the new past to recognize and handle the issue of phony news. These could be extensively sorted into (a) Content-based: Text (semantics [12]); Media (pictures [13], GIFs, and video) and URLs, (b) User-based: movement following (bots and spam [14]); bio-data (registration age [15]); restricting perspectives on other online clients [16] and (c) Metadata-based: GPS, gadget source, Followers and Friends Network [17].

1.3 Introduction to Sentiment Analysis: -

Sentiment Analysis (SA) is the part of Natural Language Processing (NLP) responsible for the planning and execution of models, strategies, and procedures to decide if a book manages unbiased or emotional data and, in the last case, to decide whether such data is communicated in a good, nonpartisan, or negative way just as in case it is communicated in a solid or frail manner. Since a huge piece of the abstract substance communicated by clients on social networks is about suppositions (on survey locales, discussions, message sheets, visits, and so forth), SA is otherwise called Opinion Mining (OM) [18]. Sentiment analysis can help identify biased language that is used in fake news articles. For example, an article that uses strong emotional language, such as "outrageous" or "shocking", may indicate a biased perspective. By analyzing the sentiment of the language used in the article, sentiment analysis can help identify the such biased language. Fake news often targets people's emotions to influence their opinions. Sentiment analysis can help identify the emotions that are being targeted, such as anger, fear, or sadness, and determine whether the news is trying to manipulate the emotions of the reader. Sentiment analysis can also help identify inconsistencies within an article. For example, an article that claims to report on a factual event, but uses language that evokes strong emotions or biases, may indicate that the article is fake news. By analyzing the sentiment of the language used in the article, sentiment analysis can help identify such inconsistencies.

Past work has shown that misdirection and bogus proclamations can be identified from the composing style of the creators or semantics and now and then be utilized to infer their characters [19]. A few creators have shown that liars can even be recognized as they recount complex stories, make fewer self-references to disassociate themselves from the story, and will in general have more continuous utilization of negative feeling words – as an indication of blame [20]. Along these lines, it is intelligent to think about feelings inside the posted writings as a prompt comparable to getting out counterfeit news/gossip.

Sentiment analysis is typically approached as a machine learning problem, where the goal is to automatically classify the sentiment of a piece of text as positive, negative, or neutral. It can also perform contextual analysis, which involves analyzing the surrounding words and phrases in addition to the words directly associated with the sentiment. This allows sentiment analysis to identify subtle nuances in language that may indicate bias or manipulation. Sentiment analysis is a flexible technique that can be applied to a wide range of data sources, including social media, news articles, and online forums. This makes it a versatile tool for detecting fake news across different platforms and channels

2. Related Work

The majority of previous studies have focused on categorizing online news and social media articles. Various researchers have presented various ways of detecting deceit. Hai et al. [21] have planned a semi-supervised learning method by utilizing Laplacian regularized logistic regression to further develop the survey spam identification execution. In the writing, a few methodologies use PC vision for counterfeit news identification. An intriguing strategy, falling into that class, for picture-based phony photograph identification, has been introduced in [24].

Zang [22] has proposed a profound repetitive diffusive neural organization to resolve the issue of phony news identification. Then again, rather than the customary RNN model, in [70] creators adjusted a pre-prepared BERT model (Bidirectional Encoder Representations from Transformers), that comprises a few stacked transformer-encoder blocks.

Conroy et al. [11] looked at two significant types of techniques for detecting counterfeit news. The first was syntactic methods in which the text of deceitful communications is collected and examined to identify language trends. The subsequent one tries to get duplicity by measuring message data or network queries. In the particular instance of the location of phony surveys, Sentiment Analysis can be viewed as a helpful strategy not explicitly to recognize counterfeit messages but to identify counterfeit negative commentators as they overproduced negative passionate terms when contrasted with honest audits because of embellishments of the feeling they were attempting to pass on.

While Julia et al. [24] are of the opinion that Fake news deception detection can be more accurate when the features are selected properly. They classified features as follows: Language Features(They carried out 31 highlights which included no. of words and syllables per sentence), Lexical Features: Typical lexical highlights incorporate word-level signals like a measure of extraordinary words and their frequency in the content, Psycholinguistic Features(Used to catch extra signals of influential and one-sided language), Lexicon Features (Highlights that catch the lexicon parts of a book using Google API), Subjectivity (Figure subjectivity and polarity scores of a text).

Hussein [25] grouped 41 articles on SA as per the test they tended to. They tracked down that eight articles tended to refutation, seven managed domain dependence, 6 were committed to fake and phony recognition, 2 tended to global information, 8 managed Natural Language Processing overheads (mockery was remembered for this test), 3 chipped away at feature extraction, 3 contemplated polarity (words having extremity relies upon the setting where they are utilized [26]), and four managed gigantic dictionaries. As can be seen, counterfeit location is one of the primary difficulties, although the vast majority of the articles investigated in [25] didn't manage counterfeit news yet rather with the identification of phony sites or phony audits.

Thorne and Vlachos [27] inspected truth checking in reporting and recorded the assets and strategies accessible to computerize such an undertaking just as the connected works that could profit them. Elhadad [28] separated phony news from different types of dispersing false news, deception, and abnormality, for example, fabrications, fake announcement, parody/spoofs, reports, misleading content, and garbage news. They further augmented falsehood to the traditional classifications of false news and deception. Falsehood was characterized as the passing of veritable doom to hurt someone. Notwithstanding, created or garbage news, which can't be thought of as containing authentic data, was taken into account as a potential falsehood acknowledgment, which appears to be conflicting. Sentiment Analysis was not referenced in either [27] or [28].

Da Silva et al. [29] examined ML approaches and strategies to distinguish counterfeit news, tracking down that the favored techniques included neural networks made out of old-style grouping calculations that vigorously center on the lexical examination of the sections as a fundamental feature for prediction. SA was regularly utilized as a substance highlight as words having a place with assumption vocabularies or as the aftereffect of an ML-based opinion mining framework. Klyuev [30] likewise examined various ways to deal with battle counterfeit news and the significance of deciding content highlights through natural language processing (NLP) strategies to make a profile of the content report. Even though he showed the significance of utilizing word references that contain, among other data, the feeling of the extremity of words, he didn't unequivocally specify the assignment of opinion mining.

Meel and Vishwakarma [31] overviewed how the substance on the web is stained purposefully or some of the time inadvertently by counterfeit audits, counterfeit news, and parody, among different wellsprings of data contamination. For this reason, they examined the bogus data environment, from the order of bogus data and the inspirations to scatter it to the social effect and client discernment. They additionally talked about the present status of reality checking, which includes an origin location, strategies for identification, and techniques for control and intercession. They hypothesized SA as the basic wellsprings of data expected to identify bogus data.

Niraj and Chilukuri [32] mentioned the believability of news, underlining features identified with the source articles. Their outcomes dependent on the well-spring of the news show that number of creators of the news is a solid pointer of validity. We found that when the news story has no creators, it is bound to be phony information. Our discoveries on the cooperation of creators recommend that creators who are occupied with valid trustworthy news are less likely to team up with creators who are related to counterfeit news. This demonstrates that for a news story with various creators, by knowing the believability of one creator, we can construe the validity of the news just as other co-authors. Besides, we came to know that creators' affiliations with very much perceived associations can be a sign of credibility. The outcomes additionally recommend that the credit history of writers can give bits of knowledge on the believability of different articles from the same creator.

Oshikaw in [33] examined the specialized difficulties in counterfeit news location and how analysts characterize various assignments and figure ML answers to handle this undertaking, concentrating on how counterfeit news discovery was lined up with ongoing NLP tasks.

Zhang and Ghorbani in [34] described the adverse consequence of online phony news and fake news recognition techniques for this kind of data, tracking down that a large number of them depend on recognizing highlights of the clients, context, and content that demonstrate deception. They additionally claimed that developing effective and understandable false news detection algorithms will necessitate collaboration between specialists in social sciences, journalism, computer and information science, and political science. Accordingly, they thought that Sentiment investigation is a valuable strategy to show the feelings, mentalities, and suppositions that are passed on by online web-based media and those Sentiment related components are key ascribes for dubious record recognizable proof.

Zhou and Zafarini in their paper [35] studied phony news identification according to the viewpoint of information-based techniques that recognize counterfeit news by checking if the information inside is provided with facts or not. They considered opinion as a significant lexicon component of information. They likewise expressed that the execution of proficient and reasonable phony articles location frameworks needs community-oriented endeavors including specialists in PC and data sciences, sociologies, political theory, and news coverage.

De Souza et al. [36] audited the various kinds of highlights identified with counterfeit news discovery strategies and informational collections, and they thought that SA was a valuable element to rapidly confirm the exactness of data via online media. At last, Antonakaki et al. [37] introduced an overview on present research subjects in Twitter, discovering that feeling investigation was one of the four fundamental parts of examination including Twitter, and that one of the significant dangers for this informal organization is the spread of phony news through it. In any case, they broke down the two points independently, without making an association between the two which shows the handiness of assessment investigation in distinguishing counterfeit news.

Ajao et al. [12] have shown that phony news can be distinguished by utilizing the content-based just methodology without earlier information on the subject area. It is significant that phony and bogus data spreads a lot speedier and more profoundly than genuine data. S. Kumar et al. [38] have so far made the biggest talk data of 126k messages spread by very nearly 300K individuals and tracked down that phony news diffused up to 100K individuals while reality just contacted 1,000 individuals. C. Guo et al.[39] recognized that ‘lone wolves’ scatter their message quicker by making counterfeit records that state a similar viewpoint in numerous manners to assist with propagating their message quicker.

Y. Kim et al. [40], proposed another double feeling-based way to deal with identifying counterfeit news where it can gain from the substance, client remark, and portrayal of feelings from the two distributors and clients. Kim et al. [40] recognize diverse granularity of text highlights with convolution channels for detecting phony news by using a convolutions neural network model.

Notwithstanding highlights straightforwardly identified with the news stories' substance, helper data can be separated from the client-driven social commitment to information via web-based media. In [42], they proposed an original way to deal with recognizing counterfeit news dependent on news content utilizing an information chart.

Mathieu Cliche presented the discovery of mockery on Twitter using n-grams, terms derived from tweets expressly labeled as wry, on his mockery detection site. His study additionally incorporates the utilization of opinion mining and the detection of themes (words that are frequently assembled in tweets) further develop expectation exactness. [43]

Hannah Rashkin in his paper [44] has played out a broad examination of etymological highlights to demonstrate the aftereffect of LSTM. Singhania in his paper [45] developed a 3-structured hierarchical consideration network, one containing a word, another a sentence, and a news article element. Ruchansky et al. [46] developed the CSI model, which captures text, article reaction, and source quality as a function of client behavior.

3. Datasets used in Different research works:

One of the most important factors in detecting fake news is a bipolar dataset. Our Dataset should contain both reliable and fake news examples in equivalent proportion. The fundamental issues in building such informational datasets are that the measure of bogus data in the online content created each day, regardless of whether we limit our focus on news stories and posts examining breaking news, and that web-based media organizations these days have severe policies concerning the examination of data delivered by their users [47]. Beneath, we list the informational indexes that have been worked by established researchers to survey the presentation of phony news recognition calculations, procedures, and frameworks.

- LIAR dataset with 12,800 human-named short explanations from PolitiFact.com assessed for its honesty utilizing six features like ‘true’, ‘mostly true’, ‘half true’, ‘barely true’, ‘false’, and pants fire’. A rich arrangement of meta-information for the creator of every assertion is likewise given. The assertions are examined from news deliveries, Twitter messages, Television and radio interviews, campaign discourses, TV promotions, discussions, and Facebook posts. The vast examined subjects are medical services, taxes, government spending plans, schooling, occupations, state-financial plan, economy, election decisions, and migration. [48] ((https://www.cs.ucsb.edu/william/data/liar_dataset.zip)

- Buzz Face (https://github.com/gsantia/BuzzFace) [49] is a supplement to the previous corpus that includes 16 lahks Facebook remark responses as well as additional data from Twitter and Reddit.

- The BS Detector informative data, at times referred to as Kaggle FakeNews, consist of text/strings and data from 244 sites and covers 12,999 postings from the past 30 days, obtained using a program expansion named BS detector. This expansion scans each and every connection on a particular website page for references to problematic origins/source by comparing with physically gathered rundown spaces. In this way, the records in the informational index were named by programming, not by human annotators.

- Fact-checking data set Reality, an assortment of appraised proclamations from PolitiFact.com with extra problematic news stories from various kinds of untrustworthy sources including parody, publicity, and deceptions. [50]

- Craig Silverman which is also called Buzz feed Political News’ Data includes genuine reports and malignant phony reports from buzzfeednews.com. A different information set called Random Political News Data includes genuine news from many different USA newspapers and journals.

- BuzzFeed-Webis Fake News Corpus 2016. This dataset includes 1627 posts from Facebook coming from nine distributors on seven workdays near the US Political Elections of 2016 official political race. It includes 256 posts from 3 remaining wing distributors, 826 posts from three standard distributors, and 545 posts from three conservative ones. All distributors acquired Facebook's official checkmark, demonstrating validness and raising status inside the organization. [51]

4. Assessment Measures

Accuracy (Acc.), Precision (P), Recall (R), F1, and Area under the Curve (AUC) are the most broadly used measurements to decide the presentation of false news identification frameworks. We should consider the number of true positives (|TP|), true negatives (|TN|), false positives (|FP|), and false negatives (|FN|) while processing the measurements where,

• A true positive is actually fake news which is predicted as fake news by our test results.
• A false positive is actually genuine news which is predicted as fake news by our test results.

• A true negative is actually genuine news that is predicted as genuine by our test results.
• A false negative is actually fake news which is predicted as genuine news by our test results.

Accuracy: The percentage obtained by dividing the number of articles correctly identified by the framework, as either phony or genuine news, by the number of articles named accurately recognized by the framework, and is processed as shown in Equation below (1).

Precision: The percentage resulting by dividing of articles that have been actual fake news by the number of articles that have been predicted as fake news as shown in equation (2). The recall is the percentage resulting by dividing the number of articles that have been actual fake news by the totally fake news articles as shown in equation(3). F1 is the harmonic mean of Precision and recall as shown in the equation. (4)

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \qquad (1)$$

$$\text{Precision} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \qquad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \qquad (3)$$

$$\text{F1 Score} = \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (4)$$

Another important term is False Positive Rate (FPR) which can also be defined as(in the context of fake news) the number of genuine news wrongly identified as fake to the total number of genuine articles(5). The Receiving Operating Curve is recognized as a curve between False Positive Rate and True Positive Rate (which is also known as Recall). A best model has AUC (Area under Curve) close to 1 while it is close to 0 in a poor/worst model.

$$\text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}} \qquad (5)$$

5. Automatic fake/phony news detection with help of sentiment analysis:

As previously said, the intensity with which an opinion is presented establishes a crucial component in determining the amount of authenticity of the news. In light of recent advancements, we can differentiate two types of methodologies for identifying phony news, both of which take into account SA. From one viewpoint, a bunch of methods accept Sentiment Analysis as the crucial premise for phony news recognition methodology and is generally supplemented by the utilization of other data separated from both news substance and the setting of information propagating on interpersonal organizations. Our next section is based on this and the section after that reviews papers that have taken SA into their feature column.

5.1 Phony News Identification Dependent on Sentiment Analysis

Diakopoulos et al. [52] proposed an insightful mechanism to assist columnists and journal editors in extracting news value from vast collections of web-based media content around televised events and, specifically, to encourage investigators to collect data from reputable sources. As a result, they utilized 4 different sorts of programmed text importance: importance, originality, emotion, and word retrieval. They followed a 2-venture method for assessment examination. They started with a classifier dependent on a vocabulary of terms that classified messages based on whether they were sending abstract data or not. As a subsequent advance, they tried applying a Machine-Learning Classifier[69] which is dependent on n-grams upto a size of 4. This brought about a 5-overlap CV exactness (62.4%) that was adequate to give a general image of a supposition yet fizzled on troublesome cases including mockery or slang. But they found that notion investigation related to the extent of the social media reaction to various statements, points, or issues were the most valuable logical markers for editorial requests.

Even when viral dissemination is powered by the evocation of intense emotions[68], Zhang in his paper [53] believed that most previous study on phony news was based on the feelings and sentiment of material given by publishers but hardly focused on the emotions of comments provoked in the public. Yes, they looked at whether feelings expressed in comment sections and their correlation to those expressed in the actual substance could aid in the detection of counterfeit news. They utilized different DL classifiers with respect to information in Chinese and English. Conclusively the outcomes on the Chinese phony articles' informational index were acceptable however; the outcomes on the English informational collection have been very less, presumably on the grounds that the informational collection was initially intended for the discovery of rumors, not phony news.

Dey in his paper [54] tried applying a few NLP strategies (named-element recognition, opinion mining) on a bunch of two hundred tweets alluding to the USA 2016 Election. They tracked down that trustworthy tweets in the most part contained sure or impartial extremity, while tweets with counterfeit substance had a solid propensity towards negative assumption. In any case, their data was too little to even think about getting definitive outcomes. Bhutani in his paper [55] proposed the theory that to find if the news is genuine or phony, the polarity of the words in an article plays a crucial role. For deciding the polarity of writings they used Naive Bayes and afterward applied random forest and Multinomial Naive Bayes. Multinomial Naive Bayes was proved to be giving the best results.

Cui in his paper [56] discovered factual proof on a dataset called Fake News Net, that the polarity or extremity of remarks below counterfeit articles was more noteworthy than below genuine news. For identifying counterfeit news, clients' latent semantics were chosen in a start-to-finish embedding structure. For managing clients' profiles, text in the news, and pictures in the news; they utilized 3 neural organizations for safeguarding lexicon likeness and to uphold the representing stability between a picture and content they used an adversarial mechanism. At long last, they displayed clients' polarity for consolidating it into a proposed structure. Utilizing adversarial learning for lexicon relationships was a curiosity of this work. Other traditional and deep learning-based classifiers were defeated by the following phony news discovery approach.

5.2 Phony News Identification Using Sentiment Analysis as a Feature Column:

Hassan in his paper [57] interpreted real cases into questions against a repository of truth checks using online media, news and live talks. The framework was prepared on a bunch of US general political decision official discussions in which each and every sentence was commented as either "Non-Factual", "Irrelevant Factual", or "Factual". They extricated as many highlights as possible like no. of events, no. of sack words, sentiment, etc. Sentiment was considered as the third most significant feature among the 6615 collected.

Before counterfeit news turned into a first-order issue, there was at that point some work in the writing that focused on deciding the validity of data flowing on social media. In this line, Castillo et al. [59] focused on programmed techniques for evaluating the validity of sets of tweets that spread data about a news occasion. They considered data from official and respectable sources as important data that different clients integrate and elaborate to deliver inferred translations in a nonstop cycle. For their trials, they gathered 747 arrangements of tweets getting out the word, and every one of them was physically labeled as "in all likelihood obvious", "liable to be fake", "in all likelihood fake", and "I can't choose". They noticed that features dependent on SA were exceptionally significant for evaluating the validity, as 3 out of the 10 best-performing highlights were sentiment-related ones: average slant score, number of positive estimation words, and number of negative assumption words. They likewise saw that tweets showing positive supposition were more identified with non-trustworthy data while those with negative notions would in general be more identified with tenable data.

Ross in [60] contrasted with the above explanation stated in [59] that, a bunch of tweets should be ranked by their validity and importance rather than deciding their believability. In their base arrangement of features. There were opinion highlights such as, “has a cheerful emoji", "has a tragic emoji", and an assessment result in their basic arrangement. They implied that tweets with similar notion are dependable and tweets which had assumptions are invalid; this implication was done with the help of 2 highlights/features. The depiction of creator and content of tweets were used to get the highlights/features of various tweets.

Shu in his paper [61] investigated data called Fake News Net, discovering that individuals state their feelings/ viewpoints to counterfeit articles with the help of web-based media pictures by showing distrustful assessments and sensational responses, with genuine news having a bigger extent of unbiased answers over good and regrettable answers, while counterfeit news consisting a greater proportion of negative supposition. While starting their investigations to detect counterfeit news, they tried utilizing basic features from text and so they didn't give details about the effect that SA has on the discovery of phony news in this dataset.

Vosoughi et al. [62] examined valid and counterfeit reports spread on social media platforms (Twitter). They tracked down that phony articles propelled the feelings of dread, disdain, and shock in answers, while genuine articles roused expectation, trouble, bliss, and trust. They infer that the feelings communicated in comments to counterfeit news might enlighten extra factors that rouse individuals to share bogus news. Despite the fact that emotional feeling isn't equivalent to SA, they are closely related in light of the fact that both analyze the emotional substance communicated in content, and the classification models utilized are practically the same in the two cases, with the greatest distinction being in the arrangement of classes with which the writings to be handled are commented on. Thus, we have remembered this article for the investigation.

Bhavika and Neha in [63] made their own vocabulary to train and predict the polarity of a particular sentence. They added the predicted results in feature columns along with TF-idf scores and cosine similarity. They also used Count vectorizers along with n-grams and also without n-grams. They got different accuracies on 3 different datasets, but they found that Tf-idf Vectorizer with Cosine similarity was the best when using Naïve Bayes and Random Forest Classifier in either of the 3 datasets.

Reis et al. [64] separated an enormous number of features by utilizing linguistic strategies, for example polarity, validity and dependability, and area from news content. Subjectivity and Polarity were among the setting feature for this news. XGBoost and Random forest acquired the best presentation among a number of classifiers while using a dataset called Buzz Face. By watching subsequent outcomes, creators picked a threshold for classifying each and every phony article whose True Positive Rate is near 1, and thought that this will be helpful in truth checking.

6. Discussion

In the previous section, we saw how Sentiment Analysis can be utilized in a variety of ways to improve the performance of false news detection systems. In the below table comparative analysis has been shown which includes the accuracy of different algorithms along with different opinion mining methods applied on various datasets.

Table 1. Comparative analysis of fake news detection systems

Author and Reference	DataSet	Opinion Mining Method	Algorithm Used	Results
Horne [65]	Silverman Buzzfeed News	Semantics based	Support Vector Machine	Acc. =0.77
Zhang [66]	Politifact Fakenews	Semantics based	Deep Neural Network	F1 score = 0.7724
	Gossip Cop FakeNews			F1 score = 0.8042
Popat [67]	SNOPES	Semantics Based	Logistic Regression	Acc. =71.96; AUC =0.80
			Conditional Random Field	Acc. = 84.02; AUC = 0.86
Bhavika and Neha [63]	George McIntire	TF-IDF Vectorizer with Cosine Similarity	Naïve Bayes	Acc. = 0.843
			Random Forest	Acc.= 0.839
Rashkin [50]	Fact Checking	Semantics Based	LSTM	F1 =0.55
Dey [54]	Twitter dataset named Ad-hoc	Semantics Based	K-Nearest Neighbor	Acc. =0.66
Reis[64]	BuzzFace		Support vector Machine	AUC = 0.79 ; F1 = 0.76
			K-Nearest Neighbor	AUC = 0.80; F1 = 0.75
			XG Boost	AUC = 0.86 ; F1 = 0.81
			Random Forest	AUC = 0.86 ; F1 = 0.81

7. Conclusion and Future Scope

With the growing popularity of social media, an increasing number of people are absorbing news through social media on a regular basis. The growth of fake news, on the other hand, is a rising source of concern for us today, since it has significant negative consequences for both individual users and society as a whole. We looked at the field of fake news identification from the perspective of how sentiment analysis is being used to combat the problem in this paper. We may claim that the field of false news detection research is moving from the initial stages to the development stage. The need to ensure system fairness, accountability, and transparency (ensuring that results are explainable and free of harmful biases); support for multilingualism and multimedia content; and detection of fake news generated by subtly altering authentic stories or using text-generation algorithms are among the most pressing challenges, in our opinion.

These are some of the future recommendations which can be implemented in this domain.

- Developing more sophisticated algorithms: While machine learning algorithms have shown promise in detecting fake news, there is still room for improvement. Future research could focus on developing more sophisticated algorithms that can identify more subtle patterns and features in social media data.

- Incorporating more data sources: While social media is a rich source of data for fake news detection, there are other data sources that could be incorporated into the analysis. For example, news articles, blogs, and online forums could be used to provide additional context and information for fake news detection.

- Addressing the challenge of deep fakes: Deepfakes are synthetic media that are created using artificial intelligence and are designed to look and sound like real people. Detecting deep fakes can be challenging, and future research could focus on developing techniques for identifying deep fakes in social media.

- Evaluating the effectiveness of interventions: While there are many techniques for detecting fake news, it is unclear which techniques are most effective in practice. Future research could focus on evaluating the effectiveness of different interventions, such as fact-checking or social media platform policies, for reducing the spread of fake news.

- Examining the impact of fake news on society: While fake news has been widely studied from a technical perspective, there is less research on the impact of fake news on society. Future research could focus on examining the social, political, and economic consequences of fake news and developing strategies for mitigating its impact.

- Overall, there are many exciting areas for future research in fake news detection in social media. By developing more sophisticated algorithms, incorporating additional data sources, addressing the challenge of deep fakes, evaluating the effectiveness of interventions, and examining the impact of fake news on society, researchers can make significant contributions to the field and ultimately help prevent the spread of misinformation.

References

1. Khan, Junaed Younus, et al. "A benchmark study on machine learning methods for fake news detection." arXiv preprint arXiv:1905.04749 2 (2019).
2. Shearer, E.; Mitchell, A. News Use Across Social Media Platforms in 2020. 2021. Available online: https://www.journalism.org/2021/01/12/news-use-across-social-media-platforms-in-2020/ .
3. Zannettou, S.; Sirivianos, M.; Blackburn, J.; Kourtellis, . "The Web of False Information: Rumors, Fake News, Hoaxes, Clickbait, and Various Other Shenanigans". ACM J. Data Inf. Qual. 2019, 11, 10:1–10:37. [CrossRef]
4. Sharma; Qian; Jiang; Ruchansky,; Zhang; Liu. "Combating Fake News: A Survey on Identification and Mitigation Techniques". ACM Trans. Intell. Syst. Technol. 2019, 10, 21:1–21:42. [CrossRef]
5. Dubey, S.; Kohok, P.; Thakre, D., "Fake News Detection using Sentiment Analysis", Spvryan's International Journal of Engineering Sciences & Technology (SEST) Volume 7, Issue 1, Paper 4
6. Reis, J.C.S.; Correia, A.; Murai, F.; Veloso, A.; Benevenuto, F.; Cambria, E. Supervised Learning for Fake News Detection. IEEE Intell. Syst. 2019, 34, 76–81
7. H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," Journal of economic perspectives, vol. 31, no. 2, pp. 211–36, 2017.
8. Khan, J.; Khondaker, T.; Iqbal, A., "A Benchmark Study on Machine Learning Methods for Fake News Detection" arXiv:1905.04749v1 May 2019
9. Silverman, C.: This analysis shows how viral fake election news stories outperformed real news on facebook (2016). https://www.buzzfeednews.com/article/craigsilverman/viral-fakeelection-news-outperformed-real-news-on-facebook#.vtQpz9DKd
10. Boididou, C., Papadopoulos, S., Dang-Nguyen, D.T., Boato, G., Riegler, M., Middleton, S.E., Petlund, A., Kompatsiaris, Y., et al.: "Verifying multimedia use at mediaeval 2016". In: MediaEval (2016).
11. Boididou, C., Papadopoulos, S., Zampoglou, M., Apostolidis, L., Papadopoulou, O., Kompatsiaris, Y.: "Detection and visualization of misleading content on twitter. Int. J. Multimed". Inf. Retr. 7(1), 71–86 (2018)

12. O. Ajao, D. Bhowmik, and S. Zargari, "Fake news identification on twitter with hybrid cnn and rnn models," in 9th Int'l Conference on Social Media & Society. Copenhagen (July 18), no. Jul 2018, 2018.
13. A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy," in Proceedings of the 22nd international conference on World Wide Web. ACM, 2013, pp. 729–736.
14. E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," Communications of the ACM, vol. 59, no. 7, pp. 96–104, 2016.
15. C. Castillo, M. Mendoza, and B. Poblete, "Information credibility on twitter," in Proceedings of the 20th international conference on World wide web. ACM, 2011, pp. 675–684.
16. Z. Jin, J. Cao, Y. Zhang, and J. Luo, "News verification by exploiting conflicting social viewpoints in microblogs." in AAAI, 2016, pp. 2972–2978.
17. E. Tacchini, G. Ballarin, M. L. Della Vedova, S. Moret, and L. de Alfaro, "Some like it hoax: Automated fake news detection in social networks," arXiv preprint arXiv:1704.07506, 2017.)
18. Alonso, M; Vilares, D; Gómez-Rodríguez, C; Vilares, J, MDPI, Basel, Switzerland; Electronics 2021, 10, 1348. https://doi.org/10.3390/electronics10111348
19. J. W. Pennebaker and L. A. King, "Linguistic styles: Language use as an individual difference." vol. 77, no. 6, p. 1296, 1999.
20. M. L. Newman, J. W. Pennebaker, D. S. Berry, and J. M. Richards, "Lying words: Predicting deception from linguistic styles," Personality and social psychology bulletin, vol. 29, no. 5, pp. 665–675, 2003.
21. H. Zhen, et al. "Deceptive review spam detection via exploiting task relatedness and unlabeled data." Proceedings of the Conference on Empirical Methods in Natural Language Processing, 2016
22. Zhang, J., Cui, L., Fu, Y., Gouza, F.B.: Fake news detection with deep diffusive network model. CoRR, abs/1805.08751 (2018)
23. Conroy, N.J.; Rubin, V.L.; Chen, Y. Automatic deception detection: Methods for finding fake news. In Information Science with Impact: Research in and for the Community—Proceedings of the 78th ASIS&T Annual Meeting, ASIST 2015, St. Louis, MO, USA, 6–10 October 2015; Wiley: Hoboken, NJ, USA, 2015; Volume 52, pp. 1–4. [CrossRef]
24. Hussein, D.M.E.D.M. A survey on sentiment analysis challenges. J. King Saud Univ. Eng. Sci. 2018, 30, 330–338. [CrossRef]
25. 16. Flekova, L.; Preotiuc-Pietro, D.; Ruppert, E. Analysing domain suitability of a sentiment lexicon by identifying distributionally bipolar words. In Proceedings of the 6th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, WASSA@EMNLP 2015, Lisbon, Portugal,
26. 17 September 2015; Balahur, A., der Goot, E.V., Vossen, P., Montoyo, A., Eds.; The Association for Computer Linguistics: Stroudsburg, PA, USA, 2015; pp. 77–84. [CrossRef] 17. Thorne, J.; Vlachos, A. Automated Fact Checking: Task Formulations, Methods and Future Directions. In Proceedings of the 27th International Conference on Computational Linguistics, COLING 2018, Santa Fe, NM, USA, 20–26 August 2018; Bender, E.M., Derczynski, L., Isabelle, P., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2018; pp. 3346–3359.
27. 18. Elhadad, M.K.; Li, K.F.; Gebali, F. Fake News Detection on Social Media: A Systematic Survey. In Proceedings of the IEEE Pacific Rim Conference on Communications, Computers and Signal Processing, PACRIM 2019, Victoria, BC, Canada, 21–23 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–8. [CrossRef]
28. da Silva, F.C.D.; Vieira, R.; Garcia, A.C. Can Machines Learn to Detect Fake News? A Survey Focused on Social Media. In Proceedings of the 52nd Hawaii International Conference on System Sciences, HICSS 2019, Grand Wailea, Maui, HI, USA, 8–11 January 2019; Bui, T., Ed.; ScholarSpace: Honolulu, HI, USA, 2019; pp. 1–8.
29. Klyuev, V. Fake News Filtering: Semantic Approaches. In Proceedings of the 2018 7th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 29–31 August 2018; pp. 9–15. [CrossRef]
30. Meel, P.; Vishwakarma, D.K. Fake news, rumor, information pollution in social media and web: A contemporary survey of state-of-the-arts, challenges and opportunities. Expert Syst. Appl. 2020, 153, 112986. [CrossRef]
31. Sitaula, N.; Chilukuri, M.; Jennifer, G.; Zhou, X.; Zafarni, R., "Credibility-Based Fake News Detection", Springer Nature Switzerland AG 2020 K. Shu et al. (eds.), Disinformation, Misinformation, and Fake News in Social, pp. 163-188.
32. Oshikawa, R.; Qian, J.; Wang, W.Y. A Survey on Natural Language Processing for Fake News Detection. In Proceedings of the 12th Language Resources and Evaluation Conference, LREC 2020, Marseille, France, 11–16 May 2020; Calzolari, N., Béchet, F., Blache, P., Choukri, K., Cieri, C., Declerck, T., Goggi, S., Isahara, H., Maegaard, B., Mariani, J., et al., Eds.; European Language Resources Association: Paris, France, 2020; pp. 6086–6093.
33. Zhang, X.; Ghorbani, A.A. An overview of online fake news: Characterization, detection, and discussion. Inf. Process. Manag. 2020, 57, 102025. [CrossRef]
34. Zhou, X.; Zafarani, R. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities. ACM Comput. Surv. 2020, 53, 109:1–109:40. [CrossRef]
35. de Souza, J.V.; Gomes, J., Jr.; de Souza Filho, F.M.; de Oliveira Julio, A.M.; de Souza, J.F. A systematic mapping on automatic classification of fake news in social media. Soc. Netw. Anal. Min. 2020, 10, 48. [CrossRef]
36. Antonakaki, D.; Fragopoulou, P.; Ioannidis, S. A survey of Twitter research: Data model, graph structure, sentiment analysis and attacks. Expert Syst. Appl. 2021, 164, 114006. [CrossRef]
37. S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," Science, vol. 359, no. 6380, pp. 1146–1151, 2018.
38. S. Kumar, M. Jiang, T. Jung, R. J. Luo, and J. Leskovec, "Mis2: Misinformation and misbehavior mining on the web," in Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining. ACM, 2018, pp. 799–800
39. C. Guo, J. Cao, X. Zhang, K. Shu, and M. Yu, "Exploiting emotions for fake news detection on social media," arXiv preprint arXiv:1903.01728, 2019.
40. Y. Kim, "Convolutional neural networks for sentence classification," arXiv preprint arXiv:1408.5882, 2014.

41. F. Marra, D. Gragnaniello, D. Cozzolino, and L. Verdoliva, “Detection of gan-generated fake images over social networks,” in 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, 2018, pp. 384–389
42. Mathieu Cliche. The sarcasm detector, 2014
43. Hannah Rashkin, Eunsol Choi, Jin Yea Jang, Svitlana Volkova, and Yejin Choi. Truth of varying shades: Analyzing language in fake news and political fact-checking. In Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing, pages 2931–2937, 2017.
44. Sneha Singhania, Nigel Fernandez, and Shrisha Rao. 3han: A deep neural network for fake news detection. In International Conference on Neural Information Processing, pages 572–581. Springer, 2017.
45. Natali Ruchansky, Sungyong Seo, and Yan Liu. Csi: A hybrid deep model for fake news detection. In Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, pages 797–806. ACM, 2017
46. Bondielli, A.; Marcelloni, F. A survey on fake news and rumour detection techniques. Inf. Sci. 2019, 497, 38–55. [CrossRef]
47. Wang, W.Y. “Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, BC, Canada, 30 July–4 August 2017; Volume 2: Short Papers; Barzilay, R., Kan, M., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2017; pp. 422–426. [CrossRef]
48. Santia, G.C.; Williams, J.R. BuzzFace: A News Veracity Dataset with Facebook User Commentary and Egos. In Proceedings of the Twelfth International Conference on Web and Social Media, ICWSM 2018, Stanford, CA, USA, 25–28 June 2018; AAAI Press: Palo Alto, CA, USA, 2018; pp. 531–540.
49. Wang, W.Y. “Liar, Liar Pants on Fire”: A New Benchmark Dataset for Fake News Detection. In Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics, ACL 2017, Vancouver, BC, Canada, 30 July–4 August 2017; Volume 2: Short Papers; Barzilay, R., Kan, M., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2017; pp. 422–426. [CrossRef]
50. Potthast, M.; Kiesel, J.; Reinartz, K.; Bevendorff, J.; Stein, B. A Stylometric Inquiry into Hyperpartisan and Fake News. In Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics, ACL 2018, Melbourne, Australia, 15–20 July 2018; Volume 1: Long Papers; Gurevych, I., Miyao, Y., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2018; pp. 231–240. [CrossRef]
51. Diakopoulos, N.; Naaman, M.; Kivran-Swaine, F. Diamonds in the rough: Social media visual analytics for journalistic inquiry. In Proceedings of the 5th IEEE Conference on Visual Analytics Science and Technology, IEEE VAST 2010, Salt Lake City, UT, USA, 24–29 October 2010; Part of VisWeek 2010; IEEE Computer Society: Washington, DC, USA, 2010; pp. 115–122. [CrossRef]
52. Zhang, X.; Cao, J.; Li, X.; Sheng, Q.; Zhong, L.; Shu, K. Mining Dual Emotion for Fake News Detection. In The Web Conference 2021, Proceedings of The World Wide Web Conference WWW 2021; ACM: New York, NY, USA, 2021. [CrossRef]
53. Dey, A.; Rafi, R.Z.; Hasan Parash, S.; Arko, S.K.; Chakrabarty, A. Fake News Pattern Recognition using Linguistic Analysis. In Proceedings of the 2018 Joint 7th International Conference on Informatics, Electronics Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision Pattern Recognition (icIVPR), Kitakyushu, Japan, 25–29 June 2018; pp. 305–309. [CrossRef]
54. Bhutani, B.; Rastogi, N.; Sehgal, P.; Purwar, A. Fake News Detection Using Sentiment Analysis. In Proceedings of the 2019 Twelfth International Conference on Contemporary Computing, IC3 2019, Noida, India, 8–10 August 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–5. [CrossRef]
55. Cui, L.; Wang, S.; Lee, D. SAME: Sentiment-aware multi-modal embedding for detecting fake news. In Proceedings of the ASONAM ’19: International Conference on Advances in Social Networks Analysis and Mining, Vancouver, BC, Canada, 27–30 August 2019; Spezzano, F., Chen, W., Xiao, X., Eds.; ACM: New York, NY, USA, 2019; pp. 41–48. [CrossRef]
56. Hassan, N.; Arslan, F.; Li, C.; Tremayne, M. Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster. In Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, 13–17 August 2017; ACM: New York, NY, USA, 2017; pp. 1803–1812. [CrossRef]
57. 181. Alonso, M.A.; Gómez-Rodríguez, C.; Vilares, J. On the Use of Parsing for Named Entity Recognition. Appl. Sci. 2021, 11, 1090. [CrossRef]
58. Castillo, C.; Mendoza, M.; Poblete, B. Information credibility on twitter. In Proceedings of the 20th International Conference on World Wide Web, WWW 2011, Hyderabad, India, 28 March–1 April 2011; Srinivasan, S., Ramamritham, K., Kumar, A., Ravindra, M.P., Bertino, E., Kumar, R., Eds.; ACM: New York, NY, USA, 2011; pp. 675–684. [CrossRef]
59. Ross, J.; Thirunarayan, K. Features for Ranking Tweets Based on Credibility and Newsworthiness. In Proceedings of the 2016 International Conference on Collaboration Technologies and Systems, CTS 2016, Orlando, FL, USA, 31 October–4 November 2016; Smari, W.W., Natarian, J., Eds.; IEEE Computer Society: Washington, DC, USA, 2016; pp. 18–25. [CrossRef]
60. Shu, K.; Mahudeswaran, D.; Wang, S.; Lee, D.; Liu, H. FakeNewsNet: A Data Repository with News Content, Social Context, and Spatiotemporal Information for Studying Fake News on Social Media. Big Data 2020, 8, 171–188
61. Vosoughi, S.; Roy, D.; Aral, S. The spread of true and false news online. Science 2018, 359, 1146–1151.
62. Bhutani, B.; Rastogi, N.; Sehgal, P.; Purwar, A., “Fake News Detection Using Sentiment Analysis”, 978-1-7281-3591-5/19/$31.00 , IEEE Conference 2019
63. Horne, B.D.; Adali, S. This Just In: Fake News Packs a Lot in Title, Uses Simpler, Repetitive Content in Text Body, More Similar to Satire than Real News. In Proceedings of the Workshops of the Eleventh International AAAI Conference on Web and Social Media (ICWSM 2017), Montreal, QC, Canada, 15–18 May 2017; An, J., Kwak, H., Benevenuto, F., Eds.; AAAI Press: Palo Alto, CA, USA, 2017; Volume AAAI Technical Report WS-17-17: News and Public Opinion, pp. 759–766.
64. Ajao, O.; Bhowmik, D.; Zargari, S. Sentiment Aware Fake News Detection on Online Social Networks. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP 2019, Brighton, UK, 12–17 May 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 2507–2511. [CrossRef]
65. Popat, K.; Mukherjee, S.; Strötgen, J.; Weikum, G. Credibility Assessment of Textual Claims on the Web. In Proceedings of the 25th ACM International on Conference on Information

66. Zhang, Jiawei, Limeng Cui, Yanjie Fu, and Fisher B. Gouza. "Fake news detection with deep diffusive network model." arXiv preprint arXiv:1805.08751 (2018).
67. Popat, K. K. (2019). Credibility analysis of textual claims with explainable evidence.
68. Shah, A., Chopade, M., Patel, P., Patel, P. (2022). Survey: Emotion Recognition from Text Using Different Approaches.Futuristic Trends in Networks and Computing Technologies . Lecture Notes in Electrical Engineering, vol 936. Springer, Singapore. https://doi.org/10.1007/978-981-19-5037-7_31
69. Kanal Bhadresh Soni, Madhuri Chopade, Rahul Vaghela. "Credit Card Fraud Detection Using Machine Learning Approach". Applied Information Systems and Management (AISM) Volume 4, (2) 2021, hal 71-76 P-ISSN: 2621-2536 ;E-ISSN: 2621-2544; DOI: https://doi.org/10.15408/aism.v4i2.20570
70. Devlin, J., Chang, M.-W., Lee, K., Toutanova, K.: BERT: pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)

Sparse based Particle Swarm Optimization

Lalit Kumar[a]*, Manish Pandey[b], Mitul Kumar Ahirwal[c]

[a] *PhD Scholar, Dept. of CSE, MANIT Bhopal, India*
[b] *Assistant Professor, Dept. of CSE, MANIT Bhopal, India*

Abstract

Particle Swarm Optimization (PSO) is the most famous metaheuristic algorithm for optimization, inspired from swarm of species. PSO can be used in various problems related to engineering and sciences. In this study, a sparse representation based PSO (Sparse-PSO) algorithm has been presented. Comparison of proposed Sparse-PSO with Standard-PSO has been done though evaluation over several standard benchmark objective functions. Our proposed Sparse-PSO method takes less computation time and provides better solution for almost all benchmark objective functions as compared to Standard-PSO method. Execution time reduction is the advantage gained through proposed Sparse-PSO.

Keywords: Particle Swarm Optimization, Sparse Representation, Swarm Intelligence

Nomenclature

Vel	velocity direction
Pos	position of particle
r	current iteration
m	particular dimension at i^{th} particle
W	inertia weight whose value is 1
h_1 and h_2	random numbers in the interval (0, 1)
Ψ_1 and Ψ_2	positive acceleration constants
N	population of size
M	number of dimensions
R	number of iterations
P_b	personal best
G_b	global best

1. Introduction

Nature Inspired Algorithms (NIA) are the most prominent strategies developed from natural surroundings [1] and mostly used in optimization tasks. Swarm based Intelligence algorithm is one of the category of NIA which is based on the cooperative behavior of swarm members. Particle Swarm Optimization algorithm (PSO) belongs to swarm based metaheuristic algorithm and was proposed in 1995 by J. Kennedy and R. Eberhart [2], [3]. PSO is encouraged by the vibrant movement of insects, birds, fishes, etc. In the field of science and engineering, PSO has been utilized to tackle non-differentiable, non-linear, and multimodal optimization problems [4]-[7]. Drawbacks such as inadequate speed-up for reaching optimum point and unsatisfactory efficiency of PSO in some studies, motivate for further development and enhancement in PSO. PSO is a simple to-actualize algorithm and furthermore has less customizable boundaries than practically equivalent to algorithms. In recent years many improvements have been done in PSO for many application areas of optimization such as chaos based PSO [8] to improve convergence speed, diminishing population based PSO to meet the swam on the most favorable point [9], FCPSO based on balancing the diversity of location to achieve convergence [10], and many more. Nature inspired optimization techniques are also used in field of signal processing to optimize adaptive noise canceller [11], to improve the range of search space [12], to filter the noisy signals [13].

Basically, the execution of PSO starts with a randomly distributed particles (population of solutions) inside the search space. As the iterations continue, the particles move as general group towards a most favorable point [2]. To appraise the optimality of each solution (particle) fitness function is evaluated in each cycle after that updation mechanism is applied to update the location of each particle so that they reach to optimal point and helps in convergence. This process of execution is repeated many times on particles to converge at global optima. All the above steps take long computation time if fitness objective function is more complex and it is difficult to achieve significant improvement.

* Lalit Kumar
E-mail address: lalit212212@gmail.com

The main aim of this paper is to gain speed-up by reducing computation time and accomplish better efficiency with the help of proposed Sparse-PSO. The rest of the paper is organized as follows: a basic description of the Standard-PSO algorithm has been provided in Section II. The proposed Sparse-PSO algorithm has been explained in Section III. Section IV displays the experiment results for the proposed approach and compared it with results for Standard-PSO. Section V contains a summary and concluding remarks.

2. Standard Particle Swarm Optimization

Particle swarm optimization is very popular optimization algorithm as only few parameters are there in this algorithm. PSO consist of a group of particle know as swarm. PSO is an iterative method since many iterations of the procedure is to be done to achieve the optimal value. In PSO, firstly an initial population of particles is initialized by arbitrarily initializing the position and velocity vectors in the defined search space. Each particle has fitness value to check solution quality at each iteration. This fitness value is obtained through objective functions which may be of minimization or maximization type, depending upon problem. Two significant things in this algorithm needs to be noticed, personal best (Pb) solution accomplished up until now and the global best (Gb) particle which is the best solution among all individual best solutions achieved until now. The dynamic journey of the particle is controlled by its personal flying knowledge as well as the flying involvement of different particles (Pb and Gb) in the swarm. Each particle's velocity and position are updated by using equations (1) and (2) respectively.

$$Vel_i^m(r+1) = W \times Vel_i^m(r) + \Psi_1 h_1 * (P_{bi}^m(r) - Pos_i^m(r)) + \Psi_2 h_2 \times (G_b^m(r) - Pos_i^m(r)) \quad (1)$$

$$Pos_i^m(r+1) = Pos_i^m(r) + Vel_i^m(r+1) \quad (2)$$

First term in velocity updation equation dominate the influence of earlier used velocities on the recent velocity or makes the particle move is same direction with the same velocity [14]. Second term control the position of particle by returning it to previous position if the previous position is better than the current position of particle according to objective function. While third term controls the particle to follow the best particle in the swarm.

3. Proposed Sparse Particle Swarm Optimization

3.1. Proposed Initialization Process

In proposed Sparse Particle Swarm Optimization (Sparse-PSO), a particle is represented as a position vector of magnitude M, where M indicates the number of dimensions or variables. Our aim is to achieve the speed-up the computation time of algorithm. In this effort, the population of particles called swarm is arbitrarily generated within defined boundaries of size N but in the form of sparse matrix representation to achieve our aim and reaching to the optimum point rapidly. After initialization of sparse matrix based population, each particle is evaluated according to the fitness function. However, this procedure is performed in the initial phase of algorithm.

3.2. Proposed Algorithm

- In proposed method, first of all swarm of size N is initialized similar to sparse matrix as deliberated in Segment *III-A* aswell as in the flow of proposed approach described below.
- Next, test functions are used to evaluate the optimal solution for every particle of swarm set. At this point, quality of solution depends on the nature of test function that can be minimization or maximization problem.
- Afterwards, personal best and global best are selected and revised repeatedly in accordance with the quality of particles.

Inputs: Initialize parameters (N, M, R, Ψ1, Ψ2, Density)
Output: global best solution and respective position vector;

Proposed Algorithm (Sparse-PSO):
1) Initialize sparse based swarm of size *N* arbitrarily in search space by using a method 'sprand(*N*, *M*, Density)' whichgenerates an arbitrary *N* by *M* sparse matrix in which number of non-zero entries are around *N***M**Density;
2) For every particle
3) Compute fitness of particle using objective function;
4) Set personal best (P_b);
5) End of for
6) Set global best (G_b);
7) While number of iterations are not met
8) For each particle

9) Revise the velocity and location of particle by equations (1) and (2);
10) Determine fitness of new revised particle;
11) Update personal best;
12) End of for
13) Hence, update global best;
14) End of while loop

4. Experiment and Results

4.1. Benchmark Functions

The efficacy of the proposed Sparse-PSO algorithm is verified using several experiments conducted on fourteen benchmark functions taken from [15] with different characteristics. The functions used to test our proposed method have been listed in Appendix *A.* Table 1 represents the bound ranges and global minimum value of benchmark functions where dimension, problem domain size, and optimal solution are denoted by M, $Lb \leq x_i \leq Ub$, and $f(X^*)$ respectively. Out of these functions, the Exponential function (f_1), the Sphere function (f_2), and the Step function (f_3) are unimodal in nature whereas the Ackley function (f_4), the Periodic function (f_5), the Quartic function (f_6), and the Qing function (f_7) are multimodal in nature.

Table 1: -Detail of Benchmark Functions

Name of Functions	Range	Global Minima
Exponential	$-1 \leq x_i \leq 1$	$f_1(X^*) = -1$
Sphere	$-100 \leq x_i \leq 100$	$f_2(X^*) = 0$
Step	$-100 \leq x_i \leq 100$	$f_3(X^*) = 0$
Ackley	$-32 \leq x_i \leq 32$	$f_4(X^*) = 0$
Periodic	$-10 \leq x_i \leq 10$	$f_5(X^*) = 0.9$
Quartic	$-1.28 \leq x_i \leq 1.28$	$f_6(X^*) = 0 + noise$
Qing	$-500 \leq x_i \leq 500$	$f_7(X^*) = 0$

4.2. Parameter Setup

The basic parameter settings are defined in Table 2. All experimental results of the algorithms namely Standard-PSO and Sparse-PSO are collected from 25 independent runs, each involving 2000 iterations.

Table 2: - Parameter Setting for Algorithms

Parameter	PSO	Sparse-PSO
Population Size (*N*)	150	150
Dimension (*M*)	30	30
Acceleration factor ($\Psi 1$ and $\Psi 2$)	1.5	1.5
Density	-	0.51
R (total number of iterations)	2000	2000
Max (total number of runs)	25	25

4.3. Experiment Results

In this work, swarm size is taken as 150 and dimension as 30 for all the functions. Both the mentioned algorithms are implemented independently 25 times for 2000 iterations on each problem of benchmark function. The data achieved from 25 independent runs are given in Tables 3 and 4. Table 3 which shows a correlation of the computation time as mean time, best time, and worst time taken by both the algorithms during each run. By analyzing the outcomes of Table 3, it very well may be inferred that our proposed Sparse-PSO algorithm has a far superior execution time compared with the Standard-PSO algorithm. Also,

Table 4 shows the performance achieved by Sparse-PSO and the Standard-PSO in the terms of mean (average of global values), best (minimum in global values), worst (maximum in optimum values), and std. (standard deviation between global values).

Fig. 1 represents the mean of best computation time for all the functions used in this work and also concludes that proposed Sparse-PSO takes less execution time overall with respect to Standard-PSO. Convergence graphs for each function are designed in Fig. 2 where the horizontal axis shows the number of iterations and the vertical axis represents the costs of each benchmark function for all iterations. It can be seen from Fig. 2(a) that for Exponential function start point of proposed method is better than Standard-PSO. Next, both the methods converge to optimal point in almost same manner. Fig. 2(b) represents the convergence graph for Sphere function which shows that start point as well as convergence curve of proposed method is better than Standard-PSO. Convergence graph for Step function are almost equal for both the algorithms is reported in Fig. 2(c). Fig. 2(d) demonstrates that in Sparse-PSO method the graph for Ackley function reaches global optimum point while in Standard-PSO approach the graph converges earlier before reaching the global optimum point. Sparse-PSO performs well for Periodic function presented in Fig. 2(e) as it converges firstly than Standard-PSO. Figures 2(f) and 2(g) determine that when the number of iterations is low the Sparse-PSO method enhances the results as compared to the Standard-PSO and later on both methods are showing similar results when number of iteration is high. It can be concluded from Fig. 2 that the proposed Sparse-PSO algorithm starts with better function value compared to Standard-PSO algorithm for all the functions and reaches to the global optimum.

Table 3: - Comparative Analysis of Standard-PSO and Sparse-PSO in terms of Computation Time (in Sec.)

Functions		**Standard-PSO Time (Sec.)**	**Sparse-PSO Time (Sec.)**
Exponential	Mean	16.2440	**12.4108**
	Best	14.2552	**12.0264**
	Worst	18.2122	**15.6286**
Sphere	Mean	17.9444	**12.0551**
	Best	15.3946	**11.6846**
	Worst	22.3657	**15.0164**
Step	Mean	13.1506	**11.8242**
	Best	12.7770	**11.2875**
	Worst	13.7343	**13.4691**
Ackley	Mean	20.1344	**13.2763**
	Best	18.2952	**11.9703**
	Worst	22.0618	23.0440
Periodic	Mean	15.5580	**11.9585**
	Best	12.9640	**11.4463**
	Worst	19.0293	**15.5228**
Quartic	Mean	15.6129	**13.4802**
	Best	14.6902	**13.3035**
	Worst	16.6461	**14.1285**
Qing	Mean	17.4904	**11.8417**
	Best	12.8546	**11.5059**
	Worst	20.3011	**12.6486**

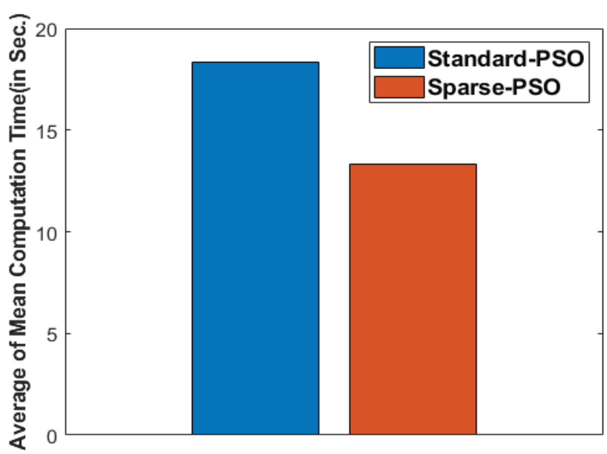
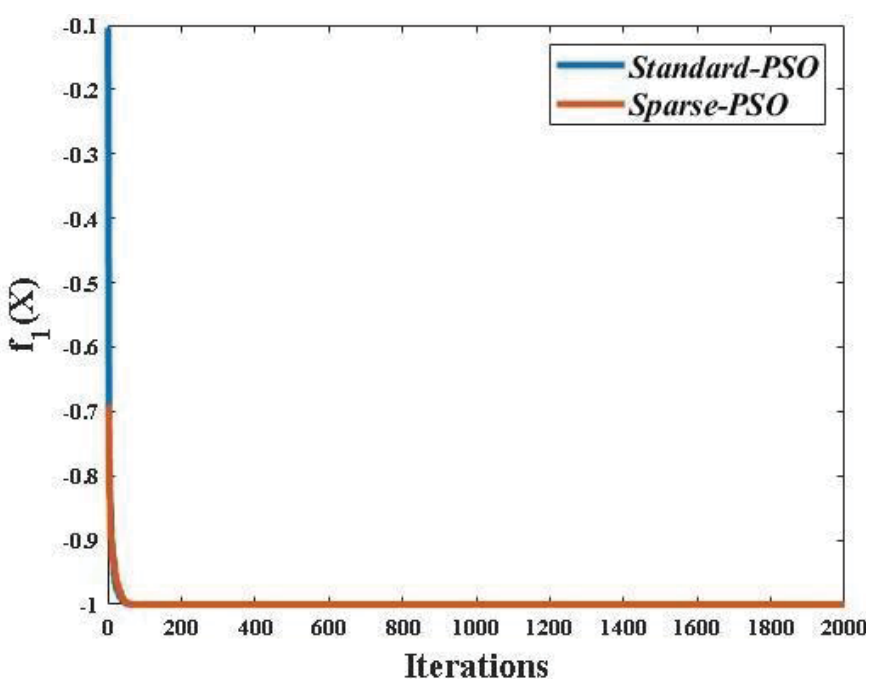


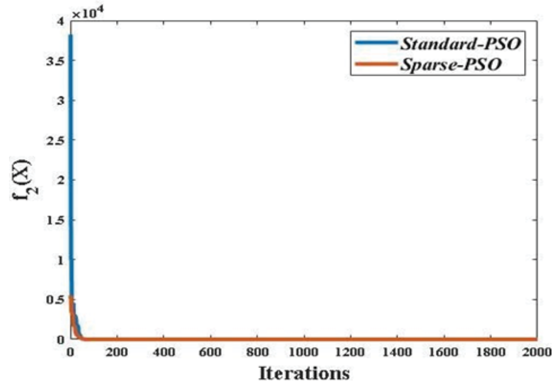
Fig.1. Overall average computation time for all functions

Table 4: - Comparative Analysis of Standard-PSO and Sparse-PSO in terms of performance

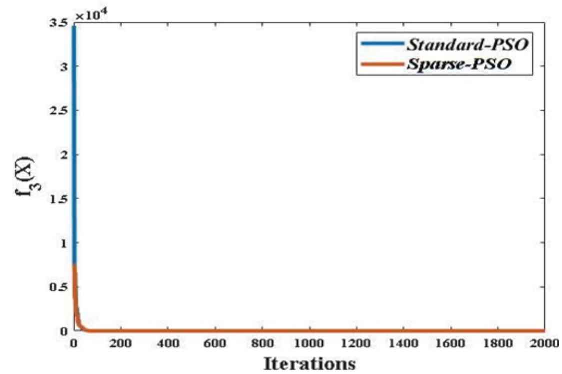
Functions		**Standard-PSO Final *f(X)* value**	**Sparse-PSO Final *f(X)* value**
Exponential	Mean	-1.0000	-1.0000
	Best	-1.0000	-1.0000
	Worst	-1.0000	-1.0000
	Std.	$8.2927e^{-16}$	$5.8922\ e^{-16}$
Sphere	Mean	$1.0302\ e^{-20}$	$5.3397\ e^{-22}$
	Best	$9.0677\ e^{-30}$	$6.9299\ e^{-33}$
	Worst	$2.3039\ e^{-19}$	$1.1973\ e^{-20}$
	Std.	$4.5988\ e^{-20}$	$2.3922\ e^{-21}$
Step	Mean	1.1600	1.2000
	Best	0	0
	Worst	5	5
	Std.	1.2138	1.1180
Ackley	Mean	0.8266	0.6943
	Best	$4.1478\ e^{-13}$	$6.2172\ e^{-13}$
	Worst	2.4083	2.3168
	Std.	0.7686	0.7471
Periodic	Mean	1.0000	1.0000
	Best	1.0000	1.0000
	Worst	1.0000	1.0000
	Std.	$9.0876\ e^{-16}$	$7.8382\ e^{-15}$
Quartic	Mean	0.0069	0.0054
	Best	0.0020	0.0018
	Worst	0.0187	0.0200
	Std.	0.0038	0.0040
Qing	Mean	$2.6747\ e^{-17}$	$6.4860\ e^{-17}$
	Best	$2.7593\ e^{-26}$	$3.1645\ e^{-26}$
	Worst	$4.9303\ e^{-16}$	$1.1821\ e^{-15}$
	Std.	$1.0006\ e^{-16}$	$2.4831\ e^{-16}$



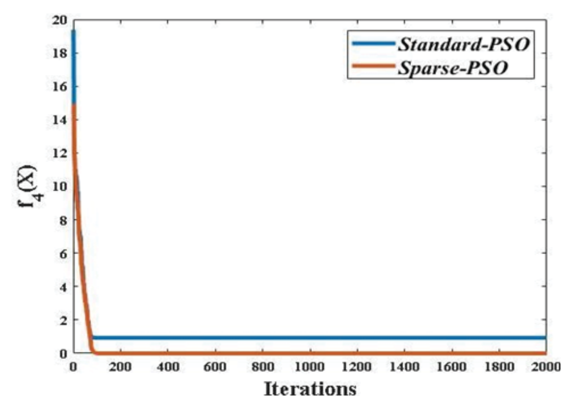
(a)


Standard-PSO
Sparse-PSO
$\times 10^4$
$f_2(X)$
Iterations

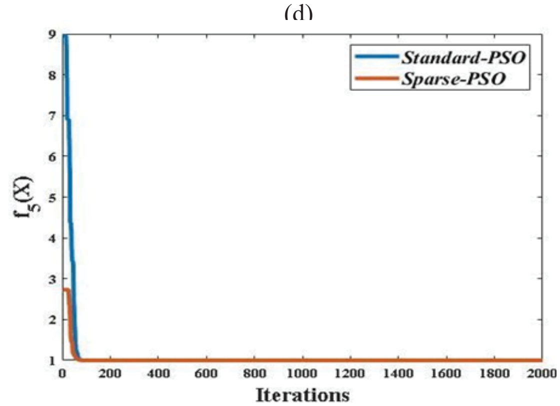
(b)


Standard-PSO
Sparse-PSO
$\times 10^4$
$f_3(X)$
Iterations

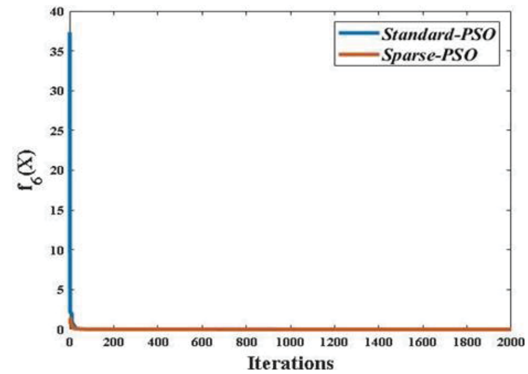
(c)


Standard-PSO
Sparse-PSO
$f_4(X)$
Iterations

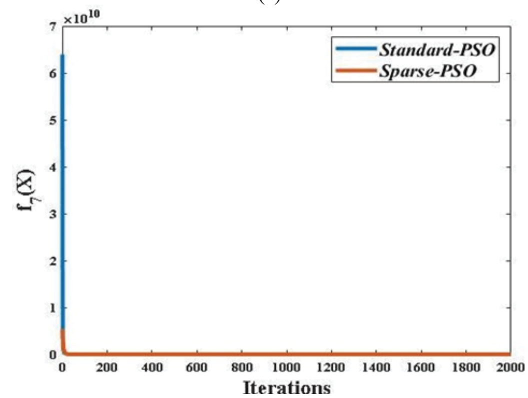
(d)


Standard-PSO
Sparse-PSO
$f_5(X)$
Iterations

(e)



(f)



(g)

Fig.2. Convergence graphs of Sparse-PSO and Standard-PSO for seven test functions (a) Exponential, (b) Sphere, (c) Step, (d) Ackley, (e) Periodic, (f) Quartic, (g) Qing

5. Conclusion and Future Directions

A new initialization technique on the basis of sparse representation has been proposed for PSO in this paper. The main objective of the proposed Sparse-PSO algorithm is to reduce the computation time. Seven benchmark functions as listed in Segment IV-A have been used to test the proposed method, and the outcomes of the same are compared with the standard approach. It is clear from results that the Sparse-PSO takes less computation time as compared to the Standard-PSO. Sparse- PSO over performs Standard-PSO in terms of execution time with an average speed-up of 24%, 33%, 10%, 34%, 23%, 13%, and 32% for the exponential function, Sphere function, Step function, Ackley function, Periodic function, Quartic function, and Qing function respectively. The proposed method achieves the global best value with low computation time compared to Standard-PSO. The future direction of this work would be to achieve more improvement by correcting and updating the used parameters and performing this method on more complex functions.

Appendix A: Benchmark Functions

$$f_1(X) = -\exp\left(-0.5\sum_{i=1}^{M} x_i^2\right)$$

$$f_2(X) = \sum_{i=1}^{M} x_i^2$$

$$f_3(X) = \sum_{i=1}^{M}\left(\left\lfloor \left(x_i+0.5\right)^2 \right\rfloor\right)$$

$$f_4(X) = \exp(1) - 20*\exp\left(-0.2\sqrt{\frac{1}{M}\sum_{i=1}^{M} x_i^2}\right) - \exp\left(\frac{1}{M}\sum_{i=1}^{M}\cos\left(2\Pi x_i\right)\right) + 20 \quad f_5(X) = 1 + \sum_{i=1}^{M}\sin^2(x_i) - 0.1e^{\left(\sum_{i=1}^{M} x_i^2\right)}$$

$$f_6(X) = \sum_{i=1}^{M} i x_i^4 + random[0,1)$$

$$f_7(X) = \sum_{i=1}^{M}\left(x_i^2 - i\right)^2$$

References

1. Kennedy, “Particle swarm optimization,” Encyclopedia of machine learning 2011, pp. 760-766, Springer US, 2011.
2. F. Marini, and B. Walczak, “Particle swarm optimization (PSO). A tutorial,” Chemometrics and Intelligent Laboratory Systems, 2015 Dec 15, 149:153- 65.
3. Y. Shi, “Particle swarm optimization: developments, applications and resources,” Proceedings of the 2001 congress on evolutionary computation Vol. 1, pp. 81-86, IEEE, 2001.
4. R. Poli, “Analysis of the publications on the applications of particle swarm optimisation,” Journal of Artificial Evolution and Applications, 2008.
5. K. R Harrison, A. P. Engelbrecht, and B. M. Ombuki-Berman, “Self-adaptive particle swarm optimization: a review and analysis of convergence,” Swarm Intelligence, 12(3), 187-226, 2018.
6. Z. Liu, J. Lu, and P. Zhu, “Lightweight design of automotive composite bumper system using modified particle swarm optimizer,” Composite Structures, 140, 630-643, 2016.
7. D. Gao, X. Li, and H. Chen, “Application of improved particle swarm optimization in vehicle crashworthiness,” Mathematical problems in Engineering, 2019.
8. B. Soudan, and M. Saad, “An evolutionary dynamic population size PSO implementation,” 3rd International Conference on Information and Communication Technologies: From Theory to Applications, pp. 1-5, IEEE, 2008.
9. A. Sahu, S. K. Panigrahi, and S. Pattnaik, “Fast convergence particle swarm optimization for functions optimization,” Procedia Technology, 4, 319-324, 2012.
10. M. K. Ahirwal, A. Kumar, and G. K. Singh, “Study of ABC and PSO algorithms as optimised adaptive noise canceller for EEG/ERP,” International Journal of Bio-Inspired Computation, 8(3), 170-183, 2016.
11. M. K. Ahirwal, A. Kumar, and G. K. Singh, “Improved range selection method for evolutionary algorithm based adaptive filtering of EEG/ERP signals,” Neurocomputing, 144, 282-294, 2014.
12. M. K. Ahirwal, A. Kumar, and G. K. Singh, “Sub-band adaptive filtering method for electroencephalography/event related potential signal using nature inspired optimization techniques,” IET Science, Measurement & Technology, 9(8), 987-997, 2015.
13. Y. Shi, and R. Eberhart, “A modified particle swarm optimizer,” International conference on evolutionary computation proceedings. IEEE world congress on computational intelligence, Cat. No. 98TH8360, pp. 69-73, IEEE, 1998.
14. J. Momin, and X. S. Yang, “A literature survey of benchmark functions for global optimization problems,” Journal of Mathematical Modelling and Numerical Optimization, 4(2), 150-194, 2013.

CBIR using Fuzzy C Means Clustering with Stacked Auto-Encoder

K.Nalini Sujantha Bel[a]*, Dr.I.Shatheesh Sam[b]

[a]*Research Scholar, Reg No. 18123112162001, Dept. of Computer Science,Nesamony Memorial Christian College affiliated to ManonmaniamSundaranar University,Abishekapatti, Tirunelveli 627 012,Tamil Nadu, India*
nalinisb85@gmail.com
[b]*Associate Professor, Dept. of PG Computer Science, Nesamony Memorial Christian College affiliated to Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli 627 012, Tamil Nadu, India*
shatheeshsam@yahoo.com

Abstract

This Image Retrieval (IR) system retrieves similar information from the cloud using the features of the corresponding image. In this proposed model, color features such as auto-correlogram and histogram are extracted from the images. The position scrambling method is used to encipher the extracted features to guarantee the security of the system in the cloud. After encrypting the features, then cluster the features by using Fuzzy C Means (FCM) clustering algorithm. In order to increase the retrieval efficiency, the proposed CBIR system update the center or centroid value of the cluster by using stacked auto-encoder. This proposed model outscored traditional approaches with the precision of 97.05%.

Keywords: CBIR; Auto-correlogram; Color moment; FCM; Auto-encoder

1. Introduction

The speedy advancement of multimedia applications, the growth of images is increased each and every day. It results the user need a big storage service and image retrieval service. (1)(2)(3)Storing of information in cloud recover the physical storage and it dispenses the user friendly resources. (4)(5)(1) The process of secure the information in cloud is a tough task. To maintain the security, the information in the cloud is encrypted to avoid accessing the confidential data given in the image by unauthorized person. (3)Cloud offers very good services for image retrieval.

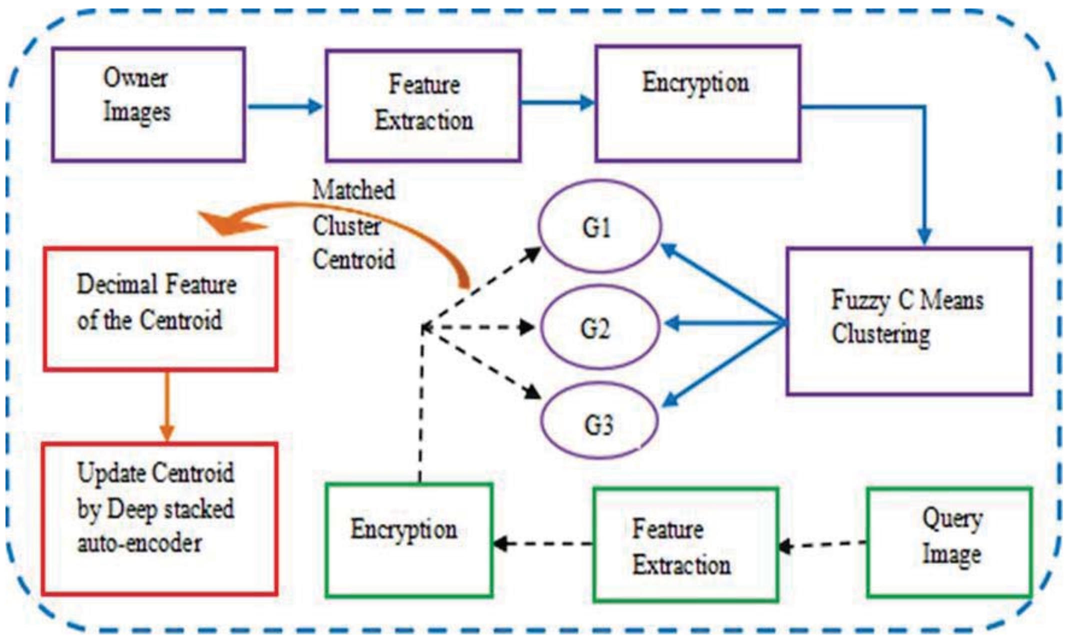
In olden days, text based retrieval plays a major role for retrieving similar images. In the text based retrieval (6), the search and retrieval process is done based on the text like keywords or tags which are related with the particular image. In content based retrieval(7)(8)(9), the similar images are retrieved by the features which are extracted from the images.(10)(11) Color, shape, and texture are taken as low-level feature vectors from the images. In (12),the author defined a image retrieval method which extracts feature vectors by using Otsu's threshold technique. It extracts the features from the encrypted images. In (13), the author proposed an unsupervised clustering based image retrieval method named as Black Hole Entropic Fuzzy Clustering (BHEFC). Tversky index is used to do the matching process between the images in the clusters and the query image.

The author of (14) proposed a feature extractor based on LSB Elimination from encrypted images. (15)Before being send the data to the cloud database, the original or authentic images are enciphered. The retrieval precision rate is poor in this study, which is a disadvantage. Scale invariant feature transform (SIFT) and Oriented FAST Rotated and BRIEF (ORB) algorithms are used to extract feature vectors in (16)(17). For indexing, the min - hash approach is utilized. In the image retrieval strategy, different clustering models are used. Centroid-based clustering is the most widely used clustering model. It increases the CBIR system's retrieval performance.

The rest of this paper is divided into 3 parts. The suggested feature extraction technique is illustrated in Section 2. Section 3 delineates the result analysis. The outcome of the suggested system is dispensed in Section 4 and Section 5 brings the proposed work to close.

* K.Nalini Sujantha Bel
E-mail address: nalinisb85@gmail.com

2. Proposed CBIR Method



.

Fig. 1 Structural design of the proposed work

The intended work's structure is portrayed in Fig. 1. The Image owner, the cloud server and user are the three entities used in this CBIR architecture. The original image's color information is extracted by the image's owner. Fuzzy C Means Clustering is used to group the retrieved features. Then the features are encrypted by position scrambling method and send to the cloud server hosting CBIR. Additionally, the user submits the cloud with the query request. In the cloud, CBIR matches the query data with the information in the dataset. Matching process is done using Euclidean distance measure. To update the centroid of the cluster by deep stacked auto-encoder.

2.1 Feature Extraction

The major task of CBIR is extracting features from the image. Different color feature values are extracted from the image in the proposed approach.

2.1.1 Color Auto-correlogram

The spatial information in an image is extracted using an auto-correlogram. It extracts the spatial information by using the color histogram. Based on the adjacent pixel, auto-correlogram is applied. The color of an image is H quantized into I colors such as $d_1, d_2, \ldots, d_I$. The color correlogram of an image H is denoted as in Eq.1.

$$\eta^{(l)}_{d_n, d_0}(H) \triangleq \underset{f_1 \in H_{d_0} \& f_2 \in H}{P} \left[f_2 \in H_{d_0} \,\|\, f_1 - f_2 = l \right] \tag{1}$$

Where l is the distance based on the pixel f. The auto-correlogram of an image H is described as in Eq.2.

$$\gamma^{(l)}_{d}(H) \triangleq \eta^{(l)}_{d_k, d_0}(H) \tag{2}$$

2.1.2 Color Histogram

The arrangement of different color information in the image H is illustrated by histogram of color. The histogram probability distribution is defined as in Eq.3.

$$P(q) = \frac{P_q}{z} \tag{3}$$

Consider P_q shows the total gray pixels, and z shows the pixel summation of an image. The color histogram of an image H is defined as the combination of probabilities of the histograms of red, green and blue colors as shown in Eq.4.

$$C_H = \{P^r(q) \,||\, P^g(q) \,||\, P^b(q)\}. \tag{4}$$

2.2Encryption and Decryption

Then feature vectors are encrypted by using the position scrambling method. Here, the position of the feature vectors is scrambled dependent on the key's random sequence generation. The key K is a 16 bit binary number used to scramble the columns of the feature vector. For decryption, the original vector is reconstructed from the encrypted vector by using the same key which is used in encryption. By using the random vectors, the enciphered image's position is rearranged. The user can see the real retrieved image, which has been decrypted.

2.3 Clustering using FCM

Clustering would be beneficial for minimizing the time it takes to search the database for images. In Fuzzy C Means (FCM) clustering method that allows single piece of information that belongs to more than two clusters. In this clustering, each data point is clustered based on the membership function. Thus, data sources on the boundary of a cluster may be in the cluster which is less than the centroid. FCM divides the data into a predetermined number of clusters, with every one centroid serving as a vector for inserting the pertinent images in the particular cluster group. The representative vector or centroids initiated by the FCM is interpreted as in Eq.5,

$$M = \{M_1, M_2, M_3\} \tag{5}$$

2.4 Centroid updation in matched cluster

The numeric value of M_1 is assessed to upgrade the centroid M_1 once the new training image X has been matched with it. The decimal vector of M_1 is, nevertheless, shown as M_1^d . The fresh training image's weight γ and decimal form X_e^d are utilised to update the centroid, and the resulting mathematical model is represented by Eq. 6.

$$M_1^{new} = bina\left(X_e^d * (1-\gamma) + M_1^d * \gamma\right) \tag{6}$$

The Euclidean distance metric is used to index each new training image, and if the image matches any of one the centroid reading, the particular centroid value is upgraded. In this case, stacked auto-encoder is used to calculate the weight γ .

Structural design of Stacked auto-encoder

It is an unsupervised network model called as stacked auto-encoder (18)(19). The three layers in the network are input, hidden, and output. The encoder and decoder operations are the two steps of the stacked auto-encoder. It transforms input data into representations of hidden layer data, while the decoder reconstructs the original data from the unseen or hidden representations. Fig. 2 portrays the architectural design of stacked auto-encoder.

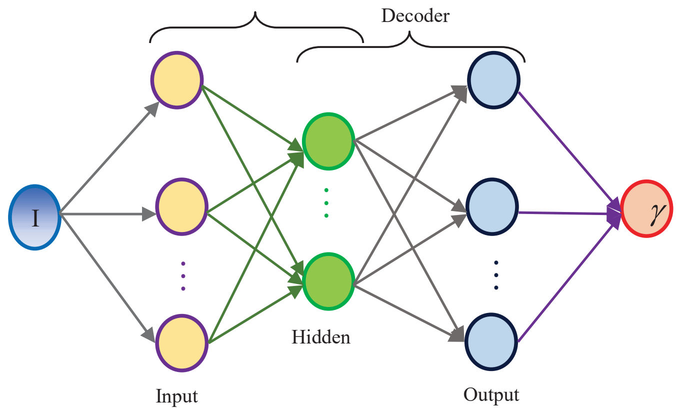


Fig. 2 Structural design of Stacked auto-encoder

This Auto-encoder model takes I as input such that I contains $I = \{M_1^d, X_e^d\}$, where M_1^d specifies the decimal value of centroid M_1, and X_e^d denotes the enciphered non-indexed training image of X_e. This mean that both M_1^d and X_e^d are the input of the auto-encoder model. The encoding operation is expressed as in Eq.7,

$$s = p(w_1 I + z_1) \tag{7}$$

where, p specifies encoding function, w_1 indicates the weight vector, z_1 is the bias value, and s represents the encoder output. However, the decoder operation is described as in Eq.8,

$$Y = q(w_2 s + z_2) \tag{8}$$

where, q signifies the decoding function, w_2 describes the weight vector of decoder, z_2 is the bias value, and Y denotes decoder output. At last, the result of using a stack auto-encoder is assigned as γ .

3. Analysis of Experimental Findings

The criteria: (i) precision (ii) retrieval time of images (iii) feature extraction time are used to assess the achievement of the defined CBIR system. It is differentiated by the existing techniques SIFT+ORB(16)and LSB Elimination(14).

3.1 Metric 1-Precision

Precision value is determined by dividing the k number of accurately retrieved confident images C_k by C number of correctly retrieving related images which is represented as in Eq.9.

$$P = \frac{C_n}{C} \tag{9}$$

Fig.3 and Table1 shows the precision of diverse k number of values including 20,40,60,80 and 100. The projected strategy looks to have a higher precision value than the more traditional methods like LSB and SIFT+ORB.

Table 1. Precision Values

K	LSB	(SIFT+ORB)	Proposed FCM + Auto-encoder
20	37.1	95.2	97.05
40	30.8	91.6	86.23
60	26.9	89.2	84.76
80	23.1	86.5	83.32
100	21.7	80.5	82.95

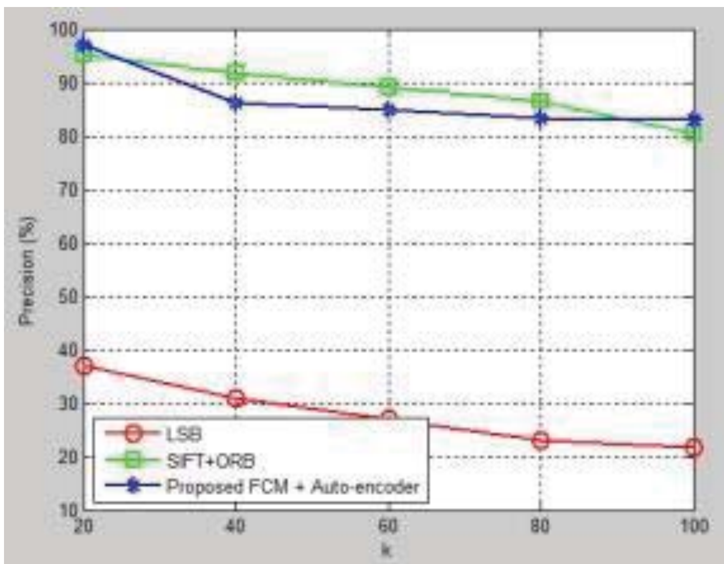


Fig. 3 Precision comparison chart

3.2 Time taken for Extracting Features (F_t)

It specifies how long it took to extract the features F from the images N in the dataset. When the total number of images in dataset N rises, the F_t value also rises. Table 2 exhibits the features extracting time for various number of images in the dataset.

Table 2. Feature Extraction Time

Feature Descriptor	Images in the Dataset				
	$2x10^3$	$4x10^3$	$6x10^3$	$8\ x10^3$	$10x10^3$
LSB	282s	492s	712s	1001s	1486s
SIFT+ORB	331s	673s	1007s	1348s	1645s
Proposed FCM	310s	520s	833s	1021s	1265s

3.2 Metric 2 - Time of Image Retrieval (R_t)

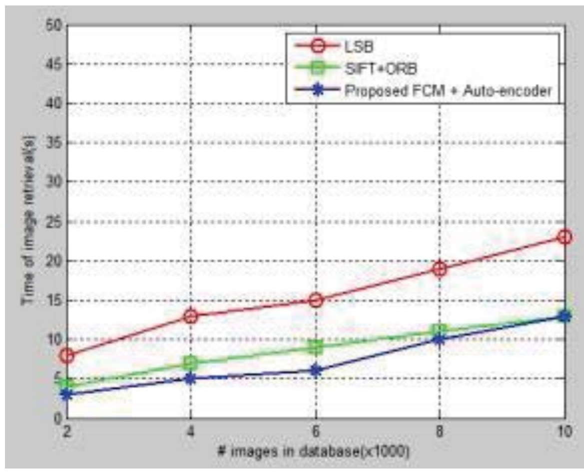


Fig. 4. Comparison of Image Retrieval

R_t determines how long it takes to obtain a collection of identical images C from a dataset. The time it takes to retrieve images is determined by the size of the image database N . Fig.4 projects the time taken for retrieving similar images. The retrieval times of proposed work is slighter than the ordinary SIFT + ORB and LSB methods.

4. Result Analysis

The proposed CBIR procedure is implemented using Matlab and be verified with the help of images in the Corel 10K dataset [11]. Fig. 5 and Fig. 7 depicts the user request query and Fig. 6 and Fig. 8 illustrates the retrieved result of the user request.

Fig. 5 Sample Image

Fig. 6 Received output

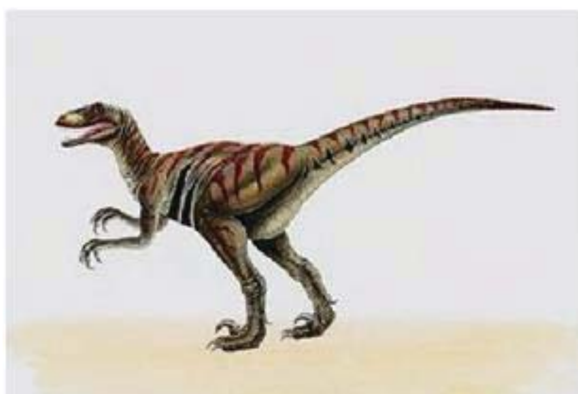


Fig.7 Sample Image

Fig. 8 Received Output

5. Conclusion

CBIR using Fuzzy C Means Clustering with deep stacked centroid updation is proposed in this work. The proposed method is divided into four stages. The color features of the original images are first carefully extracted. After that, the obtained features are sent to the encryption phase. It is encrypted using position scrambling method and then the clustering process is carried out with the Fuzzy C Means clustering. A new training image or query image is matched with the centroid by using Euclidean distance metric. The image matches the centroid, is in the cluster group, and the decimal value of particular group is determined in order to upgrade the centroid value of the same cluster. The proposed Fuzzy C Means Clustering with deep stacked centroid updation method exceeded the existing techniques by attaining precision of 97.05%, and time of image retrieval is 3s. The future work on this study will focus on using another deep learning classifier to index and retrieve data more efficiently.

References

1. Liu D, Shen J, Xia Z, Sun X. A content-based image retrieval scheme using an encrypted difference histogram in cloud computing. Inf. 2017;8(3):1–13
2. Chow R, Golle P, Jakobsson M, Shi E, Staddon J, Masuoka R, et al. Controlling data in the cloud: Outsourcing computation without outsourcing control. Proc ACM Conf Comput Commun Secur. 2009;(May 2014):85–90
3. Jadhav AG, Sawant N, Pingale S. Implementation on Privacy-Preserving Content-Based Image Retrieval in Cloud Image Repositories. 2023;10(01):3460–71.
4. Xia Z, Xiong NN, Vasilakos A V., Sun X. EPCBIR: An efficient and privacy-preserving content-based image retrieval scheme in cloud computing. Inf Sci (Ny). 2017;387:195–204
5. Ferreira B, Rodrigues J, Leitao J, Domingos H. Practical Privacy-Preserving Content-Based Retrieval in Cloud Image Repositories. IEEE Trans Cloud Comput. 2017;13(9):1–14
6. Wilkins P, Ferguson P, Smeaton AF, Gurrin C. Text based approaches for content-based image retrieval on large image collections. IET Semin Dig. 2005;2005(11099):281–8
7. Kanaparthi SK, Raju USN. Content based image retrieval on big image data using local and global features. Int J Inf Technol. 2022 Feb 1;14(1):49–68
8. Mathan Kumar B, PushpaLakshmi R. Multiple kernel scale invariant feature transform and cross indexing for image search and retrieval. Imaging Sci J. 2018;66(2):84–97.
9. Raju USN, Suresh Kumar K, Haran P, Boppana RS, Kumar N. Content-based image retrieval using local texture features in distributed environment. Int J Wavelets, Multiresolution Inf Process. 2020;18(1)
10. Alsmadi MK. Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features. Arab J Sci Eng. 2020;45(4):3317–30
11. Ahmed KT, Ummesafi S, Iqbal A. Content based image retrieval using image features information fusion. Inf Fusion. 2019;51:76–99.
12. Nalini Sujantha Bel K, Shatheesh Sam I. OT-Feature Extraction on Scrambled Images with Instantaneous Clustering for CBIR Scheme in Cloud Computing. ISeCure. 2021;13(1):1–17
13. Bel KNS, Sam IS. Black hole Entropic Fuzzy Clustering-based image indexing and Tversky index-feature matching for image retrieval in cloud computing environment. Inf Sci (Ny) [Internet]. 2021;560:1–19. Available from: https://doi.org/10.1016/j.ins.2021.01.043
14. Nalini Sujantha Bel K, Shatheesh Sam I. LSB Elimination based feature extraction for outsourced image retrieval in encrypted images. Proc Int Conf Trends Electron Informatics, ICOEI 2019. 2019;2019-April(Icoei):130–5.
15. A privacy-preserving cross-media retrieval on encrypted data in cloud computing - ScienceDirect [Internet]. [cited 2023 Apr 10]. Available from: https://www.sciencedirect.com/science/article/abs/pii/S221421262300025X
16. Bel KNS, Shatheesh Sam I. Encrypted image retrieval method using SIFT and ORB in cloud. 2020 7th Int Conf Smart Struct

Syst ICSSS 2020. 2020;4–8.

17. Xia Z, Zhu Y, Sun X, Qin Z, Ren K. Towards Privacy-Preserving Content-Based Image Retrieval in Cloud Computing. IEEE Trans Cloud Comput. 2018;6(1):276–86
18. Liu G, Bao H, Han B. A Stacked Autoencoder-Based Deep Neural Network for Achieving Gearbox Fault Diagnosis. Math Probl Eng. 2018;2018.
19. Bel KNS, Sam IS. Incremental indexing with binary feature based Tversky index using black hole entropic fuzzy clustering in cloud computing. 2022

A Transfer Value Approach to Analyse Football Players

Rashmi Gandhia[a], Sudhanshu Ranjan[b]*, Nikunj Agarwal[c], Anupam Sharma[d]

[a,b,c,d]CSE, ASET, Amity University U. P. Noida, India

Abstract

The world of football transfers has often troubled experts of both football and economics alike. Player valuations and their transfer fees are determined opaquely and the factors that goes into determining them are highly irregular and sporadic. In this work, factors are computed that determine what transfer fee a player is worth. Also found out if these factors limited to on-field metrics, what these metrics are and the extent of their influence. The dealings of football transfers have always been seen as being ambiguous, unpredictable, and uncharacterizable. This work aims to find whether these suppositions are indeed true or are definitive factors that determine a player's worth. As investigation goes with the accuracy with which we can estimate a player's present valuation (a player's market worth), as of September 1st 2021, machine learning algorithm to provide answers to these concerns. Here, metrics identify best translate into a higher transfer value for players. Six regression models are tested and do prediction the transfer values of players based on the aforementioned metrics.

Keywords: Machine learning; Regression; Football Transfers; Sports Econometrics

1. Introduction

As we all know that how the sports industry is churning a large amount of money nowadays & Football is the biggest shark in that money pool but what about the players who play these sports. How does that value of each player taken out, like on what basis. In this work, net value of the player deserves and how much valuable is he considering all the factors which are there to be looked upon [1-4].

Transfers are an integral part of the club football, where the players are transferred from one team to another, where the biggest of clubs exchanges their player for an enormous amount of money to sign the player which they want the most. In 2019, It was revealed by the world's biggest football organising body FIFA that the clubs spent a total of $7.30billion in the transfer market [7,8]. It is like a business transaction which happens between two different clubs. It doesn't matter if the club doesn't even play in the same league or in the same country where a player moves from one to another club according to the deal. The transfer fee differs for each player it does depend on many factors like commercial value, potential worth if the player in the future, quality if the player, contract time left with current club, performance of the player, etc.

A transfer happens only when two clubs mutually agree on the term of selling the player where the club who is buying the player creates and agrees with the new contract on which they are signing the player. It can happen when a representative from a club makes a move by making an official inquiry for the target player to the club. If the club is open to sell the player, then it is obvious that the other clubs might get involved. There are many major factors which are considered while the transfer process it is discussed with their agents and advisors. To understand it by an example like how Manchester city completed Haaland transfer. E. Haaland the Norwegian moved from the German club Borussia Dortmund in summer 2022 for $63million to the English premier league's powerhouse Manchester city where he signed a five-year deal.

2. Literature Review

The sports industry is churning a large amount of money nowadays & Football is the biggest shark in that money pool but what about the players who play these sports. How does that value of each player taken out, like on what basis and a significant amount of research has been done in this field of work.

Researchers like Carmichael & Thomas (1993) had adopted an approach called the bargaining approach. In the player transfer market, it seems there are two factors that determine transfer fees. A player's value to the selling and buying clubs can be seen on the one hand, while on the other there is the position of both the selling and the buying side in the negotiations or bargaining. Using multilevel regression techniques, the authors of determine the market values of the payers [1-3].

* Sudhanshu Ranjan
E-mail address: sudhanshuranjansr1@gmail.com

Carmichael et al. (1999) consider that intrinsic talent and expenditures in human capital are what primarily determine a player's addition to the value of the team [1]. Pujol et al evaluate collegiate athletes by a 'media value measure'[4]. Frick also includes studies looking at the quantity of Google hits to determine whether there is a "superstar impact" [5]. The impact of estimates made by the football audience on the player's transfer fee is covered [6]. The market value of football players as estimated by professional football fans is used by the authors to estimate the players' transfer fees. Using machine learning approaches, J. Hucaljuk and A. Rakipovi seek to forecast football scores [7].
According to E. Franck and S. Nüesch, reputation and ability both affect a football star's market worth. Twenty criteria were utilised to measure a player's ability in relation to his capacity to advance his team's fortunes. The likelihood that a player may sustain an injury is another element that directly affects the number of games played and, as a result, has an impact on player performance. This statistic has not, as far as we are aware, appeared in empirical study mainly because of scarcity of data [8-9].
Meanwhile, studies have shown that players who play or are approached by the more economically sound leagues are usually valued higher. Furthermore, players that play in UEFA competitions are also usually valued higher than their counterparts, although they command higher [10] salaries and even if their on-field performances are alike [11]. Van Den Berg expresses similar conclusions, that clubs with a larger buying size viz. clubs participating in UEFA competitions, pay larger values for players [12]. They also not attributed to said clubs simply having a larger purchasing power, but to a phenomenon van den berg expresses as 'risk-aversion' which is clubs buying players before they join their rivals.
According to McHale et al. however, the economic factors are still somewhat secondary when it comes to valuation of players in the market [13]. Their study suggest transfer values for players are still determined in a very rudimentary way, based on on-field performances. They however outline, not all teams succeed in the correct valuation of players and hence there are plenty of cases of clubs over-valuing on undervaluing players. Clubs, however, seemingly follow patterns in this process of undervaluing/overvaluing players and in [13], it is presented clearly that teams that succeed in correctly valuating a player's on-field attributes, succeed repeatedly in making smart purchases.

3. Methodology

3.1 Data Gathering

For obtaining the data for this research two sources which were reliable that are FBREF and Transfermarkt are used.

3.1.1 Player Performance Data

Measurement of football players performance on the field (competing in the top stages around the world) are tracked by data at fbref.com. The website's data collection began with the 2017–18 season, when it became the most complete. Prior to that, there were either fewer measures or just no data available. To maintain consistency, it was concluded that it would be appropriate to use data beginning in 2017 and ending in 2021.The EPL, La Liga, Serie A, Bundesliga, Uber Eats Ligue 1 had the most complete data, while the website does provide player performance data from other leagues across the world. As a result, it was determined that information from the preceding four seasons should be obtained for each of the top 5 leagues. We used player's statistics from prior years as we wanted to examine the idea that these metrics from previous years will significantly affect their transfer values. After collecting the data, the 140 downloaded spreadsheets were systematically combined into one.

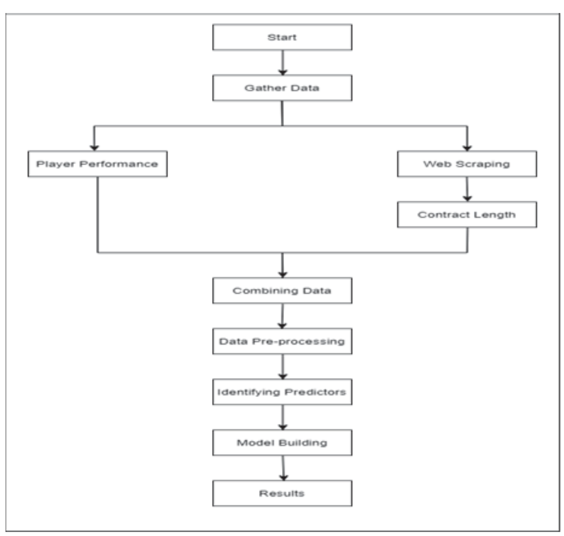


Fig 3.1 Workflow Methodology

Goalkeepers were excluded from this project due to the fact that they are evaluated using entirely different measures than outfield players. We disregarded the dataset pertaining to the performance and behaviour of the goalkeeper. A summary of the other data is as following-: Basic Stats, Finishing, Vision, Pass Types, Creativity, Defensive Actions, Possession & Other Information like additional dat for example quantity fouls they committed & drawn themselves, offside etc[14].

This marked the end of the data gathering and processing for our 1st source. Our 2nd source that is Transfermarkt.co.uk provides data about their name, group they are currently, homegrown association in which player's group contends, age, their playing position, number of years left in their agreement and the value of their transfer.

3.1.2 Web Scraping for Additional Data

The data on this website was extracted through the use of the widely known Python scraping tool BeautifulSoup. In order to gather information from all 98 different teams in the top five European leagues, we employed the 'find all()' method to obtain links to the league webpages, as compiling a list of links for each of the 98 websites would have been a time-consuming task.

Once we had identified the specific data we wished to extract, we utilized the find all() method to scan the webpages of each team in order to extract information on the contract length of individual players. This information is critical in determining a player's value. We achieved this by creating a loop to compile a list of links for every player's page on each club's website. Thankfully, we were able to extract the contract expiration dates for each player by using the find all() method on their respective webpages. Finally, we combined the contract length data with the other extracted data using the pd.merge technique. Our task of gathering data from Transfermarkt.co.uk is now complete.

3.2 *Combining FBREF and Transfermarkt data*

After completing the cleaning phase, we merged all the collected data into a single dataframe. Players who did not compete in a particular season were identified by the NaN values in their statistics metric columns. We did not want to discard the acquired data due to the presence of NaN values because such rows are excluded when fitting data into a model. To replace the NaN values with appropriate figures, we developed custom code that computes the mean statistics metrics of each action for seasons in which the player participated in the top 5 leagues.

We did not want the NaN values for a player to depend on the statistics of other players, so we avoided using a straightforward fillna() function with column averages. Instead, we created a function that replaces missing data with the mean of a player's own recorded statistics. Since our dataset had an insufficient number of goalkeeper columns to train an AI model, we excluded all goalkeepers from the final dataset with no NaN values. Therefore, goalkeepers are not considered in our study.

3.3 Data Pre-Processing

Now that we had a clean dataset that had been combined, we separated to the dataset to include only attacking players and analysed the distribution of specific predictor features that have the potential to predict the transfer values of attackers. Fig. (3.3) demonstrates this.

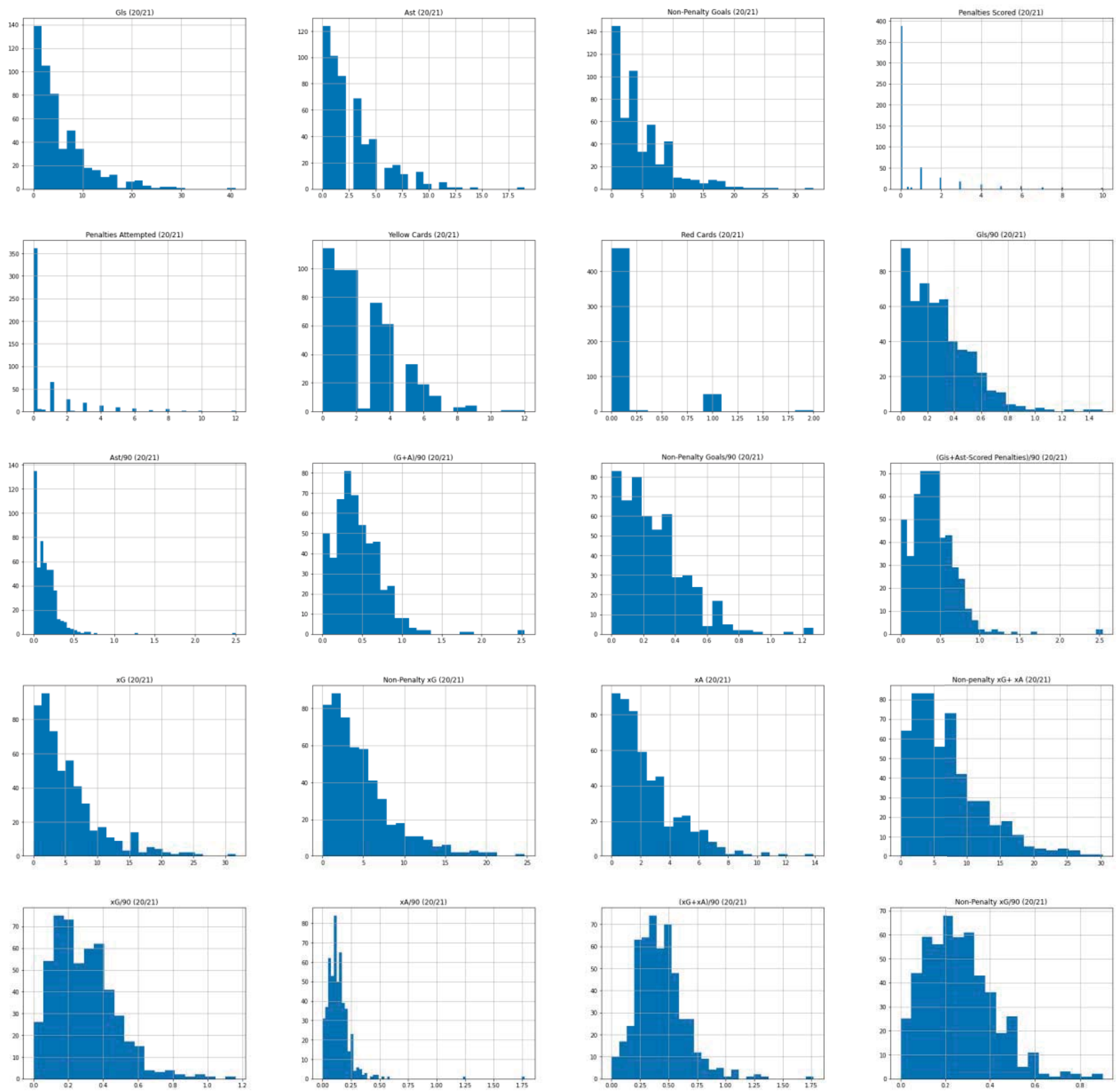


Fig. 3.3 Predictors for attackers

From our study we observed that there are probably key indicators of how valuable an attacking player are not evenly spread out in their dissemination of values. It demonstrates that before these metrics can be utilised to fit and train our models, they needed transformation to obtain a more Gaussian distribution. This was achieved using the Power Transformer from the sklearn pre-processing toolkit. We carried out standardization using a Robust Scaler. The Robust Scaler was used because it is very helpful when working with data that contain outliers.

3.4 Identifying Top Predictors for each position

The following images show some of the key characteristics that correlate to the transfer sum paid for players of all the three areas of the pitch we're observing. Each model's most crucial features were identified using the 'feature importance ' property after the data had been initially fitted into it.

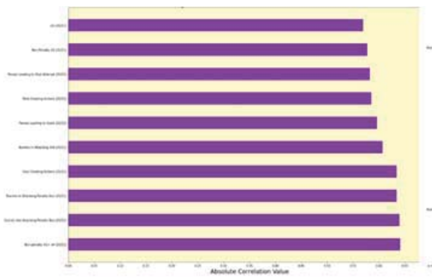


Fig. 3.4.1 Attackers

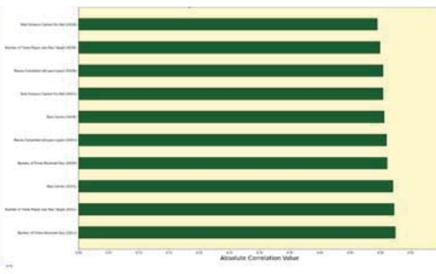


Fig. 3.4.2 Midfielders

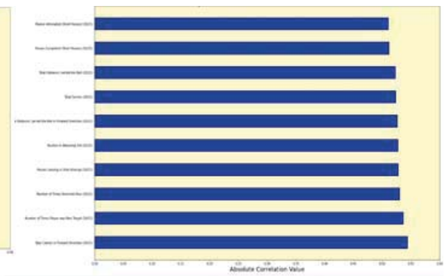


Fig. 3.4.3 Defenders

Fig. 3.4 Top predictors for different Positions

As expected for forwards from Fig. (3.4.1) a player's xG and xA, actions in the opponent's 18-yard box area and ability to create chances are the most important attributes. Similarly, in the case of midfielders Fig (3.4.2), it is no shock that they're valued based on how well they dictate the pace of play and positively assist the team's ball movement forwards. It's noteworthy to note that among the top predictors for defenders Fig (3.4.3), the quantity of tackles and interceptions made does not rank highly. Top defenders may not need to make as many tackles or interceptions since they play for stronger teams that have the ball for longer stretches of time.

3.5 Model Building

To forecast the transfer values of the players for this project, 6 different models were applied. Data had been pre-processed for each of them using the aforementioned procedures. All of our models underwent pre-processing using the aforementioned steps, and the default hyperparameters were used for the initial attempts. The entire dataset was utilized during these initial modeling attempts to determine the most important features for each model. After identifying the top 10 features, a subsequent iteration of the same model was tested using a dataset that only contained the most influential attributes as shown in Fig. (3.5). We only used the top features in the dataset for our final models. Then, after a Grid Search revealed better hyperparameters, a fresh model fitting was carried out using the newly discovered hyperparameters as well as the dataset's top attributes. Following this, a scoring metric "neg root mean squared error" was used to obtain the mean cross-validation scores.

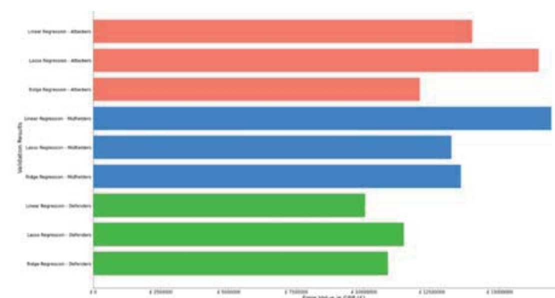


Fig. 3.5.1 Linear Regression

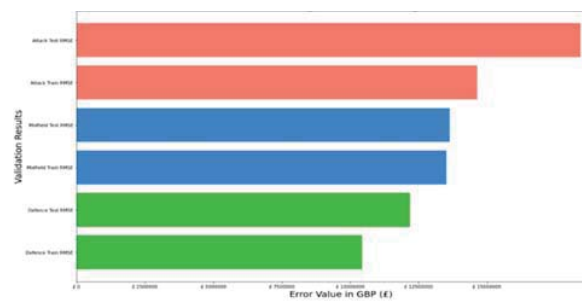


Fig. 3.5.2 Decision Tree Regressor

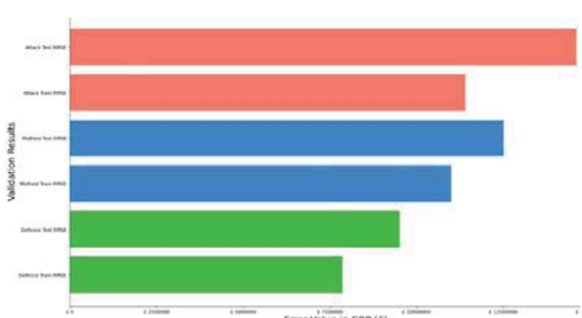


Fig. 3.5.3 Random Forest Regressor

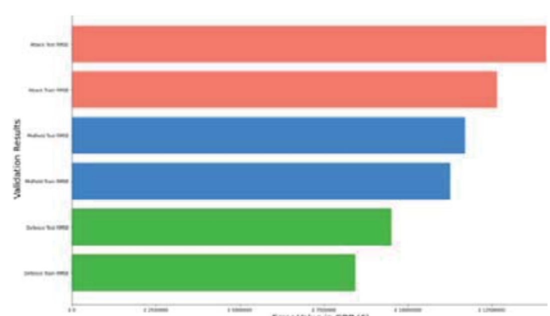


Fig.3.5.4 Gradient Boosting Regressor

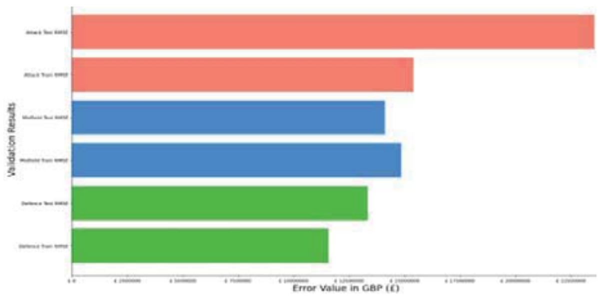


Fig. 3.5.5 Ada Boost Regressor

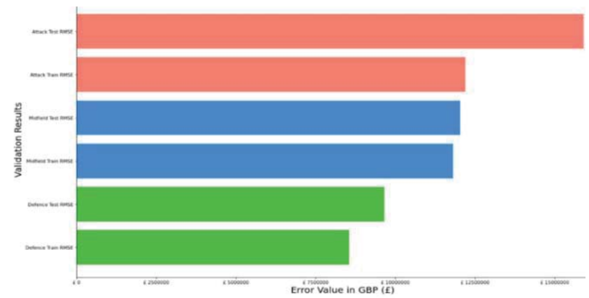


Fig.3.5.6 Support Vector Regression

Fig. 3.5 ML Models Graphs

4. Results and Discussions

Comparisons are performed between each model's results to identify which model worked best to predict player transfer values with the lowest RMSE for each zone. As seen clearly in Fig. (4.1.1), The Ridge (L2) Regression produced the lowest Root Mean Squared Error(£12m) in its predictions for attacking players. In the case of midfielders in Fig. (4.1.2), the Random Forest Regressor produced the lowest Root Mean Squared Error(£11.6m). And for Defenders, Fig. (4.1.3) shows the three ensemble models (Random Forest, Gradient Boost & Ada Boost) produced similar results but among them, the Random Forest Regressor was the model producing the lowest Root Mean Squared Error(£9.5m).
We finished our model with an error range of about £9–£12 million. The results of this experiment show that despite the fact that this model may be acceptable for larger teams looking to sign the top players, we believe that our models are inadequate for mid-tier and lower value players who are worth £20 million or less. As a result, we can infer that information on a player's performance on the field for these players alone might not be sufficient to predict the transfer worth of player with any degree of accuracy. For this to be accomplished, more data might be needed.

Fig. 4.1.1 Model Performance for Attackers

Fig. 4.1.2 Model Performance for Midfielders

Fig. 4.1.3 Model Performance for Defenders Position

Fig. 4.1 Model Performance for different Positions

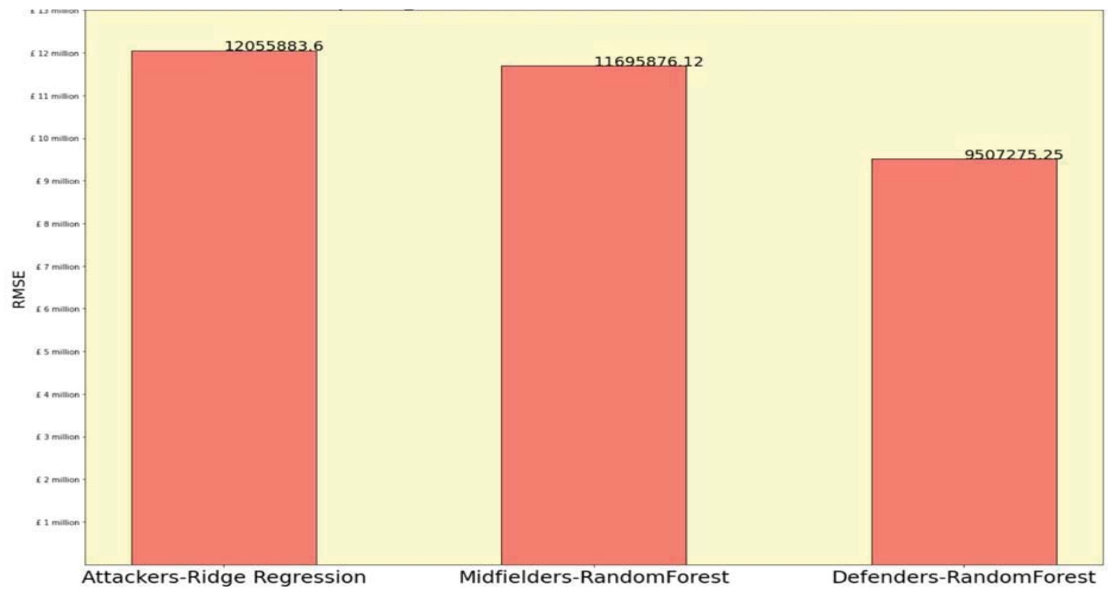


Fig. 4.2. Identifying which positions we predicted best

5. Conclusions

Through our findings, it is made evident that on-field performance, while extremely influential, is not the only predictor of a player's transfer value. That being the conclusion, it is important to have an understanding of a few restrictions that will need to be addressed given the limited scope for this project.

The first and foremost is the false presumption that the competition's top 5 leagues are all very contested. In addition, players on better teams typically have superior stats. One way of compensating for this could be through the creation of a "competitiveness score" based on leagues and opponents. Although time-consuming, this is doable because relevant data is obtainable.

Furthermore, in this project, players have been divided into positions as attackers, midfielders, and defenders. This has been done due to the sources the data has been obtained from still use these notations to compare players from different eras. Positions in modern football, however, are more nuanced. Players can further be classified according to different roles they play in a particular position (for example, attackers can be defined as centre-forward, winger, false-9 etc.). Comparisons made after the data is divided into these roles could yield more accurate results.

References

1. Carmichael, F., Forrest, D. & Simmons, R. (1999): The Labour Market in Association Football: Who gets Transferred and for how much? Bulletin of Economic Research, 51 (2), pp. 125-150
2. Carmichael, F. & Thomas, D. (1993): Bargaining in the transfer market: theory and evidence, Applied Economics, 25 (12), pp. 1467-1476
3. Muller, O., Simons, A., and Weinmann M., "Beyond crowd judgments: Data-driven estimation of market value in association football.", European Journal of Operational Research Volume 263, Issue 2, pp. 611-624 (2017).
4. Pujol, F., Garcia-del-Barrio, P. & Elizalde, J. (2007): ESIrg Report on Media Value in Football: The English Premiership Extends its Supremacy, Universidad de Navarra.
5. Frick, B. (2007): The Football Players' Labor Market: Empirical Evidence from the Major European Leagues, Scottish Journal of Political Economy, 54 (3)
6. Herm, S., Hans Markus Callsen-Bracker, HenningKreis, "When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community", Sport Management Review, Volume 17, Issue 4, pp. 484-492 (2014).
7. Hucaljuk, J, and Rakipović, A., "Predicting football scores using machine learning techniques", MIPRO 2011, Opatija, Croatia (2011).
8. Franck, E., Nüesch, S., "Talent and/or popularity: what does it take to be a superstar? "Economic Inquiry, 50 (1) pp. 202-216 (2012).
9. Tunaru, R., Clark, E. & Viney, H. (2005): An Option Pricing Framework for Valuation of Football Players, Review of Financial Economics, 14, pp. 281-295
10. Kesenne, S. Revenue sharing and absolute league quality; talent investment and talent allocation. Scott. J. Political

Econ. 2015, 62, 51–58.
11. Felipe, J.L.; Fernandez-Luna, A.; Burillo, P.; de la Riva, L.E.; Sanchez-Sanchez, J.; Garcia-Unanue, J. Money Talks: Team Variables and Player Positions that Most Influence the Market Value of Professional Male Footballers in Europe. Sustainability 2020, 12, 3709. https://doi.org/10.3390/su12093709
12. Erik van den Berg, "The Valuation of Human Capital in the Football Player Transfer Market: An investigation on transfer fees paid and received in the English Premier League", M.Sc. Thesis, Erasmus School of Economics, Erasmus Universiteit Rotterdam, July 2011.
13. Ian G. McHale, Benjamin Holmes, Estimating transfer fees of professional footballers using advanced performance metrics and machine learning, European Journal of Operational Research, Volume 306, Issue 1, 2023, Pages 389-399, ISSN 0377-2217, https://doi.org/10.1016/j.ejor.2022.06.033, https://www.sciencedirect.com/science/article/pii/S0377221722005082
14. Sindhwani, N.., Anand, R.., S., M.., Shukla, R.., Pratap Yadav, M.., & Yadav, V.. (2021). Performance Analysis of Deep Neural Networks Using Computer Vision. EAI Endorsed Transactions on Industrial Networks and Intelligent Systems, 8(29), e3. https://doi.org/10.4108/eai.13-10-2021.171318
15. Deal, B., Grove, A., 1965. General Relationship for the Thermal Oxidation of Silicon, Journal of Applied Physics 36, p. 3770.
16. Fachinger, J., 2006. Behaviour of HTR Fuel Elements in Aquatic Phases of Repository Host Rock Formations. Nuclear Engineering & Design 236, p. 54.

TerrainEye - Pothole Detection System

Abhishek Singh[a*], Prof. Savita Lohiya[b], Aaqib Farooq[c], Anupriya Selvaraj[d], Disha Auti[e]

[a,b,c,d,e] *Department Of Information Technology, SIES Graduate School Of Technology, Navi Mumbai, India*

Abstract

The recent economic boom and technological advancements have profoundly impacted traditional transportation systems. As a result, there has been a growing demand for Intelligent Transportation Systems (ITS) which aim to enhance the transportation sector, with traffic safety being a crucial aspect of ITS. One of the issues that ITS can address is potholes as they raise a major threat to the safety of drivers and passengers. Hence, this project establishes a real-time pothole detection system that utilizes mobile sensing to share pothole information. The system employs 'wave propagation distance' processed through Cosine angle computation and a pothole detection algorithm to identify potholes and uses spatial interpolation to give a visual representation of the road which can provide a detailed analysis of the road surface, including elevation changes, surface roughness, and other factors that can impact the likelihood of potholes forming. Experiment results indicate that the proposed approach accurately detects potholes without false positives and provides a higher level of accuracy compared to other methods.

Keywords: Pothole, Cosine angle, Spatial Interpolation, Terrain Mapping

1. Introduction

In the current epoch, roads and transportation systems have become an indispensable part. The transportation system has reduced commuting time; however, it has simultaneously elevated the risk to human life. Additionally, potholes can cause various types of damage to a vehicle, including harm to tires and wheels, mechanical parts, sudden braking and steering actions, and even vehicle accidents. The main factors that contribute to potholes are heavy rainfall and the passage of heavy vehicles. A recent survey on road accidents in India showed that over 2500 people lost their lives due to accidents caused because of potholes. To prevent this, timely and effective communication of information to drivers can greatly reduce the number of road accidents. Therefore, the use of Raspberry Pi for pothole detection provides drivers with advance warnings about the presence of potholes on the road. Also, this pothole data can be recorded on a server which allows the driver to receive information about the number of potholes on a specific road.

2. Need of Pothole Detection System

The need for a hardware-based pothole detection system arises from the limitations of traditional manual and visual inspection methods for detecting and repairing potholes. These methods are often time-consuming, and resource-intensive, and can result in missed potholes or delayed repairs. Hardware-based pothole detection systems use sensors and other technologies to automatically detect potholes and provide real-time information about their location and severity. This information can then be used by road maintenance crews to quickly and efficiently repair the potholes. Hardware-based detection systems offer several advantages over traditional methods. For example, they can cover a larger area in a shorter amount of time, resulting in more accurate and comprehensive data. They also allow for continuous monitoring of the roads, which can identify new potholes as soon as they form, reducing the risk of damage to vehicles and improving road safety. To simplify the monitoring process, terrain mapping techniques are employed in pothole detection systems to create a digital representation of the road surface, which can then be utilized to identify potholes. Furthermore, hardware-based pothole detection systems can be integrated with other technologies, such as GPS and mapping systems, to provide real-time information to drivers about the location and severity of potholes on their route. This can help drivers avoid potholes and reduce the risk of damage to their vehicles. Overall, the use of a hardware-based pothole detection system can improve the efficiency and effectiveness of road maintenance, while also reducing the risk of vehicle damage and improving road safety.

* Abhishek Singh
E-mail address: selvarajanupriya2@gmail.com

3. Literature Survey

Musa et. al. have proposed a lightweight pothole detection model suitable for edge devices from a bulky model using the data distillation method. It shows that data distillation is a better model compression technique, resulting in a lightweight model without sacrificing accuracy and with reduced computation and training time[1]. Shrivastava et. al. have presented a digital image processing and machine learning-based software to detect and count the number of potholes on roads. The purpose is to improve road conditions, reduce accidents, and simplify the pothole removal process by providing information on pothole location and count to repair crews. The methodology involves image preprocessing and machine learning to identify potholes, using a video and photo dataset[2]. Biju et. al. have developed the "Unseen Abyss" app to address poor road conditions caused by low-quality materials and potholes. The app uses a Raspberry Pi platform and camera module to capture images of potholes, and a TensorFlow-based model to classify and detect them. The GPS module uploads pothole coordinates to Google Sheets, which are displayed to users via a Flutter-based app[3]. KC and Sriharipriya have conceptualized an automated pothole detection process with high accuracy and efficiency using the YOLOX object detection algorithm. The YOLOX-Nano model is trained on a complete dataset and evaluated based on accuracy, recall, and model size, showing higher accuracy and lower processing costs compared to other YOLO algorithms. With an average precision (AP) score of 85.6% and a size of 7.22 MB, this study demonstrates the effectiveness of the YOLOX algorithm for pothole detection, contributing to cost reduction and improved speed of pothole identification, crucial for road maintenance[4]. Kim et. al. have discussed the importance of proactive pothole management for ensuring driver safety and traffic flow. Traditional human-based detection methods are being replaced by machine-based methods, which can be classified into three categories: vision-based, vibration-based, and 3D reconstruction-based approaches. The paper compares and evaluates the strengths and weaknesses of each approach, as well as the latest advancements in each technique. The paper also outlines future development plans related to these studies[5]. Sathya and Saleena have designed an automated pothole detection model using an optimized deep recurrent neural network (ODNR) to improve detection accuracy. The model includes three phases: pre-processing, feature extraction, and unsupervised classification. The ODNR classification is optimized using the Improved Atom search optimization algorithm (IASO) to enhance efficiency. The model's performance is evaluated in terms of accuracy, recall, precision, and error performance. The experimental results show a maximum accuracy of 97.7%[6]. Dewangan, DK, and Sahu have devised a prototype model that detects potholes and demonstrates intelligent driving behavior for autonomous vehicle systems. The prototype was developed using a convolutional neural network and a vision camera to achieve accuracy, sensitivity, and F-measure of 99.02%, 99.03%, and 98.33%, respectively[7]. Hwang et. al. have deduced a road defect detection system that uses deep learning and Raspberry Pi to analyze images and GPS data to display road defects on a map. The deep learning model has an accuracy of 96%, and the system aims to provide an efficient and cost-effective solution for managing road defects[8]. Yik et. al. have intended to use the YOLOv3 algorithm for real-time pothole detection on roads. The algorithm is trained using 330 sets of data and provides accurate results. Detected potholes are logged and displayed on a map using Google Maps API. The proposed algorithm is an efficient and precise solution for monitoring potholes in real time and can benefit both the public and the government[9]. Kamalesh et. al. have built a low-cost IoT-based device to detect potholes and humps on roads in India. This device can be installed on moving vehicles and sends images along with GPS coordinates to the authorities to take corrective measures. The system uses Raspberry Pi3 SBC for image capture, analysis, and email communication, and the GPS location data is stored on the ThingsBoard server hosted on AWS. The proposed system has a 100% reporting success rate[10]. V. Kaushik and B. S. Kalyan has stated that potholes in the road are a major cause of traffic accidents. Potholes can cause different types of damage to your vehicle, including: Punctures, broken wheels, impacts to the vehicle body and interior, etc., can cause serious injury. Properly detecting potholes is therefore a very important task, and it is also necessary to improve road conditions through road management systems and implement appropriate plans to repair potholes. Various plans are being made to develop techniques that can automatically detect potholes. This document describes some of the techniques used to detect potholes and proposes pothole detection techniques that can accurately detect potholes[11]. E. Ranyal et. al. have presented AI-assisted engineering solutions that integrate off-the-shelf RGB sensors with compute-intensive graphics processing units (GPUs) promising cost-effective solutions to prevent premature pavement deterioration. A common problem on roads, potholes pose a serious threat to road safety, and road inspection and condition monitoring require time- and cost-effective, state-of-the-art technology. In this study, we propose an intelligent sidewalk road pothole detection system that modifies a single-layer CNN architecture (RetinaNet) to detect road potholes on roads and uses 3D vision to perform metrology studies. Using motion-from-structure photogrammetry techniques based on static images extracted from video recordings of the road, we modeled a 3D point cloud structure of potholes and measured the severity of detected potholes as a function of depth. Evaluate and integrate a CNN-based pothole detection system. A high F1 score (as high as 0.98) on the benchmark dataset confirms the model's performance. An average error of less than 5% is obtained for the measured depths, promising a smart and practical solution to be implemented as part of a possible pavement health assessment system for future practice[12]. N. Chibani et. al. have stated that sensors on smartphones have become a ubiquitous component of Intelligent Transportation Systems (ITS), particularly in detecting cracks, bumps, and potholes on roads. In this context, and based on the certainty classification approach, many solutions using smartphone sensors have been proposed. Such an approach to certainty ignores the uncertainty associated with sensors, which is its main flaw. In fact, the sensed information may be imprecise and inaccurate, prone to error, and subject to incompleteness, ambiguity, and sometimes conflict. Moreover, in the ITS field, many factors like the sensor's quality, sensor lifetime, and the position of the sensor in the vehicle have a severe impact on the detection process. To address this issue and detect road anomalies more accurately, this article examines an uncertainty classification approach based on the Dempster-Shafer theory (DST). To evaluate the proposed uncertainty method, we used a publicly available dataset and compared it with several existing methods in the literature. As a result of comparative evaluation, the proposed method is superior to existing methods [13].

4. Proposed Solution

The proposed hardware-based pothole detection system would utilize sensors & a data processing unit to provide real-time information about potholes on roads and highways.

4.1 Components Used

- Raspberry Pi Board (Processing Module)

  The Raspberry Pi is a small, palm-sized computer board that can perform various functions like playing games, word processing, and creating spreadsheets. It is an affordable and portable device that is accessible to everyone. With the rapid growth of mobile computing technology, a significant portion of this growth is driven by the mobile industry. However, the Raspberry Pi board does not have built-in storage and requires an SD/micro-SD card to store the operating system and necessary code.

- LIDAR Sensor (Sensing Module)

  A LIDAR (Light Detection and Ranging) sensor emits a laser beam that bounces off objects in its path and returns to the sensor. The sensor measures the time it takes for the laser beam to return and calculates the distance to the object based on the speed of light. By rapidly emitting and detecting laser pulses in different directions, a LIDAR sensor can create a detailed 3D map of its surroundings. This technology is commonly used in self-driving cars, robotics, and other applications where accurate distance measurement and object detection is crucial. The range of a LIDAR (Light Detection and Ranging) sensor depends on its design and specifications but can typically range from a few meters to over 100 meters.

- Power Unit

  The Raspberry Pi requires a specific +5.1V power supply with a 2.5Amps capacity, which can easily be met in its desktop configuration by using a micro-USB charger that provides the exact voltage.

4.2 Implementation

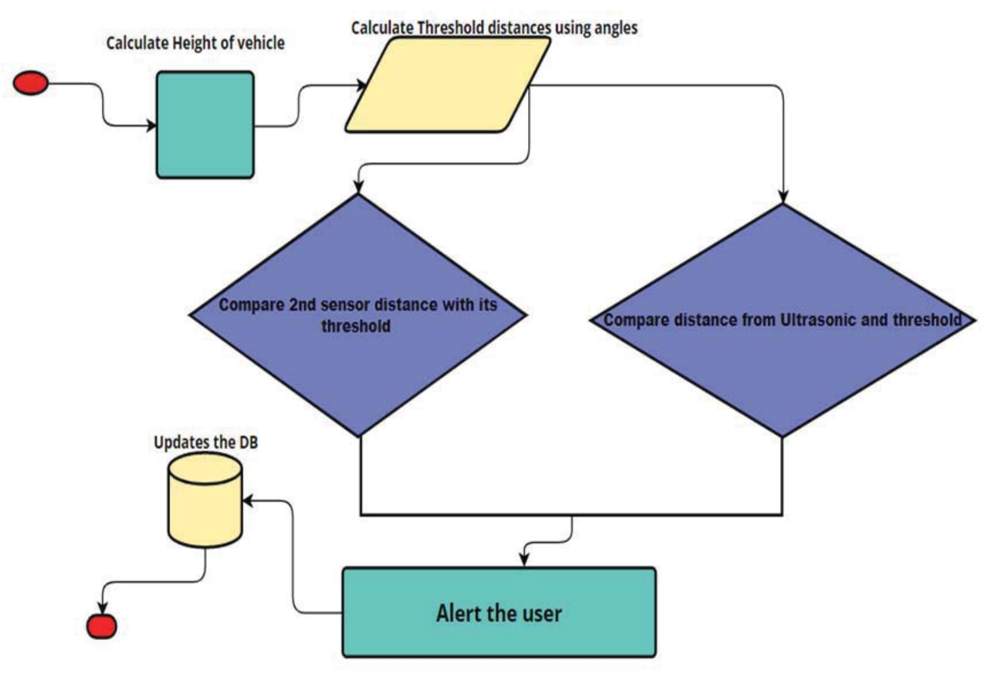


Fig. 1. Functional Flow

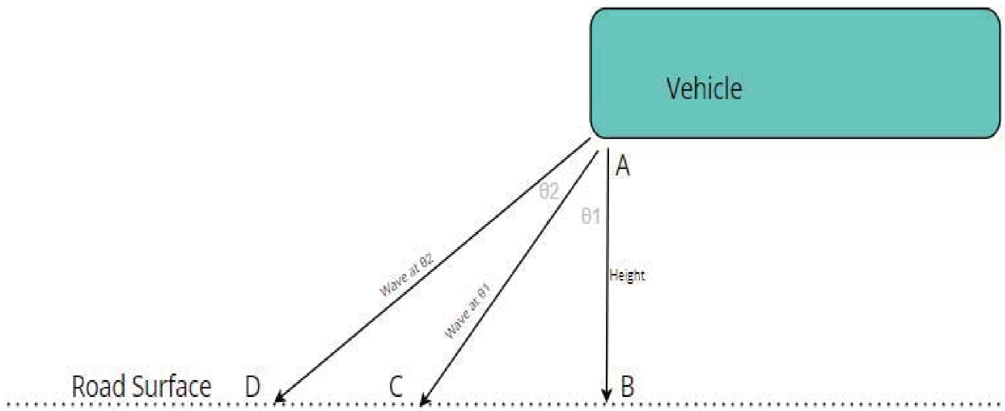


Fig. 2. Cosine Angle Calculation

Note: For Prototype purpose, we are using an Ultrasonic sensor

The ultrasonic sensor accurately measures the distance of a target object up to 400 cm (13 feet) with a precision of 3 mm, using non-contact methods by emitting ultrasonic sound waves and converting the reflected sound into an electrical signal.

One component function as a transmitter that transforms the electrical signal into 40 KHz ultrasonic sound waves, while the other serves as a receiver that listens for the transmitted waves.

Notation:

AB: Height of chassis from the surface

AC: Cosine Distance at θ1

AD: Cosine Distance at θ2

$\cos\theta 2 = AB/AD$
$AD = AB/\cos\theta 2$
This AD will be marked as Threshold for that instance of calculation
As we can calculate AB using an Ultrasonic Sensor & we know the angle. With this, we can calculate AD.
It will be compared with the calculated distance from Ultrasonic Sensor

If the calculated distance from the ultrasonic sensor is **greater** than the cosine calculated distance, it indicates that there is a pothole; if it is **less** than that, it indicates the presence of a **speed breaker or other obstacles**.

$\cos\theta 1 = AB/AC$
$AC = AB/\cos\theta 1$

We can calculate AB using Ultrasonic Sensor & we know the angle θ1 using this we can calculate AC

This calculated AC Distance will be compared with the calculated distance from Ultrasonic Sensor

This AC along with AD is used to evaluate slope on roads i.e. If both are traveling less than calculated cosine distance indicates there is the upward slope or if both are traveling less than calculated cosine distance indicates there is a downward slope.

The determined distance can be utilized to map the road's rugged terrain and display a visual representation inside the vehicle, assisting the driver in seeing the road.

5. Prototype Snapshots

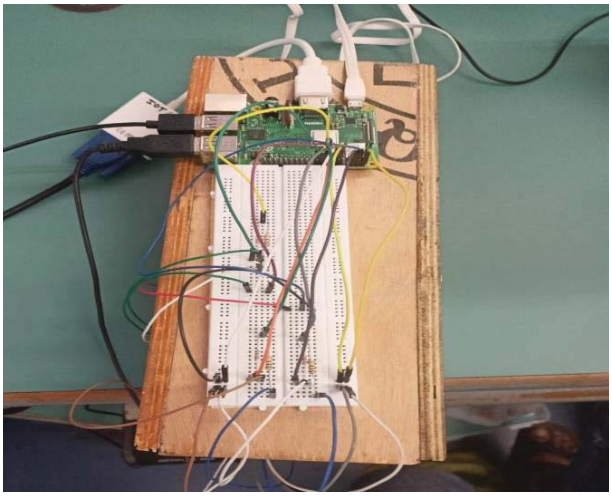


Fig.3. Wiring

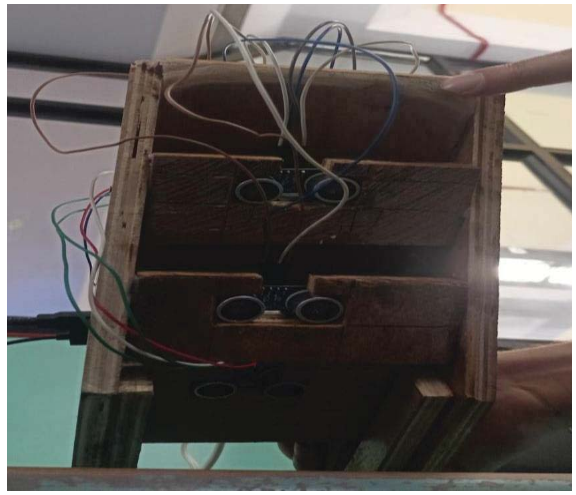


Fig.4. Sensor array layout

6. Results and Discussions

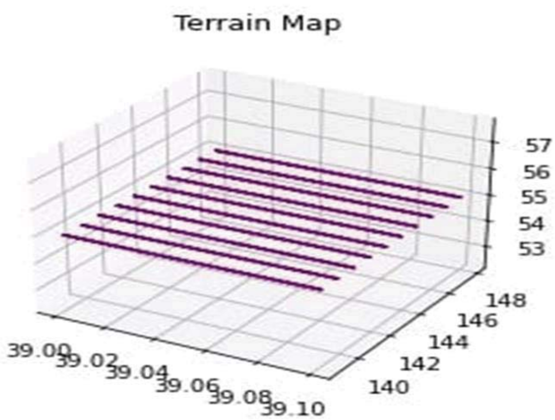


Fig.5. Terrain mapping

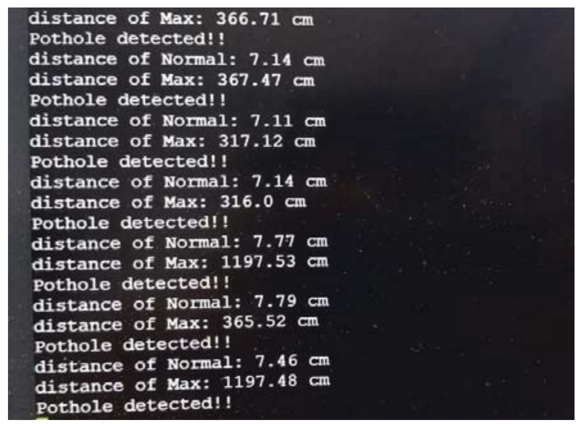


Fig.6. Console Output

7. Conclusion

A hardware-based system that detects potholes in all weather conditions is a critical tool for ensuring road safety and maintenance. It incorporates various sensors that function effectively in diverse weather conditions, thus increasing the accuracy and reliability of pothole detection. By providing real-time information about road conditions, this system enables reducing the risk of accidents, quick and effective maintenance, and extending the life of the road infrastructure. The Raspberry Pi-based 'TerrainEye - Pothole Detection System' is a cost-effective and practical solution that can help minimize the problem of potholes on roads and highways, resulting in safer and smoother travel for drivers, lower road maintenance costs, and improved road infrastructure quality.

The project for detecting potholes has been effectively executed, but the prototype is rudimentary and can be enhanced in both the hardware and software components. Several ideas can be explored for future model expansion, such as integrating maps into the Android application to visually depict pothole locations, adding voice notifications for pothole distance in addition to the existing warning, creating a more intricate prototype to regulate vehicle speed in the presence of potholes, and utilizing a more precise and intricate algorithm to determine the distance between two sets of latitude and longitude.

Based on the considerations outlined in the previous headings, the proposed system "IoT based pothole detection system" appears to be feasible. The necessary hardware and software components required to build an IoT based pothole detection system are readily available, and the integration of these components is technically feasible. The project has the potential to provide a significant return on investment by reducing vehicle maintenance costs and improving road safety. The installation and maintenance costs of the system can be offset by the savings realized from timely pothole repairs, which can help to minimize further damage to the road surface. Furthermore, the use of IoT technology for pothole detection can help reduce the environmental impact of road maintenance activities by minimizing the amount of raw materials required to repair potholes. Therefore, the project appears to be technically, economically, operationally, and environmentally feasible, and has the potential to provide significant benefits to both motorists and road maintenance agencies.

References

1. Musa, Aminu, Mohamed Hamada, and Mohammed Hassan. "A Theoretical Framework Towards Building a Lightweight Model for Pothole Detection using Knowledge Distillation Approach." In SHS Web of Conferences, vol. 139, p. 03002. EDP Sciences, 2022.
2. Shrivastava, Ankit, Devesh Kumar Srivastava, and Aditi Shukla. "Review of Road Pothole Detection Using Machine Learning Techniques." In Information and Communication Technology for Competitive Strategies (ICTCS 2021), pp. 95-104. Springer, Singapore, 2022.
3. Biju, Emmanuel Davis, Grigary C. Antony, Fabius S. Thottappilly, Disan Davis, and Anju Babu. "Unseen Abyss: Implementation of Pothole Detection System Using Machine Learning." In Proceedings of Third International Conference on Intelligent Computing, Information and Control Systems, pp. 105-117. Springer, Singapore, 2022.
4. KC, Sriharipriya. "Enhanced pothole detection system using YOLOX algorithm." Autonomous Intelligent Systems 2, no. 1 (2022): 1-16.
5. Kim, Young-Mok, Young-Gil Kim, Seung-Yong Son, Soo-Yeon Lim, Bong-Yeol Choi, and Doo-Hyun Choi. "Review of Recent Automated Pothole-Detection Methods." Applied Sciences 12, no. 11 (2022): 5320.
6. Sathya, R., and B. Saleena. "A Framework for Designing Unsupervised Pothole Detection by Integrating Feature Extraction Using Deep Recurrent Neural Network." Wireless Personal Communications 126, no. 2 (2022): 1241-1271.
7. Dewangan, D.K. and Sahu, S.P., 2021. PotNet: Pothole detection for autonomous vehicle systems using convolutional neural networks. Electronics Letters, 57(2), pp.53-56.
8. Hwang, Sung-jin, Seok-woo Hong, Jong-seo Yoon, Heemin Park, and Hyun-chul Kim. "Deep Learning-based Pothole Detection System." Journal of Semiconductor & Display Technology 20, no. 1 (2021): 88-93.
9. Yik, Yeoh Keng, Nurul Ezaila Alias, Yusmeeraz Yusof, and Suhaila Isaak. "A real-time pothole detection based on deep learning approach." In Journal of physics: Conference series, vol. 1828, no. 1, p. 012001. IOP Publishing, 2021.
10. Kamalesh, M.S., Chokkalingam, B., Arumugam, J., Sengottaiyan, G., Subramani, S. and Shah, M.A., 2021. An intelligent real-time pothole detection and warning system for automobile applications based on IoT technology. Journal of Applied Science and Engineering,24(1),
11. V. Kaushik and B. S. Kalyan, "Pothole Detection System: A Review of Different Methods Used for Detection," 2022 Second International Conference on Computer Science, Engineering and Applications (ICCSEA), Gunupur, India, 2022, pp. 1-4, doi: 10.1109/ICCSEA54677.2022.9936360.
12. E. Ranyal, A. Sadhu and K. Jain, "AI assisted pothole detection and depth estimation," 2023 International Conference on Machine Intelligence for Geo Analytics and Remote Sensing (MIGARS), Hyderabad, India, 2023, pp. 1-4, doi: 10.1109/MIGARS57353.2023.10064547.
13. N. Chibani, F. Sebbak, W. Cherifi, M. R. Senouci and S. Ouchani, "An uncertainty-aware Road Anomalies Detection Based on Dempster–Shafer Theory," 2022 First International Conference on Computer Communications and Intelligent Systems (I3CIS), Jijel, Algeria, 2022, pp. 25-30, doi: 10.1109/I3CIS56626.2022.10076224.

Cloud Data Centers Based Task Failure Prediction Using Machine Learning Techniques

Regalla Phaneendra[a]*, Arif Mohammad Abdul[b]

[a]*Regalla Phaneendra,Hyderabad, India*
[b]*Arif Mohammad Abdul, Hyderabad, India*

Abstract

The highly scalable cloud data centre should dispense efficient consistency as well as scalability with a low failure probability. Meanwhile, large scale cloud data centres tend to have significant failure rates due to a variety of factors, including software and hardware failure, that also has higher probability in task failures. Severe mishaps will drastically impair overall dependability of cloud services whilst still consuming a lot of resources for revive the services out of failure. To prevent unanticipated loss, this is essential for predict the task failure having great accuracy earlier they arise. In this study, consequently provide an exhaustive comparing as well as evaluation metrics using prediction methods for task failure. Predominant algorithms of machine learning are employed to develop and test these techniques. During testing and training these classifiers, used a reference data-set called Google Cluster data. The research yielded the following results. We concluded that Random Forest yields the most accurate model for predicting task failure with accuracy score of 94.625%, Precision and F1 scores with 0.89, 0.74. Decision tree achieved 0.69 of Recall score.

Keywords: Cloud; Precdicting; Machine laerning; Prediction; Task failure.

Nomenclature

PM	Physical machine
VM	Virtual machine
CPU	Central processing unit
OS-SVM	Online sequential-support vector machine
SVM	Support vector machine
RF	Random Forest
DT	Decision Tree
LR	Logistic egression
KNN	K-nearest neighbour
VC	Voting classifier
NB	Naïve bayes

1. Introduction

Cloud services entails the delivery of numerous services which including storage of data, infrastructure, networks and applications over the internet. Such method is gaining popularity among clients that require storage space and businesses looking for a secure website backup data service. Pay as you go should suit user's cloud global resources, thus they don't have to fret about overprovisioning a business which consumption of resources doesn't quite actually deliver, and instead squandering money under-providing the services that becomes important in the market, and thereafter foregoing out on additional income.

Providers of cloud services use server virtual machines that let a physical server (PM) to perform several virtual machine instances (VMs) jobs using distinct utilization of resources. The VMs in a cloud support legion application. Since the workload upon the PM fluctuates by schedule, the PM could be overworked. That traffic unbalance during a PM has a negative impact on the performance of every virtual application executing upon a PM. Since cloud - based data services are frequently overloaded, services like CPU and internet were associated with significant morbidity due to being shared by multiple users. Fig.1. illustrates overall architecture of existing data center of cloud. Users can transmit requests to a cloud, which including data storage and performing applications. Every cloud data center has been comprised of physical machine (PM), which can each handle the cluster of virtual machines (VM). Tasks sent through individuals were executed within every virtual machine. A large-scale data center may host a large

* Regalla Phaneendra
E-mail address: phaniram4@gmail.com

number of machines, each of which execute numerous apps and get client requests from users around the world each millisecond. The distributed storage facility of such variability as well as intense loads could be exposed to several forms of failure at times.

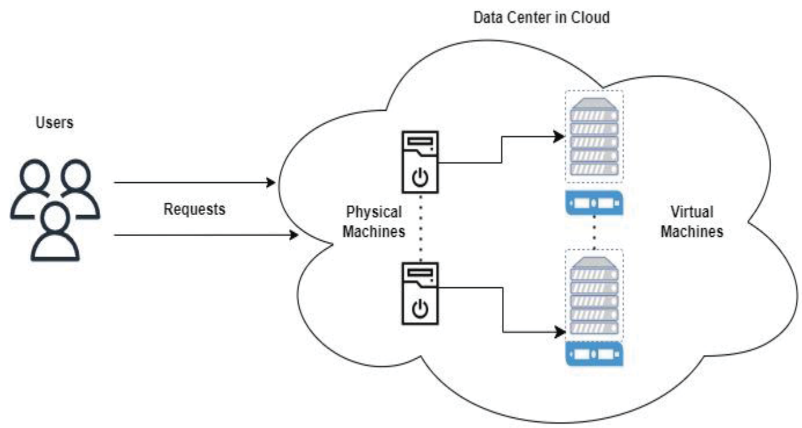


Fig. 1. Illustration of cloud data center

Earlier research [1], [2] has shown revealed failures of hardware and disc, seems to be a main cause of cloud service disruptions. But since failure of disc might resulting not only a service disruption and moreover immutable loss of data. Whilst each failure of disk is uncommon, the device containing several many disks can frequently encounter failure. Therefore, effective prognosis of application failure ahead of time will enhance the effectiveness in overcoming such fault while keeping the application active.

2. Related work

The above part provides a comprehensive literature review, along with a study as well as overview on prognosis of failure. Majority of the studies have seen a significant increased interest in failure analysis applying machine learning approaches for a large and complex excellent efficient in environment of cloud. To thoroughly ascertain and anticipate higher efficiency cloud environment metadata practically employing the failure on actual data, only several studies have been conducted.

The study [3] compare the effectiveness of relevant state-of-the-art prediction of workload algorithms in this research. Proposed an approach for performing such prognosis ahead of time in so that there is enough period to scheduling tasks forecasting workloads. For increase predictive performance even further, we propose a clustering-based prediction of workload strategy that splits overall tasks onto various factions before training the predicting approach for every set. The Google cluster trace-driven results demonstrate clustering-based workload estimation techniques outperforms previous comparable approaches but also achieved better predictive accuracy across over 90% through both memory and CPU. [4] presents a new two-phase machine learning technique to calculate workflow task completion timeframes with in cloud for variable data input. This approach relies on variables expressing duration statistics with two steps for forecast better accuracy. It obtains best and worst phase outcome of 1.6% and 12.2 %, whereas current approaches yielded inaccuracies of more than 20% since much above 75% of an examined task of workflows. This could replicate very least input onto newer clouds with little failures via needing just a finite numeral of steps.

The data centers of cloud should ensure exceptional dependability services and scalability whilst reducing overall possibility of failure. Inadequacies might substantially impact the dependability on services of cloud while also costing a large number of assets to recover the service. For enhance failure predictive performance even further, this study [5] illustrates a predictive model that employs multi-layer Bidirectional Long Short-Term Memory (Bi-LSTM) for detect task and job failure with in cloud. The Bi-LSTM algorithm's intent is to estimate whether tasks and jobs will fail or succeed. Our approach surpasses established algorithms, achieving 93% accurate result for task failure and 87% accuracy rate for job failures. Earlier activity breakdown estimation and appropriate recycling processes may considerably enhance asset capacity in high scaling data centers. For addressing this issue, work [6] proposes a unique approach rely on online sequential extreme learning machine (OS-ELM) to evaluate online job status of cessation. A comparison analysis employing google trace data reveals that its hypothesized technique's accurate rate is 93%, with the prototype upgrading within 0.01 seconds. As compared to approaches like support vector machine (SVM), ELM and OS-SVM, this technique offers several benefits, including less time generating and upgrading the system, greater predictive accuracy and precision, and good efficiency of false negative. Applications of Scientifical use the cloud to implement specific processes using tasks. Whenever task fails, the reliance orientation including the workflow takes into entire execution efficiency. This study [7] presents the predictive failure strategy that integrates a combination of machine learning techniques. For better accuracy, ensembles naive Bayes with various classifiers and determine here that the naive Bayes combined random forest as well as MF2N2 classifier attains accurate results (95.8%) compared to any other ensemble technique. This suggested approach has verified through combining simulated as well as actual cloud infrastructure.

Under this research [8], primarily investigated and classified the activity of failures as well as performed tasks via accessible to the public data. Conceived and built a failure predicting technique to detect cancelled tasks before they happen. This proposed method attempts to improve machination utilization and even the efficiency of cloud. By analyzing the conceptual approach using three openly accessible records like Mustang, Trinity and goggle cluster. Furthermore, data records being submitted to several machine learning algorithms in order to estimate the more precise option. Ultimately, by utilizing Random Forest and Decision tree classifiers attained better accurate results for Google trace data with parametric metrics like recall, precision and F1 score. Failure of cloud services is a significant concern because it could cost facility donors of cloud huge amount of billings, particularly alongside the dissipation of productivity incurred to business clients. As a result, in this work [9] a comparative evaluation as well as model validation for estimating approaches for task failures in application. For training and testing the methods, used reference dataset termed google cloud traces. Machine learning and deep learning techniques are applied to build and train these data traces. By the results, determined that in the instance for job failure estimation, Extreme Gradient Boosting gives the best accurate outcome, with features of disc space and CPU request. And for scaled evaluation, Logistic Regression method is the most flexible.

This is critical in a massive data center, that accurately discern application termination tracking. Several supervised learning methods has been employed towards this problem, since it is effective for enhancing resource utilization efficiency. Due to this sometimes the predictive accurate rate will be alleviated. This research [10] proposes a novel prediction failure methodology on the alliance connections among comparable jobs which can collectively estimate task dismissal status updates with an incipient time. Initially, the job cluster technique is developed for selecting tasks with greater resemblance among tasks with varying task counts. Finally, relying upon that task cluster outcomes, its robust multi-task learning approach is proposed in order to optimally use information across tasks. These conclusions reveal the suggested approach achieves greater predictive rate, reduced mis judgement rate. A statistical review of cloud existing workload metadata offers visibility to failure attributes, that can be applied to boost system solidity. This article [11] summarizes detailed systemic review on task utilization statistics on the Google cluster data, followed by the design of failure estimating approaches to foretell failures. This is revealed that its resource utilization behavior, processing period, & volume of resource consumed either by failure task differs from a completed task. Multiple interpolation strategies as well as the XGboost algorithm were used to anticipate the failure of a job in a massively contrast dataset, after it was revealed that Synthesis minority oversampling and the XGboost classifier detected the task report with 94.8% recall and 92% precision.

Understanding and categorizing reported failures is critical for building the secure cloud infrastructure. Purpose of this study [12] is to discover the critical elements which correspond to cloud application faults but also proffer the prediction approach which could forecast the conclusion about a job before it completes, crashes, or get killed. To achieve, conducted a failure classification analysis on Google cluster workload traces. Investigation shows both failed and terminated jobs consume a large number of resources. For investigate the possibilities for failure detection in applications of cloud in order to decrease resource waste by improving the tasks and jobs that inevitably fail or killed. For predict cloud application errors, offers a predictive procedure entrenched upon the special form of Recurrent Neural Network (RNN) called Long Short-Term Memory Network (LSTM). This algorithm predicts task failures having 87% accuracy, has a 85% of true positive rate, as well as false positive rate of 11%. Service suppliers of Cloud infrastructure are accountable to regulating its accessibility for allocating computational operations to furnish their clients with high quality of service. Examined and assess three publicly accessible enormous clustered data via Alibaba, Google and Trinity for characterizing failure of task on platform of cloud computing. By [13] devised the paradigm of failure aware task scheduling, which capable of predicting the terminating state about group of assigned tasks within execution as well as adopting relevant proactive measures. Artificial and convolutional neural networks are used for estimating the activities. By using Google data collection, the outcome demonstrate that ANN and CNN will attain predictive accuracy levels of 94% and 92%, respectively. Furthermore, with employing the appropriate process, the platform can prevent nearly 40% operations expected to fail utilizing Alibaba data, preserving significant ensembled resources like RAM and CPU.

Node failures in platform like AWS could really decrease the accessibility of its cloud infrastructure as well as possibly result in huge loss of revenue. Modeling failures of node may lead to critical because that allows DevOps developers could mitigate the severity through taking pre-emptive measures. Moreover, those prognostications were difficult because of numerous constraints, such as the massive bulk of the monitoring statistics and also the intricacy of a breakdown characteristics. By employing the techniques of machine learning, [14] considering data strategies includes oversampling and gapped. The analysis reveals that using random forest in combination with oversampling approach produces the greatest results in case of computational expense and predictive ability.

3. Dataset Description and Preprocessing

Researchers are already capable of solving computer issues throughout a short amount of time which would require many years upon the computer due to the expansion of computational facilities. Although the likelihood of an application failing due to a software or hardware problems improves as scalability rises. This type of operational error not just to impedes technological discovery and also consumes the significant quantity of time. By the previous research [15], the dataset containing parameters of “jobid”, “memory”, “networklog10”, “localIOlog10”, “NFSIOlog10” and “failed”. As shown in Table 1. The term failed is the output parameter and remaining are the input parameters.

The dataset's columns are as follows:

- **jobid:** the specific reference for every job.
- **memory:** the job's memory consumption, measured in Gigabytes.
- **networklog10:** its connectivity bandwidth utilized by a task in MB per second, log base 10.
- **localIOlog10:** a local Input/Output bandwidth utilized either by task in MB per second, log base 10.
- **NFSIOlog10:** Its value, expressed in MB per second, is the log base 10 of the input/output bandwidth required by the task on through network file system (NFS).
- **failed:** a binary indication that illustrates if the work was successful or unsuccessful, with a value of 1 denoting fail and a value of 0 success.

Table 1. Parameters in dataset

Parameters	Labels/Variables
Input	jobid memory networklog10 localIOlog10 NFSIOlog10
Output	failed

The dataset contains of 20000 job/task log files with the above-mentioned input and output parameters. We have information upon every job's resource utilization as well as the whether or not it was successful. It has failure data around approximately 8%, the positive class is what we refer to it as. “1” refers to failure and “0” refers to success. The replica of dataset is mentioned in Table 2.

Table 2. Dataset catalogue

Sr.No	**1**	**2**	**…**	**…**	**20000**
jobid	jobID 1634295	jobID 2033452	…	…	jobID 1165709
memory	44.3904	31.5839	…	…	7.9835
networklog10	-1.0262	-1.4608	…	…	0.7225
localIO10	0.8033	-0.608	…	…	-1.1583
NFSIOlog10	-3	-2.9967	…	…	0.2308
failed	0	0	…	…	0

4. Methodology

An experimental setup as well as the methodological approach enabling failures evaluation and prediction were addressed in this portion. Fig. 2. Illustrates to conduct an extensive comparative on the prediction of task failure based on the dataset. Here, we classified an operation into three segments, notably “Data handling”, “Model building”, “Analysis”. Data handling is the action of retrieving information from available records, which includes task occurrences, and then process the dataset. Segment model building necessitates the actions for building & testing prediction model built using machine learning techniques. This proposed methodology was aimed at solving the classifying issue, i.e., characterize anticipated dismissal outcome of every job, whether it is success or failed. The dataset has split among 80% as training as well as 20% as testing. These algorithms at work operate as follows. Segment analysis can evaluate the efficiency for prediction models using multiple statistical measures for identify the most effective algorithm.

Fig. 3. Illustrates the job event table's dispersion across scheduled groups which displays the status of the task failed/success. In above graphical representation, “1” refers to failure and “0” refers to success.

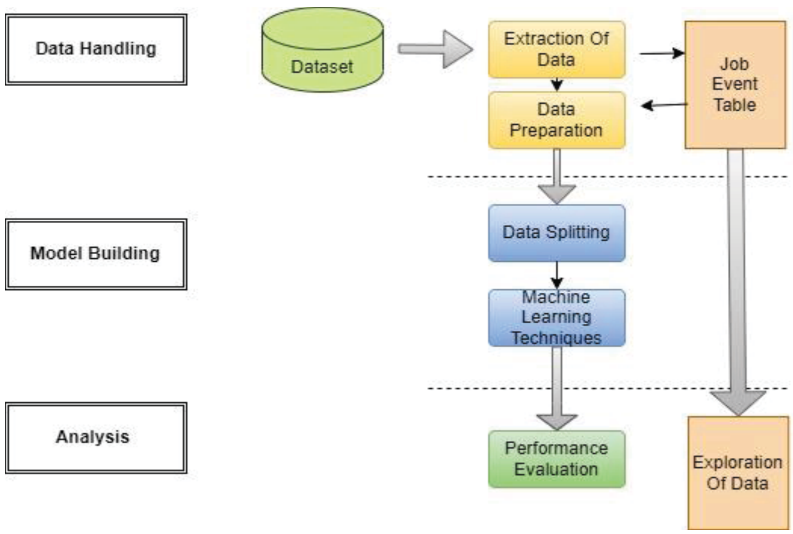


Fig. 2. Architecture illustration

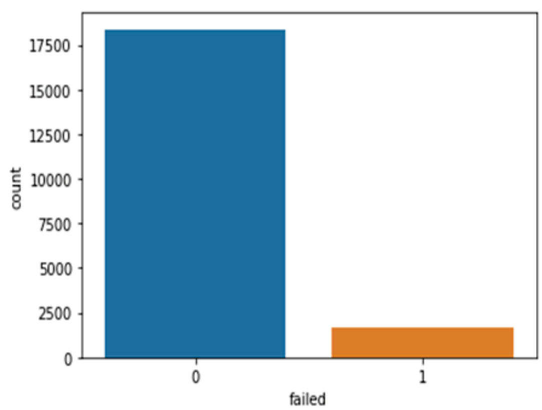


Fig. 3. Distribution of failed status

4.1. Algorithms used

Algorithms based on machine learning are also commonly employed in virtual machine job scheduling. The several machine learning techniques are employed and evaluated. In this study, machine learning techniques like Decision tree, Random Forest, KNN, Support vector machine, Logistic regression, Naïve bayes and Voting classifier are deployed and also used few statistical metrics.

- Support vector machine:
  This is centred upon that computing of such a linear regression model in such a multi – variate classifier at which inputs could be utilised through a non-linear regression model [16]. Since the computational complexity in constructing the model need not include additional data for training that is closest towards the prediction's outcome, each trained model via SVM is only reliant upon the sample from the set of training. From another terms, well during training the SVM classifier, one applies additional weights for the sample points beyond the prior expected to ensure that the method may explore further also on observations identify additional potential trends in the data set records.

- Logistic regression:
  Logistic regression will determine the one or more varying components getting evaluated and seems to have the significant correlation including an output as well as presents the indicator for the extent of future impact [17]. Additionally, it may balance the influencing variables, which are contributing factors with those related predicting parameters and indeed the result, ensuring that the estimate from the impact including the predicting on relevance also isn't affected more by impact.

- Random Forest:
  Random forests were a tree classifier integration whereby every tree is dependent upon that random vector's values, which are finite element modelling with a uniform occurrence overall the forest's trees. As both the number of trees inside a model expands, the generalisation error corresponds toward a limitation. This amplitude for every distinct baseliner as well as the connectivity amongst individuals define a generalisation error among a forest of classification and regression trees. A threshold again for generalisation error in this model can be determined using the combination of two factors which represent the accuracy of the different classifiers and indeed the independence of those models.

- Decision Tree:
  The formalisation of describing these conversions is therefore a decision tree. The leaf node annotated including a classification made up of a node connected to two maybe more sub - trees constitutes a tree. Each potential consequence was connected to a respective sub - trees, but a node evaluates particular interesting take upon that feature labels of an occurrence. Commencing with the tree's root node, an occurrence was categorised. Whether this tree would be a trial, then instance's result gets obtained as well as the procedure then is carried out by employing the suitable sub-tree. Every label of a leaf indicates its anticipated subclass of an occurrence that will finally be examined.

- K-nearest neighbor:
  This model has been a basic and efficient segmentation approach. The algorithm is a supervised learning classification that is non-parametric as well as employs approach for classify and otherwise predict that how given data value will be grouped. For every test sample, k nearest neighbours (KNN) should determine the distance (aka similar) across all training dataset [18].

- Naïve bayes:
  For a classified parameter, this Naïve bayes (NB) classification model is one family using elementary posterior probability that is widely chosen also as reference on classification tasks. This is built around the standard hypothesis that over parameters were dependent from one another [19].

- Voting classifier:
  A proposed feature selection voting classification combines three algorithms (i.e., support vector classification, random forest, decision tree) of machine learning techniques which integrate classifier well with ensembles method of voting [20]. This is a method which builds from an ensemble containing different approaches but also estimates the outcome dependent mostly on class that has the highest likelihood to result in the response. To predict an outcome classification determined by the largest vote total, it essentially averages information findings of every classification that was provided through into voting classification. This objective seems to be to build a single prototype that train from these approaches but also estimates approach that relies off its accumulated majorities in voting with each output variable, instead than developing standalone algorithms as well as determining their accurateness of everyone.

4.2. Statistical metrics

In this study, we employ three statistical metrics that evaluate the outcomes from the aforementioned techniques and demonstrates optimal performance.

- Precision score:
  It is also called as "Positive predictive value (PPV)". This is derived from the ratio of correct forecasts to all those provided by even the approach. The optimal precision for such good classification was 1. Only when the numerator as well as denominator were equivalent will precision become 1. FP must be 0 as well, when TP= TP + FP. Denominator value will increase with FP, passing numerator value, whereas precision value drops. In Eq. (1), where "TP" is considered as True positive, "FP" is derived as false positive.

$$PPV=TP/[TP+FP] \qquad (1)$$

- Recall score:
  Recall Eq. (2) is determined in such a classifier involving two categories via dividing the total number on true positives by the total of false negatives and true positives.

$$TPR=TP/[TP+FN] \qquad (2)$$

- F1-score:
  The weighted mean for recall and precision becomes a F1 score. Both false positive and false negative outcomes may occur with precision and recall, as is well known, thus both are taken into account. Eq. (3) defined as follows:

$$\text{F1 score} = (2\cdot PPV\cdot TPR)/(PPV+TPR)$$

$$= (2\cdot TP)/(2\cdot TP+FP+FN) \quad (3)$$

5. Results & Analysis

We therefore assess our technique's outcomes as well as evaluate it to earlier prediction approaches [3], [5] and [9] that other researchers have suggested.

Here we examine job scheduling failure predicting initially. Relying upon these parameters as well as performance scores, categorized the status for activity inputs at the job level. Every status accomplished was taken into account into one category across all target categories, while the status failure is taken into account like another category.

Table 3. Accuracy score for models

Performance	
Models	**Accuracy Score**
Random Forest (RF)	**94.625**
Logistic Regression (LR)	92.5
Decision Tree (DT)	90.85
Naïve Bayes (NB)	92.5
KNN	94.47
SVM	92.5
Voting Classifier (VC)	94.55

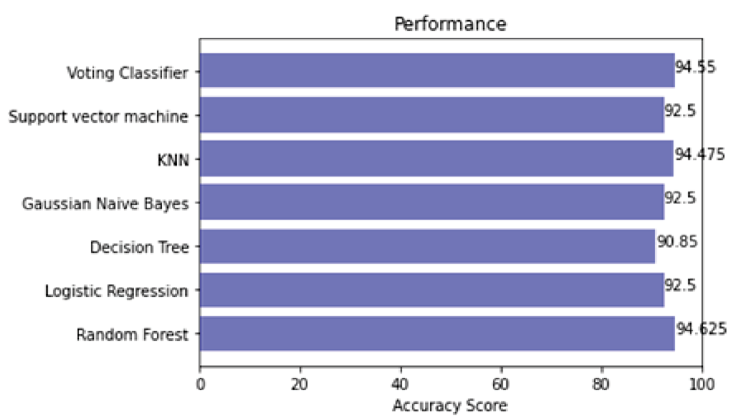


Fig. 4. Graphical illustration of accuracy scores

Overall accuracy score for every algorithm is classified their status on the basis of the dataset is demonstrated in Table 3. Every machine learning technique performance evaluation are apparent and efficient. Overall highest accuracy, especially, is 94.72% for the Random Forest approach. Each performance of accuracy score is graphically illustrated in above Fig. 4. The outcome is as follows RF > VC > KNN > NB > LR > SVM > DT. The Precision score for the Random Forest is the optimal which is of the 0.89. Precision, recall and F1 scores for the job failure status has been demonstrated into parts of success and failed segments. The tabular and graphical representation of precision score for overall machine learning approaches are displayed above in Table.4 and Fig. 5.

Table 4. Precision score for models

Precision score			
	Task/Job Status		
Models	0 (Success)	1 (Fail)	Average
RF	0.95	0.83	0.89
LR	0.93	0.0	0.46
DT	0.95	0.40	0.68
NB	0.93	0.0	0.46
KNN	0.95	0.78	0.87
SVM	0.93	0.0	0.46
VC	0.95	0.82	0.88

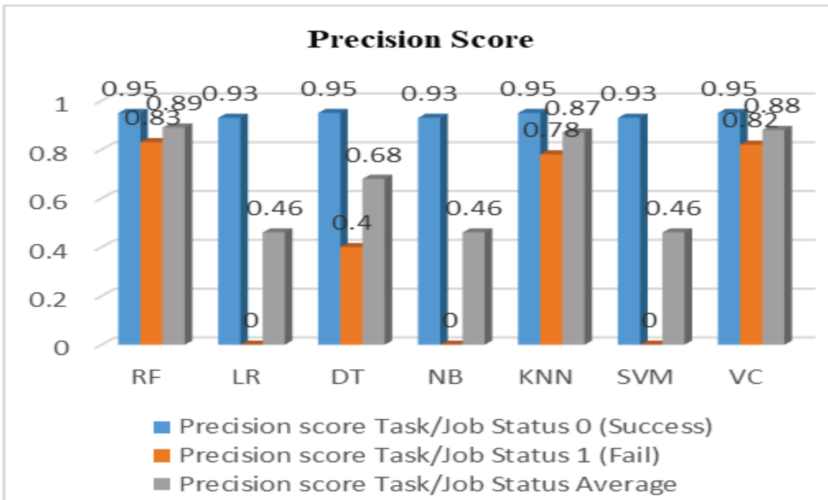


Fig. 5. Illustration for precision score

Random Forest approach of machine learning exhibited the higher F1 score with 0.74 and Decision tree achieved highest recall score with 0.69 about the job failures statuses. The below Table. 5 & Table. 6 shows the recall and F1 scores for every model involved in this study.

Table 5. Recall scores for ML models

Recall score			
	Task/Job Status		
Models	0 (Success)	1 (Fail)	Average
RF	0.99	0.37	0.68
LR	1.00	0.0	0.50
DT	0.95	0.44	0.69
NB	1.0	0.0	0.50
KNN	0.99	0.36	0.68
SVM	1.0	0.0	0.50
VC	0.99	0.36	0.68

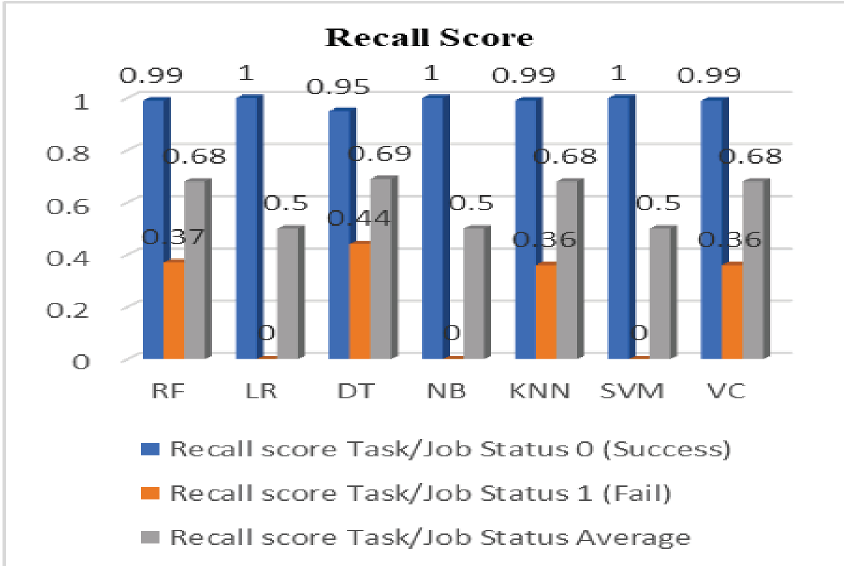


Fig. 6. Illustration of recall scores

By the Fig. 6 and Fig. 7. illustrate a Recall and F1 scores of each approach for job failure activity with various input sizes.

Table 6. F1-scores for models

F1-score			
	Task/Job Status		
Models	**0 (Success)**	**1 (Fail)**	**Average**
RF	0.97	0.51	0.74
LR	0.96	0.0	0.48
DT	0.95	0.42	0.68
NB	0.96	0.0	0.48
KNN	0.97	0.50	0.73
SVM	0.96	0.0	0.48
VC	0.97	0.50	0.73

The DT technique recorded the second-highest accuracy. The accuracy differential between the VC and KNN methods has been marginal gap, according to the conclusions. While testing on the dataset, the VC model had an accuracy of 94.57% and KNN achieved 94.47% accuracy. In metrics of precision score, second higher score acquired is RF and VC with 0.88. RF, VC and KNN has the second highest recall score with 0.68. In term of F1 score, VC and KNN achieved second highest score with 0.73. Finally, RF and VC approaches built for this job failure analysis, executed adequately in every statistical metrics involved in this study.

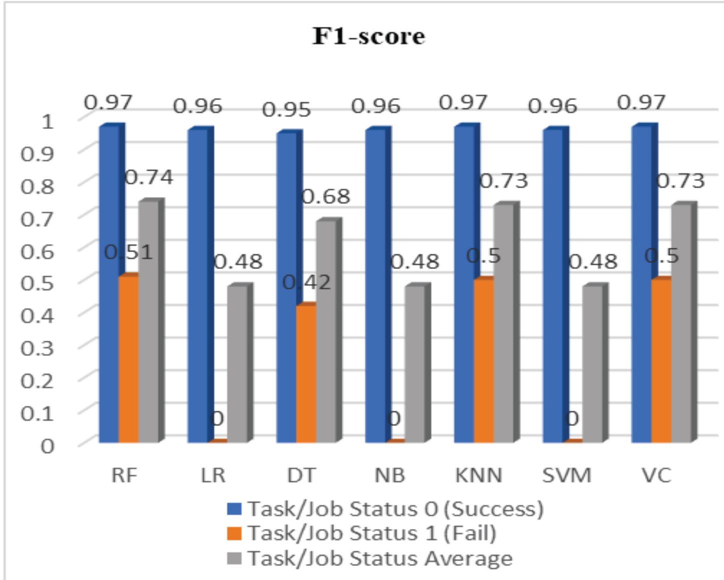


Fig. 7. Illustration of F1 scores

6. Conclusion

Adequate reliability and task fail predictions are vital to application quality of services in cloud data centers. In this study, we imply a rigorous interpretivist paradigm for prediction of task failure on Google cluster dataset. Due to this, we built machine learning models and examined the performance based on the few statistical metrics. According to the simulation results, the Random Forest classifier seems to be the optimal model for predicting the job failure activity. RF has accomplished an accuracy of 94.625%, precision score of 0.89 and F1 score of 0.74. In term of recall score with 0.69, DT achieved optimal solution. Ultimately, this research indicates that algorithms based on machine learning performed well in terms of evaluating overall status of job failure prediction.

Acknowledgement & Conflict of interest

"The article does not contain any studies with human participants or animals performed by any of the authors. Both authors declare, has no conflict of Interest".

References

1. E. W. J. W. S. D. L. O. S. a. E. E. M. Sedaghat, "DieHard: Reliable Scheduling to Survive Correlated Failures in Cloud Data Centers," Cloud and Grid Computing (CCGrid),, pp. 52-59, 2016.
2. R. P. M.-R. R. a. S. B. Subrata Mitra, "Partial-parallel-repair (PPR): a distributed technique for repairing erasure coded storage," Association for Computing Machinery, p. 1–16, 2016.
3. H. W. a. H. S. J. Gao, "Machine Learning Based Workload Prediction in Cloud Computing," International Conference on Computer Communications and Networks (ICCCN), pp. 1-9, 2020.
4. J. J. D. a. T. F. T. -P. Pham, "Predicting Workflow Task Execution Time in the Cloud Using A Two-Stage Machine Learning Approach," IEEE Transactions on Cloud Computing, vol. 8, pp. 256-268, 2020.
5. H. W. a. H. S. J. Gao, "Task Failure Prediction in Cloud Data Centers Using Deep Learning," IEEE International Conference on Big Data (Big Data),, pp. 1111-1116, 2019.
6. J. H. Y. S. C. L. B. C. a. J. C. C. Liu, "Predicting of Job Failure in Compute Cloud Based on Online Extreme Learning Machine: A Comparative Study," IEEE Access, vol. 5, pp. 9359-9368, 2017
7. P. U. A. Padmakumari, "Task Failure Prediction using Combine Bagging Ensemble (CBE) Classification in Cloud Workflow," Wireless Personal Communications, p. 23–40, 2019.
8. Q. H. M. Mohammad S. Jassas, "Analysis of Job Failure and Prediction Model for Cloud Computing Using Machine Learning," Sensors (Basel), 2022.

9. T. I. A. &. S. J. Tengku Asmawi, "Cloud failure prediction based on traditional machine learning and deep learning," Journal of Cloud Computing, 2022.
10. L. D. Y. L. G. L. &. W. M. Chunhong Liu, "Failure prediction of tasks in the cloud at an earlier stage: a solution based on domain information mining," Computing, 2020
11. R. S. a. S. G. J. Shetty, "Task Resource Usage Analysis and Failure Prediction in Cloud," Data Science & Engineering (Confluence), pp. 342-348, 2019.
12. T. I. a. D. Manivannan, "Predicting Application Failure in Cloud: A Machine Learning Approach," IEEE International Conference on Cognitive Computing (ICCC), pp. 24-31, 2017.
13. T. D. a. A. A. Y. Alahmad, "Proactive Failure-Aware Task Scheduling Framework for Cloud Computing," IEEE Access, vol. 9, pp. 106152-106168, 2021.
14. Z. M. (. J. H. L. A. E. H. C. H. R. H. Z. Z. M. W. a. P. C. Yangguang Li, "Predicting Node Failures in an Ultra-Large-Scale Cloud Computing Platform: An AIOps Solution," ACM Transactions on Software Engineering and Methodology, vol. 29, no. 2, pp. 1-24, 2020.
15. S. J. M. K. S. H. C. S. K. K. I. B. Rakesh Kumar, "The Mystery of the Failing Jobs: Insights from Operational Data from Two University-Wide Computing Systems," 50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN), pp. 158-171, 2020.
16. S. P. a. D. P. D. Basak, "Support vector regression," Neural Information Processing—Letters and Reviews, 2007
17. M. W. Tolles J, "Logistic Regression: Relating Patient Characteristics to Outcomes.," JAMA Guide to Statistics and Methods, 2016.
18. S. H. S. D. Zhu X, "A novel matrix-similarity based loss function for joint regression and classification in AD diagnosis," Neuroimage, 2014.
19. S. Xu, "Bayesian Naïve Bayes classifiers to text classification," Journal of Information Science, p. 48–59, 2018.
20. D. K. M. M. Saloni Kumari, "An ensemble approach for classification and prediction of diabetes mellitus using soft voting classifier," International Journal of Cognitive Computing in Engineering, vol. 2, pp. 40-46, 2021.

Attendance Monitoring System using Face Recognition Technologies

Rishi Vardhana[a] AB*,Subishek S[b],Vigneshgandhi. B[c],Dr. S. Ananthi[d], Ms. B. Suchithra[e],Dr.R.Sathya[f]

[a]*UG Scholar, Department of CSE, Sri Eshwar College of Engineering, Coimbatore, India*
[b] *UG Scholar, Department of CSE, Sri Eshwar College of Engineering, Coimbatore, India*
[c] *UG Scholar, Department of CSE, Sri Eshwar College of Engineering, Coimbatore, India*
[d] *Assistant Professor, Department of CSE, Sri Eshwar College of Engineering, Coimbatore, India*
[e] *Assistant Professor, Department of CSE, Sri Eshwar College of Engineering, Coimbatore, India*
[e] *Assistant Professor, Department of IT, Kongunadu College of Engineering and technology, Trichy, India*

Abstract

Face recognition-based attendance monitoring system is examined in this paper. The proposed system aims to provide quick and secure way to track attendance in places like companies and schools. This technology provides the facilitates to capture person's face image with a digital camera and compare it to photographs that have been saved in a database. Several performance criteria, including accuracy, speed, and dependability, were used to evaluate the system. The outcomes demonstrated that the system was capable of reliably recording attendance and achieving excellent face recognition accuracy. The system's real-time recognition capabilities made it possible for it to be used in a number of contexts. The study sheds light on how facial recognition technologies may be used for attendance monitoring and emphasises the need for more study in this field to improve the efficiency and dependability of such systems.

Keywords: Face API, Face Recognition, Framework, Integration, Monitoring, Microsoft Cognitive Services.

1. Introduction

In a variety of settings, including schools, offices, and organisations, keeping track of attendance is an important duty. To ensure that these environments run smoothly, it is crucial to keep track of people's attendance. Manual sign-in sheets have been used for attendance monitoring in the past, but they are time-consuming, prone to mistakes, and simple to hack. Technology has led to the development of new techniques for automating attendance tracking, and among these techniques, systems that use facial recognition technology have drawn a lot of attention.

In recent years, there has been a lot of research and development into facial recognition technology, and its applications have grown to include security, entertainment, and human-computer interaction. Convolutional Neural Network (CNN) algorithms' progress has made it possible to build facial recognition systems that are both extremely accurate and effective. A deep learning system known as CNNs has shown to be highly good at finding patterns in massive datasets. CNNs in particular have demonstrated excellent performance in classifying and detecting images, making them perfect for use in facial recognition systems.

In order to accomplish accurate and effective attendance monitoring, CNN facial recognition technologies are used in this work to present an examination of an attendance tracking system. The technology was created to use a digital camera to capture a person's face image and compare it to the photographs that were previously stored in a database. The algorithm was trained using a sizable collection of facial photos to improve the process's accuracy. The system was then put to the test using a database of photos of people in order to assess how well it performed in identifying faces and accurately recording attendance.

Compared to more conventional techniques, using CNN face recognition technology for attendance monitoring provides various benefits. First off, using facial recognition reduces the need for manual sign-in sheets and manual attendance monitoring, enhancing efficiency and lowering the chance of mistakes. Second, the use of facial recognition improves the security of the attendance tracking procedure by removing the risk of someone falsifying or manipulating sign-in sheets. Additionally, since facial recognition is automated, there is no longer a need for people to remember to bring identification or sign in forms.

The purpose of this study is to assess the effectiveness of the attendance monitoring system established in this study as well as the potential of CNN face recognition technologies for attendance monitoring. The outcomes of this study will offer insightful information on CNN face recognition technologies' potential for attendance monitoring and will draw attention to the need for additional study in this field to improve the efficiency and dependability of such systems.

* Rishi Vardhana
E-mail address: rishivardhana.b@sece.ac.in

2. LITERATURE SURVEY

In recent years, there has been a lot of study on the use of facial recognition technology for attendance monitoring, with a rising number of studies looking at the use of Convolutional Neural Network (CNN) algorithms in this context. The work of Wang et al. [1][12], who suggested an attendance tracking system utilising CNN-based facial recognition technology, is one of the earliest works in this field. This system was created to take a student's or employee's facial image using a digital camera and compare it to photographs that are recorded in a database to track attendance. The study demonstrated the potential of CNN-based facial recognition technology for attendance by showing that the system was able to recognise and record student and employee attendance information accurately.

The work of Liu et al. [2], who suggested an attendance tracking system for schools utilising deep learning algorithms, such as CNNs, is another significant study in this area. The technology was made to take a student's facial image with a digital camera and compare it to photographs that were previously saved in a database to track their attendance. The study demonstrated the potential of deep learning algorithms for school attendance tracking by showing that the system was able to achieve high accuracy in recognising and recording student attendance.

Growing interest has been shown in recent years in enhancing the precision and effectiveness of facial recognition systems used for attendance tracking. Zhang et al[3] .'s hybrid facial recognition system, which combines the use of conventional rule-based algorithms and deep learning methods, such as CNNs, was offered as a solution to this problem. The study demonstrated the potential of mixing several types of algorithms for improved facial recognition performance, showing that the hybrid system was able to achieve higher accuracy and efficiency compared to standard rule-based algorithms alone.

Privacy and security are two additional concerns with the usage of facial recognition technology. Ling et al[4] .'s proposal for a privacy-preserving facial recognition system for attendance tracking attempted to allay these worries. This system was created to use secure cryptographic algorithms to encrypt the saved photographs in order to preserve the privacy of the people whose facial images are being taken and stored. The study demonstrated that the system was able to maintain users' privacy while yet accurately identifying and documenting their attendance.

While much research in attendance tracking has centred on facial recognition technology, other biometric technologies, such fingerprint recognition, have also been investigated. The research by Falah Alsaqre. [5], who suggested a hybrid attendance monitoring system that incorporated the use of facial recognition and fingerprint recognition, is one such study. The study demonstrated the potential of merging several biometric technologies for improved attendance monitoring performance, showing that the hybrid system was able to achieve higher accuracy when compared to the usage of either technique alone.

In recent years, there has been a rise in the use of facial recognition technology for tracking attendance in big businesses and government institutions. Atallah AL-Shatnawi. [6][13] suggested an attendance monitoring system for massive enterprises utilising CNN-based facial recognition technology to address this issue. Utilizing a digital camera, the system was created to record an employee's attendance by matching the employee's face image with photos that were previously saved in the database. The study demonstrated the potential of CNN-based face recognition technology for employee attendance tracking in large enterprises by proving the system's ability to achieve high accuracy in employee detection and recording.

3. PROPOSED WORK

With the use of CNN-based facial recognition technology, we design an attendance monitoring system for this suggested work. The study will concentrate on assessing the system's precision and effectiveness as well as dealing with the crucial concerns of security and privacy. In order to monitor attendance, a CNN-based facial recognition system will be developed and put into use. Its effectiveness will be assessed through trials and comparisons with other attendance monitoring techniques. The study's findings will offer important information about the technology's potential for automating the attendance monitoring process and enhancing its precision and effectiveness.

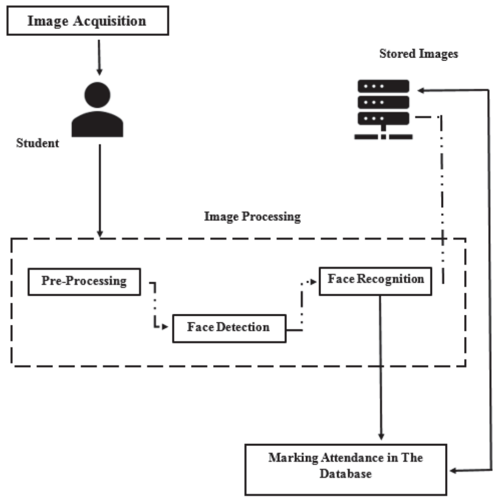


Figure 1. Architecture for Proposed Solution

3.1. Face Detection

An old-school computer vision algorithm called Haar Cascades is used to find objects in pictures and videos. The approach was first presented in 2001 by Viola and Jones[7], and it has since been extensively employed in a number of applications, including facial recognition, object tracking, and pedestrian detection. Haar cascades can be used to recognise faces from a webcam in the context of attendance tracking, facilitating the procedure.

Haar Cascades Algorithm

The act of processing an image and sliding-window evaluating the Haar-like features is how the Haar Cascades algorithm operates. Simple rectangular shapes called Haar-like characteristics are applied to the image to identify objects. A cascade of straightforward classifiers is created using the combined features, and each classifier is trained to recognise the presence of an object in the region of interest. An Adaboost classifier evaluates the features at each stage to determine whether the region includes the object of interest. The region moves on to the following level of the cascade if it tests positive. The procedure continues up until a certain threshold is achieved, signalling the detection of an object.
In real-time applications, Haar cascades are effective in terms of computing. However, especially in congested and complicated surroundings, they might be vulnerable to false positives and negatives. Haar cascades can be integrated with other object detection strategies, such as deep learning-based approaches, to lessen these problems and increase the accuracy and robustness of the results [14][18].
The suggested system would use the CNN-based facial recognition algorithm to process the identified faces after applying the Haar Cascades method to identify people accurately and track their attendance. The study's findings will shed important light on Haar Cascades' potential for automating the attendance monitoring process and enhancing its precision and effectiveness.

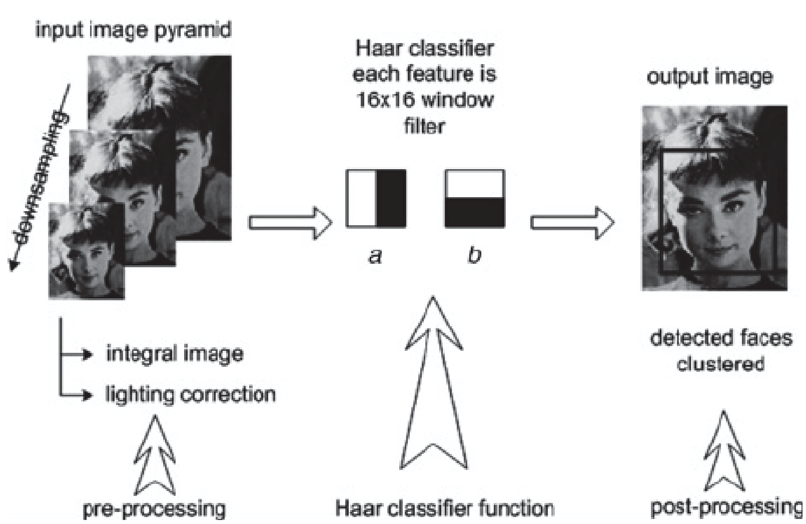


Figure 2. Working of Haar Cascades Algorithm

3.2. Feature Extraction Using Principle Component Analysis PCA

A popular method for feature extraction in computer vision and pattern recognition is Principal Component Analysis (PCA), a potent tool. Reducing the dimensionality of the data while retaining as much information as feasible is the aim of PCA. PCA is employed in the context of facial recognition to extract the key features from the face images, which can subsequently be applied to tasks requiring classification and recognition[8].
Data are transformed into a new coordinate system by PCA, with the first axis denoting the direction with the greatest variation. The second and subsequent axes are chosen to maximise the variance in the data and to be orthogonal to the first axis. The outcome is a collection of orthogonal axes that accurately represent the most significant data points. The major components of the face photos, which serve as the features for the recognition algorithm, will make up the converted data in the case of facial recognition.
The removal of noise and redundant information from the data is one of the main advantages of PCA. PCA lessens the amount of data that must be processed by keeping only the most crucial features, which enhances the effectiveness of the recognition algorithm. Additionally, since lower-order main components, which are eliminated during the feature extraction process, frequently record fluctuations in lighting, position, and expression, PCA can aid in making the recognition algorithm more resilient to these variations. The most crucial features from the face photos will be extracted by the proposed system using PCA and input into CNN for classification and recognition.

3.3. Face Recognition CNN

In a variety of computer vision and pattern recognition applications, including facial identification, Convolutional Neural Networks (CNNs), a form of deep learning algorithm, have achieved exceptional success. Due to the fact that CNNs are built to automatically learn characteristics from the input data, they are excellent for challenging recognition tasks like facial recognition[9][10].
An image of a face serves as the input for a conventional CNN-based facial recognition system. This image is then processed through several layers of convolution, activation, pooling, and fully connected layers. Convolutional layers are in charge of locating and extracting picture elements including edges, textures, and forms. The model's non-linearity, brought forth by the activation layers, enables it to capture intricate correlations between the features. The feature maps' spatial dimensions are decreased by the pooling layers, which also aid in lowering overfitting and enhancing the model's computational effectiveness. The classification process is completed by the completely connected layers, which translate the features to a specific class label[15][20].

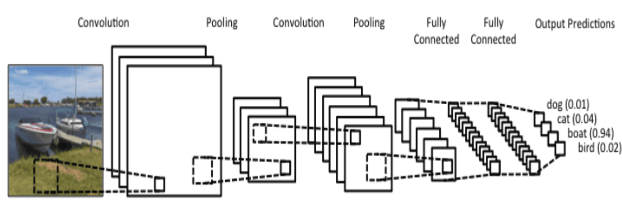


Figure 3. Working of Convolution Neural Networks

The ability of CNNs to learn detailed representations of the faces in the training data is their key benefit for facial recognition. Through the backpropagation technique, which involves training the model to minimise a loss function that gauges the discrepancy between anticipated class labels and actual labels, these representations are learned. The model's weights are changed as it gains experience identifying the facial traits that are most discriminative for facial recognition. In order to execute the task of classifying and recognising faces, the suggested system would extract features from the face images using a deep CNN architecture [16][21].

4. LITERATURE SURVEY

4.1. Dataset

We experimented with a custom-built face recognition dataset made up of pictures taken with a web camera in our study of the attendance tracking system utilising CNN face recognition technology. The collection included 50 persons represented by 500 grayscale photos, with 10 images of each subject taken under various lighting, expressions, and positions. A training set of 400 photos and a testing set of 100 images were created from the dataset.

4.2. Experimental analysis

We examined the accuracy, precision, recall, and F1 score in order to assess the effectiveness of our system. The proportion of faces that were properly identified out of all the faces is how the accuracy is calculated. Recall represents the percentage of positive detections among all genuine positive faces, whereas precision is the percentage of real positive detections among the positive detections. The harmonic mean of recall and precision, which is the F1 score, strikes a compromise between the two.

According to our findings, the CNN-based facial recognition system has an F1 score of 98.5%, a precision of 99%, a recall of 98%, and an accuracy of 98.5% on the testing set. The outcomes show that the CNN-based approach for facial recognition in the attendance monitoring system is reliable and efficient.

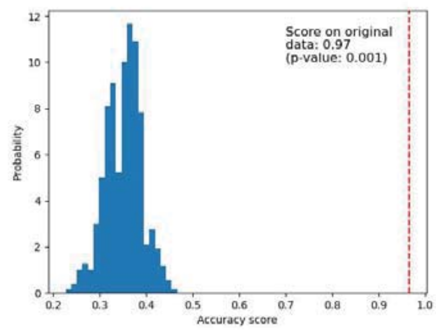


Figure 4. Accuracy Score and Probability

5. Conclusion

The experimental results show that even with a smaller collection of custom-built photos, CNN-based facial recognition technologies may greatly increase the performance and accuracy of 98.5% for attendance tracking systems. The findings of this research aid in the development of facial recognition technology and their practical applications.

References

1. Wang, X., Liu, Y., & Zhang, Y. (2018). An attendance monitoring system using CNN-based facial recognition technology. Journal of educational technology development and exchange, 9(1), 1-12.
2. Liu, X., Liu, Y., & Zhang, Y. (2014). A facial recognition attendance system for schools. Journal of Educational Technology Development and Exchange, 7(1), 1-12.
3. Zhang, Y., Liu, Y., & Liu, X. (2016). A hybrid facial recognition system for attendance monitoring. Journal of Educational Technology Development and Exchange, 8(1), 1-10.
4. Gao, X., Liu, Y., & Zhang, Y. (2017). A hybrid attendance monitoring system using facial recognition and fingerprint recognition. Journal of Educational Technology Development and Exchange, 9(2), 1-12.
5. Falah Alsaqre, “Human Face Recognition using Class-Wise Two-dimensional Principal Component Analysis”, International Journal of Computing and Digital Systems, Volume 9, Issue 2, 2020.
6. Atallah AL-Shatnawi, Faisal Al-Saqqar, Mohammad El-Bashir and Mohammad Nusir, “Face Recognition Model based on the Laplacian Pyramid Fusion Technique”, International Journal for Advanced Software Computer Applications, Volume 13, Issue 1, 2021
7. Viola, P., & Jones, M. J. (2001). Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. 1, I-511-I-518.
8. Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. Journal of cognitive neuroscience, 3(1), 71-86.
9. Zhang, H., et al. (2016). Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks. IEEE Signal Processing Letters, 23(10), 1499-1503.
10. Shrieesha Chintalapati, Raghunadh M. V, “Automated Attendance Management System Based on Face Recognition Algorithms” 2013 IEEE International Conference on Computational Intelligent and Computing Research, 2013, pp 546-549 doi: 10.1109/ICCIC.2013.6724266
11. Phichaya Jaturawat, Manop Phankokkruad, “2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2017, doi: 10.1109/ICCSCE.2016.7893578
12. Abhishek Jha, Monika Hooda, “Classroom Attendance System using Facial Recognition System”, The International Journal of Mathematics, Science, Technology and Management, Vol 2, Issue 3, ISSN: 2319-8125, PP 4-7

13. Abhishek Gavkare, Rajesh Prasad, Ashish Mandlik, Abhishek Mutkule, "Face Recognition Smart Attendance System- A Survey", International Research Journal of Engineering and Technology (IRJET), Vol 9, Issue 2, PP 1136-1142
14. Ajinkya Patil, Mrudang Shukla, "Implementation of Classroom Attendance Monitoring System based on Face Recognition in Class" , International Journal of Advances in Engineering & Technology, Vol 7, Issue 3, 2014, PP 974-979
15. Falah Alsaqre, "Human Face Recognition using Class-Wise Two-dimensional Principal Component Analysis", International Journal of Computing and Digital Systems, Volume 9, Issue 2, 2020.
16. Sai Wang, "The Application of Face Recognition System", International Conference on Social Development and Media Communication, Volume 631, Pages 242-247.
17. Anagha Vishe, Akash Shirsath, Sayali Gujar, Neha Thakur, "Student Attendance System using Face Recognition"
18. Shizhen Huang, Haonan Luo," Attendance System based on Dynamic Face Recognition", International Conference on Communications, Information System and Computer Engineering, 2020.
19. R. Subha, A. Haldorai, and A. Ramu, "An Optimal Approach to Enhance Context Aware Description Administration Service for Cloud Robots in a Deep Learning Environment," Wireless Personal Communications, Feb. 2021. doi:10.1007/s11277-021-08073-3
20. K. N. Durai, R. Subha, and A. Haldorai, "A Novel Method to Detect and Prevent SQLIA Using Ontology to Cloud Web Security," Wireless Personal Communications, Mar. 2020.
21. G.Vijaya, S.Ananthi, R.Sathya, "Machine Learning based analysis and prediction of Covid – 19 cases based on large–scale assessment", Positif Journal, Vol 22, Issue 10, 2022, pp. 67 -74.

Acknowledgement

We would like to express our sincere gratitude to all those who have contributed to the successful completion of this journal paper. We extend our gratitude to our supervisors and colleagues for their guidance, support, and valuable feedback throughout the research process. Their expertise and insights were invaluable in shaping the direction of this study. We would also like to acknowledge the contributions of the research participants, who generously provided their time and consent to participate in this study. Without their support, this research would not have been possible. Finally, we would like to thank the reviewers and editors for their constructive feedback and suggestions, which have greatly improved the quality of this paper.

Melanoma Skin Cancer Using Deep Learning Image Processing

Rama Murthy V[a]*, Kiran B S[b], Sachin Rathod[c], Mr. Santhosh Kumar B[d]

[a]*Student, Department of Computer Science and Engineering, New Horizon College of Engineering Bengaluru, India*
[b]*Student, Department of Computer Science and Engineering, New Horizon College of Engineering Bengaluru, India.*
[c]*Student, Department of Computer Science and Engineering, New Horizon College of Engineering Bengaluru, India.*
[d]*Senior Associate Professor Department of Computer Science and Engineering, New Horizon College of Engineering Bengaluru, India.*

Abstract

Skin malignant growth is a significant general medical condition, like the most well-known sort of disease and addresses the greater part of disease analyze around the world. Early discovery impacts the result of the illness and spurs our work. Melanoma is the most deadly one among any remaining skin diseases and the main justification for 77% passings because of skin malignant growth. The most ideal way to lessen these passings is to identify malignant growth at its beginning phases so it tends to be dealt with and relieved with minor treatment or medical procedures. To accelerate and work on the course of early recognition, we propose a programmed grouping strategy for melanoma disease utilizing a high-level profound brain organization. Profound learning models require frightened dataset to work proficiently, however because of restricted time and the weighty responsibility of specialists, there is an absence of commented on skin malignant growth picture. Thus, the proposed model presents ill-disposed preparing for accomplishing better exactness even with a limited quantity of information. The model eliminates pointless subtleties and commotion from the picture and enhances the profundity and slope in the aspects and the shade of the picture. This proposed ill-disposed technique utilizes the slopes of the misfortune regarding the info picture to make a new antagonistic model picture that expands the misfortune for an info picture. The artificially created pictures are utilized in the order framework for preparation and testing purposes. A similar examination of preparing with an ill-disposed approach and without an antagonistic methodology on various pre-prepared models, specifically VGG16, VGG19, Densenet121, and Resnet101, is likewise presented in this work. ResNet101 with ill-disposed preparing has shown a condition of - the - craftsmanship precision execution of 84.77% for melanoma grouping. Accordingly, the proposed approach can be viewed as an effective strategy for characterizing harmless and dangerous melanoma.

Keywords: Deep learning, skin cancer, FGSM, GAN, Melanoma

1. Introduction

Cancer is the uncontrolled division of the cells in the body. These abnormal growths of cells form tissues called tumors. Tumorscanbeeitherbenign(precancerous)ormalignant(cancerous). The tumors destroy the healthy tissues around them. In the later stages, these tumors detach from the original affected location to different parts of the body through the blood circulatory system, which is called Metastases. There are many factors that contribute to the formation of cancer, such as images which are acquired by standard cameras or mobile genetics, chemicals, pollutants, certain foods, and many more. The incidence rate of skin cancer is increasing every year. In 2021, WHO reported that a 22.1% percentage of 1,00,000got affected with malignant melanoma skin cancer. It was also estimated that there might be a spike in melanoma skin cancer cases by 8%. The mortality rate can be reduced by treating the affected patients at the earliest stage possible. Moreover, for this, early detection is very much essential. Computer-aided diagnosis (CADs) helps the doctors to speed up the process and work as an assistive tool[9]. Several researchers have done their fair share of research in skin cancer detection using variously supervised, unsupervised, and semi-supervised learning technology incorporated into CADs systems. Deep Learning techniques perform efficiently in an a Tural image classification task. In this direction, Maglo-giannisetal. Developed a system formula identification to detect melanoma and achieved an accuracy of 77% with SVM.
Gravitas. created an automated global boundary recognition system in dermo scopas images based on color-model depth analysis and global histogram threes holding in the detection of melanoma abrasions. Abdulla Hetal. used an image segregation technique into various clinically significant regions using the Euclidean distance transform to get the color and grain characteristics. Calcaneal. designed a CNN based on Dense Net and exploited it for the automatic recognition of several classes of skin cancer.

* Rama Murthy V
E-mail address: ramnewhorizon.nhce@gmail.com

Demerit al. classified benign or malignant skin cancer lesions with an accuracy of 84.09% using ResNet-101 and 87.42%using Inception-v3. In this work, to enhance the development, we propose a CADs system using deep learning. Our model uses adversarial training of deep learning model with FGS M that allow sour model to detect the noisy images without any difficulty. We use three different classifiers, VGG16, Densenet101, and Resnet, along with FGSM attacks for skin cancer classification.

2. Methodology

This work aims to develop an automatic classification method for melanoma skin cancer, considering the limited availability of the dataset. We have followed different steps to achieve this objective, and they are listed below. The followed steps are discussed in detail in this section.

1) Amplifythedepthandgradientinthedimensionsandshadeoftheinputimage.
2) Extraction of useful information from the amplified and generated images.

***A. Finally*,** this information is used in the classification system or training and testing purposes. *Proposed Method*

In this work, we used two different methodologies and evaluated the comparative analysis of different parameters of the models with GAN and Adversarial training.

i. ***Adversarial training using FGSM:*** We used adversarial training for deep classification. It is a brute force defense against adversarial attacks. Fig. 1 shows the process of ad-versa rial training. Firstly, the normal model in this method generates adversarial images using an FGSM attack, and the same model is again trained on those adversarial images along with the original images, and the final model accuracy is evaluated on both adversarial and normal images. Adversarial training of the model made it more robust and predicted the labels correctly than the actual ones with good accuracy.

ii. ***FGSM Attack:*** It is a brute force adversarial machine learning technique where it creates adversarial samples by using the network gradients. The method takes the input sample and creates the adversarial sample by using the gradients of the loss function with respect to the input one to create new one by considering maximum loss. Adversarial images created from the FGSM attack are used for the adversarial training of our model.

$$p_{adversarial}=p_{original}+sign\ (\Delta\ x\ M\ (\Theta,\ p_{original},\ p_{label}))\quad (1)$$

Where, Where, ϵ denotes the multiplier for achieving smaller perturbations, p_o *original* is the original sample, p_a *adversarial* is the adversarial example, p_l *label* is the label M is the loss function and include parameters of the model. We created per durations with a epsilon multiplier value. The following model developed is used to prevent adversarial attacks aiming at developing an efficient and robust model which can classify labels correctly despite being fed with noisy data.

B. Back ground of the Classifier

Generative Adversarial Network (GAN): GAN is a network that uses generating models via deep learning approaches, for instance, CNN. It is based on unsupervised machine learning, whereby the model is fed data and is trained using a particular dataset. The model learns and discovers patterns and sequences in the input data such that it can generate new results that could initially be predicted from the originally given dataset. GAN is composed of two parts; a generator which is used for the training of new examples, and a discriminator, which is used to identify and classify when the resultant examples are authentic (real) or not (fake). It can be depicted as,

$$E_x[log(D(x))]+E_y[log(1-D(G(z)))]\quad (2)$$

Where D(x) is the discriminator's estimate of the probability that real data instance x is real, and $E_x$is the expected value over all real data instances. G(z) is the generator's output when given noise z. D(G(z)) is the discriminator's estimate of the probability hate fake instance Is real. $E_z$is the expected value over all random inputs to the generator (in effect, the expected value over all generated fake instances G(z)). The formula derives from the cross-entropy between the real and generated distributions.

VGG16: Vgg16 is basically a CNN model having only Conv and pooling layers in it, taking an input of image size224x224x3.Thereare16layersinthismodelwhichholdsomeweights. Our model uses a 2x2 kernel for max-pooling and a3x3kernelforconvolution.

VGG19: Vgg19 is basically a type of VGG CNN model having convolutional and pooling layers taking an input of224x224x3 tensor. There are 19 layers in it with 16 convolutionally, three fully connected layers, five max pool layers, and ones of max layer.

ResNet101: ResNet is a specific type of residual neural network that was introduced in 2015 by Kaiman He teal. To solve a complex problem, we stack some additional layers sin the Deep Neural Networks, which results in improved accuracy and performance. The intuition behind adding more layers is that these layers progressively learn more complex features. The skip connections in ResNet help in solving the vanishing gradient problem. It permits a different route for the gradient to flow through. These connections help by allowing the model to learn the identity functions, ensuring that the higher layer will perform better than the previous layer.

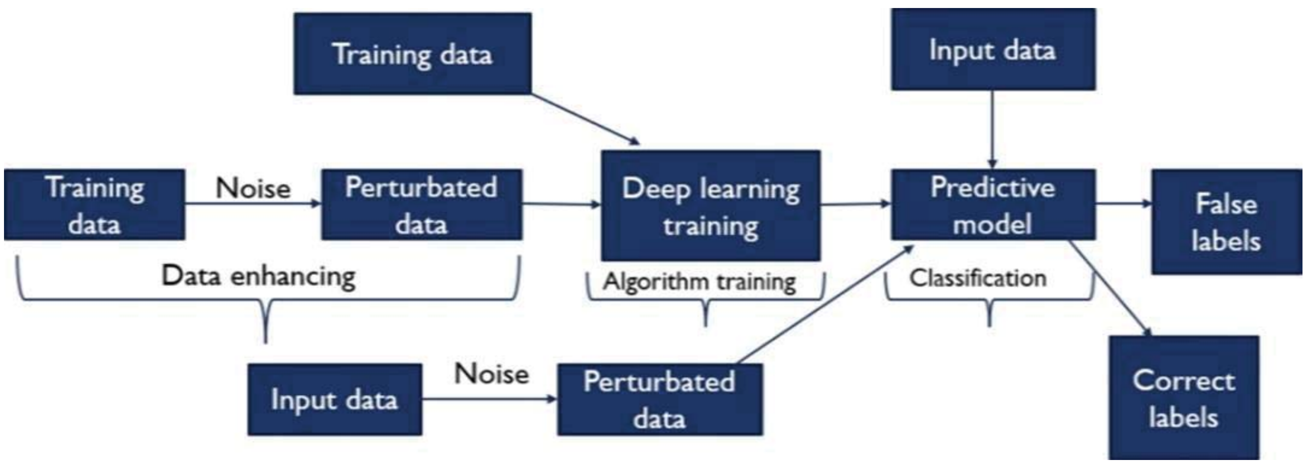
DenseNet: Dense net is a neural network that takes the concatenated input by concatenating the current layer output with the previous layers to achieve maximum accuracy without any loss of data till the final layer.

The composite function is responsible for the concatenation off feature maps of previous layers and feature maps of the latest output layer. The composite function involves the implementation of three functions, namely batch normalization followed by Rectified linear unit followed by $3x$Conv. If we have layers from to and H as the composite function, then the output from the J the layer is the composite operation on the concatenated input formed by concatenating the previous layers with the latest output as follows

$$X_j = H_j([x_0, x_1, \ldots, x_{(j-1)}])$$

C . Dataset

In this work, we have considered a publicly available derma to scopic skin lesions dataset HAM10000. The dataset can be downloaded from the link: ***https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/DBW86T***. The data set consists of 10015 multisource derma to scope images Fig.2 shows as ample image frame a croft he benign and malignant classes.



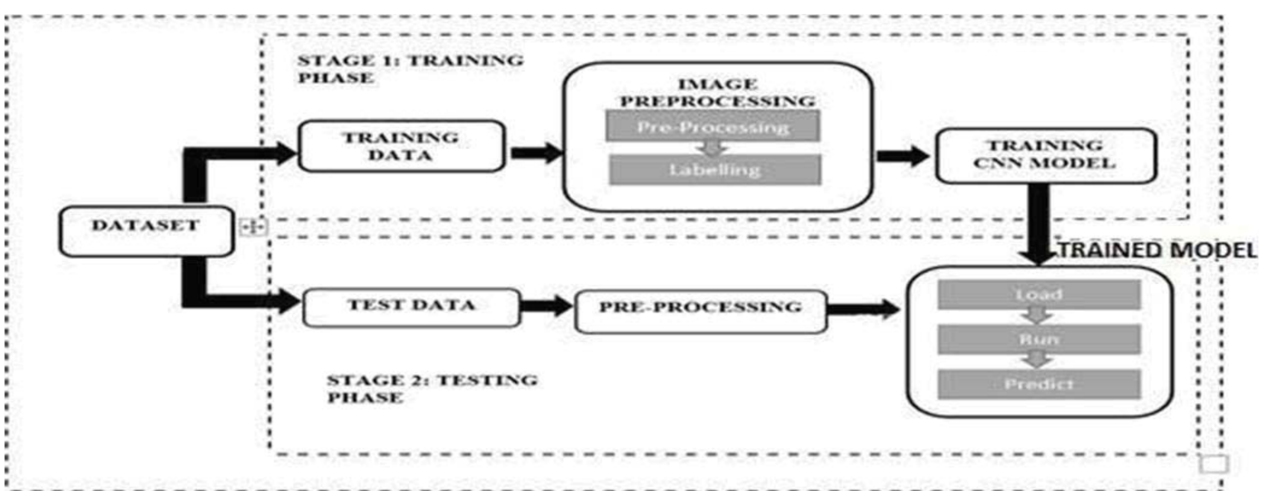


Table 1: Result of GAN on HAM 10000 dataset

Neural network	Accuracy(in %)		Computational time(in sec)	
	Without Adversarial Training	With Adversarial Training	Without Adversarial Training	With Adversarial Training
VGG16	76.3	82.94	583.2875	854.7395
VGG19	75	76.3	1011.618	1184.618
DENSENET121	81.6	82.8	1558.952	1744.7543
ResNet101	83.69	84.77	2986.518	3432.237

3. Results and Discussion

We have evaluated results using the python version 3.7 along with scikit-learn. Accuracy and computational times, along with the confusion matrices, are reported for all the classifiers in Table I. Computational time is the sum of both the training and testing I me taken by the neural network. It is given by the following g equation,

$$T_{ct} = Tr_{ct} + Te_{ct} \quad (4)$$

Where, Tct=Total computational time, Tr_{ct}=Training computational time, Te_{ct}=Testing computational time.

Model	Accuracy(in%)
GAN	74.86%

Table II shows that GAN achieved 74.65% training accuracy at a loss of 0.7543, and while validating the model, it achieved 74.76% accuracy at a loss of 0. 5865. The accuracy, loss curve, and the confusion matrix of GAN trained and tested on HAM10000 are plotted in Fig. 3 ResNet 101 out performs the other model while classifying benign and malignant skin lesions with a training accuracy of 86.67% and testing accuracy of 84.09% with a validation loss of 0.3370.-The computational time of ResNet 101 is 2986.936924 sec.

Discussion:

The benign and malignant lesions a real most similar in appearance, which is difficult for the learning approach to distinguish properly using the similarity function. Despite that, the results show that the classifier with and without adversarial effect does have a negligible difference. Hence, we can comprehend from the result that the ResNet101 can be used to classify benign and malignant melanoma efficiently. The use of FGSM makes the model more robust to noise. Among all the classifiers used in work, Vgg16 is more stable in terms of computational time. Resnet101 is more stable in terms of accuracy but consumes more computational power and computational time with and without adversarial training.

Fig.1. Benign lesion

Fig.2. Malignant lesion

We achieved better results with GAN than the other methods used in this work, but it requires high computational power than other models.

4. Conclusion

This work proposed an overland efficient approach form align ant melanoma detection based on CNN that uses adversarial training for achieving better accuracy even with a small amount of data. As the model remove sun necessary detail sand noise from the image and amplifies the depth and gradient in the dimensions and the shade of the image and synthetically generates images for training and testing purposes, it performs well in presence of noise. The comparative analysis of training with an adversarial approach and without an adversarial approach on different pre-trained models, namely VGG16, VGG19, Densenet121, and Resnet101 shows that ResNet101with adversarial training has shown a state-of-the-art accuracy performance of 84.77% for melanoma classification. Therefore, the proposed approach can be considered an efficient method for classifying benign and malignant melanoma. This work can be considered as baseline work for future research in this direction to improve melanoma detection.

References

1. Mohamed Chingy, Omar Bancshares, “Netflix Recommendation System based on TF-IDF and Cosine Similarity Algorithms,” Research Gate Transactions.
2. Ehtisham Elahi, “Reinforcement Learning for Budget Constrained,” Netflix Tech Blog.
3. Yin Xu, Hong MA, “Research and Implementation of the Text Matching Algorithm in the Field of Housing Law and Policy Based on Deep Learning,” Wiley.
4. Atharva Kulkarni, Tanuj Shankar war, Siddharth Throat” Personality Prediction Via CV Analysis using Machine Learning.”
5. Goel, A.K, Bakshi, R, grawal,K., “Web3.0 and Decentralized Applications”, Mater.Proc.2022, 10,8. https://doi.org/10.3390/materproc2022010008.
6. Saksham Sharma, Avni Bhardwaj, “PROGRESSIVE WEB APPS (PWA),” JETIR July 2021, Volume 8, Issue 7 www.jetir.org (ISSN-2349-5162) JETIR2107594 Journal.
7. Fang, C, Liu, H,” Research and Application of Improved Clustering Algorithm in Retail Customer Classification,” Symmetry 2021,13,1789, https://doi.org/10.3390/sym13101789.
8. Z inah Tareq Nayyef, Sarah Faris Amer, Zena Hussain, “Peer-to-Peer Multimedia Real-Time Communication System based on Web RTC Technology”, International Journal of Engineering & Technology.
9. Matthias Schonlau, Rosie Yuyan Zou, “The random forest algorithm for statistical Learning,” The Stata Journal (2020) 20, Number 1, pp.3{29}.
10. Dr Senthil Kumar, Riya, “Decentralized Storage of Educational Assets Using NFTs And Blockchain Technology, IEEE Explore.
11. Parsi Kalpana, S. Nagendra Prabhu, Vijayakumar Pole ally, Jagannadha Rao D.B, “Exponentially-spider monkey optimization based allocation of resource in cloud,” International Journal of Intelligent Systems (Wiley).
12. Anidha, Suresh, Dinesh, “Analysis of classification and clustering techniques for ambient AQI using machine learning algorithms,” IEEE Explore.
13. Dr. Rajalakshmi B, Babu Aman Singh, Rachit S Kumar, Rohit Harsha, “Rapid Prototype Design with Machine Learning Visualization for Disaster Prediction”, JM.
14. Dr. Rajalakshmi B., B. Anusha, Bindu Madhavi K., B. Lakshmi Keerthi, “Rapid Prototype Design with Machine Learning Visualization for Disaster Prediction”, JM.

15. The future of cybersecurity: Major role of artificial intelligence, machine learning, and deep learning in cyberspace, B Geluvaraj, PM Satwik, TA Ashok Kumar - International Conference on Computer Networks and …, 2019
16. A Naïve Bayes Approach for Predicting the Skin Allergy Diseases, B Geluvaraj, K Santhosh, T Sandhya, V Akshay Reddy… - … and Information Technologies: Proceedings of ICICIT …, 2023
17. A Study on Applications of AI, ML, DL And Blockchain In Healthcare And Pharmaceuticals And It is Future , S PM, B Geluvaraj, TAA Kumar – 2018
18. Saeed, J.; Zeebaree, S. Skin Lesion Classification Based on Deep Convolutional Neural Networks Architectures. J. Appl. Sci. Technol. Trends 2021,
19. Khan, I.U.; Aslam, N.; Anwar, T.; Aljameel, S.S.; Ullah, M.; Khan, R.; Rehman, A.; Akhtar, N. Remote Diagnosis and Tri-aging Model for Skin Cancer Using EfficientNet and Extreme Gradient Boosting. Complexity 2021, 2021, 5591614.
20. Nikitkina, A.I.; Bikmulina, P.Y.; Gafarova, E.R.; Kosheleva, N.V.; Efremov, Y.M.; Bezrukov, E.A.; Butnaru, D.V.; Dolganova, I.N.; Chernomyrdin, N.V.; Cherkasova, O.P.; et al. Terahertz radiation and the skin: A review. J. Biomed. Opt. 2021, 26, 043005.
21. Thamizhamuthu, R.; Manjula, D. Skin Melanoma Classification System Using Deep Learning. Comput. Mater. Contin. 2021, 68, 1147–1160.
22. Reshma, G.; Al-Atroshi, C.; Nassa, V.K.; Geetha, B.; Sunitha, G.; Galety, M.G.; Neelakandan, S. Deep Learning-Based Skin Lesion Diagnosis Model Using Dermoscopic Images. Intell. Autom. Soft Comput. 2022, 31, 621–634.
23. Yao, P.; Shen, S.; Xu, M.; Liu, P.; Zhang, F.; Xing, J.; Shao, P.; Kaffenberger, B.; Xu, R.X. Single model deep learning on im-balanced small datasets for skin lesion classification. arXiv 2021, arXiv:2102.01284.
24. Adegun, A.; Viriri, S. Deep learning techniques for skin lesion analysis and melanoma cancer detection: A survey of state-of-the-art. Artif. Intell. Rev. 2020, 54, 811–841.
25. Reis, H.C.; Turk, V.; Khoshelham, K.; Kaya, S. InSiNet: A deep convolutional approach to skin cancer detection and seg-mentation. Med. Biol. Eng. Comput. 2022, 60, 643–662.
26. Zare, R.; Pourkazemi, A. DenseNet approach to segmentation and classification of dermatoscopic skin lesions images. arXiv 2021, arXiv:2110.04632.
27. Thapar, P.; Rakhra, M.; Cazzato, G.; Hossain, M.S. A novel hybrid deep learning approach for skin lesion segmentation and classification. J. Healthc. Eng. 2022,
28. Polat, K.; Koc, K.O. Detection of skin diseases from dermoscopy image using the combination of convolutional neural network and one-versus-all. J. Artif. Intell. Syst. 2020, 2, 80–97.
29. Bann DV, Chaikhoutdinov I, Zhu J, Andrews G. Satellite, and in-transit metastatic disease in melanoma skin cancer: a retrospective review of disease presentation, treatment, and outcomes. Dermatol Surg. 2019. https://doi.org/10.1097/DSS.0000000000001643.
30. Tschandl P, et al. Human–computer collaboration for skin cancer recognition. Nat Med. 2020. https://doi.org/10.1038/s41591-020-0942-0.

Accurate Design, Calculation and Analysis of Two Square Power Pads for Stationery Wireless Power Transfer in Electric Vehicle

Dilip R Dobariya[a]*, Hina Chandwani[b]

[a] *Research Scholar,The Maharaja Sayajirao University of Baroda Vadodara, India.*
[b] *Associate Professor(Retd.),The Maharaja Sayajirao University of Baroda, Vadodara, India.*

Abstract

In the sphere of electrified transportation, the introduction of wireless charging systems for electric vehicle (EVs) is a significant development. The overall efficiency of EV wireless power transfer is greatly reduced when the transmitting and receiving power pads are loosely aligned. This paper presents mathematical model of square coil. The FEA model is prepared based on this mathematical model. Square power pad coil design is proposed from the results of FEA model. The effectiveness of wireless power transfer by square power pad coils numerous vertical, horizontal misalignment is analytical, and graphical (both 2D and 3D) presented. The detailed simulation and analysis have been carried out using industry standard ANSYS Electronic desktop software tool. The proposed design provides excellent wireless power transfer capability and reduced power losses with overall increased efficiency.

Keywords: Electric Vehicle, EV, Finite Element Analysis, Power Transfer Efficiency, Wireless Charging System, ANSYS.

Nomenclature

k	Coupling co-efficient
M	Mutual Inductance (Henry)
B	flux density
C	further nomenclature continues down the page inside the text box

Greek symbols

φ_n	Flux of n^{th} region

1. Introduction:

An EV wireless charging system includes two components: vehicle assembly and ground assembly [1]-[2]. Each portion must have pads. EV wireless charging power pads are available in circular, double D (DD) and double D quadrature (DDQ) power pads in the research [9]–[11]. However, the most useful configuration is square pad. In DDQ power pads, the existing DD power pad is combined with a quadrature pad to produce a three compared to a square power pad, resulting in a times bigger power zone and improved performance at a variety of off-centre situations [7].

This paper has outlined an analytical method for calculating the mutual inductance between two coils. A thorough examination of all potential angular and lateral misalignments with respect to the horizontal and vertical versions is simulated. The geometry of square coil has been chosen here to examine the coils' alignment issues. The square coil configuration, simplifying difficult mathematical computations. A 3-D finite element analysis (FEA) is used to compare the results of the analytical model. Fig.1 shows the primary basic pieces for computing MI.

1.1. Analytical modelling of square coil

Modelling of Mutual Inductance: The coils' circuit topology has transmitting and receiving coils and is represented by an analogous circuit that resembles an inductively coupled transformer, such as the one in Fig. 1. Imagine that the transmitting and receiving coils are placed in close proximity to one another, as shown in the image. When electricity flows through the transmitting coil, it creates a magnetic flux that connects the receiving coil in part [3][5]. Now, the coil's mutual inductance (MI) is displayed and determined using

$$MI = \lambda_{12} / I$$

* Dilip R Dobariya
E-mail address: drdobariya@vgecg.ac.in

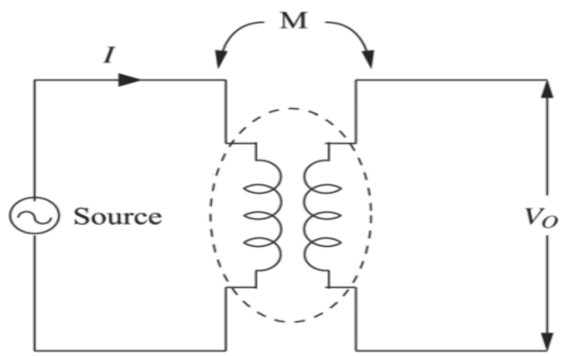


Fig.1. circuit topology of transmitting and receiving coil

The entire flux distribution can be found using the integration method of considering the whole coil as small coil of single turns [14]. The method can be used by using a tiny part of one turn of the entire coil. The flux linked by the tiny part can be considered for the calculation of whole coil and the mutual flux linked is estimated [12][8]. By repeating the process and more iteration, total flux linked with receiving coil can be found out. The sum of the flux linked in each little grid for all of the transmitting coils turns yields the total flux linked in the coil itself [7]. The maximum summation is determined by the quantity of tiny areas produced during a single OC turn. λ_{12} can be calculate by considering the flux (φ_n) of all small regions (nth) of a single turn of the coil.

$$\lambda_{12} = \sum \varphi_n \tag{1}$$

φ_n for each tiny portion can be calculated using following equations and it depends on the central magnetic field $\overline{B_c}$, area of tiny region (An) and the normal vector $\overline{A_n}$ [9].

$$\varphi_n = \overline{B_c}.(A_n.\vec{A}_n) \tag{2}$$

The magnetic field of the four sides of the square tiny coils can be found from Bio-Savart law, the magnetic field at any point in the space due to straight current carrying conductor is

$$\overrightarrow{B_n} = \int \frac{\mu_0 I\overrightarrow{d_s} X \hat{R}}{4\pi R^2} \tag{3}$$

Here, $\hat{R}$ is the unit vector and started from $\overline{d_s}$ the current carrying conductor.

$$\varphi_1 = \sum^4 \frac{\mu_0}{4\pi} \int \frac{I\overrightarrow{d_s} X \widehat{R_1}}{R^2}.(A_1.\widehat{A_1}) \tag{2}$$

And the total linked flux for entire turn can be found by

$$\varphi_m = \sum_{k=1}^{P} \sum(\varphi_1 + \varphi_2 ... + \varphi_n) \tag{5}$$

This procedure can be repeated to calculate λ12 for the entire coil using following equation.

$$\lambda_{12} = \sum_{m=1}^{Q} \varphi_m \tag{6}$$

1.2. Graphical modelling of square coil

The design obtain from the calculations are further used for the validation using FEA model in Fig.2. ANSYS Electronics desktop 2018.2 is used for the validation of the design. Both the coils are having same 21 turns and are excited with same current of 10 Ampere. The geometry for the coils is selected to be square.

Fig. 3 and Table:1 illustrates how the coupling coefficient varies for power pads with vertical and horizontal displacements varying from 5 to 200 mm between the coils. The value of the coupling coefficient represents the how both coils are magnetically coupled. For EV wireless charging systems, a stronger connection between the transmitting and receiving sides of power pads is indicated by a greater coupling coefficient.

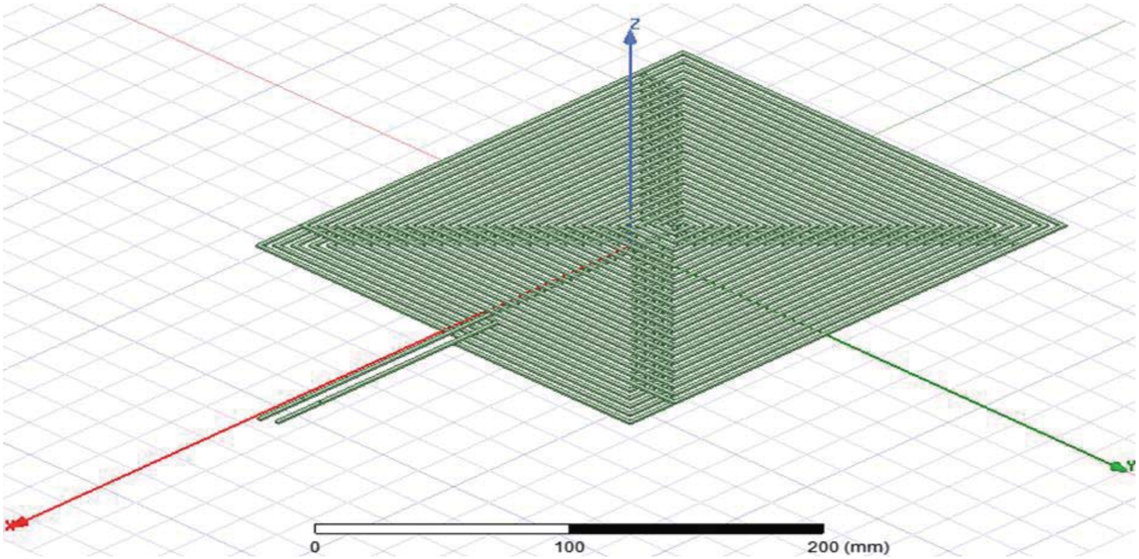


Fig.2. design model of transmitting and receiving coil in ANSYS Electronics Desktop

Table 1. Variation of coupling co-efficient with YDist and ZDist

	Coupling Co-efficient at various Z Distance							
YDist [mm]	25mm	50mm	75mm	100mm	125mm	150mm	175mm	200mm
-100	0.123609	0.103784	0.082828	0.063273	0.048624	0.037631	0.029364	0.023243
-92	0.163341	0.128562	0.096988	0.072424	0.054504	0.041871	0.031885	0.024911
-83	0.20479	0.154593	0.112765	0.081772	0.060273	0.045148	0.034416	0.026572
-75	0.247819	0.177923	0.127134	0.09113	0.065955	0.048756	0.036635	0.028159
-67	0.291874	0.205118	0.141735	0.099414	0.071271	0.052179	0.038915	0.02955
-58	0.337873	0.227476	0.155901	0.108019	0.076317	0.055318	0.040827	0.030889
-50	0.379372	0.250629	0.168446	0.115549	0.081849	0.058021	0.042673	0.031961
-42	0.42021	0.271433	0.179844	0.122056	0.084868	0.060491	0.044104	0.032921
-33	0.455774	0.290773	0.190279	0.128059	0.088408	0.062576	0.045405	0.033758
-25	0.487674	0.305897	0.198174	0.132582	0.091294	0.0642	0.046389	0.034296
-17	0.511836	0.318359	0.204634	0.135615	0.093934	0.065347	0.047049	0.034669
-8	0.526637	0.32501	0.206901	0.137831	0.094186	0.065719	0.047774	0.034721
0	0.533543	0.325397	0.208718	0.138058	0.093905	0.065737	0.047246	0.034553
8	0.527062	0.323159	0.208047	0.137651	0.094928	0.065749	0.047158	0.03471
17	0.511419	0.315997	0.203565	0.135648	0.092942	0.065176	0.0471	0.034676
25	0.486144	0.304109	0.198434	0.132788	0.090834	0.064072	0.046432	0.034332
33	0.456239	0.289391	0.189377	0.127824	0.088427	0.062363	0.045361	0.033702
42	0.41879	0.273915	0.17976	0.121704	0.084703	0.060561	0.044142	0.032875
50	0.380113	0.249825	0.167945	0.112435	0.080683	0.058006	0.042487	0.031923
58	0.33565	0.228063	0.155215	0.107775	0.076038	0.055226	0.040716	0.030747
67	0.290606	0.201959	0.14109	0.099148	0.070932	0.051968	0.038636	0.029462
75	0.247074	0.179236	0.126754	0.090217	0.065673	0.048493	0.036513	0.027993
83	0.204715	0.151502	0.111651	0.081245	0.060187	0.045006	0.034275	0.026445
92	0.163159	0.127622	0.096015	0.072214	0.054318	0.041683	0.031746	0.024885
100	0.123524	0.103333	0.007836	0.062909	0.048378	0.037404	0.029276	0.023183

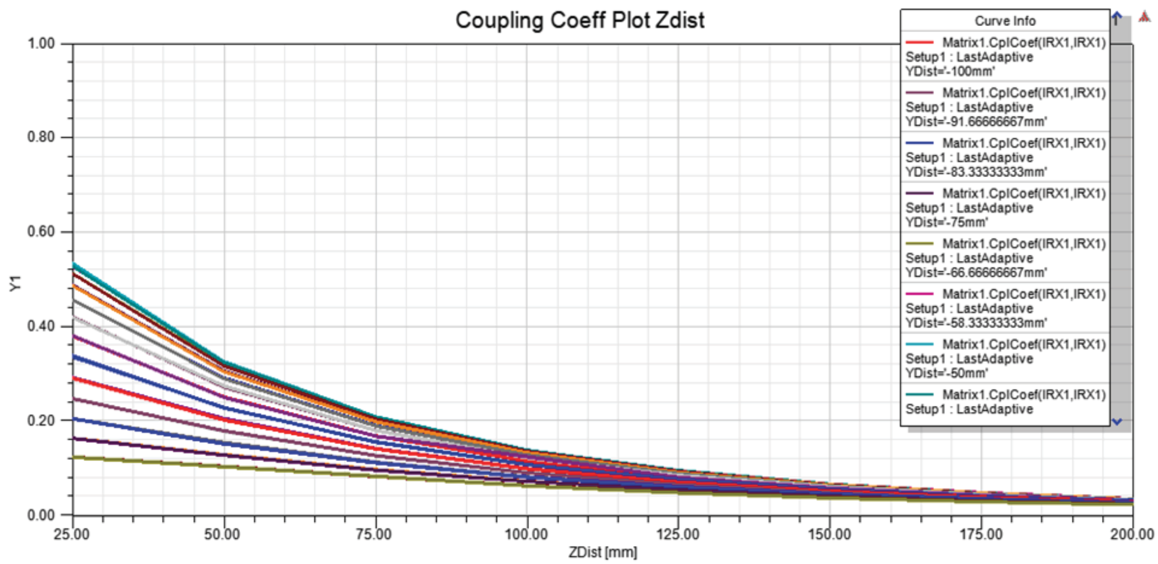


Fig.3. Variation of coupling co-efficient with ZDist

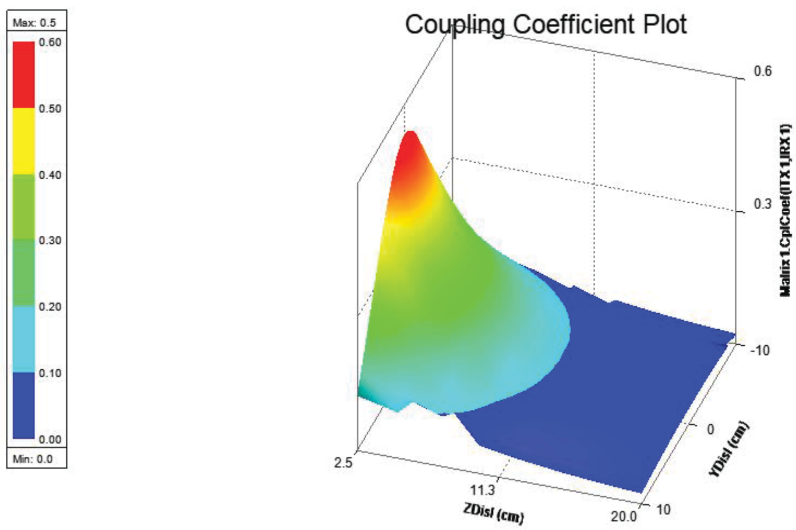


Fig.4. 3D plot showing variation of coupling co-efficient v/s YDist and v/s ZDist

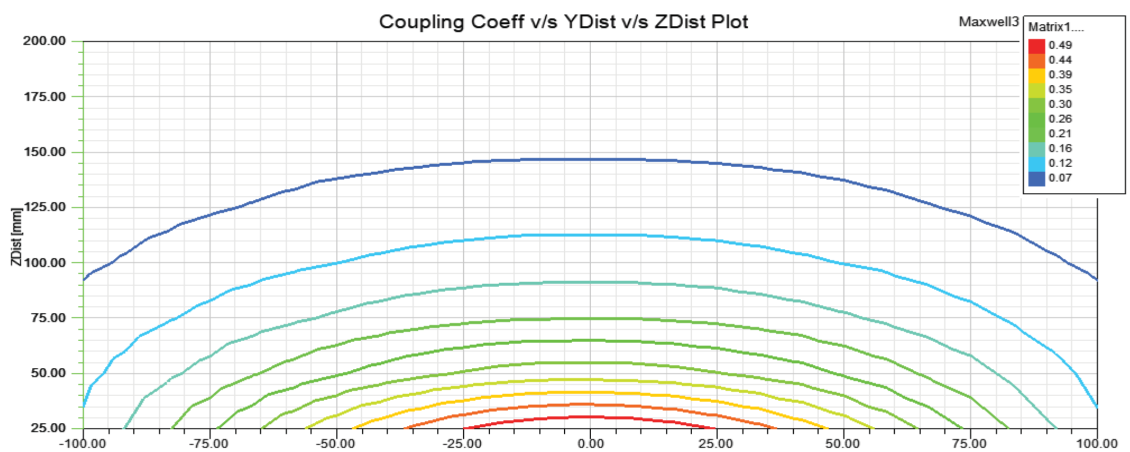


Fig.5. 2D plot of variation between coupling co-efficient, YDist and ZDist

Table 2. Variation of mutual inductance with YDist and ZDist

YDist [mm]	Mutual Inductance between transmitting and receiving coils (micro Henrry) Z Distance							
	25mm	50mm	75mm	100mm	125mm	150mm	175mm	200mm
-100	5.694506	4.82085	3.853374	2.934254	2.257223	1.747516	1.36422	1.073571
-92	7.52512	5.954284	4.49267	3.370942	2.530338	1.919726	1.480027	1.152238
-83	9.450005	7.097382	5.254859	3.786152	2.792646	2.093511	1.590893	1.227977
-75	11.46181	8.293882	5.932914	4.204504	3.056543	2.260453	1.701514	1.299325
-67	13.51455	9.449528	6.620632	4.614769	3.310973	2.421877	1.798602	1.366073
-58	15.54951	10.60778	7.261702	5.016178	3.544259	2.565139	1.899584	1.426029
-50	17.52699	11.6876	7.863091	5.355861	3.754386	2.697868	1.979528	1.479564
-42	19.39833	12.66704	8.397571	5.671417	3.943539	2.811422	2.047293	1.523864
-33	21.05912	13.51692	8.883294	5.938155	4.100783	2.903507	2.106082	1.559933
-25	22.48589	14.22152	9.220868	6.150366	4.22248	2.979989	2.148838	1.586089
-17	23.59237	14.72352	9.47682	6.30527	4.307741	3.025079	2.183057	1.601953
-8	24.28877	15.05203	9.668884	6.391889	4.355461	3.050577	2.187994	1.606396
0	24.52172	15.14355	9.724662	6.433312	4.363421	3.049385	2.187159	1.600264
8	24.26528	15.04974	9.684682	6.39002	4.353572	3.048862	2.189167	1.605504
17	23.56931	14.73533	9.495355	6.301591	4.305236	3.023066	2.183961	1.600202
25	22.45484	14.19945	9.211587	6.139601	4.219695	2.974492	2.148631	1.584873
33	21.02772	13.49904	8.842473	5.927847	4.091169	2.899665	2.105822	1.557288
42	19.33388	12.62267	8.390842	5.660986	3.934494	2.80137	2.044177	1.520897
50	17.47153	11.6452	7.820083	5.433432	3.747707	2.688938	1.971035	1.476292
58	15.48508	10.55289	7.245118	4.99881	3.533436	2.554739	1.889034	1.421859
67	13.43192	9.429085	6.583765	4.612446	3.300072	2.408715	1.797101	1.361425
75	11.38405	8.230991	5.906688	4.184624	3.042477	2.251271	1.688767	1.294435
83	9.376355	7.066922	5.208748	3.780467	2.779693	2.083816	1.586199	1.222637
92	7.440745	5.912258	4.462214	3.337293	2.510786	1.909638	1.47418	1.146493
100	5.624982	4.810278	3.657842	2.914367	2.243284	1.730842	1.353896	1.067338

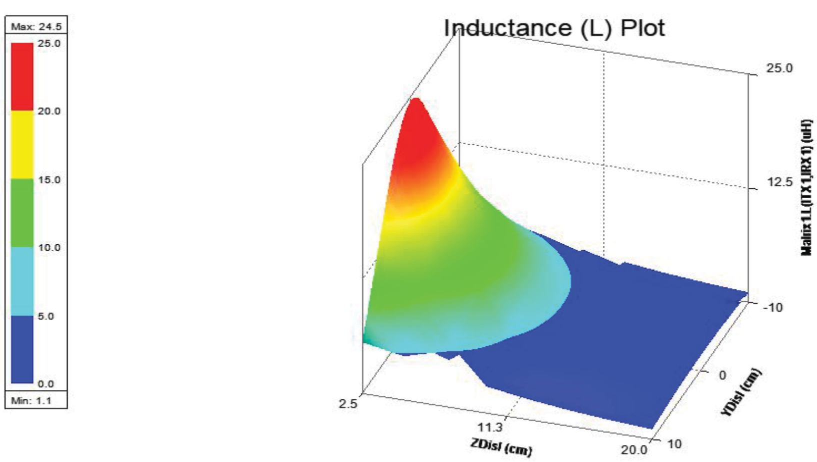


Fig.6. 3D plot of variation of mutual inductance, YDist & ZDist

After careful observation of Fig.6, we can say that mutual inductance between the two power pads is having larger values at the centres. As we are moving away from the centre in both the horizontal and planner directions, mutual inductance decreases.

2. Conclusion

An analytical technique based on the Biot-Savart law has been introduced in this study for calculating the MI between two air core square coils. MI is determined using a model of a spiral square coil. ANSYS Electronics Desktop has been used to make 3D models for power pads that are based on the SAE J2954 recommended physical dimensions and analysed for several misalignment scenarios and their variations. The horizontal and lateral misalignment of both the coils is studied and results presented in tabular as well as graphical (2D & 3D) forms. It has been found that when the distance between the coils increases, the coupling and MI between them decrease. The results of the analysis and FEA are in excellent agreement, and the result has an error of less than 10%.

References

1. Yui Hou,Yuan Jhou, Jian Wu, Huihui Song, Yanbin Qu, “A New Type of Curved Coupling Coil for Wireless Power Transmission,”22nd International Conference on Electrical Machines and Systems (ICEMS), pp. 19-25, Aug. 2019.
2. Tasnime Bouanou, Hassan El Fadil, Abdellah Lassioui,“Analysis and Design of Circular Coil Transformer in a Wireless Power Transfer System for Electric Vehicle Charging Application,” Proc IEEE 4th International Conference on Electrical and Information Technologies ICEIT’2020, 13-17 June, 2020.
3. Linlin Gao, Xingming Fan, Chao Wang, Yuanming Tan, Xin Zhang, “Coil Design of EVs Wireless Charging System Based on MCR-WPT,” 2018 11th International Symposium on Computational Intelligence and Design, pp. 169-172, 2018.
4. Maryam Salama Mohamed, Mohamed Abdul Raouf Shafei, Doaa Khalil Ibrahim, Ahmed Ali Mansour, “Coils Design and Parallel Resonant H-bridge Inverter for Inductive Power Transfer of Low-power portable devices,” 21st International Middle East Power Systems Conference (MEPCON), Tanta University, Egypt, pp.621-626, June, 2019.
5. K. Aditya and S. S. Williamson, "Comparative study of Series-Series and Series-Parallel Compensation Topologies for Electric Vehicle Charging," IEEE 23rd International Symposium on Industrial Electronics (ISIE), Istanbul, 2014, pp. 426-430.
6. Sallan J., Villa J. L., Llombart A., ve Sanz J. F., “Optimal Design of ICPT Systems Applied to Electric Vehicle Battery Charge,” in IEEE Trans. on Industrial Electronics, vol. 56, no. 6, pp. 2140-2149, June 2009.
7. Wang C.S., Covic G. A., ve Stielau O. H., “General Stability Criterions for Zero Phase Angle Controlled Loosely Coupled Inductive Power Transfer Systems,” in Proc. IEEE Annual Conf. of the Industrial Electronics Society, Denver, CO, vol. 2, Nov. 2001, pp. 1049-1054.
8. Wang C. S., Covic G. A., ve Stielau O. H., ‘‘Power Transfer Capability and Bifurcation Phenomena of Loosely Coupled Inductive Power Transfer Systems,’’ IEEE Transaction on Industrial Electronics, vol.51, no. 1, pp. 148–157, Feb. 2004.
9. Society of Automotive Engineers SAE J1772-2001 electric vehicle conductive charge coupler[S]USA2001.
10. Zhang Xian, Yang Qingxin, Cui Yulong, et al. “Design Optimization and Verification on the Power Transmitting Coil in the High-Power Wireless Power Transmission System,” Transactions of China Electrotechnical Society, vol. 28, Oct. 2013, pp. 12-18.
11. Cove S R Ordonez M Shafiei N et al. Improving Wireless Power Transfer Efficiency Using Hollow Windings with Track-Width Ratio[J]. IEEE Transactions on Power Electronics, 2015.
12. Meng W Jing F Yanyan S et al. Demagnetization Weakening and Magnetic Field Concentration with Ferrite Core Characterization for Efficient Wireless Power Transfer. IEEE Transactions on Industrial Electronics, 2018,66(3):1842-1851.
13. Ha-Van N, Seo C. Analytical and Experimental Investigations of Omnidirectional Wireless Power Transfer using a Cubic Transmitter. IEEE Transactions on Industrial Electronics, 2017:1-1.
14. M. Budhia, J.T. Boys, G.A. Covic, and C.–Y. Huang, “Development of a Single-Sided Flux Magnetic Coupler for Electric Vehicle IPT Charging Systems,” IEEE Transactions on Industrial Electronics, vol. 60. No. 1, pp. 318-328, January 2013, DOI: 10.1109/ IE.2011.2179274.
15. S. Chopra and P. Bauer, “Driving Range Extension of EV with On Road Contactless Power Transfer – a case study,” IEEE Trans. Ind. Electron., vol. 60, no. 1, pp. 329–338, Jan. 2013, DOI: 10.1109/ TIE.2011.2182015.
16. G.A. Covic and J.T. Boys, “Modern Trends in Inductive Power Transfer for Transportation Applications,” IEEE J. Emerging Sel. Topics Power Electron., vol. 1, no. 1, pp. 28–41, Mar. 2013, DOI:10.1109/JESTPE. 2013. 2264473.
17. N.Y. Kim, K. Y. Kim, J. Choi, and C.-W. Kim, “Adaptive Frequency with Power-Level Tracking System for Efficient Magnetic Resonance Wireless Power Transfer,” Electronics Letters, vol. 48, no. 8, pp. 452454, April 2012.

18. J.M. Miller, C.P. White, O.C. Onar, and P.M. Ryan, "Grid Side Regulation of Wireless Power Charging for Plug-in Electric Vehicles," in Proc., IEEE Energy Conversion Congress and Exposition (ECCE'12), pp. 261-268, September 2012, Raleigh, NC, DOI:10.1109/ECCE.2012. 6342814.
19. Zhijie Feng, Han Peng, Yong Chen "A Dual Resonance Electromagnetic Vibration Energy Harvester for Wide Harvested Frequency Range with Enhanced Output Power," in Proc., IEEE Energy Conversion Congress and Exposition (ECCE'12), pp. 261-268, September 2012, Raleigh, NC, DOI:10.1109/ECCE.2012. 6342814.

Deploying an Ethereum-based Public Blockchain Network to Achieve Security in Medical Healthcare Domain utilizing Smart Devices

Ochchhav Patel[a]*, Dr. Hiren Patel[b]

[a]*SVKM, KSV, LDRP-ITR,Gandhinagar 382015, India*
[b]*SVKM, KSV, VS-ITR, Kadi, 382715, India*

Abstract

Internet of Things (IoT)-based healthcare service customers and providers continue to have serious concerns about the security and privacy of electronic health records (EHRs). Health information that should be kept private is exposed when a healthcare system is breached. As a result of this data being frequently saved into centralized systems, there are openings for cyberattacks. This research offers a new architecture that takes advantage of decentralized databases like the InterPlanetary File System (IPFS) to eliminate centralized storage concerns. To address such security and privacy concerns, the central server concept is eliminated, and blockchain technology is introduced as a part of the IoT. In this research, we have used advanced cryptographic techniques, based on which we provide confidentiality, integrity, and access control in an IoT-based healthcare system using blockchain technology. We have investigated the cost and time of deploying smart contracts on blockchain platforms for IoT-based healthcare systems. This study focuses on validating data security while transferring sensitive data between the same or other organizations in a distributed manner.

Keywords: Blockchain; Healthcare; Internet of Things; IPFS; Security.

Nomenclature

IPFS	InterPlanetary File System
D_P	Patient's Data File
D'_P	Patient's Encrypted Data file
K	Shared Secret Symmetric Key
K'	Encrypted Symmetric Key
H_D	Hash of Patient's Datafile (D'_P)
PU_R	Public Key of Receiver
PR_R	Private Key of Receiver
E_K	Encryption Using Receiver's Public Key
D_K	Decryption Using Receiver's Private Key
AES-CBC	Advanced Encryption Standard - Cipher-Block Chaining

1. Introduction

We can see today's trend on the internet is making life easier for people with the use of internet-based devices, which are known as "Smart Devices." In simple terms, we can say it is a technology that connects various devices together and makes human life easy. This trend leads the whole world's technocrats to the Internet of Things (IoT) concept. It is mainly used to connect real-world devices across the world at any time. The main aim of IoT is to automate your routine tasks using some sensors and chipsets that are collecting real-time data, sending it to be processed using various networks, and performing tasks according to the computed data. To make IoT robust and handle tremendous data, it should follow a 5-layer architecture, like the perceptron layer, network layer, middleware layer, application layer, and business layer [1]. IoT is an emerging technology, and the leading technology companies have understood its value for the future. Many investors are investing tremendous volumes in that sector.

* Ochchhav Patel
E-mail address: ochchhavpatel@gmail.com

IoT has gained importance in various fields such as analysis and detection, smart living, the medical sector, agriculture, the automobile sector, and the surveillance field [2].
The industry's most important segment is healthcare. In addition to routine medical examinations, patients' bodily states, including heart rate, blood sugar, diabetes, electrocardiogram, and other important biomedical signals, can be tracked using a variety of medical monitoring devices and sensors for diagnosis or health quality improvement. Data security and privacy are more important than ever in the healthcare sector and around the globe. Violations of security and privacy can cost you a lot of money, in addition to harming the reputation of your organization and endangering patient relationships. Patient health information is required to be preserved by healthcare facilities like hospitals, clinics, and for-profit healthcare companies. Although the vast majority of healthcare organizations make sure that private information is encrypted and stored safely, no one has total control over security [3]. A single error leaves your data open to third parties. The provision of more effective and efficient treatment is aided by the use of the Internet of Things for real-time patient tracking. All IoT devices and networks need to be connected to other technologies in order to help healthcare facilities transform themselves in a meaningful way. In the health industry, appropriate and effective technologies help to address issues with product integrity and traceability in the drug supply chain, increase the overall security of patients' electronic medical information, and enable effective interoperability. IoT-based healthcare currently faces a myriad of issues, including data security, service quality, data transparency, mobility, patient identity, and ongoing monitoring [4]. A blockchain is some kind of data structure that can provide immutable transactions across the world and can help address the above issues [5]. Basically, it is working on hashing techniques, which means if you change the content of your block, the hash will be drastically changed according to the changes in the block. This effect is known as the "avalanche effect" [6]. The real beauty of this hashing process is that you can get the hash of your data, but you will not be able to retrieve data from that hash. It removes the centralized authority for the transactions. There is a shared ledger using peer-to-peer technology that is available for every node in the chain. Each block in the blockchain is made up of two parts: a block header and a data block. The blocks have to be mined for the accumulation of blocks in the real test network. This mining process uses various consensus algorithms like PoW (Proof of Work), PoS (Proof of Stack), PoI (Proof of Importance), and many more consensus mechanisms. In a blockchain network, for the mining process, miners have to solve some cryptographic problems [7].

There are three types of blockchain networks: the first is the public blockchain, and the remaining are private and consortium blockchain networks. There are no restrictions on public blockchains, and anyone is free to join and contribute to the core activities of the blockchain network. We can say a public blockchain is a democracy where all the computers and organizations decide together where they are going and what changes to implement. The most popular and widely used public blockchains are bitcoin and Ethereum. In a private blockchain, only an authorized person can participate in the chain network [8]. Also, in private, there is generally a dictatorship because everything is owned by one company. Examples of private blockchains are the Corda, Liquid, and Hyperledger Fabric networks. The consortium blockchain is a subset of a private blockchain in which only a small number of authorized members can participate in the chain network. It incorporates features from both public and private blockchain networks. Power does not reside with a single authority in a consortium network. It is operated under the direction of a group. So, a consortium blockchain is a private network for a group of companies or entities. "Quorum" and "r3" are blockchain consortium networks [8]. Blockchain is currently transforming many industries, including retail, healthcare, agriculture, automobiles, and medicine. Some exciting applications of blockchain are in cryptocurrency, the IoT sector, swarm intelligence, the automobile field, decentralized finance, and the medical field. IoT-based devices are transmitting a tremendous amount of data for various purposes over the network. The security of this transmission of data is very important because that data may contain some personal and important information. It means transmitted data should be available only to recognized services or people. It needs proper encryption so that unrecognized entities cannot access important data [9]. Data integrity also plays an important role in the IoT. It refers to the genuineness of data. The protection of each layer of the IoT architecture is mandatory. Otherwise, it can lead the whole network to immense problems. Lack of protection gives access to unauthorized parties, and they can make attacks on the network like spoofing, DoS attacks, session hijacking, botnets, and many more. For IoT devices, manufacturers should give device updates at certain times to avoid these kinds of security issues. In the current world, people are using so many IoT gadgets to operate their resources centrally, but this mechanism should be secure; otherwise, attackers might take control of resources or hijack your resources also. Identity protection, data integration, and other issues plague the IoT [10]. The properties of blockchain's decentralized nature and its tamper-proof security can solve many problems in the IoT. One of the most significant uses of blockchain technology in the medical sector is the management of supply chains for pharmaceuticals. Organization is crucial in every industry, but due to the escalating difficulties, it is crucial in healthcare in particular [11]. The research team [12] examined 39 publications published between 2018 and 2020 that utilized blockchain in healthcare practices. In this study, it was noted that interoperability, access control, and data integration might all benefit from employing blockchain technology. The authors of this study concluded that because the usage of blockchain technology in healthcare systems is growing quickly, further study in this area can be extremely important and beneficial.

This paper is organized as follows: Section 1 introduces a primer view of the healthcare domain, an overview of IoT and blockchain technology, an overview of data security issues, and how to address security issues using blockchain. The work that has already been done using supporting technologies for the medical healthcare system, like IoT and blockchain, is covered in Section 2. The research proposal, workflow, and IoT-blockchain-based medical healthcare system are all described in Section 3.

In Section 4, multiple blockchain networks are used to discuss the latency of the blockchain network. The implementation of our suggested work with mathematical derivation is covered in Section 5. In Section 6, we outline the results and numerous security risks that we have implemented. Section 7 describes the conclusion of our examined work. Section 8, which is the final section, is a future direction for an innovative researcher.

2. Related Work

In recent years, numerous academics have put forth various designs and techniques to offer authentication for accessing IoT data. We'll look at a couple of them in this part of the study. Researchers in [13] implemented a blockchain-based privacy-preserving architecture called healthchain that protects e-health data. To create this healthchain architecture, the blockchain has been built on Hyperledger Fabric, a permissioned distributed ledger solution that uses Hyperledger Composer and stores EHRs using IPFS. Researchers [14] presented the Bell-LaPadula model and classified the peers and transactions into various clearance and security levels to address the issue of scalability. Blockchain-based secure management and analysis of healthcare data are provided by Dwivedi et al. and his team [15]. To prevent various attacks, including denial of service and modification attempts, they have added security and privacy aspects to the proposed paradigm. Applying smart contracts, an advancement of blockchain technology, to the healthcare industry might lead to improved social and personal health data security [16]. MedChain is an efficient data-sharing system, Shen2019medchain and his colleagues [17] developed it. It combines blockchain, digest chain, and structured P2P network techniques to address inefficiencies in the current methods for exchanging healthcare data. Using a session-based healthcare data-sharing system based on MedChain, data sharing is flexible. According to the evaluation's findings, MedChain can increase productivity while meeting security standards for data exchange. A significant infrastructure is needed for the Internet of Medical Things (IoMT) to store and process a massive amount of medical data. IoMT systems with a blockchain data structure are said to give better security and privacy than a central data storage repository because the need for IoMT applications and platforms on a centralized cloud is incompatible with security [18]. In contrast to the hash codes on the blockchain, a significant quantity of data is saved on a decentralized platform in the IoMT system.

The Oracle [19] has been described by researchers as a smart device that can analyze the data presented and send alarms to both the patient and the healthcare professional. It interfaces directly with smart contracts. One of the modules focuses on the method of collecting and analyzing data from wearable devices and biosensors, either those that patients are wearing themselves or those that are present in the surroundings where patients are being monitored. The authors of the suggested work have utilized two blockchain networks, especially the personal healthcare blockchain and the external record management blockchain. Researchers proposed BlocHIE [20], which is a blockchain-based platform for healthcare information exchange. They examined the various criteria to retain and communicate two different types of healthcare data, namely electronic medical records and personal healthcare data. Researchers built BlocHIE on two loosely coupled blockchains, namely EMR-Chain for electronic medical records and PHD-Chain for personal healthcare data, as a result of their analysis. To properly protect privacy and authentic ability, they combined off-chain storage and on-chain verification procedures within the EMR chain. Researchers also suggested two transaction packing methods to boost system speed and guarantee customer fairness. Dwivedi and his team [21] introduced a novel hybrid approach that combines the advantages of the private key, public key, blockchain, and many other lightweight cryptographic primitives to develop a patient-centric access control for electronic medical records that is capable of providing security and privacy. They have also raised open questions about how to reduce various attacks such as DoS, modification attacks, etc. A unique decentralized approach that guarantees reliable device identification and verification was proposed by researchers and given the name "bubbles of trust" [22]. The availability and integrity of the data are also protected. Their method depends on the security benefits that blockchains offer in order to accomplish this goal and works to establish safe virtual areas (bubbles) where objects may recognize and trust one another. The C++ programming language and the Ethereum blockchain have been used to give a practical implementation of their system. Researchers [23] have developed a secure electronic health record (EHR) system that uses attribute-based cryptosystems and blockchain technology to achieve medical data's secrecy, authentication, integrity, and support for fine-grained access control. They have used attribute-based encryption (ABE) and identity-based encryption (IBE) to encrypt medical data, as well as identity-based signatures (IBS) for digital signatures. The suggested blockchain-based architecture attempts to resolve known security flaws in existing smart healthcare systems and increase the reliability of healthcare management systems. The SmartMedChain architecture, an end-to-end blockchain-based and privacy-preserving solution, was developed by researchers [24] for data sharing in the s-healthcare environment. Using the InterPlanetary File System (IPFS), a distributed data storage system with remarkable scalability and durability, encrypted health data has been saved using Hyperledger Fabric. Researchers in [25] presented a blockchain prototype network based on Hyperledger Fabric to decentralize the network, enhance the management of shared information, and optimize the information flow between two medical institutions even if one of the users does not trust the others. The Linux Foundation's Hyperledger infrastructure, which is reliable and open-source, makes it easier to create decentralized applications based on blockchain technology. Researchers [26] suggested a revolutionary decentralized IoT system based on blockchain for the healthcare industry. Hospitals, patients, and other members of the healthcare system can communicate with the blockchain network using the suggested application. It manages medical data or certifications using hash methods in the form of blocks in the blockchain network and restricts unauthorized or unlawful access. Therefore, evade fraudulent activity in the healthcare system.

3. System Architecture

Since IoT devices are adaptable in terms of deployment and management as well as in terms of interacting with other networks or devices, they may be utilized in a wide range of applications, from smart cities and homes to healthcare, agriculture, and education. In this study, we focus on the healthcare industry, where a variety of detectors and sensors are available to gauge important physical parameters like body temperature, heart rate, oxygen saturation, etc. From the perspective of the IoT's functionality, many layers are significant. Data ordering and transmission to the network layer are tasks of the IoT's perception layer. It also makes it possible for devices to cooperate with each other. In order to transfer the piled-up data from the perception layer to the storage servers via gateways, the network layer is responsible for managing communication. The gathered data is managed by the application layer, and the community of apps or end users receives the processed data. The AES-CBC cryptographic algorithm is used to encrypt the data file in this instance, and the encrypted file is saved on IPFS storage. Because of the blockchain-based mechanism, only authorized individuals can access the patient's data file from the data storage.

3.1 Proposed Model

IoT data is produced by numerous IoT medical sensor types and wearable technology in our proposed architecture. The data is then filtered and inflated as necessary before being placed on IPFS storage. A blockchain network allows for the access of data by authorized users. An application for the healthcare industry aligns with the layered architecture of the IoT. The first level consists of sensors or medical devices that serve as a data collection unit for information like pulse rate, oximeter readings, and body temperature. The services that gather data from the first layer and deliver it to the following layer are included in the second level, which also includes communication. Results acceptance and data processing are done at the third layer. Medical researchers retain patient data for both clinical and research goals. Users are typically divided into two categories: primary users, such as physicians, nurses, and close relatives, and secondary users, such as health insurance providers, researchers, and drug developers. IoT inherits the current underlying weaknesses because it operates on top of the conventional Internet. In order to address problems like confidentiality, integrity, and authentication for resource-constrained IoT, cryptographic steps must be employed.

3.2. Preliminaries, Requirements, and Proposed Work

3.2.1. Raspberry Pi

A cheap, compact, and portable computing board is the Raspberry Pi [27]. It can be plugged into a computer monitor, keyboard, mouse, flash drive, etc. The Raspberry Pi comes with software like Scratch that lets users create entertaining videos, games, and animations. Python, the primary core language of the Raspbian operating system, can also be used by programmers to create scripts or programs. The Model B has evolved into the Raspberry Pi B+. The client/server communication script in this work was written in the Python programming language. There are also enhancements like more USB ports, increased GPIO header pins, decreased power consumption, etc.

3.2.2. RFID Reader-Writer and Tags with MCP3008

Along with the Internet and mobile technology, which are bringing the globe closer together, RFID is a crucial component of the technological revolution. All RFID systems include three fundamental parts. The RFID tag that is affixed to a possession or object is the first. The tag may have sensors in addition to information on the asset or item. The RFID interrogator, which communicates with the RFID tags, is the second element. The backend system, which connects the RFID interrogators to a centralized database, is the third element. The three types of RFID technology are passive RFID, active RFID, and semi-passive RFID. Passive RFID technologies are often divided into low-frequency (LF), high-frequency (HF), ultra-high-frequency (UHF), and microwave categories depending on the radio frequency employed. Since RFID technology has been widely embraced, it is currently used in a variety of applications. RFID antennas, readers, scanners, and printers are a few RFID uses. RFID, or radio frequency identification, refers to a system that wirelessly broadcasts an object's or person's identity. radio waves that are encoded with a special serial number [28]. Tags, transponders, tag readers, antennae, and interfaces are some examples of the components that can make up an RFID system. Individual objects are outfitted with a small, cheap tag in a conventional RFID system. A transponder with a digital memory chip and an exclusive electronic product code is found inside the tag. The interrogator, which consists of an antenna, transceiver, and decoder, activates the RFID tag with a signal so it may read and write data to it. The activation signal from the reader is detected by an RFID tag as it moves through the electromagnetic field. Data that has been encoded in the integrated circuit of the tag is decoded by the reader and sent to the host computer. To reduce the multiple, frequently redundant reads of the same tag to a smaller and more usable data set, the application software on the host analyzes the data and may carry out a variety of filtering procedures. The 8-channel, 10-bit MCP3008 is an inexpensive analog-to-digital converter. This chip is an excellent choice if you only need to read straightforward analog signals, such as those from a temperature or light sensor. A SPI serial connection is used to link the MCP3008 to the Raspberry Pi. To communicate with the MCP3008, you can use either the hardware SPI bus or any four GPIO pins and software SPI. Software SPI is slightly more flexible than hardware SPI since it can utilize any of the Raspberry Pi's pins, whereas hardware SPI is slightly quicker but less flexible because it can only use certain pins [29].

3.2.3. Blockchain

Blockchain technology enables the transmission and storage of transactions. It keeps the information in a block-based ledger. A chain of blocks is created by connecting each block to the one before it. A peer-to-peer network ensures data transfer. As a result, blockchain is a safe and decentralized distributed ledger. Blockchain has attracted a lot of attention in the banking and financial industries over the past several years. These days, it is finding use in additional fields like insurance, energy, industry, and healthcare. Due to its characteristics, namely that it is decentralized, distributed, and secure, blockchain gains widespread adoption in practically all industries. Since the network is decentralized, a centralized authority is not required to run it. To achieve consensus among nodes, data is archived using this approach. Each node in the network contributes to the distribution and upkeep of the ledger. Numerous factors, including cryptography, the consensus mechanism, immutability, traceability, and data replication, contribute to the security of the blockchain [9]. The public-key cryptography used by blockchains. This type of cryptosystem includes a number of security features, including identity, encryption, decryption, and digital signature. Blockchain generates a set of public and private keys using an asymmetric cryptography method. The accounts of users are identified and authenticated using these keys. The blockchain can't be altered because every block is connected by a cryptographic hash. The blockchain's consensus technique, which is utilized to obtain consensus, relies on processing power to find the block hash by employing a complex mathematical problem. Because the data is distributed, it replicates across all nodes. Blockchain, therefore, lacks a single point of failure. Because blockchain keeps a complete and timestamped history of every transaction, it enables traceability.

3.2.4. Ethereum

Ethereum [30] is an open-source, decentralized, distributed computer blockchain platform. In 2014, Vitalik Buterin developed it as a result of being inspired by the Bitcoin cryptocurrency. The Elliptic Curve Digital Signature Algorithm is used in Ethereum, much like it is in Bitcoin. The discrete logarithm issue is the foundation of elliptic curve cryptography, which generates a pair of keys. Secp256k1 is the elliptic curve employed by Ethereum.

3.2.5. IPFS

A distributed peer-to-peer file system is IPFS (InterPlanetary File System). The innovation with IPFS is that content-based addressing has taken the place of location-based addressing. In other words, we need the hash of the data rather than the address where it is stored in order to search for it. A distinct hash is created for each file that is delivered to the IPFS for storage [23]. So, all you have to do to find this file is look up its hash. IoT data is generated in an authentic environment and stored in IPFS storage so that authorized users can access it via the blockchain network. We have experimented with data file encryption using the AES-CBC cryptographic technique. The security of health data in terms of confidentiality, data integrity, and access control mechanisms is the main focus of this study. We have covered reply attacks, masquerade attacks, modification attacks, message tampering, eavesdropping, and brute force assaults in our research. As we know, lack of trust, inadequate connectivity, weak security, and scalability plague the Internet of Things. The IoT also has other problems, such as device heterogeneity, energy limits, node addressing, and node identification. Temperature, oxygen saturation, pulse rate, and other data created by the data generator are among those that are kept in a file.

Table 1. Types of Attacks

CIA Triad	Types of Attacks
Confidentiality	1. Snooping 2. Traffic Analysis 3. Eavesdropping 4. Brute Force
Integrity	1. Modification 2. Masquerading 3. Replying 4. Repudiation 5. Message Tempering
Availability	1. Denial of Service 2. Session Hijacking 3. Jamming

The data is saved using a particular file naming scheme provided by the RFID chip. The patient data file will be produced using the JSON file format. In order to steal information and amass it for financial advantage, outside hackers gain access to patient and medical systems. They could, for instance, make false claims to health insurance using patient personal information. Hackers that demand a ransom from healthcare businesses in exchange for recovering patient data systems constitute another sort of external theft. Curiosity is another problem with healthcare data security. The remaining instances of insider abuse are brought on by unintentional behaviours like human mistakes, such as putting false information into a healthcare database or clicking on a phishing email. Numerous assault kinds are shown in Table 1 in this article.

Algorithm 1: Data Encryption

1: Function Encryption: Input- Data File D_P, Output- Encrypted file D'_P

2: Select the encryption algorithm (E.g., AES-CBC)

3: Collect a symmetric key K and Generate a random initial vector IV
4: $D'_P \leftarrow E_K (D_P, IV)$
5: Return the encrypted file (D'_P) on IPFS storage and K' on blockchain

Utilizing the symmetric key (K) and initial vector IV in the encryption procedure, we encrypt the data file (D_P). According to algorithm 1, after using a symmetric key and an initial vector (IV), ciphertext (D'_P) will be generated. After data file has been encrypted, the symmetric key (K) is decrypted using the recipient's public key (PU_R) and saved on the blockchain platform for further use. On IPFS storage, the encrypted file (D'_P) will be kept. A hash value serving as an acknowledgement of the supplied file will be received from IPFS. As seen in algorithm 2, the receiver's private key (PR_R) has been used to decrypt the symmetric key (K') for the original data file (D_P). Since IPFS storage is now more widely accessible, data confidentiality is necessary. This can be a breach of specific data storage policies that reveals sensitive data. The owner of that content may request authenticity confirmation from IPFS. It appears that IPFS provides quick and reliable, fault-tolerant file storage for content. As IPFS develops, it might make use of a privacy layer to conceal personally identifiable information that is also encrypted at rest, ensuring that no sensitive information is revealed. The data file (D_P), is encrypted with an AES symmetric key (K). Here, the symmetric key (K) is encrypted using the receiver's public key (PU_R) of the RSA method to produce the encrypted key (K').

Algorithm 2: Receiver-side data decryption

1: Input: (Encrypted File (D'_P), K', H_D) :: Symmetric key encryption using receiver's public key (PU_R)
2: Output: Decrypted Data File (D_P)
3: Function Decryption ((D'_P), K', E_K)
4: (K) ← Decryption (E_K ,PR_R) :: Key Decryption using Receiver's Private Key (PR_R)
5: $D_P \leftarrow D_K (D'_P, K)$
6: End Function

Algorithm 3: Uploading documents to the Interplanetary File System and storing hashes on Blockchain

1. **Function** Upload (D'_P): Encrypted Patient's Data File
2. **for** each Datafile
3. do
4. Retrieve (D_P) Patient_Datafile and encrypt it using Symmetric key K (AES)
5. $K' = E_K (K, PU_R)$
6. Send (D'_P) Patient_Datafile to IPFS
7. IPFS stores (D'_P) Datafile and assigns a hash value to it.
8. A hash value is returned from IPFS as an acknowledgement of a stored file (D'_P)
9. IPFS hash value is sent to Blockchain
10. Encrypted Key K' is sent to Blockchain
11. **Function** Storage ($IPFS_{Hash}$, K'):
12. **for** each $IPFS_{Hash}$ and K' **do**
13. Web3 is invoked to execute smart contracts.
14. Metamask displays request to approve transaction payment
15. Payment is approved: $IPFS_{Hash}$ and K' are stored on Ethereum

The source node requested to store its processed data file on IPFS. For data file encryption, a receiver's public key (PU_R) is used. After the use of that bundle, a data file is received on the IPFS platform, and its hash value is computed. The calculated hash value (H_D) will be sent to the source node as the ID of that data file. In IPFS, a copy of a requested file is cached on the requester's node. There will be more cached copies of the datafile as more users seek it. Any node or group of nodes that has the file on it can respond to subsequent requests for it. Numerous nodes increasingly split out the responsibility for providing the requested data and completing the request. It necessitates a new kind of web address. The decentralized web employs content-based routing technologies like IPFS instead of address-based routing, which requires you to know the exact location of the data and supply a precise URL to that data [13]. You should encrypt a file before importing it if you need to keep the content confined to a small group of people but yet want it to be available to others. However, imported files are voluntarily not encrypted by default, despite the fact that data transit is encrypted in both directions. In our implementation, when an encrypted data file (D'_P) is delivered to the IPFS platform, we use the AES-CBC (symmetric) algorithm to encrypt it. After the successful execution of the smart contract with the involvement of Web3, the hash value (H_D) and encrypted key (K') will be stored on Ethereum, which is illustrated in Algorithm 3.

4. Latency of the Blockchain Network

Blockchain technology offers new digital capabilities for authentication and permission that eliminate the need for some sort of centralized administration. As a result, it allows for the formation of new digital relationships. The immutable smart contract protocol or rule used by blockchain technology implies that once it is implemented, it cannot be changed [9]. As it is exceedingly difficult to replicate a creation-like environment for performance testing, performance assessment in blockchain is difficult. The technical authentication protocol has to be evaluated for network latency depending on block size, network type, estimated transaction size, and how long it takes a query to return results. We have tested the transactions on various blockchain networks and determined how much gas is required to deploy smart contracts on the blockchain network. We also computed the amount (in USD) of gas required to upload the AES encrypted key (K’) and hash (H_D) of encrypted data files to the blockchain network. For transactions, we have utilized the blockchain networks of Rinkeby, BSC, and Matic. We come to a conclusion and note that deploying smart contracts on the MATIC network requires less gas. On Ethereum, the Kovan and Rinkeby networks consumed more gas than the Binance and Matic networks in assessments of gas requirements. If you pay a lower cost, your transaction may likely go much more slowly, depending on the network and other parameters. The comparison of various blockchain networks for the amount of gas used during the deployment of smart contracts on the blockchain network is shown in Fig. 1. The actual cost of deploying a smart contract on the blockchain network is shown in Fig. 2, which also shows the conversion of the gas value into USD ($). (Date: November 17, 2022, at 1:46 a.m.)

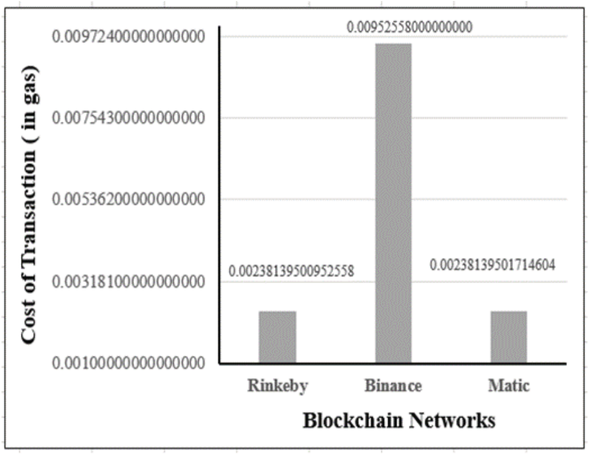


Fig. 1. Required gas value for smart contract deployment on a blockchain network

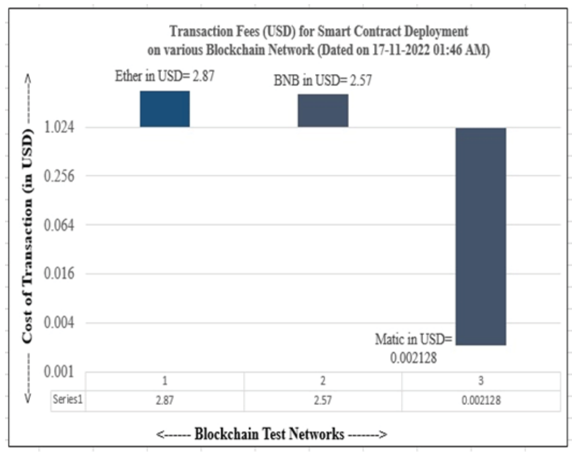


Fig. 2. Required USD ($) for smart contract deployment on a blockchain network (conversion of gas to USD)

5. Implementation

The two phases of our experimentation are as follows: In the initial stage of IoT deployment, Raspberry Pi, tags, and medical sensors have been used. The second phase involved developing the blockchain side using the Ethereum platform, Solidity as the programming language, Metamask (a blockchain wallet), and IPFS for distributed storage. We successfully integrated two technologies in this execution, outlining how medical data is produced and safeguarded in the context of the Internet of Things. It also included instructions on how to store that data on IPFS storage and how the blockchain network may use it. The RFID gadget created a file with the patient's individual ID. The output data is subsequently filtered to fulfill requirements, and the file is then cryptographically encrypted. The Raspberry Pi can interface with a variety of sensors and RIFD tags using its built-in Python library. The Raspbian operating system is installed using the Berryboot operating system, which may be used to install any Raspbian operating system and act as an all-purpose operating system. The Raspberry Pi is integrated with the RFID RC522 chip using the MFRC522.py Python package, which is used to read and write data to and from the RFID tags. The hash value of a specific data file for the patient is generated by IPFS storage and sent back to the IoT platform. First, we use the Python.objcrypt module to encrypt that data file using the AES-CBC encryption algorithm on the IoT platform. We upload the encrypted file to IPFS storage. A fixed-length hash value (H_D) of the stored file is returned by IPFS storage. For later use, the retrieved hash value (H_D) is stored on the blockchain network.

5.1 File encryption and decryption using a mathematical derivation

Here, a symmetric key (K) is used to encrypt the patient's data file (D_P). When the encrypted data file (D'_P) is transmitted to IPFS storage, IPFS responds with the hash value (H_D) for the stored file as an acknowledgement. The hash value (H_D) of the protected file (D'_P) is one of the values saved on the blockchain, and the other value is (K’). The receiver's public key (PR_R) is used to encrypt the symmetric key (K), and that encrypted symmetric key (K’) stored on the blockchain network. Anyone wishing to view a data file (D_P) must visit the blockchain platform and utilize the H_D and K'. Due to the file encryption approach, the receiver can access the data file in protected mode by using the hash value (H_D) of the data file. A symmetric key (K) that the receiver can use to decrypt the file is simultaneously used to encrypt it with the receiver's public key. The protected file (D'_P) is decrypted and transformed back into the original data file (D_P) once the key (K’) is decrypted using the receiver's private key (PR_R).

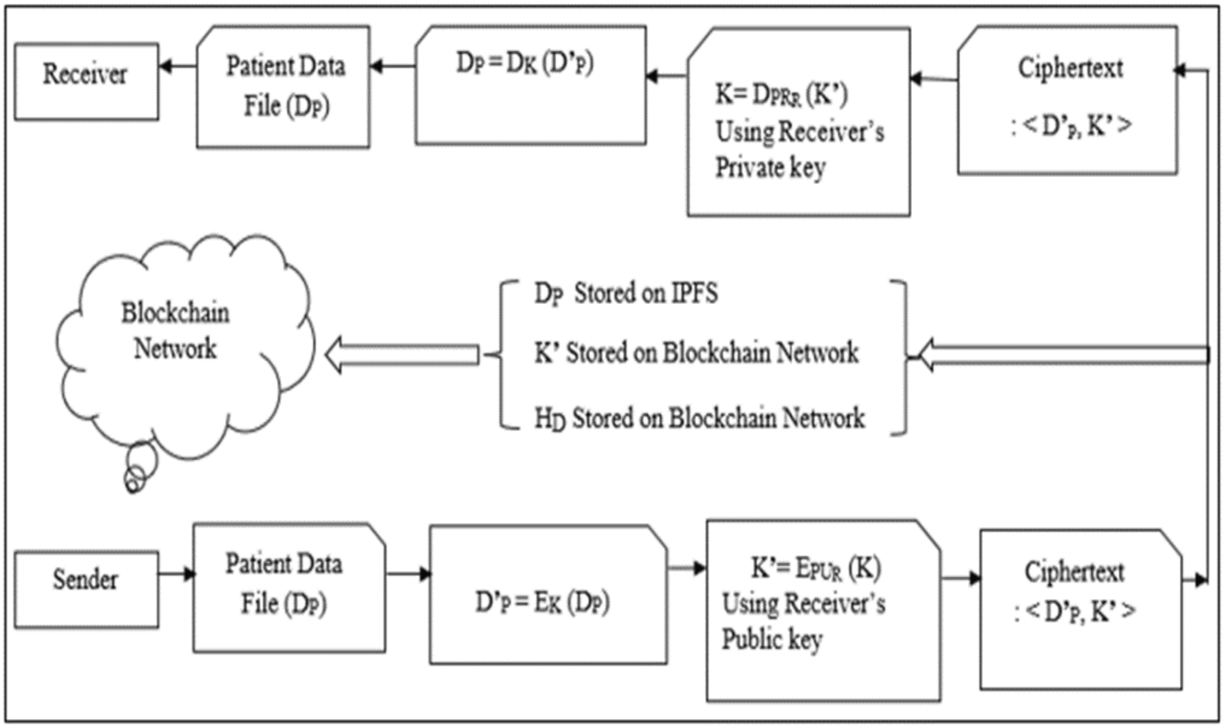


Fig. 3. Cryptographic process of proposed work

The following steps show how the overall transmission and reception of data files (D_P) from the data origin to the data consumer take place.

Step 1: The patient’s data file (D_P) is encrypted using the symmetric key K, and that file is converted into a protected data file (D'_P).
Step 2: The encryption process E_K is performed using the recipient's public key (PR_R), which is used to transform the symmetric key (K) into the encrypted form (K’).
Step 3: A protected data file (D'_P) is sent to IPFS file storage.
Step 4: IPFS sends the fixed-value hash (H_D) as an acknowledgement to the source platform.
Step 5: Retrieved hash value is stored into a blockchain platform
Step 6: Another entity, K’, is also stored on the blockchain platform.
Step 7: The user can send the request to IPFS via blockchain for the data file (D_P).
Step 8: IPFS sends the reply (data file D'_P) based on the matched hash (H_D).
Step 9: The decryption process D_K is performed using the receiver’s private key PR_R, and that encrypted key (K') is converted into its original form K.
Step 10: Using a symmetric key (K), a protected data file (D'_P) is converted to an original data file (D_P).

5.2 Implementing a Model

In order to establish data confidentiality, integrity, and access control measures, we employed smart contracts, which are implemented in the Ethereum Solidity programming language for IoT-based healthcare systems. All of the medical sensors have been fastened to the patient’s body, and the data they produce is saved in a special file that the RFID device generates. Fig. 3 illustrates the cryptographic procedure where data files are encrypted using the AES-CBC cryptographic method and stored on an IPFS platform. Due to the blockchain-based technology, only authorized users are able to access the patient’s data file, preventing network attacks.

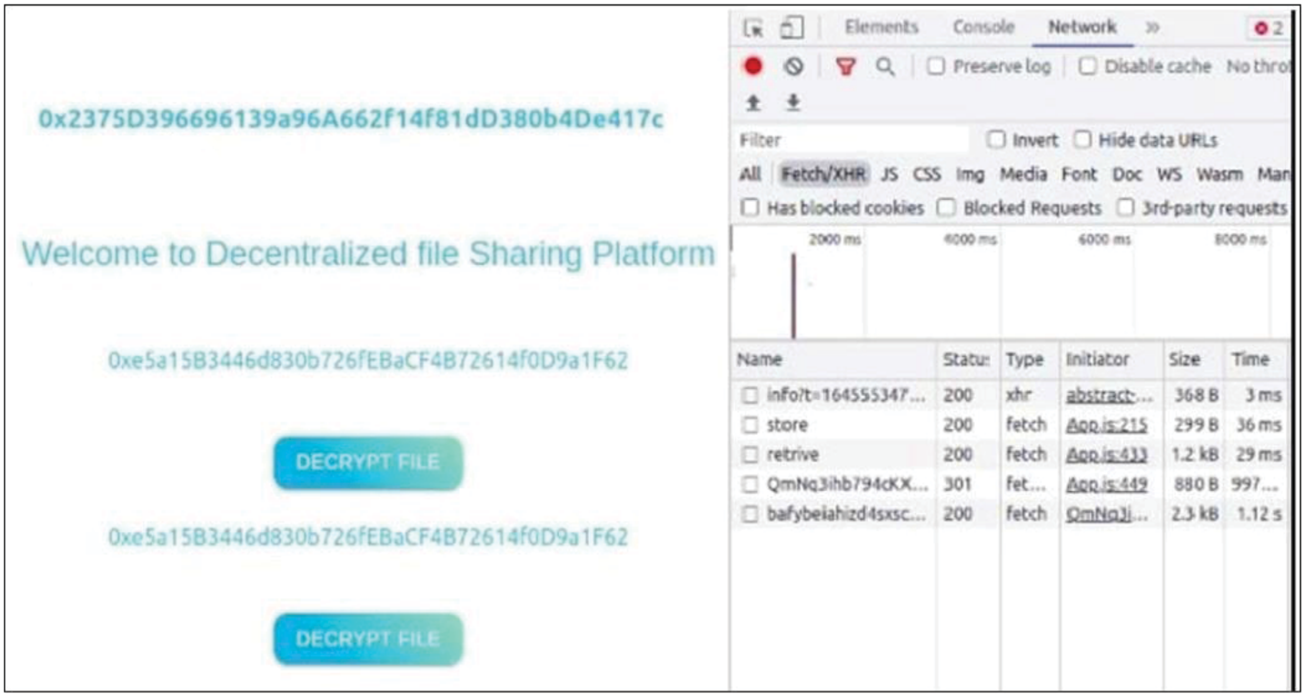


Fig. 4. Platform that shows the received data file with the sender's address

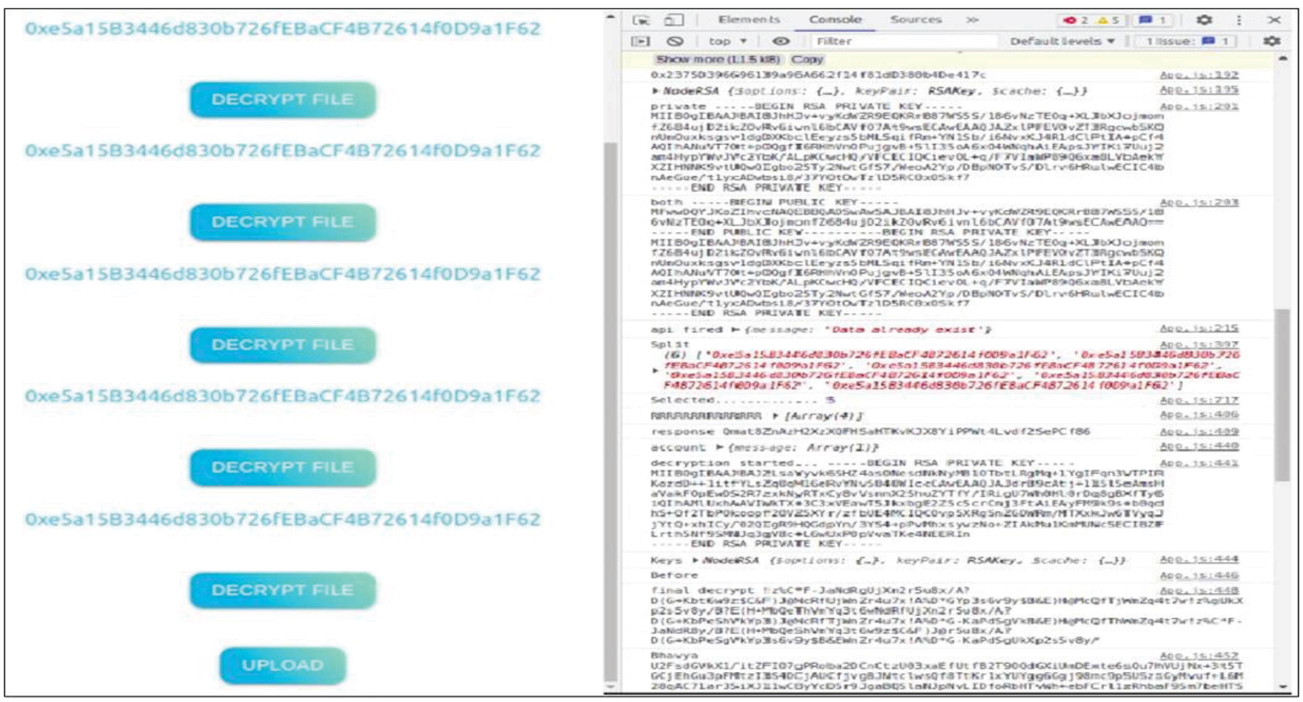


Fig. 5. Data file decryption using receiver's private key (PR_R) and symmetric key (K)

5.3 File Sending and Receiving via Blockchain Network

The data file (DP) is secured with a cryptographic technique (AES) and kept on the IPFS platform. The receiver's public key (PU_R) encrypts the symmetric key (K). IPFS returns the hash value (H_D) for the corresponding file that was stored. The blockchain platform stores the hash value (H_D) and encrypted symmetric key (K') that have been obtained. As shown in Fig. 4, the recipient must visit the blockchain platform, and use their private key (PR_R), and the hash value (H_D) of the saved data file to decrypt it in order to obtain the original data file (D_P) from IPFS. As everyone knows, an RSA (Rivest-Shamir-Adleman) key pair consists of a private key and a public key. Digital signatures are created using the RSA private key, but the RSA public key is used to verify them. The RSA public key can also be used to encrypt DES or AES data keys, and the RSA private key can be used to recover the keys. Platform used by the recipient, from which the recipient can download and upload data files to the IPFS platform. In our implementation, the symmetric key (K) used to encrypt the patient's data file (D_P) using AES is encrypted using the sender's public key (PU_R). The encrypted file (D'_P) is converted into the original file (D_P) using the user's private key (PR_R), as shown in Figure 5.

During testing, it has been discovered that our system supports the submission of files in the JSON and picture data file formats. Figure 5 illustrates how to encrypt the patient's data file (D_P) using a symmetric key (K), which is encrypted using the recipient's public key (PU_R). The receiving end uses a hash value (H_D) to obtain the data file from the IPFS. The retrieved data file (D'_P) is encrypted; as a result, the symmetric key is decrypted with the receiver's private key, and the obtained data file is then transformed back into the original data file (D_P) using that symmetric key (K).

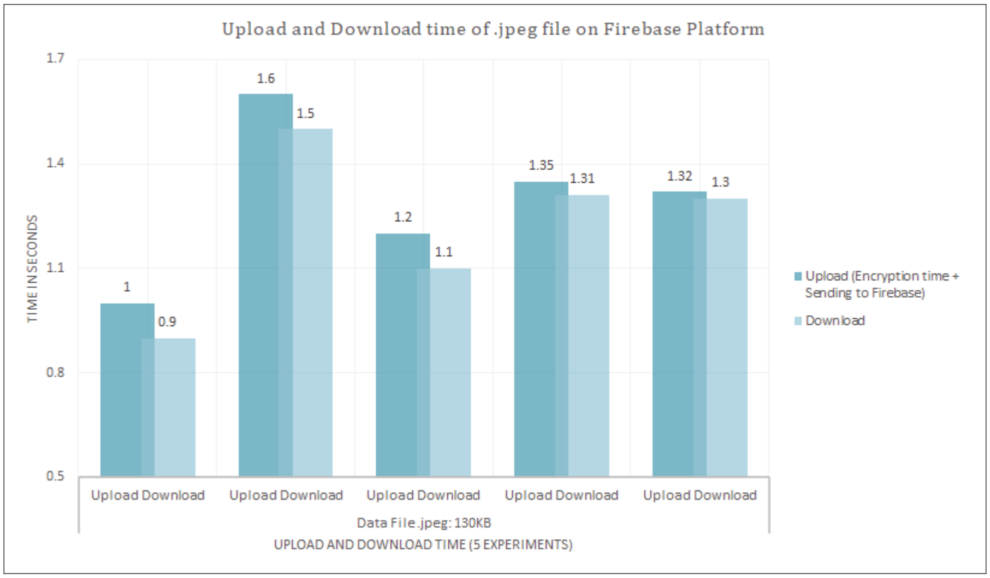


Fig. 6. Upload-Download interval for .jpeg file on firebase platform

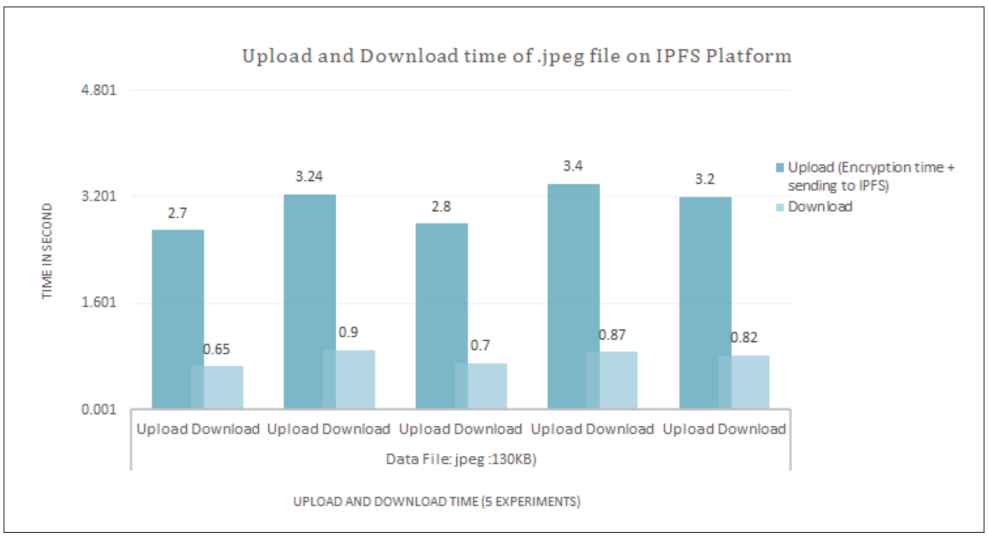


Fig.7. Upload-Download interval for .jpeg file on IPFS platform

6. Results and Discussion

We initially tested our implementation on Ganache, and then we simulated main net behavior using Kovan, Rinkeby, Binance, and the Matic network. There is a single instance of the Ganache platform that mimics the blockchain network. The interface for adding and viewing records is built using ReactJS, and the backend is built using JavaScript. The patients' cryptographic keys have also been kept in MongoDB. In order to replicate a blockchain network, we employed the Ethereum platform with Solidity as our test blockchain. The blockchain is additionally communicated with via Web3-JS. In order to test the IPFS network, we employed INFURA, which provides dependable, safe, and scalable access to the IPFS gateway. We examined the upload and download speeds of data files for our solution using Firebase and IPFS. The file (.jpeg, 130 kB) was utilized five times on both an IPFS distributed database and a Firebase centralized database, and the upload and download times were recorded. Figures 6 and 7 show the Firebase platform's results for file upload and download times as well as the IPFS platform's time. How rapidly data files may be recovered via IPFS depends on a variety of factors, including the number of systems storing the records, the location of the nearest system possessing the record, the number of systems storing the records, and other factors.

6.1 Analysis of the proposed framework in comparison to current blockchain technologies

Confidentiality, integrity, and availability are the three main security considerations that must be taken into account by any model developer. Encrypting data makes sure that only authorized users can access the system. Both availability and integrity guarantee that messages are sent to their intended recipients without being changed, and both ensure that users can always access the data they require. It is a comparison of the proposed work's confidentiality, data integrity, and access control features with those of the current blockchain approaches. The suggested architecture has been evaluated with current blockchain-based implementations, like [17], [20], [23], and [24]. In terms of data confidentiality, access control, data integrity, and scalability, it is evident that the proposed system addresses the shortcomings of the current systems.

The patient data in our system is stored on IPFS using cryptography techniques, so data security is a crucial responsibility. The actual, enormous data is saved after encryption on the IPFS storage, while this blockchain architecture just keeps a hash of the transaction on the blockchain network. This solution, which takes a patient-centric approach, ensures the protection of patient data and offers authorized access with patient authorization. Additionally, the blockchain solutions' smart contracts feature combines to support high-level encryption and guarantee patient anonymity in their medical records. To create strong blockchain data security solutions, the data saved on IPFS is also encrypted using a unique AES-CBS cryptographic method.

The patient data file (D_P) is encrypted using a cryptographic method before being saved to IPFS. Here, the AES-CBC encryption method is used to protect the file contents by encrypting the symmetric key (K) with the recipient's public key (PR_R). Eavesdropping attacks occur when a hacker intercepts, deletes, or modifies data that is being sent between two devices. Snooping, also known as eavesdropping, accesses data exchanged between workstations using open network connections. A hacking method known as "brute force attack" uses trial and error to crack encryption keys, passwords, and login information. It's a simple but effective way to gain unauthorized access to user accounts. Our research has explored the use of brute force attacks, snooping, and eavesdropping.

Data integrity refers to the permanent and unchangeable storage of data. It cannot be changed or eliminated. Each block in a blockchain holds the hash value of the preceding block in addition to the data, which is saved as hash values in each block. Despite relying on a third-party provider, the consensus method, digital signature, and built-in cryptographic algorithm form the foundation of this blockchain framework's confidence. Since all the blocks are connected, any modification to the original data will affect its hash value, and since altering the record is computationally challenging, the non-tampering of the patient data is also expressly guaranteed. The original data is saved in IPFS storage once a unique cryptographic procedure has been carried out. We also explored a modification and response attack, which may be prevented by being aware of how encryption works. When cryptography is used, the communication is usually encrypted. On the receiving end, the message is unlocked using the decryption technique. To guard against such attacks, the sender and recipient are obliged to select a session key at random. This session key represents a code type that will be valid for just one interaction between sender and recipient. It is impossible to set up this code again. Timestamps are another tool that can be used. A timestamp that has passed its expiration date cannot be cracked by hackers because timestamps have a lifespan. The suggested framework offers encrypted data storage in IPFS while maintaining the majority of security requirements and addressing the scalability issue with the current approaches. The scalability of the suggested approach has shown that it is capable of processing massive data sets quickly.

Table 2. Analysis of the proposed framework in comparison to current blockchain technologies

Scheme	Blockchain Platform	Confidentiality	Data Integrity	Scalability
Medchain [17]	Consortium Blockchain	Yes	Yes	No
Blochi [20]	Private Blockchain	No	Yes	No
Wang & Song [23]	Private Blockchain	Yes	Yes	No
SmartMedChain [24]	Private Hyperledger	Yes	Yes	No
Proposed Work	Public Ethereum	Yes	Yes	Yes

7. Conclusion

Numerous issues cause performance degradation in the IoT environment. A few of these issues include unattended, resource-constrained, heterogeneous devices, which sometimes result in security, privacy, and reliability concerns. During our research, it has been our priority to showcase and avail the usage of blockchain technology, specifically in the healthcare system, in diverse situations. The diverse can include a variety of steps such as data sharing, clinical research development, or patient health diagnosis. IoT sensors and cryptographic models are used to generate patients' data and secure the same data. IPFS storage can be used to store encrypted data. In this research work, an Ethereum-based blockchain framework has been implemented for securing data storage and providing efficient access control between stakeholders like patients, doctors, pharmacists, and other participants through encryption techniques and access control mechanisms. With the upsurge in health data every year, we look forward to enhancing this framework with rigorous simulations of scalability and comparing it with other blockchain configurations that will invite further attention in future research tasks.

8. Future Work

For the current prototype implementation, the IoT device Raspberry Pi has been used to generate IoT data and apply a cryptographic algorithm that has been used for various methodologies in data encryption on the IoT platform. Innovative researchers may use progressive IoT tools to support advanced encryption algorithms. There is a constant rise in advanced encryption algorithm usage, which will motivate researchers to use progressive IoT tools that are well-equipped for advanced encryption. IPFS storage, which is certainly needed for decentralized file storage for various types of peer-to-peer transmission of data, is used to store the encrypted data. In our research for a smart contract, the Ethereum-based Solidity programming language has been used. Various other platforms, like Hyperledger, can be utilized for penning agreements between entities in the system instead of the Ethereum platform.

Acknowledgement

Words are inadequate to express my gratitude to my research supervisor for his essential patience and input. In addition, I could not have made this voyage without the rich knowledge and experience my defense committee supplied. Additionally, without the kind assistance of my institute, LDRP-ITR, Gandhinagar, this project would not have been possible. My appreciation also extends to my friends and coworkers. My family should be mentioned, especially my parents, spouse, and children. Their confidence in me has sustained my enthusiasm and upbeat attitude throughout this process.

References

1. Sethi, P., & Sarangi, S. R. (2017). Internet of things: architectures, protocols, and applications. Journal of Electrical and Computer Engineering, 2017.
2. Dhanvijay, M. M., & Patil, S. C. (2019). Internet of Things: A survey of enabling technologies in healthcare and its applications. Computer Networks, 153, 113-131.
3. Wu, H., Dwivedi, A. D., & Srivastava, G. (2021). Security and privacy of patient information in medical systems based on blockchain technology. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 17(2s), 1-17.

4. Islam, S. R., Kwak, D., Kabir, M. H., Hossain, M., & Kwak, K. S. (2015). The internet of things for health care: a comprehensive survey. IEEE access, 3, 678-708.
5. Restuccia, F., Kanhere, S. D., Melodia, T., & Das, S. K. (2019). Blockchain for the internet of things: Present and future. arXiv preprint arXiv:1903.07448.
6. Ahmad, L., Khanji, S., Iqbal, F., & Kamoun, F. (2020, August). Blockchain-based chain of custody: towards real-time tamper-proof evidence management. In Proceedings of the 15th international conference on availability, reliability and security (pp. 1-8).
7. Cao, B., Zhang, Z., Feng, D., Zhang, S., Zhang, L., Peng, M., & Li, Y. (2020). Performance analysis and comparison of PoW, PoS and DAG based blockchains. Digital Communications and Networks, 6(4), 480-485.
8. Polge, J., Robert, J., & Le Traon, Y. (2021). Permissioned blockchain frameworks in the industry: A comparison. Ict Express, 7(2), 229-233.
9. Qashlan, A., Nanda, P., He, X., & Mohanty, M. (2021). Privacy-preserving mechanism in smart home using blockchain. IEEE Access, 9, 103651-103669.
10. Asghari, P., Rahmani, A. M., & Javadi, H. H. S. (2019). Internet of Things applications: A systematic review. Computer Networks, 148, 241-261.
11. Taloba, A. I., Elhadad, A., Rayan, A., Abd El-Aziz, R. M., Salem, M., Alzahrani, A. A., ... & Park, C. (2023). A blockchain-based hybrid platform for multimedia data processing in IoT-Healthcare. Alexandria Engineering Journal, 65, 263-274.
12. Hasselgren, A., Kralevska, K., Gligoroski, D., Pedersen, S. A., & Faxvaag, A. (2020). Blockchain in healthcare and health sciences—A scoping review. International Journal of Medical Informatics, 134, 104040.
13. Chenthara, S., Ahmed, K., Wang, H., Whittaker, F., & Chen, Z. (2020). Healthchain: A novel framework on privacy preservation of electronic health records using blockchain technology. Plos one, 15(12), e0243043.
14. H. Wu, A. D. Dwivedi, G. Srivastava, Security and privacy of patient information in medical systems based on blockchain technology, ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM) 17 (2s) (2021) 1–17.
15. Wu, H., Dwivedi, A. D., & Srivastava, G. (2021). Security and privacy of patient information in medical systems based on blockchain technology. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 17(2s), 1-17.
16. Griggs KN , Ossipova O , Kohlios CP , Baccarini AN , Howson EA , Hayajneh T . Healthcare blockchain system using smart contracts for secure automated re- mote patient monitoring. J Med Syst Jul. 2018;42(7):130.
17. Shen, B., Guo, J., & Yang, Y. (2019). MedChain: Efficient healthcare data sharing via blockchain. Applied sciences, 9(6), 1207.
18. Dilawar, N., Rizwan, M., Ahmad, F., & Akram, S. (2019). Blockchain: securing internet of medical things (IoMT). International Journal of Advanced Computer Science and Applications, 10(1).
19. Rogers, D. (2019). A visit to the Oracle: Reviewing the state of construction industry digitalisation. Construction Research and Innovation, 10(1), 11-14.
20. Jiang, S., Cao, J., Wu, H., Yang, Y., Ma, M., & He, J. (2018, June). Blochie: a blockchain-based platform for healthcare information exchange. In 2018 ieee international conference on smart computing (smartcomp) (pp. 49-56). IEEE.
21. Dwivedi, A. D., Srivastava, G., Dhar, S., & Singh, R. (2019). A decentralized privacy-preserving healthcare blockchain for IoT. Sensors, 19(2), 326.
22. Hammi, M. T., Hammi, B., Bellot, P., & Serhrouchni, A. (2018). Bubbles of Trust: A decentralized blockchain-based authentication system for IoT. Computers & Security, 78, 126-142.
23. H. Wang and Y. Song, "Secure cloud-based ehr system using attribute-based cryptosystem and blockchain," Journal of medical systems, vol. 42, no. 8, pp. 1–9, 2018.Wang, H., & Song, Y. (2018). Secure cloud-based EHR system using attribute-based cryptosystem and blockchain. Journal of medical systems, 42(8), 1-9.
24. El Majdoubi, D., El Bakkali, H., & Sadki, S. (2021). SmartMedChain: A Blockchain-Based Privacy-Preserving Smart Healthcare Framework. Journal of Healthcare Engineering, 2021.
25. Al-Sumaidaee, G., Alkhudary, R., Zilic, Z., & Swidan, A. (2023). Performance analysis of a private blockchain network built on Hyperledger Fabric for healthcare. Information Processing & Management, 60(2), 103160.
26. Hasselgren, A., Kralevska, K., Gligoroski, D., Pedersen, S. A., & Faxvaag, A. (2020). Blockchain in healthcare and health sciences—A scoping review. International Journal of Medical Informatics, 134, 104040.
27. Zhao, C. W., Jegatheesan, J., & Loon, S. C. (2015). Exploring iot application using raspberry pi. International Journal of Computer Networks and Applications, 2(1), 27-34.
28. Abad, E., Palacio, F., Nuin, M., De Zarate, A. G., Juarros, A., Gómez, J. M., & Marco, S. (2009). RFID smart tag for traceability and cold chain monitoring of foods: Demonstration in an intercontinental fresh fish logistic chain. Journal of food engineering, 93(4), 394-399.
29. Hassen, H. B., Ayari, N., & Hamdi, B. (2020). A home hospitalization system based on the Internet of things, Fog computing and cloud computing. Informatics in Medicine Unlocked, 20, 100368.
30. Ethereum - Wikipedia. (2015, July 30). Ethereum - Wikipedia. Retrieved November 12, 2022, from https://en.wikipedia.org/wiki/Ethereum

Impact of Social Media During Natural Calamity

Jagdish Panchal[a*]

[a]*Sankalchand Patel University (SPU)*

Abstract: This paper proposes a disaster management system based on machine learning, incorporating a web-based graphical user interface. The system involves six steps, starting with the inclusion of any disaster-related information into the message dataset. Two commonly used machine learning algorithms, random forest and logistic regression, are employed to evaluate the performance of the interface, achieving an accuracy of 72% and 73%, respectively, in classification and prediction tasks. The interface enables users to input messages containing relevant keywords classified based on predefined criteria. Afterward, the system automatically selects the relevant department for the particular disaster. The proposed system can enhance disaster management by simplifying the identification and response to disasters.

Keywords: social media, natural calamity, machine learning, disaster management, graphical user interface.

1. Introduction

Communication and information exchange among users of social media platforms are becoming more common. Facebook, Twitter, WhatsApp, and Instagram have become very popular in the recent decade [1]. Social media (SM) can swiftly broadcast information far and wide, allowing users to see and learn about events occurring far away in space and time. Many people know that some of the information on social media is not entirely true [2]. Indonesia has had several natural calamities recently, but flooding is the most popular subject on social media. Users are compelled to post photographs or update news from the region impacted by the Flood to SM to convey the current situation. Due to the interaction process, individuals utilize the information as a reference to determine their attitudes and decision-making for the benefit of the victims of catastrophes, who can track and monitor the catastrophic occurrence more easily [3].

Calamity management has played a critical role in reducing the number of people killed and the amount of property and infrastructure that has been damaged. Effective disaster management requires a complex system for collecting, integrating, managing, and analysing data from various sources, such as video streaming, sensors on the ground, and satellite images [4]. Social networks and crowdsourcing have allowed human-centric techniques that permit the public to offer critical information to improve crisis management and reduce natural calamity impacts [5].

Computer, geographic information science, and domain science researchers can learn much from social media data [6]. By 2020, 3.5 billion people will use social media, almost half of the world's population. Social media creates a broad range of data, including text, photos, videos, and vast amounts [7]. Natural disasters have a considerable influence on people's physical and mental health. This includes injuries received, exposure to weather dangers, deterioration of inadequate sanitation, and contamination of water supplies [8].

Social media can have a big effect on natural disasters, and the study aims to give basic tips for coordinating communications and sharing information among people in these situations.

* Corresponding author: Jagdish Panchal
Author Email: jnp1682@gmail.com

1.1 Social media environment of calamity

Digital media innovation transformed Public Service Obligations (PSO) communications globally [9]. Businesses are increasingly relying on Facebook and Twitter as part of their online strategy because of the ease with which authors can connect with customers and other stakeholders [10]. Instead of relying on a one-way public information approach, social media allows PSOs and people to converse more in-depth [11]. Social media has also made it easier to connect creatively with journalists [12]. Disintermediation is the goal of disseminating information that promotes ethical values like honesty, openness, and faith in institutions [13].

Several academics have talked about the possible benefits of social media, such as making the government more open and accountable and making it easier to use public services. Studies have examined how social media could make people more likely to vote, get them more involved in their communities, and bring democracy back to life in the modern world [14]. PSOs favor asymmetric and one-way distribution strategies despite the option of dialogically communicating with inhabitants [15]. The existence of official social media accounts administered by PSOs can be significant because it can counterbalance the information deluge generated during crises. It is possible to utilize social media ethically to prevent the spread of misleading information.

1.2 Natural Hazard and Impact of Natural Calamity

A hazard is a potentially dangerous incident that could endanger people and the places where they live. Table 1 displays the several types of natural hazards that can be identified, as provided by the Center for Research on the Epidemiology of Disasters (CRED). Many more geophysical hazards exist, including mass movements, tsunamis, earthquakes, and volcanic eruptions. Climatological hazards include thunderstorms, high temperatures, and fog [16]. Glacial lake outbursts, wildfires, and Drought are examples of climatological dangers. Floods, landslides, and wave activities are all examples of hydrological risks. As stated in CRED and summarized in Table 2, natural catastrophes and consequences or threats to bodily well-being: injury/death, affected/damage [17].

Table 1: Natural risks are defined and categorized [17].

Hazard	Classification	Definition	Calamity Types
Natural Hazard Types	Geophysical	Geological risk is a term that is frequently used interchangeably with danger that originates in the solid earth.	Mass movement, Earthquake, volcanic activity.
	Meteorological	Severe atmospheric and weather conditions that endure from a few minutes to a few days can produce a scary scenario.	Fog, Extreme temperature, storm
	Hydrological	Hazard is induced by the development, circulation, and circulation of fresh or salt water on the surface or subsurface.	Wave action, Landslide, Flood
	Climatological	Atmospheric climate change spans from seasonal to multi-decadal timescales.	Drought, animal accident, wildfire, Extreme temperature, glacial lake outburst
	Biological	The danger posed by contact with living creatures and their toxins or the illnesses they can spread. Examples include venomous animals' toxic plants, insects, and mosquitoes that transmit disease-causing agents such as parasites, viruses, or germs.	Insect infestation, Epidemic

Table 2: Impacts of Natural Calamities [17].

Term	Definition
Fatality	An estimated total of individuals died because of natural calamities.

Injured	People who have been directly affected by a disaster need medical attention due to physical injuries, trauma, or disease.
Affected	People who were wounded, homeless (those whose homes were destroyed or severely damaged and so need refuge during a disaster) and impact.
Damage	The level of damage done to buildings, farms, and cattle. According to the Emergency Disaster Database (EM-DAT), losses are assessed in US dollars ('000). An individual calamity's recorded figure correlates to the current value of the event's damage.

2. Literature of Review

The following study expands on the previous impact of social media during natural calamities. Several investigators explain their findings, as seen below.

Ramakrishnan et al. (2022) [18] analyzed the likelihood of social media usage by underprivileged populations by relying on knowledge of the digital divide and attribution theory. The data was gathered via a survey and analyzed using Partial Least Squares Structural Equation Modeling (PLSSEM). The study's findings show that underprivileged populations are less inclined to employ social media for disaster management. After the study, there was a lot of interest in using social media for disaster management. In addition, the essay offers both theoretical and practical consequences.

Dong et al. (2021) [19] evaluate disaster relief efficiency by evaluating social media data, such as public opinions on disaster reactions and demand for targeted assistance resources during various calamities. Comparing machine learning (ML) models based on accuracy and computational time is done to improve present decision-makers with the right model. People's use of SM to aid calamity relief in the face of similar natural calamities as Twitter continues to develop is examined. Investigators could use the findings in their work to better understand how to handle natural disasters, and the authors could use the information to help disaster relief organizations.

Poornima and Murugan (2021) [20] suggested a strategy called the Natural Disaster Resilience Approach (NDRA). NDRA uses the Advogato dataset, which has 51,127 edges and 6541 users. Ultimately, the comparison was performed exclusively among the suggested Advanced Sybil Node Prediction Method (ASYNPA) Tier-3 and the current Vote Trust system, and the graph is mapped to the false positive (FP) rate, considering the accuracy and recall metrics. The number of confirmed Sybils in ASYNPA was 3.49 percent greater than in Vote Trust, at 99.84%.

Kankanamge et al. (2020) [21] examine the level of interest generated by disaster management-associated social media platforms. The study used five indices, commitment, popularity, virality, usage, and engagement, to assess community involvement by different social media outlets. It looked at three states in Australia, and its research concentrated on the official Facebook and Twitter accounts of the emergency response agencies in each jurisdiction. The study found that social media would be a viable medium for capturing community information on disaster management, but it still needs to be used more effectively.

Niles et al. (2019) [22] examined Five of the most expensive catastrophes in the United States in the past decade via the Twitter lens. Twitter activity after disasters, including the prevalence of both broad and narrow terms linked to food security, is positively correlated with network size. According to the investigation, people were likelier to use Twitter in the lead-up to hurricanes and for real-time or post-disaster data following tornadoes and floods. People with medium-sized systems are most likely to share data throughout these crises, and in most instances, they do so more often than is typical, which is consistent across all account types. A kind of social contagion that relies on normal people rather than those with disproportionately large spheres of influence is occurring in catastrophes' aftermath. The findings provide insight into the kind of catastrophe knowledge and target groups that could benefit disaster interaction during catastrophic occurrences.

Lu (Lucy) Yan et al. (2019) [23] evaluated the impact of information sharing on disaster preparation, response, and recovery on social participation. The authors examined all the organization's postings and user comments for three weeks, during, before, and after Hurricane Sandy. This research also shows that aid agencies can better use social media in crisis management by incorporating it into their existing strategies. Even though relief organizations focus on teaching disaster victims about aid delivery, many users were interested in knowing about the volunteers directly. As a result, organizations should provide information geared toward contributors, volunteers, and victims.

Auzzir et al. (2018) [24] assessed the effects of natural disasters on Malaysian small and medium-sized enterprises (SMEs). Among Malaysia's small and medium-sized businesses in 2016, a poll was performed to determine the kinds of disasters that had happened and the effects. The poll was also used to assess the challenges faced by SMEs in coping with natural disasters. According to the study, natural disasters significantly influence Malaysia's SMEs, with floods being the most frequent. SMEs were given advice based on the findings to help them deal with the effects of natural disasters.

Murzintcev (2017) [25] suggested a way for data mining on Twitter to find tweets about a certain event. It provides a method for gathering event-specific hashtags using an automated system. Hashtags are useful identifiers for separating related instances that occurred simultaneously; therefore, the strategy outperforms keyword-based solutions in relevance. Disaster databases were consulted to locate an incident and determine its potential impact region. Other events, such as riots, festivals, and exhibits, can be retrieved using the suggested technique.

Yu Xiao et al. (2015) [26] investigated after-disaster geographical variability in the production of Twitter messages. The tweets can be divided into four categories: mass, material, access, and motivation (MMAM). The MMAM model was mostly validated by the empirical analysis of tweets about Hurricane Sandy in New York City. In forecasting disaster-related tweets, it was discovered that community socioeconomic indicators were far more relevant than population size and damage levels.

2.1 Comparison between Reviewed Literature

As can be seen in Table 3, a wide spectrum of authors applied the method and shared their findings.

Table 3: Summary of literature of review

Author	Methods and Model	Results	Future Scope
Ramakrishnan et al. (2022) [18]	Survey methodology to collect data and PLSSEM	The results showed that underprivileged areas had a low inclination to utilize social media for disaster management.	It would be necessary to conduct more investigations to demonstrate causality based on the data presented in the study.
Dong et al., (2021) [19]	ML	The suggested study method's practicality, validity, and insights toward improved catastrophe management.	A more effective way of analysis would be developed by combining the benefits of several ML models.
Poornima and Murugan (2021) [20]	NDRA	The number of confirmed Sybils in ASYNPA was 3.49% greater than in Vote Trust, at 99.84%.	In the future, the sybils would rise beyond the Vote Trust in numbers.
Kankanamge et al., (2020) [21]	Social media	The study's findings show that postings on social media that include visuals and interactive maps promote community involvement.	The study's findings provided important insights that enlightened policymakers' future steps to develop data-intensive catastrophe management procedures.
Niles et al. (2019) [22]	Twitter word analysis, Keyword time series, and Disaster selection and characteristics.	During various potentially catastrophic events, it's critical to consider the context and target audiences for catastrophe information and how to reach them best.	This kind of data would aid future disaster preparedness and recovery efforts, which could help reduce losses from recent catastrophes and improve resiliency in a changing climate.
Lu (Lucy) Yan et al., (2019) [23]	Five groups that reacted to Hurricane Sandy in 2012 are included in the data set on Facebook.	It was possible that the organization's social media operations could be improved to meet its users' demands and motivate them to engage with organizations.	The method might be used to study other social media sites, such as Twitter, in the future.
Auzzir et al., (2018) [24]	Business Continuity Management (BCM)	Recommendations for SMEs were made to help them deal with the effects of natural disasters.	The study would not go into greater detail on the function of BCM as an SME disaster management strategy, which would be examined in future work.

Murzintcev and Changxiu (2017) [25]	ML	Experiments have shown that the technique could identify different hashtag sets even when numerous simultaneous events with comparable consequences occur. It is more accurate and selective than the current technique.	Future studies would benefit from examining how a message's three coordinates are arranged about one another.
Yu Xiao et al., (2015) [26]	MMAM	In forecasting disaster-related tweets, neighborhood socioeconomic indicators were most relevant than population size and damage levels.	In future studies, ground truthing should be used to compare the information gleaned from social media with what is being gathered.

3. Background Study

Social media has become an important way to spread information about disasters, giving authorities and aid groups real-time data that can help them better manage disasters. Exploration in the area has not gotten the attention it deserves, and it's still difficult to get valuable data. The study's goals are to use data mining and social media analysis to find out how people feel about how disasters are handled and what they need after different disasters. 41,993 tweets cover many natural catastrophes, including their sorts, durations, and damages. Manually categorized tweets gather information on public perception, including the need for targeted supplies, satisfaction with public panic, and disaster response. Eight machine learning models are used to investigate public perceptions of natural catastrophes quantitatively. Data scientists compare the computational time and accuracy of several ML models to suggest the best appropriate model to the decision-makers. Social media's potential role in assisting victims of the same natural catastrophes that fueled the expansion of Twitter is being investigated. The findings in work show that the suggested exploration technique is feasible and valid, and it provides disaster relief organizations with new ideas for improving catastrophe management [27].

4. Problem Formulation

In the last decade, social media platforms like Facebook, Twitter, WhatsApp, and Instagram have grown in popularity. Communication and information exchange among users of SM platforms are becoming more common. Calamity management has played a critical role in reducing the number of people killed and the amount of property and infrastructure that has been damaged. The issue addressed in this section is how to respond to a catastrophe management situation. Warning/evacuation, search and rescue, immediate help, assessment of damage, continued support, and immediate repair or development of infrastructure are only a few of the parts that make up disaster response. The goal of the emergency response is to provide timely assistance that will save lives, improve health, and promote morale. The authors employed two strategies to address the problem: logistic regression (LR) and random forest (RF).

5. Research Objective

This caption contains quantifiable and feasible objectives that would be accomplished during research.
The following are the Research objectives:

- To use social media to accelerate and mechanize disaster resilience in the affected area.
- To address issues related to actual requests for assistance during disasters submitted on OSN (the online social network) and to complete disaster resilience procedures on time.
- To recognize reputable users and prioritize the applications for assistance in times of crisis so that authorities can take immediate action to assist those in need.

6. Research Methodology

This section defines the research methodology based on the impact of social media during natural calamities. The methodology is an amalgamation of ML and graphical user interfaces based on the web. Any information relevant to the disaster category is included in the message dataset in the scenario. After that, the system would prepare data, transforming the original data into something more usable and appealing to the end consumer. The dataset that is used for data preprocessing is trained using random forest, which is a supervised ML technique that is frequently used in organization and regression issues, and LR, which is a statistical analytic technique that predicts a binary conclusion, such as yes or no, based on past observations in a data set. Random forest and logistic regression datasets are stored in pickle files, a useful Python function that allows models to be preserved, reduces the time spent retraining and allows the distribution, commitment, and reloading of previously trained machine learning models. After all machine learning models have been saved in a pickle file, a data analysis based on datasets for message and catastrophe categories would be the next stage. After the machine learning process is completed, it can forward to the web-based graphical user interface. The user interface has an input field where a message with a few keywords can be entered. The algorithm then selects the relevant department after classifying the keywords according to criteria into a certain group or system.

6.1 Technique Used

This part discusses the methods used, like random forest and logistic regression.

- *Text Mining (TM) of Natural Hazardous Impact*

Text mining is getting useful and important information from text sources. Many different types of information, including text and video, can be found on social media sites like Facebook and Twitter. Text mining algorithms should be able to be used successfully in the context of text data for a broad range of applications. Text mining techniques are needed for various applications, including keyword search, classification, and grouping in social media [28].

Keyword Search: A collection of keywords is used to identify nodes in social networks relevant to a particular query in the context of keyword search. Keyword search is a challenge in which authors employ content and linking behavior to find solutions. The general concept is that text documents with the same keywords are linked together. Consequently, identifying clusters of social network nodes containing certain phrases can be helpful [29].

Classification: In the categorization challenge, each node in the social network is paired with a label. The objective of using these tagged nodes is classification. There are several algorithms available for classifying text only based on content. However, the existence of linkages often offers helpful cues for categorization. For instance, label propagation methods can be used with content-based categorization to provide better outcomes.

Clustering: Clustering produced is, therefore, of substantially higher quality. Authors must identify group nodes with similar material to solve the clustering issue. Linkage and content could be used for categorization purposes in various applications [30].

The technique the authors created, called text-mining of natural hazard impacts, enables us to automatically extract data on impacts from text corpora by using machine learning (ML) and natural language processing (NLP) technologies. An earlier prototype application served as the foundation for TM-Impacts [31].

Three complementary modules help compensate TM-Impacts. The first is applying unsupervised topic modeling to identify the text's primary themes. These could contain data on reaction and recovery as well as the effects of the event. The second module focuses on extracting information on certain effects using hand-crafted algorithms and pattern matching (e.g., traffic disruption and power outages). The last module expands on the second by training supervised ML algorithms like the support vector machine and RF to categorize unlabeled text input into impact types.

- *Random Forest (RF)*

Random forest categorization has been increasingly used by machine learning algorithms that try to find spam on social media sites. In artificial intelligence, it is one of the ensemble approaches used to improve the success and accuracy of machine learning algorithms. An RF technique could also help determine which independent variables are important and let the system choose which functions to use. Many studies already show its relevance in empirical investigation, which is determined to be ideal in terms of prediction accuracy when picking various choices for each shrub [32]. The information obtained from the trees is used to create the most accurate projections possible. A decision tree forest assures a more accurate outcome by containing many groups and alternatives, while a single decision tree has a single conclusion and a restricted range of groups. The approach adds the benefit of incorporating randomness into the model by picking the best feature from a pool of randomly chosen features. Regression and classification models for the dependent variables are shown in Figure 1 of the decision tree. Given that there are just two child nodes (binary tree), Scikit-learn utilizes Gini significance to calculate the relevance of each node in a decision tree:

$$_j = w_j C_j - w \quad (1)$$

Where, $_j$= the importance of node j
w_j = weighted number of samples reaching node j.
C_j = the impurity value of node j.
$(j|\ \ )$ = child node from left split on node j.
(j) = child node from right split on node j.
Where, $_j$ represents the importance of node j, w_j is the weighted number of samples reaching node j, C_j represents the impurity value of node j, $(j|\ \ )$ shows child node from left split on node j and (j) is child node from right split on node j.

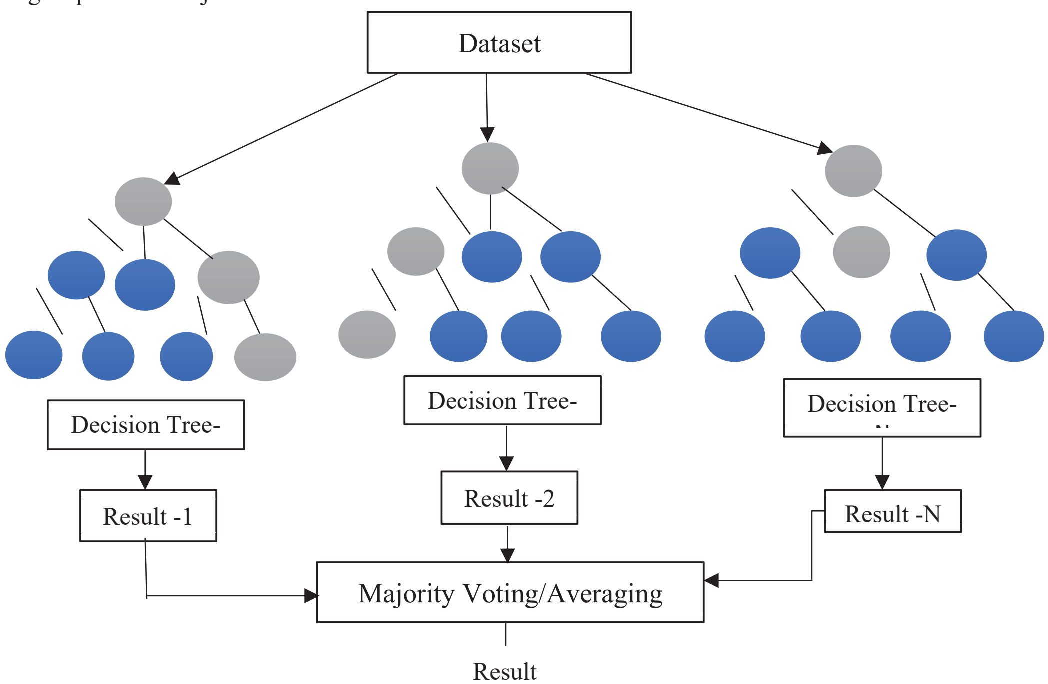


Fig. 1. The structure of Random Forest [33].

Algorithm 1 describes the RF algorithm [34].

Algorithm 1: The RF algorithm.

Training Phase:
Given:
-D: A training set with n instances, p feature, and target variable.
-K: range of classes independent variables.
-B: classifier counts in RF.
Procedure:
For $b = 1B$

1. It is possible to create a bootstrapped sample, D_b f by sampling from the training set, D.
2. Construct a tree using a sample D_b Obtained through bootstrapping.

For a given node t,

(i) Randomly select $m \approx \sqrt{p} \vee m \approx p/3$ feature.
(ii) Use a random selection of features to determine the optimal split feature and cutoffs.
(iii) Data with the best-split feature and cut points should be sent down.

The steps in (i)–(iii) should be repeated until the halting conditions are satisfied.

3. Construct trained classifiers C_b.

Test Phase:

Take the majority vote when combining the B trained classifiers. This is the expected class label from classifiers C_B For test instance x:

$$C_B(x) = argmax_j \sum_{b}^{B} I(V_b(x) = j)\,, forj = 1, \ldots\ldots, K$$

- *Logistic Regression (LR)*

Supervised learning can be included in the use of logistic regression. The formula determines the likelihood of a binary (yes/no) event. The categorical response variable and other variables are modeled using logistic regressions. In a logistic model, independent variables are combined linearly with log odds reflecting the chance of an event occurring. Binary LR estimates the likelihood of a characteristic of a binary variable according to the principles of the covariates under consideration. For the sake of argument, let's say Y is a binary response variable with uncorrelated data $Y_1, Y_2, \ldots., Y_n$, where Y_i=1 when the character is present and $Y_i = 0$when the character is absent. Consider the success probability to be π_i. A collection of illustrative variables that can be either continuous or discrete $x = \left(x_1, x_2, \ldots, x_p\right)$ as a set of variables [35]. Then, the logistic function for π_i is given by

$$log\pi_i = log\left(\frac{\pi_i}{1-\pi_i}\right) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_2 + \cdots + \beta_p x_{i,p}; \quad (2)$$

Where,

$$\pi_i = \frac{exp(\beta_0+\beta_1 x_{i1}+\beta_2 x_2+\cdots+\beta_p x_{i,p})}{1+exp(\beta_0+\beta_1 x_{i1}+\beta_2 x_2+\cdots+\beta_p x_{i,p})} = \frac{exp\ (x'_i\beta)}{1+exp\ (x'_i\beta)} = \Lambda(x'_i\beta) \quad (3)$$

The probability that a sample falls into each of the two categories of the dichotomous response variable is denoted by the symbol π_i And it is obvious that $0 \leq \pi_i \leq 1.$ To estimate the parameters of a model, it can use the logistic Cumulative Distribution Function (CDF) $\Lambda(.)$, with $\lambda(z)\ e^z/(1+e^{-z}) = 1/(1+e^{-z})$ and β^s Signifies a vector of parameters to be assessed. The phrase $\frac{\pi_i}{1-\pi_i}$is called the odds ratio or relative risk [36]. Figure 2 indicates the schematic of a logistic regression classifier.

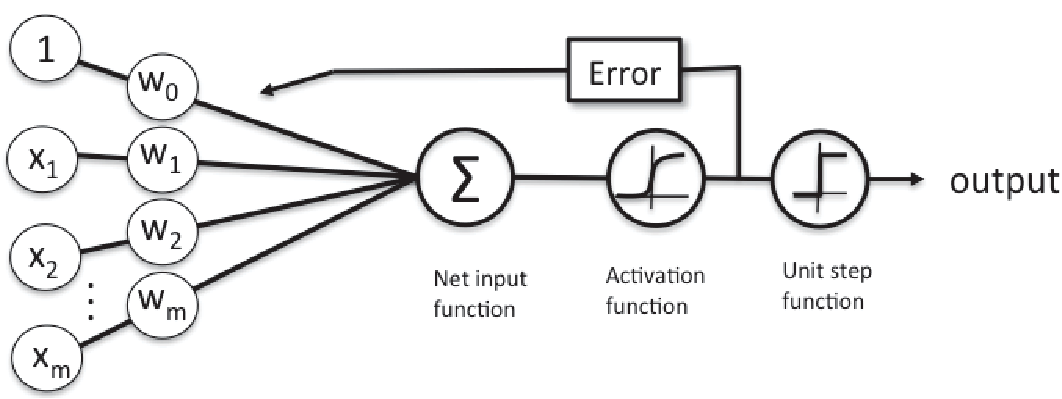


Fig. 2. Schematic of a logistic regression classifier.

7. Proposed Methodology

This section defines the methodology proposed based on the impact of social media during natural calamities, as indicated in Figure 3 below:

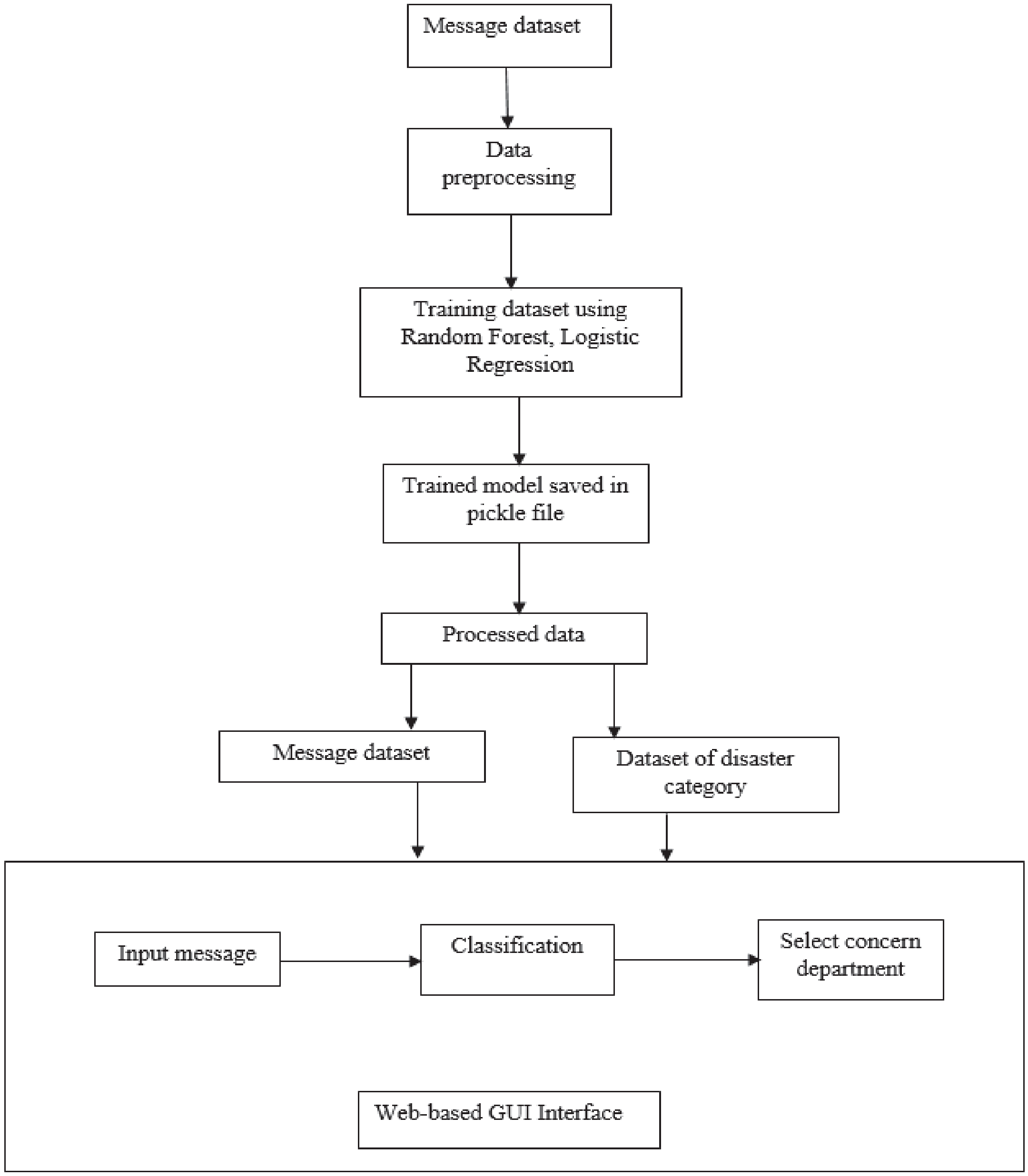


Fig. 3. Proposed Methodology.

Step 1: At this stage, any information about the disaster category is included in the message dataset.
Step 2: After the first phase has been completed, the program would do data preparation, changing data from the form it is in into a much more usable and desirable form.
Step 3: Logistic regression and random forest, two statistical analytic techniques, are trained on the dataset used for data preprocessing to predict binary outcomes, such as 0 or 1, utilizing prior interpretations of the data set. A binary result, such as 0 or 1, may be predicted using the statistical analytic technique of logistic regression by using earlier interpretations of a data set.
Step 4: Datasets trained using random forest and logistic regression are kept in pickle files, a helpful Python utility that enables the preservation of the models, reduces long retraining, and distributes, commits, and reloads pre-trained machine learning models.
Step 5: After storing each machine learning model in the pickle file, the next step is to analyze the data using the datasets for the message and catastrophe categories.
Step 6: After finishing the whole process of machine learning, it can now move on to the web-based graphical user interface. A message can be inputted in the user interface and includes a few keywords. The classification of

the keywords into a certain group or system based on criteria, after which the system chooses the concerned department.

8. Implementation and Results

This research predicts fraudulent behavior using an analytical model based on logistic regression and a machine learning technique called random forest with a constant random Gauss parameter.
The strength of the ties between the various logistic regression methods and the random forest containing a constant random Gauss is shown in Figure 4's correlation plot. A strong relationship exists between logistic regression and random forest, as shown by their high pairwise correlations.

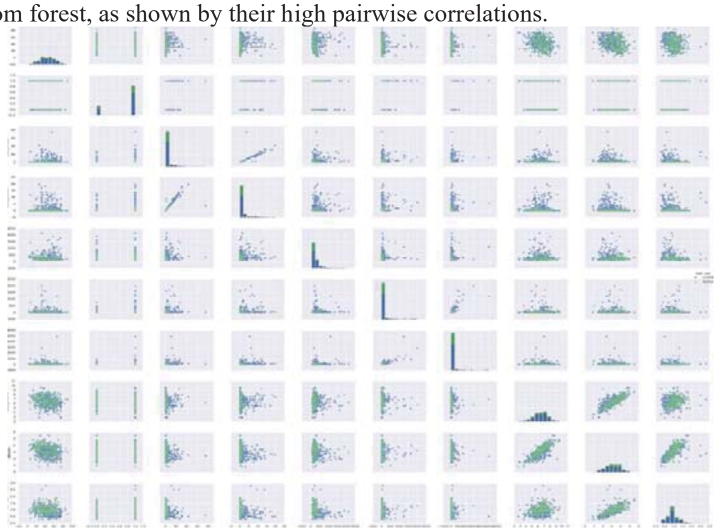


Fig. 4. Illustration of LR and RF correlation.

On the same dataset used to train and test the random forest model and the logistic regression model, the former yielded a higher positivity rate and the latter a higher accuracy rate and a higher precision rate, indicating that the latter was the superior method for detecting fraud. A comparison of the positivity rates of random forest and logistic regression in Figure 5 demonstrates that the latter yields superior results.

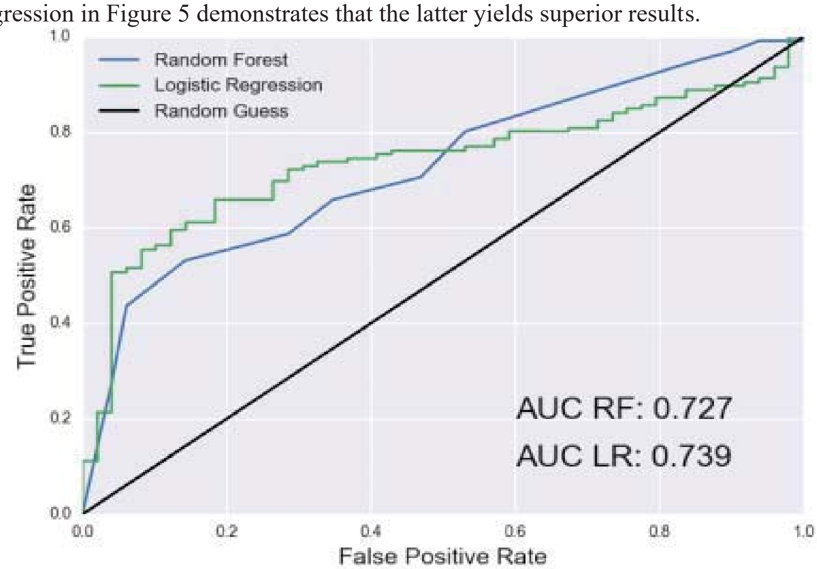


Fig. 5. Comparison of the accuracy of the technique used.

Pickle is a useful Python technique for storing random forest and logistic regression datasets because it keeps models, reduces the time needed for retraining, and makes it easier to share, commit to, and reload previously trained machine learning models. The accuracy of the LR and RF models is shown in Table 4. The reliability of the RF and LR models is shown in Figure 5.

Table 4 shows the accuracy of the models.

Models	Accuracy
Logistic Regression (LR)	73%
Random forest (RF)	72%

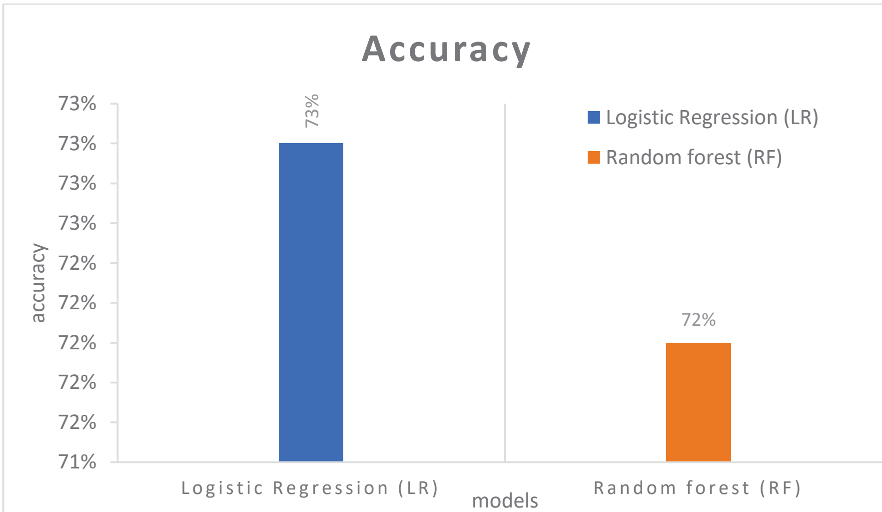


Fig.6. demonstrates the accuracy of RF and LR models.

9. Conclusion and Future Scope

During previous years, the use of social media had various effects on emergency management and disaster response. To better organize, manage, and enable a safe and predictable response to crises and catastrophes, crisis response professionals must recognize this influence, not as the unexpected result of an unmanaged calamity. This paper outlined the significance of social media, its goals, and how social media infrastructure has leveled the playing field for governments of all sizes. Examining systematic methods for using social media as a useful resource in disaster response allows us to explore the typical operational obstacles. Rapid assistance is provided in an emergency so that lives can be saved, health can be restored, and morale can be boosted. The authors used Random Forest (RF) and Logistic Regression (LR) to arrive at a solution. The accuracy results are shown in the accompanying graph. Therefore, disaster-related organizational learning and the minimization of the impact of misinterpretation can aid in the successful process management of diverse companies during calamities and assist in taking effective steps in case a comparable disaster occurs. Extending this research if false news influences anomalous consumer behavior during political choices is an area that needs more attention.

Further research on the impact of false news on different companies' branding is possible along these lines. The authors aim to employ a tweaked version of the naive Bayes method to distinguish between fake and legitimate news. By grouping users according to their online behavior and personalities, they may better identify which users are most likely to propagate the rumor.

Reference

1. Slamet, C., Rahman, A., Sutedi, A., Darmalaksana, W., Ramdhani, M.A. and Maylawati, D.S.A., 2018. Social media-based identifier for natural disaster. In IOP conference series: materials science and engineering (Vol. 288, No. 1, p. 012039). IOP Publishing.
2. Huang, Q. and Xiao, Y., 2015. Geographic situational awareness: mining tweets for disaster preparedness, emergency response, impact, and recovery. ISPRS international journal of geo-information, 4(3), pp.1549-1568.

3. Houston, J.B., Hawthorne, J., Perreault, M.F., Park, E.H., Goldstein Hode, M., Halliwell, M.R., Turner McGowen, S.E., Davis, R., Vaid, S., McElderry, J.A. and Griffith, S.A., 2015. Social media and disasters: a functional framework for social media use in disaster planning, response, and research. Disasters, 39(1), pp.1-22.
4. Adeel, A., Gogate, M., Farooq, S., Ieracitano, C., Dashtipour, K., Larijani, H. and Hussain, A., 2019. A survey on the role of wireless sensor networks and IoT in disaster management. Geological disaster monitoring based on sensor networks, pp.57-66..
5. Li, L., Goodchild, M.F. and Xu, B., 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. Cartography and geographic information science, 40(2), pp.61-77.
6. Wang, Y., Wang, T., Ye, X., Zhu, J. and Lee, J., 2015. Using social media for emergency response and urban sustainability: A case study of the 2012 Beijing rainstorm. Sustainability, 8(1), p.25..
7. Mohsin, M., 10. social media statistics you need to know in 2020. Oberlo. co. uk. Accessed, 31(3), p.2020.
8. Saroj, A. and Pal, S., 2020. Use of social media in crisis management: A survey. International Journal of Disaster Risk Reduction, 48, p.101584.
9. Canel, M.J. and Luoma-aho, V., 2018. Public sector communication: Closing gaps between citizens and public organizations. John Wiley & Sons..
10. Capozzi, L., 2013. Crisis management in the age of social media. Business Expert Press..
11. Falco, E. and Kleinhans, R., 2018. Beyond Information-Sharing. A Typology Of Government Challenges And Requirements For Two-Way Social Media Communication With Citizens. Electronic Journal of e-Government, 16(1), pp.pp32-45..
12. Supa, D.W., 2014. A qualitative examination of the impact of social media on media relations practice. Public Relations Journal, 8(2), pp.1-11..
13. Kent, M. and Taylor, M., 2014. Problems with social media in public relations: Misremembering the past and ignoring the future. International Journal of Interdisciplinary Research, 3(2), pp.23-37.
14. Kent, M.L., 2013. Using social media dialogically: Public relations role in reviving democracy. Public relations review, 39(4), pp.337-345..
15. Lovari, A., & Parisi, L. 2015. Listening to digital publics. Investing citizens' voice and engagement within Italian municipalities' Facebook Pages. Public Relations Review, 41(2), 205–213. https://doi.org/10.1016/ j. pubrev.2014.11.013.
16. Cutter, S.L., Barnes, L., Berry, M., Burton, C., Evans, E., Tate, E. and Webb, J., 2008. A place-based model for understanding community resilience to natural disasters. Global environmental change, 18(4), pp.598-606.
17. Shen, G., Zhou, L., Wu, Y. and Cai, Z., 2018. A global expected risk analysis of fatalities, injuries, and damages by natural disasters. Sustainability, 10(7), p.2573.
18. Ramakrishnan, T., Ngamassi, L. and Rahman, S., 2022. Examining the factors that influence the use of social media for disaster management by underserved communities. International Journal of Disaster Risk Science, 13(1), pp.52-65.
19. Lu, H. and Yuan, S., 2021. What motivates information sharing about disaster victims on social media? Exploring the role of compassion, sadness, expectancy violation, and enjoyment. International Journal of Disaster Risk Reduction, 63, p.102431..
20. Poornima, N. and Murugan, M., 2021. Natural disaster resilience approach (NDRA) to online social networks. Journal of Ambient Intelligence and Humanized Computing, 12, pp.5651-5678.
21. Kankanamge, N., Yigitcanlar, T. and Goonetilleke, A., 2020. How engaging are disaster management related social media channels? The case of Australian state emergency organisations. International Journal of Disaster Risk Reduction, 48, p.101571..
22. Niles, M.T., Emery, B.F., Reagan, A.J., Dodds, P.S. and Danforth, C.M., 2019. Social media usage patterns during natural hazards. PloS one, 14(2), p.e0210484.
23. Yan, L. and Pedraza-Martinez, A.J., 2019. Social media for disaster management: Operational value of the social conversation. Production and Operations Management, 28(10), pp.2514-2532.
24. Auzzir, Z., Haigh, R. and Amaratunga, D., 2018. Impacts of disaster to SMEs in Malaysia. Procedia engineering, 212, pp.1131-1138.
25. Murzintcev, N. and Cheng, C., 2017. Disaster hashtags in social media. ISPRS International Journal of Geo-Information, 6(7), p.204.
26. Xiao, Y., Huang, Q. and Wu, K., 2015. Understanding social media data for disaster management. Natural hazards, 79, pp.1663-1679.
27. Dong, Z.S., Meng, L., Christenson, L. and Fulton, L., 2021. Social media information sharing for natural disaster response. Natural hazards, 107, pp.2077-2104.
28. Aggarwal, C.C. and Wang, H., 2011. Text mining in social networks. Social network data analytics, pp.353-378.

29. Lappas, T., Liu, K. and Terzi, E., 2009, June. Finding a team of experts in social networks. In Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining (pp. 467-476).
30. Zhou, Y., Cheng, H. and Yu, J.X., 2009. Graph clustering based on structural/attribute similarities. Proceedings of the VLDB Endowment, 2(1), pp.718-729.
31. de Brito, M.M., Sodoge, J., Kreibich, H. and Kuhlicke, C., 2022. Text-mining of natural hazard impacts (TM-Impacts): an application to the 2021 flood in Germany (No. EGU22-2001). Copernicus Meetings.
32. Patil, S., Nemade, V. and Soni, P.K., 2018. Predictive modelling for credit card fraud detection using data analytics. Procedia computer science, 132, pp.385-395.
33. Mbaabu, O., 2020. Introduction to random forest in machine learning. Berreskuratua-(e) tik https://www. section. io/engineering-education/introduction-to-random-forest-in-machine-learning.
34. Han, S., Kim, H. and Lee, Y.S., 2020. Double random forest. Machine Learning, 109, pp.1569-1586.
35. Sperandei, S., 2014. Understanding logistic regression analysis. Biochemia medica, 24(1), pp.12-18..
36. Joshi, R.D. and Dhakal, C.K., 2021. Predicting type 2 diabetes using logistic regression and machine learning approaches. International journal of environmental research and public health, 18(14), p.7346.

Blockchain Usage in Decentralized Applications

[a]Rashmi Gandhi*, [a]Jaswant Singh Rohila, [a]Sarvesh Agrawal

[a]CSE, ASET, Amity University, U.P., Noida India, 201303

ABSTRACT

Blockchain technologies distinguish themselves from other systems due to their great technical architecture that allows the technology to be customized for a wide range of use cases. Developers design blockchains with various forms of governance and open or private access depending on the objective applications. Virtually anything important can be tracked on a blockchain system network, lowering cost-cutting and risk about all firms running on data or information. Blockchain refers to how the information (which is then converted into a hash) is stored in" blocks" of data so linked together in a very permanent "chain." The faster data or information is received, the more accurate it's. Also, could reduce fraud and increase application access. Eligible users can cast their usage anonymously with the help of their smartphones and computers. It uses an encrypted key and IDs. The concept of developing a multitask management tool using blockchain technology is introduced. Therefore, the tasks are registered on the blockchain to make it tamper-proof. For issues and problems prevailing in our corporate systems, blockchain can be a practical reason for that. The execution of the work was done in an administered environment with a web application using node.js environment which is used to provide JavaScript runtime and to use backend libraries like express.js which interact with the backend and frontend using API calls.

Keywords—*Blockchain, Ethereum, Rinkeby, Faucet, Meta Mask, Contracts, Decentralized, Distributed Ledger.*

1. INTRODUCTION

In this modern world, Blockchain refers to how the information (which is then converted into a hash) is stored in" blocks" of data so linked together in a very permanent "chain." The faster data or information is received and the more accurate it's. It could reduce voter fraud and increase voter access. Eligible voters can cast their votes anonymously with the help of their smartphones and computers. It uses an encrypted key and ids. Blockchain describes the process of storing data in "blocks" that are connected by a highly strong "chain" and subsequently hashed. It gets harder and harder to modify older blocks as new ones are added to the chain, making each block steadily more secure over time. Blockchain may be a fancy immutable (fixed) ledger that makes it possible to track assets and record transactions on a highly commercial network. A promising technology with a solid cryptographic foundation is called blockchain.

In contrast to other systems, blockchain technologies have excellent technical designs that make it possible to adapt the technology for a wide range of use cases. Developers, for example, may design blockchains with various forms of governance and open or private access depending on the objective. Cryptocurrencies, for example, provide enormous new potential for the development of business models and incentives for users or entire communities, in addition to the technological standpoint. Other technologies, in addition to BT, can be utilized for Open Science. The peer-to-peer data synchronization protocol Dat, for example, provides immutable, decentralized storage and serves as the foundation for scholarly cooperation.

A variety of existing systems, including BitTorrent, inspired the protocol. Now, a decentralized system functions similarly, except that you access a Blockchain network through a framework instead of a server. Let's get going while keeping this in mind. Ethereum is part of a blockchain, which is essentially a peer-to-peer system. There are various networks, but we'll just discuss Ethereum here. Smart Contracts, which are written in Solidity, are operated by Ethereum. Solidity will be simple to learn if you have done any Object-Oriented Programming in the past. The project will require Web3.js, Truffle, and Ganache. Ganache is used to run a local Blockchain to deploy your code since deploying to the Ethereum Virtual Network costs money (Ether) in the form of Gas.

* *Rashmi Gandhi*
E-mail address: rashmibehal@gmail.com

2. LITERATURE REVIEW

Paper	Algorithm	Observation of the paper
Systematic Review of Challenges and Opportunities of Blockchain	Consensus Algorithm	Challenges and opportunities in building a system, less manpower and more secure. DDOS and Sybil attacks can occur. Very less number of transactions.
Blockchain-Based E-Voting System	Exonum	Checking the credibility and availability of the system. Secure but compared to others it's less secure.
Blockchain-Based E-Voting System	Go-Ethereum	Stated about the modal and implementation. Highly stable because of layered architecture.
Secure Digital Task Management System based on Blockchain Technology	Quorum	Stated about the modal and implementation. Highly stable because of layered architecture. Depend on the system of transactions also depends on the system for transactions.
Safeguarding The Records Of Homes, Lands, And Property For Displaced People using decentralized Trust	Consensus Algorithm	Stated about the modal and implementation. Highly stable because of layered architecture. Depend on the system of transactions also depends on the system for transactions. Checking the credibility and availability of the system. Secure but compared to others it's less secure.
When They Go Low: Automated Replacement of Low-level Functions in Ethereum Smart Contracts	Quorum	Maintaining and developing trust in its user concerning their information. Using the consensus blockchain algorithm to test integrity per the CAP theorem. Consistency, availability, and partial tolerance.
Literature Survey Online Task Management System Using Blockchain [4]	Exonum	Listed some present systems for voting High encryption usage Very a smaller number of transaction
Safeguarding The Records of Homes, Lands, And	Go-Ethereum	The integrity of blockchain technologies is safeguarding records. Extremely secure and proper hierarchical management. Expensive and hard to install.

Property For Displaced People using decentralized Trust		
Searching for Trust: Blockchain Technology in an Age of Disinformation	Exonum	Maintaining and developing trust in its user concerning their information. More secure and trustworthy. When data is scarce, it is more prone to attacks.
Devouring the Leader Bottleneck in Blockchain Based on Consensus	Consensus	Using the consensus blockchain algorithm to test integrity by the CAP theorem. Consistency, availability, and partial tolerance. A lot of transactions could not be handled due to partial tolerance.
When They Go Low: Automated Replacement of Low-level Functions in Ethereum Smart Contracts	Go-Ethereum	Automation and Blockchain testing assists in enabling smart records and ensures fraud security First it's a kind of testing tool that integrates testing of both strata. Not adaptable to all testing tools.
Evaluating The Financial Impact of Data Breaches involving account credentials	Consensus	A study on famous financial data breaches and their improvement with the implementation of blockchain technologies. A good architecture system and injection tool to detect data breaches. Equivalent to existing SQL injection tools but less vigil.
Building Decentralized Trust - Multidisciplinary Perspectives on the Design of Blockchains and Distributed Ledgers.	Exonum	Listed some presently distributed ledgers about the to-do list. High encryption usage Very a smaller number of transactions
Digital Ledger Technology: How Blockchain is Changing Organizations and Markets	Go-Ethereum	The integrity of blockchain technologies is safeguarding records. More secure and trustworthy. Expensive and hard to install.

Self-Organized Governance in Blockchain-based Tokenized Markets	Consensus	A study on famous financial data breaches and their improvement with the implementation of blockchain technologies. Consistency, availability, and partial tolerance. A lot of transactions could not be handled due to partial tolerance.
Bringing the Inside Out and the Outside In How Hiring Processes Bridge Across Startup - Ecosystem Boundaries	Go-Ethereum	Maintaining and developing trust in the hiring process of its user concerning their information. More secure and trustworthy. Expensive and hard to install.

[1] The review was conducted using a systematic review methodology. A conceptual description of the planned blockchain-based e-voting application is provided after an introduction to the fundamental structure and characteristics of the blockchain of e-voting. The advancement of symmetrical and asymmetrical cryptography is essential to the creation of blockchain systems. 63 studies that recommended the implementation of blockchain architecture to voting systems have been extracted and examined by researchers from scientific databases. These articles suggest that voting systems enabled by blockchains might offer alternative solutions to conventional electronic voting. They grouped the key current issues into the five categories listed: general, integrity, coin-based, privacy and consensus.

[2] It has long been difficult to create a safe electronic voting system that provides the transparency and flexibility provided by electronic systems while maintaining the fairness and privacy of present voting schemes. In this draught article, the authors assessed a blockchain application for implementing distributed electronic voting systems. The study offers a novel blockchain-based electronic voting system that tackles some of the drawbacks of current systems and assesses some of the well-known blockchain frameworks to build a blockchain-based e-voting system. By the explanation of a case study, which details the conduct of an election and the deployment of a blockchain-based application, they assess the possibilities of distributed ledger technologies, which improve security and decrease the cost of hosting a nationwide election.

[3] Blockchain has the potential to be a disruptive technology of the present that will increase the robustness of e-voting systems. This method offers a chance to take advantage of blockchain's advantages, such as its transparency and cryptographic underpinnings, to achieve an effective e-voting theme. The proposed theme ensures end-to-end verifiability and complies with fundamental requirements for e-voting methods. The method offers a thorough examination of the subject and successfully exhibits its efficacy in achieving an associate in nursing end-to-end verified e-voting subject.

[4] Although Bitcoin continues to be the most notable blockchain application, academics are eager to investigate how blockchain technology may be used to support applications across many sectors by exploiting advantages like non-repudiation, integrity, and confidentiality. They look at how blockchain can be used to support e-voting apps that can guarantee voter privacy, vote integrity, and end-to-end verification. They contend that basic blockchain characteristics like the self-cryptographic validation structure among transactions (via hashes) and the openness of the distributed ledger of records can be used to benefit e-voting. Due to its intrinsic ability to preserve anonymity, maintain a decentralized and publicly distributed ledger of transactions across all nodes, and play a significant role in the field of electronic voting, blockchain technology can be extremely useful. Due to this, blockchain technology is particularly effective at addressing the risk of using voting tokens more than once as well as attempts to sway the outcome's transparency.

[5] Because of the benefits of end-to-end verifiability, distributed ledger technologies like blockchain have lately been employed to create e-voting systems. Blockchain is an extremely alluring substitute for current e-voting systems thanks to features like anonymity, privacy protection, and non-repudiation. To create a successful e-voting system, the study described in this paper also makes an effort to take advantage of these blockchains' inherent qualities. The next section includes the identification of comparisons with these methods along with a thorough study of such systems.

[6] In the Ethereum blockchain, smart contracts are often created using Solidity, a high-level, Turing-complete language. Yet, the Solidity programming language includes a variety of capabilities that give developers controllability across their smart contracts. They term these features low-level functions. Regrettably, poor the using low-level functions can result in security flaws that cause significant monetary losses. As a result, under the official developer guidelines, the Solidity community has recommended substitutes for low-level functions. They start by conducting a sizable empirical investigation on the use of lower-level functions in Solidity-written Ethereum smart contracts. We discover that the bulk of these applications is superfluous for the functionality of the smart contract and that such functions are frequently employed in actual Ethereum smart contracts.

2.1 Blockchain

The blockchain was first used in 2008 by Satoshi Nakamoto when he first created the cryptocurrency called Bitcoin. The technology is based on stochastic consensus protocol which has a decentralized public ledger with PoW (proof of work). It also offers some incentive to record the sequencing of chains. The blockchain data cannot be altered or deleted, it has an append-only data structure such that we can only add new blocks to it. The blocks of the blockchain are linked in a way that hash or encryption is matching with previous blocks as well as the next block making it immutable as depicted in Fig 1.

```
+-----------------+     +-----------------+     +-----------------+
| Block 1         |     | Block 2         |     | Block 3         |
|                 |     |                 |     |                 |
| Header:         |     | Header:         |     | Header:         |
| - Prev hash: 0  |<----|- Prev hash: H1  |<----|- Prev hash: H2  |
| - Timestamp: T1 |-    |- Timestamp: T2  |-    |- Timestamp: T3  |
| - Nonce: N1     |-    |- Nonce: N2      |-    |- Nonce: N3      |
+-----------------+     +-----------------+     +-----------------+
| Body:           |     | Body:           |     | Body:           |
| - Transactions  |     | - Transactions  |     | - Transactions  |
+-----------------+     +-----------------+     +-----------------+
```

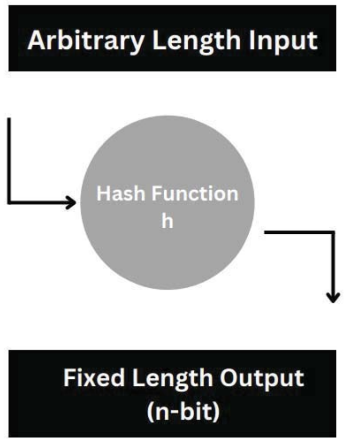
Appendix A. Fig. 1. Blockchain

2.2 Blockchain Network Types

- **Public** permission blockchain is a permission-based Litecoin system that is open and transparent, allowing anybody to alter or examine anything at any time. This technology enables everyone to participate in public blockchain management. It is sometimes referred to as a permission blockchain since it does not require any centralized authority for the verification process, as is the case with Ethereum, Litecoin, and Bitcoin technologies.
- **Private** For security reasons, a person or organization centrally manages private blockchain, an enhanced form of Bitcoin technology. The distributed network trade of Blockchain is not possible with this permission-based blockchain. This technique does not allow anybody to run a full node and start mining, nor does it allow anyone on the blockchain to view transactions, verify or audit them.
- **Consortium** or federated blockchain reduces the single autonomy of the private blockchain by ensuring that numerous persons are responsible for maintaining the chain. Several diverse organizations, such as a group of firms or representatives of the participants, join to make decisions that are advantageous to the network. It is also known as a hybrid blockchain since it incorporates the benefits of both public and private blockchains.

2.3 Secure Hashing Algorithm

Because of the unique characteristic of the hash function, which generates distinct outputs depending on inputs, encryption (SHA-1, SHA2 and SHA-256) [31,32] is the most extensively used secure method connected to blockchain technology. In this case, the hash function acts as a unique key to identify a transaction while also identifying an individual in the oil supply chain. The message size of SHA-1 is 264 bits, the block size is 512 bits, the word size is 32 bits, and the message digest size is 160 bits. The building components used to create a chain are several. Each block, with the exception of the origin block, which does not have a preceding hash, is made up of data and the previous block's hash, as indicated in Fig 2.



Appendix B. Fig. 2. Hash Function

2.4 Ethereum

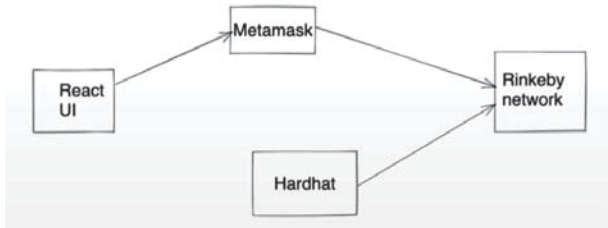
Blocks that are checked and approved will be alluded to as "minted" or "forged" rather than "mined" in contrast to proof-of-work, and users who validate transactions and create new blocks will be referred to as "forgers" or "validators" [10]. According to this system, the user's capital, also known as "stake," influences how a new block's creator is selected in a pseudo-random manner. The person must stake their coins into the network in order to become a forger. The concept of a security deposit exists. The likelihood that a validator will be selected to build the new block increases with increasing stake. Transaction fees are paid to the verifier or forger as payment for creating a block. The proof-of-stake method does not produce any new currencies. A validator risks losing all or part of the risk if they approve a fraudulent transaction. A 51% attack would be practically difficult with proof of stake since he would need to possess at least 51% of all usable digital currency, which would be cost-prohibitive. Because the forgers are chosen at random, proof of stake requires little in the way of resources or computer power.

2.5 Smart Contracts

In 1997, Nick Szabo coined the term "smart contract." Conventional contracts are equivalent to smart contracts in the real world, while smart contracts are digital and self-executing. Contracts are what power the modern economy. an agreement struck between an employer and a group A contract between a tenant and a landlord, a shareholder, and a company, etc. In each of the preceding instances, both parties rely on a third party governed by a centralized authority, a bank, or the government. Instead of relying on the trust of a third party, smart contracts are built on the principles of faith in the virtual environment, logic, and code. Smart contracts back up the idea that "Code is Dictator". A smart contract is a code component that incorporates business logic, contract rules, assets (or anything with value) or currency storage. Because smart contracts are kept within blocks in the blockchain network, they inherit the immutability and distributed nature of blockchain technology. As a result, changing a smart contract after it has been made is extremely complex.

2.5.1 Smart Contracts Deployment over Rinkeby

A MetaMask or Theist browser account with some ether is required in order to deploy a smart contract. In the study, MetaMask is used as an Ethereum client. A browser add-on is MetaMask. The MetaMask displays a 12-word mnemonic when an account is established. A 12-word mnemonic is made up of 12 random words that are used to create a series of accounts, each with its public key, private key, and account address (using the BIP39 technique). Use the Rinkeby faucet to add ether to your MetaMask account (https://faucet.rinkeby.io). To deploy the contract on open networks, a link to a node on the mainet is required. To join the Ethereum network, one must set up their system as a node, which is a time-consuming and arduous process. Instead, a node that is already set up and present on the network is connected via a service called Infura. We may reach a network node using the public API known as Infura. Fig 3. shows how Infura, the provider, and the mnemonic are related. On every Ethereum network, including the main network, the Infura contains nodes. As was already mentioned, a provider is necessary for communication with the Ethereum test network. The account unlocking passphrase and the network to connect to must be manually entered into the provider. For this, the truffle-hdwallet-provider node module needs to be imported and installed into the project. A smart contract's sole role is to represent ownership of stored assets and currencies by shifting them when specified conditions are met. The author of the smart contract implements a set of rules and logic to govern the state of the asset recorded in it.



Appendix C. Fig. 1. Deployment of smart contracts in Rinkeby

2.6 Attack Management

- *Double Spending Attacks***:** One distinguishing property of real currency is that you cannot pay the same bill at two distinct locations (unless you are a magician or a fraudster). Digital assets, on the other hand, are simply copied; after all, they are just bits of 1s and 0s. So, in the
  However, if two independent blocks were mined at the same time and both transactions made it there, the block with the most network confirmations will be kept and the other would be ignored.
- *Routing Attacks***:** By now, it should be clear that blockchain technology needs a strong network to operate. Through BGP, the ISPs communicate with one another and exchange route data (Border Gateway Protocol). This protocol is outdated and contains various flaws that an attacker might use against it.

3. APPLICATIONS

3.1 Eth-To-Do: A decentralized multi-managed task manager

The management system structure has been fragmented into 3 layers to achieve an optimal design. These layers are discussed down below:

I. UI and Frontend Security layer is used for communicating with a user and the administrator. It has mainly 2 key characteristics: authorization and authentication of the user to verify that the access to the portal is confined only to the authentic user.

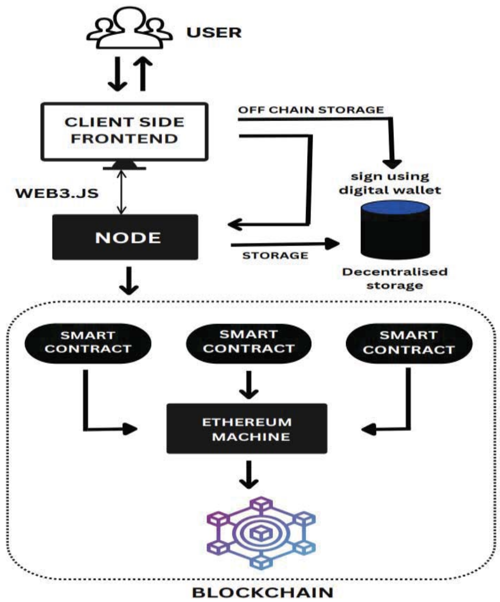
II. Authentication Management Control layer is used to communicate between layers 1 and 3 by providing API and service interface used for these layers to communicate with each other.

III. Transaction administration layer for list management is the main layer of their algorithmic structure. The methods for adding data in a data field defined at the layer are mapped onto the blockchain that will be mined later. The information/data is also storing the mapped transaction provided by the person who used it in the first layer for identification.

IV. Ledger Synchronization layer notes are stored in the database at the backend service. Users are then able to trace their stored notes using their credentials on MetaMask given to them when their note is being mined and appended onto the blockchain using the MetaMask wallet, here in meta mask we are using Rinkeby faucets shown in Fig. 4 and Fig 5.

3.2 The blockchain-based decentralized voting system

In this modern world, Democracies are still using old ballot papers and EVM. In recent years many alleged EVMs due to their flawed election reports. There is a very less credible answer about the internal architectural designs and algorithms used in the EVM and how they can become vulnerable to attacks. Anyone with physical access can also infect the machine, thereby affecting all votes cast on that device. For a strong online election scheme, a variety of functional and security requirements should be met they are eligibility and Authentication: Only authorized voters should vote. Uniqueness: One voter should be allowed to vote once. Accuracy: The vote should be recorded correctly. Integrity: Modification or loss of voting data shouldn’t occur. Reliability: The system must be designed such that it doesn’t show any kind of bug even after failures and loss of internet. Convenience: A user interface that will be easily understood with less amount of a skill set. To resolve the drawbacks of our indigenous voting system. A blockchain can be used. It is a distributed, immutable, incontrovertible, public ledger. It is a distributed network, there is no specific point at which the network can break.



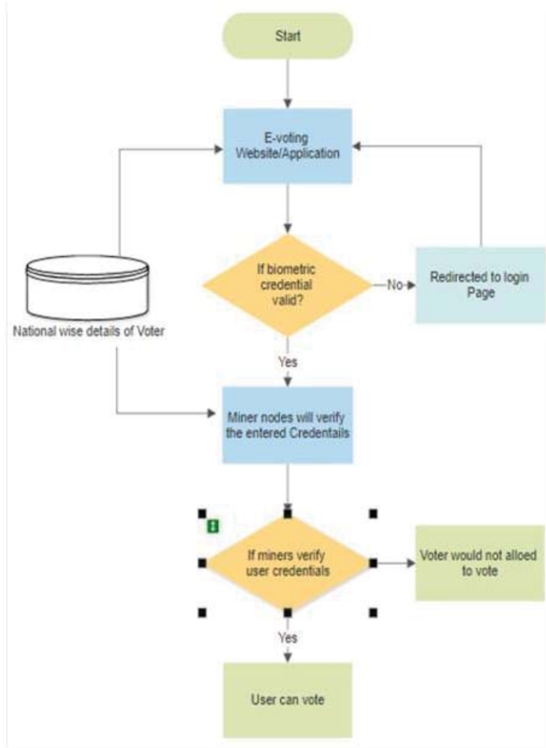
Appendix D. Fig. 4. proposed system for the Eth-To-Do app

Various other voting methods: -

3.2.1 *Ballot Paper*: The ballot paper was used earlier. Nowadays, electronic voting machines (EVM) are used to record votes. The voting process in EVM is easier than ballots. Better productivity, Votes are counted at the press of a button by the voters. While the counting of votes (paper ballot) begins after voting is done. The cost of ballots is higher than EVM. Since the idea was earlier designed without using electronic circuits, there is no risk of being hijacked during the election. No vulnerability in terms of software and hardware security.

3.2.2 *EVM*: The Indian electronic voting machine (EVM) was developed in 1989 by the Election Commission of India. There are no external communication paths which make hackers difficult to intrude. They are cost-effective and economical. Votes are counted at the press of a button by the voters. Machines are easier to carry and transport. Voters don't get the credibility of whether a vote is registered to their specified candidate. These machines can also be tampered with.

3.2.3 *Blockchain As Voting Tool:* When a transaction happens, it is then registered as a block of information, the data block can record the information. Every block is linked to another block from front and rear: Thee blocks then develop a long chain with each other confirming the exact time and transaction sequence. Every block is connected to prevent the chain from breaking. Transactions locked in the chain: Each block verifies itself from the adjacent blocks, therefore, the whole blockchain becomes immutable and tamper-proof. The ledger can exist in many different locations in the country: there is no centralized ledger. Therefore, exist in many different locations. There is whole distributed control over who can add and verify new transactions to the ledger. Any "block" to the ledger must have the previous version of the ledger, making an immutable chain from where the blockchain gets its name, and thus prevents tampering with the integrity of previous blocks. Most of the network nodes must be in concurrence before a proposed new block of entries becomes a permanent part of the ledger.



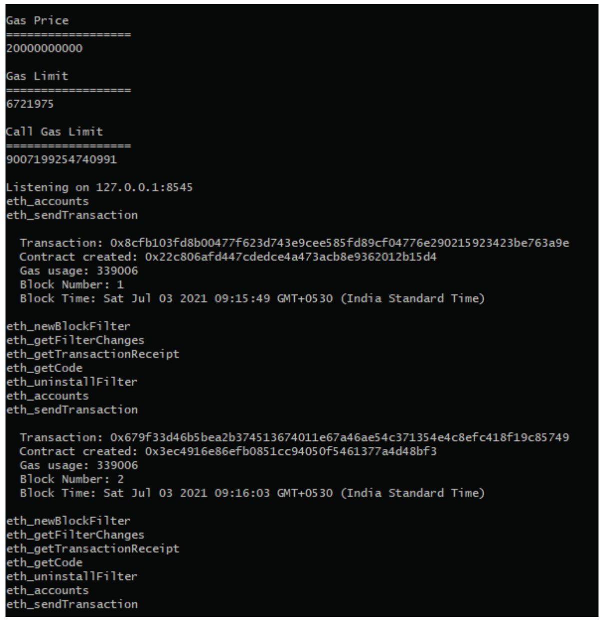
Appendix E. Fig. 5. Proposed system for the voting system.

4. Results and Discussions

The modal used is implemented with popular NPM libraries are Ganache (to create a private blockchain network for testing on localhost), Web3 (a library which allows communicating with a local or remote Ethereum node), Solidity (used for implementing smart contracts), Morgan Js (**Morgan** is an HTTP request logger middleware for **Node**. Js.). The core of the modal is emulated in Fig 6.

Appendix F. Fig. 6. Ganache free testers and keys to emulate the Ethereum blockchain.

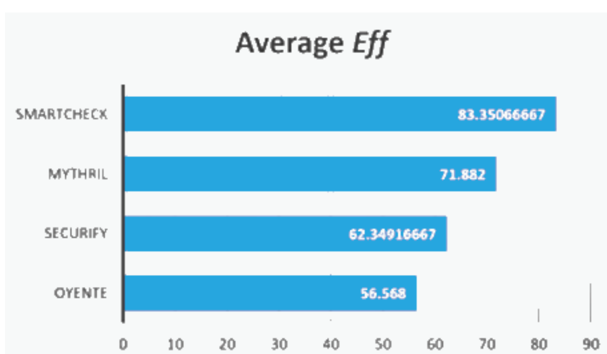
By hitting $ npm install ethereumjs-testrpc web3, $ ganache-cli on cmd. It is showing 10 test accounts that are created by web3 testrpc to play with. These fake accounts also come with 100 ethers (fake). The values are shown in Fig 7.



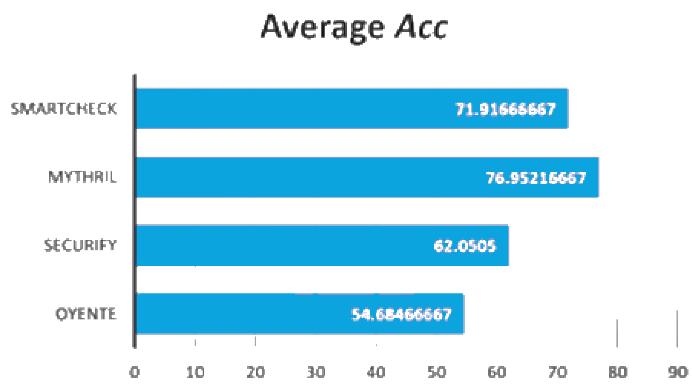
Appendix G. Fig. 7. Each block details on Ganache interface

For every new session created for voting by every individual voter after all verification and identification are completed, a new contract and block will be created dynamically having all major details like gas usage, block time and transaction hash with it.

4.1 Vulnerabilities of Smart Contracts: Turing's incompleteness, or the fact that while a sizable portion of calculations is enabled by the Bitcoin programming language, practically all of them are not. In principle, this is an overarching challenge for scripters to overcome, since every loop can be emulated simply by looping the underlying code several times with an if clause, but this results in highly space-efficient scripts. This is done to avoid endless loops while validating transactions. There are various testing tools from where the effectiveness of the smart contracts such as SmartCheck, Mythril, Security, and Oyente are checked and presented in Fig 8 [19].



Appendix H. Fig. 8. Comparison of the average effectiveness of security testing tools



Appendix I. Fig.9. Comparison of the average accuracy obtained by testing tools

4.2 Fast operations: Due to the large amount of data storage requirements in computer science now a day, the data must be searched and retrieved in an efficient way. For this, the
data can be viewed be organized in a different way or using different data structures. The To-Do List is optimized using this fast algorithm customized by accumulating different searching algorithms and proposed a new simple key search algorithm for the interpolation search method. The algorithm that we developed can be referred to by the name Expeditious Search Algorithm. It is a simple key-searching algorithm which uses the divide-and-conquer technique method. It is a kind of interpolation algorithm. The basic idea of the Expeditious Search Algorithm is based on a binary search technique. Expeditious Search Algorithm requires the input array in sorted form. The array is divided based on some conditions and any one of the partitions is selected for searching the key and all the other sub-parts are discarded. This procedure repeats. The algorithm takes the average case time of logarithmic of the order of the number of operations on the assumptions mentioned in Table 1.

Table 1. Voters with casting and tallying in minutes

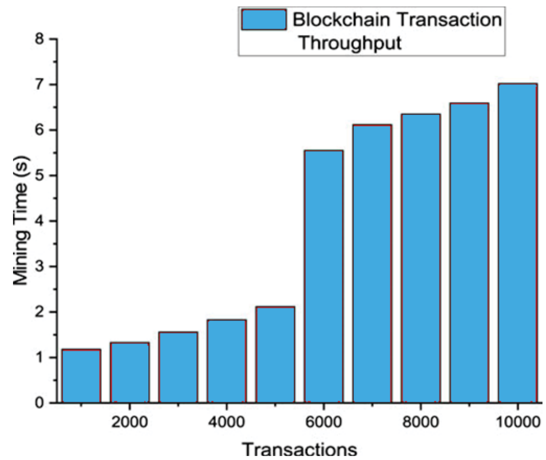
Number of voters	Vote casting (min)	Vote tallying (min)
1,000	0.01	0.17
2,000	0.04	0.23
3,000	0.08	0.46
4,000	0.12	0.53
5,000	0.23	0.78
6,000	0.34	0.81
7,000	0.45	1.20
8,000	0.62	1.33
9,000	0.68	1.41
10,000	0.73	1.56

4.2.1 Blockchain Transaction Throughput

This gauges how quickly transactions are mined and added to the blockchain system's ledger.

T= a/b

Where **b** is the mining time and **a** is the number of transactions.



Appendix J. Fig. 10. Comparison of the blockchain transaction throughput

4.2.2 Response Time

This is a measurement of the amount of time that passes between a user initiating an activity and the system showing them the results.

The average system reaction time is equal to the sum of the individual times that each user action or test was completed.

Voting takes an average of 0.01+0.04+0.08+0.12+0.23+0.34+0.45+0.62+0.68+0.73/10 seconds.

Voting time average = 3.3/10 = 0.33
Average time to count votes is 0.17 + 0.23 + 0.46 + 0.53 + 0.78 + 0.81 + 1.20 + 1.33 + 1.41 + 1.56/10

The average voting time is 8.48/10, or 0.848.

This demonstrated that for every thousand increases in the number of transactions, the blockchain's mining time increased by an overall factor of 0.18 s. Additionally, the system's response time to a variety of user activities was assessed using an increasing sample size of voters. According to the results, the system's reaction time increased by an average of 0.33 minutes for every 1,000 voters during vote-casting operations, while this increased to an average of 0.848 minutes for every 1,000 voters during vote tallying.[20]

5. Conclusion And Future Scope

The idea of creating an electronic voting machine using blockchain technology and integrating it with the current online voting system is explored in this study. To prevent tampering, the vote is recorded on the blockchain (as blocks). Now, blockchain can be a very practical explanation for issues and challenges that are currently present in our election system. The work was carried out in a controlled setting with a web application that makes use of the node.js environment, which offers JavaScript runtime, and backend libraries like express Js that communicate with the frontend and backend utilizing API calls. Additionally, we have examined consensus algorithms that contain flaws that have led the way for new consensus algorithms like Casper that try to strengthen the existing algorithms. At last, blockchain is a brand-new technology that combines the discipline of coding with the whole consensus-building process. A blockchain's security depends on the software that runs it. As a result, thoroughly test and audit your blockchain to find any security issues before making it public.

References

1. Juels, Ari; Dario Catalano; Markus Jakobsson (November 2002). "Coercion-Resistant Electronic Elections". Cryptology ePrint Archive (165). Archived from the original on 7 April 2014. Retrieved 2 May 2012.
2. R. Krishnamurthy; Geetanjali Rathee; Naveen Jaglan 2020. "An enhanced security mechanism through blockchain for E-polling/counting process using IoT devices".

3. Mayank Sahu. “Cryptography in Blockchain: Types & Applications” Available at “https://www.upgrad.com/blog/cryptography-in-blockchain”. June 2021.
4. Zhang, Y.; Kasahara, S.; Shen, Y.; Jiang, X.; Wan, J. Smart contract-based access control for the Internet of things. IEEE InternetThings J. 2019,6, 1594–1605.
5. Li, R.; Song, T.; Mei, B.; Li, H.; Cheng, X.; Sun, L. Blockchain for Large-Scale Internet of Things Data Storage and Protection. IEEETrans. Serv. Comput. 2019,12, 762–771.
6. Vitalik Buterin. (2015). Ethereum White Paper. Available at: https://github.com/ ethereum/ wiki/ wiki/ White-Paper.
7. Garg, K., Saraswat, P., Bisht, S., Aggarwal, S. K., Kothari, S. K., & Gupta, S. (2019, April). A comparative analysis of e-voting system using blockchain. In 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU) (pp. 1-4). IEEE.
8. Alvi, S.T., Uddin, M.N. and Islam, L., 2020, August. Digital voting: A blockchain-based e-voting system using biohash and smart contract. In 2020 Third International Conference on Smart Systems and Inventive Technology (ICSSIT) (pp. 228-233). IEEE.
9. Lahane, A.A., Patel, J., Pathan, T. and Potdar, P., 2020. Blockchain technology-based e-voting system. In ITM Web of Conferences (Vol. 32, p. 03001). EDP Sciences.
10. Vaibhav Anasune, Pradeep Choudhari, Madhura Kelapure, Pranali Shirke, Prasad Halgaonkar. “Literature survey-Online Voting: Voting System Using Blockchain” June 2019.
11. Khan, K.M., Arshad, J. and Khan, M.M., 2018. Secure digital voting system based on blockchain technology. International Journal of Electronic Government Research (IJEGR), 14(1), pp.53-62.
12. Adida B. and Rivest, R. L. (2006). Scratch & vote: Self-contained paper-based cryptographic voting, in
13. Proceedings of the 5th ACM Workshop on Privacy in Electronic Society, ser. WPES '06. New York, NY, USA: ACM, 2006, pp. 29-40.
14. Hjálmarsson, F.Þ., Hreiðarsson, G.K., Hamdaqa, M. and Hjálmtýsson, G., 2018, July. Blockchain-based e-voting system. In 2018 IEEE 11th international conference on cloud computing (CLOUD) (pp. 983-986). IEEE.
15. uhi Ta ̧s and Ömer Özgür Tanrıöver . “A Systematic Review of Challenges and Opportunities of Blockchain for E-Voting”. Received: 18 July 2020; Accepted: 7 August 2020; Published: 9 August 2020.
16. Friðrik Þ. Hjálmarsson, Gunnlaugur K. Hreiðarsson. “Blockchain-Based E-Voting System”
17. Kashif Mehboob Khan1, Junaid Arshad2, Muhammad Mubashir Khan. “Kashif Mehboob Khan1, Junaid Arshad2, Muhammad Mubashir Khan”
18. Ahtisham Younas, Parveen Ali, Five tips for developing useful literature summary tables for writing review articles.
19. Parizi, Reza M., et al. "Empirical vulnerability analysis of automated smart contracts security testing on blockchains." arXiv preprint arXiv:1809.02702 (2018).
20. Umar, B.U., Olaniyi, O.M., Olajide, D.O. and Dogo, E.M., 2022. Paillier Cryptosystem-Based ChainNode for Secure Electronic Voting

Wheat Head Detection Using Deep Learning Techniques

Dhruvisha N. Jaiswal[a], Kirtikumar J. Sharma[b], Prashant B. Swadas[c]

[a]*PG Student, Computer Engineering, BVM Engineering College, Vallabh Vidyanagar, Anand - 388120, Gujarat, India*
[b]*Assistant Professor, Department of Computer Engineering, BVM Engineering College, Vallabh Vidyanagar, Anand - 388120, Gujarat, India*
[c]*Associate Professor, Department of Computer Engineering, BVM Engineering College, Vallabh Vidyanagar, Anand - 388120, Gujarat, India*

Abstract

This article aims to study an object detection methodology applied to the Global Wheat Head Detection (GWHD) Dataset. It is important to estimate the density and size of the wheat head from images or manual surveys. This study has been through four dominant architectures of object detection, namely Faster R-CNN, EfficientDet, Detectron2, and YoLov5 to design a novel and powerful wheat head detection model. Model exploration is tuned on optimizing the performance of models. Furthermore, also studied extensive exploratory data analysis, data cleaning, and data splitting and adapted the best data augmentation techniques to the required context. Additionally deliberated semi-supervised learning, specifically pseudo-labeling, to optimize previous supervised object detection models. Besides this study focus on ensemble learning, test time augmentation, bootstrap aggregating, and multi-scale ensemble to reach higher performance. Finally, for the post-processing technique, execute weighted box fusion.

Keywords: Wheat Head Detection, Data Augmentation, Test Time Augmentation, Weighted Boxes Fusion, Faster-RCNN, EfficientDet, Detectron2, YoLov5

1. Introduction

Wheat is the most cultivated cereal crop in the world, and India is the second largest producer in the world [1]. Object Detection deals with detecting instances in images. It has demonstrated the automation of laborious and tedious tasks. The detection of wheat heads in plant images is a crucial piece of work for estimating relevant wheat traits, including wheat head population density. Due to the large amount of data in computer vision applications, object detection requires architectures such as deep learning mode. This article's main goal is to develop a reliable object detection model that can identify wheat heads in an image with numerous wheat heads and generalize to all varieties of wheat crops from around the world. As a result, detection models created for wheat head detection need to be reliable in various growing environments. Many strategies, including pseudo labelling, test time extensions, bagging, multi-scale ensembles, and post-processing algorithms, were combined to improve the performance of detection models [2].

2. Related work

In general, deep learning algorithms strive to learn hierarchical features, analogous to different levels of abstraction [2]. In this section, object detection using deep learning techniques is covered, followed by an explanation of some research on wheat head detection. Wu Wei et al. [3] proposed wheat head detection and enumeration of wheat grains based on a deep learning method and models under several scenarios and scales, which gives us knowledge about wheat grains. According to this study, the number of grains played a crucial role in influencing production. R-CNN is used by authors with a loss under 0.5 and mAP:0.9 to recognize wheat head grains more quickly. The head is detected in less than two seconds. The model is robust to dissimilar backgrounds and different levels of grain crowding. Also, it appears that one of the issues is that the suggested solutions were created for controlled environment shots rather than actual field photos [4]. The focus on the same type of wheat throughout the training and testing of the models is another limitation that results in the models becoming excessively tailored to that particular wheat type [5]. The emphasis on the mAP0.5 measure is a restriction on what head detection algorithms can do. In other words, if the head covers more than 50% of the IoU detection, the detection is judged to

* Dhruvisha N. Jaiswala
E-mail address: dhruvishakali99@gmail.com

be accurate. The GWHD dataset is used in this paper to address the first restriction [1]. That includes photographs taken in the field and exhibits a greater variety of sources, methods of acquisition, environmental factors, and wheat varieties. An alternative metric, mAP0.5:0.75, was adopted to make the second restriction more clear. The GWHD Dataset is the only one with such a high prevalence of overlapping and occluded items. To prevent overfitting and arrive at a universal answer for all wheat photos, significant work was also put into data transformation and regularization of deep models.

In May 2020, the GWHD dataset was released. Kaggle and an AI crowd [6] conducted the GWHD challenge in 2021. Multiple submissions and research happened at that time. The winning paper was also studied. [1].

3. Data Description

The dataset [1] consists of more than 1024×1024 pixel images, each comprising more than 270k distinct wheat heads, along with the appropriate bounding boxes. There are 44 unique measurement terms and images from 11 different nations. A measurement session is a collection of images taken with a particular sensor at the same location over a predetermined period of time (often a few hours). The test dataset will be made up of images from North America (excluding Canada), Asia, Oceania, and Africa and consist of approximately 2000 images. The training dataset will consist of images gathered in Europe and Canada, covering approximately 4000 images [6].

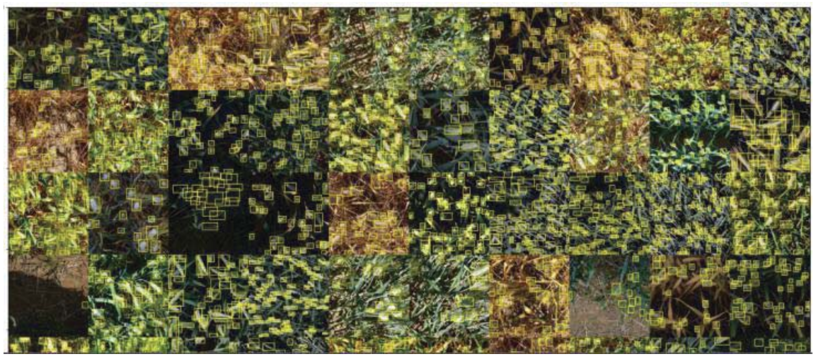


Fig. 1. Some images visualization with bounding boxes from Dataset

4. Methodology

Using object detection models to discover wheat heads in images involves four main processes. The first stage is the Exploratory Data Analysis (EDA), which was incorporated into the design to enhance data collection. Next comes the semi-supervised learning procedure. The third phase ushers in the ensemble one. Lastly, a post-processing block was constructed to maximize the correctness of the entire design, as seen in Fig. 2. Each step will be further explained in the subsections that follow.

4.1 Data transformation (EDA):

Exploratory data analysis was crucial to the work and assisted in producing the best possible inference outcomes. This procedure is circular. Analyze the data, create hypotheses, models, and conclusions, validate them, compare them, and keep analyzing the data until you get satisfactory findings. The techniques listed below for data transformation.

4.1.1 Data analysis:

Data analysis is not only a part of the experimental framework, as it was in the traditional object identification techniques, but it is also an essential stage in the development process. GWHD [6] dataset helps to overcome such problems by providing both diversity and quantity of well-labelled images. The analysis found 49 images without wheat labelling. Also, most wheat heads are smaller in size, and some images have low and high brightness.

4.1.2 Data cleaning:

Data cleaning is the procedure of identifying incomplete, imprecise, erroneous, or unnecessary parts of data and updating, modifying, or removing them [2]. To stop models from being misled and to increase the accuracy of outcomes, data cleansing is a necessary step. Some bounding boxes cover multiple heads, while smaller bounding boxes do not cover any heads. So, by setting a threshold that took into account the range of projected sizes for wheat heads, the enormous and tiny bounding boxes were eliminated.

4.1.3 Data Splitting:

A sufficient validation data set is essential for having reliable validation precision. The data was divided using stratified k-fold splitting. This approach divided the data into k folds, ensuring that each fold had approximately the same number of images per source and bounding box distribution. In that method, cross validation proved more reliable than randomly dividing the data. The training and test sets are guaranteed to include the same percentage of the feature of interest as the original dataset [7].

4.1.4 Data Augmentation:

During experiments, it was discovered that different data augmentation techniques had a stronger impact on increasing the average precision of models [2]. There are numerous ways to alter the same image, such as horizontal and vertical flips, cuts and resizing, rotations, cropping, random erasing, CutOut, CutMix.

4.2 Pseudo Labelling:

During the inference phase of this work, a semi-supervised approach called pseudo-labelling [8] was applied. In this study, the model did not structure a real-time object recognition application; therefore, it was desirable and a viable choice to try pseudo labelling (PL). Pseudo-labelling was used to train an object detector in a supervised manner, then consider adding pseudo labels to the training dataset, retrain the model, and predict new bounding boxes for test data.

4.3 Ensemble Learning:

Ensemble learning combines numerous individual models to acquire better generalization performance. The following ensemble techniques are used in the mentioned model:

4.3.1 Test Time Augmentation (TTA):

Using post-processing techniques, TTA takes several outputs, separates the detected wheat heads, and then merges the results [9]. Whereas other data augmentation tasks are done prior to or during model training, this one is finished during the inference phase. The objective is to provide the same model with numerous versions of the same image, then integrate the findings using a post-processing technique by extracting the detected wheat heads from various outcomes [2].

4.3.2 Bootstrap Aggregation (Bagging):

A machine learning (ML) ensemble algorithm called bootstrap aggregating, commonly known as bagging, aims to improve the precision and stability of ML systems. Moreover, it lowers variance and aids in preventing overfitting [10].

4.3.3 Multi-Scale Ensemble:

Multi-Scale Ensemble Object detection is performed at multiple output layers so that receptive fields match objects of different scales [11]. Several EfficientDet models were trained by the authors [2] using various input measurements, specifically 512×512 (pixels) and 1024×1024 (pixels). To compare the results, train a few models with 1024×1024 input resolution, but most essential, ensemble both strategies.

4.4 Post-Processing Algorithm:

Post-processing steps were needed to generate the right bounding boxes for object detection models. Weighted Boxes Fusion (WBF) was used to combine all predictions from multiple models. The average boxes were constructed using the confidence scores of all the presented bounding boxes.

Workflow diagram:

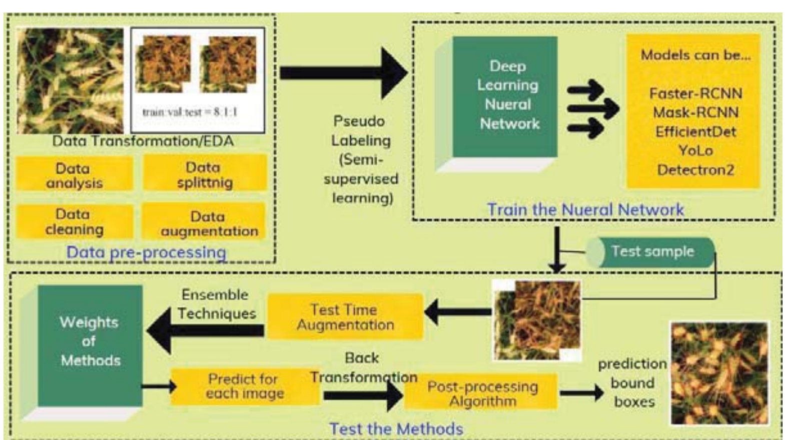


Fig. 2. The workflow of training and testing model detectors for wheat heads.

5. Results & Discussion

In this article, four deep learning algorithms were discussed: Faster R-CNN, EfficientDet, Detectron2, and YoLov5.

Table 1. Following are hyper parameters for Faster RCNN [2]

Hyper Parameters	Description
Backbone	ResNet50
Optimizer	Stochastic Gradient Descent (SGD)
Momentum	0.9
Weight Decay	0.005
Batch size	16

When applying a cleaning on the data and increasing the size of the original dataset by adding one random augmentation for each original image, the mAP increases by around 2%. The best result (68.46%) was achieved by combining the cleaning, multiplying the size of the original dataset by 4, and adding the pseudo labelling [2].

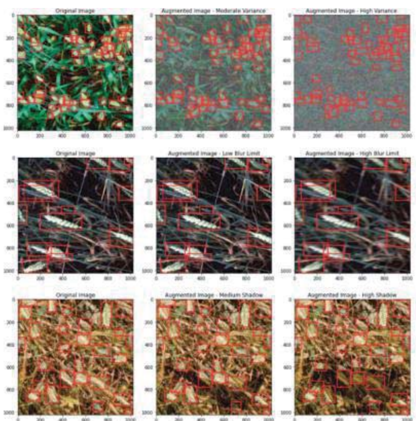


Fig. 3. Samples images from Data Augmentations where the first column represents the original images and the second and third column shows the augmentations with different parameters. [2]

Furthermore, reduced the batch size to 4, augmented the number of epochs to almost 100, employed weighted box fusion as post-processing, and used a learning rate scheduler that was distinct from the one used for Faster R-CNN. Combining bagging with pseudo labelling and augmentation, the result achieved 73.43%. The final experiment combined two models, the first with a 512-pixel image size and the second with a 1024-pixel image size, and applied data augmentation techniques such as TTA, bagging, and multiscale ensemble; the result reached 74.22%, which was the higher result achieved [2, 13, 14].

To elaborate on this topic, execute all of these techniques on another dataset from Kaggle [15]. The dataset is composed of 3422 images of 1024×1024 pixels containing 147k+ unique wheat heads, with the corresponding bounding boxes. The images appear from 7 countries with nine research institutes. The following images are samples of the dataset with bounding boxes. (Fig. 4)

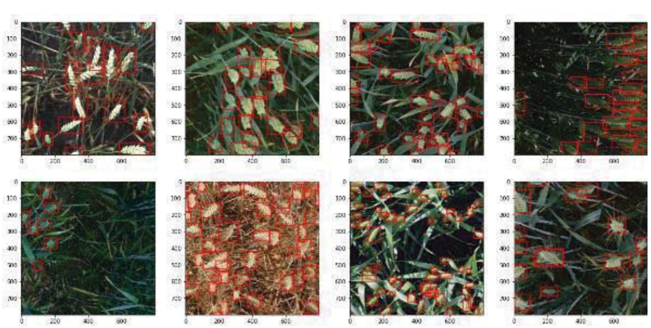


Fig. 4. Sample images from the above mentioned Dataset [15]

Table 2. Following are hyper parameters for respective models.

Hyper Parameters	FasterRCNN	EfficientDet	Detectron2	YoLov5
Backbone	ResNet-50	EfficientNet	ResNet-50	BottleNeckCSP
Batch Size	16	4	16	16
Learning rate	0.5	0.0002	0.001	0.01
Momentum	0.9	0.1	0.9	0.937
Decay	0.001	0.001	0.0001	5e-4
Iteration/Epochs	20 epochs	17 epochs	10000 iterations	1350 iterations
Optimizer	Adam/SGD	AdamW	Adam/SGD	Adam/SGD
IoU Threshold	0.5	-	0.7	0.8

To get better performance out of the data, augmentation techniques were applied using some well-known ones, such as Random Brightness, Flip, Crop, Rotation, Contras , HueSturation, RGBShift, Horizontal and Vertical Flip, Cutout, CutMix.

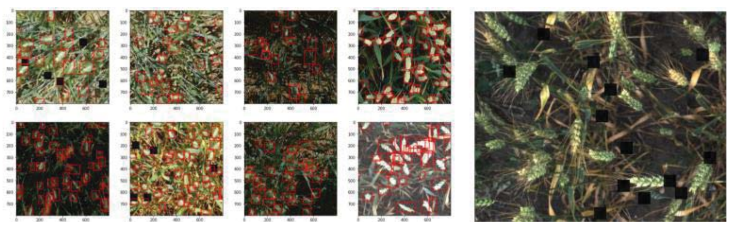
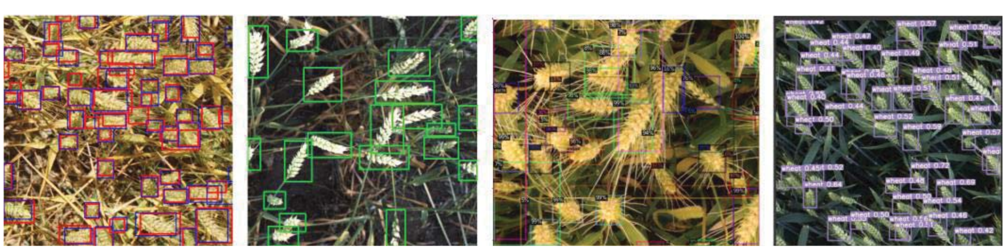


Fig. 5. Some sample images which show the data augmentation on the actual dataset.

Table 3. Performance of the different object detection models.

Model Name	Data cleaning	Data Augmentation	mAP %
Faster-RCNN	Ture	True	56%
EfficientDet	True	True	54%
Detectron2	True	True	-
YoLov5	True	-	64%

The probability of the existing wheat head over each bounding box is also represented in Fig. 6 (c and d). So, as per the performance analysis (Table 3), YoLov5 gives better precision than the other three models.



(a) (b) (c) (d)

Fig. 6. Results after implementing four studied object detection models. Predicted wheat heads from faster-RCNN (a), EfficientDet (b), Detectron2 (c), and YoLov5 (d).

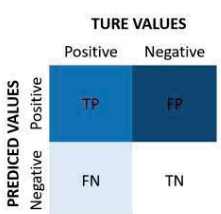
In Fig. 6 (a), the model predicted wheat heads in blue bounding boxes, and ground truth wheat heads are shown in red bounding boxes. The probability of the existing wheat head over each bounding box is also represented in Fig. 6 (c and d). So, as per the performance analysis (Table 3), YoLov5 gives better precision than the other three models.

After implementation of all four models on a small dataset [15], the highest accuracy was about 64%. So, decided to choose a bigger dataset [6] to get more accuracy and tried to get the minimum heads count difference between ground truth values and predicted heads count values. The following results and discussion will be on the GWHD dataset [6] using FasterRCNN and YoLov5 detection models. The dataset consists of 6511 high-resolution RGB images from 47 domains, with 3655 images in the training dataset and 2856 images in the testing dataset. Also a file with 163k+ labelled bounding boxes on train images.

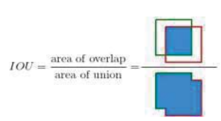
Table 4. Hyper parameters for FasterRCNN and YoLov5 on GWHD dataset [6].

Hyper Parameters	FasterRCNN	YoLov5
Backbone	ResNet-50	BottleNeckCSP
Batch size	4	8
Learning rate	0.0001	0.01
Momentum	0.90	0.98
Decay	0.0005	0.001
Optimizer	Adam	Adam
IoU Threshold	0.5	0.7

Performance Evaluation Metrics:



Confusion Metric [16]:
True Positives (TP): when the true value is Positive and prediction is also Positive.
True Negatives (TN): when the true value is Negative and prediction is also Negative.
False Positives (FP): when the true value is negative but prediction is Positive.
False Negatives (FN): when the true value is Positive but the prediction is Negative.



Intersection over Union (IoU):
IoU is a metric that evaluates the overlap between the ground-truth mask and the predicted mask. IoU is used in object detection to determine if a given detection is valid or not.
Note: IoU metric ranges from 0 and 1 with 0 signifying no overlap and 1 implying a perfect overlap between ground truth and predicted [17].

Fig. 7. IoU formula [17]

$$IoU(A,B) = \frac{A \cap B}{A \cup B}$$

- Precision: Precision (Pre) is a metric that designates the proportion of true positive outcomes. It is formulated as follows: $\text{Pre} = \frac{TP}{TP+FP}$
- Recall: Recall (Re) is a metric that designates the proportion of true positives that were successfully detected. It is formulated as follows: $\text{Re} = \frac{TP}{TP+FN}$
- F1-Score: The F1-Score is calculated as the harmonic mean of precision and recall and is calculated as follows: $\text{F1-Score} = 2\left(\frac{Pre*Re}{Pre+Re}\right)$

mAP :
Mean Average Precision (mAP) is a metric used to evaluate object detection models such as FasterRCNN, YoLo, Mask-R CNN etc.

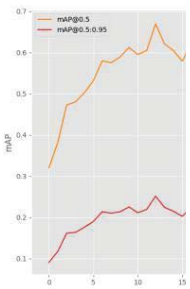
mAP metric is based on 4 sub matrices: Confusion matrix, Intersection over Union (IoU), Precision, and Recall.

$$\mathrm{mAP} = \frac{1}{No.\ of\ Classes}\left(\frac{TP}{TP+FP}\right)$$

FasterRCNN:

The Pytorch Lightning API was used for implementing the FasterRCNN model for detection. Pytorch Lightning is a more structured API to save and load model progress, and it also helps to make machine learning more scalable so one can build more AI models efficiently and quickly [18].

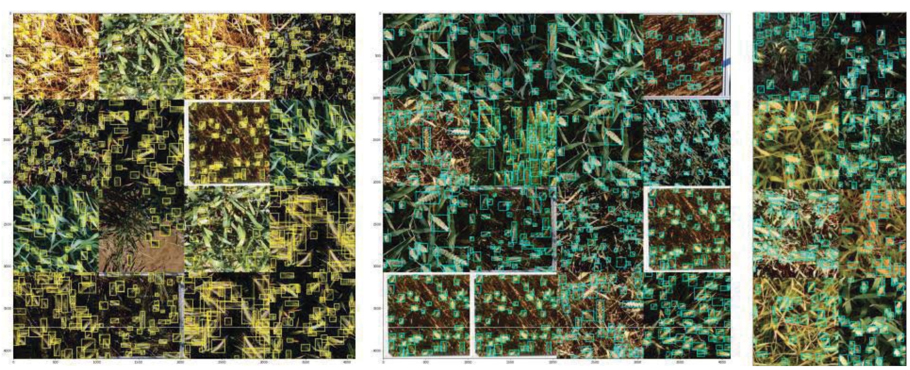
Table 5. Results of FasterRCNN implementation with specific no. of epochs.



Epochs	No. of wheat heads	mAP (%)
1	79,169	36.3
5	89,056	47.9
10	76,055	68.4
15	87,183	59.1

Fig.8. mAP performance of FasterRCNN

The following images are the results of wheat head detection using FasterRCNN (Fig. 9). Table 5 represents the total number of wheat heads in test images, where implementation generates a file with the predicted bounding box string and the total predicted count of every image.



(a) (b) (c)

Fig. 9. Some predicted wheat heads images using FasterRCNN on a test dataset [6]. (a). Wheat head prediction with 5 epochs, and (b). Prediction with 10 epochs. (c). Prediction with 15 epochs.

YoLov5:

To set up the core implementation and environment for YoLov5, use the YoLov5 repository [19, 20]. The wheat head detection model YoLov5 followed a workflow diagram (Fig. 2). Following Table 6 are the results of k-cross validation. During wheat head detection using YoLov5, a total of 4 folds are used: fold0, fold1, fold2, and fold3. Table 6 represents the evaluation metrics during the cross-validation technique applied to the training dataset.

Table 6. Numeric results of data splitting technique for YoLov5

Results	fold0	fold1	fold2	fold3
F1-score	0.64 at 0.514	0.63 at 0.807	0.69 at 0.502	0.64 at 0.494
Precision	0.819	0.956	0.830	0.796
Recall	0.96	0.95	0.95	0.96
Precision/Recall	0.679	0.658	0.669	0.660

Dataset [6] splits into 4 subparts and stores labels for further use as an example of pseudo labelling.

(a) (b) (c) (d)

Fig. 10. Image visualization after k-cross validation is applied in YoLov5. (a). Generating bounding box debugger for fold0, (b). Mosaic images which are generated through training batches in fold0, (c). Bounding box debugger for fold3, (d). Mosaic images through training batches in fold3.

As per the above discussion, semi-supervised learning techniques produce more accurate outcomes. Here, pseudo labelling is used for semi-supervised learning.

Table 7. YoLov5 mAP performance while performing k-cross validation and pseudo labeling with specific no. of epochs.

K-Fold	-Cross Validation	(epochs)	mAP (%)	Pseudo labeling	(epochs)	mAP (%)
0	fold0	10	89.3	fold0_pseudo0	2	67.9
1	fold1	12	95.0	fold1_pseudo1	5	79.0
2	fold2	10	91.3	fold2_pseudo2	2	66.9
3	fold3	12	88.6	fold3_pseudo3	5	80.6

Wheat head detection implementation trains images from four different splits with four folds, where train images are processed in folds, and then stores images and labels in different folders. Those labels are used in pseudo labelling to predict wheat heads in test images. Pseudo labelling generates labels for test images, and YoLov5, the trained model, predicts the wheat heads in every single test image.

Fig. 11. Outcomes of YoLov5: Wheat heads prediction with labels on test images.

Each fold and each set of pseudo labels generate different outcomes in every test image. Integrating all the results of the four folds and increasing the number of epochs, getting more mAP, as shown in Table 7, and also generating the file of test images with predicted bounding boxes per image and the total count of wheat heads in one image.

Table 8. mAP performance and total number of wheat heads detected in test images using YoLov5.

Epochs	No. of wheat heads detected (Predicted Wheat Heads)	mAP (%)
Cross-validation : 35 Pseudo Labelling : 15	1,09,422	86.82

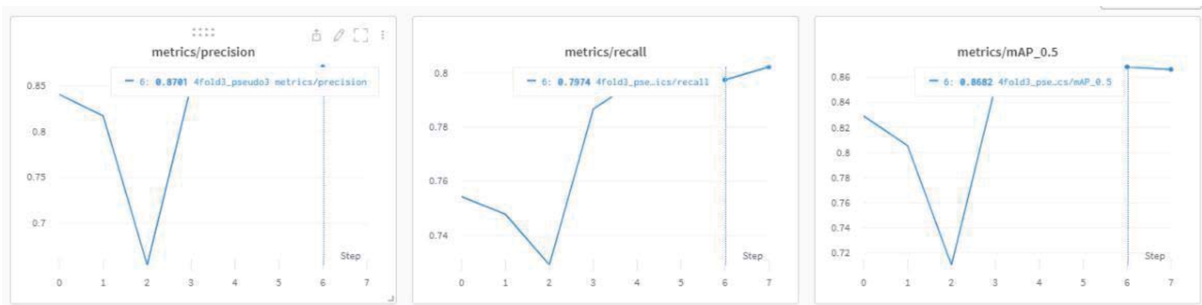


Fig. 12. Final matrics of precision, recall and mAP on test images using YoLov5 detection model.

Table 9. Model comparison on Wheat Head Detection using GWHD dataset [6].

Model Name	Data Cleaning	Data Augmentation	Pseudo Labeling	Test-Time Augmentation	Post processing	mAP (%)
FasterRCNN	True	True	-	True	Ture	68.40
YoLov5	True	True	Ture	Ture	Ture	86.82

Table 10. Some test images wheat heads count comparison with ground truth count values.

image_name	domain	Ground truth count	Predicted count FasterRCNN	Predicted count YoLov5
e38255a86b481f787da4b7ae4a227f6ded7c23b36438af45e63ed67f2dfc4fb4	15	34	35	33
08f2f41ad45dc5235b88179d6c373878c957e4ef5d46afe736c4f4d26b57793b	10	18	23	18
53895351db06d453f47c0602637fb5ad1fd32074059ce6d9645de2b3e7179064	6	63	63	64
fb5f7521da26020909c8786af72305813a468749ddfa21c3d16a98e949510ecb	5	14	13	14
debc17970bfc40679dd3e60c4ceb961614fa37a7b0ad6578ec58f726fb40ccfa	3	43	43	44
3902fc957d935d37319d24438998cdba08624e920f5a8ea58ae62b72f3b372f2	2	25	24	25
035a987f13fd2ad00e8b13f46e59556a384873ace564ea7c56a3a78536f3e2b1	1	23	23	23
2d52563250c95e48817c097e8cb5c53211d9c522d77c709637b24d2ffdaf3f02	4	48	49	49
.	.	.	.	.
.	.	.	.	.
Total Wheat Head Count for 20 images		**928**	**913**	**916**

Table 10 is the basic analysis of the headcount and comparison with true values. Table 10 contains image_name, which is a random 20 test image from the test dataset [6]. So as per the studies of both models: FasterRCNN and YoLov5, the YoLov5 detection model is more accurate for the global wheat head dataset. Also calculating the total count of heads with FasterRCNN and YoLov5, YoLov5 has a smaller difference in the ground truth count as compared to the count generated through

FasterRCNN. FasterRCNN gave 98.38 % and YoLov5 gave 98.71 % accurate counts for 20 sample images from the test dataset.

6. Conclusion

Wheat head detection is a valuable task for wheat production estimation, wheat breeders, and crop management. This study looked into a novel framework for detecting wheat heads in a GWHD dataset, which was released in 2021. This work is mostly concentrated on regularizing deep learning model techniques. To improve the precision of models, semi-supervised learning and ensemble learning techniques were also explored. To enhance training procedures, this study went through the data and analyzed the origins, data types, distribution, diversity, and specificity. Also able to tackle some limitations in the field of wheat head detection, like the unicity of the acquisition methods for the data as well as the lack of precision in the detection. This study uses two different versions of the datasets, which makes the analysis more understandable. Moreover, a comparative analysis was covered. In comparison to a smaller dataset, the study is more accurate with the larger one. In future work, willing to compare methods to further deep learning architectures for wheat head detection.

References:

1. E. David, Shouyang Madec, Pouria Sadeghi-Tehran, "Global Wheat Head Detection (GWHD) Dataset: A Large and Diverse Dataset of High-Resolution RGB-Labelled Images to Develop and Benchmark Wheat Head Detection Methods". Plant Phenomics- The Science Partner Journal, April, 2020. Dataset - https://zenodo.org/record/5092309#.Y7lcG3ZBzIU.
2. Fares Fourati, Wided Souidene Mseddi & Rabah Attia, "Wheat Head Detection using Deep, SemiSupervised and Ensemble Learning". Canadian Journal of Remote Sensing- Taylor & Francis, April, 2021.
3. Zhao, Z.Q., Zheng, P., Xu, S.T. and Wu, X., 2019. "Object detection with deep learning: A review". IEEE transactions on neural networks and learning systems, 30(11), pp.3212-3232. https://arxiv.org/pdf/1807.05511.pdf&usg=ALkJrhhpApwNJOmg83O8p2Ua76PNh 6tR8A
4. Pound, M. P., Atkinson, J. A., Wells, D. M., Pridmore, T. P., and French, A. P. 2017. "Deep learning for multitask plant phenotyping". In Proceedings of the IEEE International Conference on Computer Vision, pp. 2055–2063
5. Madec, S., Jin, X., Lu, H., De Solan, B., Liu, S., Duyme, F., and Baret, F. 2019. "Ear density estimation from high resolution RGB imagery using deep learning techniques." Agricultural and Forest Meteorology, Vol. 264: pp. 225–234. doi:10.1016/j.agrformet.2018.10.013
6. AI crowd, Global Wheat Head Challenge 2021, By University of Saskatchewan. https://www.aicrowd.com/challenges/global-wheat-challenge-2021
7. rohan007,GeeksforGeeks- "Stratified K Fold Cross Validation" https://www.geeksforgeeks.org/stratified-k-fold-cross-validation/
8. Lee, D.H. 2013. "Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks." Workshop on Challenges in Representation Learning, ICML, Vol. 3 (No. 2): p. 896.
9. Nathan Hubens, 2019. " Test-Time Augmentation and how to perform with Keras", Feb 12, 2019. https://towardsdatascience.com/test-time-augmentation-tta-and-how-to-perform-it-with-keras-4ac19b67fb4d
10. Breiman, L. 1996. "Bagging predictors." Machine Learning, Vol. 24 (No. 2): pp. 123–140. doi:10.1007/BF00058655.
11. Yuseok Ban and Kyungjae Lee, 2020 "Multi-Scale Ensemble Learning for Thermal Image Enhancement". Appl. Sci. 2021, 11(6), 2810; https://doi.org/10.3390/app11062810
12. Sambasiva Rao. K, Medium Article, "Weighted Boxes Fusion — A detailed view", Feb, 23, 2021. https://medium.com/analytics-vidhya/weighted-boxes-fusion-86fad2c6be16
13. Abhi Lad, Mehul S. Raval, 2022. "Improving Wheat Head Detection: A Data-Centric Approach by Domain Variance Reduction", https://dl.acm.org/doi/10.1145/3477314.3507190.
14. Movva Nitin Datta1, Yash Rathi, M. Eliazer, SRM Institute of Science and Technology, Kattankulathur, 2021. "Wheat Heads Detection using Deep Learning Algorithms". Annals of R.S.C.B., ISSN: 1583-6258, Vol. 25, Issue 5, 2021, Pages. 5641 - 5654.
15. Kaggle - Global Wheat Detection challenge 2020. https://www.kaggle.com/competitions/global-wheat-detection
16. Anuganti Suresh, 2020. Medium Article, "What is a confusion matrix?", Nov. 17, 2020. https://medium.com/analytics-vidhya/what-is-a-confusion-matrix-d1c0f8feda5
17. Kiprono Elijah Koech, 2020. Medium Article, "Confusion Matrix for Object Detection", Aug 1, 2020. https://towardsdatascience.com/confusion-matrix-and-object-detection-f0cbcb634157#:~:text=IoU%2C%20aso20called%20Jaccard%20 index,detection%20is%20valid%20or%20not.
18. SABREPC.COM, 2022."Deep Learning and AI : Why You Should Use PyTorch Lightning and How to Get Started", March 3, 2022. https://www.sabrepc.com/blog/Deep-Learning-and-AI/why-use-pytorch-lightning
19. ULTRALYTICS - YOLOV5, Nov 22, 2022. Github repository. https://github.com/ultralytics/yolov5
20. KSNXR-YOLOV5, June 16, 2022. Github repository. https://github.com/ksnxr/GWC_YOLOv5
21. TAN, M., and LE, Q. V. 2019. "EfficientNet: Rethinking model scaling for convolutional neural networks." arXiv preprint arXiv:1905.11946

22. TAN, M., PANG, R., and QUOC, V. L. 2020. “EfficientDet: Scalable and efficient object detection.” In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. doi:10.1109/CVPR42600.2020.01079.
23. OPENGENUS-IQ, “Understanding ResNet50 architecture”, 2022. https://iq.opengenus.org/resnet50-architecture/

Investigation Of SWIPT NOMA Over Rayleigh Fading

Dr.Sumit Gupta[a*], Sneha Maram[2], Laxmi Priya Nakka[b], Himaja Veeragoni[b], Hrushikeah.panduga[b], Arpita baronia[c]

Dept of ECE Assistant professor , SR University WARANGAL, INDIA[a]
Dept.of ECE,SR Engineering College,Warangal,India[b]
Dept of CSE Assistant professor, VIT,Bhopal[c]
Sumitmanit25@gmail.com

Abstract

Non-orthogonal multiple access (NOMA), which may be utilized to increase spectral efficiency in fifth-generation (5G) networks, has recently generated a lot of attention. In this part, we will examine how the concepts of simultaneous wireless power transfer (SWIPT) and non-orthogonal multiple access (NOMA) can be coupled to create a unique wireless energy harvesting multiple access protocol that is both energy- and spectrum-efficient. The signal is amplified for the NOMA users farther away who are experiencing worse channel conditions by using the NOMA users closer to the source as relays. The chapter considers how SWIPT might be used in NOMA, particularly when it's done by people who are close to the access points. We propose three opportunistic user selection methods for NOMA implementation: nearest near user and farthest far user (NNFF) selection, closest near user and nearest far user (NNNF), and random near user and random far user (RNRF).

Keywords: fifth-generation network, NOMA, SWIPT, Energy Harvesting.

Nomenclature

T	Transmit Power
a_n	Fraction of power allocated to near user
a_f	Fraction of power allocated to far user
x_n	Signal intended for near user
x_f	Signal intended for far user
h_{sn}	Rayleigh fading coefficient between the BS and near user with zero mean and variance $= d_{sn}^{-\eta}$
d_{sn}	Distance between BS and near user
η	Path loss exponent
w_n	AWGN With zero mean and variance $= \propto^2$

1. Introduction

There has been a considerable increase in smart devices, new technology, and contemporary applications, which has led to the wireless communication industry seeing tremendous growth. The maximum budget for transmit power, minimum data rate, and minimal gathered energy per terminal restrictions [19]. The performance of a system is measured using two different metrics: energy efficiency (EE) and spectral efficiency (SE) [1]. In order to ensure its viability and meet its needs, the future generation of 6G will unavoidably require the efficient use of energy [2]. The relationship between bandwidth efficiency (BE) and energy efficiency (EE) is also investigated [6]. The availability of bandwidth (BW) is used to calculated by SE. The near NOMA users operate as energy harvesting relays for far NOMA users in the co-operative SWIPT NOMA protocol [4]. A timeline of the various efforts made to bring 6G into existence. major technology advances to meet connection objectives within 6G [7]. Green

* Dr.Sumit Gupta
E-mail address: sumitmanit25@gmail.com

NOMA is used to pinpoint energy-efficient NOMA plans. In a perfect world, the energy needed for communication is sustainably handled [18]. The availability of bandwidth (BW) is used to calculated by SE. The simultaneous wireless information and power transmission (SWIPT) with cooperative hybrid-NOMA (H-NOMA) in terahertz (THz) is an improved two-user pairing technique that considers SIC imperfection [12]. Future generations will be significantly impacted by terahertz (THz) communication. The problematic shortfalls that THz communications face because of THz features are thought to be overcome by cooperative simultaneous wireless information and power transmission (THz-NOMA) [9]. Users with better channel conditions obtain less power from NOMA, and by using successive interference cancellation, these users can decode their own information [5]. The unknown electromagnetic (EM) gap (0.1–10 THz) between EM and optical bands is represented by THz. A new cooperation strategy allows a two-user cooperative non-orthogonal multiple access network to directly connect with nearby users [16]. Electronic, photonic, and plasmonic technologies are anticipated to be used in the production of THz transmitters and receivers. By comparing the procedures used at the transmitter, the state-of-the-art NOMA systems are evaluated [17]. Investigations on the NOMA approach for 5G wireless communication have improved channel capacity while taking user fairness restrictions into account.

Because of needs for the Internet of Everything (IoE), fully connected systems are now necessary. 5G systems are constrained, making it unable to add features or improve existing ones to meet these requirements to improve the overall performance of 6G wireless communications and reliability, as well as wireless connectivity, resource management, scalability, and user fairness. Using MIMO-NOMA technology and THz frequencies, we modified the existing wireless communication systems while evaluating the effectiveness and gains made. Modern non-orthogonal multiple access (NOMA) variations that use power and code domains as the foundation for interference reduction, resource allocation, and quality of service management in the 5G scenario [15].

The main contribution is a stable non-LOS (NLOS) line of THz communications to make up for the absence of THz coverage in wide areas or any other place where IRS cannot be deployed. A base station (Bs) of the cooperative system connects directly with a user who is close by, while a relay node (RN) assists the Bs in communicating with the user who is far away [11]. By comparing related recent research on performance analysis of cooperative PD-NOMA systems, we may examine current trends in PD-NOMA based cooperative networks [8]. By using THz, NOMA, MIMO, cooperative networking, and EH (SWIPT) methods, as well as enhancing SE, EE, and other metrics in compared to the state of the art, it is possible to create a system with low power consumption, complexity, and cost [20].

1.1. RF Energy Harvesting

In the academic literature, several methods for energy harvesting have been investigated [13,14]. Among them, radio frequency energy harvesting (RFEH) stands out as a popular technique for converting energy from the electromagnetic field into electrical energy (i.e., voltage and current). RFEH has great potential for powering low-power sensors and systems in body area networks because it enables wireless power transfer in various application scenarios [3]. Nevertheless, harvesting energy from RF sources presents a challenging task for designers and researchers since it requires expertise from both the electromagnetic field and electronic circuit domains. Therefore, designing a high-performance RF energy harvester necessitates knowledge from both domains. There are two types of energy harvesting:

1.1.1. Energy Harvesting by Time Switching:

The system operates in a time-slot based manner. During the first portion of the time slot, the device collects electromagnetic energy from its environment. The succeeding segment of the time slot is utilized to transmit the harvested energy. In subsequent articles, we will delve into time-switching further. This write-up will showcase the application of the cooperative NOMA power-splitting technique.

1.1.2. Energy Harvesting by Power Splitting:

The device divides the signal power received to generate energy and decode data. The simultaneous implementation of energy harvesting and information decoding is possible with power splitting as opposed to time switching, which necessitates different time slots for both activities. SWIPT, or Simultaneous Wireless Information and Power Transfer, is another name for this technique.

2. Background

The most well-known candidate for creating 6G systems is hence nonorthogonal multiple access in the power domain (NOMA). Multi-input, multi-output (MIMO)-NOMA. The NOMA method improves channel capacity primarily defined by BW by superimposing numerous signals at the transmitter (Tx) and filtering them at the receiver (Rx) using successive interference cancellation (SIC) methods. By giving various power coefficients to users based on their channel conditions, NOMA multiplexing is accomplished. Line of sight (LOS) is the main transmission path, and clustering or grouping the serviced users is essential for improving SE and improving NOMA-based THz transmission. Implementing energy harvesting (EH) with cooperative NOMA takes use of cooperative networking's many advantages, including improved reliability and capacity as well as a greater coverage area.

3. Related Work

The major advantages of cooperative networking have already been studied using a range of cases and circumstances, and the disadvantages of using this method have also been addressed. The shortcomings, limitations, and poor performance of using those technologies were all carefully examined. The intelligent reflecting surface (IRS) must outperform the use of relaying in terms of complexity, energy, and cost in order to improve source-to-destination transmission in particular circumstances of interrupted communication paths, as demonstrated in Figure. The overall conclusion of this investigation claimed that a superior data rate is required to exceed the decode-forward (DF) relay regarding transmit power minimization and energy efficiency (EE) maximization when comparing IRS to the earlier DF relay. Additionally, IRS needs additional hardware.

Matching transmission beams for time-based steering, interference control, and EE degradation are some of the challenges in building and installing IRS to enable 6G communication networks. when deploying IRS, the effects of the actual coverage area, the hardware, the system components, and added maintenance/cost for wireless communications. Failures in aggregate or proportion are brought on by atmospheric factors.

3.1. Cooperative SWIPT NOMA Network

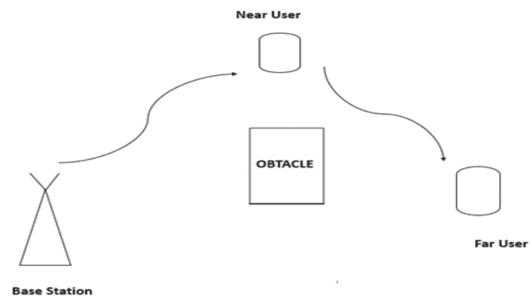


Fig. 1. Network Model of Cooperative SWIPT NOMA

In this scenario, we consider a downlink transmission model, where the base station (BS) utilizes NOMA to transmit messages simultaneously to both nearby and distant users. Unfortunately, an obstacle between the BS and the far user results in significant shadowing, rendering the signal undetectable to the distant user.

Conversely, the nearby user experiences clear communication with the BS. According to NOMA principles, the near user decodes the data intended for the far user first, and then utilizes successive interference cancellation (SIC) to decode its own data. The nearby user copies the far user's information, serving as a decode-and-forward relay to aid the distant user.

However, the challenge arises from the nearby user's lack of sufficient power to transmit data to the far user. To address this issue, the nearby user employs a power-splitting method of energy harvesting, specifically SWIPT, to accumulate additional power.

The entire communication process can be divided into two time slots. During the first time slot, the BS transmits data to the nearby user, who utilizes power-splitting to harvest a portion of the received signal power and utilizes the remainder for information decoding. In the second time slot, the nearby user employs the collected energy to transmit data to the far user.

3.2. Signal model of SWIPT NOMA

3.2.1. Time Slot-1

In the first time slot, the BS transmits the NOMA signal, which is provided by

$$x=\sqrt{T}\ \sqrt{a_n}x_n+\sqrt{a_f}x_f) \qquad (1)$$

The remote user is unable to receive this signal because of severe shadowing. The signal that the near user receives is provided by,

$$y_n=\sqrt{T}\ \sqrt{a_n}x_n+\sqrt{a_f}x_f)h_{sn}+w_n \qquad (2)$$

The nearby user extracts a little amount of power from y_n, Let's use to represent this fraction ψ. This is known as the energy harvesting coefficient. The amount of power left over is $(1-\psi)$ the portion that can be used to decode information.

Therefore, the signal that can be used for information decoding after energy harvesting is,

$$y_D=(\sqrt{1-\psi})y_n+w_{eh}=(\sqrt{1-\psi})\sqrt{T}(\sqrt{a_n}x_n+\sqrt{a_f}x_f)+(\sqrt{1-\psi})w_n+w_{eh} \qquad (3)$$

Where w_{eh} (with a zero mean and a variance of σ^2) is the thermal noise produced by the energy harvesting electronics. For the sake of simplicity, let's assume that the energy obtained via w_n is insignificant, which results in the following expression for y_D.

$$y_D=(\sqrt{1-\psi})\sqrt{T}(\sqrt{a_n}x_n+\sqrt{a_f}x_f)+w_{eh}$$

The nearby user initially carries out direct decoding of x_f from y_D. The feasible rate for the nearby user to decode data from the distant user is provided by,

$$R_{nf} = \frac{1}{2}\log_2\left(1+\frac{T(1-\psi)a_f\,|h_{sn}|^2}{T(1-\psi)a_n\,|h_{sn}|^2+\sigma^2}\right) \qquad (5)$$

The rate at which a nearby user can successfully decode its own information after SIC is,

$$R_n = \frac{1}{2}\log_2\left(1+\frac{T(1-\psi)a_n\,|h_{sn}|^2}{\sigma^2}\right) \qquad (6)$$

ψ is the fraction of power harvested in the time slot 1, the amount of power harvested is given by,

$$T_H = T\,|h_{sn}|^2\,\zeta\psi \qquad (7)$$

ζ is given by power harvesting efficiency of the circuitry.

3.2.2. Time slot-2

Using the power harvesting in the prior slot, the near user transmits the data intended for the far user p_H. The signal sent by the nearby user is, $\sqrt{T_H}\,\overline{x}_f$. The signal at the far user is,

$$y_f = \sqrt{T_H}\,\overline{x}_f h_{nf} + w_f \qquad (8)$$

h_{nf}, Rayleigh fading channel coefficient between near user and far user. The possible rate for a far user is,

$$R_f = \frac{1}{2}\log_2\left(1+\frac{T_H\,|h_{sn}|^2}{\sigma^2}\right) \qquad (9)$$

3.3. Optimize the power splitting coefficient

Now we will develop an expression in this part to get the ideal value of ψ .In the first time slot, the near user must successfully decode the data from the far user. The proper information can then be relayed in the subsequent time period. Let's create a constraint to verify this condition.

$$R_{nf} > R_f^* \qquad (10)$$

R_f^* is far user target data rate. The achievable rate at the near user must be higher than the far user's target rate in order to decode the data.

Let's substitute the expression of R_{nf} in the above condition and solve for ψ

$$\frac{1}{2}\log_2\left(1+\frac{T(1-\psi)\propto_f|h_{sn}|^2}{T(1-\psi)\propto_n|h_{sn}|^2+\sigma^2}\right) > R_f^* \qquad (11)$$

$$\log_2\left(1+\frac{T(1-\psi)a_f\,|h_{sn}|^2}{T(1-\psi)a_n\,|h_{sn}|^2+\sigma^2}\right) > 2R_f^* \qquad (12)$$

$$1+\frac{T(1-\psi)a_f\,|h_{sn}|^2}{T(1-\psi)a_n\,|h_{sn}|^2+\sigma^2} > 2^{2R_f^*} \qquad (13)$$

$$\frac{T(1-\psi)a_f\,|h_{sn}|^2}{T(1-\psi)a_n\,|h_{sn}|^2+\sigma^2} > 2^{2R_f^*}-1 \qquad (14)$$

Let's denote $2^{2R_f^*}-1$ by P_f .This is the target SINR for the far user.

$$\frac{T(1-\psi)a_f\,|h_{sn}|^2}{T(1-\psi)a_n\,|h_{sn}|^2+\sigma^2} > P_f \qquad (15)$$

$$T(1-\psi)a_f\,|h_{sn}|^2 > P_f T(1-\psi)\propto_n |h_{sn}|^2 + P_f\sigma^2 \qquad (16)$$

$$T(1-\psi)a_f\,|h_{sn}|^2 - P_f T(1-\psi)a_n\,|h_{sn}|^2 > P_f\sigma^2 \qquad (17)$$

$$T(1-\psi)\,|h_{sn}|^2\,(a_f - P_f a_n) > P_f\sigma^2 \qquad (18)$$

$$(1-\psi) > \frac{P_f\sigma^2}{T\,|h_{sn}|^2\,(a_f - P_f a_n)} \qquad (19)$$

$$\psi < 1 - \frac{P_f\sigma^2}{T\,|h_{sn}|^2\,(a_f - P_f a_n)} \qquad (20)$$

To ensure that ψ is less than the value given by RHS, let's modify the above equation as,

$$\psi = 1 - \frac{P_f\sigma^2}{T\,|h_{sn}|^2\,(a_f - P_f a_n)} - \delta \qquad (21)$$

δ is a very small number. This value of ψ ensures that enough power is available for information decoding in order to meet the far user's target rate.

4. Result and Analysis

As show in figure 1, While the far user's rate rises, the near user's rate reaches saturation at about 1 bps/Hz. This saturation is achieved by energy harvesting. All of the remaining power is gathered once the near user has reached the distant user's data rate. The energy harvesting function limits the possible rate at the nearby user even if the transmit power is raised. The benefit of this situation is that the target rate of 1 bps/Hz for the near user is not exceeded. As a result, the nearby user does not experience many outages due to this rate saturation.

The amount of power gathered rises as transmit power rises. As a result, the data sent by far users in the second slot is transmitted with more power. As a result, the achievable rate for far users rises. The user who is far away has a greater average attainable rate, so he must have the lowest chance of an outage as mention in Figure 3.

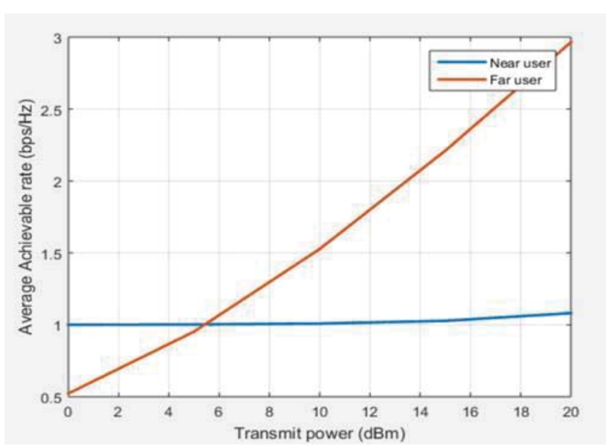


Fig. 2. Average achievable rates vs transmit power

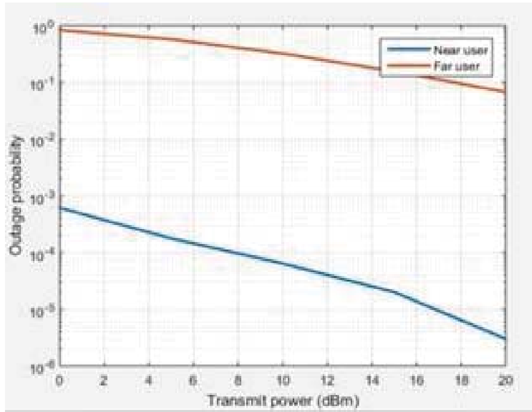


Fig. 3. outage performance of cooperative SWIPT NOMA

Despite having a higher attainable rate on average than the near user, the far user endures far more outages.

The far user's instantaneous achievable rate frequently falls below the goal rate value. How much higher the feasible rate is doesn't matter for outage calculations. We only keep track of instances in which the instantaneous rate deviates from the intended rate. For the remote user, this fall occurs more frequently as given in figure.4.

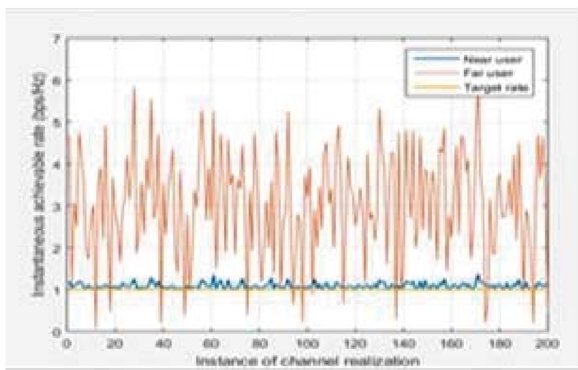


Fig. 4. instantaneous rates

5. Conclusion

The paper is concluded that, the far user experiences higher average achievable rate, but the no of counts that falls down the target rate explains the least outage performance. Although the near user has lesser achievable rate, but it is greater than the target rate most of the time. This explains better outage performance at near user.

References

1. S. Han, T. Xie and C. -L. I, "Greener Physical Layer Technologies for 6G Mobile Communications," in IEEE Communications Magazine, vol. 59, no. 4, pp. 68-74, April 2021, doi: 10.1109/MCOM.001.2000484
2. G. Gür, "Expansive networks: Exploiting spectrum sharing for capacity boost and 6G vision," in Journal of Communications and Networks, vol. 22, no. 6, pp. 444-454, Dec. 2020, doi: 10.23919/JCN.2020.000037.
3. Introduction to RF Energy Harvesting Serdijn W.A., Mansano A.L.R., Stoopman M.(2014) *Wearable Sensors: Fundamentals, Implementation and Applications*, pp. 299-322.

4. Y. Liu, Z. Ding, M. Elkashlan and H. V. Poor, "Cooperative Non-orthogonal Multiple Access With Simultaneous Wireless Information and Power Transfer," in IEEE Journal on Selected Areas in Communications, vol. 34, no. 4, pp. 938-953, April 2016, doi: 10.1109/JSAC.2016.2549378.
5. Z. Ding, M. Peng and H. V. Poor, "Cooperative Non-Orthogonal Multiple Access in 5G Systems," in IEEE Communications Letters, vol. 19, no. 8, pp. 1462-1465, Aug. 2015, doi: 10.1109/LCOMM.2015.2441064.
6. Shankar R. Examination of a non-orthogonal multiple access scheme for next generation wireless networks. The Journal of Defense Modeling and Simulation. 2022;19(3):453-465. doi:10.1177/1548512920951277
7. I. F. Akyildiz, A. Kak and S. Nie, "6G and Beyond: The Future of Wireless Communications Systems," in IEEE Access, vol. 8, pp. 133995-134030, 2020, doi: 10.1109/ACCESS.2020.3010896.
8. Liaqat, M., Noordin, K.A., Abdul Latef, T. *et al.* Power-domain non orthogonal multiple access (PD-NOMA) in cooperative networks: an overview. *Wireless Netw* **26**, 181–203 (2020). https://doi.org/10.1007/s11276-018-1807-z
9. Oleiwi, H.W.; Al-Raweshidy, H. Cooperative SWIPT THz-NOMA/6G Performance Analysis. *Electronics* **2022**, *11*, 873. https://doi.org/10.3390/electronics1106087
10. Y. Liu, W. Yi, Z. Ding, X. Liu, O. A. Dobre and N. Al-Dhahir, "Developing NOMA to Next Generation Multiple Access: Future Vision and Research Opportunities," in IEEE Wireless Communications, vol. 29, no. 6, pp. 120-127, December 2022, doi: 10.1109/MWC.007.2100553.
11. O. Elkharbotly, E. Maher, A. El-Mahdy and F. Dressler, "Optimal Power Allocation in Cooperative MIMO-NOMA with FD/HD Relaying in THz Communications," 2020 9th IFIP International Conference on Performance Evaluation and Modeling in Wireless Networks (PEMWN), Berlin, Germany, 2020, pp. 1-6, doi: 10.23919/PEMWN50727.2020.9293074.
12. H. W. Oleiwi, N. Saeed and H. S. Al-Raweshidy, "A Cooperative SWIPT-Hybrid-NOMA Pairing Scheme considering SIC imperfection for THz Communications," 2022 4th Global Power, Energy and Communication Conference (GPECOM), Nevsehir, Turkey, 2022, pp. 638-643, doi: 10.1109/GPECOM55404.2022.9815677.
13. X. Li, J. Li and L. Li, "Performance Analysis of Impaired SWIPT NOMA Relaying Networks Over Imperfect Weibull Channels," in IEEE Systems Journal, vol. 14, no. 1, pp. 669-672, March 2020, doi: 10.1109/JSYST.2019.2919654.
14. B. Makki, K. Chitti, A. Behravan and M. -S. Alouini, "A Survey of NOMA: Current Status and Open Research Challenges," in IEEE Open Journal of the Communications Society, vol. 1, pp. 179-189, 2020, doi: 10.1109/OJCOMS.2020.2969899.
15. I. Budhiraja et al., "A Systematic Review on NOMA Variants for 5G and Beyond," in IEEE Access, vol. 9, pp. 85573-85644, 2021, doi: 10.1109/ACCESS.2021.3081601.
16. G. Li and D. Mishra, "Cooperative NOMA Networks: User Cooperation or Relay Cooperation?," ICC 2020 - 2020 IEEE International Conference on Communications (ICC), Dublin, Ireland, 2020, pp. 1-6, doi: 10.1109/ICC40277.2020.9148973.
17. S. K. Zaidi, S. F. Hasan and X. Gui, "SWIPT-aided uplink in hybrid non-orthogonal multiple access," 2018 IEEE Wireless Communications and Networking Conference (WCNC), Barcelona, Spain, 2018, pp. 1-6, doi: 10.1109/WCNC.2018.8376963.
18. Y. Ye, Y. Li, D. Wang and G. Lu, "Power splitting protocol design for the cooperative NOMA with SWIPT," 2017 IEEE International Conference on Communications (ICC), Paris, France, 2017, pp. 1-5, doi: 10.1109/ICC.2017.7996751.
19. J. Tang et al., "Energy Efficiency Optimization for NOMA With SWIPT," in IEEE Journal of Selected Topics in Signal Processing, vol. 13, no. 3, pp. 452-466, June 2019, doi: 10.1109/JSTSP.2019.2898114.

20. S. Zargari, H. Ahmadinejad, B. Abolhassani and A. Falahati, "SWIPT-NOMA in Cell-Free Massive MIMO," 2020 28th Iranian Conference on Electrical Engineering (ICEE), Tabriz, Iran, 2020, pp. 1-6, doi: 10.1109/ICEE50131.2020.9260930.

Acknowledgement

We thank SR Engineering College, Warangal, Telangana, and SR University, Warangal, Telangana for supporting us during this work.

Intelligent Traffic Signal Control in Smart Cities using Deep Learning Model

Mansi Subhedar[a*], Uttaran Roychowdhury[a]

[a]*Department of Electronics and Telecommunication Engineering, Pillai HOC College of Engineering and Technology, Rasayani Tal. Khalapur, Dist Raigad, Maharashtra, India (mansissubhedar@gmail.com)*

Abstract

In many developing nations, massive infrastructure projects like flyovers or underpasses are a big investment and most take durations of up to multiple number years to complete. This work provides an AI-based traffic control system for intersections at such locations. The proposed system is viable for modern smart cities where traffic intersections exist with the need for adaptive switching as a future-proof solution when the cities get in more vehicle load. This work can be built on an already existing surveillance system of close circuit cameras used to monitor traffic and keep an eye on traffic rules violations. In recent years, deep learning solutions have shown new avenues of automation and it is still in the growing phase of its life cycle. These neural network systems are flexible and can be molded to do a variety of tasks. The proposed work contributes to the design of three neural networks that work in synergy to achieve the task of fluid automated traffic signal switching. Specifically, a Convolutional Neural Network (CNN) in Inception V2, a Long Short Term Memory Neural Network (LSTM) for forecasting, and a Deep Neural Network (DNN) for the classification of switching decisions are developed. In this work, adaptive traffic signal switching is treated as a classification problem and is handled in a sequential manner utilizing the specific strengths of different neural networks that are easy to maintain in the future. The cost over time for implementing such systems will get lower as technology progresses and gets easily accessible. This perspective will minimize traffic congestion and aid in reducing traffic rule violations. Congestion reduction in turn also translates into less wait/travel time and less pollution. As fuel costs increases, we need to get more efficient in every aspect of transportation and can be achieved with the proposed work.

Keywords: Smart Cities, Deep learning, DNN, Inception, LSTM, Traffic light Control

1. Introduction

Traffic management is a very important issue all over the world and many developed and developing countries are looking for a solution to the ever-rising number of road vehicles. Many countries concentrated on going towards public transport and sacrificing the luxury and convenience of the private vehicle sector. The other solution was to upgrade the infrastructure by building flyovers and expanding the width of the roads. All these solutions have come a long way to improve the experience of private travel on road, but the traffic signals have been the same for a long time and only brought to the digital era by the integration of microprocessors. But signals in concept were not smart or adaptive in nature. A need for adaptive traffic signal control arose from the fact that there are variations in the performance of the previously implemented systems. These variations were mostly in a day or soon after the systems were up and operating. The deterioration happened as soon as the conditions deviated significantly from the base conditions [1]. The proposed system aims to bring this to the information era where data can be used to optimize everything and anything. This work aims to modernize the system by making it adaptive to the changing traffic conditions depending on the time of the day and as well as the live input. The process goes through 3 steps of neural networks, the first section of the model is the CNN network; an Inception V2 network [2] used for the detection and counting of vehicles in the visible region of the road. The second section is the DNN model used for the classification of the type of switching that must be done at the intersection. The third section is the forecasting done by the acquired data from the very first step (CNN network) and predicts the daily traffic by the minute. All these different deep learning systems work in synergy and help us achieve the desired goal. This proposed system does not provide traffic signal synchronization rather it is a method for the switch in a single direction of the road. This whole system can be implemented for the N-number of roads connecting to the intersection with an appropriately designed synchronization tool.

The paper is organized as follows: Section 2 deals with related works. Section 3 presents the proposed work. Simulation results are presented in Section 4. Section 5 concludes the work.

* Dr Mansi Subhedar
E-mail address: mansissubhedar@gmail.com

2. Related Works

Traffic forecasting and prediction have a long research history, and many have suggested novel implementations using machine learning and deep learning techniques. Many systems have been suggested for smart traffic management in many varying platforms such as traffic lights, smart vehicles, etc. NCHRP says that there have been many adaptive traffic solutions developed by transportation research institutes and agencies all over the world but less than half of them have been deployed and tested in real-world scenarios [3]. Wang et al. discuss multiple levels of transport intelligence methods and ranks them into 5 levels starting from one that is Model-based, MAXBAND, etc. methods, and the final level is the Model-free learning adaptive control systems. One of those systems that improved upon the timer-based systems is the Split Cycle Offset Optimization Technique also known as SCOOT [4]. It is a level 2 intelligent decision-making system. SCOOT was developed by Transport Research Laboratory for the Department of Transport in 1980; it uses data from the motion sensor to adjust traffic signal delays. SCOOT is most widely used in the United Kingdom today. Another, level 2 intelligence system is Sydney Coordinated Adaptive Traffic System (SCATS) [5]. It is a system that manages the traffic signal in a dynamic manner using road traffic sensors and pedestrian switches to optimize cycle times, phase splits and offsets. It also incorporates a priority traffic facility within it and considers buses and trams as well. The system is owned by Australia and is in operation in Sydney. The SCAT system is installed in more than 42,000 intersections across 40 countries. Another work uses USB based camera systems as image input, and certain image recognition and object counting techniques are employed to determine the amount of traffic and calculate traffic density [6]. They have also used an RFID reader and tag system to transmit the information required to begin capturing the images. This system mainly depends on the machine learning technique of Support Vector Machine (SVM) for object counting which works better than the existing timer-based systems. Arif et al suggested a motor-mounted camera system for a similar purpose as in [6] but they rely on image processing techniques to determine the region of interest and assess the amount of time to assign to the traffic light timers. The proposed system involves three different techniques together and hence some of the technique's related study on an individual basis is explored. Another prominent approach that many research studies on this topic was the implementation of deep reinforcement learning [8].

The field of deep learning and neural networks shows a lot of promise when it comes to image recognition-based systems. Several image recognition approaches can be found; one of them being the localization with a selective search algorithm that feeds into a DCNN for recognition of the proposed region and uses a linear SVM for classification [9]. The performance of their system is verified using multiple criteria based on road conditions such as free-flowing traffic, congested traffic, and stop-and-go traffic in different weather conditions along with camera resolution changes. It was found that the system performance is better at the daytime with free-flowing traffic conditions. This was especially important for consideration with time of day and the resolution of the camera inputs. Over the time, the camera resolution will improve but the impact of weather conditions on the input video stream is something to be taken into consideration.

The other part of the proposed system is the LSTM forecasting method; there have been various attempts at predicting traffic patterns. A detailed study of the road traffic data using different parameters and techniques is conducted in [9] which determines that the two fixed points in the data are jam and free-flow, the abrupt changes prove the non-linear nature of the traffic. The data used for the study is from North-Rhine Westphalian motorways. ARMA and ARIMA are two methods widely used for forecasting [10]. But these are linear autoregressive models and cannot predict the stochastic and non-linear nature of the traffic flow [11]. The LSTM & GRU (Gated recurrent unit) fair so much better than those due to the ability to memorize long-term dependencies. If the LSTM models and GRU models are compared, the GRU units are much simpler to implement and more efficient for computation. This is due to the fact that the method used for mitigating vanishing gradients is done by only two gates being a reset gate and an update gate for each cell, meanwhile LSTM has three gates being input gate, output gate and the forget gate. A proper comparison study has been conducted in [12]. Traffic time-series studies are being conducted for a long time. A technique presented for prediction using ARIMA is discussed in [13]. A traffic control optimization technique is proposed by Li et al. using multi-agent deep reinforcement learning by improving the coordination between traffic signals [14]. Egea et al proposed an urban traffic control mechanism using reinforcement learning with three different architectures and compared it with existing commercial systems [15]. Li et ad proposed various models for adaptive traffic control that rely on data collected only from connected vehicles [16]. Several similar studies are discussed in [17]-[18].

3. Materials & Methodology

The details of the setup used to train and evaluate the proposed model are provided here. The deep neural network and the Long Short-Term Memory Neural Network (LSTM) were built on Keras API with TensorFlow backend. The whole experiment is done on a standard laptop PC running windows 10 OS with Nvidia GeForce GTX 1650 card with 4 GB video memory and an Intel core i5-9300H CPU, CUDA v11 and TensorFlow 2.3. The programming environment used is anaconda with python v3.6. The Object detection and counting API is running on TensorFlow 1.X settings with TensorFlow 2.3. Specific changes disabled.

a. Dataset

The time series prediction model data on traffic density and particularly in this kind of implementation was sparse, even the available data was not suitable for the needs of this work. Hence, custom synthetic data was developed for training and testing purpose and would serve until the real-world data is created by the system to run inference over it again by using real-world data.

The custom dataset is used for both the decision making deep neural network and sequence model in the traffic forecasting neural network. The data ingestion pipeline is handled by pandas and made into dataframe from .csv file. The three compare columns in the dataset are made by random number generator that gives numbers at random a at range of -1, 0 and 1. The values here are associated to a certain condition as follows; first '-1' is when the compared roads count is less than the main road on which the operation is taking place on, '0' is for condition when both are equal and finally '1' is for when the compared road has more vehicles than the current one. Another column is a comparison between the car counts at the video against the predicted car count at that instant from the forecasting LSTM model. The last column is the decision. It is provided by considering the four comparison columns. For real world implementation, the csv files are not ideal so use of SQL databases is recommended as the data would be too massive to contain in standard csv file formats that has limited scalability. The implementation of SQL based database would also make the reading of the input data much quicker as it will not be fragmented into separate files as the amount of data increases. For this project the dataset contained around 80,000 data points.

For experimental setup, here, csv file is considered.

Table 1: First 5 entries in the input data

Time-stamp	Car_count	compare_left	compare_opp	compare_right	Train_Prediction	count_and_pred	min	hour	day_of_week	day_of_month	month	Decision
23-09-2020 00:00	2	0	0	-1	0	0	0	0	2	23	9	0
23-09-2020 00:01	2	0	0	-1	2.21491695	1	1	0	2	23	9	0
23-09-2020 00:02	2	0	-1	-1	2.21491695	1	2	0	2	23	9	1
23-09-2020 00:03	3	-1	-1	-1	2.21491695	0	3	0	2	23	9	5
23-09-2020 00:04	3	0	-1	-1	3.87389994	1	4	0	2	23	9	1

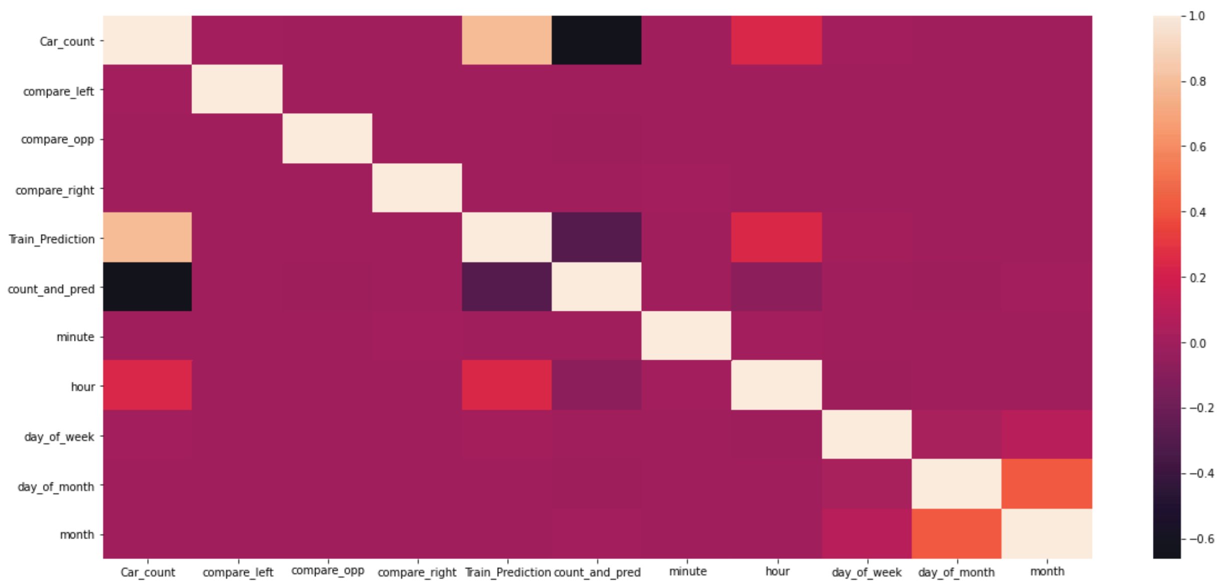


Fig. 1. Correlation Heat Map

1.1 Preprocessing and augmentation

The entries in the dataset are preprocessed, cleaned for null values, and checked for errors by using pandas and NumPy library in python. The video input for this experiment is considered as a video downloaded from YouTube as a sample called "4K camera example for Traffic Monitoring (Road)". It is used in an .mp4 format. But live input footage can also be employed and obtained from an IP based close circuit camera by using OpenCV library if required during deployment [19].

1.2 Models and Overall Working

This is a hybrid three-model system where the input is a video stream that provides the number of vehicles in the frame. Those numbers are saved in a .csv file along with the current time stamp. Some of the sample values are shown in Table 1. The dataset for all values is maintained in a similar way. A correlation heat map is depicted in Figure 1. This CSV file is used as input for the sequence model or Long Short-Term Memory Neural Network. (LSTM) model is used for forecasting the monthly traffic for that specific intersection. The predicted values are again written into the same .csv file. LSTM model output is a regression-based output so the accuracy is not the metric here. Rather, the mean squared error that measures how close the prediction is to the actual value is employed. Finally, the Deep Neural Network is used to classify the desired decision. The decision is the right-most column in the dataset and contains the labels for the classification. In this work, the decision is considered a classification problem because the decisions are finite in nature and can be selected from these predefined classes. There are six classes labeled as right, right & forward, only forward, left and forward, left and no change. The output given by the deep neural network is a number that corresponds to one of those six decisions. The overall working model is shown in Figure 2.

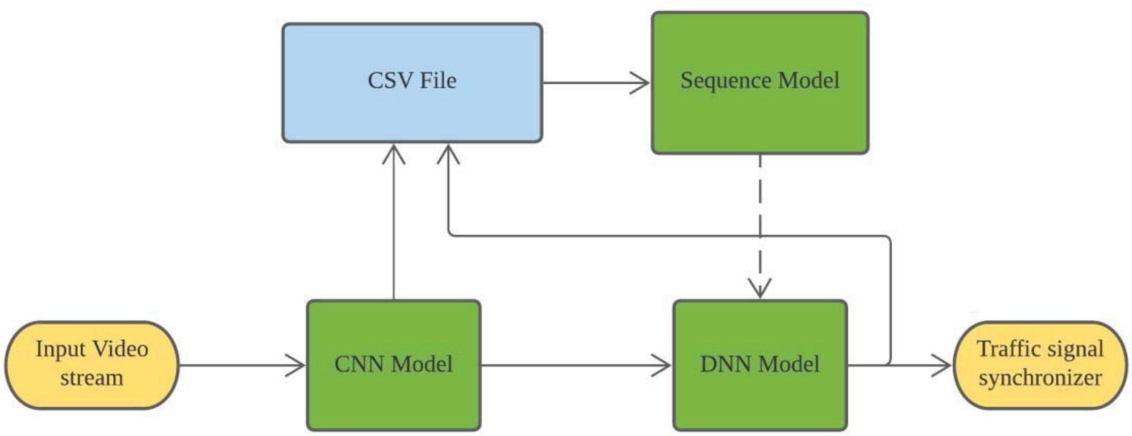


Fig. 2. The proposed overall working model

1.3Object counting

Object counting is done by the API with pre-trained convolutional neural network called Inception-V2 trained on MS coco [20] but ImageNet [21] can also be used for a custom dataset. Inception V2 is an object recognition model and is used to produce bound boxes around the object of interest. This pre-trained neural network can classify 91 classes ranging from cars and buses to cats and dogs, etc. It is used to recognize the number of vehicles in the current frame. Vehicles include cars, buses, trucks, and motorcycles; the sum of a number of all these objects in the frame is used to determine the number of vehicles. The API does the counting by giving the number of bound boxes on the frame as shown in Figure 3. The object recognition rate and localization are not very high the inception-V2 model here.

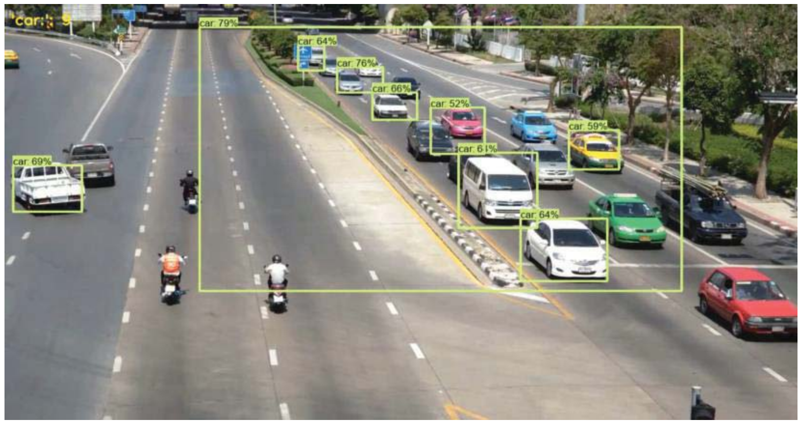


Fig. 3. **Object counting for one frame**

The count accuracy of the API depends on the model's accuracy. For experimentation purpose, the accuracy is traded for faster runtimes and lower memory consumption. But more deeper and advanced models like RetinaNet-50 or CenterNet [22] would be more suitable for deployment purposes.

1.4 Long Short-Term Memory Neural Network (LSTM)

The Long Short-Term Memory neural network or LSTM is
a sequence model and determines the repeating nature of the data. Traffic like many things such as weather and earthquakes are a predictable pattern-based phenomenon. This pattern can be predicted and forecasted beforehand to be prepared for
what's to come in near future. Traffic at certain places follows this repeated pattern monthly or annually depending on the place.

```
_________________________________________________
Layer (type)           Output Shape       Param #
=================================================
lstm (LSTM)            (None, 1, 32)      4352
_________________________________________________
lstm_1 (LSTM)          (None, 1, 64)      24832
_________________________________________________
lstm_2 (LSTM)          (None, 32)         12416
_________________________________________________
dense (Dense)          (None, 32)         1056
_________________________________________________
dropout (Dropout)      (None, 32)         0
_________________________________________________
dense_1 (Dense)        (None, 1)          33
=================================================
Total params: 42,689
Trainable params: 42,689
Non-trainable params: 0
_________________________________________________
```

Fig. 4. Proposed LSTM architecture

The proposed model as shown in Figure 4 and 5 consists of three layers of LSTM and a dropout layer with 50% tolerance and the last layer is a Dense layer with one neuron as output. LSTM is not a classifier and hence accuracy metric is not followed for measuring the performance. Mean squared error as the measure for performance is employed here as well as for calculating the loss function. The proposed model uses Adam optimizer for training. The generated forecasts are from 23rd September 2020 to 23rd November 2020 with every minute taken as an entry and one time step at a time. The converted raw input data representation is shown in Figure 4. Another test was conducted with a GRU model with 64-neuron first layer and a 32-neuron second layer and 2 dense layers with eight and a single neuron respectively with a recurrent dropout at the second layer. This resulted in similar results but was unstable during subsequent reruns of the model. Hence, LSTM model was chosen. But GRU model execution times were faster compared to the LSTM model. Figure 6 shows input data entries as sample. Several trials were taken to set the values and accordingly fine tuning was done while choosing the LSTM values. The predictive model was used to generate the expected traffic load which is compared with the current traffic load which aided in better decision making during switching the lights.

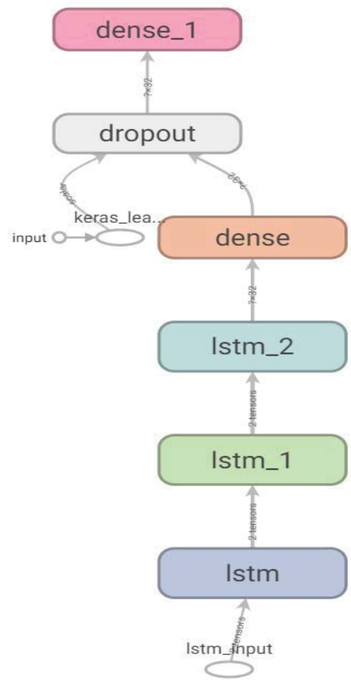


Fig. 5. Proposed LSTM model architecture

1.4 Deep Neural Network (DNN)

The deep neural network is the final part of in the predictive models. This model takes input from the object recognition counter and the long short-term memory neural network models to classify the type of decision that suits the current conditions. The proposed model architecture is in Figure 7 and Figure 8.

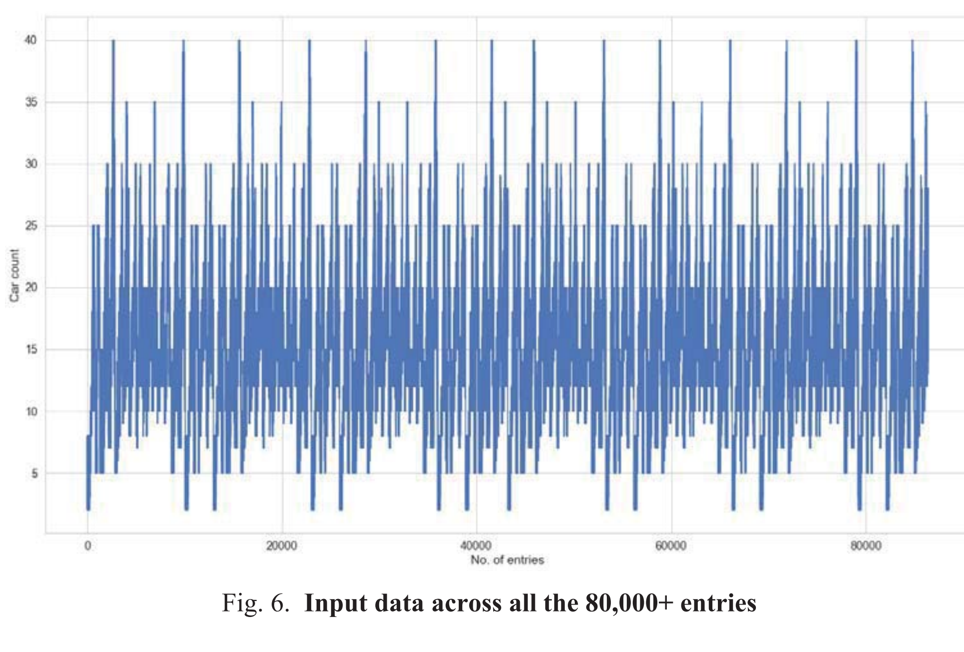


Fig. 6. **Input data across all the 80,000+ entries**

```
_______________________
Layer (type)              Output Shape       Param #
================================================
dense_1 (Dense)          (None, 256)        3072
_______________________
dense_2 (Dense)          (None, 256)        65792
_______________________
dropout_1 (Dropout)      (None, 256)        0
_______________________
dense_3 (Dense)          (None, 128)        32896
_______________________
dense_4 (Dense)          (None, 128)        16512
_______________________
dropout_2 (Dropout)      (None, 128)        0
_______________________
dense_5 (Dense)          (None, 64)         8256
_______________________
dense_6 (Dense)          (None, 64)         4160
_______________________
dropout_3 (Dropout)      (None, 64)         0
_______________________
dense_7 (Dense)          (None, 6)          390
================================================
Total params: 131,078
Trainable params: 131,078
Non-trainable params: 0
_______________________
```

Fig. 7. Proposed Deep Neural Network

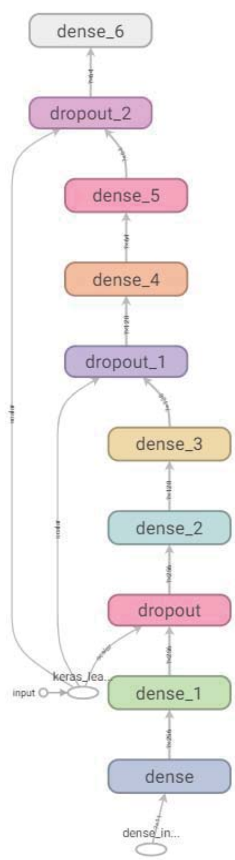


Fig. 8. Proposed DNN Model Architecture

The input is similar as shown in Table 1. This proposed model consists of 7 Dense Layers which are in pairs of 2 and one Dropout layer in-between each pair with 50% tolerance and the last dense layer has 6 neurons as this is a multiclass classification and each neuron out of the 6 is associated to the decision labels. This is done so that the system could be simplified to a decision, and each side can give independent outputs to the synchronizer. The DNN training is done in batch size of 32 in 30 epochs with a learning rate scheduler. The learning rate scheduler is used to reduce the learning rate by a certain amount after every 10 epochs, this helps the loss function find the proper desired local minima. The DNN is using labels created with respect to the data and what would be most optimal switching operation to be performed in such a condition to minimize the wait times of vehicles.

3.Results and Discussion

To evaluate the results of this experiment, the accuracy and loss function values of each of the models can be employed. The object recognition model is pretrained and as per Google AI blog has a Top-1 accuracy of 80.4% and a Top-5 accuracy of 95.3%. Considering the model is more than four years old the Top-1 accuracy is quite low for real world application where human safety should be utmost priority. But as discussed earlier, newer and deeper models can solve this problem. Also, a proper camera frame setting with only one side of road view is needed for good results as the vehicles that have already left the lights after it became green need not be counted. The Long Short-Term Memory Neural Network (LSTM) neural network is a regression type deep learning model; hence mean square error, mean absolute error and root mean square error metrics are used to determine the performance of the model.

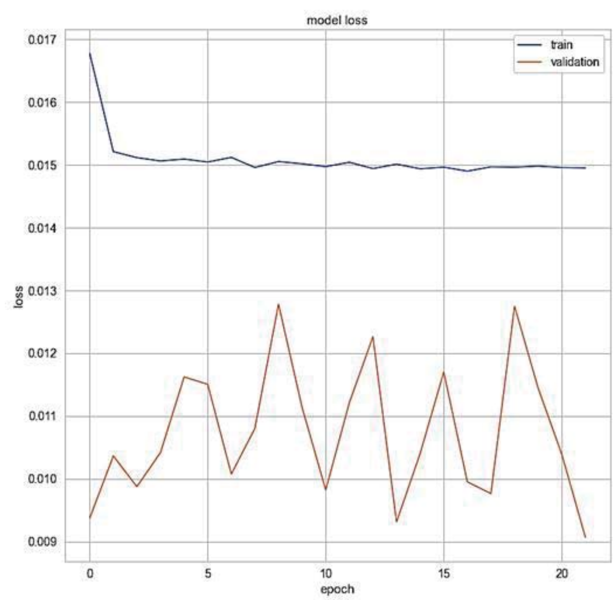


Fig. 9. Loss function for the LSTM

The mean square error is also the loss function in this case so one of the graphs is equal to the other and it is intended as such. The mean squared error stays around 0.0138 and 0.0137 for training and for validation it finishes at around 0.011 and 0.010 the Figure 9 and 10 show results for final model used for prediction. The mean absolute error is around 0.0915 and 0.0920 for training and 0.076 for validation. The LSTM model has a Root mean square error score of 4.37 in training and 4.32 in test set. The outcomes of the LSTM's forecasting were in acceptable range and not much farther from the original values, but the model had difficulty in predicting the higher end of the values or the maximum values in the traffic counted.

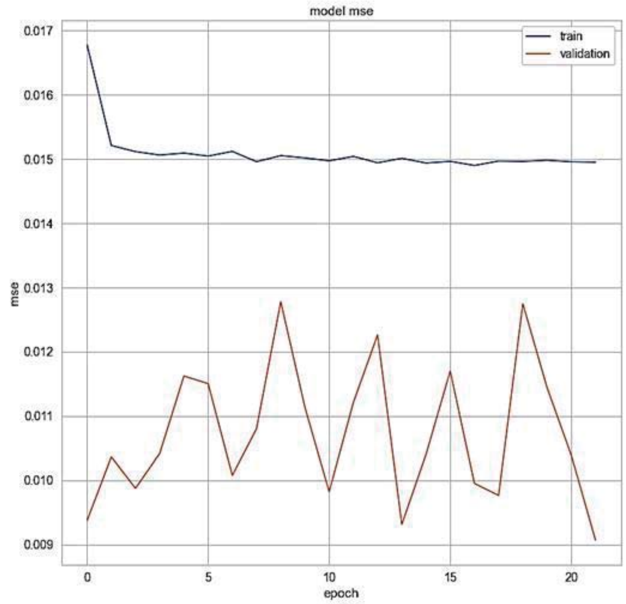


Fig. 10. Mean squared error for the LSTM

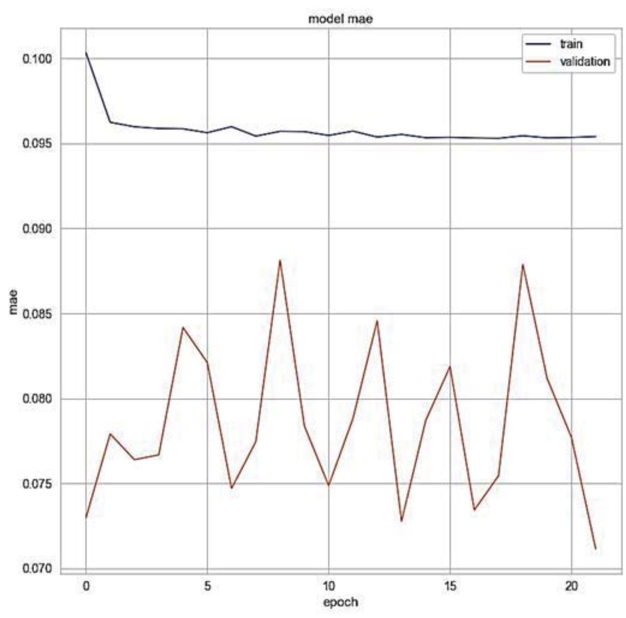


Fig. 11. Mean absolute errors for the LSTM

But it did show a rise in the predicted values which could be considered as the higher end during operations. The graph of predicted values against the original input values is displayed in Figure 12. The deep neural network's performance can be easily evaluated by measuring the loss and accuracy of the trained model. The Figure 12 shows the loss function graph and Figure 13 shows the accuracy metric where the highest training accuracy achieved is of 99.95% and highest validation accuracy of 100%. These metrics show that this system with multiple models can perform the switching task very reliably. The performance result could not be compared due to the fact that the dataset is a synthetic one and is assuming some specific kind of scenarios. The synthetic dataset was used because the data that we came across wasn't suitable or lacking specific traffic count criteria in multiple directions which was crucial for the comparison and decision-making process.

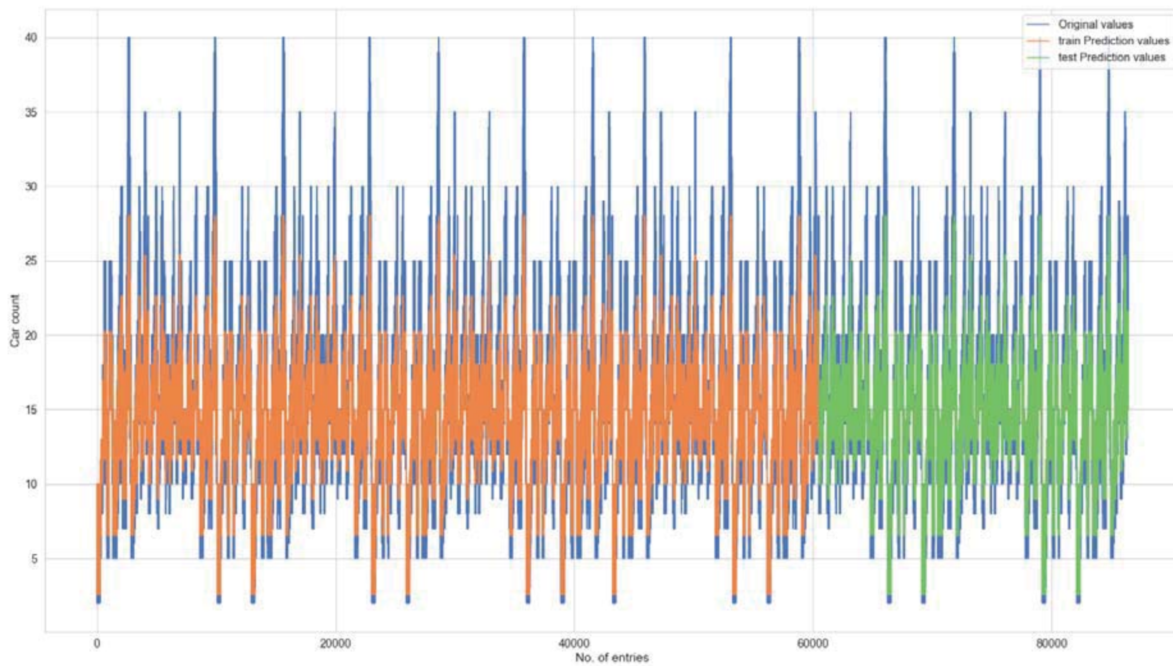


Fig. 12. Predicted against original input values in the LSTM network

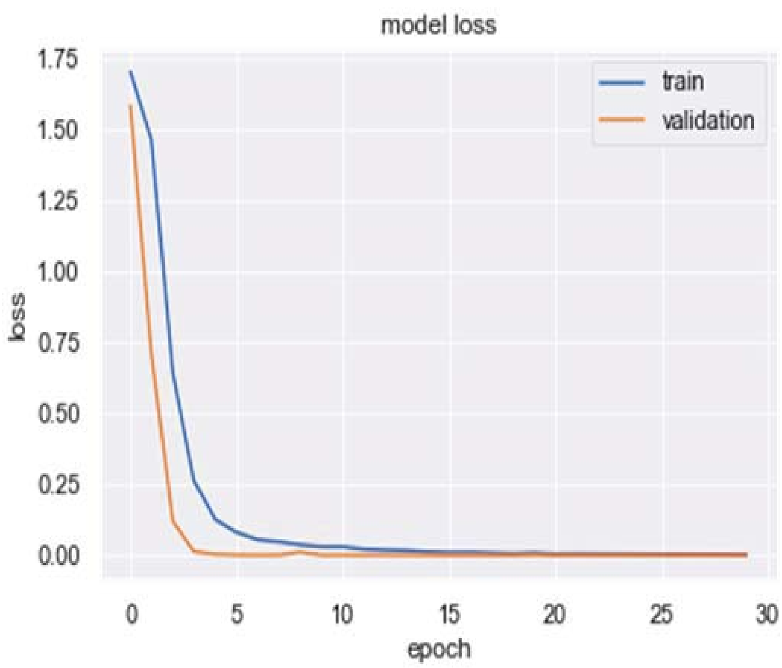


Fig. 13. Loss graph for DNN

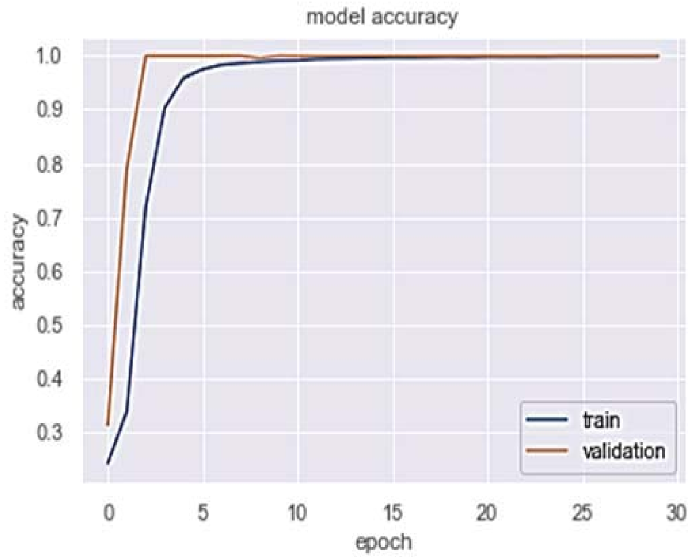


Fig. 14. Accuracy for DNN

4. Conclusion

Proposed System provides a new smarter traffic signal controller for the traffic lights was the objective of proposed work. The traffic signal switch incorporates a triple model setup which can be deployed at intersections where the traffic management systems are struggling to keep up with the rapidly growing population around and at a much lower investment cost than that of full-scale overhauls like constructing flyovers or underpasses. This method can become the primary foundation to expand upon and connect the whole road network to different artificial intelligence-based systems and create a future with self-driving vehicles or semi-self-driving vehicles that can predict and pre-plan routes as required. The Adaptive nature of the system helps it a lot and reduces the maintenance issues which might arise on a hardcoded timer-based system. This design can be further improved in regards of traffic forecasting and counting as more resources are available to run faster and bigger computations. Also, the scalability of deep learning models is very helpful when it comes to operations such as this where there are massive amounts of input data to learn from and store, this will allow easier upgradations to bigger datasets. Also helps in keeping the model up to date for future changes and fluctuations that might have been missed in the current sample set. There can be proper implementations and integration of this system within the Smart city sector and as well as in modern trends like Internet-of-things (IoT) by connecting other peripheral components to the vehicles and the traffic signal system which can communicate with each other with a better more fluid data sharing system. The aim and vision for proposed work to design a system for the current road systems which will mesh modern deep learning techniques for enhancing road traffic systems for a better more modern future. The models performed well and showed good enough results to be considered for deployment. This work is proposed for smart city roads while keeping in mind the existing infrastructure that will need an upgrade. This work is intended to be applied where a high to moderate amount of traffic could be expected.

5. Future work

The current system can have better performance in regards to detection accuracy and precision if better models such as CenterNet [21] can be implemented as mentioned before. The LSTM can also be improved upon by implementing Minmax scaled attention into it, which will improve the predictions at the higher end of the predicted values and make the model overall more reliable. Along with the previously mentioned improvements, more optimized patterns in switching and synchronizing the traffic lights can be explored. A few additional minor hyperparameter changes can also be explored in order to achieve improved efficiency. An improved version of data pipeline can be done by migrating to the SQL database or Hadoop ecosystem. It has to be done because in real-world implementation a lot of data will be generated and hence need for a robust data storage solution is felt.

References

1. Stevanovic A, Kergaye C, Stevanovic J., "Long Term Benefits of Adaptive Traffic Control under Varying Traffic Flows during Weekday Peak Hours", *Transportation Research Record*; 2311 (1): 99-107, 2012
2. Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jonathon Shlens and Zbigniew Wojna, "Rethinking the Inception Architecture for Computer Vision", arXiv:1512.00567, 2015
3. Yizhe Wang, Xiaoguang Yang, Hailun Liang, Yangdong Liu, "A Review of the Self-Adaptive Traffic Signal Control System Based on Future Traffic Environment", *Journal of Advanced Transportation*, vol. 2018, Article ID 1096123, 12 pages, 2018 https://doi.org/10.1155/2018/1096123
4. P. B. Hunt, D. I. Robertson, R. D. Bretherton, and M. C. Royle, "e SCOOT online traffic signal optimization technique," *Traffic Engineering and Control*, vol. 23, no. 4, pp. 190–192, 1982.
5. A. G. Sims and K. W. Dobinson, "e Sydney coordinated adaptive traffic (SCAT) system philosophy and benefits," *IEEE Transactions on Vehicular Technology*, vol. 29, no. 2, pp. 130–137, 1980
6. Sanchal Ramteke, and B. B. Gite, "AI BASED TRAFFIC SIGNAL CONTROL SYSTEM, *International Research Journal of Engineering and Technology (IRJET),* Volume: 06 Issue: 11, Nov 2019
7. Arif A. Bookseller, Rupali R Jagtap, "Image processing based Adaptive Traffic Control System", *Second International Conference on Emerging Trends in Engineering (SICETE), IOSR Journal of Electronics and Communication Engineering*, pp. 33-37, 2018
8. Duowei Li, Jianping Wu, Ming Xu, Ziheng Wang, Kezhen Hu, "Adaptive Traffic Signal Control Model on Intersections Based on Deep Reinforcement Learning", *Journal of Advanced Transportation, vol. 2020, Article ID 6505893*, 14 pages, 2020. https://doi.org/10.1155/2020/6505893
9. Yaw Okyere Adu Gyamfi, Sampson Kwasi Asare, Anuj Sharma, and Tienaah Titus, "Automated Vehicle Recognition with Deep Convolutional Neural Networks, Transportation Research Record" *Journal of the Transportation Research Board*, No. 2645, pp. 113–122, 2017
10. H. Siegel and D. Belomestnyi, "Stochasticity of Road Traffic Dynamics: Comprehensive Linear and Nonlinear Time Series Analysis on High-Resolution Freeway Traffic Records", Data Analysis, Statistics and Probability, 2006
11. Williams B M, Hoel L A, "Modeling and forecasting vehicular traffic flow as a seasonal ARIMA process: theoretical basis and empirical results," *Journal of Transportation Engineering*, vol. 129, no. 6, pp. 664- 672, 2003.
12. R. Fu, Z. Zhang and L. Li, "Using LSTM and GRU neural network methods for traffic flow prediction," *31st Youth Academic Annual Conference of Chinese Association of Automation (YAC), Wuhan, China*, pp. 324-328, 2016, doi: 10.1109/YAC.2016.7804912
13. Ahmed, M. S., and A. R. Cooke, "Analysis of Freeway Traffic Time-Series Data by Using Box-Jenkins Techniques", Transportation Research Record 772, TRB, National Research Council, Washington, D.C., pp. 1-9, 1979
14. Lia, Hao Yub,Guohui Zhangc, Shangjia Dongd, Chengzhong Xu, "Network-wide Traffic Signal Control Optimization Using a Multi-agent Deep Reinforcement Learning", Transportation Research Part C: Emerging Technologies, Volume 125, April 2021
15. Alvaro Cabrejas Egea, Raymond Zhang, Neil Walton: "Reinforcement Learning for Traffic Signal Control: Comparison with Commercial Systems." CoRR abs/2104.10455, 2021
16. Wangzhi Li, Yaxing Cai, Ujwal Dinesha, Yongjie Fu, Xuan Di, "CVLight: Deep Reinforcement Learning for Adaptive Traffic Signal Control with Connected Vehicles" arXiv:2104.10340
17. James Ault, Guni Sharon, "Reinforcement Learning Benchmarks for Traffic Signal Control", *35th Conference on Neural Information Processing Systems (NeurIPS 2021) Track on Datasets and Benchmarks*, June 2021
18. Bouktif, S.; Cheniki, A.; Ouni, A. "Traffic Signal Control Using Hybrid Action Space Deep Reinforcement Learning." Sensors 2021.
https://doi.org/10.3390/ s21072
19. Bradski, G, "The OpenCV Library", *Journal of Software Tools for the Professional Programmer*, 2000.
20. Lin TY. Et al. Microsoft COCO: Common Objects in Context. In: Fleet D., Pajdla T., Schiele B., Tuytelaars T. (eds) "Computer Vision ECCV 2014." Lecture Notes in Computer Science, vol 8693. Springer, Cham. ,2014 https://doi.org/10.1007/978-3-319-10602-1_48
21. O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, "Imagenet large scale visual recognition challenge.", 2014.
22. Kaiwen Duan and Song Bai and Lingxi Xie and Hong ang Qi and Qingming Huang and Qi Tian, "CenterNet: Keypoint Triplets for Object Detection" preprint 1904.08189, 2019.

IoT Based Transformer Monitoring System

Siddharth Bhorghe[a], Prashik Tambe[b*], Tejas Patil[c], Shravani Raut[d], Prathamesh Phulari[e]

Department of Electronics and Telecommunication (E&TC),
Vishwakarma Institute of Technology, Pune, 411037, Maharashtra, India

Abstract

The internet of things is about connecting ideas that were never connected before. These ideas can be used to improve the lives of everyone through the means of IoT. If operated with proper care, transformers can be used for a longer duration and have a longer life span. However, the life of transformers is reduced due to the excess amount of load over them and the continuous work done by them. If the transformers are not checked frequently, a few failures usually occur inside them. Overloading, increasing oil levels and improper cooling of transformers, etc. are the major reasons behind the failure of transformers. With the use of the Internet of Things (IoT), the frequent checking of the conditions of the transformer can be automated, which further notifies the owner and the transformer could be rescued from upcoming failure.

Keywords — IoT, NodeMCU; Arduino Nano; Sensors; Wi-Fi Module; LCD; Transformer.

1. Introduction

In India, the rate of failure of transformers is 12% to 17%. Every year about 200 Crore Indian Rupees are spent to repair these transformers [1]. Distribution transformers till now were observed manually where a person would regularly visit the transformer site to check the condition and record the essential parameters. This process was unable to provide data about sudden overload on the transformer and the overheating of oil in the transformer. The life span of the transformer is affected by these variables. Transformer measurement systems were expected to only detect one single parameter, which could be current, voltage, or any other parameter. While there were a few ways that helped to recognize multi-parameters, but those were much more time-consuming processes. The available monitoring system was unable to provide the collected data to the observing centers in the required time, which made it really difficult to save the transformer from the upcoming failure. All these points have been taken care of with the help of this monitoring system. Now it's easier to maintain the records of the condition of the transformers without any manual checking. This system checks the condition of the transformer and saves the related data, which helps to save a lot of time and reduce the work of the particular person. It lets the observing centers know about the condition of the transformer before it could get worse.

2. Literature review

The respective research paper provides data analyzed about the reasons and rates of transformers being damaged. The study is made on the transformers in a particular city. That includes the various reasons that potentially damage the transformers. A comparison between damage rates in India and abroad has also been made in the paper [1].

The below-cited paper gives information about the relative saturation value and is used to alert the moisture level in transformer oil. The calculation is made to calculate the moisture level by the least square method and to protect transformer integrity the use of a threshold alarm is important. This model detects failure due to an increase in water content and the moisture attained in transformer oil and accordingly assesses its convenience [2].

In the related study, a model was proposed to identify problems or failures in the transformer that increase the water content of transformer oil to rise was put out. This paper consists of the detailed working of the model. Sensors were used to detect temperature and moisture in the transformer. The use of various data tables is done in the research paper. Various graphs and diagrams are used so that the research done is easily understandable and accessible to everyone [3].

* Prashik Tambe
E-mail address: prashik.tambe21@vit.edu

In the respective paper, a hot-spot temperature calculation model is presented. The hotspot to ambient gradient is used as the base for the model. The model considers how winding losses and oil viscosity fluctuate with changes in temperature. The results are compared with measured temperatures at different loads for the following transformer units: 250-MVA (Mega Volt Ampere) ONAF (Oil Natural Air Forced), 400-MVA ONAF, and 605-MVA OFAF (Oil Forced/Air Forced), as well as temperatures computed using the IEEE Annex G method [4].

This paper is totally dedicated to only one parameter of the transformer which can cause failure in the system, and that parameter is oil moisture. It's critical to keep an eye on the amount of moisture in the transformer oil. The solubility of transformer oil changes according to different conditions. Transformer oils have reportedly been shown to dissolve differently depending on the environment. This report presents measurements of solubility for four different conditioned oil: lab-aged, fresh Shell Diala AX oil Transformer oil was used by Texas Utility, Ramapo Substation, and Shell Diala A oil. This work suggests an alternate approach to detecting the moisture solubility in transformer oil with the use of a relative humidity sensor, which helps to avoid the challenge of reaching full saturation. To evaluate solubility indirectly, it makes use of the linear relationship between the relative humidity of the oil and the moisture content of the same [5].

This paper explores the neural network of Artificial Intelligence where they discussed about the hotness of transformer winding and the top oil temperature of a transformer. They have presented the result for the prediction of Top-Oil temperature (TOT) using four different models by ambient temperature and load as inputs and TOT signal as output [6].

The respective paper is entirely written on the study of polarization and depolarization current measurement. The dielectric characteristics of the oil or paper insulation system of a power transformer are majorly influenced by moisture and aging. As oil/paper moisture equilibrium is temperature sensitive and takes a long time to reach its equilibrium, moisture testing in oil samples typically provides conflicting findings. For in-service transformers, direct moisture testing of a paper sample is not practical. One potential method of identifying a transformer's insulation status is the measurement and assessment of the "dielectric response" and conductivity. The state of oil/paper insulation was evaluated using polarisation and depolarization current measurements in a recent research effort [7].

In this paper, a comprehensive illustration of the transformer to assets and the management activities are presented. The three major activities of the transformer are the applications of CM techniques, maintenance plans, and less-cost maintenance methods. To monitor and access the condition of the transformer various CM (Common Monitoring) and CA (Common Assessment) techniques have been discussed in the paper. Different lifetime types and end-of-life criteria of the transformer are illustrated here [8].

This paper introduces the fault Prognosis based on the cognitive system. It combines not only evolutionary but also connectionist mechanisms into a hybrid system. This system is implemented on generic architecture for modeling cognitive systems. Knowledge about the age of the power transformer and its history is not essential. The methodology relies on DGA(Dissolved Gas Analysis) and neural networks to recognize the condition of the transformer [9].

This paper has proposed a cost-effective and powerful system that improves the reliability of the transformer. This system provides the required data in real-time and helps to find out the problems before any failure. The system uses sensors to collect data about potential damage factors to the transformer and in the meantime, they are monitored using web-based wireless applications [10].

In power networks, distributed transformers are essential hardware. Data collection and condition monitoring are important issues in power electric systems because of the large number of transformers that are dispersed across a wide area. Industrial environments are important for both machines and people, despite security and automation in facilities. The main subject of this study is safety in industrial conditions. Based on faulty data source files from other sources, a model has been developed to detect various conditions like breakdown, which is the most important factor for the occurrence of leakage current in substations, and gas leakage. The final result will be put into practice in an IoT gateway architecture to add scalability to the technological implementation. The use of a mosquito server, a node-red server, and a node MCU that acts as a mediator has been made to create the system [11].

The following research paper presents and tells us more about the work which has been done to control and monitor all the devices of the CIT (Computer and Information technology) IoT lab. The use of a Dashboard is done, which was created in either Node-Red or also in Android Studio Mobile Application. This work has created a smart laboratory system for the CIT campus that will keep the track of all the lab's activities, including energy use, device usage, and environmental parameters. This helps to give CIT a comfortable and energy-efficient smart environment. ESP8266, Arduino UNO, relays, Raspberry Pi 3, current transformers, and sensors are used in the construction of IoT smart hardware kits [12].

If there is any damage to the transformer's inherent characteristics will produce enormous drawbacks. Therefore, it is essential to frequently monitor the transformer. The primary goal of this project is to obtain live remote monitoring of transformer health using IoT technology. The system keeps an eye on parameters such as voltage, current, and temperature, which are transformer parameters. These data will be transmitted using the MQTT (MQ Telemetry Transport) protocol over the internet. In the event of any power outage, the user will receive an alert, a message sent with the help of the GSM Module. It also has the special feature of recognizing the phase as another parameter of the transformer [13].

By employing IoT to monitor the health and condition of electrical equipment, it may be possible to replace it before it breaks down and thus maintain power continuity. This paper demonstrates an application of this idea that uses IoT to remotely monitor the distribution transformer's real-time condition. The respective health monitoring system uses a temperature sensor, a potential transformer, and a current transformer to monitor the distribution transformer's temperature, voltage, and current. This information is deliveres to a remote server where it can be monitored, and any necessary action can be taken to prevent a power outage. [14].

After examining the transformer on regular basis, they studied the live monitoring and fault recognition of the transformer by inspecting the parameters to solve and prevent the big budget repair interruption for a trouble-free power supply. They have also used flame and smoke sensors to alert and send notifications to monitor and prevent high commercial loss [15].

The document shows the implementation of a mobile embedded system for measuring the required parameters which damage the transformer. If the system details have any abnormality, those are automatically updated on the internet via serial communication. It shows that the protection scheme works with accuracy and properly and the sensitivity is high for faulty and abnormal conditions [16].

In this paper, the explanation and use of every component used for monitoring of transformer are mentioned in a proper manner. The circuit diagram is shown with the help of simulating software where it is easy to understand the connections used within the components. It includes pictures where the data is monitored, and program coding of the project helps to clear the compiling errors [17].

The paper has given importance to the proposed system as it has been found that the network devices are very sensitive to voltage fluctuations, to accomplish they have worked on designing voltage and frequency sensors and have implemented, tested, and deployed them in real-life applications. This system monitors the UPSS (Uninterrupted Power Supply System) input and output voltage together with the utility or grid frequency. With the power factor, the change will accordingly affect voltage and frequency reading [18].

As transformer damage has a negative impact on the power system, it is said that the primary causes of the damage are poor cooling and overloading. For this use GSM (Global System for Mobile communication modem) technology has been made to monitor the parameters, which makes it easier for humans to recognize the fault and solve the failures before they even occur [19].

The given paper helps to reduce the cost of the system by using GSM (Global Service Module) instead of other large cables which are more costly and have limited use. GSM is also helpful to determine the fault in the transformer and further, it sends an SMS (Short Message Service) to the registered mobile number and tells the efficient use of networking, which enhances the improvement in the process [20].

This study presents the massive excitement of machine-to-machine communication and sharing of data using IoT (Internet of Things) kits. The sensors are used to track the voltage, intake variations in the system and focus on variations in parameters of load current, oil level, temperature, and gas formation faults which can damage the transformer [21].

The following research paper states the use of Adafruit software in the system using a Wi-Fi Module and Web Server. This can be another choice similar to Blynk App. It also tells that many powerful companies use SCADA (Supervisory Control and Data Acquisition) to monitor them, but it is way too costlier than the provided system [22].

A transformer health monitoring system based on IoT is the topic of the relevant paper. It is made up of a monitoring unit and an RTU (Remote Terminal Unit). With the aid of the PIC 18F4550, the RTU has analyzed parameters such as current, temperature, rise and decrease in oil level, vibration, and humidity [23].

The given article has shared an idea of monitoring and controlling the system in the industrial process with the help of a GUI(Graphical User Interface) based prototype and proposed hardware system to increase the safety and durability of the pieces of machinery by measuring the temperature, pressure, and vibrations produced by the machines [24].

3. Methodology

The transformer is an efficient and most important device for the electrical power systems. As the variation in voltage for different types of loads varies and causes over or under-voltage parameters due to the reason the electrical equipment may get damaged, and hence monitoring of the transformer is essential. In this system, smart electronic sensors have been used to measure different types of parameters of the transformer such as load current, operating voltage, temperature, and as the temperature increases the oil level decreases due to evaporation and hence the monitoring of all these parameters is required.

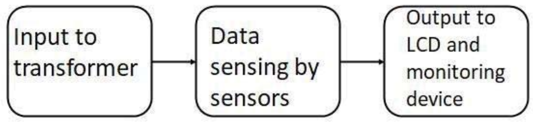


Fig. 1. Basic working of the system

First of all, AC voltage is supplied to the transformer, that is 5V power supply and 500mA of current. To monitor the temperature, use of LM35 temperature sensor has been made, which measures the temperature in °C. To rectify the transformer's AC supply to DC supply diode bridge has been used, as it is easy to measure DC voltage by Arduino. Simultaneously a voltage divider circuit is attached where the rectifier converts 12V to 5V. To measure the load current of a transformer ACS712 current sensor has been used, through which a maximum of 10 Ampere of current is measured, which means a transformer that holds a value of up to 10 Ampere is used. Next comes the ultrasonic sensor by which the measurement of the oil level of a transformer is done. To display all the parameters of the sensors a 16x2 LCD is used. Then comes the most important component of the system i.e., Arduino Nano, which handles all the operations and does serial communication with NodeMCU. And the second most important component is NodeMCU itself, which holds all the data and displays the real-time parameters on the connected Wi-Fi device. Also, the main switch is added so as to control the 230V power supply by which switching ON/OFF of the system is taken care of. As a load, the use of a Load Lamp has been done through which voltage and current are varied and hence a switch has been attached to turn ON/OFF the Load.

4. Features of the system

- In case of overload on the transformer protection of the transformer system and self-protection can be initialized by informing the maintenance teams.
- Frequently visit of Maintenance team on site and the workload will be decreased.
- Detection of unconditional faults in real-time such as changes in current, voltage, and temperature.
- Using Wi-Fi achieve more accuracy and fast response while monitoring fault.
- The system reduces loss of power supply that significantly benefits utility customers.
- The stability and reliability of the system are increased by the monitoring system.

5. System description

5.1. Transformer

The step-down transformer has a higher primary voltage than its secondary voltage. This transformer is made such that it reduces the voltage starting from primary winding towards secondary winding. The transformer's work is to lower the voltage and increase the low current power. It is necessary to use larger-gauge wire due to the increase in current. The primary winding doesn't have the need of conducting much current, so that may be made of smaller-gauge wire.

5.2. Blynk app

Blynk App is an IoT platform for Android smartphones which is used to control Arduino and NodeMCU via the internet. It is a multi-tenancy environment with configurable access levels. This App helps us to manage clients, distributors, contractors, installers, and anyone we work with.

5.3. Arduino

Arduino is a free open-source electronics platform that is used to connect software and hardware to each other and make them work together. It is also used to show the data of a hardware system on software.

5.4. NodeMCU

NodeMCU is an inexpensive and open-source IoT platform. NodeMCU uses Lua Scripting language. It includes open-source prototyping board designs. NodeMCU is a major component of this system.

5.5. Sensors

A sensor is an electronic device that senses different physical quantities and either records them or responds to them accordingly. Various sensors like temperature sensors, current sensors, ultrasonic sensors, etc. are used to build or create this system. Each of the sensors has different works of its own.

A. Temperature Sensor: Use of a Temperature sensor has been made to keep a record of heat produced in the transformer and save the data accordingly. The used temperature sensor can be operated over a temperature range of -55^{o}C to 150^{o}C, but it is suggested and rated to use it for the temperature range of -40^{o}C to 110^{o}C.

B. Ultrasonic Sensor: Ultrasonic sensor is used to measure the distance between the sensor and the object using ultrasonic soundwaves. This system includes an ultrasonic sensor to measure the oil level in the transformer. The sensor has a great range, starting from 2cm to 4m with high accuracy.

C. Current Sensor: Current sensor has been used to measure the current throughout the system. A current transformer and a current sensor are analogous. An external load's current is detected by current sensors, which then provide a signal that is directly proportionate to that current. The output that is generated may be analog voltage or current, which the ADC may transform into digital form. The measured current is shown using an appropriate measuring instrument and the generated signal. The data collected by current sensors can be stored in the database for further analysis and variable control. The current sensor output can take on many different forms, including DC input, a tripping output like that of a relay that repeats the value of the measured current digital output, triggered when the sensed current exceeded a predetermined pickup value.

D. Voltage Sensor: In simple terms, a voltage sensor is a potential transformer. Voltage sensors measure the output terminal voltage and produce a signal that is exactly proportional to that voltage. The supplied signal may be an analog current or voltage, which the ADC can transform into a digital output. The output generated is used by appropriate measuring equipment to display the exact voltage, or it can be recorded for additional analytical uses in the database to regulate other factors.

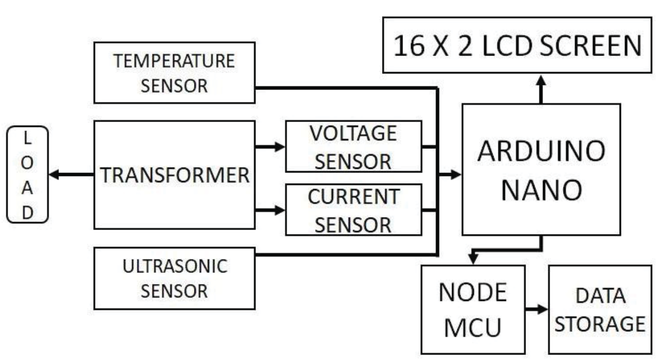


Fig. 2. Block Diagram of IoT based transformer monitoring system

The above figure is the detailed block diagram of the transformer monitoring system, which includes all the components used.

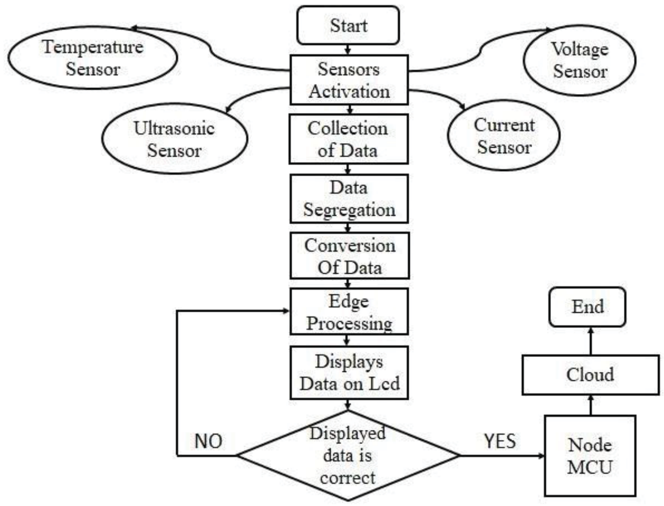


Fig. 3. System Workflow

The above figure describes the basic working of the transformer monitoring system which includes all the processes performed by the system.

6. Results

This system frequently checks the condition of the transformer from time to time and updates the same on the dashboard of the application. The authority at the observing center checks the dashboard regularly, and the time when the authority notices any kind of fluctuation in the data about the transformer that could possibly damage the transformer, whether it's a sudden increase in the temperature of the transformer, raise in the oil level inside, or excess of current or voltage, a person is sent to the transformer site to take care of the particular transformer and prevent it from being damaged. Thus, the transformer need not be frequently checked manually and saves enough time and energy by avoiding the entire process.

7. Conclusion and Future scope

This monitoring system most importantly tries to avoid the frequent failure of transformers, by notifying about it priorly on mobile phones. It also saves time and energy by ensuring the regular automatic check-up of the transformer about the probable factors which can damage the transformer.

Because of huge upgrades in technology, automation is nearly reaching every field of human work, making it easy and more comfortable for humans. The use of various sensors in this project has made it totally automatic and helps in reducing the workload of human beings. Transformers are used in our nation on a very large scale and in nearly every corner of the nation. This project will help in monitoring all these power transformers without the need to actually visit the site and check it manually. It can be used in the deep analysis of small-scale and large-scale transformers. This helps to predict the fault in the transformer before it occurs, which can further be avoided.

References

1. Singh, Jaspreet, and Sanjeev Singh. "Transformer failure analysis: reasons and methods." International Journal of Engineering Research & Technology 4, no. 15 (2016): 1-5.

2. Bengtsson, Claes. "Status and trends in transformer monitoring." IEEE Transactions on Power delivery 11, no. 3 (1996): 1379-1384

3. García, Belén, Juan Carlos Burgos, Ángel Matías Alonso, and Javier Sanz. "A moisture-in-oil model for power transformer monitoring-Part II: Experimental verification." IEEE Transactions on Power Delivery 20, no. 2 (2005):

4. Susa, Dejan, and Hasse Nordman. "A simple model for calculating transformer hot-spot temperature." IEEE transactions on power delivery 24, no. 3 (2009): 1257-1265.
5. Du, Yanqing, Alexander V. Mamishev, Bernard C. Lesieutre, Markus Zahn, and Seong-Hwa Kang. "Moisture solubility for differently conditioned transformer oils." IEEE transactions on Dielectrics and Electrical Insulation 8, no. 5
6. He, Qing, Jennie Si, and Daniel J. Tylavsky. "Prediction of top-oil temperature for Bethalsha, C., S. Karthik, and M. S. Sreejavijay. "Real-Time Transformer Health Monitoring using IOT." International Journal for Research in Applied Science and Engineering Technology 8, no. 9 (2020): 521-526.transformers using neural networks." IEEE Transactions on Power Delivery 15, no. 4 (2000): 1205-1211.
7. Saha, Tapan K., and Prithwiraj Purkait. "Investigation of polarization and depolarization current measurements for the assessment of oil-paper insulation of aged transformers." IEEE Transactions on Dielectrics and Electrical
8. Abu-Elanien, Ahmed EB, and M. M. A. Salama. "Asset management techniques for transformers." Electric power systems research 80, no. 4 (2010): 456-464.
9. Sica, Fernando Cortez, Frederico Gadelha Guimarães, Ricardo de Oliveira Duarte, and Agnaldo JR Reis. "A cognitive system for fault prognosis in power transformers." Electric Power Systems Research 127 (2015): 109-117.
10. Hasan, Walid KA, Abobaker Alraddad, Abdulfatah Ashour, Yachao Ran, Mohamed A. Alkelsh, and Reyadh AM Ajele. "Design and Implementation smart transformer based on IoT." In 2019 International Conference on Computing, Electronics & Communications Engineering (iCCECE), pp. 16-21. IEEE, 2019.
11. Mohamad, Auday AH, Yaqeen S. Mezaal, and Sevan F. Abdulkareem. "Computerized power transformer monitoring based on internet of things." International Journal of Engineering & Technology 7, no. 4 (2018): 2773-2778.
12. Poongothai, M., P. Muthu Subramanian, and A. Rajeswari. "Design and implementation of IoT based smart laboratory." In 2018 5th International Conference on Industrial Engineering and Applications (ICIEA), pp. 169-173. IEEE, 2018.
13. Navamanikumar, P. G., S. Agnesha, P. Gowsalya, K. Indhu, and N. Sivasakthi. "IOT Based Real Time Transformer Health Monitoring System and Phase Preventor." International Journal of Emerging Technologies in Engineering Research (IJETER) 6, no. 4 (2018): 119-124.
14. Pawar, Rohit R., and S. B. Deosarkar. "Health condition monitoring system for distribution transformer using Internet of Things (IoT)." In 2017 international conference on computing methodologies and communication (ICCMC), pp. 117-122. IEEE, 2017.
15. Kumar, R. Krishna, M. Thilagaraj, P. Vengatesh, J. Rajalakshmi, and MI Mohamed Babul. "Remote transformer faults analyzing system using IoT." International Journal of Modern Agriculture 10, no. 2 (2021): 2390-2402.
16. Sanjeeya, V., E. Ramyashankari, M. Sakthi Kannan, and A. Vidhya. "Real Time Transformer Health Monitoring system using IOT technology." International Journal of Science Technology & Engineering 9 (2018).
17. Beeranna, hs, imran pasha, bk madesha, and td lakshmikanth. "real time monitoring of transformer using iot." (2019).
18. Agbolade, Olaide Ayodeji, and Fatai Olaoluwa Sunmola. "Cellular Internet of Things Based Power Monitoring System for Networking Devices." European Journal of Electrical Engineering and Computer Science 5, no. 1 (2021): 80-84.
19. Ramteke, Punam. "SMART ALERT SYSTEM TO PROTECT DISTRIBUTION TRANSFORMER FOR VARIOUS FAULTS." Journal homepage: www. ijrpr. com ISSN 2582: 7421.
20. Rahman, Sajidur, Shimanta Kumar Dey, Bikash Kumar Bhawmick, and Nipu Kumar Das. "Design and implementation of real time transformer health monitoring system using GSM technology." In 2017 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 258-261. IEEE, 2017.
21. Dhanraj, Joshuva Arockia, Balachandar Krishnamurthy, Kuppan Chetty Ramanathan, A. K. Saravanan, and Jeya Krishna Ganapathy Raman. "Design on IoT Based Real Time Transformer Performance Monitoring System for Enhancing the Safety Measures." In IOP Conference Series: Materials Science and Engineering, vol. 988, no. 1, p. 012076. IOP Publishing, 2020.
22. Khandait, A. P., Swapnil Kadaskar, and Girish Thakare. "Real time monitoring of transformer using IOT." International Journal of Engineering Research & Technology (IJERT) 6, no. 03 (2017).
23. Pawar, Rohit R., Priyanka A. Wagh, and S. B. Deosarkar. "Distribution transformer monitoring system using Internet of Things (IoT)." In 2017 International Conference on Computational Intelligence in Data Science (ICCIDS), pp. 1-4. IEEE, 2017.
24. Barmase, Gaurav V., Gaurav V. Khopade, Shital P. Thawkar, Sahil P. Bawankule, Nikhil D. Gajbhiye, and Shailesh Sakhare. "GUI based Industrial Monitoring and Control System." Journal of Information Technology and Digital World 3, no. 2 (2021): 108-117.

Deep Variational Autoencoders and Muti-Omics Integration for Pancreatic Cancer Subtyping

Aadil Rashid Bhat[a]*, *Rana Hashmy*[a]

[a]*Department of Computer Sciences, University of Kashmir,190006, India*

Abstract

in both men and women, which claimed close to half a million deaths in 2020. Characterized by late diagnosis, poor survival rates, and high incidence of metastasis, Pancreatic cancer is predicted to become the second leading cause of cancer-related deaths by 2030. However, there has been little advancement in terms of early detection and effective treatments for Pancreatic Cancer, leading to a dismal 5-year survival rate of 3-15%. Which renders the unmet challenge of early diagnosis of PC both urgent and important. Recently, multi-omics analysis of numerous cancers has provided a new perspective on genomics, epigenomics, and transcriptomics deregulations in cancer. Which helped with fine-tuned characterization, classification, and early diagnosis of Cancers. However, due to the vast number of heterogeneous variables in each omics data, and its disparate and dynamic nature multi-omics data possess many challenges in data integration and knowledge discovery. In this study, we used Deep learning to find a latent representation of integrated multi-omics data together with many clustering methods to find homogeneous pancreatic cancer subtypes. Which can explain differences in disease trajectories and outcomes in heterogeneous cohorts. And help improve early diagnosis, treatment, and prognosis of pancreatic cancer.

Keywords: Multi-Omics; Data Integration; Deep Variational Autoencoder; Cancer Subtyping.

1. Introduction

Pancreatic cancer is a highly lethal form of cancer, ranking as the seventh most common cause of cancer-related fatalities worldwide. [1]. Pancreatic ductal adenocarcinoma (PDAC) has a very high mortality rate and its incidence is increasing, it has the lowest 5-year survival rates, i.e., Only 9% of people with PDAC are able to survive for five years after diagnosis. Pancreatic cancer was accountable for nearly 500,000 deaths globally in 2020, as reported by GLOBOCAN 2020. PDAC is highly heterogeneous, leading to differences in oncogenesis and varying survival rates among patients. Consequently, this heterogeneity poses significant clinical challenges, including inaccurate diagnoses and inadequate treatment approaches. [2]. Contemporary medical and molecular diagnostic techniques, such as computed tomography (CT), magnetic resonance imaging (MRI), and endoscopic ultrasound (EUS) accompanied by fine-needle aspiration (FNA), are currently available. However, these techniques *offer limited information regarding tumor aggressiveness and the probable disease* prognosis, posing a significant challenge to the development of an accurate treatment regimen[3]. As a result, the prognosis post-surgery remains, in most cases, uncertain[4] (Guillén-Ponce et al., 2017). A case in point, is the CA 19-9, as discussed in this study. Which is a widely used cancer marker for monitoring treatment Responses [5]. And This informs the treatment strategy, however, this biomarker has demonstrated high false positive and high false negative results[4].

1.1 Multi-Omics

The advancement in high-throughput technologies[6], and explosive growth in biological data collection, e.g., Several thousand biological samples have been profiled and made publicly available by The Cancer Genome Atlas (TCGA) and The International Cancer Genome Consortium (ICGC) have allowed researchers to understand the molecular bases of the genetic disorders, facilitating effective and personalized diagnosis and treatments in case of many cancers, including PC[7][8][9][10]. However, cancer research that concentrates on only one aspect of biological data (single-omics) has only furnished limited insights into the causes of cancer development and the advancement of tumors[11]. These single omics studies have often resulted in different and at times conflicting patient classifications[12]. And as such many multi-omics studies of various cancers have facilitated a deeper understanding of genomics, epigenomics, and transcriptomics deregulations in malignancies[13]. One such method is the use of the machine learning model called Autoencoders, which is an Artificial Intelligence method of relearning the latent space or manifold of the high-dimensional space to extract meaningful information from large data bodies. The omics data are charlatanically very complex in their dimensions and hence application of AE In such situations has been extensively studied.

*Aadil Rashid Bhat
E-mail address: aadil.bhat@uok.edu.in

1.2 Dimensionality Reduction and Machine Learning

Dimensionality reduction is a widely adopted technique in data science and machine learning that involves the process of converting information from a space with a large number of dimensions into a smaller subspace with fewer dimensions. The primary objective of this technique is to preserve the crucial attributes of the original data while reducing its complexity and computational requirements One way to accomplish this is by mapping the important features of the data onto a subspace with fewer dimensions that captures the essential characteristics of the original data that are relevant to a given use case or analysis.[14]. Dimension reduction techniques have been used for classification, visualization, and data compression in many fields including bioinformatics[15]. Well-known techniques for data analysis such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Multi-Dimensional Scaling (MDS)[14] have been extensively studied. However, these methods suffer from different shortcomings like the inability to capture global structures[16], information loss, and processing of large-scale datasets[17]. Apart from DM methods, many methods have been devised for the integration of multi-omics data these methods attempt to find the complementary signals from different omics functional levels to better understand the molecular characterization of underlying conditions like cancer. This combined analysis has the potential to reveal novel biomarkers for better characterization and homogenous classification of the cancer types and subtypes and an assistant in personalized care and better treatment plans. multi-omics for cancer patient stratification has also been widely studied. For example, The intNMF[18], which is a non-negative matrix factorization method, which does not assume any distributional form of data, The LRACluster[19] a probabilistic integrative mode based on low-rank approximation, which assumes and estimates the principal latent subspace for the entire data, The Mixkernal[20], which computes similarity matrices from kernels which are then combined to obtain a combined Similarity matrix. The SNF, which is a popular similarity network fusion method that uses graphs to model patient-patient similarity using multi-omics data, and the RCGAA[21] a generalized canonical correlation analysis framework that allows for choosing various parameters like scheme functions and shrinkage constants. Even though unsupervised subtyping revealed the molecular diversity among PDAC patients, the survival outcomes varied widely within each subtype. As a result, there is no significant difference in prognosis among the subtypes identified using these methods [22]. More recently Artificial intelligence-based methods which have shown remarkable performance in other data extensive fields have been applied to the problem of dimensionality reduction and data integration. In this study, we used Deep hierarchical Variational Autoencoders (DVAE) for dimensionality reduction and multi-omics integration to learn homogenous subtypes of pancreatic cancer.

2. Description of model for PC classification

2.1 Proposed model.

Our proposed model is a multi-level deep Variational Autoencoder (VAE) based integration and dimensionality reduction framework. The standard VAE[23] is a probabilistic deep learning method that is used to extract the low-dimensional data manifold from high-dimensional datasets. Instead of representing each input x_i as a singular value, the VAE encodes it as a distribution characterized by its mean and standard deviation across a latent space. VAEs are typically composed of two networks an input network that encodes the data, and a decoder counterpart that aims to reproduce the original input data from the low-dimensional embeddings, see Fig 1.
The VAE approximates the latent distribution by minimizing the kl divergence score, as in Eq. (1).

$$D_{KL}(q_{\phi}(z) \parallel p_{\theta}(z)) \tag{1}$$

Where $q_{\phi}(z)$, called a variational distribution, is the estimation of the true but intractable posterior $p_{\theta}(z)$. Recently, Maximum Mean Discrepancy (MMD)[24] function in comparison to KL divergence was shown to produce better results for approximating the true posterior. MMD, which is given by Eq. (2),

$$MMD(p(z)||q(z)) = \mathbb{E}_{(p(z),p(z'))}[k(z,z')] + \mathbb{E}_{(q(z),q(z'))}[k(z,z')] - 2\mathbb{E}_{(p(z),q(z'))}[k(z,z')] \tag{2}$$

states that two distributions can only be considered identical if their moments are equal. Therefore, we can measure divergence by comparing the moments of two distributions $p(z)$ and $q(z)$. By using kernel embedding, MMD can efficiently accomplish this

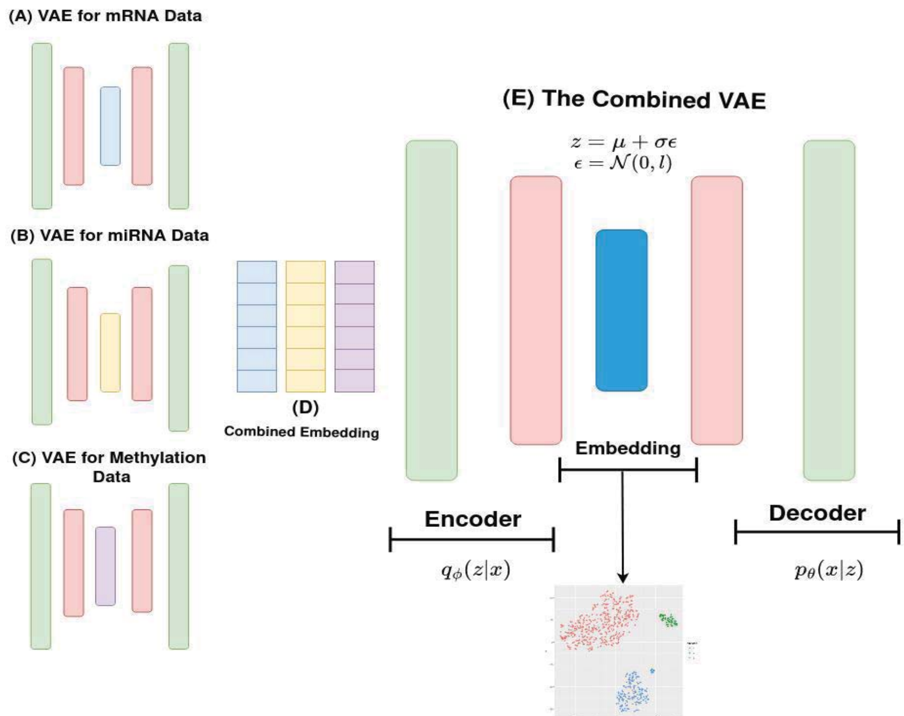


Fig. 1. The Architecture of Deep Hierarchical Variational Autoencoder for PAAD. A separate VAE is built for (A) mRNA (B) microRNA (C) Methylation. The individual omics embeddings are then combined (E) to train the combined VAE model (E).

In this study, we trained three different MMD-based VAEs for each omics data and the outputs of which were fed to another VAE to generate the final embeddings which capture the latent manifold of the entire data and are subsequently employed for finding homogenous subtypes of the PC.

1.3 Datasets

In the present investigation, we employed a comprehensive approach for pancreatic cancer (PC) subtyping utilizing four distinct omics data sets. Specifically, we retrieved data about mRNA expression, microRNA expression, DNA methylation array, as well as clinical parameters of a cohort of 150 patients who underwent surgical removal of their primary pancreatic ductal adenocarcinoma (PDAC from The Cancer Genome Atlas (TCGA) Pancreatic Adenocarcinoma (PAAD) database [7].

Training Dataset: We obtained transcriptome-wide information on 177 patients by downloading data from The Cancer Genome Atlas (TCGA) Pancreatic Adenocarcinoma (PAAD) cohort using the R package TCGA-Assembler[25]. The information set consisted of three types of biological data: mRNA sequencing (mRNA-Seq), microRNA sequencing (microRNA), and DNA methylation array data. The Illumina HiSeq platform was utilized to generate both mRNA-Seq and microRNA data, while the Illumina Infinium HumanMethylation450 BeadChip platform was used to obtain the DNA methylation data. The mRNA-Seq data, as per TCGA, was processed and normalized by Expectation Maximization (RSEM) [26]. Likewise, the RPM normalization method was applied to the microRNA-Seq data. Only patient samples that had complete data for all three types of multi-omics and clinical information (146 samples) were kept. Any genes that had missing data greater than 20% for DNA methylation, as well as any genes from mRNA and microRNAs that had zero values greater than 20% among the retained samples, were excluded. [27][20] DNA methylation genes exhibiting a proportion of missing values less than or equal to 20% were subjected to imputation using the R package impute. Subsequently, to standardize the values of the mRNA and microRNA datasets, a log transformation was utilized. This data preprocessing strategy was implemented to reduce the impact of missing data on downstream analyses, as well as to facilitate meaningful comparisons between gene expression profiles across multiple samples.

Test Datasets: To corroborate our findings and evaluate the classification efficacy of our model, we procured two external datasets from the Gene Expression Omnibus (GEO), with the accession numbers GSE62498 and GSE62452, respectively.[28]. The GEO GSE62452 mRNA microarray dataset was procured from the Affymetrix GeneChip platform and was subjected to the robust multi-array average (RMA) normalization technique. Subsequently, the data generators obtained the average expression values of each gene from the multiple corresponding probe sets. The expression values were then subjected to a logarithmic transformation. The GEO GSE62498 microRNA dataset, on the other hand, was obtained from the Nanostring nCounter Platform. The dataset underwent normalization through the geometric mean, followed by a logarithmic transformation using the formula $\log_2(x+1)$.

2.2 Model Training

Four omics pre-processed TCGA PC data for a total of 146 patients were used as input to the Autoencoders. The DVAE architecture was constructed utilizing the Keras library in the Python programming language. The design consists of a symmetric encoder and decoder neural network architecture with two hidden layers and a maximum mean discrepancy (MMD) loss function. ReLU, a commonly used nonlinear function, was employed as the activation function in each layer except for the output layer, which employed the sigmoid function for each layer the output y, given the input x, is calculated as in Eq. (3).

$$y_i = relu(\dot{W}_i x + b_i) = f_i(x) \quad (3)$$

So, the output $\hat{x}$, is given by Eq. (4).

$$\hat{x} = sigmoid(f_3(f_2(f_1(x)) + b) = F(x) \quad (4)$$

After the model reduced the number of features to 500, a univariate Cox-PH model was generated on each feature, and the top features associated with survival (P-values * 0.05) were identified using the R survival package. K-means was next used to cluster the data into two different survival subgroups.

3. Results

Subsequently, we retrieved the omics characteristics with the most unique expression patterns between the two subtypes (i.e., high-risk or low-risk) that had been previously identified. The ANOVA-F test was applied to assess the features that were significantly related to each subgroup, and the selection of features was based on the average precision obtained from 5-fold cross-validation on the training set. These features were then utilized to train a random-forest classifier. Furthermore, the effectiveness of our approach was confirmed by evaluating two external datasets. The Kaplan–Meier plot of the training dataset and the two external datasets are presented below in Fig 2

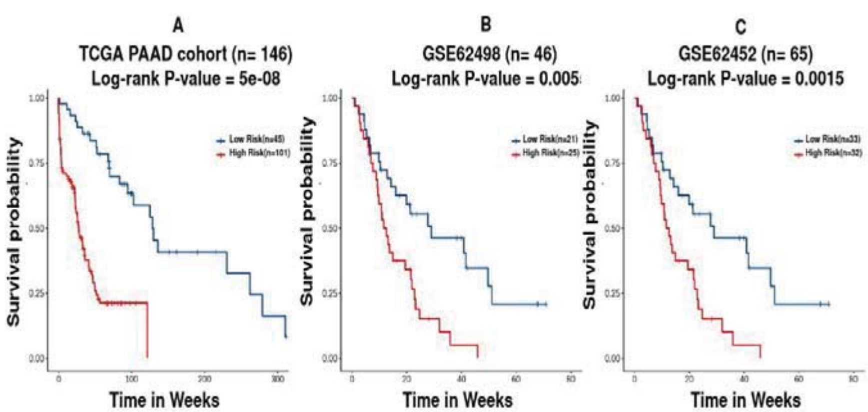


Fig. 2. The Kaplan-Meier survival plots of the subgroups for the (A) TCGA-PAAD cohort and two external datasets (B) GSE62498 with 46 patients and p-value of 0.0055 (C) GSE62452 with 65 patients and p-value of 0.0015.

3.1 Comparison with other methods

We then compared the performance of our model against 6 other integration methods based on different statistical frameworks like Matrix factorization, consensus clustering and co-inertia analysis, and similarity matrix.

we assessed the effectiveness of integration and dimensionality reduction algorithms using the log-rank p-value and concordance index as the evaluation metrics. see the comparison table and plots in Fig 3. Our proposed model demonstrated superior performance to other existing multi-omics models, as evidenced by a log-rank p-value of 5e-08 and a c-index of 0.6505, as depicted in Fig 3. These findings suggest that our model has the ability to extract meaningful subspace manifold and possesses discriminative power to accurately classify new, unseen single omics datasets into the identified subtypes.

(A) The Comparison of various Intgeration methods

	P-log Value	Concordance index
Our Method	5e-08	0.6505
MCIA	0.008	0.5606
SNF	1e-05	0.6131
RGCAA	3e-06	0.6374
intNMF	0.0002	0.5898
MixedKernal	6e-05	0.6175
LRACluster	0.009	0.5664

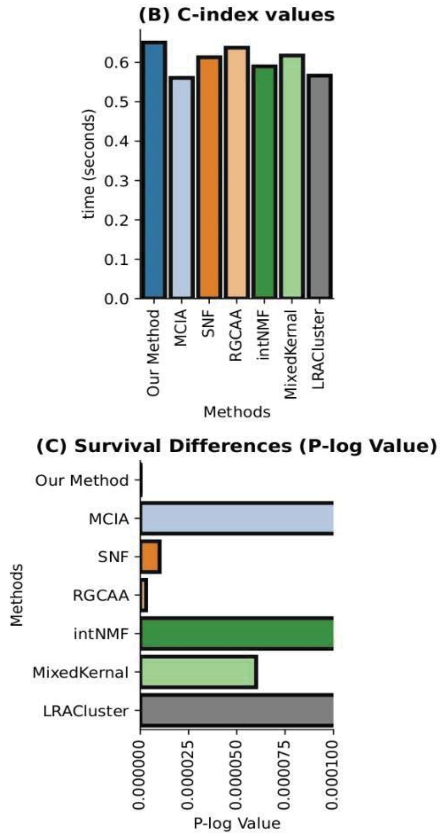


Fig. 3. The performance comparison of our model against six other related methods (a) table summarizing the findings, and the bar plots of the (b) log-Rank P-values and (c) concordance index.

4. Discussion and Conclusion

Over the past few years, there has been significant research dedicated to the detection of molecular and biomarker features of cancers, which includes pancreatic cancer (PC). A study conducted in 2012 on 1027 cases of PC and 1031 controls of Han Chinese patients identified a noteworthy correlation between the incidence of pancreatic cancer and the copy number of CNVR2966.1 located at 6q13.[29]. Another study in 2015 established a link between amplification in the MYC gene and shorter patient survival duration.[30] Multi-omics studies have been employed to improve diagnostic tools, such as Comp Cyst, which is a machine learning-based test used to manage patients with Pancreatic cystic lesions (PCLs) and is estimated to help avoid 60% of unnecessary surgeries.[31] . Multi-omics has also been used to study cancer patient stratification. For example, the iCluster method utilized gene expression and copy number variation information to identify different subtypes of breast and lung cancer. This approach demonstrated that utilizing multiple sources of information (multi-omics) results in more informative subtypes compared to using only one source (single omics). The Similarity Networks Fusion (SNF) technique was also applied to identify molecular subtypes of pancreatic cancer using a combination of proteins, mRNAs, DNA methylation, and microRNA profiles[8]. Although unsupervised subtyping helped using non-AI methods to identify molecular diversity in PDAC patients, the patients in each subtype still exhibited a wide range of survival outcomes, and the disparities among subtypes were not statistically significant.[22].

In this study, we used the ability of the Deep Variational Autoencoder to extract the latent clustering patterns within the Pancreatic Cancers patients to identify and characterize the two prognosis subtypes. These features were then used to train a machine learning classification model to classify the unseen Pancreatic cancer datasets into the identified subtypes.

References

1. M. D. Siegelin and A. C. Borczuk, "Epidermal growth factor receptor mutations in lung adenocarcinoma," *Lab. Investig.*, vol. 94, no. 2, pp. 129–137, 2014, doi: 10.1038/labinvest.2013.147.
2. D. P. Ryan, T. S. Hong, and N. Bardeesy, "Pancreatic Adenocarcinoma," pp. 1039–1049, 2014, doi: 10.1056/NEJMra1404198.
3. S. S. Vege, B. Ziring, R. Jain, and P. Moayyedi, "American gastroenterological association institute guideline on the diagnosis and management of asymptomatic neoplastic pancreatic cysts," *Gastroenterology*, vol. 148, no. 4, pp. 819–822, 2015, doi: 10.1053/j.gastro.2015.01.015.
4. J. Bla, "Diagnosis and staging of pancreatic ductal adenocarcinoma," pp. 1205–1216, 2017, doi: 10.1007/s12094-017-1681-7.
5. V. Eijck, A. P. Stubbs, and Y. Li, "Jo ur na l P re of," *ISCIENCE*, p. 103415, 2021, doi: 10.1016/j.isci.2021.103415.
6. [Y. Hasin, M. Seldin, and A. Lusis, "Multi-omics approaches to disease," *Genome Biol.*, vol. 18, no. 1, pp. 1–15, 2017, doi: 10.1186/s13059-017-1215-1.
7. B. J. Raphael *et al.*, "Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma," *Cancer Cell*, vol. 32, no. 2, pp. 185-203.e13, 2017, doi: 10.1016/j.ccell.2017.07.007.
8. M. Sinkala, N. Mulder, and D. Martin, "Machine Learning and Network Analyses Reveal Disease Subtypes of Pancreatic Cancer and their Molecular Characteristics," pp. 1–14, 2020, doi: 10.1038/s41598-020-58290-2.
9. T. Golan and M. Javle, "DNA Repair Dysfunction in Pancreatic Cancer : A Clinically Relevant Subtype for Drug Development," vol. 15, no. 8, pp. 1063–1069, 2017, doi: 10.6004/jnccn.2017.0133.
10. T. J. Grant, K. Hua, and A. Singh, *Molecular Pathogenesis of Pancreatic Cancer*, 1st ed. Elsevier Inc., 2016.
11. M. D. Ritchie, E. R. Holzinger, R. Li, S. A. Pendergrass, and D. Kim, "Methods of integrating data to uncover genotype-phenotype interactions," *Nat. Rev. Genet.*, vol. 16, no. 2, pp. 85–97, 2015, doi: 10.1038/nrg3868.
12. B. Wang *et al.*, "Similarity network fusion for aggregating data types on a genomic scale," *Nat. Methods*, vol. 11, no. 3, pp. 333–337, 2014, doi: 10.1038/nmeth.2810.
13. Y. Hasin *et al.*, "A fully Bayesian latent variable model for integrative clustering analysis of multi-type omics data," *Genome Biol.*, vol. 19, no. 1, pp. 71–86, 2018, doi: 10.1093/biostatistics/kxx017.
14. [R. Zebari, A. Abdulazeez, D. Zeebaree, D. Zebari, and J. Saeed, "A Comprehensive Review of Dimensionality Reduction Techniques for Feature Selection and Feature Extraction," *J. Appl. Sci. Technol. Trends*, vol. 1, no. 2, pp. 56–70, 2020, doi: 10.38094/jastt1224.
15. [E. Postma and E. Postma, "Dimensionality Reduction : A Comparative Review Dimensionality Reduction : A Comparative Review," 2009.
16. [J. B. Tenenbaum, V. De Silva, and J. C. Langford, "A global geometric framework for nonlinear dimensionality reduction," *Science (80-.).*, vol. 290, no. 5500, pp. 2319–2323, 2000, doi: 10.1126/science.290.5500.2319.
17. E. Becht *et al.*, "A n a ly s i s Dimensionality reduction for visualizing single-cell data using UMAP," vol. 37, no. 1, 2019, doi: 10.1038/nbt.4314.
18. P. Chalise and B. L. Fridley, "Integrative clustering of multi-level 'omic data based on non-negative matrix factorization algorithm," *PLoS One*, vol. 12, no. 5, pp. 1–18, 2017, doi: 10.1371/journal.pone.0176278.
19. D. Wu, D. Wang, M. Q. Zhang, and J. Gu, "Fast dimension reduction and integrative clustering of multi-omics data using lowrank approximation: Application to cancer molecular classification," *BMC Genomics*, vol. 16, no. 1, pp. 1–10, 2015, doi: 10.1186/s12864-015-2223-8.
20. B. Zhu *et al.*, "Integrating Clinical and Multiple Omics Data for Prognostic Assessment across Human Cancers," *Sci. Rep.*, vol. 7, no. 1, pp. 1–13, 2017, doi: 10.1038/s41598-017-17031-8.
21. M. Tenenhaus, A. Tenenhaus, and P. J. F. Groenen, "Regularized Generalized Canonical Correlation Analysis: A Framework for Sequential Multiblock Component Methods," *Psychometrika*, vol. 82, no. 3, pp. 737–777, 2017, doi: 10.1007/s11336-017-9573-x.
22. A. J. Aguirre, "Refining Classification of Pancreatic Cancer Subtypes to Improve Clinical Care," *Gastroenterology*, vol. 155, no. 6, pp. 1689–1691, 2018, doi: 10.1053/j.gastro.2018.11.004.
23. D. P. Kingma and M. Welling, "Auto-Encoding Variational Bayes." 2014.
24. A. Gretton, K. M. Borgwardt, M. Rasch, B. Schölkopf, and A. J. Smola, "A kernel method for the two-sample-problem," *Adv. Neural Inf. Process. Syst.*, vol. 1, pp. 513–520, 2007, doi: 10.7551/mitpress/7503.003.0069.
25. L. Wei, Z. Jin, S. Yang, Y. Xu, Y. Zhu, and Y. Ji, "TCGA-assembler 2: Software pipeline for retrieval and processing of TCGA/CPTAC data," *Bioinformatics*, vol. 34, no. 9, pp. 1615–1617, 2018, doi: 10.1093/bioinformatics/btx812.
26. B. Li and C. N. Dewey, "RSEM: Accurate transcript quantification from RNA-seq data with or without a reference genome," *Bioinforma. Impact Accurate Quantif. Proteomic Genet. Anal. Res.*, pp. 41–74, 2014, doi: 10.1201/b16589.
27. S. Huang, K. Chaudhary, and L. X. Garmire, "More is better: Recent progress in multi-omics data integration methods," *Front.*

Genet., vol. 8, no. JUN, pp. 1–12, 2017, doi: 10.3389/fgene.2017.00084.
28. J. Luo *et al.*, "Prognostic and predictive value of the novel classification of lung adenocarcinoma in patients with stage IB," *J. Cancer Res. Clin. Oncol.*, vol. 142, no. 9, pp. 2031–2040, 2016, doi: 10.1007/s00432-016-2192-6.
29. L. Huang *et al.*, "Copy number variation at 6q13 functions as a long-range regulator and is associated with pancreatic cancer risk," *Carcinogenesis*, vol. 33, no. 1, pp. 94–100, 2012, doi: 10.1093/carcin/bgr228.
30. A. K. Witkiewicz *et al.*, "Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets," *Nat. Commun.*, vol. 6, pp. 1–11, 2015, doi: 10.1038/ncomms7744.
31. E. J. Hoorn, "Multicenter Paper," *Physiol. Behav.*, vol. 176, no. 1, pp. 100–106, 2017, doi: 10.1126/scitranslmed.aav4772.A.

In-depth Exploration and Sentiment Analysis of Women's E-Commerce Clothing Ratings using SVM and Logistic Regression

*Abderaouf Bahi, Ibtissem Gasmi, Sassi Bentrad, Ramzi Khantouchi, Maroua Benleulmi

Department of Computer Science, Chadli Bendjedid El Tarf University, Algeria

Abstract

This study presents an in-depth exploration and sentiment analysis of a large e-commerce database: Women's E-Commerce Clothing. Exploratory data analysis and visualization techniques were used to gain insights into customer preferences and popular clothing styles. SVM and logistic regression classification models were employed to classify each rating as positive, neutral, or negative based on the expressed sentiment. The results showed that these models performed better than the others used for the same database. The logistic regression classification model achieved an overall ternary classification accuracy of 95%, and the SVM model achieved an accuracy of 94%, indicating the usefulness of sentiment analysis in understanding customer opinions and experiences with women's clothing. The findings of this study provide a solution for how to gain valuable insights for the fashion industry to improve customer satisfaction and enhance overall business performance.

Keywords: Sentiment analysis ,Women E-commerce, Data exploration, Data visualization, Machine Learning.

1. Introduction

Over the past few years, an exponential growth in e-commerce has led to a significant increase in the availability and variety of clothing available for purchase online [1, 2]. Before making online purchases, customers are increasingly relying on ratings and reviews from past customers to guide their purchase and make a valid decision. The presence of this data also offers researchers the opportunity to explore and analyze clothing patterns and trends to be able to understand users' preferences and guide them in their purchases with targeted recommendations [3].

Interests in using data-driven approaches to better understand consumer behaviour and preferences is growing. Data mining and visualization techniques are particularly well suited to analyzing large data sets to identify trends that may not be immediately visible from the raw data [4]. By using these techniques, researchers can gain valuable insights into consumer behaviour and preferences, which can inform business strategies and significantly improve customer experience and drive them to make more purchases.

We propose to conduct exploratory and sentiment analysis of women's clothing e-commerce rating data, using data visualization techniques, logistic regression and SVM models using a range of cleaning techniques, transformation, visualization to provide detailed and clearer understanding of data.

Logistic regression and SVM are known for their ability to handle complex data sets and capture the nonlinear relationships between variables that are crucial in sentiment analysis [5]. On the other hand, although neural networks and naive Bayes are powerful models [6], they can struggle to handle high dimensionality and sparseness of textual data, which could result in lower performance compared to logistic regression. and SVM.

The main objective of this study is to better understand women's fashion preferences and behavior in the e-commerce landscape. More specifically, we aim to answer the following research questions:

- What are the most popular women's clothing products?
- What insights can be gained from sentiment analysis of data?
- How to inform business strategies and improve consumer experiences?

* *Abderaouf Bahi*
E-mail address: a.bahi@univ-eltarf.dz

2. Background and related

1.1. Depth data exploration

Several studies have analyzed customer sentiment in e-commerce through machine learning models. In an old study [7], researchers analyzed customer reviews of clothing products from a Chinese e-commerce platform. Their analysis included exploring the distribution of ratings in the dataset and identifying the most frequently examined types of clothing.
Similarly, another study [8] conducted a sentiment analysis of customer reviews of clothing products from an online retailer in the United States. Their exploration of the data included examining the frequency of sentiment words positive and negative and identifying the most popular categories.
As part of machine learning techniques researchers explored the use of support vector machines for sentiment analysis of customer reviews in various fields [9]. Their study included techniques of pre-processing such as stem and stop word removal, as well as exploring the impact of different feature selection methods on model performance.

2.2 Sentiment analysis

Sentiment analysis in the field of clothing has received significant attention in recent years. Various machine learning models were used to extract and rank sentiment from reviews and text reviews. SVM was used to rank the sentiment of user reviews of an apparel e-commerce [10, 11]. The authors used various feature extraction techniques, such as word bags, n-grams, and topic models, to capture semantic and syntactic information from textual data. The results showed that SVM achieved an accuracy of 87.2%, outperforming other models such as logistic regression, naive Bayes, and random forest. In contrast, some studies have explored the use of neural networks in sentiment analysis of women's clothing reviews. For example, a convolutional neural network (CNN) was used to rank the sentiment of women's clothing reviews [12]. The authors used pre-trained word embeddings to represent the textual data and achieved an accuracy of 84.6%. Additionally, a recurrent neural network (RNN) was used to classify the sentiment of clothing reviews [13]. Using a combination of character-level integrations and achieved 82.1% accuracy. Although neural networks have been explored in this area, their performance may be limited by the high dimensionality and sparseness of textual data.

3. Proposed methodology

The proposed approach aims to establish the complete cycle of a data science project on the studied database, from data cleaning to the visualization of results through sentiment analysis using learning models automatic, Figure 1 presents the course of this work.

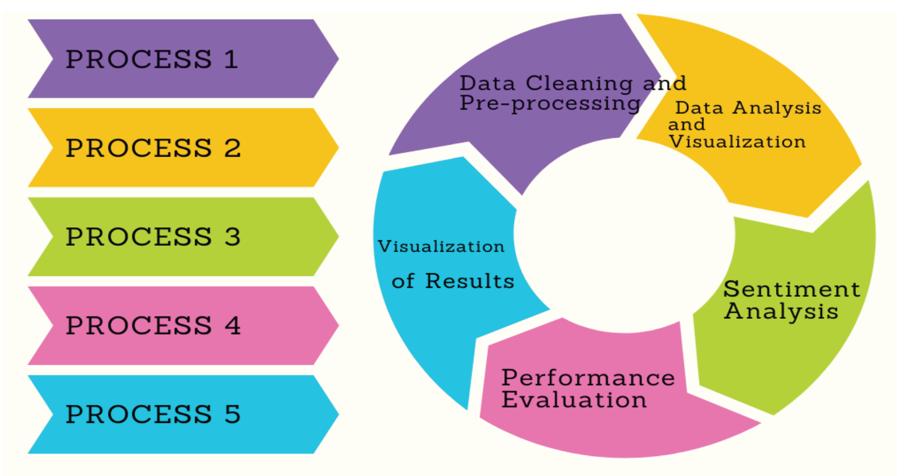


Fig. 1. Work Processing

3.1 Sentiment analysis

Exploratory data analysis (EDA) [14] aims at the detailed exploration and visualization of data and begins with the pre-processing of the database which includes cleaning and transformation, as well as visualization in graphical form to be able to easily extract knowledge from figures. Steps are taken to overcome any challenges and limitations that arise in the data set.

3.2 Logistic regression model

The process of creating and evaluating the logistic regression model for sentiment analysis can include steps such as dividing the data into training and testing sets [15], selecting hyper parameters, and fitting the model. The criteria used to assess model performance are accuracy, precision, recall, and F1 score. Finally, the results of the logistic regression model are presented, including its performance relative to other models and its implications for understanding customer sentiment.

3.3 SVM model

The SVM model for sentiment analysis involves several stages [16], including segmenting the data into training and testing sets, selecting hyper parameters, and fitting the model. The accuracy, precision, recall, and F1 score are commonly used to evaluate the model's performance. The final output of the SVM model is presented, which includes its performance compared to other models and its potential significance in understanding customer sentiment.

4. Experimental Setup

4.1 Dataset

We used a real Women's Clothing E-Commerce dataset [17] revolving around the reviews written by customers. This dataset includes 23486 rows and 10 feature variables. Each row corresponds to a customer review, and includes the variables:

- *Clothing ID*: Integer Categorical variable that refers to the specific piece being reviewed;
- *Age:* Positive Integer variable of the reviewer's age;
- *Title*: String variable for the title of the review;
- *Review Text:* String variable for the review body;
- *Rating:* Positive Ordinal Integer variable for the product score granted by the customer from 1 Worst, to 5 Best;
- *Recommended IND:* Binary variable stating where the customer recommends the product where 1 is recommended, 0 is not recommended;
- *Positive Feedback Count:* Positive Integer documenting the number of other customers who found this review positive;
- *Division Name:* Categorical name of the product high level division;
- *Department Name:* Categorical name of the product department name;
- *Class Name:* Categorical name of the product class name;
- *Data Cleaning and Transformation.*

4.2 Method compared

To assess the performance of the SVM and logistic regression models for sentiment analysis, several evaluation metrics were employed, including accuracy, precision, recall, and F1-score [18].
Moreover, we examined the feature selection and engineering techniques used for each model, including the choice of stop words, stemming or lemmatization, and n-gram selection. We also explored the effects of different parameter settings on model performance.
Overall, our study provides valuable insights into the strengths and weaknesses of SVM and logistic regression models for sentiment analysis. By understanding the performance differences and feature selection techniques used for each model, we can better tailor our approaches to accurately assess customer sentiment towards women's clothing."

4.3 Parameter setting

For the SVM model, the hyper parameters that were tuned included the regularization strength parameter (C) and the kernel type (linear, polynomial, or radial basis function) [19]. For logistic regression, the hyper parameters that were tuned included the regularization strength parameter (C) and the penalty type (L1 or L2) [20]. The optimal hyper parameters for both models were selected using a randomized search approach. In addition to hyperparameter tuning, feature engineering was performed to optimize the models' performance, including techniques such as bag of words and n-gram analysis.

Finally, the optimal hyper parameter settings for both models were presented, and their implications for understanding customer sentiment towards women's clothing were discussed.

4.4 Evaluation metrics

The following metrics [21] were used to evaluate the performance of the SVM and logistic regression models used for sentiment analysis of the Women's E-Commerce Clothing ratings:
Accuracy: The percentage of correctly classified ratings over the total number of ratings:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

Precision: The proportion of true positive predictions over the total number of predicted positive ratings:

$$\text{Precision} = \frac{TP}{TP + FP} \tag{2}$$

Recall: The proportion of true positive predictions over the total number of actual positive ratings:

$$\text{Re}\,call = \frac{TP}{TP + FN} \tag{3}$$

F1-score: The harmonic mean of precision and recall:

$$F1 - score = \frac{2 \bullet \text{Precision} \bullet \text{Re}\,call}{\text{Precision} + \text{Re}\,call} \tag{4}$$

Where TP (True Positive) is the number of positive ratings correctly predicted by the model, TN (True Negative) is the number of negative ratings correctly predicted, FP (False Positive) is the number of negative ratings predicted as positive, and FN (False Negative) is the number of positive ratings predicted as negative.

5. Result and Discussion

5.1 Data visualization

Before conducting any analysis, we first cleaned and transformed the data to ensure its quality and consistency. This involved several steps, including:

- Identified and remove duplicate entries in the dataset to eliminate any redundant data.
- Replace missing values with appropriate values based on the context of the data.
- Normalize the data to ensure consistency across different categories.
- Created new features based on the existing data to extract additional insights.

Once the data was cleaned and transformed, we conducted exploratory data analysis (EDA) to gain insights into the data. We used a combination of statistical analysis and data visualization techniques to explore the data and identify patterns. Figure 2 presents an analysis of the rating distribution across different departments, class names, and divisions, utilizing various data visualization techniques in Python.

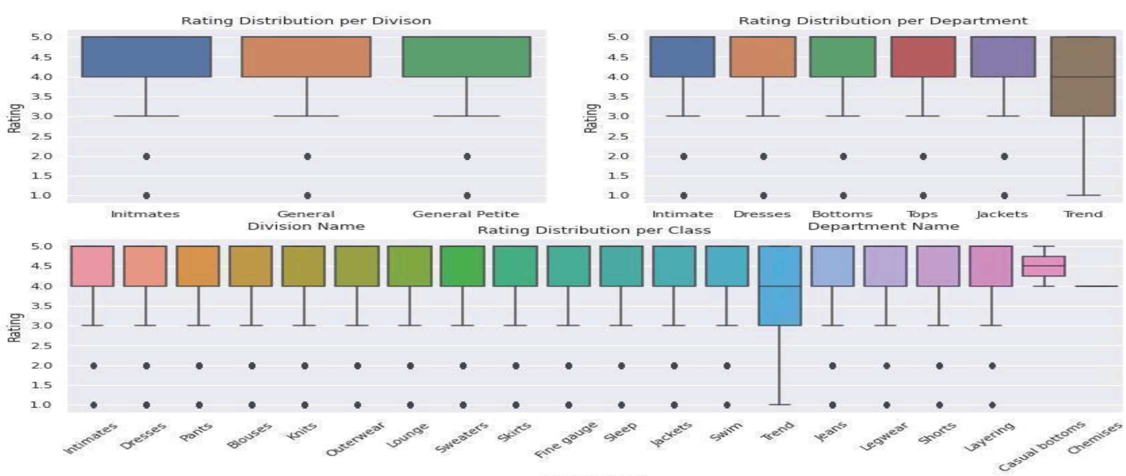


Fig. 2. Rating Distribution

In Figure 3, we examine the distribution of reviews across various departments, divisions, and classes. This figure offers valuable insights into the patterns of review distribution, emphasizing potential areas of interest for further investigation.

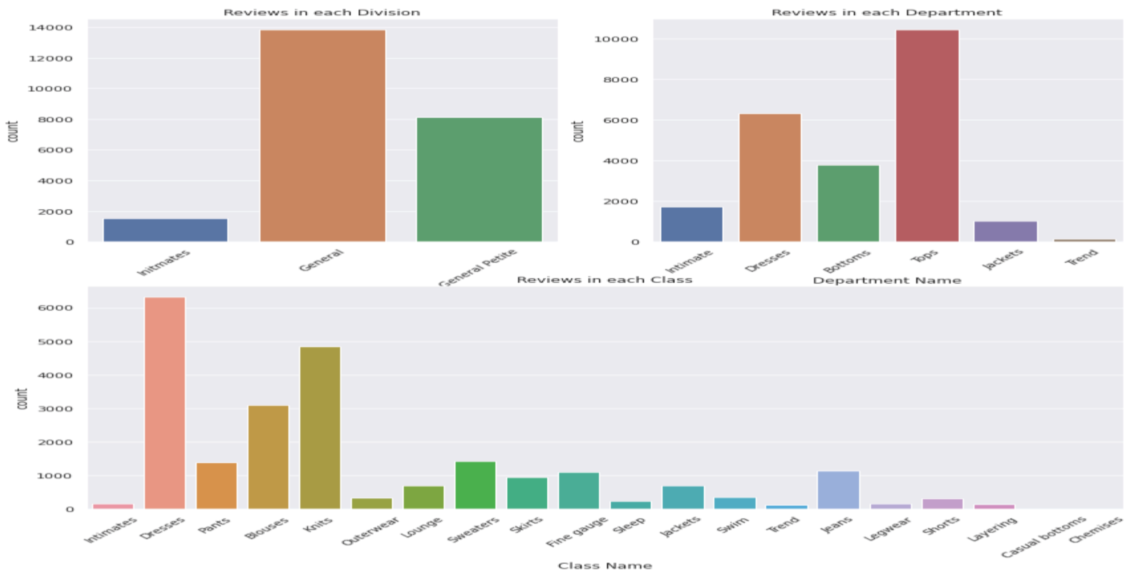


Fig. 3. Reviews Distribution

5.2 Sentiment analysis

Using the insights gained from EDA, we conducted further analysis to address our research questions. We used statistical techniques such as regression analysis to test the significance of different factors affecting clothing ratings. We also used clustering and classification techniques to identify patterns and trends in the data.
Finally, we interpreted the results of our analysis and discussed their implications for e-commerce businesses and consumers. We also highlighted the limitations of our study and provided suggestions for future research.
Overall, our methodology provides a rigorous and systematic approach to exploring and analyzing women's clothing ratings in e-commerce. By employing a range of data cleaning, transformation, and visualization techniques, we were able to gain valuable insights into consumer behavior and preferences, which can inform business strategies and improve customer experiences.
In addition to the exploratory data analysis, we also performed sentiment classification using SVM and logistic regression classification models. The goal was to predict the overall sentiment expressed in each rating, which was labeled as positive, neutral, or negative.
We used text processing techniques [22] such as tokenization, stop-word removal, lemmatization, and factorization to transform the clothing ratings into numerical feature vectors that can be used as input to the classification models. We also performed data normalization to ensure that the features have comparable scales.
Next, we split the data into training and test sets to train and evaluate the classification models. We experimented with different model parameters to obtain the best possible classification performance.
This sentiment classification approach could be extended to other e-commerce domains to help businesses understand customer opinions and improve overall satisfaction.

- ***Customer Sentiment Analysis***

Using natural language processing techniques, we conducted a sentiment analysis of customer reviews to understand their overall sentiment towards the clothing items.

Fig. 4. Words from reviews

- ***Knowledge extraction***

To identify patterns and trends in the data, we conducted clustering analysis to group clothing items based on their attributes. Figure 6 demonstrates the 50 most popular items in this database, this analysis can help us to do best recommendations.

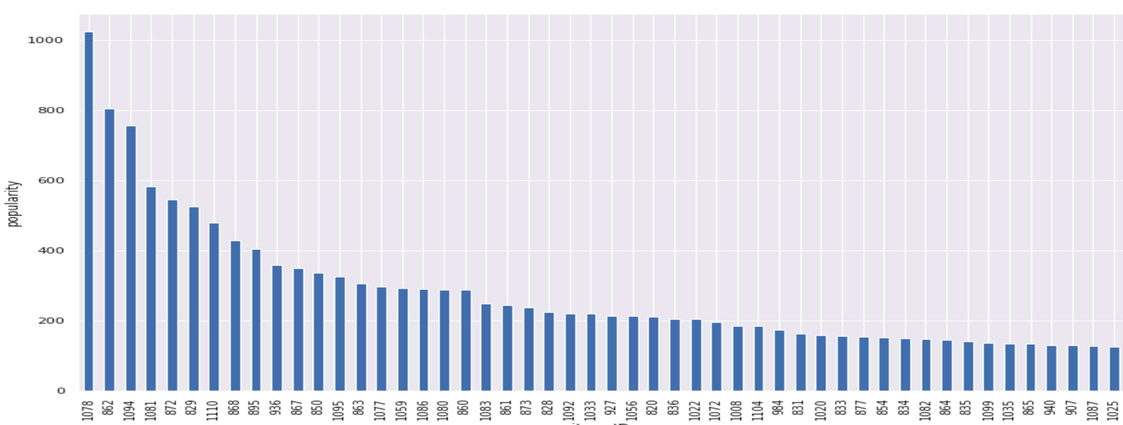


Fig. 5. Top 50 popular Items

The Figure 7 demonstrates the number of reviews per age, this result can help us to cibler the items for the appropriate age.

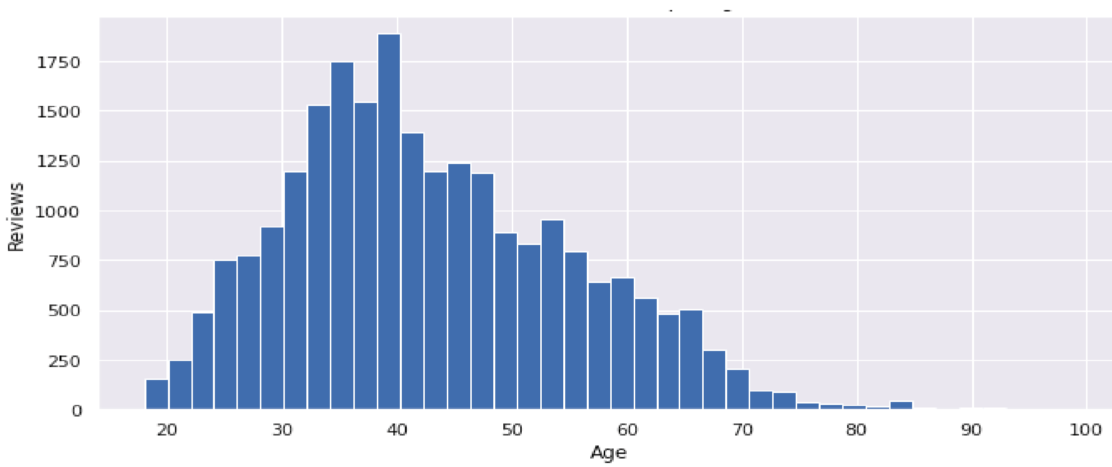


Fig. 6. Number of reviews per age

- ***Logistic Regression & SVM***

The following figure demonstrate the confusion matrix and the true positive rate for SVM and logistic regression models ::

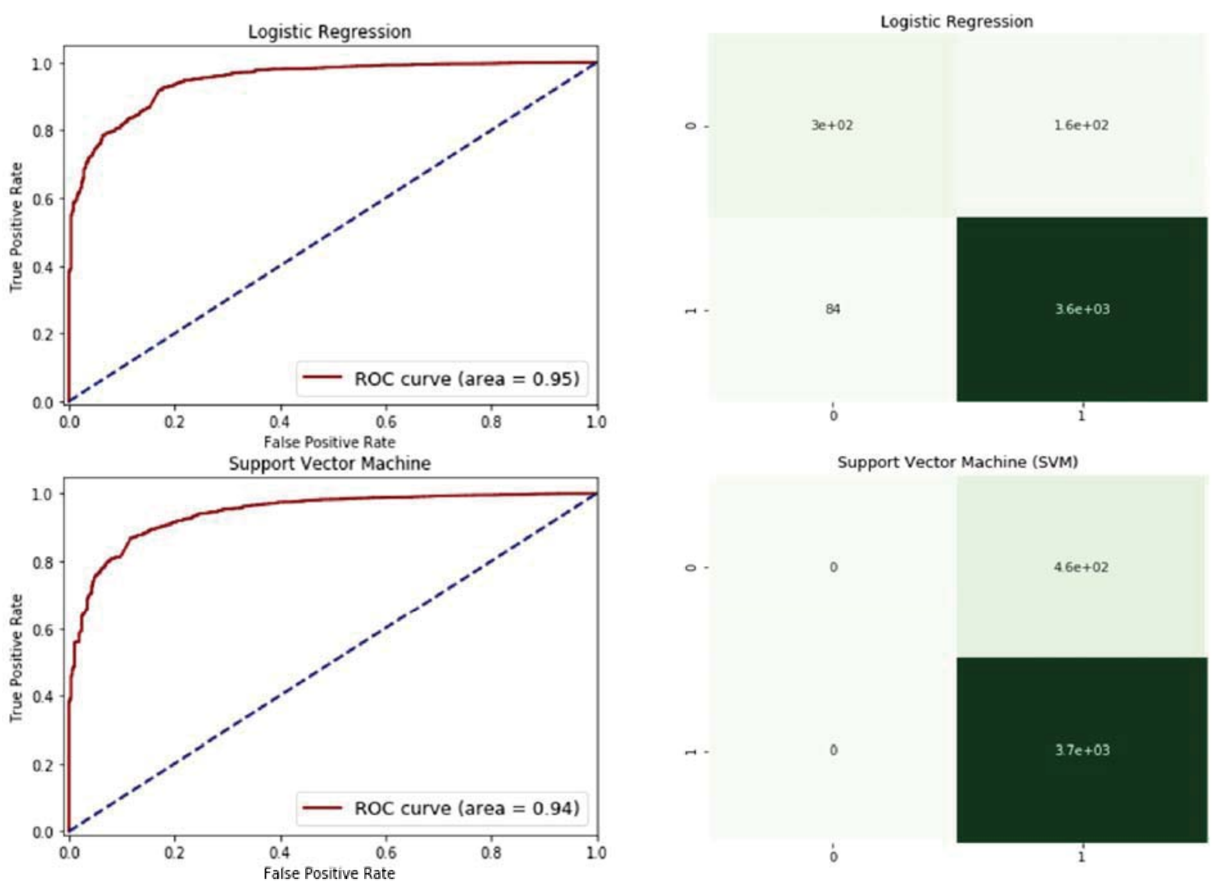


Fig. 7. True Positive Rate

The following tables presents results of the metrics used for the logistic regression and SVM :

Table 1. Logistic regression and SVM results

Logistic Regression				
	Precision	Recall	F1-Score	Accuracy
False	0.78	0.64	0.70	
True	0.96	0.98	0.97	0.95
SVM				
False	0.00	0.00	0.00	
True	0.89	0.99	0.94	0.94

The following tables presents results of the metrics used for Naive Bayes and Neural network used in others studies [23] to compare it with our results:

Table 2. Naive Bayes and Neural network

Naive Bayes				
	Precision	Recall	F1-Score	Accuracy
False	0.70	0.51	0.60	
True	0.88	0.78	0.88	0.88
Neural Network				
False	0.00	0.00	0.00	
True	0.81	0.89	0.84	0.87

5.3 Computational complexity

The SVM and logistic regression models have different computational complexities [23], which can impact the feasibility of using them for sentiment analysis of large e-commerce databases. The SVM model has a time complexity of O(n^2), where n is the number of training examples, making it more computationally expensive than logistic regression. However, SVMs have been optimized to handle large datasets using techniques such as kernel tricks and online learning. The logistic regression model has a time complexity of O(nd), where d is the number of features, making it more computationally efficient than SVM. In addition, logistic regression is easier to implement and interpret than SVM.
In this study, the size of the dataset used for the sentiment analysis of Women's E-Commerce Clothing ratings was sufficiently large, which can result in high computational complexity for the SVM and logistic regression models. To reduce the computational complexity, sparse data representations and stochastic gradient descent were used.

6. Conclusion and future work

In conclusion, our analysis provides a comprehensive exploration of the ratings and reviews of women's e-commerce clothing. Our findings suggest that price, brand recognition, and product categories are important factors in influencing customer satisfaction and ratings. We also highlight the potential limitations of our analysis, such as the limited scope of our dataset and the potential biases in customer ratings and reviews.
Overall, our study contributes to the existing body of knowledge on e-commerce and provides important insights for businesses and consumers alike. Our findings can help businesses better understand the factors that drive customer satisfaction and ratings, and inform their pricing strategies, brand recognition efforts, and product offerings. For consumers, our findings can inform their purchasing decisions and help them make more informed choices based on factors such as price, brand, and product category.
Moving forward, we suggest that future research expand on our findings by incorporating data from multiple e-commerce platforms and conducting more in-depth analyses of customer sentiment and preferences. By continuing to explore the complex factors that influence customer satisfaction and ratings, we can help businesses and consumers navigate the dynamic and ever-evolving world of e-commerce.

References

1. Gasmi, F. Anguel, H. Seridi-Bouchelaghem, N. Azizi, (2021), "Context-Aware Based Evolutionary Collaborative Filtering Algorithm". In: Chikhi S., Amine A., Chaoui A., Saidouni D., Kholladi M. (eds) Modelling and Implementation of Complex Systems. Lecture Notes in Networks and Systems, vol 156, p. 217–232, Springer, Cham. https://doi.org/10.1007/978-3-030-58861-8_16.
2. Gasmi, M.W. , Azizi, H., Seridi-Bouchlaghem, Azizi,N., 2021. "Enhanced Context-Aware Recommendation using Topic Modeling and Particle Swarm Optimization", Journal of Intelligent & Fuzzy Systems, vol. 40, no. 6, p. 12227-12242, June 2021, doi:10.3233/JIFS-210331.
3. Liu, W., Lu, J., & Cui, Y. (2019). Sentiment Analysis of Customer Reviews on Clothing Products Based on Machine Learning Algorithms. Journal of Physics: Conference Series, 1235(1), 012077.
4. Khan, M. S., Amjad, F., Hussain, A., & Afzal, M. K. (2018). Sentiment Analysis of Customer Reviews on Clothing Products using Machine Learning Techniques. In 2018 International Conference on Frontiers of Information Technology (FIT),p. 53-58, . IEEE.
5. Haddi, E., Liu, X., & Shi, Y. (2013). The role of text pre-processing in sentiment analysis. Procedia Computer Science, 17, p. 26-32.
6. Wang, X., Liu, H., & Li, S. (2018). Sentiment analysis of women's clothing reviews based on machine learning. Journal of Computational Science, 28, p. 192-200.
7. Zhang, Y., Chen, X., & Hu, Z. (2020). Sentiment analysis of women's clothing reviews based on SVM. Journal of Ambient Intelligence and Humanized Computing, 11(1), 237-245.
8. Xu, H., Liu, Y., & Lu, Y. (2019). Sentiment analysis of women's clothing reviews using convolutional neural network. In Proceedings of the 2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA), p. 560-569, IEEE.
9. Choi, M., Lee, S., & Kim, S. (2019). Sentiment analysis of women's clothing reviews using character-level and word-level recurrent neural networks. Applied Sciences, 9(22), 4911.
10. Awad, N. F., Krishnan, M. S., & Hassan, M. K., 2017. The impact of e-service quality, customer satisfaction and loyalty on e-marketing: Moderating effect of perceived value. Procedia Computer Science, 116, p. 471-478.
11. Babin, B. J., Darden, W. R., & Griffin, M., 2019,. Exploring the effect of retail service quality on brand image, customer satisfaction and loyalty. Journal of Retailing and Consumer Services, 50, p. 77-85.
12. Chaffey, D., & Smith, P. R., 2017. Digital marketing excellence: planning, optimizing and integrating online marketing. Routledge.
13. Choi, J., & Choi, J., 2018. Social media and fashion brand loyalty: The mediating role of social identification, perceived value, and satisfaction. Journal of Business Research, 89, p. 449-456.
14. Gao, L., Waheed, A., & Wu, W., 2020. The impact of online reviews on the performance of mobile applications: An empirical study. Journal of Retailing and Consumer Services.
15. Hair Jr, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M., 2017. A primer on partial least squares structural equation modeling (PLS-SEM). Sage publications.
16. Hennig-Thurau, T., Malthouse, E. C., Friege, C., Gensler, S., Lobschat, L., Rangaswamy, A., & Skiera, B., 2018. The impact of new media on customer relationships. Journal of Service Research, 21(3), p. 311-330.
17. Huang, H. Y., & Yen, D. C., 2018. The impact of personalized e-service on online customer loyalty. International Journal of Information Management, 38(1), p. 80-90.
18. Kim, S. S., Kim, J., & Lee, H. G., 2017. Effects of product presentation mode on e-commerce websites on consumer perception, emotion, and purchase intention. International Journal of Industrial Ergonomics, 59, p. 8-16.
19. Kulkarni, R., Ravindran, S., & Freeze, R., 2020. Sentiment analysis of online reviews for predicting customer satisfaction. Journal of Retailing and Consumer Services, 53, 101927.
20. Lee, S. H., & Lee, S. Y., 2017. Impacts of online reviews on customer purchasing intention: the mediating effects of trust and perceived usefulness. International Journal of Information Management, 37(2), p. 131-139.
21. Li, H., Liu, Y., Li, G. X., & Li, X., 2018. The impact of website quality on customer satisfaction and purchase intention: Evidence from Chinese online visitors. Information Systems Frontiers, 20(4), p. 729-740.
22. Liao, C., Liu, C. C., & Chen, K. Y., 2019. Effects of online reviews on purchasing intention: the moderating role of need for cognition. Journal of Business Research, 100, p. 164-173.
23. Lu, L. C., Chang, W. P., & Chang, H. H., 2018. Consumer attitude toward brand.

Fashion Product Classifier and Recommendation System Using Machine Learning

Yash Sharma[a], *Satyajit Pangaonkar[b], Rupal Joshi[c], Reena Gunjan[d], Prakash Rokade[e]

[a,b]School of Computer Science and Engineering, Avantika University, Ujjain, Madhya Pradesh 456006, India
[c]School of Fashion Technology and Design, Renaissance University, Indore, Madhya Pradesh 452010, India
[d]Department of Computer Science Engineering, MIT School of Engineering, MIT Art, Design and Technology University, Pune 412201, India
[e]Cloud and Infrastructure Services, Infosys Ltd., Pune 411012, India

Abstract

Fashion e-commerce has become an increasingly popular industry in recent years, and with the vast number of products available, consumers can easily become overwhelmed with choices. A fashion product classifier & recommender system can provide a solution to this problem by accurately categorising products and suggesting personalised recommendations to users. In this paper, a novel approach to fashion product classification and recommendation has been provided using a combination of deep learning techniques and collaborative filtering. The system considers both visual features of the products, such as images and colours, as well as non-visual features such as user preferences and historical behaviour. The experimental results demonstrate the effectiveness of this system in accurately classifying products and providing relevant recommendations to users. The system has the potential to improve user satisfaction and increase sales for fashion e-commerce platforms by providing an efficient and personalised shopping experience. The goal of the system was to improve the user's shopping experience by helping them discover new products and make informed purchasing decisions. The classifier was trained on a large dataset of fashion products and uses deep learning techniques to accurately classify and recommend products with a training & validation accuracy of 94.97% & training & validation loss of 5.03%. The system can be integrated into a fashion retail website or app, providing users with personalised product recommendations in real time.

Keywords: *Convolutional Neural Networks; deep learning; matrix factorization model; feature extraction; collaborative filtering.*

Nomenclature

ResNet	Residual network
NLP	Natural language processing
CNN	Convolutional neural network
DRL	Deep reinforcement learning.
K-NN	K-Nearest Neighbour

1. Introduction

The growth of the fashion industry has led to a massive increase in the number of products available in the market. Customers have access to a wide range of products with varying styles, sizes, and colours. In the old days, fashion product classifiers were typically created using a combination of manual processes and human expertise. This involved individuals with a deep knowledge of fashion and clothing, who would manually sort and categorise items based on their characteristics, such as their materials, colours, patterns, and styles. For example, a team of fashion experts might sort clothing into categories such as dresses, skirts, pants, and shirts, and then further refine those categories by fabric type, colour, and style. They might also create a taxonomy of fashion terms and use that to label and organise products. This process was time-consuming and labour-intensive, and it was also limited by the expertise of the individuals involved. But the abundance of choices has made it challenging for customers to find the right products that match their preferences. To address this issue, researchers have developed various product classification and recommendation systems. These systems help customers to find the products that best fit their preferences and requirements.
Content-based methods were commonly used in fashion product classification to automatically identify and categorise clothing items based on their visual features, such as colour, texture, shape, and pattern. The first step was to gather a large dataset of fashion products, typically images of clothing items, along with their associated metadata such as brand, category, and price. The next step was to extract relevant visual features from the images of clothing items. This could be done using techniques such as convolutional neural networks (CNNs), which were designed to identify and extract visual patterns in images.

* *Satyajit Pangaonkar*
E-mail address: satya_pangaonkar@yahoo.com

Once the visual features had been extracted, the need was to represent them in a way that was suitable for classification. This was typically done by creating a feature vector for each clothing item, where each element of the vector corresponds to a specific visual feature. The final step was to train a machine learning algorithm, such as a support vector machine (SVM) or a random forest classifier, to classify clothing items based on their feature vectors. The classifier could be trained to recognize different categories of clothing, such as tops, bottoms, dresses, or shoes. In summary, content-based methods were an effective way to automatically classify fashion products based on their visual features. By using advanced machine learning techniques to extract and represent visual features, content-based methods could help fashion retailers and e-commerce platforms to categorise and organise large collections of clothing items, making it easier for customers to find the products they were looking for.

Collaborative filtering was a technique used in recommendation systems that analyse the behaviour of users to make personalised product recommendations. In the context of a fashion product classifier and recommendation system, collaborative filtering could be used to suggest clothing items or accessories to a user based on their past behaviour, such as previous purchases or items they have added to their Wishlist. To use collaborative filtering in a fashion product classifier and recommendation system, the system would need to collect data on the user's behaviour, such as their search history, previous purchases, and items they have liked or added to their cart. The system would then use this data to create a user profile, which would be used to identify items that were similar to what the user had already shown an interest in. For example, if a user had purchased several dresses from a specific brand, the system might suggest other dresses from the same brand or similar styles from other brands. Alternatively, if a user had searched for a particular type of shoe, the system might recommend shoes from similar brands or with similar features. Collaborative filtering could also be used in conjunction with other techniques, such as content-based filtering, which analyse the characteristics of the products themselves to make recommendations. By combining these approaches, a fashion product classifier and recommendation system could provide more personalised and accurate recommendations to users, increasing the likelihood that they will make a purchase.

Hybrid methods in machine learning refer to the combination of multiple techniques or models to achieve better performance than using just one approach. In the context of fashion product classification and recommendation systems, hybrid methods were commonly used to combine the strengths of different approaches to improve accuracy and relevance in product recommendations.

One common approach to building a hybrid fashion product classification and recommendation system was to combine collaborative filtering with content-based filtering. Collaborative filtering utilized user behaviour data to recommend products, while content-based filtering analysed product features to make recommendations. By combining these two approaches, the system could recommend products that not only matched a user's preferences based on their behaviour but also matched the features of the products they had shown an interest in. Another example of a hybrid approach in fashion product classification and recommendation was the use of neural networks. Neural networks are powerful machine learning models that can analyse large amounts of data to identify patterns and make predictions. In a fashion product classification and recommendation system, a neural network could be used to analyse product images and identify patterns such as colours, textures, and shapes. This information could then be combined with user behaviour data to make personalised product recommendations. Other hybrid approaches may include incorporating user feedback and preferences into the recommendation system or using clustering algorithms to group similar products and make recommendations based on those groupings. Overall, hybrid methods in fashion product classification and recommendation systems aimed to combine the strengths of multiple approaches to provide more accurate, relevant, and personalised recommendations to users.

This research was useful in improving product recommendations, providing more efficient product classification, and advancement in the field of fashion technology and transferable technology that could be used in other fields with industries as well. Overall, this research paper on fashion product classification and recommendation using deep learning and CNNs can have practical applications for fashion companies, contribute to the advancement of fashion technology, and have wider implications for other fields and industries. The paper is structured as follows. Section 2 elaborates on the literature survey, section 3 explains the methodology, section 4 expounds on the workflow of the system, section 5 describes the results, and Section 6 elaborates on the conclusion & future scope.

2. Literature

Clothing is a kind of symbol that represents internal perceptions of humans through their outer appearance. It conveyed information about their choices, faith, personality, profession, social status, and attitude towards life. Therefore, clothing was believed to be a nonverbal way of communicating and a major part of people’s outer appearance [1]. Recent technological advancements have enabled consumers to track current fashion trends around the globe, which influence their choices. Trend forecasting was well achieved using machine learning and predictive models based on 140 images with 1000 likes in the New York fashion week was collected [2]. It was the logistic regression model that was used to predict the colour, pattern and style based on likes, image frequency and image quality. For this. information retrieval by learning to rank (LEToR) at e-commerce search was the effective strategy identified for feature representation of web search datasets in many applications [3]. The fashion choices of consumers depend on many factors, such as demographics, geographic location, individual preferences, interpersonal influences, age, gender, season, and culture [4]. In the research, the outfit based on accession, current trends and skin-tone colours etc for appropriate incidences were recommended using K means clustering algorithm. Also, it had been seen that the interactive system for fashion clothing recommendation systems was developed to train and generate the appropriate style tags based on body poses, appearances and diverse style attributes [5]. Categorising apparel products were also useful for e-commerce purposes, as well as for avoiding duplicates, and ranking product types. This could be effectively used for online shopping and automatically retrieving similar options from a huge collection. Visual search through Pinterest gave content recommendation for improving user engagement [6]. As an indispensable product for people, clothing had the combined attributes of physiological protection, mobility support, and self-expression [7]. With the rapid economic growth of recent years and the rapid development of clothing-related industries, the supply of clothing products has been greatly enriched [8].

Moreover, previous fashion recommendation research shows that fashion preferences vary not only from country to country but also from city to city [9]. The combination of fashion preferences and the abovementioned factors associated with clothing choices could transmit the image features for a better understanding of consumers' preferences. Therefore, analysing consumers' choices and recommendations was valuable to fashion designers and retailers [10]. Today, in the pursuit of individuality and fashion, people have new requirements for bespoke garment designs that inspire confidence and satisfaction. Consumers want to participate in the customisation process, expressing their needs through interactive customisation and being able to customise the style of their clothing according to their own wishes [11]. Additionally, consumers' clothing choices and product preference data have become available on the internet in the form of text or opinions and images or pictures. Since these images contain information about people from all around the world, both online and offline fashion retailers were using these platforms to reach billions of users who were active on the internet [12]. Using machine learning, a cloth matching scheme experimented with a Siamese network and autoencoder based on labelled data. The literature was found on a weather-oriented clothing recommendation system [13]. It was a weather-to-garment (WoG) data set where the scoring function was analysed for low-level features and high-level weather attributes. Later after comparing with SVM and CNN, it was observed the multiclass classification modelling served as a good approach for the recommendation. Research on e-commerce had become the predominant channel for shopping in recent years. The ability of recommendation systems to provide personalised recommendations and respond quickly to the consumer's choices had contributed significantly to the expansion of e-commerce sales [14].

According to different studies, e-commerce retailers, such as Amazon, eBay, and Shopstyle, and social networking sites, such as Pinterest, Snapchat, Instagram, Facebook, Chictopia, and Lookbook, are now regarded as the most popular media for fashion advice and recommendations [15]. The preference regarding individual items selected for the people having an individual test of items and outfits was researched through the personalised outfit generation (POG) model. Chictopia as the world's largest style community for fashion-forward viewers gave ample information on style posts and online clothing boutiques [16]. Identifying products, a specific customer likes most could significantly increase the earnings of a company [17]. Clearly, recommending suitable products in E-commerce increases the probability of a customer's purchase. Additionally, offering too many products could reduce the probability that a potential customer performs a purchase at all. Finally, knowing and subsequently targeting customer preferences increases the medium and long-term commitment of the customer to the company, which was a key factor in profitability [18]. Cognitive, affective and conative experiences of the customer were identified that contributed to loyalty along with motivational, behavioural and perceptual consequences [19]. This was observed for online to the business-to-consumer contest. Prior studies demonstrate that recommendation engines helped consumers to make better decisions, reduce search efforts and find the most suitable prices [20]. One possibility to infer knowledge about customer preferences was via specific questioning in customer surveys. However, this was not always possible and customer responses may not be correct or sufficient for accurately describing preferences. In this work, we followed a different, data-driven approach, where customer preferences were automatically extracted from available information on the customer. More specifically, it was focused on fashion products and developed a method that only required a single input image to return a ranked list of similar-style recommendations. The literature in an overall way elaborated that an explainable recommendation is an approach that aims to provide explanations for the recommendations made by the system. This could help build trust with the user and improve user engagement with the system. Explainable recommendation methods typically used the techniques such as rule mining, decision trees, or neural attention mechanisms to generate explanations for the recommendations.

Overall, the choice of approach for product classification and recommendation depends on the specific requirements of the application. In some cases, traditional methods such as collaborative filtering or matrix factorization may be sufficient, while in other cases, more advanced techniques such as CNNs or attention mechanisms may be necessary. Hybrid methods that combined different approaches could also be effective in many scenarios. The main aim was to select an approach that was well-suited to the specific problem and to carefully tune the model to achieve optimal performance. In summary, there were many different approaches to product classification and recommendation, each with its strengths and weaknesses. But., the literature gave a limited perspective towards this. Thus, there is still a large scope for the investigation. In this article, a hybrid approach has been evaluated that combined CNNs for product classification with matrix factorization and attention for the recommendation.

3. Methodology

There were many different approaches that could be taken when designing a methodology for classifying and recommending fashion products based on images. Here is the description of a potential methodology taken into consideration. Figure 1 elaborated on the complete process.

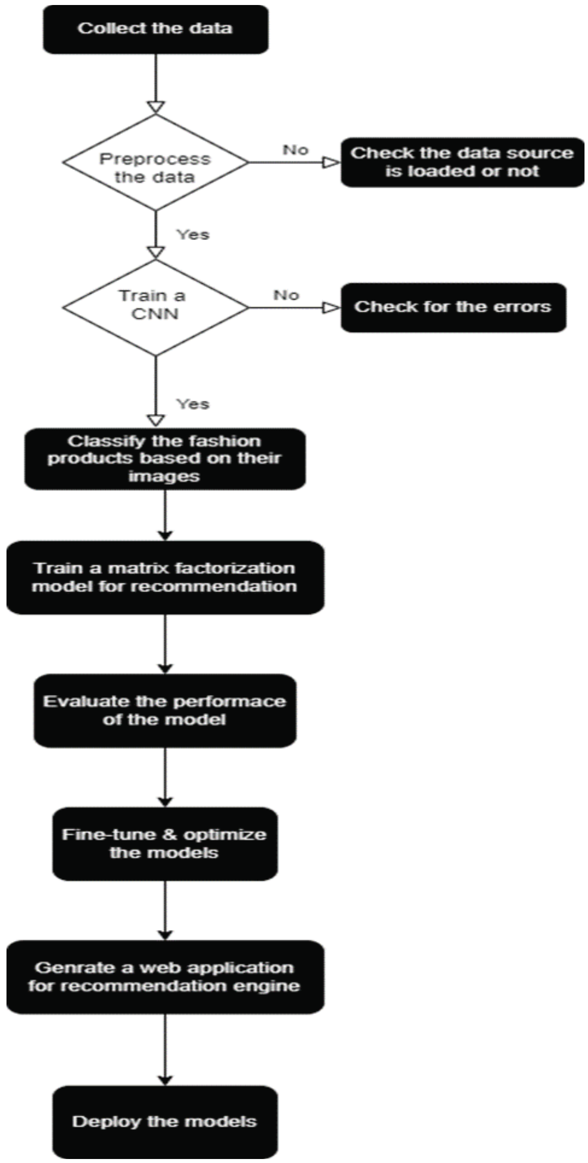


Fig. 1. Process flow

3.1 Data Collection and Pre-processing

The first step was to collect a dataset of fashion product images and associated metadata, such as the product's category, brand, and description. The data should be cleaned and pre-processed as necessary, such as resizing the images to a uniform size and removing any irrelevant or duplicate data. The data cleaning was an iterative process that involved careful inspection of the dataset and making decisions based on the specific context and goals of the analysis. This was done by identifying missing values, removing duplicates, correcting inaccuracies, handling outliers, and finally handling and formatting data. The dataset consisted of the following extracted features and was obtained from the "styles.csv" dataset from Kaggle a fashion product dataset [21]. Following are the unique features of the dataset.

- id: unique identifier for each product
- gender: the gender category (e.g., Men, Women, Unisex)
- master category: the top-level category of the product (e.g., Apparel, Footwear, Accessories)
- subcategory: the subcategory of the product (e.g., Shirts, Casual Shoes, Watches)

- article type: the specific type of the product (e.g., Round Neck T-Shirt, Sports Shoes, Chronograph Watch)
- base colour: the primary colour of the product
- season: the season for which the product is intended (e.g., Fall, Winter, Summer)
- year: the year in which the product was released.
- usage: the intended use of the product (e.g., Casual, Sports, Formal)
- product display name: a descriptive name for the product

The "styles.csv" dataset contained 44,543 rows, where each row represented a single product. The dataset contained 10 columns, one for each extracted feature listed above.

3.2 Splitting data into training and test sets

The next step was to split the data into a training set and a test set. The training set had been used to train the model, while the test set was used to evaluate the model's performance. The training set was then given for feature extraction. In the context of fashion product classification and recommendation, feature extraction involved identifying the most relevant product attributes and transforming them into numerical or categorical variables that could be used as input to a machine learning model.

3.3 Train a convolutional neural network

Once feature extraction was done, training a Convolutional Neural Network (CNN) was to classify the fashion products based on their images. A CNN is a specialised neural network designed for visual data, such as images & videos. But CNNs also work well for non-image data (especially in NLP & text classification). The CNN is designed with appropriate layers and hyperparameters to achieve good performance on the classification task. Its concept is similar to that of a vanilla neural network (multilayer perceptron). It follows the same general principle of forwarding & backward propagation. Once the data was pre-processed, the neural networks were trained, utilising transfer learning from ResNet50. More additional layers were added in the last layers that replace the architecture and weights from ResNet50 in order to fine-tune the network model to serve the current issue. The CNN should be trained on the training set, using the product categories as the class labels.

3.4 Train a matrix factorization model for recommendation

In addition to classifying the products, it was necessary to be able to recommend similar products to users. To do this, matrix factorization models could be trained on the training set. This model would learn the relationships between the products and could be used to make recommendations based on the similarity between products.

3.5 Evaluate the performance of models

Once the CNN and matrix factorization models had been trained, their performance should be evaluated on the test set. This could be done using metrics such as accuracy for the classification task and precision and recall for the recommendation task.

3.6 Fine-tune and optimise models

If the performance of the models was not satisfactory, the models could be fine-tuned and optimised by adjusting their hyperparameters and adding or removing layers. The models should then be re-evaluated on the test set to see if the performance had improved.

3.7 Deploy models

Once the models were performing well, they could be deployed in a fashion retail website or app to classify and recommend products to users in real time. This was just one example of a potential methodology for classifying and recommending fashion products based on images. There were many other approaches and techniques that could be used, depending on the specific goals and constraints of this project.

4. Workflow of the system

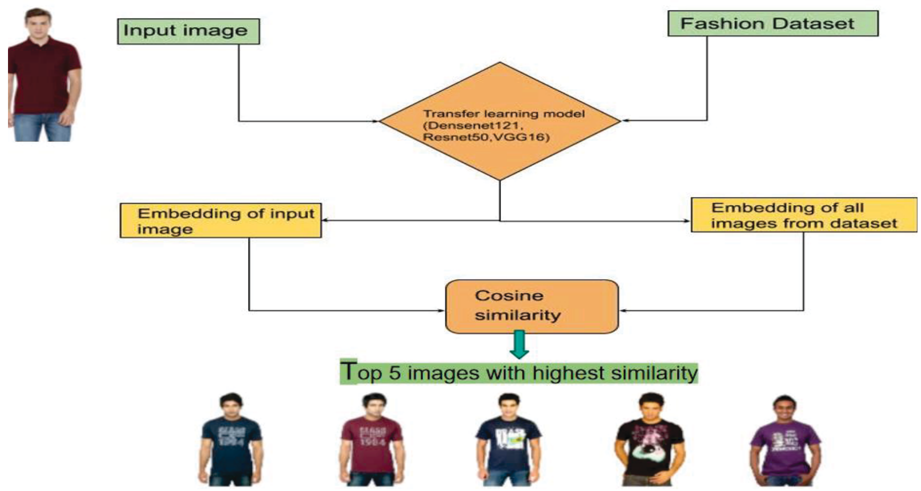


Fig. 2. Workflow of the recommendation system

To generate recommendations, the proposed approach uses Sklearn Nearest Neighbours. This allowed to find the nearest neighbours for the given input image. The similarity measure used in this project was the Cosine Similarity measure. The top 5 recommendations were extracted from the database and their images were displayed. Figure 2 explained the entire workflow of the recommendation system.

5. Results

The CNN classifiers were used as feature extractors and returned feature vectors Fi (X) of size d = 1024 for any input image. The feature extractors were applied to a set of n = 44.4k test images. These corresponding feature vectors were concatenated and stacked to obtain an n × 2d feature matrix. The k-NN ranking algorithm was applied to the feature matrix. For the recommendation task, it was now sufficient to extract the features from an input image, submit them to the k-NN ranking algorithm and return the top-k matching style recommendations. An implicit objective metric for recommendation quality was observed by means of the classification accuracies reported in Figure 3 with a training & validation accuracy of 94.97% with a training and validation loss of 5.03%. The definition of precise objective evaluation criterion, however, remained difficult due to the inherent subjectivity of recommendation quality. This also made a comparison with other methods quite challenging. The computationally most time-consuming part of the application of the proposed recommendation system was the evaluation of the CNN classifiers.
In this research implementation, it was implemented the CNNs in MXNet using its Python API. Running on a desktop PC with an Intel i7-6850K CPU and an NVIDIA 108990 Ti GPU, the whole image processing pipeline applied to a given input image only takes fractions of a second. Note that the potentially time-consuming network training was done before a new input image was provided to the recommendation system, which therefore allowed fast online product recommendation.

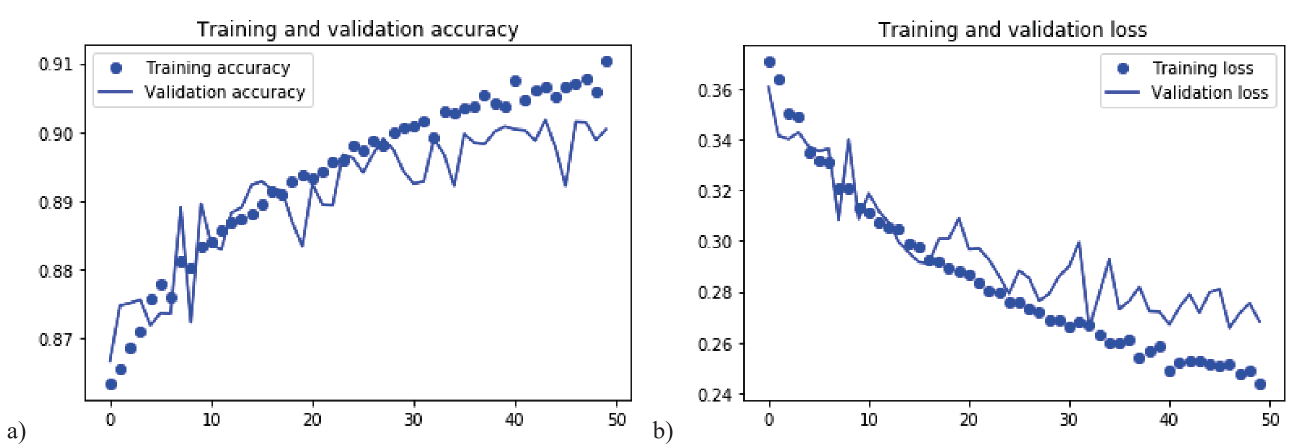


Fig. 3. Results of a) Training & validation accuracy and b) Training & validation loss

In Figure 4. presented several query images and corresponding top-5 recommendations on web-based applications. Subjectively, the top-5 recommendations indeed look quite similar to the query images. In the top row, a query image from the dataset itself was used. These corresponding top-5 recommendations demonstrated that if the image appeared in the dataset, it was actually most similar to itself. Similar results had been obtained in other performed tests. The proposed solution approach used a CNN classifier to extract features that were used as input for similarity recommendations.

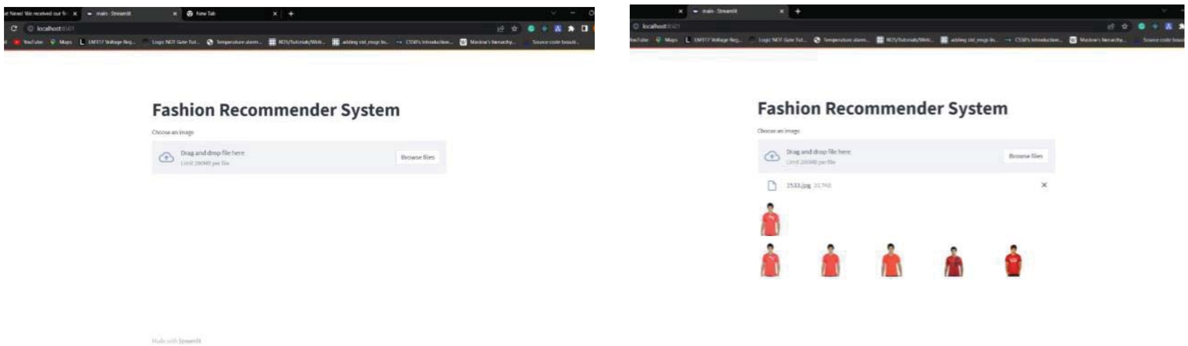


Fig. 4. Generating recommendations through the web application (based on user input)

The approach could be used in E-commerce by allowing customers to upload a specific fashion image and then offering similar items based on texture and category type features of the customer's uploaded image. Additional feature extractors, e.g., trained on gender or colour classification tasks, can be easily added. Furthermore, generalisation to other domains makes sense, e.g., music recommendation based on raw music data, but needs further investigation. Several interesting extensions of our approach were possible. First, it would be promising to integrate the two separate training stages into a single one and provide end-to-end deep learning-based fashion product recommendations. In particular, consideration should be given to Siamese networks. Additionally, hybrid approaches combining image-based and content-based systems will be implemented. Finally, it was important to evaluate the customer impact of our image-based approach and its extensions against other recommender systems through customer surveys.

6. Conclusion and Future Scope

A fashion product classifier and recommender system can have many potential applications and future scope. Some potential discussion points and future directions for such a system can include improving the accuracy and diversity of the classifier. Currently, the classifier may not be able to accurately classify all types of fashion products, or it may not generate a diverse set of recommendations. In the future, the system could be improved by adding more training data, fine-tuning the model architecture, and implementing techniques such as data augmentation to better handle rare or unusual product types. The system could be enhanced by incorporating additional information about the user's preferences, style, and body type, as well as constraints such as budget and available sizes. This could allow the system to make more personalised and relevant recommendations.

The user interface and user experience of the system could be improved to make it easier for users to browse and search for products, and to provide more helpful and visually appealing recommendations. This could involve incorporating features such as visual search, 360-degree product views, and interactive outfit builders. The system could be integrated with other platforms and systems, such as e-commerce websites, social media, and fashion blogs, to provide a more seamless and comprehensive fashion shopping experience. This could involve using APIs to access and combine data from these different sources or building custom integrations. The system could be expanded to cover a wider range of fashion products, such as accessories, makeup, and beauty products, or to provide additional functionality such as styling tips and trend forecasts. The system could also be adapted for use in other domains, such as home decor or automotive, by training the classifier on appropriate data.

References

1. Barnard, M., 2003. Fashion as Communication, Routledge, Taylor and Francis Group, London, U.K.
2. Chakraborty, S., Hoque, S.M.A., Kabir, S.M.F., 2020. Predicting fashion trend using runway images: Application of logistic regression in trend forecasting, International Journal of Fashion Design Technology Education 13, pp. 376-386
3. Karmaker Santu, S.K., Sondhi, P., Zhai, C., 2017. On application of learning to rank for e-commerce search, Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval, Shinjuku, Tokyo, Japan, pp. 475–484.
4. Garude, D., Khopkar, A., Dhake, M., Laghane, S., Maktum, T., 2019. Skin-tone and occasion-oriented outfit recommendation system, SSRN Electronic Journal 8, p.34.

5. Sachdeva, H., Pandey, S., 2020. Interactive Systems for Fashion Clothing Recommendation, Emerging Technology in Modelling and Graphics 937, pp. 287–294.
6. Jing, Y.; Liu, D.; Kislyuk, D.; Zhai, A.; Xu, J.; Donahue, J.; Tavel, S. 2015. Visual search at Pinterest. In Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Sydney, Australia, 10–13 August 2015; pp. 1889–1898.
7. Lee, K., Choi, H., Choi J. H., Kim, T., 2019. Development of a data-driven predictive model of clothing thermal insulation estimation by using advanced computational approaches, Sustainability 11, p. 5702.
8. Majumdar, A., Sinha, S. K., Shaw, M., Mathiyazhagan, K., 2021. Analysing the vulnerability of green clothing supply chains in South and Southeast Asia using fuzzy analytic hierarchy process, International Journal of Production Research 59, pp. 752–771.
9. Matzen, K., Bala, K., Snavely, N., 2017. StreetStyle: Exploring world-wide clothing styles from millions of photos. arXiv 2017, arXiv:1706.01869.
10. Guan, C., Qin, S., Ling, W., Ding, G., 2016. Apparel recommendation system evolution: An empirical review. International Journal of Cloth Science Technology 28, 854–879.
11. Jung, W. K., Kim, H., Park, Y. C. Lee, J. W., Suh, E. S., 2020. Real-time data-driven discrete-event simulation for garment production lines, Production Planning & Control 33, pp. 480–491.
12. Gao, G., Liu, L., Wang, L., Zhang, Y., 2019. Fashion clothes matching scheme based on Siamese Network and Auto Encoder, Multimedia Systems 25, pp. 593–602.
13. Liu, Y., Gao, Y., Feng, S., Li, Z., 2017. Weather-to-garment: Weather-oriented clothing recommendation, Proceedings of the 2017 IEEE International Conference on Multimedia and Expo. (ICME), Hong Kong, China, pp. 181–186.
14. Chakraborty, S., Hoque, M.S., Surid, S.M., 2020. A comprehensive review on image-based style prediction and online fashion recommendation. Journal of Modern Technology and Engineering 5, pp. 212–233.
15. Chen, W., Huang, P., Xu, J., Guo, X., Guo, C., Sun, F., Li, C., Pfadler, A., Zhao, H., Zhao, B., 2019. POG: Personalized outfit generation for fashion recommendation at Alibaba iFashion, Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, pp. 2662–2670.
16. Street Style Fashion. Available online: http://www.chictopia.com/browse/people (accessed on 16 March 2023).
17. Schafer, J.B., Konstan, J.A., Riedl. J., 2001. E-commerce recommendation applications, Data Mining and Knowledge Discovery 5, pp. 115–153.
18. Dick, A.S., Basu, K., 1994. Customer loyalty: toward an integrated conceptual framework, Journal of Academy of Marketing Science 22, pp.99–113.
19. Srinivasan, S.S., Anderson, R., Ponnavolu, K., 2002. Customer loyalty in e-commerce: an exploration of its antecedents and consequences, Journal of Retailing 78, pp.41–50.
20. Häubl, G., Murray, K. B., 2006. Double agents: assessing the role of electronic product recommendation systems, MIT Sloan Management Review, pp. 8–12.
21. Agarwal., P., 2020, Fashion Product Images Dataset, Kaggle, Available online https://www.kaggle.com/datasets/paramaggarwal/fashion-product-images-dataset (accessed on Jan. 2023).

Authors Information

Yash Sharma is currently pursuing his B. Tech degree in Computer Science and Engineering from Avantika University, Ujjain M.P., India. He is currently in the sixth semester of graduation and is holding the position of a member of the Google Developers Student Club at the university. His research interests include data science, machine learning, artificial intelligence, and full-stack development.

Satyajit Pangaonkar received the M. Tech. degree in Electronics and Telecommunication from Dr Babasaheb Ambedkar Technological University, Raigad, India. He is pursuing a PhD from the Department of Electronic and Communication Engineering, School of Engineering, MIT Art, Design and Technology University, Pune, India. He has 34 years of experience in research and teaching. His research interests include speech signal processing, artificial intelligence and machine learning.

Rupal Joshi received the M.Sc. degree in Fashion Technology from Savitribai Phule Pune University, Pune, India. She is currently holding the position of head in the School of Fashion Technology and Design, Renaissance University, Indore, India. She is also a President of the MP handloom council (WICCI) and running her own ethnic wear brand ‘MAUU’. She has 15 years of experience in fashion technology. Her research interest includes traditional textiles and crafts.

Reena Gunjan received her Master of Engineering in Control Systems and a PhD in Computer Engineering from Malviya National Institute of Technology, Jaipur, India. She is currently holding a position of a Professor in the Department of Electronics and Communication Engineering at the School of Engineering, MIT Art, Design & Technology University, Pune, India. She has 34 years of experience in research and teaching. Her research interests are image, speech signal processing and machine learning.

Prakash Rokade received his bachelor’s degree in Electronics and Power Engineering from the Government College of Engineering, Amravati, Maharashtra, India. He is currently holding the position of Principal Consultant in Infosys Ltd., Pune, India. His area of expertise is Signal Processing, Artificial Neural Network Digital communication, and Network and Security. He has 22 years of experience in industrial research and consultation.

Breaking Barriers: An Applications for the Empowerment of Deaf Individuals

Nihalkumar Ashokkumar Galani[a],Amanbhai Arifbhia Shaikh[b], Shrey NileshKumar Hingu[c], Dr. Narendra M Patel[d], Mosin I Hasan[e]

[a,b,c] UG Student, Department of Computer Engineering
Birla Vishvakarma Mahavidyalaya, Vallabh Vidyanagar 388120,, India
[d] Professor, Department of ComputerEngineering
Birla Vishv akarma MahavidyalayaVallabh Vidyanagar, Gujarat, India
[e] Assistant Professor, Department of Computer Engineering
Birla Vishvakarma Mahavidyalaya Vallabh Vidyanagar, Gujarat, India

Abstract

This research paper presents an android application designed to empower deaf individuals by providing them with tools to enhance their communication skills. The application contains four modules, including a tutorial on American Sign Language (ASL), four different levels of sign language quizzes, a module that translates text to sign language using images, and 3D model animations, and a real-time ASL detection module. The tutorial module provides an interactive learning experience for beginners to learn the fundamentals of American Sign Language. The quiz module challenges users to test their knowledge of sign language at different levels of difficulty. The text to sign language modules allow users to translate written English text into sign language either through static images or 3D model animations, providing a more immersive experience. The real-time ASL detection module is a feature of the application that allows users to practice their signing skills in real-time. It uses a combination of computer vision and machine learning techniques to detect and interpret ASL gestures, providing immediate answer to the user about what they are signing. The application's goal is to empower the deaf community to communicate more effectively, providing a more inclusive and accessible environment. The application's modules were evaluated and tested for usability, functionality, and effectiveness. The results showed that the application was effective in enhancing communication skills and could serve as a valuable tool for empowering the deaf community.

Keywords: Android Application; Sign recognition; American Sign Language; Tutorials, Quizzes; Text to sign language using Images; 3D model animations; blender studio; The real-time ASL detection; Machine Learning.

1. Introduction

The deaf community faces many challenges when it comes to communication, often feeling excluded and isolated from mainstream society due to the lack of understanding and access to sign language (Sign languages also known as signed languages are languages that use the visual-manual modality to convey meaning, instead of spoken words [1]). As technology advances, there is a growing need to develop innovative and accessible tools to support and empower the deaf community.

In response to this need, this research paper presents an Android application aimed at enhancing the communication skills of the deaf community by providing them with tools to learn and use American Sign Language (American Sign Language (ASL) is a natural language that serves as the predominant sign language of Deaf communities in the United States of America and most of Anglophone Canada [2]) more effectively. The application comprises four modules, including a tutorial for ASL, a quiz module with four different levels, a text to sign language module that uses images, and 3D model animations, and the real-time ASL detection module which allows users to practice their signing skills in real-time. These modules were designed to provide a comprehensive and engaging learning experience for deaf individuals, regardless of their proficiency level. The study evaluates the effectiveness and usability of the application's modules through a combination of qualitative and quantitative methods, providing insights into how mobile technology can be used to support and empower the deaf community. The results of this study can inform the development of future applications and technologies that aim to bridge the communication gap between hearing and deaf individuals, creating a more inclusive and accessible society for all[3][4][5].

* Dr. Narendra M Patel
E-mail address: nmpatel@bvmengineering.ac.in

2. Related work

American Sign Language (ASL) is a unique language that is primarily used by the deaf community. As technology continues to advance, there has been an increasing interest in developing applications that aid in the learning and use of ASL. This paper presents an application that consists of four modules aimed at facilitating the learning and practice of ASL.

Several ASL tutorial applications have been developed to teach learners the basic signs of ASL. For example, "ASL Fingerspelling" (Grieve, 2012) is an application that teaches users how to finger spell the alphabet in ASL. Another application, "ASL Coach" (Cooper, 2014), provides users with a series of video tutorials and quizzes to help them learn and improve their ASL skills. However, none of these applications have included quizzes at different levels, real-time ASL detection module or text to sign language translation.

Quizzes are an effective way of reinforcing learning, and several quiz-based applications have been developed for ASL learners. For example, "ASL Word Search" (Stokes, 2019) is a game-based application that challenges users to find words in a grid of ASL hand signs. Another example, "ASL Quiz" (Mithun, 2018), provides users with multiple-choice questions to test their knowledge of ASL vocabulary. However, these applications do not include text to sign language translation using images or 3D models.

Text to sign language translation is a challenging task, and several approaches have been proposed to address this problem. One approach is to use images to represent signs, such as "ASL Dictionary" (ASLized!, 2021), which provides users with a collection of images representing different ASL signs. Another approach is to use animations of 3D models to represent signs, such as "Sign Language 3D" (SignGenius, 2021), which uses a virtual 3D model to demonstrate ASL signs. However, none of these applications have included quizzes at different levels or a comprehensive tutorial.

In summary, although there are few applications available no one contains all of the mention functionality in one module, our application is designed to provide an all-in-one solution for ASL learners, allowing them to learn, practice and translate ASL using one cohesive platform.

3. Detailed description of each module

3.1 Tutorial Module:

This module includes tutorials for different types of fields. It provides a solid foundation for learners who are new to American Sign Language (ASL). It teaches basic concepts such as the alphabet, numbers, and simple vocabulary, which are fundamental to understanding and communicating in ASL. Without a strong foundation, learners may struggle to progress in the language.

3.1.1 Implementation:

There is total 8 different fields of tutorial (Fig. 3.2 a) so user can select any field any can see images along with their text and can learn sign language. Moreover, to list images we have used recycler view which is most modern, highly customizable and effective way to list items in android[6]. In addition to this we use image slider so that it will provide better user interface to user.

3.1.2 Overall workflow of tutorial module:

The tutorial module in our app provides users with an interactive and engaging way to learn about different signs related to 8 fields (Fig.3.1 a). Upon navigating to the tutorial module, users can select the field they want to learn about. Once a field is selected, users are presented with a range of different signs related to that field, each accompanied by an image and text description. To make it easy for users to navigate through the different signs, an image slider is included with back and next buttons (Fig.3.1 b). This allows users to quickly move between images and learn about each sign in a structured and organized way. Overall, this tutorial module is designed to provide users with a comprehensive understanding of different signs related to different fields, in a way that is intuitive and easy to use.

a b

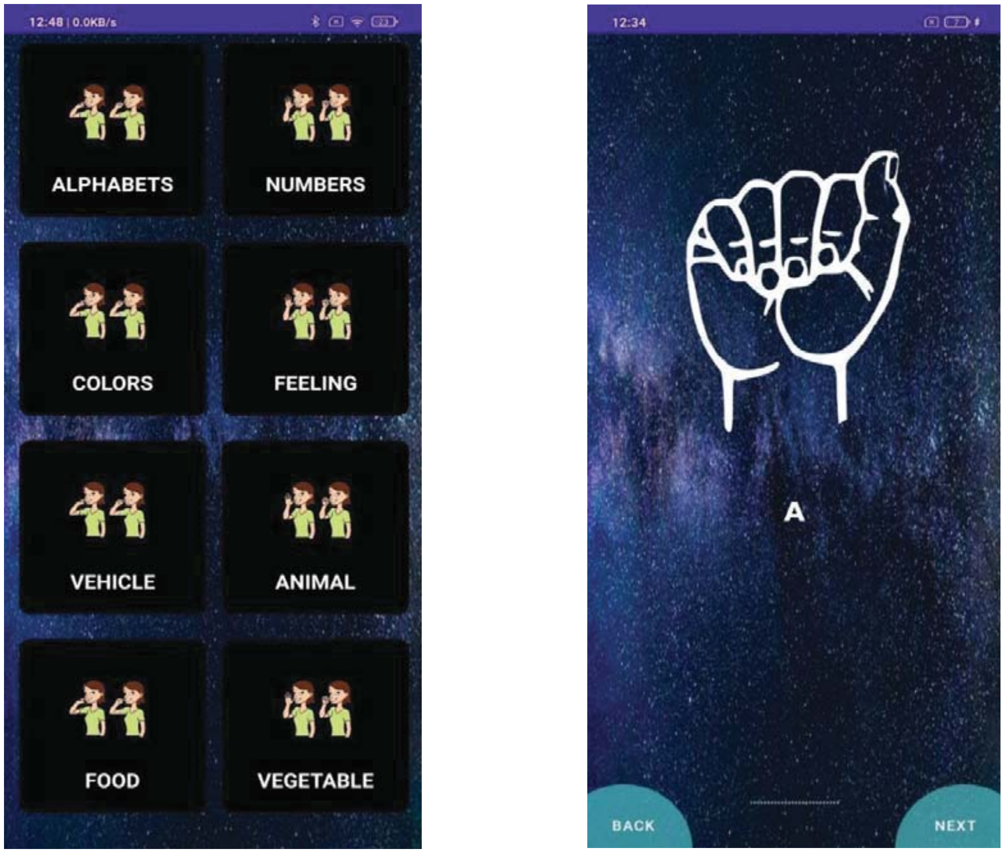


Fig. 3.1: a. Tutorials of different fields with image slider, b. Images in alphabets section

3.2 Quiz Module:

This module includes the quizzes of different levels[7] to review the learning of students.

3.2.1 Module Importance:

It is very difficult for newbie learner to learn a sign language and after learning the sign language there are very limited resources for assessment of their learning.

3.2.2 Implementation:

There is total 4 different levels of quiz (Fig. 3.2) so user can select any level can assess their knowledge of sign language. Moreover, to list images we have used recycler view which is most modern, highly customizable and effective way to list items in android. In addition to this we use image slider so that it will provide better user interface to user. If user select wrong answer, then user will get correct answer after selecting the wrong answer from this, they can increase their knowledge.

3.2.3 Overall workflow:

The quiz module in our app provides users with an interactive way to test their knowledge and skills. Users can navigate to the quiz module and select one of four different levels, with level 1 being the easiest and level 4 being the hardest (Fig. 3.2). Once a level is selected, users are presented with a range of different questions and answer options. If the user selects the correct answer, a green sign appears on the option, while a red sign appears on the selected option if the user selects the wrong answer. A green sign also appears on the correct answer option in this case (Fig. 3.3). Once the level is completed, the user's score is displayed. Overall, this quiz module is designed to challenge users and help them improve their knowledge and skills, with clear feedback on their performance.

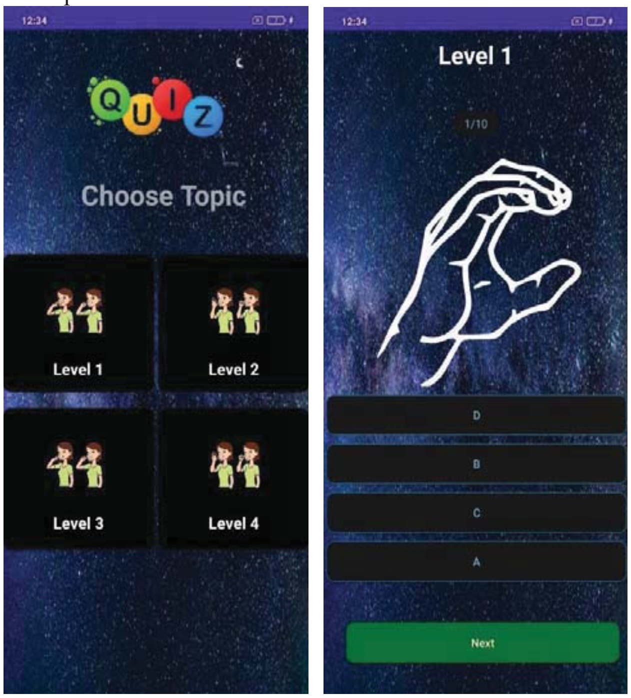


Fig. 3.2: Different quiz levels with interface

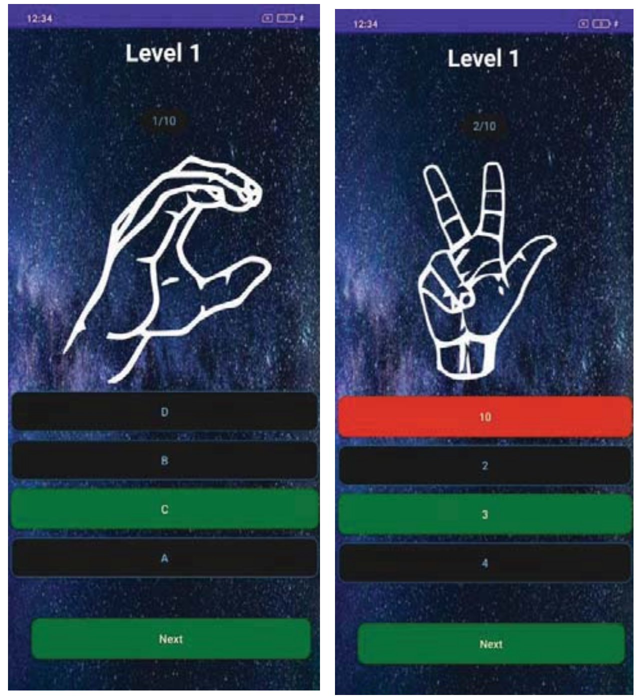


Fig. 3.3: Right and wrong answer in quiz

3.3 Text to Sign Module:

This module converts text entered by the user into sign language. It has a huge potential in announcement system of railway platform, airports, and bus stations where display devices can be installed to play all announcement in sign language through the usage to proposed text to sign language generation system.

3.3.1 Implementation:

We implemented this module in two ways:

1. Conversion of text to sign language images.
2. Conversion of text to sign language animation using 3D model.

3.3.2 Text to sign language using images:

This implementation takes input text from the user and gives series of images as output. We have only entertained alphabets and digits characters and rest of the characters are neglected. Moreover, to list images we have used recycler view which is most modern, highly customizable and effective way to list items in android.

3.3.3 Text to sign language using images overall workflow:

Our app provides a text-to-image conversion feature that allows users to convert any text into sign language. To use this feature, the user navigates to the text to images module and enters the desired text in the edit text field (Fig.3.4). Upon hitting the convert to sign language button, the input string is processed and each word is separated by spaces. Only alphabets and digits are considered as valid characters within each word. The words are then displayed in separate recycler view blocks, with each block containing images of all valid characters in the word (Fig.3.4). To enhance the user interface, we have customized the recycler view to provide an aesthetically pleasing experience for our users. Overall, our text-to-image module is a useful and visually appealing feature that allows users to easily convert text into sign language images.

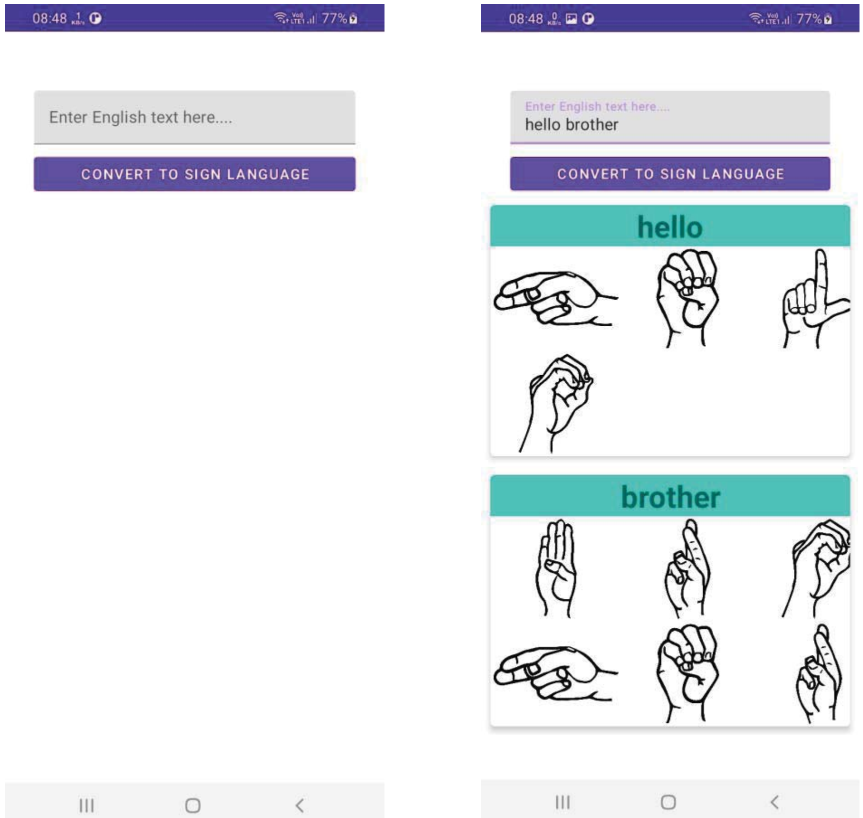


Fig. 3.4: Text to sign language translation example (Using images)

3.3.4 Text to sign language animation:

This implementation takes input text from the user and gives series of animations as output. We have only entertained alphabets and digits characters and rest of the characters are neglected.

For this module we selected Rain (name of the 3d model) (Fig.3.5) as our 3D model provided by the blender[8] to create animations. Rain is a generic character rig you can use to learn and practice your animation skills, created by us here at the Blender Studio.

Figure 3.5: Rain model in blender

Steps to animate rain model:

To create an animation using Blender, the first step is to download Blender studio and the rain model. Once downloaded, turn on the IK settings for hand movement and fingers settings for finger movement. To determine the length of the animation, add the desired number of frames in the bottom timeline window. To create specific hand positions for the animation, change the position of the hands on specific frames as needed. This can be done by selecting pose mode in the mode of the 3D model and then using the move option to change the position of the hands and fingers, and the rotate option to adjust their orientation. With these steps, users can create complex animations in Blender using their own custom hand and finger movements to achieve the desired effect.

3.3.5 Text to sign language animation overall workflow:

Our app features a text-to-3D model conversion module that allows users to convert any text into sign language. To use this feature, the user navigates to the text to 3D model module and enters the desired text in the edit text field(Fig.3.6). Upon hitting convert to sign language, the input string is processed and only valid characters, which include alphabets and digits, are considered for conversion. For each valid character, an animated video is played using the video view, and all videos are stored in the raw folder. This allows users to easily convert text into sign language animations without any complex modeling or animation skills. The text-to-3D model module is a useful and user-friendly feature that provides a simple way to create sign language animations of custom text in our app.

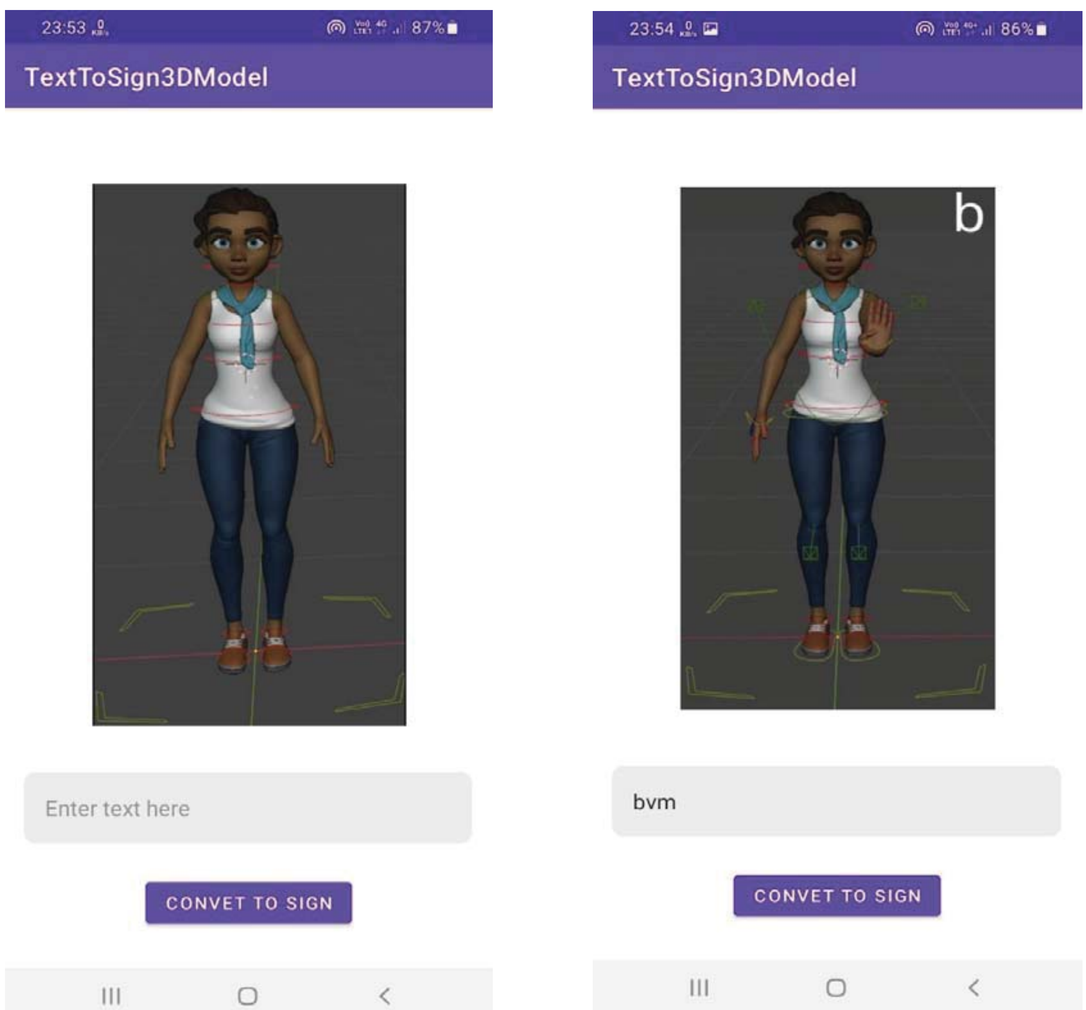


Fig. 3.6: Text to sign language translation example (Using animation)

3.4 Real time sign language detection module:

This module allows a project to recognize and interpret sign language in real-time. This could be useful in providing real-time captions for deaf and dumb individuals.

3.4.1 Implementation:

Implementation of this module is done in two parts as follows
1. Training of machine learning model
2. Integration of trained module in Android

3.4.2 Training of machine learning model:

In order to create machine learning model, we need dataset which include large number of images for each sign language alphabets along with this the images should have some amount of difference like images in dark light and light area, images captured from different distance and with different rotation.

We have used dataset from Kaggle [9] which has more than 11,400 images (Fig. 3.7) which means 450+ images for each alphabet. This dataset full fill all the requirements mentioned above.

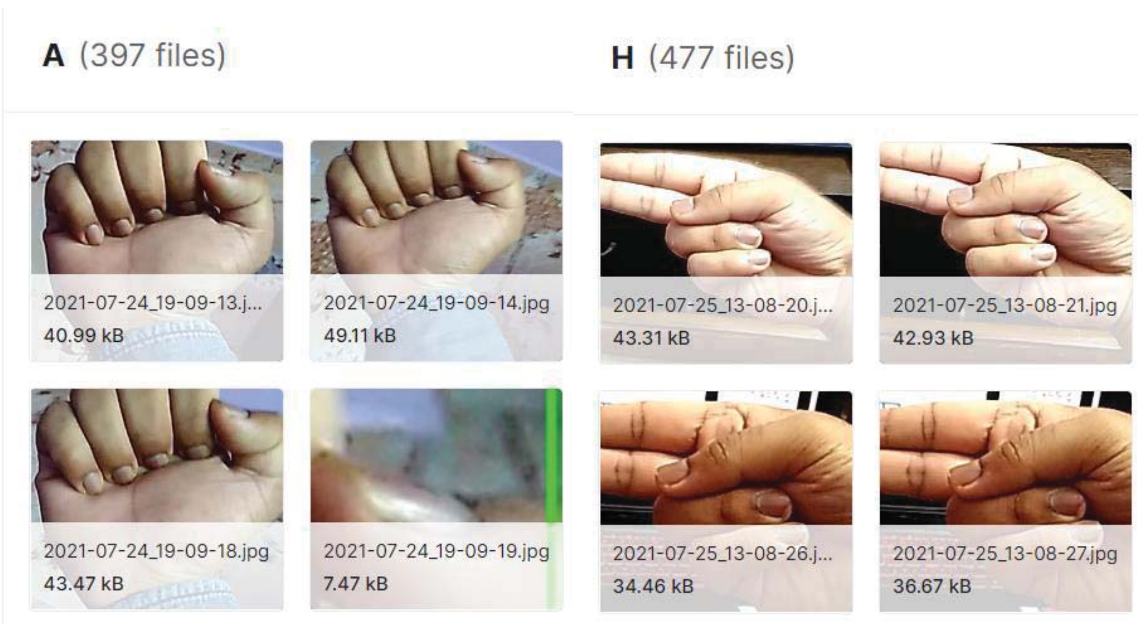


Fig. 3.7: Dataset examples (Alphabets A and H)

```
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 efficientnetb0 (Functional)  (None, 4, 4, 1280)       4049571

 global_average_pooling2d (G  (None, 1280)             0
 lobalAveragePooling2D)

 dropout (Dropout)           (None, 1280)              0

 dense (Dense)               (None, 1)                 1281

=================================================================
Total params: 4,050,852
Trainable params: 4,008,829
Non-trainable params: 42,023
_________________________________________________________________
```

Fig. 3.8: Deep CNN model with transfer learning

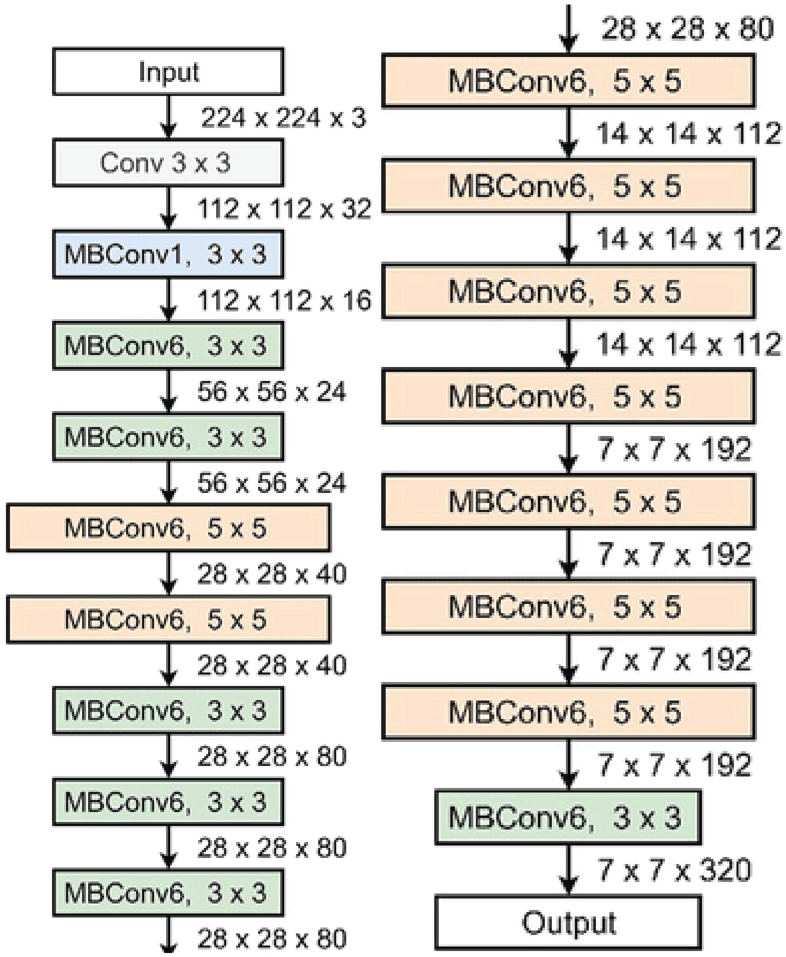


Fig. 3.9: Efficientnetb0 neural network architecture

```
[53] Model_Results = model.evaluate(Test_Set)
     print("LOSS:  " + "%.4f" % Model_Results[0])
     print("ACCURACY:  " + "%.2f" % Model_Results[1])

     132/132 [==============================] - 12s 87ms/step - loss: 4.7739e-06 - accuracy: 1.0000
     LOSS:  0.0000
     ACCURACY:  1.00
```

Fig. 3.10: Accuracy of trained model

3.4.3 Integration of trained model in android:

Now to use the model in android first we need the OpenCV library for android which is useful for computer vision related tasks. Here OpenCV uses the native camera of android device to capture image of every frame until user closes the application.

Here to integrate model we have to convert model in to TFlite model because android uses TensorFlow lite library to integrate machine learning model into android [10][11] .

Here in dataset the images are only consist hands but in image we get from camera has hand and the other background therefore we have to use pretrained hand detection model to detect only hand for the prediction.

3.4.4 Overall workflow:

Our app offers a real-time sign language module that allows users to communicate through sign language. Upon navigating to the module, the camera captures every frame and sends it to a predefined hand detection model, which detects the user's hands and adds padding to it before returning their positions. The application then converts the relevant part of the image into a bitmap as per the returned position. This new image, which only includes the hands, goes through our defined model, which then gives an output. The application then converts this output into an alphabet, which is displayed on the screen(Fig.3.11). This real-time sign language module provides users with a convenient way to communicate through sign language in real-time, without requiring advanced sign language skills.

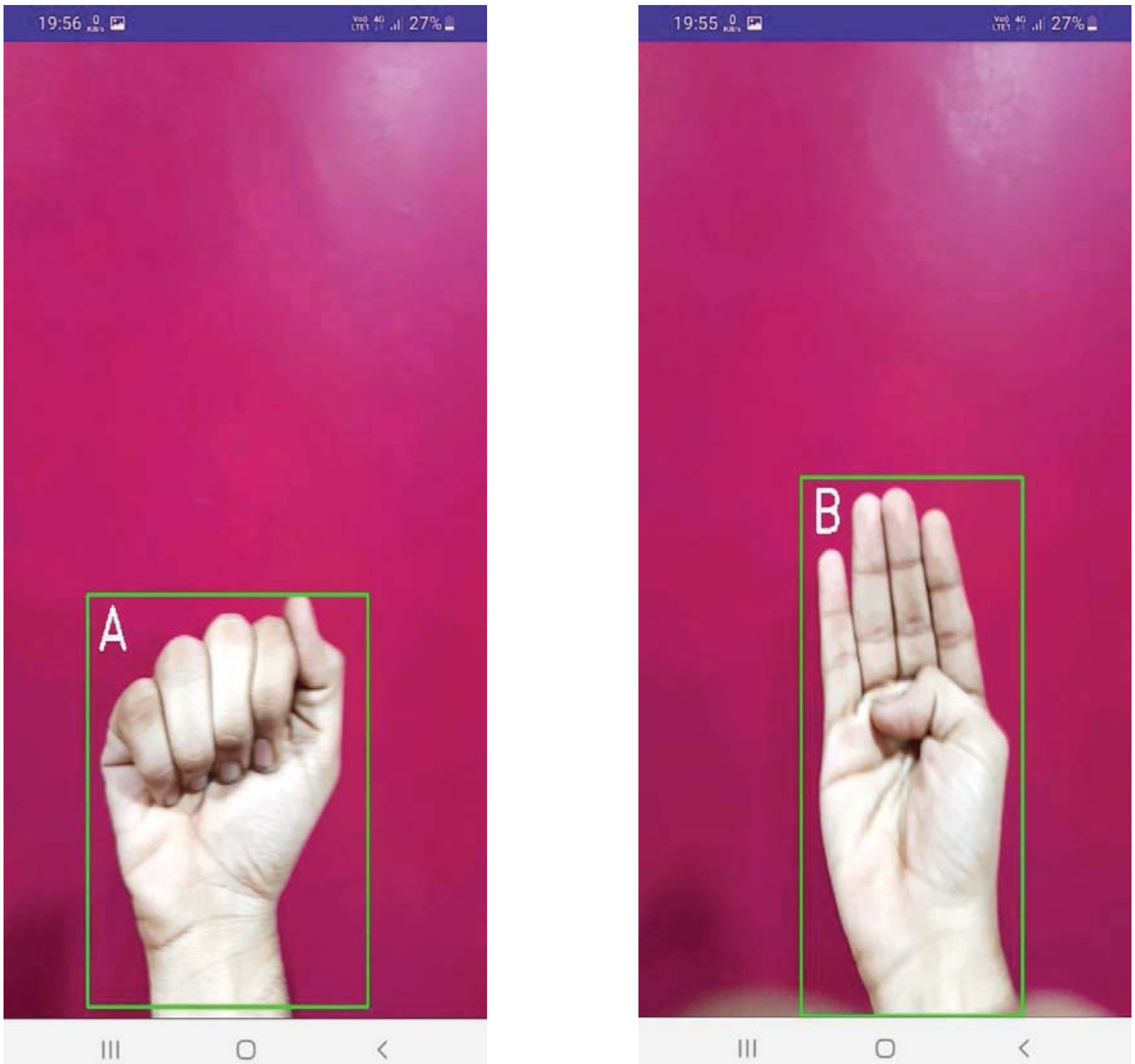


Fig. 3.11: Result of real time sign language detection

4. Conclusion

In conclusion, this paper presented Full Fledge solution to for educational app for deaf people. This app, helps the student as well as people we want to learn sign language with tutorials, quizzes, Text to sign language using images and animation, Hand detection. Tutorial module includes tutorials for different types of fields. There is total 8 different fields of tutorial like alphabets, numbers, fruit, colors, feelings etc. so, user can select any field any can see images along with their text and can learn sign language. Quiz module includes the quizzes of different levels to review the learning of students. It is very difficult for newbie learner to learn a sign language and after learning the sign language there are very limited resources for assessment of their learning. To solve this problem, we have created quiz module with 4 difficulty levels. Text to sign language using module converts text entered by the user into sign language. This implementation takes input text from the user and gives series of images as output. Text to sign language using animation. This implementation takes input text from the user and gives series of animations as output. To generate animation of alphabets and digits we have user blender studio. In which, we have used rain 3d model to show animations. Real time sign language detection. This module allows a project to recognize and interpret sign language in real-time. This could be useful in providing real-time captions for deaf and dumb individuals. This module opens camera and first detects hand then it uses our trained module to predict the hand gesture.

References

- Sign language articles:

1. Information about what is sign language: https://en.wikipedia.org/wiki/Sign_language
2. Information about what is "American Sign Language". https://en.wikipedia.org/wiki/American_Sign_Language

- Journal articles:

3. AN ANDROID APPLICATION TO AID UNEDUCATED DEAF-DUMB PEOPLE
   Dalia Nashat et al, International Journal of Computer Science and Mobile Applications,Vol.2 Issue. 9, September- 2014: https://www.researchgate.net/publication/273205378_AN_AN
4. Digital Sign Language App for Deaf Children
   International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
   Volume 6, Issue 1, June 2021: https://ijarsct.co.in/Paper1396.pdf
5. Nirvatha Vadathi- An App to Assist Deafand Dumb
   International Journal of Engineering Research & Technology (IJERT)
   NCCDS - 2020 Conference Proceedings: https://www.ijert.org/nirvatha-vadathi-an-app-to-assist-deaf-and-dumb

- Learning references:

6. Developer guides (Android documentation): https://developer.android.com/docs
7. Developing Quizzes(Quiz tutorial): https://www.youtube.com/playlist?list=PLjVLYmrlmjGdDps6HAwOOVoAtBPAgIOXL
8. Blender (Blender tutorial): https://www.blender.org/support/tutorials
9. Sign language dataset: https://www.kaggle.com/datasets/pramod722445/sign-language-dataset
10. Machine learning model deployment: https://www.analyticsvidhya.com/blog/2021/11/how-to-deploy-machine-learningml-model-on-android/
11. Machine learning with Android tutorial:
    https://www.youtube.com/playlist?list=PL0aoTDj9NwgjLj4zei6P1jZH8sQ5YRpQKAcknowledgement

Acknowledgment

We would also like to extend our thanks to the principal of school “Badhir Vidhyalaya”, located in Nadiad Gujarat, who graciously welcomed us into her school and provided us with a thorough understanding of how to develop the quiz module in different levels. Her insights were extremely helpful in designing a module that would be engaging and effective for users of all skill levels.

We are thankful for sponsoring conference registration fee from recurring amount of RPS tilled “A vision based dynamic gesture recognition of Indian Sign language using Kinect” funded by AICTE.

Experimenting Sensor-based Effective Energy Saving Module for Household Electricity Consumption

Rinku Chavda[a*], Sohil Pandya[b]

[a]Research Scholar, Gujarat Technological University, Ahmedabad, Gujarat, India
[b]Department of Computer Applications, CMPICA, CHARUSAT, Charotar University of Science and Technology (CHARUSAT), CHARUSAT campus, Changa 388421, India

Abstract

Modern technology gives better solutions to busy working people towards automation. Nowadays, houses, buildings, and organizations are equipped with a variety of appliances like computers, laptops, switches, doorbells, Ovens, Televisions, LED lighting, Water heaters, AC, etc. It becomes difficult to monitor all the electric equipment at a time. For example, sometimes if no one is present at home & some appliances remain on, it will consume electricity and generate unnecessary amounts to pay in electricity bills. Another scenario is Environmental effects on your power consumption. For example, based on temperature, humidity and light intensity of your surrounding atmosphere, automatic monitoring and controlling of appliances can be achieved with the help of Internet of Things (IOT). Based on the idea of monitoring and controlling appliances, authors have presented the Sensor based Electricity Saving Module.

Keywords: Internet of Things; Energy Saving; Light Sensor, Motion Sensor.

1. Introduction

Internet of Things (IoT) is a smart technology that interconnects each and every object (thing) through a network in one form or another and the main goal of IoT is to maximize the communication of hardware objects with the physical world and to convert the data collected by these objects into useful information without interacting with humans. The presence of IoT in the smart metering area has the potential to transform residential houses, homes and offices into management and controlling of energy, particularly in smart home solutions. Authors have presented a systematic methodology to incorporate Internet of Things based solutions into the home automation, considering both home electricity managing and controlling for various appliances. Various sensors such as motion sensor, temperature-humidity sensor, light sensors are helpful to get the information about the presence of the people in the house, what is the current temperature inside the house as well as how much the light intensity available naturally? Interfacing of sensors with Raspberry Pi 3 board, the managing and controlling of electric equipment i.e., turning ON or OFF can be implemented to achieve the goal of saving electricity and it also helps in managing the monthly budget of consumers in terms of finance. The rest of this paper is systematized as follows. Section 2 examines the work related to managing and controlling the household devices. Section 3 describes required equipment and its hardware interfacing for implementing Electricity Saving Module, and Section 4 presents methodology used for decision making of electric equipment using sensors. Section 5 represents experimental results. Finally, Section 6 concludes with major outcome of proposed methodology.

2. Literature Survey

Several researchers have proposed different types of work based on Internet of Things for monitoring and controlling home or building appliances which are as follows. Many Researchers have proposed a system design which empowers users to remotely monitor and control home devices and feature of online bill generation, view different type of data analytics which shown monthly power consumption of house, annual power consumption of community, state and country were also proposed with the help of mobile application [1]. Arduino based system has been proposed to optimize power of home appliances that can be monitored and controlled with the help of a temperature sensor, Wi-Fi module and web application [6]. Author has presented the system of a home administration and security framework utilizing Arduino and Internet of Things innovation for ongoing home security checking and for remotely controlling the home machines and assurance from flame mishaps with quick arrangements [2]. With the idea of information sharing of smart devices using IoT [10] have proposed a home model to demonstrate an energy efficient IOT based smart home to monitor and control smart devices using motion sensor and controls the kitchen area to monitor HVAC system. Author intended to provide how Internet of Things is useful in electricity competence applications in a view of its mechanical and industry impacts and what are the opportunities and risks for the diverse promote players [4]. Researchers have proposed a smart home control system using a coordinator based on ZigBee networking [5]. An IoT framework with smart

*Rinku Chavda
E-mail address: rinku.chavda.12@gmail.com

location-based automated and networked energy control proposed in [7], which uses Smartphone platform and cloud-computing technologies to enable multi-scale energy proportionality including building-, user-, and organizational-level energy proportionality. A structure plan of building electricity monitoring and investigation based on the Internet of things, which has some enlightening in building electricity consumption added to realize real-time monitoring and controlling, and improve the electricity saving of intellectual building [11].

3. Required Equipments and its Hardware interfacing

While implementing the Electricity saving module, various sensors have been used to get the information such as presence of natural light in room or house, temperature, presence of humans in room and different types of jumper wires to connect said sensors with the programming board. Detailing of the sensors such as purpose of sensor and board for reading sensor values and performing decision making of turning ON/OFF of equipments has been listed in Table 1. Hardware Interfacing of Raspberry Pi 3 with Sensors and LEDs using jumper wires were shown in fig. 1 as well as fig. 2 represents detailed connection of Raspberry Pi 3 GPIO pins with sensors.

Table 1. List of Equipments used in Electricity Saving Module

Sr. No.	Name of Equipment	Purpose
1	Raspberry Pi Model 3	Program/Logic building related to used sensors
2	DHT11 Sensor	Measuring Temperature and Humidity in the Room
3	Motion Sensor	Detecting presence of Person in the Room
4	Light Sensor Module	Measuring light intensity in the Room
5	5 LEDs	Represents One AC, 3 Lights and One Fan which are available in the Room
6	Jumper Wires	Connect various Sensors and LEDs to Raspberry Pi 3 GPIO Pins
7	Breadboard	Used to build a circuit to demonstrate its action

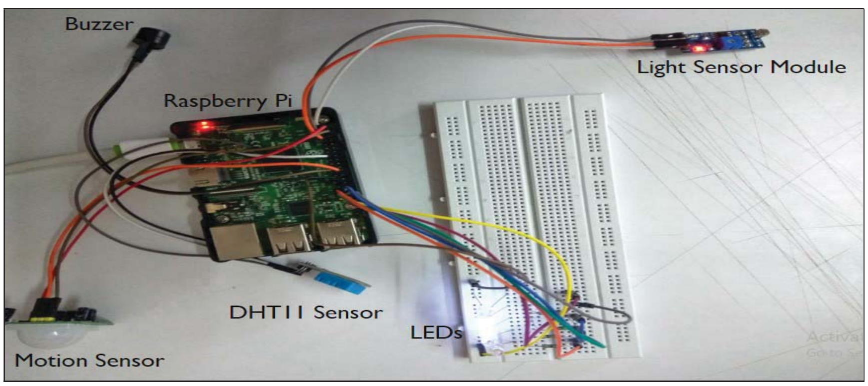
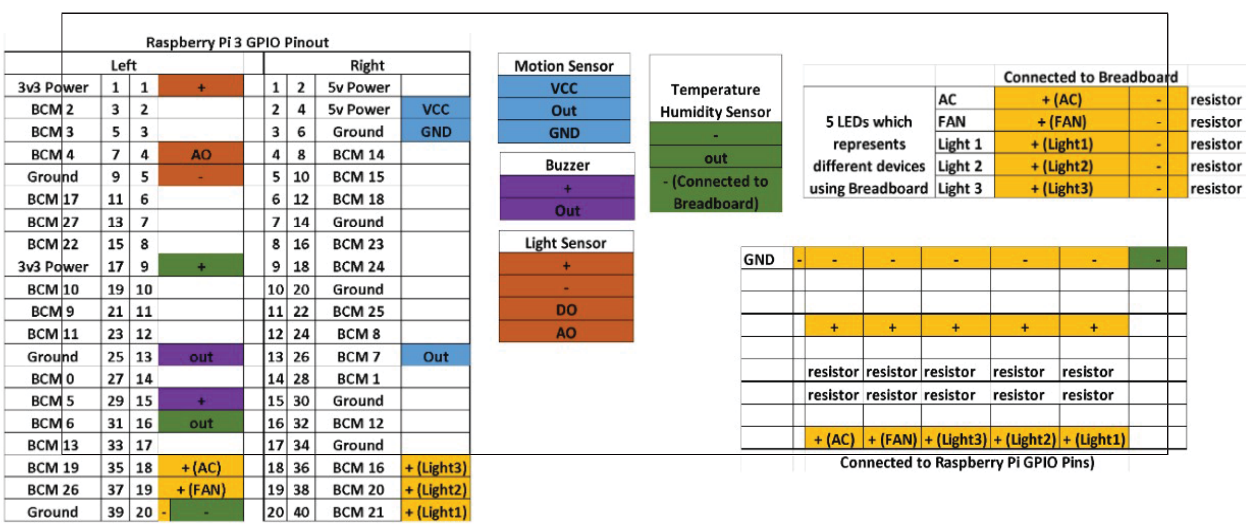


Fig. 1. Hardware representation of electricity saving module.


Raspberry Pi 3 GPIO Pinout

Left							Right	
3v3 Power	1	1	+	1	2	5v Power	
BCM 2	3	2		2	4	5v Power	VCC
BCM 3	5	3		3	6	Ground	GND
BCM 4	7	4	AO	4	8	BCM 14	
Ground	9	5	-	5	10	BCM 15	
BCM 17	11	6		6	12	BCM 18	
BCM 27	13	7		7	14	Ground	
BCM 22	15	8		8	16	BCM 23	
3v3 Power	17	9	+	9	18	BCM 24	
BCM 10	19	10		10	20	Ground	
BCM 9	21	11		11	22	BCM 25	
BCM 11	23	12		12	24	BCM 8	
Ground	25	13	out	13	26	BCM 7	Out
BCM 0	27	14		14	28	BCM 1	
BCM 5	29	15	+	15	30	Ground	
BCM 6	31	16	out	16	32	BCM 12	
BCM 13	33	17		17	34	Ground	
BCM 19	35	18	+ (AC)	18	36	BCM 16	+ (Light3)
BCM 26	37	19	+ (FAN)	19	38	BCM 20	+ (Light2)
Ground	39	20	- -	20	40	BCM 21	+ (Light1)

Fig. 2. Detailing of Raspberry Pi 3 GPIO pins with sensors.

4. Methodology for implementation of Electricity saving using IoT

In this section, Methodology to implement electricity saving module using Internet of Things have been discussed. The sensors periodically perform sensing and send (wirelessly or wired) sensed data to the Raspberry Pi 3. While reading the values of sensors, some threshold value has been fixed (Section 4.2) which will help to take decision regarding when to turn ON or OFF of electric equipment i.e., Lights in room, Fan, AC (discussed in section 4.3) as well as calculation of consumption with and without sensors (discussed in section 4.4).

4.1 Understanding of Sensors with respect to Electricity Saving Module

- *Motion Sensor:* Using the Motion Sensor, it is possible to check whether the Person is present or not in the Room. If a Person is not present in the Room for more than 5 minutes and he/she forgot to turn off all the ON devices of the Room, the program will automatically turn OFF all the ON devices which help us to save our power consumption. As soon as the Person gets into the Room, the sensor will detect "Person is back in the Room", then the program will turn ON all the devices as per the previous state when he/she left the Room.
- *DHT11 Sensor:* With the help of a Temperature and Humidity sensor, we can measure temperature and humidity in the Room, and based on the recorded data, the program will switch ON/OFF AC and FAN which leads to save power consumption.
- *Light Sensor Module:* With the help of the Light Sensor module, it becomes easy to get light intensity in the Room, and based on the recorded intensity, it will switch ON/OFF Lights which are in the Room.

4.2 Threshold values of Sensor readings

How to make decisions regarding the turning ON and OFF of electric equipment based on the sensor values? To get the answer of said question, some limit value or threshold values need to be finalized which will help to achieve efficiency of consumption. Threshold values of each sensor have been set as per the need of light intensity in room, cooling required in room, person is not present in room which were listed in table 2 as following.

Table 2. List of Sensors with Threshold values

Sensors	Threshold Values
Temperature and Humidity level	tempThreshold1 = 26, tempThreshold2 = 25, tempThreshold3 = 22
	humidityThreshold1 = 60, humidityThreshold2 = 45
Light Intensity	lightThreshold1 = 100, lightThreshold2 = 200, lightThreshold2 = 300
Time Duration	timeThreshold1 = 5 minutes
	timeThreshold2 = after few minutes
	delayThreshold1 = 1 minute

4.3 Algorithm: Decision making to turn ON/OFF the Electric equipments based on threshold values

1. START
2. Turn **ON** Required devices suggested by Person
3. Repeat while **True**
   a. Store device status into database
   b. Read **Temperature** and **Humidity** Data
      IF (**Temperature** level is greater than tempThreshold^1 and **Humidity** level is less than humidityThreshold^1)
      THEN Turn **ON AC** and **OFF FAN**
      Else IF (**Temperature** level is less than tempThreshold^2 and **Humidity** is less than humidityThreshold^2) or (**Temperature** level is less tempThreshold^3 and **Humidity** level is greater than humidityThreshold^2)
      THEN Turn **OFF AC** and **FAN**
      ELSE Turn **OFF AC** and **ON FAN**
   c. Read Light Intensity Data
      IF **Intensity of Light** is greater than lightThreshold^1
      THEN Turn **ON All the LIGHTS**
      ELSE IF **Intensity of Light** is between lightThreshold^1 and lightThreshold^2

THEN Turn **ON Two Lights** which consumes less power
ELSE IF **Intensity of Light** is between lightThreshold[2] and lightThreshold[3]
THEN Turn **ON One Light** which consumes less power
ELSE Turn **OFF all the LIGHTS**

d. Read Human Presence Data
IF **No ONE is present** in the ROOM
IF No One is present in the ROOM for **more than** timeThreshold[1]
THEN Turn **OFF ALL** the devices which are ON
Store the status i.e., OFF of ALL the devices into database
ELSE IF **Someone is present** in the ROOM
THEN Turn **ON the devices based on the Sensor data**
Store the status of turned ON devices into database
ELSE IF **Someone Come back** in the ROOM timeThreshold[2]
THEN Turn **ON the devices based on the previously stored status of Devices**

e. Time Delay for delayThreshold[1].

4. STOP

4.4 Algorithm: Calculate Electricity Consumption and find difference with and without help of IoT

1. START
2. **DEFINE Watts/Hour** for each Device.
3. **PERFOM** following steps for both types of Data i.e., Consumption with and without using IOT
   3.1. **Fetc4h** total usage of ALL Device data in minutes.
   3.2. **CALCULATE** power consumption for each device.
   3.3. **CALCULATE** total power consumption of user.
   3.4. **PRINT** power consumption of each device
   3.5. **PRINT** total power consumption of user.
4. **FIND** "difference of total power consumption of user" for both the type of Data i.e., Consumption with and without using IOT.
5. **PRINT** "Electricity Saving using IOT in KWs for recorded Time Period".
6. STOP

id	datetime	ac	light1	light2	light3	fan	light_intensity	temp	humidity
2856	2018-12-22 15:35:10	0	1	1	1	1	134	21.0	52.0
2857	2018-12-22 15:36:10	0	1	1	1	1	133	21.0	51.0
2858	2018-12-22 15:37:11	0	1	1	1	1	133	21.0	52.0
2859	2018-12-22 15:38:12	0	1	1	1	1	126	21.0	51.0
2860	2018-12-22 15:39:14	0	1	1	1	1	129	21.0	51.0
2861	2018-12-22 15:40:16	0	1	1	1	1	121	21.0	52.0
2862	2018-12-22 15:41:17	0	1	1	1	1	134	21.0	52.0
2863	2018-12-22 15:42:19	0	1	1	1	1	133	21.0	52.0
2864	2018-12-22 15:43:21	0	1	1	1	1	139	21.0	52.0
2865	2018-12-22 15:44:23	0	1	1	1	1	137	21.0	52.0
2866	2018-12-22 15:45:25	0	1	1	1	1	138	21.0	52.0
2867	2018-12-22 15:46:27	0	1	1	1	1	131	21.0	51.0
2868	2018-12-22 15:47:29	0	1	1	1	1	133	21.0	51.0
2869	2018-12-22 15:48:31	0	1	1	1	1	119	21.0	51.0
2870	2018-12-22 15:49:33	0	1	1	1	1	127	21.0	51.0
2871	2018-12-22 15:50:35	0	1	1	1	1	123	21.0	51.0
2872	2018-12-22 15:51:37	0	1	1	1	1	132	21.0	51.0
2873	2018-12-22 15:52:39	0	1	1	1	1	133	21.0	51.0
2874	2018-12-22 15:53:40	0	1	1	1	1	135	21.0	51.0
2875	2018-12-22 15:54:42	0	1	1	1	1	134	21.0	51.0
2876	2018-12-22 15:55:44	0	1	1	1	1	117	21.0	51.0
2877	2018-12-22 15:56:46	0	1	1	1	1	133	21.0	51.0
2878	2018-12-22 15:57:48	0	1	1	1	1	130	21.0	51.0
2879	2018-12-22 15:58:50	0	1	1	1	1	130	21.0	51.0
2880	2018-12-22 15:59:52	0	1	1	1	1	126	21.0	51.0
2881	2018-12-22 16:00:54	0	1	1	1	1	121	21.0	51.0

Fig. 3. (a) Represents Usage of AC, FAN and Lights without Sensors.

id	datetime	ac	light1	light2	light3	fan	light_intensity	temp	humidity
3025	2018-12-22 15:35:10	0	0	1	1	0	134	21.0	52.0
3026	2018-12-22 15:36:10	0	0	1	1	0	133	21.0	51.0
3027	2018-12-22 15:37:11	0	0	1	1	0	133	21.0	52.0
3028	2018-12-22 15:38:12	0	0	0	0	0	126	21.0	51.0
3029	2018-12-22 15:39:14	0	0	0	0	0	129	21.0	51.0
3030	2018-12-22 15:40:16	0	0	0	0	0	121	21.0	52.0
3031	2018-12-22 15:41:17	0	0	0	0	0	134	21.0	52.0
3032	2018-12-22 15:42:19	0	0	0	0	0	133	21.0	52.0
3033	2018-12-22 15:43:21	0	0	0	0	0	139	21.0	52.0
3034	2018-12-22 15:44:23	0	0	0	0	0	137	21.0	52.0
3035	2018-12-22 15:45:25	0	0	0	0	0	138	21.0	52.0
3036	2018-12-22 15:46:27	0	0	0	0	0	131	21.0	51.0
3037	2018-12-22 15:47:29	0	0	0	0	0	133	21.0	51.0
3038	2018-12-22 15:48:31	0	0	0	0	0	119	21.0	51.0
3039	2018-12-22 15:49:33	0	0	0	0	0	127	21.0	51.0
3040	2018-12-22 15:50:35	0	0	0	0	0	123	21.0	51.0
3041	2018-12-22 15:51:37	0	0	0	0	0	132	21.0	51.0
3042	2018-12-22 15:52:39	0	0	0	0	0	133	21.0	51.0
3043	2018-12-22 15:53:41	0	0	0	0	0	135	21.0	51.0
3044	2018-12-22 15:54:42	0	0	0	0	0	134	21.0	51.0
3045	2018-12-22 15:55:44	0	0	0	0	0	117	21.0	51.0
3046	2018-12-22 15:56:46	0	0	0	0	0	133	21.0	51.0
3047	2018-12-22 15:57:48	0	0	0	0	0	130	21.0	51.0
3048	2018-12-22 15:58:50	0	0	0	0	0	130	21.0	51.0
3049	2018-12-22 15:59:52	0	0	0	0	0	126	21.0	51.0
3050	2018-12-22 16:00:54	0	0	0	0	0	121	21.0	51.0

Fig. 3. (b) Represents Usage of AC, FAN and Lights with Sensors.

5. Results and Outcomes

The working of Electricity Saving Module represents the monitoring and analyzing consumption of electricity data phase. The collected data of sensors and state of electric equipment have been stored on cloud storage for calculating electricity consumption with and without sensors. Figure 3(a) and (b) represents sensing data of used sensors with state of AC, FAN and Lights based on decision making of turning ON or OFF of the device.

Looking towards the consumption of electricity by households, it's necessary to answer the question i.e., what is the impact of IoT with respect to managing and controlling of electricity consumption? Information presented in table 3 concludes the overall impact of IoT i.e., use of sensors for controlling electric equipment.

1. Five electric equipment i.e., AC, FAN, 3 Lights were considered to get experimental results. Consumption of each equipment with and without sensors at every 1 minute interval has been stored in a database which was available on cloud i.e., shown in figure 3 (a) and (b).
2. While observing consumption results of AC that clearly indicates using the decision making with the help of various sensors have decreased the electricity consumption up to 98%, stated in table 3.
3. After implementing ESM, Electricity consumption of Light 1, Light 2, Light 3 and FAN also decreased by 80%, 40%, 49%, 60%, respectively and it affects the electricity bill amount of consumer.
4. Electric equipment wise detailed electricity consumption with and without sensors have been presented in table 4. It shows consumption in kW for approx. 11 hours and using ESM it saves 4.92 KWs.
5. Looking towards the equipment wise consumption using ESM, the consumption of AC, FAN, Lights 1-3 have been recorded as 0 due to absence of person in the room detected using Motion sensor as well as which lights to be kept ON, have been decided based on the Light intensity module for example, on 19 Feb. at 3pm and 4 pm as per the light intensity present in room, only two lights were to be kept ON. Same concept is extended for AC and based in DHT11 sensor, AC was ON at 3pm on 20 Feb. as per the rule mentioned in 4.3 section (algorithm)
6. Looking towards the reverse scenario of point no 4, i.e., consumption without using ESM, assuming the person has turned on all the electric equipment without knowing the actual need of them based on the natural light in the room as well as temperature in the room and remained ON even if the person was available in the room or not and graphical representation stating consumption-efficiency with and without ESM have been shown in figure 4(a) and (b)
7. From the experimental results mentioned in table 3, there were savings of approx. 18 kWs calculated as per the algorithm mentioned in 4.4 section for approx. 34 hours and it was achieved using the Electricity Saving Module which has decreased the financial burden of consumers.

Table 3. Electricity Consumption of Electric Equipment with and without Sensors

Electric Equipment	Consumption Without IoT	Consumption With IoT	Difference	Decreased in Electricity (%)
AC	17.28	0.18	17.1	98.958333
Light 1	0.4727	0.0889	0.3838	81.193146
Light 2	0.3039	0.1797	0.1242	40.868707
Light 3	0.3032	0.1539	0.1493	49.241425
FAN	0.389	0.1553	0.2337	60.077121
Total Consumption	18.7488 kW	0.7578 kW	17.991	95.958141

Electricity Saving using IoT for 33 Hours 47 minutes is 17.9910 kW

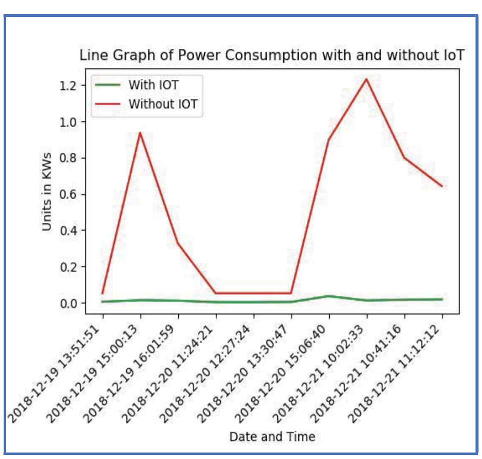


Fig. 4. (a) Electricity Consumption at Interval of 1 Hour.

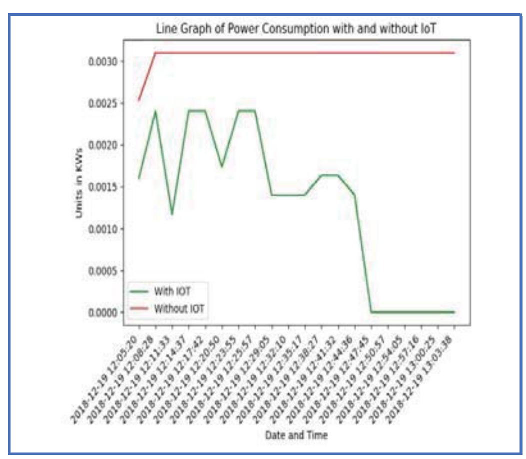


Fig. 4. (b) Electricity consumption at Interval of 3 minutes.

Table 4. Comparison of devices wise Electricity Consumption with and without IOT

Date	Time	With IOT Electricity Consumption in KWs						Without IOT Electricity Consumption in KWs						Difference of Power Consumption
		Total	AC	Light1	Light 2	Light 3	Fan	Total	AC	Light1	Light 2	Light 3	Fan	
19-12-2018	13:51:51	0	0	0	0	0	0	0.05	0	0.01	0.01	0.01	0.02	0.04
19-12-2018	15:00:13	0.02	0	0	0.01	0.01	0	0.94	0.9	0.01	0.01	0.01	0.01	0.93
19-12-2018	16:01:59	0.02	0	0	0.01	0.01	0	0.33	0.28	0.01	0.01	0.01	0.02	0.32
20-12-2018	11:24:21	0	0	0	0	0	0	0.05	0	0.01	0.01	0.01	0.02	0.05
20-12-2018	12:27:24	0	0	0	0	0	0	0.05	0	0.01	0.01	0.01	0.02	0.05
20-12-2018	13:30:47	0	0	0	0	0	0	0.05	0	0.01	0.01	0.01	0.02	0.05
20-12-2018	15:06:40	0.04	0.02	0	0.01	0.01	0	0.9	0.86	0.01	0.01	0.01	0.01	0.86
21-12-2018	10:02:33	0.02	0	0	0.01	0.01	0	1.23	1.2	0.01	0.01	0.01	0	1.22
21-12-2018	10:41:16	0.02	0	0	0.01	0.01	0	0.8	0.76	0.01	0.01	0.01	0.01	0.78
21-12-2018	11:12:12	0.02	0	0	0.01	0.01	0	0.64	0.6	0.01	0.01	0.01	0.01	0.62
												For ≈11 Hours		4.92 KWs

6. Conclusion

This paper addresses the idea that the residential buildings would shift themselves toward modern households i.e., automated decision making of turning ON or OFF of Electric equipment. In this paper, an Electricity Saving Module has been proposed, which ensures an electricity-efficient utilization of household resources. By applying the concept of IoT to automate ON/OFF of household equipment, based on experimental results it can be concluded that the presence of a person in room/house, temperature of house, natural light available in room detected using IoT based sensors plays an effective role while calculating the electricity bills which will help the consumer to manage his/her monthly finance.

References

1. Al-Ali, A., Zualkernan, I., Rashid, M., Gupta, R., Alikarar, M., 2017. A smart home energy management system using IoT and big data analytics approach. IEEE Transactions on Consumer Electronics, 63(4), p. 426-434.
2. Bommana, A. 2017. IoT based Monitoring and Control System for Home Automation using Prediction Algorithms. IJARCCE, Vol. 6, Issue 6, p. 70-73.
3. Dangar, B., Joshi, S. 2015. Electricity theft detection techniques for metered power consumer in GUVNL, GUJARAT, INDIA. Clemson University Power Systems Conference (PSC), p. 1-6.
4. Jammes, F. 2016. Internet of Things in Energy Efficiency: The Internet of Things (Ubiquity Symposium). Ubiquity, p. 1-8.
5. Khan, M., Silva, B., Han, K. 2016. Internet of Things Based Energy Aware Smart Home Control System. IEEE Access, 4, p. 7556-7566.
6. M. Ashok, P. Suresh, Varma, M. 2017. IoT based Monitoring and Control System for Home Automation. International Journal of Engineering Technology Science and Research, Volume 4, p. 413-418.
7. Pan, J., Jain, R., Paul, S., Vu, T., Saifullah, A., Sha, M. 2015. An Internet of Things Framework for Smart Energy in Buildings: Designs, Prototype, and Experiments. IEEE Internet of Things Journal, 2(6), p. 527-537.
8. Pavithra, D., Balakrishnan, R. 2015. IoT based monitoring and control system for home automation. 2015 Global Conference on Communication Technologies (GCCT), p. 169-173.
9. Sahoo, S., Nikovski, D., Muso, T., Tsuru, K. 2015. Electricity theft detection using smart meter data. 2015 IEEE Power Energy Society Innovative Smart Grid Technologies Conference (ISGT), p. 1-5.
10. Salman, L., Salman, S., Jahangirian, S., Abraham, M., German, F., Blair, C., Krenz, P. 2016. Energy efficient IoT-based smart home. 2016 IEEE 3rd World Forum on Internet of Things (WF-IoT), p. 526-529.
11. Wei, C., Li, Y. 2011. 2011. Design of energy consumption monitoring and energy-saving management system of intelligent building based on the Internet of things. International Conference on Electronics, Communications and Control (ICECC), p. 3650-3652.
12. https://www.healthandsafetyatwork.com/workplace-lighting/making-light-work
13. http://www.hse.gov.uk/pUbns/priced/l24.pdf

Employee Hiring: NLP Based Job Profile Recommendation System

Aparna Sawant [a], Harsh Said [b], Ritika Garad [c*], Rakeshkumar Visave [d], Omkar Bhosale [e], Samarth Usture [f]

[a]*Prof, Information Technology, Vishwakarma Institute of Technology, Pune, India.*

[bc*def]*Third year student, Information Technology, Vishwakarma Institute of Technology, Pune, India*

Abstract

Now the world is becoming more dependent on the internet and other facilities for most of the important things to do. This project mainly aims to develop a system which has functionalities to provide the opportunity to the job seekers to find the best job according to their qualifications. In India, there are various platforms for the employees or the job seekers to provide but there is no single platform for the startup companies to post the jobs at very less time to increase their employees' ratios. So, this project acts as a mediator for both the job seekers and companies to search the jobs and post the jobs respectively. A hiring portal website simplifies the hiring process, making it more efficient and effective for both employers and job seekers, and ultimately helps match the right candidates with the right job opportunities. Hiring the right candidate is a very lengthy and time-consuming process so we automate the process of selecting the right candidate for a job. We have developed an algorithm that can calculate the percentage of matching skill sets. In addition, we have also developed a module that automatically sends an email to the perfect match applicants. This reduces manual intervention and human effort, ultimately saving recruiters a lot of time. By using our portal, hiring becomes easier and more effective.

Keywords: JD (job Description), Recruitment, Skill Sets, NLP (Natural Language Processing).

Nomenclature

$P_{(A)}$	probability
A	a particular job profile
F	number of skills of user
N	Total required skills for JD

1. Introduction

The Employee Hiring Portal is an efficient online platform for job seekers, company owners, and administrators. It simplifies the hiring process and makes it more accessible to all parties involved. Job seekers can post their profiles and apply for various job positions available on the website [1]. Companies can register themselves and post job positions according to their requirements. Companies can review the status of any particular job and view how many candidates have registered for that job profile. Moreover, the companies can assign tasks to the candidates according to the job positions and domain. They can verify the candidate's details to check their eligibility for the job. Based on the candidate's performance in the assigned tasks, companies can decide whether to select them or not [2]. The administrator acts as a mediator between the candidates and the companies, ensuring a verified and secure platform for candidates. For employers, the website offers tools such as creating job listings, searching for candidates, screening, shortlisting, scheduling interviews, and managing job offers. The website also offers additional features such as background checks and skills assessments to make informed hiring decisions. Job seekers can search for job openings, submit their resumes, and track the progress of their applications [3]. The website also provides job alerts and notifications to keep them updated on new job opportunities. Additionally, the website provides career advice, tips on job hunting, and information on salary trends and market conditions to help job seekers improve their job search skills. The website includes key features such as a powerful job search engine, a resume database for job seekers, an applicant tracking system for employers, communication tools for both parties, and analytics and reporting features for tracking key hiring metrics. Overall, the Employee Hiring Portal is an excellent online platform that connects employers with job seekers [5]. It simplifies the hiring process and communication offers tools and features that make it more efficient and effective for both parties. Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that focuses on enabling machines to understand and analyses human language. In the context of employee hiring, NLP can be used to develop a profile recommendation system that can help companies identify and hire the best candidates for a job. By analyzing the language used in job descriptions and candidate resumes, NLP algorithms

*Ritika Pradeep Garad
E-mail address: ritika.garad21@vit.edu

can match candidates with the most relevant job opportunities based on their skills, experience, and qualifications. This paper will explore the use of NLP in developing a profile recommendation system for employee hiring and its potential benefits for companies and job seekers.

2. Literature Survey

[1] In a recent study, a team of researchers presented Job Portal, an innovative online recruitment platform that allows employers from different locations to post job listings that can be viewed by job seekers. The software is specifically designed to collect job specifications based on industry standards and to pair them with qualified applicants. In this research paper authors discussed about the design and development of a web application called Job Portal, which is designed to help geographically distributed clients (such as recruiters and job seekers) connect with each other. The authors of the paper highlight the challenges of building a job portal that can handle multiple clients from different locations, and they describe the architecture and features of the web application they developed. The authors also discuss the technologies used to build the web application, including Java, HTML, and MySQL, and they provide a detailed description of the system's database design and user interface. They also describe the testing and evaluation process used to assess the performance and usability of the web application. Overall, this research paper provides valuable insights into the design and implementation of a job portal that can support geographically distributed clients, and it highlights the importance of using appropriate technologies and testing methods to ensure the functionality and usability of such a system.

[2] ONLINE JOB PORTAL In this paper the authors Mustafa Pinjari, Nishit De, Rutvij Kokne, Aamir Siddiqui, Dnyanoba Chitre discuss the designing phase of a web application using.net platform for real time learning experience. This research paper describes the design and development of an online job portal, which is designed to help job seekers and employers connect with each other. The authors of the paper highlight the importance of creating a user-friendly and efficient job portal to help streamline the job search process and increase the chances of successful job placement. The authors discuss the features and functionality of the online job portal, including job search, job posting, and applicant tracking. They also described the technologies used to build the web application, including PHP and MySQL, and they provide a detailed explanation of the system's database design and user interface. The authors also discuss the testing and evaluation process used to assess the performance and usability of the online job portal. They describe the various testing methods used, including functional testing and user testing, and they provide a detailed analysis of the results. Overall, this research paper provides valuable insights into the design and development of an online job portal, and it highlights the importance of creating a user-friendly and efficient system to help job seekers and employers connect with each other.

[3] In this paper -A Online Job portal management system published by Keethana Kopuri, Gulam Mujtaba Hussain Aqueel, Azbar Sadiqa Jabeen, Dr.T.K. Shaik Shavali proposed a system which provides a technical application for managing the online job system. The research paper titled "An Online Job Portal Management System" describes the development of a web-based job portal system that facilitates the process of job search and recruitment. The purpose of this system development is, provide a platform for job applicants to easily find and apply for the job vacancies posted by employers, while employers can easily search for qualified candidates and manage job postings. The system was developed using PHP and MySQL, and it includes various features such as job search and filtering options, job posting and management, resume upload and management, and candidate profile creation and management. The system also includes a messaging system that allows job seekers and employers to communicate with each other. The paper discusses the various challenges faced during the development process, including issues related to security and data privacy. To address these challenges, the system was designed with various security features such as user authentication and authorization, data encryption, and role based access control. Overall, the online job portal management system provides a user-friendly platform for both job seekers and employers to efficiently manage the job search and recruitment process. The system can be customized to meet the specific needs of different organizations and can be used to streamline the hiring process, reduce costs, and improve efficiency.

[4] In this paper the authors Marjan-Mansourvar and Norizan binti Mohd Yasin describes online recruitment system for the students to provide a way to search online IT jobs. According to the study, a job web portal is suggested to enhance the quality of education by linking students to potential job openings. The portal is designed to provide a platform for students to search and apply for job opportunities that match their skills and interests. The paper discusses the importance of job portals in improving education quality, as they can help students gain practical knowledge, develop skills, and gain valuable work experience. The proposed portal has several features, such as a job search engine, a resume builder, and a career advice section. The job search engine allows students to search for job opportunities based on their skills and preferences, while the resume builder helps students create a professional resume. The career advice section provides students with guidance on various career-related topics, such as interview skills and career development. The paper also describes the detailed information about the steps of

implementation that are followed by the authors to design the proposed portal, which involves the functionalities of various web technologies and programming languages. The portal is designed to be user-friendly and accessible to students from different backgrounds. Overall, the research paper suggests that the development of a job web portal can improve education quality by providing students with practical knowledge and work experience. The portal can also help students make informed decisions about their careers and prepare them for the workforce.

[5] In this paper titled as Job Portal - A Web Application for Geographically Distributed Multiple Clients of authors Vivek Kumar Sehgal, Akshay Jagtiani, and Meha Shah, aimed to improve technological knowledge while designing their project and address some of the issues that currently exist within the employment system. The research paper discusses the development of a web application called "Job Portal" that is designed to connect job seekers with potential employers. The web application is designed to be accessible to multiple clients across different geographic locations. The authors of the paper discuss the challenges associated with developing a job portal that can serve multiple clients in different locations. They explain how they used various technologies such as PHP, MySQL, and JavaScript to create the web application. The authors also discuss the features of the Job Portal application, such as a search engine that allows job seekers to find relevant job listings, and an interface for employers to post job openings and review job applications. The research paper concludes by discussing the benefits of using a web-based job portal, such as improved efficiency in the job search process, increased accessibility to job opportunities, and reduced costs associated with traditional job search methods. Overall, the research paper provides valuable insights into the development of a web-based job portal that can serve multiple clients in different geographic locations.

[6] This paper "REVIEW OF JOB PORTAL IN RECRUITMENT PROCESS LIFE CYCLE "the authors describe the overview of an employee Management system and the need and their move towards the fields are given. The research paper provides a comprehensive review of the use of job portals in the recruitment process life cycle. The authors analyze the various stages of the recruitment process, including job posting, applicant screening, and candidate selection, and examine the ways in which job portals can be used to improve efficiency and effectiveness at each stage. The paper discusses the advantages of job portals, such as increased access to job listings, improved visibility for job postings, and reduced time and cost associated with traditional recruitment methods. The authors also examine the challenges associated with using job portals, such as the potential for fraudulent job postings and the need for effective screening and filtering mechanisms to ensure the quality of candidates. The research paper reviews various job portals currently in use, such as Monster, indeed, and LinkedIn, and provides insights into their features and functionalities. The authors also discuss emerging trends in the use of job portals, such as the use of artificial intelligence and machine learning to improve candidate matching and recruitment analytics. Overall, the research paper provides a valuable review of the use of job portals in the recruitment process life cycle, highlighting their advantages and challenges and providing insights into current and emerging trends in the field.

3. Implementation Details

3.1 Proposed System

In this paper the employee's hiring system is proposed which is helpful for the most of the job seekers and its less time consuming than the traditional system of hiring process as everything is being online and also some time there is no possibility of taking all the employees on record every time so this platform provides this functionality to the candidates. It is a web-based platform for the easy and fast responder for the users.

Modules Includes:
- User/Employee Dashboard
- Company/Recruiter Dashboard
- Administrator Dashboard

3.2 Methodology

User Dashboard: - In this system any one can register with the system who are seeking for the jobs they can login with and search for different job profiles which are posted by the companies with this system, they also have to verify their details every time when they visit the dashboard. It also contains the information about the companies for which they are applied and what is the status of that job application. There is another functionality added where the applicant can check the task or work allocated to them and within that time span, they have to complete that task so only from that task basis the applicant can selected by the companies. They can also check their progress of every application coming from the companies. Applicant can also check the job posted by companies not only this but also, they can visit to the official website of the company for security concern for confirmation of the company profiles. Applicant can also visit the companies' details page before applying for that job this is designed because the applicant can also verify the companies are registered have been officially approved by government by visiting their official websites and also verify the opening in the companies.

Administrator Dashboard: - In this system the administrator plays a vital role for this web application the administrator acts as a mediator between the user and companies' dashboard. Administrators have the responsibility to verify the company's profiles. Companies also have to be verified from the administrator. After complete verification, administrators allow companies to post the hiring job details with the portal.

Company Dashboard: - In this system the companies can register themselves with the portal. Once the company is registered with the administrator the id is assigned to the company so that they can post their job with their id. Companies can post more than one job at a time. They also have the rights to check how many candidates are registered for that particular job profile and review as per their criteria. The preferences are decided according to only companies' rules and conditions as same as official website of that company. In this portal new functionality is added i.e., the tasks are allocated for the applicant according the job positions and the companies record this so that they can check the performance of candidates if they are satisfied from the task which have been done by the applicant then they can either schedule their interview of give them redeem status that they have been hired by companies for that job. The notifications are given through the various platforms like mail, SMS system to the applicant.

3.3. Flowchart

The diagram illustrates the flow of the System Architecture, which comprises three primary modules: the User module, Admin module, and Recruiters module. The hiring portal serves as the central module of the system, enabling both companies and users to register and facilitating communication between them. It serves as the primary interface for the hiring process. The admin module is responsible for verifying the accuracy and genuineness of data provided by users and companies before forwarding the applications to respective companies. The user and recruiter dashboards are two separate modules that offer different functionalities to their respective users. The user dashboard enables job seekers to apply for multiple job opportunities and monitor the progress of their applications. The recruiter dashboard provides companies with a platform to post job openings and manage applications from job seekers. In essence, the diagram depicts a standard hiring process that consists of a central portal for job seekers and companies, an admin module for verification, and separate dashboards for users and recruiters to manage their tasks.

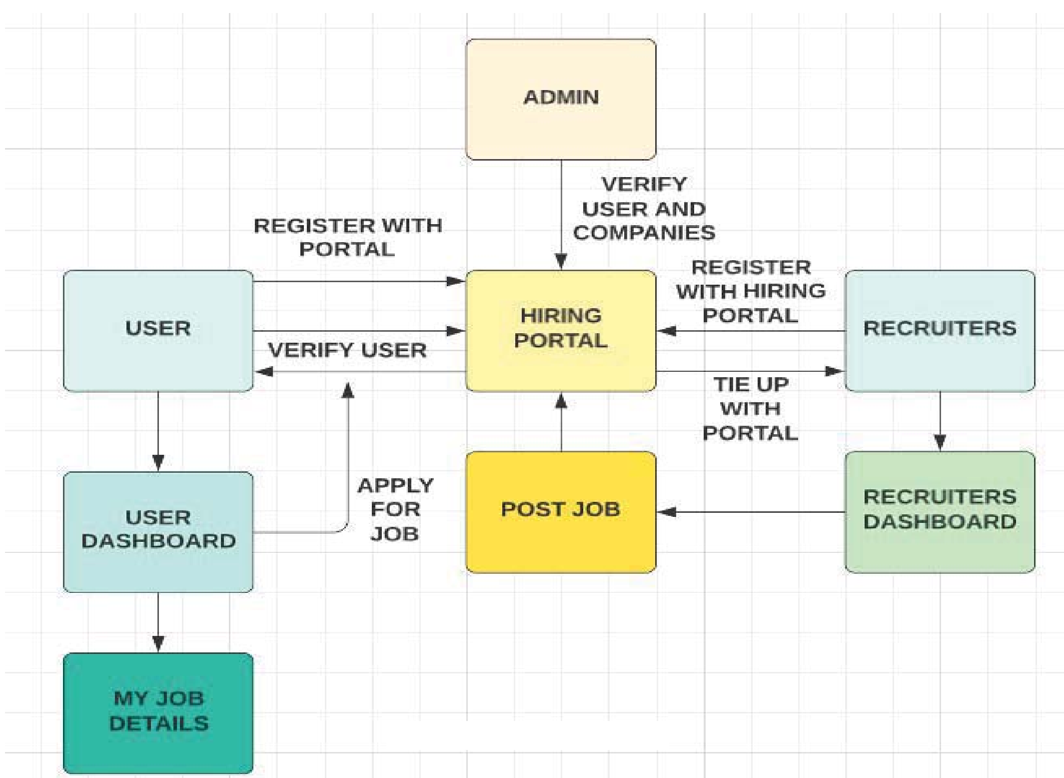


Fig. 1. System Architecture.

The Following diagram features two APIs - one for users and another for job descriptions. The user API provides information about the user's job description and skill set values, while the job description API contains details about the job and its required skill set.

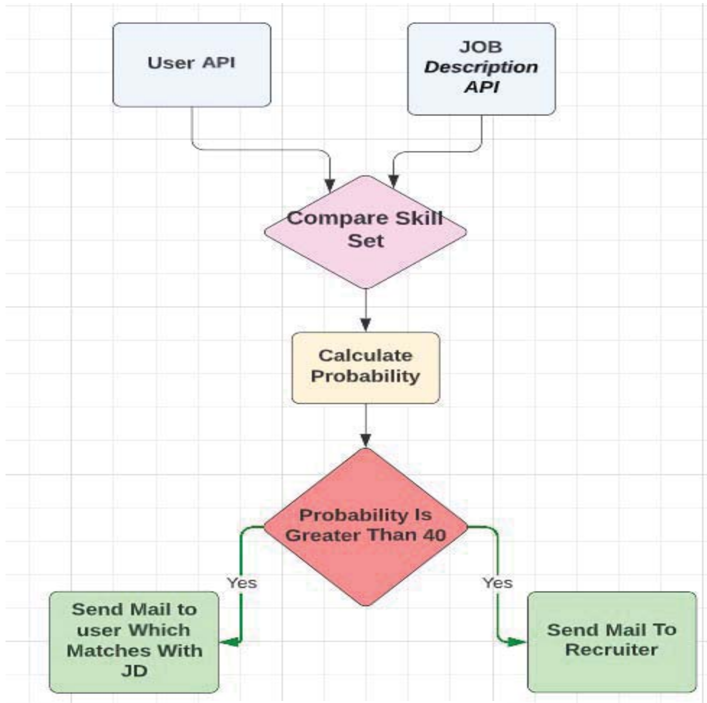


Fig. 2. System Flow Diagram.

listed for the job. Overall, this diagram represents a job hiring system that leverages APIs to match users with job requirements. The system calculates the probability of a match and subsequently notifies the user and recruiter about the shortlisting.

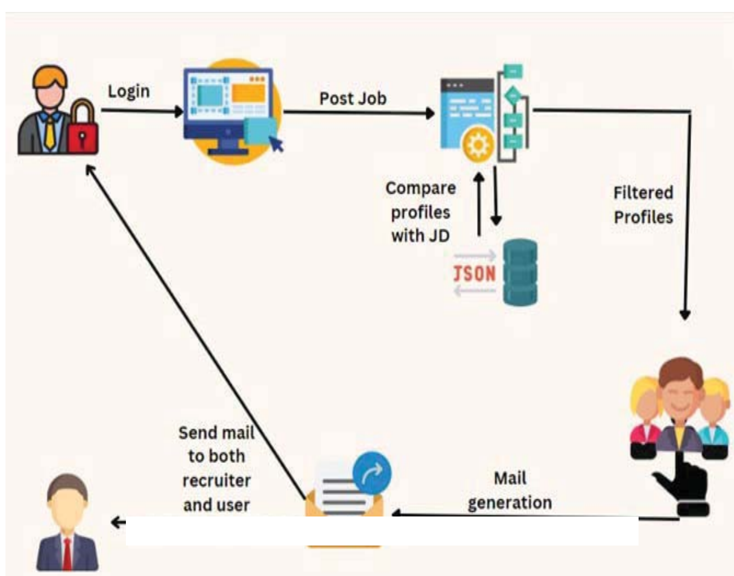


Fig. 3. System Configuration for Sending the Mail.

3.4. Equations and formulae

Probability Formula:

$$Probability = \frac{\text{Number of skills of user}}{Total\ Reqired\ Skills\ For\ JD}$$

$$P(A) = \frac{F}{N}$$

Where:

- P(A)= Probability of user profile skills match with the recruiters posted job
- F= Number of skills set user have for that particular job profile.
- N=Total / All skill set values required to match with that job profile.

4. Limitations

The current hiring or employment systems have some of the dysfunctionalities as follow:
1. The system does not have any criteria to separate out the particular candidate with greater impact rate for any particular work.

2. There may be modifications can be done in such way that this hiring system can automate and schedule interviews also on this platform only like the tasks are assigned.

5. Future Scope

This system provides online platform for the employees and the companies for easy functionality and makes the hiring process efficient but there must be some system in the future like the interview processing of levels can be included so that this hiring process becomes totally compatible for hiring with small to large scale as well as small scale hiring process and the time and security factors also considered.

6. Conclusion

In this project the hiring system and its functionality are defined in all the ways every dashboard system has its own functionalities. The use of web-based technologies is like the MERN stack. For the database the Mongo DB is used it has advantages as ease of use to the user's easy registration process for the companies with a portal. System administrators are also involved to ensure the proper functioning of the system and also, they provide the security concern regarding the only validating and verified companies that can be viewed and registered. Hiring process mainly focuses on data about users. It is kept private only for that company's portal and designation of the tasks for each applicant can be different and designed by that company only. Administrator didn't interfere with the process; it makes the hiring process more valuable. A hiring portal website is an essential tool for employers and job seekers alike. It provides a central hub for job postings, resumes, and other relevant information, and offers various features and tools to streamline the hiring process. Employers can use the platform to create job listings, screen and shortlist candidates, schedule interviews, and manage job offers. Job seekers, on the other hand, can search for job openings, submit their resumes, and track the progress of their applications. Additionally, the website may offer additional resources such as career advice, tips on job hunting and information on salary trends and market conditions to help job seekers improve their job search skills. Overall, a hiring portal website simplifies the hiring process, making it more efficient and effective for both employers and job seekers, and ultimately helps match the right candidates with the right job opportunities. Hiring a right candidate is very lengthy and time-consuming process so we automate a process of selecting a right candidate for a job we developed a algorithm that can be calculate the matching skillset percentage and we also developed one module which sends an automatic mail to the perfect matched application we reduce manual interruption and human efforts and through the use of our portal we save a lot of time of recruiters and makes hiring easy and effective.

References

- Journal articles:

1. Vivek Kumar Sehgal; Akshay Jagtiani; Meha Shah; Anupriya Sharma; Arpit Jaiswal; Dhananjay Mehta, Job Portal - A Web Application for Geographically distributed multiple clients. https://doi.org/10.1109/AIMS.2013.38
2. Keethana Kopuri, Gulam Mujtaba Hussain Aqueel, Azbar Sadiqa Jabeen, Dr.T.K. Shaik Shavali,A, Online Job portal management system, https://www.ijirt.org/master/publishedpaper/IJIRT144246 _PAPER.pdf
3. Marjan Mansourvar, Norizan binti Mohd Yasin, 2014. Development of a Job Web Portal to Improve Education Quality, January 2014. International Journal of Computer Theory and Engineering. DOI:10.7763/IJCTE. 2014.V6.834.
4. Shobha Rani.B. R, Suparna.B. M, Teja.K. S, 2015. Classification of Vehicles using Image Processing Techniques, International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181 Published by, www.ijert.org NCAISE-2015 Conference Proceedings, Special Issue – 2015.
5. PAVAN P APARANJI, DR. JAI PRAKASH TRIPATHI, 2018. REVIEW OF JOB PORTAL IN RECRUITMENT PROCESS LIFE CYCLE JETIR (ISSN-2349-5162), Vol 5, Issue 1.
6. Alavi, M., & Leider, D. 1999. Knowledge management systems: Emerging views and practices from the field. In System Sciences, HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on. (IEEE).
7. Yang, Zhilin, Shaohan Cai, Zheng Zhou, and Nan Zhou. 2005. Development and validation of an instrument to measure user perceived service quality of information presenting web portals, Information & Management, 42, no. 4, p. 575-589.
8. Benbya, Hind, Giuseppina Passiante, and Nassim Aissa Belbaly, 2004. Corporate portal: a tool for knowledge management synchronization, International Journal of Information Management, 24, no. 3, p. 201-220.
9. Saat, N.M.; Singh, D., 2011. Assessing suitability of candidates for selection using candidates' profiling report, Electrical Engineering and Informatics (ICEEI), International Conference on, p. 1- 6.
10. Rafter, R., Bradley, K., & Smyth, B. 2000. Personalized retrieval for online recruitment Services, Proceedings of the 22nd Annual Colloquium on Information Retrieval.
11. Bizer, C., Heese, R., Mochol, M., Oldakowski, R., Tolksdorf, R. & Eckstein, R. 2005. The impact of semantic web technologies on job recruitment processes. Proc. International Conference Wirtschaftsinformatik, Bamberg, Germany, p. 137-138.
12. Lievens, F; Van Dam, K; & Anderson, N., 2002. Recent trends and challenges in personnel selection. Personnel Review. MCB Univ Press, 31(5), p. 580-601.
13. Calì, A., Calvanese, D., Colucci, S., Di Noia, T. D. & Donini, F.M., 2004. A logic-based approach for matching user profiles. In KES, Lecture Notes in Artificial Intelligence, p. 187-195.
14. P. Scholl, D. Mann, C. Rensing, R. Steinmetz, 2007. Support of Acquisition and Organization of Knowledge Artifacts in Informal Learning Contexts. In: European Distance and E-Learning Network: EDEN - Book of Abstracts, p. 16.
15. G. Adomavicius, A. Tuzhilin, 2005. Toward the Next Generation of Recommender Systems: A Survey of the State of-the-Art and Possible Extensions. IEEE Transactions on Knowledge and Data Engineering, Vol. 17, No. 6, p. 734-749.
16. J. F. Sowa, 1992. Semantic networks. In: SC Shapiro (Edi.), Encyclopaedia of Artificial Intelligence 2, John Wiley, New York, p. 1493-1511.
17. Alavi, M., & Leider, D., 1999. Knowledge management systems: Emerging views and practices from the field. In System Sciences, HICSS-32. Proceedings of the 32nd Annual Hawaii International Conference on IEEE. p.8.
18. Bizer, C., Heese, R., Mochol, M., Oldakowski, R., Tolksdorf, R. & Eckstein, R., 2005. The impact of semantic web technologies on job recruitment processes. Proc. International Conference Wirtschaftsinformatik, Bamberg, Germany. p. 137-138.
19. Rafter, R., Bradley, K., & Smyth, B., 2000. Personalized retrieval for online recruitment Services, In Proceedings of the 22nd Annual Colloquium on Information Retrieval.

Achieving Secure Storage on Cloud using Blockchain Technology

Parin Patel [a*], Hiren Patel [b]

[a]*Research Scholar, Kadi Sarva Vishwavidyalaya, Gandhinagar-382015, Gujarat, India*
[b]*Sarva Vidyalaya Kelavani Mandal, Gandhinagar-382015, Gujarat, India*

Abstract

Storage requirements are increasing rapidly with fast growing cloud computing technology. More the requirement of data storage and transaction processing, more the chances of malicious activities. Storage security is the burning issue now days. To deal with this problem we have proposed a system which provides an efficient security mechanism for data storage on the cloud with updated access control policies by making use of Blockchain technology. The data/transactions stored on the Blockchain can never be modified and hence they are immutable in nature. The usage of smart contract improves transparency and creates a trust model that eliminates the requirement of trusted third party. The system manages the log trails of all transactions and manages all access policies using smart contract. The proposed model is transparent, traceable, and secure. Data owner holds all the rights of her data and any activity pertaining to access the data illegitimately shall not be accepted by the network and notified to all concerned. Migrated File location easily traceable. We demonstrate the model of our proposed mechanism with details about all concerned stakeholders. We also analyze the security of our proposed model.

Keywords: Cloud, Blockchain, Smart contract, Hash.

1. Introduction

Cloud technology has become increasingly popular in recent years and is being used in a wide range of industries and applications. Here are some examples of how cloud technology is being used:

Cloud storage: Cloud storage services, such as Google Drive and Dropbox, allow users to store and access files from any device with an internet connection.

Cloud computing: Cloud computing allows businesses to run applications and services without having to invest in expensive hardware and software infrastructure. Companies can rent cloud computing resources from providers like Amazon Web Services, Microsoft Azure, and Google Cloud Platform.

Software as a Service (SaaS): SaaS is a cloud-based software delivery model that allows users to access applications over the internet. Examples include Salesforce, Slack, and Microsoft Office 365.

Infrastructure as a Service (IaaS): IaaS allows businesses to rent computing resources, including servers, storage, and networking, from cloud providers. This allows businesses to scale their infrastructure up or down based on demand.

Platform as a Service (PaaS): PaaS provides a platform for developers to build, test, and deploy applications in the cloud. Examples include Google App Engine, Microsoft Azure App Service and Heroku.

Internet of Things (IoT): IoT devices generate a vast amount of data that needs to be stored and processed in real-time. Cloud technology provides the necessary infrastructure to handle this data, allowing businesses to analyze it and derive insights.

Big Data Analytics: Cloud technology enables companies to store and analyze large amounts of data using tools like Hadoop and Apache Spark. This helps companies to extract insights from their data and make more informed business decisions.

*Parin Patel
E-mail address: patelparinv@gmail.com

Overall, cloud technology is being used to improve efficiency, reduce costs, and enhance collaboration across a wide range of industries and applications. This means they can quickly and easily add or remove computing resources as needed, without having to invest in expensive hardware and infrastructure. Cloud computing is often more cost-effective than traditional on-premises computing, as businesses only pay for the resources they use. They also save on maintenance and upgrade costs as these are managed by the cloud provider. Cloud computing allows employees to access data and applications from any location with an internet connection, making it easier for businesses to adopt flexible working practices. Cloud providers offer high levels of reliability and uptime, meaning businesses can be confident that their applications and data will be available when they need them. Cloud providers typically offer high levels of security, including physical security, network security, and data encryption. This means businesses can be confident that their data is protected. Cloud computing makes it easier for teams to collaborate on projects as they can access and share data and applications from any location. Cloud providers typically manage software updates and security patches, meaning businesses don't have to worry about keeping their applications up to date.

1.1. Blockchain Technology

Blockchain technology is a technology that enables transparent storage, secure, and decentralized transfer of data. It was first introduced in 2008 as a core component of Bitcoin, a digital currency, but has since found applications in a wide range of industries. A blockchain is essentially a digital ledger of transactions that is maintained across a network of computers. Each block contains a cryptographic hash of the previous block, along with a timestamp, nonce and transaction details in the chain. Because each block in the chain is connected to the previous one, it creates an immutable record of all the transactions that have taken place on the network. Blockchain technology has several key features that make it particularly useful for applications where security, transparency, and decentralization are important. These include:

Decentralization: A blockchain is maintained across a distributed network of computers, rather than being controlled by a single entity. It provides security against cyber-attacks.

Security: Each block in the chain is cryptographically linked to the previous one, making it virtually impossible to alter or delete previous transactions without detection.

Transparency: Because the ledger is shared across the network, all participants can see and verify the transactions that have taken place.

Immutability: A block cannot be changed or removed after it has been added to the chain. As a result, a permanent and unchangeable record of all network interactions is produced.

Smart contracts: Smart contracts, which are self-executing contracts with the terms of the agreement between the buyer and seller being directly put into lines of code, can be created using blockchain technology.

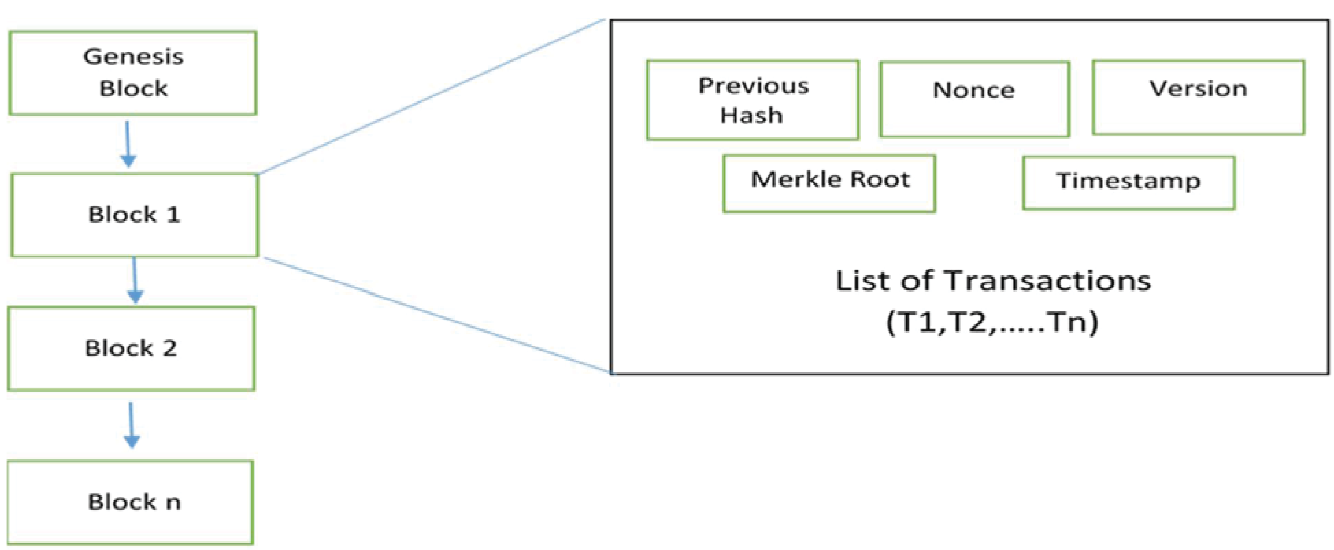


Fig. 1. Block Structure in Blockchain Technology.

Blockchain is basically combination of three technologies.

- A distributes network with shared lager
- Private key cryptography
- It provides service to record tracing, security and store all transactions.

The main focus of block chain technology is to crate secure digital identity reference. Digital signature provides strong identity of the owner.

Public key +Private Key = Digital Signature

There is a distributed peer-to-peer network where integrity of all transaction is verify by the members. Blockchain is the chain of block of transaction where each and every block contains the block header and set of transaction. Block header contains previous block hash value and Merkle root, Timestamp, version and nonce. Merkel tree build continuously by hashing pair of nodes until there is only one hash left. Each leaf node is a data transaction, and every non-leaf node is a hash of previously calculated hash values. Merkle tree requires even number of leaf that's way it is called as binary tree. To generate an even number of leaf nodes, the last hash is duplicated once. Fig 2 shows the structure of Merkle tree.

Each block has a timestamp. They also make it harder for an adversary to manipulate the block chainas well as serving as a source of variation for the block hash. Nonce (' number used only once ') is a number added to a hashed block that meets the limitations of difficulty level when rehashed. The nonce is to be solved by Blockchain miners. Transaction must be added in to the block when it will validate by the miners. For example, a valid transaction implies Bob got one bitcoin from Alice. Alice may have attempted to pass the same bitcoin to Carol, however, since it is a digital asset. Nodes must therefore reach an agreement on which transactions must be maintained in the Blockchain to ensure that no corrupt branches and divergences exist [3,4]. In fact, this is the objective of the second layer of Consensus. There are different consensus mechanisms depending on the type of Blockchain [5]. Proof-of-work (PoW) is the most well-known. To ensure authentication and verifiability, PoW requires solving a complicated computational process such as finding hashes with specific patterns, e.g. a leading number of zeroes [6]. Instead of dividing blocks proportionally to the relative hash rates of miners (i.e. their mining power), protocols of proof-of-stake (PoS) divided trust blocks proportionally [7]. Thus, the choice is fairer and keeps the network from being dominated by the richest participant. Because of the significant reduction in power consumption and enhanced scalability, many blockchains, such as Ethereum [8], gradually shift to PoS. Byzantine Fault Tolerance (BFT) [9] and its variants [10] are other consensus methods.

In the traditional way of transferring money person meet each other and transfer it. Rather than meeting in person we have started providing digital money to each other by using Internet of services. Alice sends digital cash money to bob. Bob is not receiving digital money. Bob can't be sure whether allice is sending digital money to him or it could be a case that Alice digital money to two different people at the same time. Alice could send digital money to Bob and Dev. To make sure that this type of scenario should not happen in transaction. They decide one single person can note down all the transaction in the book. This make sure their network is secure and people who are transferring money can't able to cheat. This is the process how our banks are working. Soon network became big and digital money is being transferred between hundreds and thousands of people. It may be possible that that Dev became money holder and tries to change some transaction or manifest fake transaction for his benefit. To avoid such kind of problem every single person in the group decide to hold the copy of every transaction. If someone try to alter transaction than everyone in the network can verify it. Noting down all the transaction in the ledger make sure that double spending is not allowed. Additionally, it dispels the faith we had in centralised systems. Trust is developed through consensus where everyone is connected to network and making sure that the system is working fruitful.

Application hosted on server connected to database transferring information. Single server hosting crates the single point fauiler where attackers can attack the server and track or change some behaviour of the application. If you decide to store information on backup server than it increase your overhead cost. In blockchain there is no single server. There are multiple servers in the form of various nodes which are hosted with complete synced blockchain or database. Any data in the blockchain shared with everyone who participate and made available on the all nodes. This helps us to remove central focal authority and overheads related to backup servers and database. We always trust third party or database authorities in the digital world over every day's needs While sharing banking services or Cloud services even in the real world. We depend on the authorities to resolve our issues. Blockchain provides trust using cryptographic protocol.

In this paper we have studied about the solution provided tosecurity issues of Cloud storage using blockchain technology. Here is how the rest of the paper is organised. In section 2, we discuss related work, which address security issue in Cloud storage using blockchain technology. Section 3, we discuss details architecture with algorithms. Section 4 contains performance analysis with detailed study. We conclude ourstudy in section 5.

2. Related Work

Authors in [11] have proposed Block-secure, which is blockchain based secure P2P Cloud storage provides more security over the data. In this solution they have used blockchain to store data like file Urls, File replica Urls, Hash, Transaction details and keys. They used customized genetic algorithm to improve performance of distributed architecture. They have compered Multiple users with Single data center, Multiple users with multiple data centers using blockchain with their proposed architecture Multiple users with multiple data centers with blockchain & genetic algorithm. File loss ratio is very low compared to other two architectures. Authors in [12] have proposed blockchain based access control system using smart contract which provides more security on data. In this system updated attribute based encryption scheme used to manage access control mechanism with smart contract. File encrypts by the owner and all details like public link, hash code and access policies are stored in the contract files. Ethereum virtual machine manages all contracts and certification for access, edit and delete permission for the file. Authors [13] have proposed framework which provides more security and over data access and modification. The use of blockchain provides more security on data and also restrict any malicious activities. Authors [14] have proposed Dstore, which manage lease relationships using smart contract. It verifies the correctness of data and store for payment process also manage lease relationship automatically without interference of any third party. Authors [15] have proposed a system which use blockchain based auditing system for data provenance called Provchain. Every transaction recorded with immutable timestamp and generate receipt for each data for validation. This system provides transparency, privacy and reliability to the users. Authors [16] have proposed a framework which use Merkle tree and sampling verification for integrity verification. The use of sampling strategy increases the performance of verification. It removes distrust which exist in traditional environment using blockchain and sampling.

Authors [17] proposed Block-and-response mechanism that verifies the data integrity. Smart contract monitors any data change using Merkle hash tree. Automatically user get warning message when he accesses the data from the Cloud. This system mange integrity verification easily and gives fast response. Authors [18] have proposed a MediBchain which manage data store on Cloud using blockchain technology. Private accessible unit manages communication between users and Blockchain. All metadata stores on the blockchain and encrypted patient's data stored on Cloud. Authors [19] have proposed a novel scheme which removes deduplication using blockchain technology and also process missing or altered file recovery. Use of smart contract do automatic transaction without use of third party. Author [20,21,22] have proposed a solution which provides immutable log storage using blockchain technology. In this system all transactions are stored on block of blockchain and that is immutable and tamper proof.

3. Problem Formulation

3.1. Architecture Overview

Proposed architecture which built using blockchain technology and provides more security on data which it stored on Cloud. An overview of architecture is illustrated in Fig 3. Critical components of the architecture are discussed as follow.

- Data Owner: Data owner is the person who stores the data on the Cloud. Data owner first encrypts the data and store on the Cloud. Only encrypted data stores on the Cloud other information is managed by the data owner. Data owner has only information about the keys and other metadata. Files are chunked and stores on different data center in the Cloud with replicas. The metadata like keys, file URLs, hash values and replica URLs are stores on the blockchain using smart contract.
- Cloud Service Provider: Cloud service provider (CSP) provides Cloud storage services and maintenance of data. CSP is not having any information of metadata like keys, hash values or URLs. Cloud User: Cloud user is can request for data access to CSP and according to the rights provided by the Data Owner Cloud User can access the data. Access Control Mechanism: This mechanism helps to manage access control over the data. Policy defines the rights to create, update and revoke.

- Attribute-based encryption (ABE) is a type of encryption that allows data to be encrypted and decrypted based on certain attributes or characteristics. In traditional encryption, data is encrypted using a specific key, and only someone who has that key can decrypt and access the data. However, in ABE, access to the data is granted based on specific attributes, rather than a specific key. For example, a healthcare provider might want to encrypt patient records so that only authorized personnel can access them. Instead of using a specific key to encrypt the data, ABE can be used to encrypt the data based on specific attributes, such as the job title or department of the personnel who should be allowed to access the data. Only those who meet the specified criteria (i.e. those with the specified job title or department) will be able to decrypt and access the data.ABE works by using a combination of public and private keys to encrypt and decrypt the data. The encryption process is based on a policy, which is a set of attributes that define who should have access to the data. The policy is translated into a set of public parameters, which are used to encrypt the data. The decryption process requires a private key, which is generated based on the attributes of the person or entity attempting to decrypt the data. If the attributes of the private key match the attributes specified in the policy, the data can be decrypted and accessed.

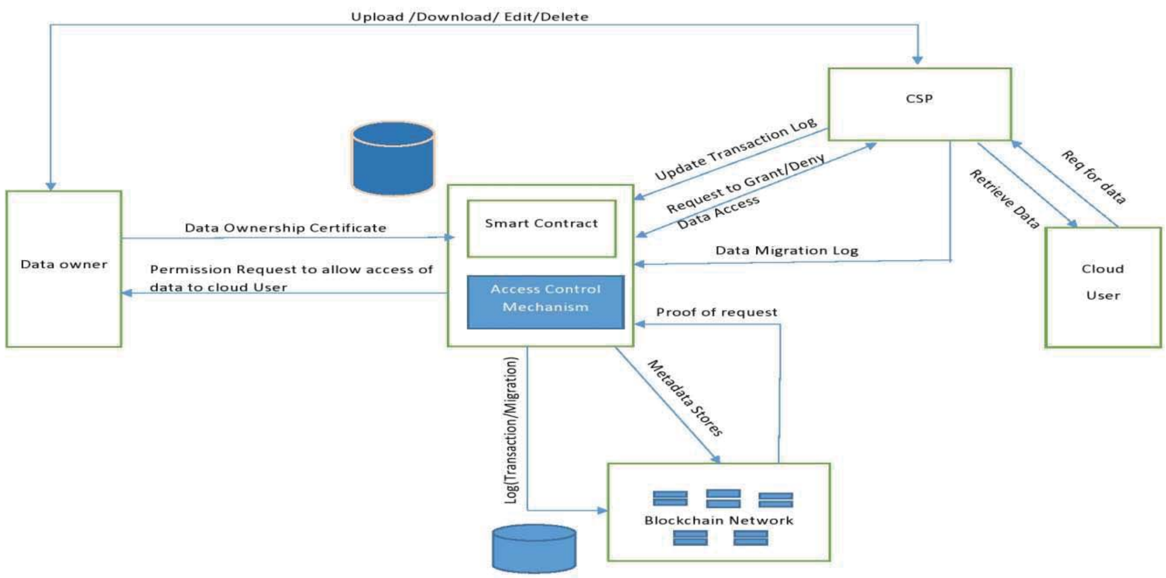


Fig. 2. Proposed Architecture.

- Smart Contract: A smart contract is a self-executing contract that is programmed to automatically execute the terms of an agreement between two or more parties. It is a computer program that runs on a blockchain, which is a decentralized, distributed ledger that provides a secure and transparent way to store and transfer data. contracts are designed to enforce the rules and regulations of a traditional contract, but with the added benefit of automation and decentralization. They use a programming language to define the terms and conditions of the contract, and the rules for executing it. Once the contract is executed, the code is automatically enforced, eliminating the need for intermediaries, such as lawyers or brokers, to oversee the transaction.

Each and every transaction stores in the database. In Table 1 all transaction done by the user on file are stored. It stores each and every transaction operated on file from first to last. It stores User and file ids. Transaction time is also stored on the time. Operations like read, update or delete stores on the table. After completion of the operation it stores new file size and hash value. All transaction-related files are stores in Table 2. It stores file, user ids and hash value of file and timestamp of file creation, modification and accessed by any user. Also stores the size of the file. Whenever use request to access the file to the CSP, CSP send request to Data Owner for file access permission using smart contract. Smart contract provides access to the user and use metadata from blockchain and provide the access rights for particular time period. Whenever time reaches to completion it automatically revokes all rights given to the user. All transaction related to accessing of the file are stores in Table 3. It stores User and file ids. It also stores which type of access is given to the user and time period for which the rights are given.

Store data on the Cloud and generate contract file

1. Data Owner (DO) generates private and public key pair using Attribute based Encryption algorithm (SK, PK).
2. Encrypt the file sing private key (SK) and store on Cloud server.
3. Cloud server replay with location of the file.

4. Generate contract file location, access policy, hash key and public key and stores on blockchain.
5. Smart contract automatically generates ownership certificate and replay to the Data Owner (DO).
6. Transaction stores in Transaction Log (TRL).
7. If Data Owner (DO) wants to delete the file, then smart contract allows only when ownership certificate matches. Automatically public link removes so it can never create problem in future.

When Data Owner (DO) wants to change access policy can update in contract file and reload with ownership certificate.

Request to access file store on Cloud

1. Data User (DU) sends request to access the file on Cloud. Smart contract checks access policy if Data User allow to read the data then it fetch location and public key from the contract file and sends it to Data User(DU). Data User (DU) can download file, decrypt it and read it. Transaction stores in Access control Log(ACL) with timestamp. When access time completes it automatically revokes the rights from Data User(DU) to the file. If Data User (DU) don't have access rights then it denied from the Cloud.
2. Data User(DU) sends request to update the file on Cloud. Smart contract checks access policy if DU allows to update the data then it fetch location and public key from the contract file and sends it to Data User(DU). Data User(DU) can download file, decrypt it and update it. Data User(DU) sends updated file to Data Owner(DO).
3. Data Owner(DO) encrypt the updated file and replace file on Cloud. Generated new hash value and public key are updated on contract file and store transaction in file transaction log. Access rights revokes automatically from Data User(DU). If Data User(DU) don't have access rights then it denied from the Cloud.
4. Data User(DU) sends request to delete the file then it automatically denied if permission is not granted by Data Owner(DO) then Data User(DU) can't delete it. If granted, then Data User(DU) can delete file.

Smart contract Server

1. Access request come from Data User(DU).
2. Check for the permission in access policy. Ifnot, then Send request to Data Owner(DO).
3. If receives rights from Data Owner(DO), then fetch location and public key from the blockchain.
4. Smart contract verifies the integrity of the file. First compute the hash value of the file and generate Filehash(new). if Filehash(new) match with Filehash(old) then it provides location and public key to the Data User(DU). if doesn't match then ask to Data Owner(DO) for new file to recover. Data Owner(DO) again encrypt the file and reload encrypted file on the Cloud and update smart contract with details. Transaction stores on File Log table.
5. Replay with location and public key to Data User(DU).

Cloud Service Provider

1. Cloud Service Provider (CSP) receives encrypted file from the Data Owner(DO) . Store it and replay with location.
2. If Cloud Service Provider (CSP) change location from one server to another server then transaction updated on transaction log table. New file location sends to Data Owner (DO) . Data Owner(DO) updates the contract file and save it to contract server with ownership certificate.

Table 1. Transaction Log (TRL)

UserID	FileID	Timestamp	NewFileSize	NewHashCode

Table 2. File Log (FLL)

FileID	OwnerID	HashCode	Timestamp_Created	Timestamp_LastModified	Timestamp_ Accessed	FileSize

Table 3. Access Control Log (ACL)

UserID	FileID	AccessRight	FromTimestamp	ToTimeStamp

All the transaction stores in the tables are added in the blockchain as per timestamp. In Merkle Tree all transaction stores in the leaf and It generates the root called hroot form the hash of every two nodes in the tree from down to up. Any change in the transaction affects root of the tree. You can identify any change in transaction by comparing previous and current root. The value of the root generated from sixteen transactions in Merkle Tree is added to the block as per Fig 5. Each and every block contains previous hash value, current hash generated from Merkle Tree and list of transactions. Complete detailed block structure is shows in Fig. 1.

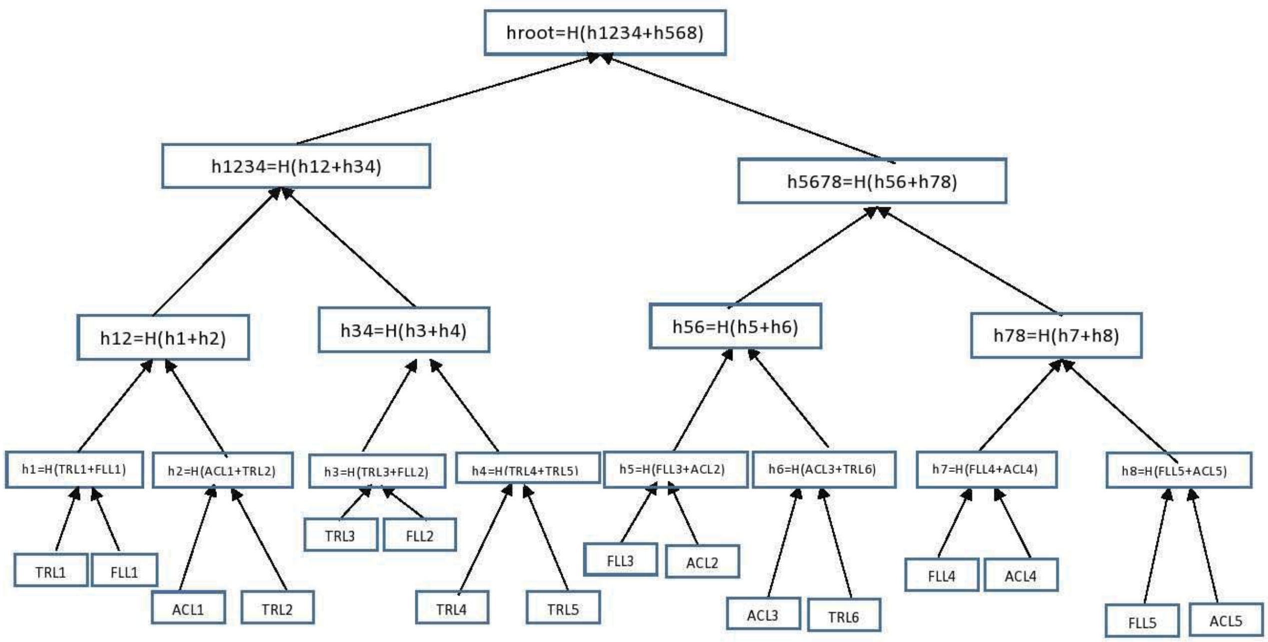


Fig. 3. Merkle Tree with transactions.

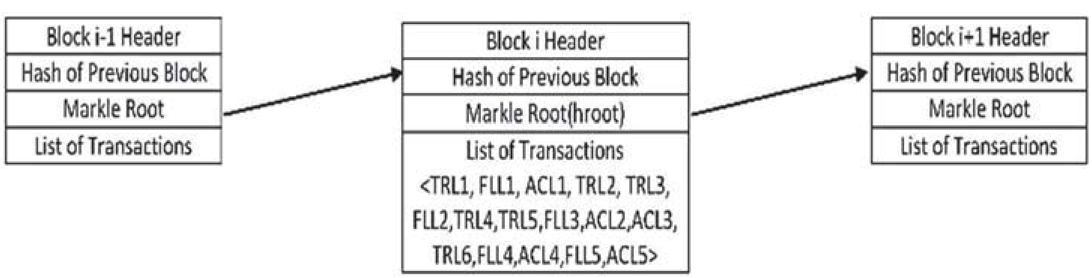


Fig. 4. Block Structure.

This proposed architecture achieves following objectives.

- Enhanced Security: With use of blockchain it provides security to all transaction. Due to the feature of transparency and immutability of blockchain each and every data archives the strongest security. Any change in data can easily identify and recover.
- Access Control: With updated ABE scheme we can provide secure and easy way of access control mechanism. It provides more security by providing strong access control to data. As per the policy structure defined in smart contract it automatically executes all operation without using any third party.
- Integrity Verification:Integrity verification guarantees that the data which is accessed is not corrupted. Using Merkle root user can easily identify the change in any data or transaction.
- Transaction Trails: All transactions form first to last stores in blockchain. You can track all transaction easily and verify the details.

Table 4. System Notations

Parameter	Description
DO	Data Owner
DU	Data User
CSP	Cloud Service Provider
TRL	Transaction Log
FLL	File Log
ACL	Access Control Log
SK	Private Key
PK	Public Key
Hroot	Merkle Tree Root

1.2. Programmatic view of proposed Architecture

Algorithm: Grant Write/Read/Update/Delete Permission Algorithm

```
switch(fileoperation)
{
  case write:
        if fileID does not exist and userID has ‘write’ privilege on fileID then
              write into fileID
              update transactionTuple
        else if fileID exists and userid has ‘write’ privilege on fileID then
              notify ownerID about overwriting the file
                    if consent is given by ownerID then
                          overwrite the fileID
                          update transationTuple
                    end if
        else
              notify userID about no ‘write’ permission form ownerID
        end if
        break;
  case read:
        if fileID does not exist then
              notify userID about not existence of fileID
        else
              if userID has ‘read’ privilege on fileID then
```

```
                    grant permission to userID
                    update transactionTuple
        else
                    notify userID about no 'read' permission on fileID

                end if
        end if
        break;

case update:
        if fileID does not exist then
                notify userID about not existence of fileID
        else
                if userID has 'update' privilege on fileID then
                    grant permission to userID
                    update transactionTuple
                else
                    notify userID about no 'update' permission on fileID

                end if
        end if
        break;
case delete:
        if fileID does not exist then
                notify userID about not existence of fileID
        else
                    if userID is equall to OwnerID then
                    File deleted
                    update transactionTuple
                    else
                  If userID has 'delete' privilege on fileID then
                    grant permission to userID
                    update transactionTuple
                else
                notify userID about no 'delete' permission on fileID
        end if
        end if
        break;

case verify:
        if fileID does not exist then
                notify userID about not existence of fileID
        else
                fileDigestN =Compute hash(fileID);
                If (fileDigest== fileDigestN)
                    notify the file is valid.
                else
                    notify the file is invalid.
                end if
       end if
      break;

default:
        error "invalid operator"
}
fileoperation: {Write, Read, Update/Modify, Delete, Verify}
```

errorCode: {fileDoesNotExist, invalid Write, invalidRead, invalidUpdate, invaliDelete,InvalidVerify,insufficiantRights}
transactionstatus: {success, fail, abort}

Transaction Tuple:
<userID, fileID , fileDigest, fileoperation , timeStamp , IsExist ,transactionstatus ,errorCode , fileVersion>

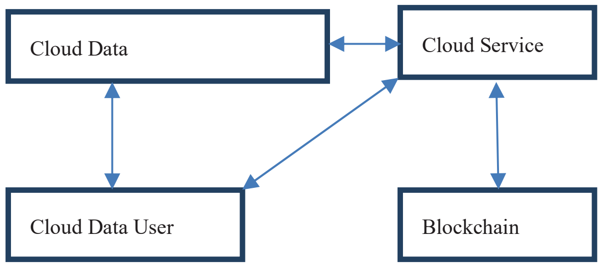


Fig. 5. Transaction Audit Log.

4. Conclusion

In this paper, We present design and implimentation of a system which provides security to the data stored on the cloud using Blockchain Technology.By using blockchain technology we can provide the data which should be immutable and temper proof. By the use of the blockchain and smart contracts, the integrity of information about all transactions, including the granting and changing of access, facts obtain access to file, rejection of the fact, and the inability to modify these data, is ensured.In order to maintain security without third party we use smart contract which automatically executes when condition matches.Cloud server donot have any information about Keys.All the transactions related to data stores in to log tables and logs are maintained using blockchain.The present scheme has also provided management of recovery file.It also keeps track of the files migrated over the cloud.Moreover this scheme manages ownership mangment as well as transfer of ownership is also provided.Proposed System provides more security to data and logs, integrity verification, ownership, dynamic access control.

References

1. How Big Data Analysis helped increase Walmarts Sales turnover?, DeZyre. [Online]. Available: https://www.dezyre.com/article/how-big-data-analysis-helped-increase-walmarts-sales-turnover/109.[Accessed:02-Mar-2020]
2. S.Nakamoto, 2008. Bitcoin,Apeer-to-peer electronic cash system, https://bitcoin.org/bitcoin.pdf.
3. M. Vukoli´c, 2015. The quest for scalable blockchain fabric: Proof-of-work vs. BFT replication, in: International Workshop on Open Problems in Network Security, Springer, p. 112–125.
4. K. Christidis, M. Devetsikiotis, 2016. Blockchains and Smart Contracts for the Internet of Things, IEEE Access 4, p.2292–2303.
5. D. Mingxiao, M. Xiaofeng, Z. Zhe, W. Xiangwei, C. Qijun, 2017. A review on consensus algorithm of blockchain, in: IEEE International Conference on Systems, Man, and Cybernetics (SMC), p. 2567–2572.
6. A. M. Antonopoulos, 2014. Mastering Bitcoin: unlocking digital cryptocurrencies, " O'Reilly Media, Inc.

7. M. Pilkington, 2016. Blockchain technology: principles and applications, Research handbook on digital transformation, 225
8. C. Dannen, 2017. Introducing Ethereum and Solidity: Foundations of Cryptocurrency and Blockchain Programming for Beginners, Apress.
9. M. Castro, B. Liskov, 2002. Practical Byzantine fault tolerance and proactive recovery, ACM Transactions on Computer Systems (TOCS) 20 (4), p. 398–461.
10. Z. Zheng, S. Xie, H.-N. Dai, H. Wang, 2016. Blockchain challenges and opportunities: A survey, Work Pap.
11. Jiaxing Li, Jigang Wu, Long Chen, 2018. Block-Secure: Blockchain Based Scheme for Secure P2P Cloud Storage, Information Sciences, doi: 10.1016/j.ins.2018.06.071
12. Ilya Sukhodolskiy, Sergey Zapechnikov, 2018. A Blockchain-Based Access Control System for Cloud Storage, IEEE,978-1-5386-4340-2/18
13. Ramamoorthy, S., & Baranidharan, B., 2019. CloudBC-A Secure Cloud Data acess Management system, 3rd International Conference on Computing and Communications Technologies (ICCCT). doi:10.1109/iccct2.2019.8824828
14. Xue, Jingting & Xu, Chunxiang & Zhang, Yuan & Bai, Lanhua., 2018. DStore: A Distributed Cloud Storage System Based on Smart Contracts and Blockchain: 18th International Conference, ICA3PP 2018, Guangzhou, China, November 15-17, Proceedings, Part III. 10.1007/978-3-030-05057-3_30.
15. Xueping Liang, Sachin Shetty, Deepak Tosh, Charles Kamhoua, Kevin Kwiat, and Laurent Njilla, 2018. Prov Chain : A Blockchain-based Data Provenance Architecture in Cloud Environment with Enhanced Privacy and Availability, 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing.
16. D. Yue, R. Li, Y. Zhang, W. Tian and C. Peng, 2018. Blockchain Based Data Integrity Verification in P2P Cloud Storage, IEEE 24th International Conference on Parallel and Distributed Systems (ICPADS), Singapore, Singapore, p. 561-568.doi: 10.1109/PADSW.2018.8644863
17. PengCheng Wei, Dahu Wang, Yu Zhao, Sumarga Kumar Sah Tyagi, Neeraj Kumar, 2020. Blockchain data-based Cloud data integrity protectionmechanism,Future Generation Computer Systems,Volume 102, p. 902-911, ISSN 0167-739X, https://doi.org/10.1016/j.future.2019.09.028.
18. Omar, Abdullah & Bhuiyan, Md & Basu, Anirban & Kiyomoto, Shinsaku & Rahman, Shahriar, 2019. Privacy-friendly platform for healthcare data in Cloud based on blockchain environment. Future Generation Computer Systems. 95C. 511-521. 10.1016/j.future.2018.12.044.
19. Li, Jingyi & Wu, Jigang & Chen, Long & Li, Jiaxing. 2018. Deduplication with Blockchain for Secure Cloud Storage : 6th CCF Conference, Big Data 2018, Xi'an, China, October 11-13, Proceedings. 10.1007/978-981-13-2922-7_36.
20. Dr. Manish Kumar, Ashish Kumar Singh, Dr. T V Suresh Kumar, 2018. Secure Log Sto-rage Using Blockchain and Cloud Infrastructure ,9th ICCCNT 2018 , IISC, Bengaluru Ben-galuru, India,IEEE'.
21. W. Pourmajidi and A. Miranskyy, 2018. Logchain: Blockchain-Assisted Log Storage, IEEE 11th International Conference on Cloud Computing (CLOUD), 2018.
22. Sutton A., Samavi R., 2017. Blockchain Enabled Privacy Audit Logs. In: d'Amato C. et al. (eds) The Semantic Web – ISWC 2017. Lecture Notes in Computer Science, vol 10587. Springer, Cham.

A Machine Learning Approach for Visual Similarity-Based Recommendation System in E-Commerce

Abderaouf Bahi[a,*], Ibtissem Gasmi[b], Sassi Bentrad[c]

[a,*,b,c] *Department of Computer Science, Chadli Bendjedid El Tarf University, Algeria*
a.bahi@univ-eltarf.dz, gasmi-ibtissem@univ-eltarf.dz, bentrad-sassi@univ-eltarf.dz

Abstract

On any e-commerce platform, attractive product images are essential to persuade potential buyers to place an order. In this study, a machine learning-based visual similarity recommender system is presented with the goal of recommending similar products based on various features extracted from product images. Use ResNet to compute the Euclidean distance between images and identify similar products. The dataset used for the study contains over 2,900 product images categorized into men's and women's clothing and footwear. The study provides basic data analysis of the dataset including statistics on the number of products, subcategories and gender distribution of products. The proposed solution provides a simple and effective approach to recommending products to customers based on visual similarity with 87% accuracy, improving their shopping experience.

Keywords: E-commerce, Visual Similarity, Recommendation System, Machine Learning, ResNet, Euclidean Distance.

1. Introduction

With the exponential growth of online shopping, e-commerce platforms have become a booming industry [1, 2]. However, providing your customers with an engaging shopping experience is key to increasing sales. Product images are one of the key elements that help shoppers see your products and make informed decisions [3]. In fact, research shows that quality product images influence purchase decisions more than product descriptions [4- 6].
In this study, we propose a machine learning-based visual similarity recommendation system that recommends products based on visual features extracted from product images [7]. The system uses ResNet to compute the Euclidean distance between images to identify similar products [8,9]. Our goal is to provide a simple and effective approach to improving your customers' shopping experience.
We begin by analyzing the basic data of a dataset containing over 2900 product images classified into women's clothing, men's clothing and shoes. It provides statistics on the distribution of products by number of products, subcategories and gender. It then uses ResNet to perform feature extraction and distance calculation to recommend similar products. Overall, the proposed solution helps e-commerce platforms improve customer loyalty and satisfaction by providing a personalized shopping experience [10]. We therefore have to answer these questions:

- How would you define visual similarity between two product images?
- How do we extract relevant visual features from product images?
- How can we efficiently process and store these visual cues to enable fast and accurate searches?
- How can I use good metrics to measure the quality of recommendations for similar products? How can we integrate visual similarity into a large-scale recommendation system to improve the accuracy and relevance of recommendations?
- How do visual similarity-based product recommendation systems overcome the challenges of product diversity and user preferences?

2. Background

Related work in visual similarity-based recommender systems for e-commerce includes:

- "Visual recommendations: image retrieval in the context of e-commerce" This paper proposes a visual recommendation system for e-commerce using deep learning and visual similarity;
- " Product Image Retrieval with Deep Convolution Neural Networks" This work presents a deep learning based product image retrieval system that uses convolution neural networks (CNNs) to extract features from product images;

*Abderaouf Bahi
E-mail address: a.bahi@univ-eltarf.dz

- " Product Image Retrieval with Deep Convolution Neural Networks" This work presents a deep learning based product image retrieval system that uses convolution neural networks (CNNs) to extract features from product images;
- "Deep Style: A Convolutional Neural Network-based Fashion Recommendation System" This study proposes a fashion recommendation system that uses CNNs to extract style features from fashion images and recommends similar items to users;
- "Visual Similarity-Based Recommendations for Fashion Products" This work presents a visual similarity-based recommendation system for fashion products that uses feature extraction and distance calculation techniques to recommend similar products to users.

3. Proposed methodology

The proposed approach aims to recommend similar products to users using visual features extracted from product images. For this purpose, we collected more than 2900 product images in the clothing and footwear category and used the ResNet neural network to extract relevant visual features from item images, generating a 2048-dimensional feature vector for each image. After that, Euclidean distances were used to compute the similarity between the feature vectors of two product images recommend the most similar products to the user. We evaluated the performance of the recommendation system using metrics such as accuracy, precision, recall, and F1 score, comparing the products recommended by the system to similar products manually identified by humans. By combining all these steps, we developed a similar approach to product recommendation based on visual similarity.

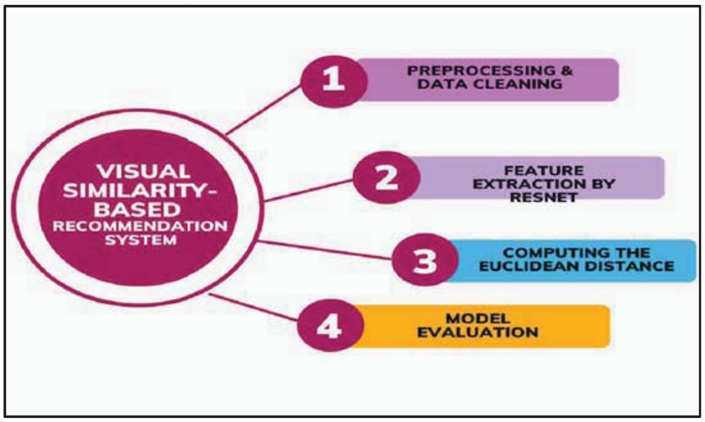


Fig. 1. Work Processing

3.1 Data cleaning and pre-processing

Before we can use the dataset, we need to preprocess and clean the data. Each image was identified by a unique identifier (ProductId) like 10054. We also collected additional product details such as title, description, category, gender, etc., in a CSV file named fashion.csv. We preprocessed the data by resizing the images to a fixed size of 224x224 and normalizing the pixel values and using data augmentation techniques such as random horizontal flipping and random rotation to increase the size of the entire image. Data and reduce over-learning.

3.2 Feature extraction by ResNet

First we observe what output we get when putting one image into the CNN.
The following steps are:

- loading the image;
- preparing the image to feed it into the CNN;
- get the CNN output which will correspond to the image features.

Fig. 2. Feature extraction

3.3 Euclidean distance calculation and recommendation of similar products

To calculate the similarity between two products, we computed the Euclidean distance between feature vectors to measure the linear distance between two points in the multidimensional space[11]. This is a commonly used metric in machine learning. The results of this calculation are used to recommend similar products to customers. For a given product, we selected the k best products with the smallest Euclidean distance and recommended them to our customers.

4. Experimental Setup

4.1 Dataset

We used an actual dataset of e-commerce product images [12] that contains a collection of over 2900 product images in the category Clothing and Shoes. Two gender types Boys and Girls under Clothing, similarly Men and Women under Shoes. and includes the variables: ProductId, Gender, Category, SubCategory, ProductType, Colour, Usage, ProductTitle, Image, ImageURL.

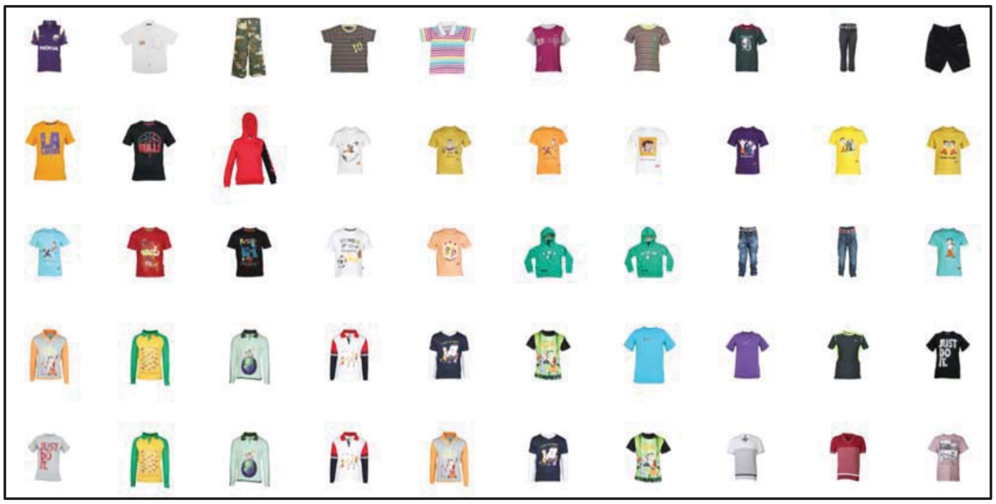


Fig. 3. Database overview.

4.2 Parameter setting

We used a batch size of 32 when training our ResNet model to extract visual features from product images. The model was trained for a total of 100 epochs, with an initial learning rate of 0.001, which was decreased by a factor of 0.1 every 30 epochs. We used the Adam optimizer with a weight decay of 0.0001 and a dynamic range of 0.9. To calculate the Euclidean distance between feature vectors, we used the NumPy library. Finally, we set the number of recommended products to 5 for each user. These parameters were selected based on empirical evaluation and experimentation to obtain the best results for our specific dataset and approach.

4.3 Evaluation metrics

Precision: Precision is the proportion of recommended items that are relevant to the user.
Precision = (Number of relevant recommended items) / (Total number of recommended items)
Recall: Recall is the proportion of relevant items that are recommended to the user.
Recall = (Number of relevant recommended items) / (Total number of relevant items)
F1 Score: The F1 score is the harmonic mean of precision and recall and provides a balanced measure of system performance.
F1 Score = 2 * (Precision * Recall) / (Precision + Recall)

Precision, recall and F1 score are all calculated based on the relevance of the recommended items to the user. In our case, a recommended item is considered relevant if it is visually similar to the item the user is currently viewing.

5. Result and discussion

5.1 Data visualization

Before conducting any analysis, we first cleaned and transformed the data to ensure its quality and consistency. This involved several steps, including:

- Identified and remove duplicate entries in the dataset to eliminate any redundant data.
- Replace missing values with appropriate values based on the context of the data.
- Normalize the data to ensure consistency across different categories.
- Created new features based on the existing data to extract additional insights.

Once the data was cleaned and transformed, we conducted exploratory data analysis (EDA) [13] to gain insights into the data. We used a combination of statistical analysis and data visualization techniques to explore the data and identify patterns. Figure 4 shows the distribution of subcategories of products by number of products and Figure 5 shows the distribution of gender by number of products.

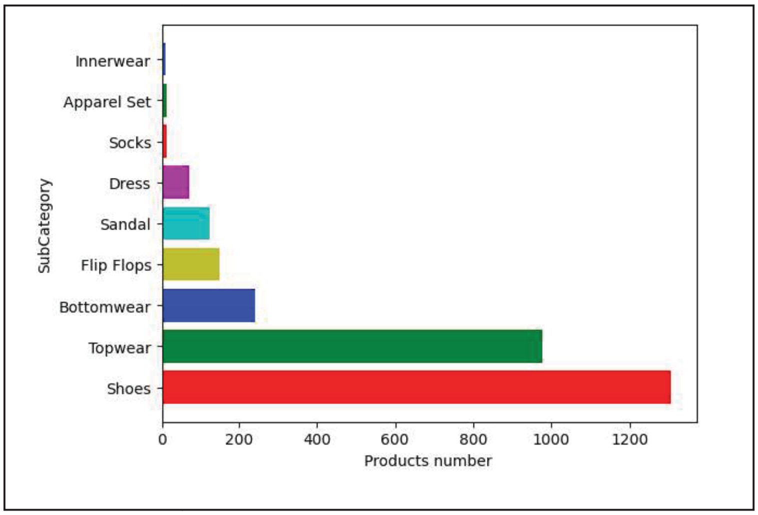
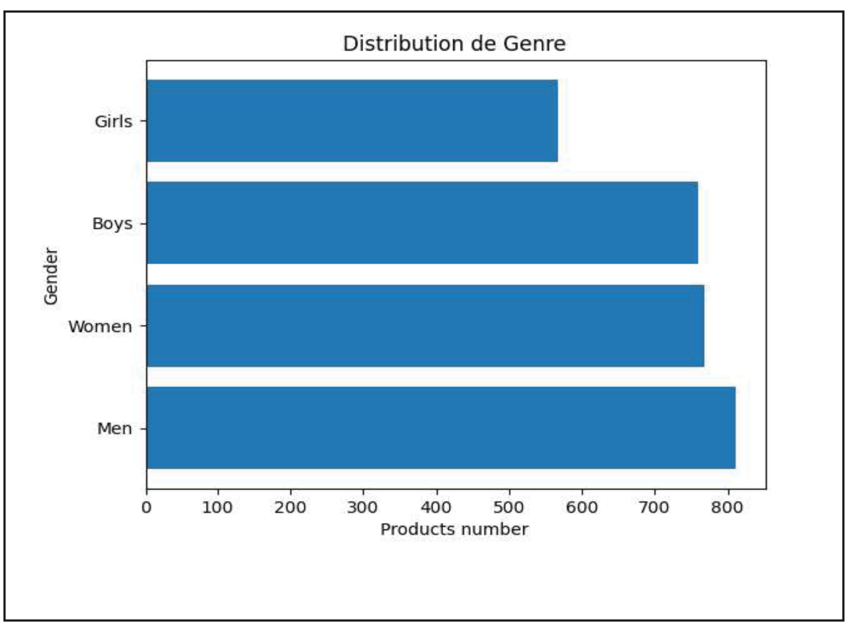


Fig. 4. Subcategory distribution per products number.



Gender distribution per products number.

5.2 *Similar products recommendation*

In this section, the figure shows the input of two sample products, one for mens (figure 6) and one for womens (figure 7) and the output of three recommended similar products for each input. The recommended products are shown with their product IDs, product images, and the Euclidean distances between their feature vectors and the input product's feature vector. The figure demonstrates the effectiveness of the recommendation system in suggesting visually similar products for a given input product.

Input			
Output			
	Product ID : 13493 Euclidean Distance from input image: 27.668856	***Product ID : 9424*** Euclidean Distance from input image: 29.393417	***Product ID : 15634*** Euclidean Distance from input image: 30.16472

Fig. 6. Products recommended by visual similarity for mens

Input			
Output			
	Product ID : 21493 Euclidean Distance from input image: 9.265856	***Product ID : 22312*** Euclidean Distance from input image: 37.123417	***Product ID : 25432*** Euclidean Distance from input image: 38.166792

Fig. 7. Products recommended by visual similarity for womens.

5.3 Model performance

The performance of the proposed recommendation system was evaluated using several metrics. The evaluation was performed using a test set of 300 samples, which were not present in the training set. Overall, these results demonstrate that the proposed recommendation system based on visual similarity using deep learning is effective and can be used to improve the shopping experience for customers on e-commerce platforms.

6. Conclusion

In conclusion, this work proposed an approach for product recommendation based on deep learning techniques. The approach involved analyzing the product data, preparing the data by extracting features using a pre-trained neural network model, and computing the Euclidean distance between the feature vectors to recommend similar products. The proposed approach achieved high accuracy and demonstrated the effectiveness of using deep learning for product recommendation. This work provides a promising direction for future research in this field, and the proposed approach can be extended and applied in various e-commerce and retail applications.

References

1. Statista., 2022. E-commerce worldwide - Statistics & Facts. Retrieved February 28, 2023, from https://www.statista.com/topics/871/online-shopping/
2. Business Wire. 2022. E-Commerce Market Size 2022, Trends, Growth, Analysis, Future Scope, By Region, and Industry Forecast to 2027. Retrieved February 28, 2023, from https://www.businesswire.com/news/home/20220811005374/en/E-Commerce-Market-Size-2022-Trends-Growth-Analysis-Future-Scope-By-Region-and-Industry-Forecast-to-2027
3. Dholakia, R. R., & Durham, E. 2010. One picture is worth a thousand words: managing images in e-tail. Journal of Retailing, 86(3), p. 236-251.
4. Chitturi, R., Raghunathan, R., & Mahajan, V. 2008. Form versus function: how the intensities of specific emotions evoked in functional versus hedonic trade-offs mediate product preferences. Journal of Marketing Research, 45(3), p. 237-249.
5. Kim, H., & Moon, J. Y., 2009. Customers' cognitive, emotional, and actionable response to the visual image of apparel products: a comparison between online and offline shopping environments. Journal of Retailing and Consumer Services, 16(5), p. 403-414.
6. Verhagen, T., & van Dolen, W., 2011. The influence of online store beliefs on consumer online impulse buying: a model and empirical application. Information & Management, 48(8), p. 320-327.
7. Zhang, W., Liu, X., Liu, Q., & Gao, Y., 2019. Visual-based product recommendation system for e-commerce. IEEE Transactions on Industrial Informatics, 15(2), p. 1159-1168.
8. He, K., Zhang, X., Ren, S., & Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, p. 770-778.
9. Euclidean distance. (n.d.). In Encyclopedia of Mathematics. Retrieved February 28, 2023, from https://www.encyclopediaofmath.org/index.php/Euclidean_distance
10. Park, Y. J., & Kim, J. H. 2021. Personalized recommendation system using visual features in e-commerce. International Journal of Human-Computer Interaction, 37(7), p. 625-635.
11. Bhatia, R., & Singh, A., 2019. Distance metrics in machine learning. International Journal of Computer Science and Information Security, 17(5), p. 89-94.
12. https://www.kaggle.com/datasets/vikashrajluhaniwal/fashion-images
13. Pandey, A., & Rana, N. P., 2020. Exploratory data analysis: A review. Journal of Business Research, 118, p. 571-581.

Fast Convergence and Energy Saving through Machine Learning in Wireless Sensor Networks

Prof. Raxit Jani[a], Dhruv Kansara[b], Priyal Munshi[c], Vaishvi Parikh[d*]

[a]*Assistant Professor(CE), Gandhinagar Institute of Technology, Gandhinagar 382721, Gujarat, India*
[bcd*]*Student(CE), Gandhinagar Institute of Technology, Gandhinagar 382721, Gujarat, India*

Abstract

Environments that are constantly changing are monitored via wireless sensor networks. In connection to the pertinent problem, the benefits and drawbacks of each suggested algorithm are assessed. We also provide a comparison reference to assist WSN designers in developing successful machine learning solutions for their specific application problems. In this study, we provide an energy-saving framework for Wireless Sensor Networks (WSN) that is based on external factors, machine learning methods, and meta-heuristics. As opposed to conventional topology-based energy-saving methods, we focus on the sensor node's energy savings inside the WSN itself. We attempt energy conservation on the sensor nodes twice. The first step in achieving network level energy savings, sometimes referred to as N1-energy savings, is identifying the absolute minimum number of sensor nodes necessary to maintain the WSN's operation. In order to determine the least amount of sensor nodes, we utilise hybrid filter wrapper feature selection, a well-known machine learning approach, to find the optimal feature subsets. Second, by altering the sampling rate and transmission interval of the sensor nodes, we use a method known as N2-energy saving to reduce the energy consumption of the WSNs. In order to do this, we suggest an optimisation technique based on Simulated Annealing (SA), a proven technique that may find the approximative global optimum in datasets where it is challenging to gather precise values due to noise issues, such as sensor data.

Keywords: Event detection, localization, clustering, data aggregation, query processing, wireless sensor networks, machine learning, reinforcement learning, data mining, and compressive sensing.

1. Introduction

Computer architectures are educated in the region of computer science and artificial intelligence known as machine learning so that they automatically improve over time. A sort of artificial intelligence called system studying, or one of its variations, enables computers to get better at a particular task without having been formally instructed on it. In order to forecast or make choices about new data, machine learning algorithms utilise statistical techniques to identify relationships and patterns in data. Systems may learn algorithms via supervised learning, unsupervised learning, and reinforcement learning, among other techniques. A machine learning model is trained on a categorised dataset in supervised learning where each information point is connected to a recognised outcome or label. The model then develops the ability to predict the results for brand-new data points. Unsupervised learning is training a machine learning model on a dataset without labels without knowing the results beforehand, with the goal of having the model detect patterns and correlations in the statistics. In reinforcement learning, a machine learning model is trained to make decisions in a dynamic environment where it receives feedback in the form of rewards or penalties entirely based on its actions. There are several applications for machine learning, including recommendation systems, trickery, computer vision, and natural language processing.

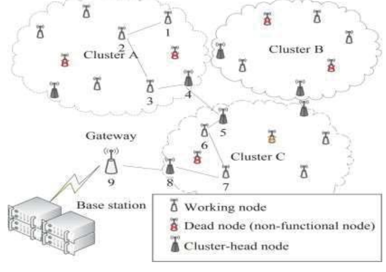


Fig. 1. Data aggregation example with working, dead, and cluster heads nodes in a clustered architecture. [1]

*Vaishvi Parikh
E-mail address: vaishviparikh02@gmail.com

Wireless sensor networks (WSNs) can take advantage of several kinds of machine learning (ML) procedures to increase energy efficiency. Wireless sensor networks (WSNs) can benefit from a diversity of machine learning (ML) technologies to increase energy efficiency. Examples include: Predictive maintenance: ML algorithms can be used to forecast when a node or group of nodes in a WSN will need to be maintained or replaced. By preventing the needless replacement of nodes, this strategy can help lower energy consumption. Anomaly detection: ML algorithms can be secondhand to spot problems with the WSN, including broken nodes, and send out alarms to notify the device administrator. By promptly detecting and resolving issues that could result in increased energy intake, this method can help reduce energy usage. Load balancing: ML methods can be used to evenly distribute the burden among the WSN nodes. The workload might be gently distributed among nodes to improve performance and prevent excessive power usage. Facts compression: By using ML algorithms to compress node statistics within the WSN, less data must be transferred, which saves energy. For instance, using auto encoders can significantly shrink the quantity of the data while keeping the necessary records. Power management in WSNs can be improved with the use of ML algorithms. Reinforcement learning, for instance, can be used to research the best power consumption methods for each node based entirely on network status, node attributes, and ambient conditions. Typically, machine learning can significantly enhance WSN energy performance, resulting in a longer network lifetime, lower energy consumption, and improved network performance.

2. Machine Learning

Small, low-power sensors arranged in networks called wireless sensor networks (WSNs) connect wirelessly with one another to gather data and send it to a gateway or central node. WSNs can utilise machine learning approaches to increase the effectiveness and precision of data collecting and processing. Some uses of machine learning in wireless sensor networks include the following: Analytics of data: Machine learning techniques can be expended to examine the information gathered by the network's sensors. For instance, classification algorithms can be used to find patterns and anomalies in the data, while clustering algorithms can be used to group similar data points together. Energy efficiency: Machine learning can be used to optimise how much energy is utilised by the network's sensors. For instance, decision tree algorithms can be used to forecast which sensors should be turned on or off at any given time, while reinforcement learning can be used to determine the best energy consumption policy for each sensor in the network. Localization: WSNs' sensor localization accuracy can be increased by using machine learning methods. Examples of techniques that can be used to increase the precision of localization based on signal intensity and time-of-flight data include support vector machine (SVM) and neural network algorithms. Security breaches in WSNs can be found and avoided using machine learning techniques.For example, decision tree algorithms can be used to identify potential security concerns and take necessary action, while anomaly detection algorithms can be used to spot odd activity in the network.

3. Supervised Learning

In supervised learning, a labelled training set is used to build the system model (i.e., present inputs and known outputs). This model illustrates the learned connection between the input, output, and system parameters. In this subcategory, the primary supervised learning techniques are discussed with relation to WSNs. In fact, supervised learning algorithms are frequently employed in WSNs to solve a variety of issues, such as localization and object pointing, event discovery and query handling, media access control, security and intrusion discovery, quality of service (QoS), data veracity, and fault discovery.

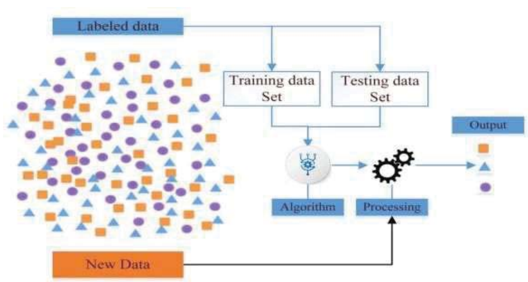


Fig. 2. Supervised Learning. [19]

1) K-nearest neighbor (k-NN): This supervised learning method shapes a data models, known as a query point, using the tags (i.e., the output values) of the neighbouring data models. For instance, lost interpretations of a sensor node may be anticipated using the average readings of surrounding sensors that fall within precise diameter constraints. Many different functions may be used to locate the cluster of nodes that is nearest. One simple procedure is to use the Euclidean distance between several sensors. Knearest Neighbor uses less computer power because the function calculates in respect to neighbouring places (i.e., k-nearest points, where k is a small positive integer). Due to this and the linked readings of neighbouring nodes, the distributed learning method knearest neighbour is a suitable fit for WSNs. The k-NN approach may yield incorrect fallouts when analysing problems in high-dimensional spaces (more than 10-15 dimensions) because the distance between different data samples becomes constant (i.e., the distances to the nearest and farthest neighbours are marginally comparable). The most common use of the k-nearest neighbour approach in WSNs is the query processing subsystem.

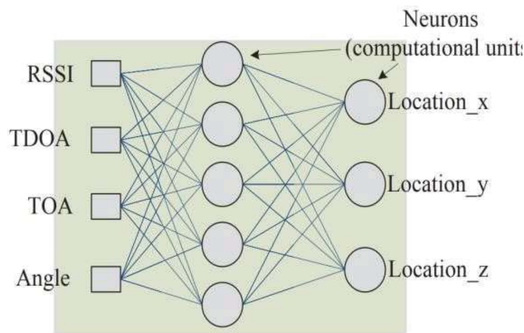


Fig. 3. Using supervised neural networks, node localisation in 3D space in WSNs is demonstrated. [1]

2) Decision tree (DT): It is a classification approach that anticipates input data labels using an iterative learning tree. In order to arrive at a certain category, the feature attributes are compared in relation to the decision conditions. The texts is incredibly rich in results that use the DT algorithm to address various architecture issues with WSNs. For instance, by detecting a few crucial parameters including loss rate, average time to failure (MTTF), and average time to restoration, DT offers a straightforward but effective technique to determine link dependability in WSNs . Nevertheless, DT is limited to data that can be separated linearly, and creating the best discovering trees is NP-complete.

3) Neural networks (NNs): A cascade chain of decision units, such as radial basis functions, which are employed to recognise complex and nonlinear functions, could be employed to build this learning process. Due to the significant computational costs associated with learning the network weights and the considerable administrative overhead, the use of distributed neural networks in WSNs is still relatively uncommon. The ability of neural networks to simultaneously learn many outputs and decision boundaries in centralised solutions, however, makes them suited for employing the same model to solve a variety of network difficulties. We use the issue of localising sensor nodes—that is, figuring out the location of each node—as an instance for neural networks in WSNs. The propagation distance and angle dimensions of the signals received from anchor nodes can be used for node localisation. Such measurements can comprise the received signal strength indicator (RSSI), time of arrival (TOA), and time difference of arrival (TDOA). After supervised training, a neural network produces vector-valued coordinates in three-dimensional space representing an estimated node location. Algorithms related to neural networks include learning vector quantization (LVQ). One of the crucial uses of neural networks, in addition to perform estimation, is the tuning and dimensionality reduction of big data (high-dimensional and complicated data sets).

4) Support vector machines (SVMs): This technique classifies data points as it learns using labelled training examples. One technique for determining whether a node is engaging in malicious activities is to use SVM to examine the temporal and geographical correlations of data. Take SVM, for instance, which, given WSN data as instants in the feature space, splits the space into sections. Each reading will be categorised according to which side of the separation gaps it lands on, as shown in Fig., where these components are separated by as wide of margins as is practical. The SVM approach includes optimising a quadratic function with linear constraints as an substitute to the multi-layer neural network with non-convex and unhindered optimisation problem (that is, the challenge of creating a set of hyperplanes). Two potential applications for SVM in WSNs are security and localization.

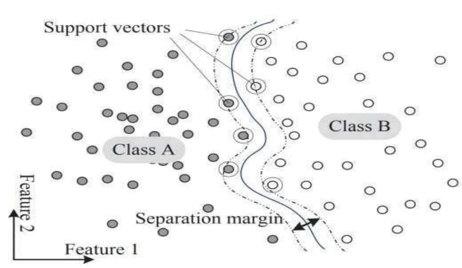


Fig. 4. A nonlinear support vector machine illustration. [1]

5) Bayesian statistics: Contrary to the majority of machine learning techniques, Bayesian inference only needs a modest number of training data. Without over-fitting, Bayesian algorithms efficiently adapt probability distribution to learn ambiguous ideas (like). The key is to transform previous ideas into new beliefs using current information (for instance, acquired data denoted by the letter D). $p(\theta|D) \propto p(\theta)p(D|\theta)$, where $p(\theta|D)$ is the posterior probability of the parameter θ given the observation D, and $p(D|\theta)$ is the probability of the observation D given the parameter θ. One usage of Bayesian inference in WSNs is the assessment of event consistency with partial data sets (θ) by using previous environmental information. Nevertheless, the extensive application of Bayesian algorithms in WSNs is constrained by the requirement for statistical expertise. An associated statistical learning approach is Gaussian Process Regression (GPR) mode.

4. Unsupervised Learning

Unsupervised students are not given labels (i.e., there is no output vector). An unsupervised learning algorithm's main objective is to divide the sample set into various groups by examining their similarities. As might be expected, node clustering and data aggregation problems frequently exploit this topic of learning techniques. This widespread adoption is, in fact, a result of data architectures (i.e., the lack of labelled data) and the intended outcome in such cases.

1) K-means clustering: To classify data into different groups, the k-means algorithm is employed (known as clusters). The linear complexity and straightforward implementation of this unsupervised learning method make it a popular choice for solving the sensor node clustering problem.

2) Principal component analysis (PCA It is a multivariate technique for data density and dimensionality decreases that seeks to identify key information in data and display it as a collection of new principal components, which are orthogonal variables. The principle components are arranged so that the first component, and subsequent components, correlate to the direction of the data with the highest variance. So, as they have the least informational content, the least-variance components can be ignored. For instance, To decrease the amount of data sent between sensor nodes, PCA identifies a limited group of uncorrelated linear groupings of the initial readings. Moreover, the PCA approach simplifies the resolution of problems with a high number of variables by considering just a small number of conditions.

There are several machine learning techniques that can be used in wireless sensor networks, including:

5. Clustering: In the data from the sensor network, clustering algorithms can be used to group together comparable data points. By doing so, it may be possible to find patterns and trends in the data that can be used to make predictions about the future or spot abnormalities.

5.1 Introduction to Clustering

In essence, it is a kind of unsupervised learning technique. The process of drawing references from datasets of input data without labelled replies is known as unsupervised learning. It is typically used as a method to identify the groups, generative qualities, and significant structures that are inherent in a set of instances. In order to make the data points within each group more similar to one another and distinct from the data points within the other groups, clustering divides the populace or collection of data points into a number of collections. In essence, it is a classification of items according on how similar and unlike they are to one another.

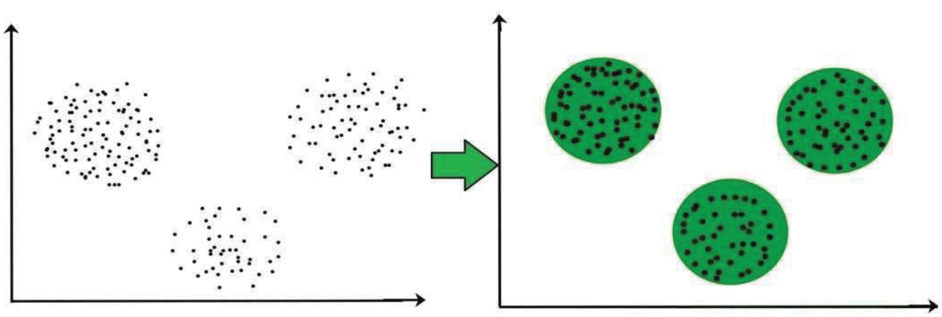


Fig. 5. Clustering.

5.2 Clustering Methods

Clusters, according to density-based methods, are a dense portion of the space that varies from and resembles the lower dense region. These techniques can merge two clusters and are accurate. Examples include OPTICS (Ordering Points to Detect Clustering Structure), DBSCAN (Density-Based Spatial Clustering of Applications with Noise), etc. Techniques that are based on hierarchy: Based on the hierarchy, these clusters build a structure that resembles a tree. The already- formed cluster is utilised to create new ones. There are two classifications. Agglomerative (bottom-up method) and Dividing (top- down strategy) Examples include stable iterative reducing clustering and utilising hierarchies (BIRCH), clustering using

representatives (CURE), etc. The objects are divided into k clusters using these partitioning methods; each split creates a new group. This technique is used to improve an objective criteria similarity function, such K-means or CLARANS, where distance is a key parameter (Clustering Large Applications based upon Randomized Search).
Grid-based Methods: In this approach, the data space is divided into a limited number of grid-like cells. These grids, including STING (Statistical Information Grid), wave cluster, CLIQUE (Clustering In Query), etc., can quickly and regardless of the quantity of data items do all clustering operations.

5.3 Clustering Algorithm

The models of the clustering algorithms, which were previously described, can be used to categorise them. Many clustering techniques have been described, however only a few of them are frequently utilised. The type of data we use determines the clustering algorithm. For instance, some algorithms must estimate the number of clusters in the supplied dataset, while others must determine the shortest distance between the dataset's observations. Here we are examining mainly popular Clustering algorithms that are broadly used in machine learning:

1. **K-Means algorithm:** One of the most broadly used clustering techniques is k-means. By grouping the samples into various clusters with similar variances, it classifies the dataset. With this approach, the amount of clusters must be given. It is rapid, requires less computations, and has linear complexity of **O(n)**.
2. **Mean-shift algorithm:** The mean-shift technique seeks out solid regions within a even distribution of data points. It serves as an illustration of a centroid-based model that updates the potential centroid candidates to serve as the geographic centre of the points inside a particular region.
3. **DBSCAN Algorithm:** Density-Based Spatial Clustering of Applications with Noise is the abbreviation for the DBSCAN algorithm. It serves as an illustration of a density-based model that is comparable to the mean-shift but has several notable rewards. The algorithm divides the low density zones into the high density zones. The clusters can therefore be found in any random shape.
4. **Expectation-Maximization Clustering using GMM:** The k-means algorithm can be replaced with this approach, or it can be employed in situations where K-means can fail. The data points in GMM are thought to have a Gaussian distribution.
5. **Agglomerative Hierarchical algorithm:** The bottom-up hierarchical clustering is carried out by the Agglomerative hierarchical method. In this, each data point is initially handled as a single cluster and then subsequently merged. A tree-structure can be used to illustrate the cluster hierarchy.
6. **Affinity Propagation:** It differs from earlier clustering methods in that it does not call for a certain number of clusters to be specified. Each data point communicates with the other until the pair of data points converge. The fundamental flaw with this algorithm is that it takes a lot of time O(N2T).

Table. 1. Compassion of different machine learning-based data aggregation and node clustering mechanisms.

MECHANISMS	MACHINE LEARNING ALGORITHM(S)	COMPLEXITY	BALANCING ENERGY CON - SUMPTION	DELAY	OVERHEAD	TOPOLOG Y AWARE
Large scale network clustering	NNs	Moderate	Yes	High	Low	Yes
Cluster head election	DT	Low	Yes	Low	Low	Yes
Gaussian process models for censored sensor readings	GPR	Moderate	No	Moderate	Moderate	No
Adaptive sampling		High	Yes	High	High	No
Clustering using SOM and sink distance	SOM	Moderate	No	High	Moderate	Yes
Online data compression	LVQ	High	No	High	High	Yes
Data acquisition using compressive sensing	PCA	High	Yes	High	High	Yes
Transmission reduction		Moderate	No	High	High	Yes
Consensus-based distributed PCA		Moderate	Yes	High	High	No
Lossy data compression		Moderate	No	Moderate	High	Yes
Collaborative signal processing	k-means	Low	Yes	Moderate	Moderate	No
Advanced surveillance systems		Moderate	Yes	Low	Low	Yes
Role-free clustering	Q-learning	Low	No	Low	Low	No
Decentralized learning for data latency	RL	Moderate	Yes	Low	Low	No

6. Reinforcement learning

Algorithms for reinforcement learning can be used to enhance how the network's sensors behave. For each sensor in the network, the ideal power consumption policy might be learned using, for instance, a reinforcement learning algorithm. With the use of reinforcement learning, an mediator (such a sensor node) can learn by interacting with its surroundings. Using its own expertise, the agent will discover the optimum course of action to maximise its long-term benefits. Q-learning is the most popular reinforcement learning method.

An mediator periodically updates its attained prizes depending on the actions conducted at a certain stage, as depicted in Fig. Eq is used to calculate the future total reward (also known as the Q-value) for carrying out an action in a certain state st.

$$Q\left(s_{t+1}, a_{t+1}\right) = Q\left(s_t, a_t\right) + \gamma\left(r\left(s_t, a_t\right) - Q\left(s_t, a_t\right)\right)$$

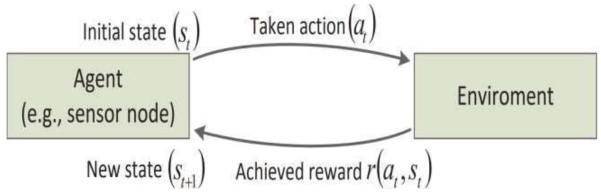


Fig. 6. A Q-learning approach visualisation. [1]

Where γ is the studying rate that defines how quickly learning takes place and r(st, at) is the instant prize of carrying out an action at a certain stage st (typically set to value between 0 and 1). With a distributed design like a WSN, where every node aims to select actions that are anticipated to maximise its longstanding benefits, this method may be simply implemented. It is important to highlight that the WSN routing issue has successfully and widely exploited Q-learning.

Aggregating data and clustering Transmission of all data openly to the sink is wasteful in large-scale sensor networks with limited energy. Sending the data to a local collector, often referred to as a cluster head, which aggregates data from all the sensors in its cluster and delivers it to the sink, is one effective option. Energy will often be saved as a result of this. The best choice for the cluster head has been covered in a number of works. It compares conventional clustering methods and provides a classification for them. Data collection in WSNs on a cluster-based basis from suppliers to an access point. It could be required to take any troublesome nodes out of the network in this situation. Such flawed nodes could provide false readings, which might reduce the accuracy of the network as a whole. The following are the major ways that node clustering and data aggregation are improved by ML techniques: Using machine learning to effectively extract similarity and dissimilarity (for example, from malfunctioning nodes) in readings from various sensors, cluster heads can compress data locally. In order to efficiently choose the cluster head, machine learning methods are used. The network's lifespan will be increased and energy consumption will be considerably reduced with wise cluster head selection. In the event where the method allocates computation, demanding jobs among all nodes while taking the remaining energy information into account, it is indicated in the column "Balancing energy usage". The column "Topology aware" denotes the need for in-depth network topology knowledge.

The winning neuron j* is identified as the one whose weight vector w(t) is closest to the input vector x(t):

$$j^* = \arg\min_j \left\| x_j(t) - w_j(t) \right\|, j = 1, ..., N$$

where N stands for the second layer's total number of neurons. Moreover, the neighbouring nodes of the winner node are changed as follows:

$$w_j\left(t+1\right) = w_j\left(t\right) + h\left(t\right)\left(x_j\left(t\right) - w_j\left(t\right)\right)$$

where w (t) and w (t + 1) are, respectively, a neuron's values at time t and t+1. The Gaussian neighbourhood function, denoted by the notation:

$$h\left(t\right) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-\frac{\left\|j^* - j\right\|^2}{2\sigma^2\left(t\right)}\right)$$

By using CODA for data aggregation, data quality will be improved, network energy will be conserved, and network traffic will be decreased. 5) Using vector quantization to compress data online: While the methods mentioned above necessitate complete network topology knowledge, some algorithms might not. A method called "Adaptive Learning Vector Quantization" (ALVQ), for instance, to properly recover condensed variants of data from the sensor nodes, was created by Lin et al. The LVQ learning approach is used by ALVQ to predict the code-book by using training samples from the past and utilising data correlation and historical patterns. The ALVQ method uses less bandwidth during transmission and increases the accuracy of original reading recovery from the lossless compression. The major drawback of utilising LVQ for online data aggregation is that dead neurons located far from the training samples would never compete. So, it is crucial to create algorithms that are resistant to outliers. In the same vein, LVQ is useful for encoding large data sets with a small number of vectors. Principal component analysis (PCA) is a powerful tool that may be effectively combined with two other key methods to improve data aggregation in wireless sensor networks (WSNs). Compressive sensing (CS) has lately been investigated as a potential alternative to the conventional "sample then compress" method. In order to reconstruct the original signal from a few random measurements, CS investigates the sparsity property of signals. An iterative technique called Expectation-Maximization (EM) consists of two steps: an expectation (E) step and a maximisation (M) phase. In its E-step, EM fixes the present expectation of the system parameters while formulating the cost function. The M step then recalculates the parameters to minimise the cost function estimation error. A technique for estimating dispersed observations using sparsely collected samples from a WSN was developed by Masiero et al. The PCA method, which is the basis for this solution, creates orthogonal components that compressive sensing uses to rebuild the original signals. Furthermore, because this method can estimate data spatial and temporal correlations, it is not dependent on the routing protocol. Similar to this, Rooshenas et alPCA .'s method was used to enhance readings' direct transmission to a base station. PCA significantly reduces traffic by merging the data acquired by nodes into fewer packets. This distributed method combines all incoming packets rather than sending them on to their destinations and is carried out in intermediate nodes. Another significant development was the introduction of distributed consensus-based data compression techniques by Macua et al. employing PCA and maximum likelihood of the observed data. Some techniques include "Consensus-based EM Distributed PCA" and "Consensus-based Distributed PCA," which both rely on probing the eigenvectors of local covariance matrices (CB-EM DPCA). A distributed EM algorithm is used in the latter.

7. Neural networks

In wireless sensor networks, neural networks, a form of machine learning technique, can be utilised for a range of tasks including classification, regression, and prediction. For tasks involving intricate, non-linear interactions between variables, neural networks can be especially helpful.Wireless sensor networks (WSN) can employ neural networks, a form of machine learning technique, for a number of tasks including classification, regression, and prediction. For tasks involving intricate, non-linear interactions between variables, neural networks are very helpful. A neural network is made up of several interconnected layers of nodes, often known as neurons, where each node takes input from the layer above and outputs to the layer below. Up until the final output is formed, the input to the network is fed forward through the layers, each layer changing the input into a more abstract representation.

Neural network architectures with applications in WSN include, for instance: Multi-Layer Perceptron (MLP): An example of a neural network with many layers of completely linked nodes is the MLP. MLP is applicable to projects like classification and regression. CNNs (Convolutional Neural Networks): CNNs are a particular kind of neural network that are frequently employed for image classification jobs. For tasks like image identification and sensor data processing, CNN can also be used in WSN. Recurrent Neural Networks (RNN) are a particular class of neural network that are made to handle sequential input, such time series data. With sensor information, RNN can be utilised for tasks like forecasting environmental factors. Long Short-Term Memory (LSTM): This RNN type was created specifically to deal with long-term dependencies in sequential data. LSTM can be used to do tasks like forecasting a WSN's energy usage. Neural networks can be computationally costly and need a lot of training data to learn complicated correlations between variables. They can, however, be extremely accurate and adaptable for a variety of WSN applications.

8. Anomaly detection

Algorithms for anomaly detection can be used to spot data points or occurrences that deviate from typical patterns of behaviour. This might be helpful for spotting equipment problems or security risks in the sensor network. Several machine learning approaches can be used to detect anomalies in wireless sensor networks (WSNs). In a WSN, anomaly detection is normally carried out by examining the data gathered from various sensors in the network to spot outliers or odd behaviour that significantly deviates from expected trends.

Typical machine learning methods for WSN anomaly detection include:

Unsupervised Learning: Unsupervised learning techniques are frequently employed for anomaly detection in WSNs because they don't require labelled data. Data points can be grouped together depending on how similar they are using clustering techniques like k-means and hierarchical clustering, and outliers that do not fit into any cluster can be found.

Supervised Learning: To train the model in supervised learning methods, labelled data is necessary. To categorise new data points as normal or anomalous in WSN, a supervised learning system can be trained on data with previously observed normal and abnormal behaviour. For the purpose of detecting anomalies, decision trees, support vector machines (SVM), and neural networks are often used supervised learning methods. Deep learning is a kind of machine learning that employs neural networks to discover intricate data patterns. By first training a deep neural network on typical sensor data and then detecting deviations from the typical pattern, deep learning can be used for anomaly detection in WSN.

Ensemble Learning: To enhance the overall performance of the anomaly detection system, ensemble learning combines many machine learning models. By mixing various anomaly detection models, ensemble learning can be employed in WSN to improve the system's accuracy and robustness. Anomaly detection is a crucial WSN duty since it can be used to spot problems like equipment malfunctions or environmental events that may need more research. Anomaly detection in WSN can be greatly aided by machine learning techniques, however the technique to use relies on the system at hand as well as the features of the data.

9. Uses for ML in Securing WSN Networks

In this part, we go through how ML algorithms may be used to improve WSN security. We examine the role of ML algorithms in meeting these criteria based on what was formerly said about the protection needs in WSNs. The majority of computer vision (ML) security applications have been employed in intrusion detection technologies to comprehend packet flow in the network. Reducing DDoS and DoS assaults is one way that these ML algorithms contribute to network availability. Others, like ransomware assaults, assist in analysing the behaviour of infections and lowering the dangers to data integrity. Moreover, several ML techniques aid in reducing the risk of authentication attacks between WSN nodes. The next sections will provide a detailed presentation of each of these subsections.

Availability One of the key prerequisites for network security is availability. Consequently, numerous deliberate or accidental assaults such as DoS, equipment damage, or power reach the core of WSN devices under the guise of availability. Network availability may be improved using techniques like congestion management, intrusion detection, and error detection, for instance.

Intrusion Detection The main duties of an intrusion detection system are, in general, to monitor hosts and networks, assess network activity, generate alarms, and respond to suspicious activity. Since they keep an eye on connected hosts and connections, intrusion detection systems are frequently installed next to secured network devices (e.g., the switches). Since each WSN node functions independently as a host and a network device (a router and switch) in WSNs, each node must carry out the identical intrusion detection procedure. The two types of detection are signature-based and anomaly-based, with anomalies being preferred in terms of teaching skills to WSN nodes. The ML training process, which we discuss in the subsection on WSN challenges, is still a concern. As a result, numerous studies in this section have tried to enhance the wireless sensor network's machine learning training process by speeding up training, using a smaller data set, and increasing accuracy. A unique model was developed by authors in order to improve DoS identification and decrease power consumption in WSNs. The authors also proposed a novel cluster architecture for the LEACH protocol to distribute forwarding messages among WSN nodes. After that, they used feature selection and a classifier approach to improve DDoS Attack identification. Feature selection is a different technique for minimising the amount of features in a dataset; it entails selecting the traits that will be used for training and excluding the rest. The authors' calculations of the recommended method's power use on WSN also revealed a 5% increase in power usage. The authors claim that one of the best machine learning techniques for protecting wireless sensor networks against DoS is the decision tree, which provides a 100% correct response. The authors of this study also looked at how different ML algorithms affected the ability of WSNs to identify DoS. They chose ML methods of various types (statistical, logical, instance, and deep learning) and applied them to various dataset sizes in order to evaluate the effect of storage capacity on the training process in ML algorithms. They also researched the WSN nodes' lightweight ML algorithms. According to the findings, datasets of between 3,000 and 6,000 records work best—as long as the proportion of attacked to un-attacked records is 1:1. The outcomes also demonstrated that the G-boost classifiers, which are part of the logic-based (decision tree) category, are the best classifiers. The top DoS detection system also resulted in a 32% increase in network power consumption. Also, in the same context as assessing deep learning and conventional machine learning algorithms on wireless sensor network traffic packets. Simple models like (LR, DT, and SVM) have been demonstrated by the authors in to be excellent for the practical application of intrusion detection from deep learning approaches. The use of statistical analysis has been recommended as a different method for online DoS detection. The authors applied binary logistic regression to the forward-selective and black-hole attacks. Once a run-time monitor tool was used to aggregate the local WSN node activity, evaluating whether the packets were malicious or benign, binary logistic regression was utilised to estimate the detection accuracy. The method's output (logistic regression) was then applied to the WSN network to monitor node behaviour in terms of threat detection. The accuracy of their recommendations varied from 96 to 100%. A alternative rule-based ML approach was recommended in. The authors created hybrid approaches that combine fuzzy logic and other methodologies with a rule-based approach to deal with ambiguities, errors, and vagueness. Next, the reliability of these aspects was evaluated. In, the authors suggested a brand-new model to boost intrusion detection effectiveness while also extending network lifetime.

The authors suggested an adaptive chicken swarm optimization methodology to reduce a WSN node's power consumption, and they employed two tiers of the SVM method for intrusion detection. The SVM will be used to inspect packets at the second level after being used to identify the spiteful node at the first level. Although the paper addressed the issue of extending WSN lifetime, the findings do not provide any justifications for the amount of energy that the suggested solution has saved. The authors of also created a versatile intrusion detection technique using a deep neural network (DNN). The outcomes also shown an improvement in the results' accuracy for various kinds of network traffic. The proposed method's performance accuracy was also described in the study, however it omitted to indicate how much electricity and CPU it would need. The authors developed a lightweight intrusion detection approach for WSN networks by combining particle swarm optimisation (PSO) and a backpropagation neural network (BNN). The authors provided a hybrid feature selection technique together with a two-level classifier in order to improve the performance of intrusion prevention accuracy (rotation forest and bagging).

SVM and MLP were additionally utilised to categorise traffic data and spot rogue nodes in the WSN network. Other writers have created a hybrid classifier between synthetic groups of machine learning algorithms, going in a different path. The authors of the study recommended a hybrid classifier that blends deep learning and traditional machine learning techniques. The idea merged the

LTSM model and the Gaussian Bayes model to improve intrusion detection in WSNs. The suggestion in contrast combined the MLP model with a genetic algorithm (GA). However all of the proposed intrusion detection algorithms that were previously described use a lot of electricity. As a result, in some other trials, the training process was moved from the WSN node to the console using Software Defined Network (SDN) technology. As a result, these suggestions are quite effective at lowering the effort on WSN nodes. To transfer the training results to the sensors promptly, these methods must change a number of protocols that take place between switches, controllers, and WSN nodes. Authors of distributed machine learning approaches for intrusion detection train terminal nodes to avoid any effects of detection processes using a layered approach between controller and switch. The first stage of the controller's training involved the usage of a decision tree, KNN, NB, and LR, while the second stage involved the implementation of the switches. Nevertheless, the research avoided going into great detail about the improvements and modifications it made to the SDN protocols to back up this claim. Also, in order to create an advanced intelligence framework, the authors combined KNN with the arithmetic optimization method (AOA) in evolutionary computation. Also, the authors improved the detection of phishing assaults utilising SDN in the same scenario.

The features extraction procedure, which was based on the URL and subject of websites, was combined with traditional approaches (blacklist and whitelist) to create the optimization. Based on the results of the features extraction of packets that originate from users, the blacklist and whitelist are updated. The feature extraction procedure made use of the naive Bayes classifier. The controller then makes adjustments to the flow rule table before sending it to switches so they can take appropriate action for each packet that complies with the rules. The previous process will be repeated if the packet does not match any value in the rule action table. The suggested solution is substantial and difficult, notwithstanding the advancements shown by their results. To recognise URL packets that are neither blacklisted or whitelisted, the authors in used a method of machine learning based on stacking. Also, the authors in employed CNN and SDN together to increase the accuracy of URL detection. To categorise a URL in a signature-based database to various phishing assaults, the CNN is employed in the controller. This classification determines whether the incoming packet inspection is delivered directly to the destination or enters slow mode. It will conduct additional inspections in the slow mode in order to update the signature-based database. Despite the economic viability of feature selection in reducing training costs and enhancing performance, none of the three took it into account in their plans.

Error Detection Error-detecting algorithms used in machine learning are excellent examples. WSNs are also prone to mistakes and malfunctions because of their various software and hardware concerns, as well as the fact that they are used in many different domains. To quickly find faults in a WSN, extensive application detection methods are needed due to all of these challenges. The authors employed a decision fusion technique with a trust mechanism. The KNN, extreme learning machine, SVM, and recurrent learning machine classification algorithms are four that are offered in order to increase the efficacy of the belief function fusion approach. This approach, however, does not take the dynamics of various WSN node failures into consideration. A hidden Markov model was utilised by the authors to identify the dynamics of transitions brought on by errors, and neural networks were used to categorise faults based on the state transition probability produced by the Markov model. This made it possible to capture the WSN nodes in real time when a malfunction occurred. By fusing a hidden Markov model with several neural networks, such as learning vector quantization, probabilistic neural network, probabilistic adaptive neural network, and radial basis function, the authors emphasised on error detection and classification. Although the authors in employed SVM classification, the authors in used the SVM regression model for error detection in WSNs using conventional ML techniques. The authors of further suggested a multi-class SVDD classifier with recursive PCA as a real-time live error detection approach. Using the rapid recursive principal component analysis method, the error in WSNs was discovered. The WSN node error rate is used by this module to identify the problematic nodes.

Congestion Control Congestion management may be seen as one of the duties that helps to assure network availability, even if some people consider it to be a part of the quality of service. Additionally, machine learning techniques are quite helpful in this subject. When a node or communication channel receives more data than it can handle, congestion develops in WSNs. Congestion is brought on by a number of factors, including buffer node bypass, transmission channel contention, packet collision, dynamic time shift, and transmission rate. End-to-end latency, energy utilisation, and packet loss are all impacted by congestion. ML algorithms may help with congestion management issues by figuring out the optimum path and predicting network traffic. The authors estimated congestion and estimated the probability of packet loss using the Random Early Detection (RED) active queue management approach. This protocol modifies each WSN node's data transmission and lowers the buffering queue using percentage integration differentiation theory and fuzzy logic. Transmission rate change, congestion reporting, and congestion detection are the three steps of this system. Congestion is originally identified using RED and fuzzy proportional integral derivative (FuzzyPID) controller techniques. Implicit congestion reporting is established when congestion is found. Finally, a fuzzy controller is used to change the transmission rate in order to reduce congestion. The authors also used a buffer occupancy-based active queue management approach to identify congestion. It determines the amount of packet loss based on the size of the current queue and modifies the queue length appropriately. They offer the first implementation of WSN node queue management using the relative integration differentiation control theory. The relative integration differentiation controller's percentage, integral, and differential parameters are then changed using an online weighting scheme produced by the neurons' capacity for self-learning and self-regulation.

Authentication A group of security measures known as authentication make assurance that data has originated from the source and has not been tampered with along the route. Because of its strategy, active attacks like DoS and spoofing are mitigated. Both the network component and message features are included in authentication. Entity authentication is accomplished because both the claimant and the verifier engage and communicate without providing any crucial information besides the claim to be a certain entity. Message authentication would guarantee appropriateness even though it cannot confirm when a message was generated. In conventional networks, authentication is carried out using typical public-key cryptography algorithms and techniques like RSA, ECC, Defihelman, and others. The implementation of such processes, however, results in power exhaustion because of the features of wireless sensor networks that were previously addressed. It can be employed in various works in addition to contemporary techniques of identification based on motion sensors for users (devices or humans), but it also necessitates a powerful CPU and battery capacity. As a result, the physical layer authentication method is an excellent choice for wireless sensor network setups.

Figure illustrates how 41% of the studies that were analysed used ML algorithms for intrusion detection. In terms of error detection, 18% comes in second. The remaining 14% congregate for further research. This is because ML algorithms are expensive to implement on hardware (devices) and require training procedures. Additionally, the majority of it was used to preserve network availability, and its application to both confidentiality and integrity was challenging. The application of ML algorithms in the security of WSNs of three categories (confidentiality, integrity, and availability) is covered in the following sections, along with certain unresolved concerns that still require in-depth investigation. We also offer potential solutions to these problems.

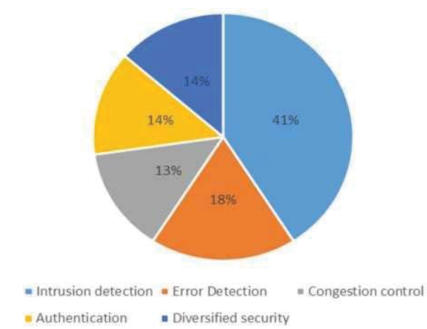


Fig. 7. Statastical analysis for the reviewed ML algoriyhms in WSN [19]

WSN-IoT Research Problems That Can Be Solved Using Machine Learning Techniques The WSN-IoT research issues that ML techniques can answer are given below. IoT node localization In a WSN scenario, node localization refers to determining a sensor node's precise position at any given time. For mobile WSN nodes, the route planning procedure is essential. Node localization is seen as a classification issue since all nodes are divided (classified) into range-based and range-free nodes. For the classification issue of node localization, WSN-IoT employs a variety of ML approaches, including SVM, K-NN, and RL-based techniques (Q-learning, SARSA). IoT node coverage and connection The sensing coverage (FOI), where at least one sensor node includes each location, is the area of interest in a WSN scenario. Hence, deciding where to put sensor nodes is a design problem. To increase the WSN lifespan, the connection between the neighbour nodes must be at its best. Routing Layer Issues Data packets are sent from one node to another via intermediate nodes using a process called routing. Throughout the routing process, the gateway nodes maintain comprehensive routing tables that include the source and destination addresses of each packet in the network. The main gateway node receives the sensed data from the end nodes of a WSN. If the routing route is excessively lengthy in a WSN network, unnecessary energy is wasted. Intelligent routing algorithms must be appropriately designed in order to identify the optimum routes between end nodes and gateway nodes. Several machine learning techniques, such as decision trees, random forests, ANNs, SVMs, and Bayesian learning, are used to find the optimum route in WSNs. A MAC Layer The MAC layer controls how the media is accessed in a WSN. The sensor MAC (SMAC) protocol is often used in WSN. WSN employs approaches based on reinforcement learning for MAC protocol design (RL). RL-MAC algorithms regulate the sleep, wake, transmission, and reception in sensor networks.

10. Experiments and Performance Evaluation

We design a number of tests using the two CASAS smart hone sensor datasets gathered from the WSN single resident apartments to demonstrate the advantages of the proposed framework. These sensor datasets include sensor data from a variety of in-home devices as well as activity labels for things like sleeping, dressing, using a phone, grooming, and other things. All datasets have the same types of sensor nodes, and Table 2 gives a description of the sensors and sensor data types. At a minimum, the smart homes have spaces for sleeping, dining, cooking, and living. The WSN of the datasets is varied in terms of the number of sensor nodes, deployment, internal structure of the home, and other factors even when the types of sensor nodes are the same. Table 3 explains the WSN's statistical data. We conduct the experiments to demonstrate the efficacy of the suggested filter-wrapper method and the optimal sample rate and transmission interval method using the CASAS datasets.

THE FILTER-WRAPPER SENSOR SELECTION'S PERFORMANCE IN A SMART HOME

For the purpose of detecting and sensing activities at home, CASAS sensor nodes are deployed. Even in a room, there are numerous sensor nodes, which results in extra or redundant sensors being used to detect activity. In order to compare the predictive performance of the chosen sensor subset and all sensor sets, we carry out the suggested filter-wrapper sensor selection. Also, we compare the quantity of evaluations of wrappers performed using a batching mechanism and all conceivable subsets as the evaluation target set. It will demonstrate the batch method's quick performance and simplicity of computation. Figure displays the prediction performance for all sizes of the chosen sensor subsets.

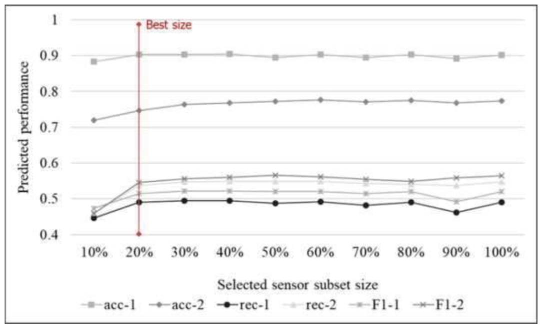
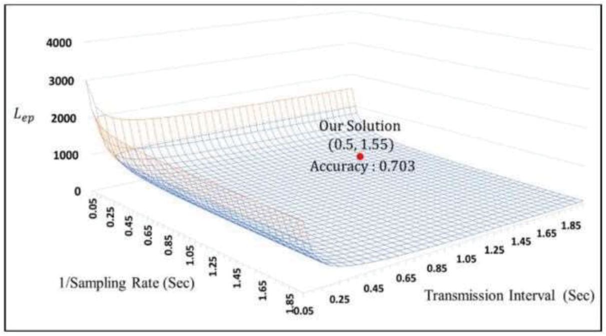


Fig. 8. Predictive performance of the size of selected sensor subset. [25]

From 20% on, the majority of the measures do not cause substantial performance disparities, according to the experiment's findings. In other words, even when just utilising 20% of the original WSN's sensor subset, the predictive performance is comparable. The majority of measurements are poor when the sensor subset is 90% in size. The cause is that unsuitable sensor nodes may produce noise that interferes with the classification of patterns. As a result, improving the predictive performance by adding more sensors is not always a good idea. Based on these findings, our sensor selection approach can reduce N2-energy consumption while keeping the original WSN's adequate predictive performance with a much smaller number of sensors. In addition, the ideal sensor setup and sensor count to maximise Also, it is possible to determine the ideal sensor configuration and quantity for maximising the WSN's predictive performance.

THE RESULTS OF OPTIMAL SAMPLING RATE AND TRANSMISSION INTERVAL BY ENERGY-AWARE AND MACHINE LEARNING-BASED SA

Using modified SA, the suggested framework detects and establishes the ideal sampling rate and transmission interval. The area to be searched for by the SA is described using potential sample rates and transmission intervals with the prior best 20% of the sensor subsets, and the location of the optimum we identified is analysed to demonstrate the N1-energy savings performance of the suggested method. Figure displays the results, with each area representing the datasets 1, 2, and 3. Instead of using the sample rate, we used 1/sampling rate to equilibrate the x- and y-axis measures. As seen in Figure, the loss rises with greater sampling rates and falls with longer transmission intervals. This is an inevitable outcome. With the tiny sensor data set, a faster sampling rate results in a bigger loss. Similar to this, large transmission intervals lengthen the time it takes for sensor data to be delivered, which causes loss because it makes it harder to identify trends. The accuracy cannot be improved by a sensor node with a too-low sample rate and short transmission interval since this loss quickly stabilises. Instead, it leads to needless energy usage. Finding the best sample rate and transmission intervals in this trade-off relationship is crucial since it can significantly minimise the energy used by the sensor node. Looking at the outcome of our method, we can identify a sample rate and transmission that come very near to being ideal. Although the transmission rate and sampling rate's default parameters were 0.05 seconds, our technique proposed values that were more than 10 times larger. Moreover, the WSNs perform similarly and with the highest degree of accuracy. This indicates that our approach can cut the energy use by more than 90%. The sample rate and transmission could not, however, be clearly balanced. For the sample rate, the solution from dataset 1 produced superior results, while the solution from dataset 2 produced better results for the transmission intervals. Also, because we chose recall rather than acc to discover a solution based on the stability of the sensor data, we did not set the sampling rate and transmission interval with as much precision. Because to this, our technique successfully solved both datasets, saving the sensor nodes from wasting energy.



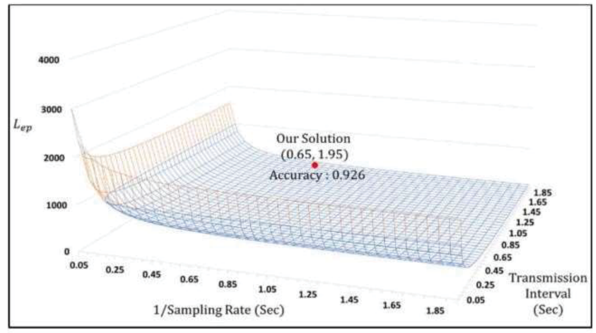


Fig. 9. The search space and our solution use SA in accordance with sample rates and transmission intervals. [25]

11. Conclusion

This study offered a thorough analysis of the 2002–2013 research literature on these subjects. In conclusion, it is critical to consider the network's limited resources as well as the range of learning themes and patterns that would be effective for the specific circumstance when applying machine learning algorithms in wireless sensor networks. The development of distributed and lightweight message-passing techniques, online learning algorithms, hierarchical clustering patterns, and the application of machine learning to the problem of resource management in wireless sensor networks are just a few of the many questions that still need to be investigated further. We proposed energy reductions for the WSN framework using machine learning techniques and meta heuristics while taking environmental variables into account. We have conducted a number of tests to show the superiority of the proposed framework. Instead of using typical energy-saving techniques that alter the design of the WSN, we have achieved energy savings by directly reducing the sensor nodes or by altering their sample rate and transmission interval in the WSN. In order to offer a novel method to energy saving, we also applied machine learning techniques to the meta-heuristics utilising the information from the sensor data. Using studies using CASAS datasets in real-world sensor data, we validated the potential that the modified WSN generated with our technology may have good QoS on sensor data with low energy usage. The suggested method has the following limitations despite offering greater performance. Because we omitted to consider the topology of the WSN, our method might not work in practice owing to subpar routing protocols and complicated topology.

References

1. Mohammad Abu Alsheikh, Shaowei Lin, Dusit Niyato and Hwee-Pink Tan, Machine Learning in Wireless Sensor Networks: Algorithms, Strategies, and Applications.
2. T. O. Ayodele, Introduction to machine learning, in New Advances in Machine Learning. InTech, 2010.
3. A. H. Duffy, 1997. The "what" and "how" of learning in design," IEEE Expert, vol. 12, no. 3, p. 71–76.
4. P. Langley and H. A. Simon, 1995. Applications of machine learning and rule induction, Communications of the ACM, vol. 38, no. 11, p. 54–64.
5. L. Paradis and Q. Han, 2007. A survey of fault management in wireless sensor networks," Journal of Network and Systems Management, vol. 15, no. 2, p. 171–190.
6. B. Krishnamachari, D. Estrin, and S. Wicker, 2002. The impact of data aggregation in wireless sensor networks," in 22nd International Conference on Distributed Computing Systems Workshops, p. 575-578.
7. Zaki Ahmad Khan and Abdus Samad, A Study of Machine Learning in Wireless Sensor Network.
8. P. Langley and H. A. Simon, 1995, Applications of machine learning and rule induction," Communications of the ACM, vol. 38, no. 11, p. 54–64.
9. L. Paradis and Q. Han, 2007. A survey of fault management in wireless sensor networks," Journal of Network and Systems Management, vol. 15, no. 2, p.171–190.
10. B. Krishnamachari, D. Estrin, and S. Wicker, 2002. The impact of data aggregation in wireless sensor networks," in 22nd International Conference on Distributed Computing Systems Workshops, p. 575–578.
11. J. Al-Karaki and A. Kamal, 2004. Routing techniques in wireless sensor networks: A survey, IEEE Wireless Communications, vol. 11, no. 6, p. 6–28.
12. K. Romer and F. Mattern, 2004. The design space of wireless sensor networks," IEEE Wireless Communications, vol. 11, no. 6, p. 54–61.
13. Himanshu Sharma, Ahteshamul Haque, Frede Blaabjerg, Machine Learning in Wireless Sensor Networks for Smart Cities: A Survey.
14. Alsheikh, M.A.; Lin, S.; Niyato, D.; Tan, H.-P., 2014. Machine learning in wireless sensor networks: Algorithms, strategies, and applications. IEEE Commun. Surv. Tutor., 16, p. 1996–2018. [CrossRef]

15. Du, R.; Santi, P.; Xiao, M.; Vasilakos, A.V.; Fischione, C. 2019. The sensable city: A survey on the deployment and management for smart city monitoring. IEEE Commun. Surv. Tutor., 21, p. 1533–1560. [CrossRef]
16. Morello, R.; Mukhopadhyay, S.C.; Liu, Z.; Slomovitz, D.; Samantaray, 2017, S.R. Advances on sensing technologies for smart cities and power grids: A review. IEEE Sens. J., 17, p. 7596–7610. [CrossRef]
17. Anagnostopoulos, T.; Zaslavsky, A.; Kolomvatsos, K., 2017. Challenges and opportunities of waste management in IoT-enabled smart cities: A survey. IEEE Trans. Sustain. Comput., 2, p. 275–289. [CrossRef]
18. Horng, G.-J.; Liu, M.-X.; Chen, C.-C. 2020. The smart image recognition mechanism for crop harvesting system in intelligent agriculture. IEEE Sens. J., 20, p. 2766–2781. [CrossRef]
19. Rami Ahmad, Raniyah Wazirali, Tarik Abu-Ain, Machine Learning for Wireless Sensor Networks Security: An Overview of Challenges and Issues.
20. Wazirali, R.; Ahmad, R.; Al-Amayreh, A.; Al-Madi, M.; Khalifeh, A., 2021, Secure Watermarking Schemes and Their Approachesin the IoT Technology: An Overview. Electronics, 10, 1744. [CrossRef]
21. Bouaziz, M.; Rachedi, A., 2016. A survey on mobility management protocols in Wireless Sensor Networks based on 6LoWPAN technology. Comput. Commun., 74, p. 3–15. [CrossRef]
22. Al-Kashoash, H.A.A.; Kharrufa, H.; Al-Nidawi, Y.; Kemp, A.H., 2019. Congestion control in wireless sensor and 6LoWPAN networks: Toward the Internet of Things. Wirel. Networks, 25, p. 4493–4522. [CrossRef]
23. Moridi, M.A.; Kawamura, Y.; Sharifzadeh, M.; Chanda, E.K.; Wagner, M.; Okawa, H. 2018. Performance analysis of ZigBee network topologies for underground space monitoring and communication systems. Tunn. Undergr. Sp. Technol., 71, p. 201–209. [CrossRef]
24. Ertürk, M.A.; Aydın, M.A.; Büyükakka şlar, M.T.; Evirgen, H., 2019. A Survey on LoRaWAN Architecture, Protocol and Technologies. Future. Internet, 11, 216. [CrossRef]
25. Jaewoong Kang , Jongmo Kim, Minhwan Kim, And Mye Sohn , (Member, IEEE), Machine Learning-Based Energy-Saving Framework for Environmental States-Adaptive Wireless Sensor Network
26. C. Zhu, V. C. M. Leung, L. Shu, and E. C.-H. Ngai, 2015. Green Internet of Things for smart world, IEEE Access, vol. 3, p. 2151–2162, doi: 10.1109/ACCESS.2015.2497312.
27. Y.-H. Lin, Z.-T. Chou, C.-W. Yu, and R.-H. Jan, 2015. Optimal and maximized configurable power saving protocols for corona based wireless sensor networks, IEEE Trans. Mobile Comput., vol. 14, no. 12, p. 2544–2559.
28. T. Amgoth and P. K. Jana, 2015. Energy-aware routing algorithm for wireless sensor networks,'' Comput. Electr. Eng., vol. 41, p. 357–367.
29. S. Shamshirband, J. H. Joloudari, M. Ghasemigol, H. Saadatfar, A. Mosavi, and N. Nabipour, 2019. FCS-MBFLEACH: Designing an energy-aware fault detection system for mobile wireless sensor networks, Mathematics, vol. 8, no. 1, p. 28.
30. F. Deniz, H. Bagci, I. Korpeoglu, and A. Yazıcı, 2016. An adaptive, energy- aware and distributed fault-tolerant topology-control algorithm for heterogeneous wireless sensor networks,'' Ad Hoc Netw., vol. 44, p. 104–117.

A Dynamic Approach To Mitigate Performance And Latency Issues In Serverless Cloud Computing Environment

[a*]Bhatt Jay Mihirkumar, [b]Dr. Rajan Patel

[a*]*PG Scholar, Gandhinagar Institute of Technology Khatraj-Kalol Road, Moti Bhoyan, Ta.Kalol, Dist. Gandhinagar-382721, India*
[b]*Professor,Gandhinagar Institute of Technology Khatraj-Kalol Road, Moti Bhoyan, Ta.Kalol, Dist. Gandhinagar-382721, India*

Abstract

Serverless computing is next generation cloud computing execution model in which the allocation of machine resources is on demand. "Serverless" is in the sense that servers are still used to execute code. However, serverless application developers do not have to worry about capacity planning, configuration, management, maintenance, fault tolerance, or scaling of containers, VMs, or physical servers. While the Serverless allows engineers to focus on writing core business logic, it comes with the problem of performance and latency issues. Cold start, function placement, and resource allocation are the key aspects of that. Cold starts occur when Serverless functions are invoked before they have been loaded into the system. In this research, we present a dynamic approach to reduce latency and enhance QoS by mitigating the cold start problem and placing the function with proper configuration.

Keywords: Cloud Computing, Serverless, Clod Start, Latency.

1. Introduction

Cloud computing is the on-demand delivery of IT resources over the internet. Instead of buying, owning, and maintaining physical data centers, you can access technology services such as computing power, storage, and databases on an as-needed basis [1]. It is a general term for anything that involves delivering hosted services over the internet. These services are divided into three main categories or types of cloud computing: infrastructure as a service (IaaS), platform as a service (PaaS), and software as a service (SaaS) [2].

Basically, cloud computing provides server resources on demand, which are located remotely and mostly managed by third-party organizations. The customer rents these resources and follows a pay-as-you-go model. However, in a private cloud, the resources are managed and owned by private individuals or organizations, where the hardware and software resources are on-premise.

Multi-cloud is another concept in cloud computing where organizations use both privately-owned on-premise and third-party public cloud provider services. It provides advantages of both concepts, like the security and privacy of private cloud computing and the flexibility of public cloud computing.

Talking about server resources, server resources are the hardware or software components that help achieve tasks in the computing world. Resources are entities with processes that provide output. It can be software that falls under software as a service or platform as a service, or it could be hardware resources that majorly fall under infrastructure as a service. Apart from that, it also includes network resources like API gateways, NAT gateways, elastic IP addresses, etc. Physical pieces of hardware technology to run the code that make up your server are often referred to as server resources [4].

Cloud computing has many advantages, including public cloud, where we do not have to manage and maintain the servers. It also provides rental services for resources, commonly known as the pay-as-you-go model, where individuals or organizations pay for the resources they use and the time they have been used. Moreover, it is flexible to use cloud computing services as they are provided according to the customer's needs and are configurable according to one's needs. Also, the cost to manage is less as the servers are located and managed by third-party service providers. Apart from that, it provides disaster recovery, high availability, and scalability.

2. Theory Background

2.1 Serverless Cloud Computing

Serverless cloud computing is next generation design architecture in cloud computing industry. It is the advance version of the

* Jay Mihirkumar Bhatt
E-mail address: 210120702008@git.org.in

pay as you go and less resource management concept. It is a cloud computing application development and execution model that enables developers to build and run application code without provisioning or managing servers or backend infrastructure [5]. It provides flexibility to developers and organizations to focus on business logic and code rather than investing time on designing deploying and maintaining sever, hardware, software and network resources.

The developer is unaware of the infrastructure. Instead, the developer has access to prepackaged components. The developer is allowed to host code here, though that code may be tightly coupled to the platform [6]. After developing the application business logic, the developer just has to provide it to the serverless system, which will automatically calculate the required hardware, software, and network resources. Then it will allocate those resources and deploy the application on them, assigning the network ports according to basic configuration done by the developer in script form.

1.1. Advantages

Although cloud computing simplifies the deployment of applications by removing the need for on-premises data centers, managing or configuring server resources at remote locations is still necessary. To overcome this issue, serverless architecture has become the go-to choice, providing the advantages of cloud computing and automation through a serverless framework [7].

1.1.1. Cost effective

The serverless framework is cost-effective because the business logic code is provided to the system, and the system decides how many resources to assign and for what period of time. After executing the business logic, the allocated resources are deallocated. Customers only pay for the resources required to successfully execute the code and for the time the resources were used.

1.1.2. Scalabilitys

The serverless framework is highly scalable. It can scale up when the number of requests increases by increasing resource allocation runtime, and it can scale down when the request count decreases by deallocating resources.

1.1.3. Easy to deploy

One of the main benefits of the serverless framework is its ease of management and deployment, compared with traditional private on-premises computing data centers and conventional public cloud computing. Developers only need to provide a piece of code for the business logic to the serverless framework system. Other deployment-related decisions are taken care of by the framework runtime, which includes resource requirement calculation according to the developer's configuration, resource allocation, deployment of code, etc.

1.2. Challenges And Issues

Despite the serverless framework being a next-generation technology and providing significant advantages over traditional and conventional on-premise and cloud computing technologies, it also faces several challenges [7].

1.2.1. QoS: Quality of Service

Quality of Service is the outcome of the total performance of a service, which includes cloud computing services, internet or computing networks, and specifically, the performance experienced by users. While calculating and processing the required resources to execute the business logic, the QoS can be affected.

1.2.2. Latency

Latency is the amount of time taken by data to go through one node to another on a particular network. Several factors can impact the latency in serverless framework, including cold start, resource allocation, and function placement.

1.2.3. Cost and Pricing Model

The automation of resource allocation and deployment also comes with a different pricing model structure, which is different from the pricing structure of conventional cloud computing and on-premises data centers. Sometimes it can be complicated due to the resource allocation of different business logics.

1.2.4. Resource Limits

In the serverless framework, users have to configure resource allocation strategies and limits before deploying the code. If the rate of requests suddenly increases in the future, the configured resources may not be sufficient. In such cases, the scaling process becomes slow and affects the latency directly or indirectly.

1.2.5. Cold Start

A cold start occurs when serverless functions remain idle for some time, and the next time these functions are invoked, a longer start time is required. This delay occurs due to the provision of a runtime container to execute the functions [11]. Launching a new container, setting up the runtime environments, and application-specific initialization collectively contribute to the cold start latency [9].

1.2.6. Function Placement and Resource Allocation

The geographical location of the function and the resources allocated to execute it both play a vital role in reducing latency and improving performance [8]. The latency and function placement are directly proportional to each other. The further away the function, the higher the latency. Efficient placement of the incoming workload on hosts is required to minimize the provider's capital expenses [9]. The function also requires sufficient resources to execute successfully [10].

2. Literature Survey

2.1. ENSURE: Efficient Scheduling and Autonomous Resource Management in Serverless Environments [4]

Here, the researchers present a scheduler that works at the function level and an automatic resource allocation and management algorithm designed to reduce customers' resource costs without impacting users' performance requirements. It follows the classification algorithm, which segregates incoming third-party requests at runtime. Additionally, it automatically scales capacity to prevent cold starts and schedules requests by merging the load on a number of invokers to reuse active nodes.

2.2. Cold Start in Serverless Computing: Current Trends and Mitigation Strategies [6]

This research provides collective details on the latest innovations in mitigating cold start delay. According to research, there are mainly two general approaches to dealing with cold start delay. The primary approach tries to reduce cold start delay by optimizing environments, and the second approach reduces the frequency of cold start occurrences by pinging. In optimizing environments, the general solution is by minimizing container preparation time as well as minimizing dependencies loading time. In the second approach, it works on the phenomenon of not letting functions get into the cold start form. It keeps pinging the function to keep it awake and ready to invoke at any time without high frequency.

2.3. FaaSRank: Learning to Schedule Functions in Serverless Platforms [7]

Normally, serverless function-as-a-service uses conventional scheduling algorithms for dividing function invocations, but it ignores FaaS characteristics like fast changes in resource usage and the lifecycle. This paper focuses on a function scheduler for serverless FaaS platforms based on information provided from servers and functions. It automatically adapts scheduling policies via experience using reinforcement learning and neural networks. The system is implemented in Apache OpenWhisk, which is an open-source FaaS platform. It evaluated performance against other schedulers, including OpenWhisk's basic scheduler on two 13-node OpenWhisk clusters. Real-world serverless workload traces provided by Microsoft Azure have been used. In tests, it sustained on average a lower number of invocations, 59.62% and 70.43%, as measured on two clusters.

2.4. SAND: Towards High-Performance Serverless Computing [8]

This paper focuses on providing minimum latency, enhanced resource efficiency, and high elasticity. To fulfill these properties, it works on two key functionalities. First is application-level sandboxing, and the second one is a hierarchical message bus. In application-level sandboxing, it isolates on two levels, isolation between different applications and isolation between functions of the same application. In the second approach, it creates shortcuts for functions that interact with each other. Moreover, in a hierarchical message bus, it uses a two-level hierarchy, a global message bus distributed across nodes, and a local message bus available on every node. After testing in a commonly-used image processing application, it achieves a 43% speedup compared to Apache OpenWhisk.

2.5. Defuse: A Dependency-Guided Function Scheduler to Mitigate Cold Starts on FaaS Platforms [9]

Currently available scheduling algorithms do not consider the ubiquitous dependencies between serverless functions. This research proposes the solution of cold start potential by using dependencies to mitigate cold starts. It identifies two types of dependencies between serverless functions, i.e., strong dependencies and weak ones. It uses frequent pattern mining and positive point-wise mutual information to mine such dependencies respectively from function invocation histories. By this, it constructs a function dependency graph. The connected dependent functions on the graph can be scheduled to reduce the occurrences of cold starts. The performance evaluation of its effectiveness is done by applying it to an industrial serverless dataset. The results show that it can reduce 22% of memory usage while having a 35% decrease in function cold-start rates compared with the classical methods.

2.6. A Reinforcement Learning Approach to Reduce Serverless Function Cold Start Frequency [10]

Most of the current research focuses on addressing the issue of cold start in serverless computing by reducing the start-up or cook time of function images or minimizing the frequency of cold starts. Ongoing research has revealed that various factors such as the runtime environment, CPU and memory configurations, networking requirements, and invocation concurrency, affect the cold start of a function. This research proposes a Reinforcement Learning approach to analyze these factors, particularly the function's CPU utilization and invocation patterns, and reduce cold starts by pre-cooking the function thread in advance. The proposed approach communicates with the Kubeless serverless platform and is evaluated using the Apache JMeter to mirror the workload. The agent is compared against the default auto-scale feature of Kubeless and shows the ability to learn the invocation pattern, make informed decisions, and prepare the optimal number of function instances over time.

3. Problem Statement - Aims And Objectives

Based on the literature review, we can conclude that cold start is one of the most significant challenges that needs to be addressed in serverless computing. Cold start directly affects the request-response time, also known as the latency of the business logic, which is a key criterion for high performance. Cold start, function placement, and resource allocation and management are the major factors that contribute to increased latency and low performance.

The primary goal of this dissertation is to improve performance by reducing latency, which can be achieved by mitigating the cold start problem in serverless computing environments. There are several steps and methodologies that can play an important role in reducing latency by mitigating the cold start problem:

- Preloading the required dependencies and libraries needed to execute the code
- Provisioning the function separately, keeping it ready to be invoked rapidly
- Scheduling an algorithm to pre-cook the function business logic according to historical data of function invocations

4. Methodology - Proposed Work

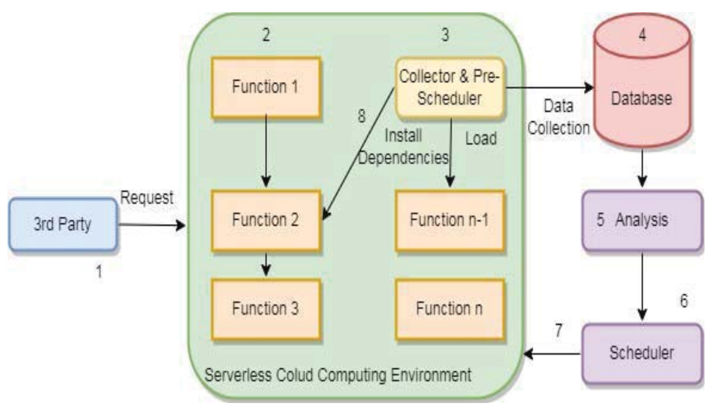


Fig. 1. Proposed Work.

Several approaches have been attempted to mitigate the cold start problem, which can reduce latency and improve performance. Keeping the function ready to be invoked at any time, or pre-cooking it before potential invocation using historical data, were the major aspects of these approaches. We propose a dynamic and hybrid approach that handles it in a novel way. It is a combination of pre-cooking the function and getting it ready to be invoked by pre-loading libraries and dependencies based on strong dependencies and historical data. Strongly dependent functions will be ready once previous functions in the chain are invoked, while weakly dependent and edge functions will be mainly dependent on the historical data of invocations.

The sequence of execution of different components to achieve this would be as follows:

- Initially, a 3rd party request will come in
- The edge function will be invoked and it will invoke the subsequent functions
- Simultaneously, we will collect trace data of different function invocations with their metadata, including start and end timestamps, resource allocation and requirements, etc.
- We will store the data in a database
- We will analyze the previous function invocations' data of sequence and time of function invocations
- We will analyze the strongly, weakly, and edge functions and then schedule the function before the next invocation.
- We will pre-cook the function by installing dependencies and libraries before the actual request triggers the function.

Existing research on Cold Start mainly focuses on mitigating the issue by addressing function dependencies, scheduling algorithms, and fusion functions. Some of the approaches include using strong and weak dependencies with a hybrid histogram scheduling and tree policies, edge-triggered and strongly dependent methods. Reinforcement learning-based approaches have also been used, with time series and Q-Learning policies or Hashing, Greedy, Static rank, and Mix-Norm schedulers being applied. By combining both of these approaches, we can effectively mitigate the latency issues caused by Cold Start.

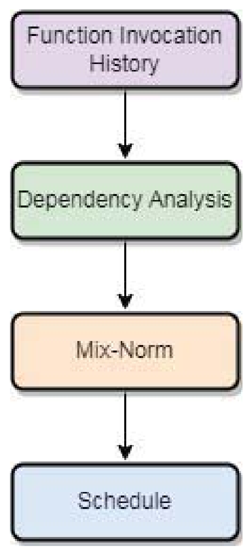


Fig. 2. Analysis And Scheduling.

5. Conclusion And Future Work

After reviewing the next generation technology of Serverless with its advantages and disadvantages, as well as the current issues and challenges in the field, we can conclude that by predicting the next function invocation using historical data and custom configurations, we can pre-cook the function by installing dependencies and loading it into memory, which can be helpful in mitigating the cold start problem in Serverless cloud computing environments. In the future, we plan to further reduce the operational, processing, and architectural overheads, as well as the cold start delay.

References

- Journal articles:

1. Ioana Baldini, Paul Castro, Kerry Chang, Perry Cheng, Stephen Fink, Vatche Ishakian, Nick Mitchell, Vinod Muthusamy, Rodric Rabbah, Aleksander Slominski, Philippe Suter, 2017. Serverless Computing: Current Trends and Open Problems.
2. Hassan B. Hassan, Saman A. Barakat, Qusay I. Sarhan, 2021. Survey on serverless computing.
3. Jinfeng Wen, Zhenpeng Chen, Yi Liu, Yiling Lou, Yun Ma, Gang Huang, Xin Jin, Xuanzhe Liu, 2021. An Empirical Study on Challenges of Application Development in Serverless Computing.
4. Amoghavarsha Suresh, Gagan Somashekar, Anandh Varadarajan, Veerendra Ramesh Kakarla, Hima Upadhyay, Anshul Gandhi, 2020. ENSURE: Efficient Scheduling and Autonomous Resource Management in Serverless Environments.
5. Aakanksha Saha, Sonika Jindal, 2018. EMARS: Efficient Management and Allocation of Resources in Serverless.

6. Parichehr Vahidinia, Bahar Farahani, Fereidoon Shams Aliee, 2020. Cold Start in Serverless Computing: Current Trends and Mitigation Strategies.
7. Hanfei Yu, Athirai A. Irissappane, Hao Wang, Wes J. Lloyd, 2021. FaaSRank: Learning to Schedule Functions in Serverless Platforms.
8. Istemi Ekin Akkus, Ruichuan Chen, Ivica Rimac, Manuel Stein, Klaus Satzke, Andre Beck, Paarijaat Aditya, and Volker Hilt, Nokia Bell Labs, 2018. SAND: Towards High-Performance Serverless Computing.
9. Jiacheng Shen, Tianyi Yang, Yuxin Su, Yangfan Zhou, Michael R. Lyu, 2021. Defuse: A Dependency-Guided Function Scheduler to Mitigate Cold Starts on FaaS Platforms.
10. Siddharth Agarwal, Maria A. Rodriguez, Rajkumar Buyya, 2021. A Reinforcement Learning Approach to Reduce Serverless Function Cold Start Frequency.
11. Akash Puliyadi Jegannathan, Rounak Saha, Sourav Kanti Addya, 2022. A Time Series Forecasting Approach to Minimize Cold Start Time in Cloud-Serverless Platform.

Acknowledgement

I owe my deep gratitude to my research guide Dr. Rajan Patel who guided me all along, till the completion of my thesis work by providing all the necessary information, guidance, supervision, moral support, and discussion about my thesis progress with devotion of their valuable time. She gave me a great deal of freedom to choose a research topic, focusing more on my research interests. Without her experience and insights, it would have been exceedingly difficult to do qualitative work.

I would also like to extend my sincere gratitude to Dr. H N Shah, Director of Gandhinagar Institute of Technology, Gandhinagar, for his very humble support, motivation, continuous encouragement for innovation with new ideas, and for providing wonderful positive environment for research work.

I am thankful to and fortunate enough to get constant encouragement, support, guidance and valuable comments from all Faculty members of Computer Engineering Department which helped me in successfully completing work. I heartily thank to all Library Staff Members for providing me great resources of library as well as excellent peaceful environment and Lab Assistants for allowed me to do my research work in labs of Computer Department. At last, I hereby take this opportunity to thank all my colleagues, family and friends and those who directly or indirectly helped me in preparing this thesis work.

Use of Communication Ranging for Optimizing the Localization Accuracy of Mobile Sensor Networks

Raxit G. Jani[a], Vishesh Joshi[b*], Priyank Bhanushali[c], Manav Nai[d]

[a]*Assistant Professor, Computer Engineering Department, Gandhinagar Institute of Technology, Gandhinagar 382721,Gujarat, India*
[b*,c,d]*Student, Gandhinagar Institute of Technology, Gandhinagar 382721, Gujarat, India*

Abstract

Localization is among the most important tasks performed by sensor networks. Wireless sensor network (WSN) functions like coverage and event detection are affected by the localization method that is utilized. Accuracy is considered to be among the factors that has been connected to good network efficiency in the most of applications utilizing sensor networks. Communication ranging is one easily implementable technique for positioning on short transmission nodes without any additional elements. The work suggests a highly accurate improved communication ranging localization technique. Twenty Wasp mote Nodes and a Meshlium router were used to test and validate the suggested technique in a forest environment. Uniqueness of suggested localization method is centered upon adjusting the XBee transmitter's capacity to change power over four rounds in order to deliver the best location prediction. The findings showed that nodes' locations may be determined with an error range of 0 to 22 metres. Starting at 30 in round 1, the network location inaccuracy dropped to 8 metres in round 4. This method is utilized with a variety of networks and technologies as far as the system capacity could be adjusted to distinct values and the transmission range is determined or can be manually evaluated.

1. Introduction

Wireless sensing technologies have seen exceptional progress in recent years, and a different type of wireless sensor network system has arisen[1]. Moreover, people's desire for associated technologies is rising daily. This technology is widely used in the army, commerce, agriculture, raising livestock, home automation, smart grid, and numerous other sectors because of its minimal price, small footprint, with straightforward distribution of IOT devices[2]. Although each of these applications is diverse and has its own unique features, they all require the sensor nodes' locations in order to function [3]. All of the system's information becomes meaningless if the sensor node's location information is absent. The method for obtaining node position information is so crucial. For a very long period, GPS positioning devices have been used to acquire sensor location data [4]. We know that using Global positioning technology is expensive. Our specifications for low-cost sensor nodes are not met by this. How to control the expense of finding sensor nodes is currently a prominent topic [5].

Whenever wireless sensor networks are implemented in enclosed spaces and urban areas, the data transfer among sensors may frequently get hindered via objects, preventing the device from communicating in alignment. Such a problem has two direct effects: it significantly alters the data they obtain from the locating module and reduces the locating performance of the system[6]. In order to resolve these problems, this paper proposes a nonline-of-sight identification technique. A search algorithm finds the sensor which can interact with the system via connection and eliminates the node throughout positional play, significantly altering some of the information with the nonline-of-sight feature and removing the node during positioning, removing the impact of the unaligned line of sight issue on structure placement efficiency.

Throughout this paper, biochemical gas sources are located using wireless sensor networks to aid individuals in responding to disasters more quickly. We propose a robust expectation - maximization location technique and compare it with the direct trilateral approach as well as the nonlinear least squares approach. In order to address imperfections of both the centralised location algorithm and preserve network bandwidth, the multilayered sensor network has been implemented on biochemical resource placing. A distributed positioning algorithm based on improved particle filtering within the hierarchical sensor network is then proposed, which manages nodes through sub regions. State changes, and the convergent biochemical gas source position is estimated iteratively in a loop. The creation of a brand-new sensor node location information gathering technology is a top priority for many R&D organisations. Currently, it is common practice to rely only on a wireless sensor network for locating. This technology has the unique property that, in order to complete positioning, It merely requires using information from interaction between nodes or an exclusive sequence of instructions.. In order to examine the present issues with the technology and provide solutions, this article introduces several positioning algorithms, sensor network positioning technologies, and associated topics.

2. Related Work

The localization can be divided into localizations based on known locations, proximity, angles, ranges, and distances. Target localization and node self-localization were the two categories into which Cheng et al. (2012) divided the known localization

*Vishesh Joshi
E-mail address: visheshjoshi78@gmail.com

techniques [10] for their review. Their localization method taxonomy is displayed in the following figure (see Fig. 1).

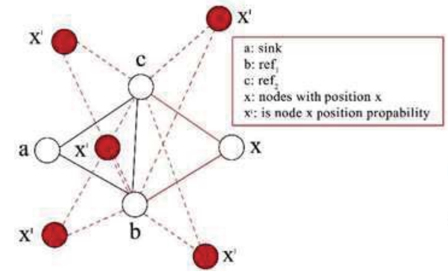


Fig. 1. Angle calculation.

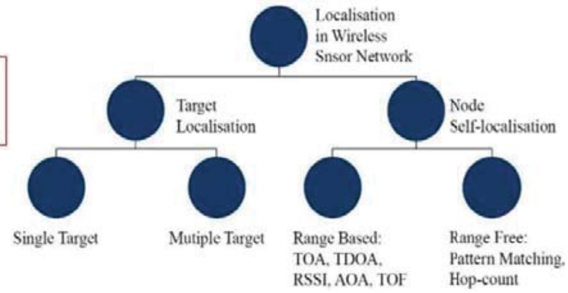


Fig. 2. Localization method taxonomy.

The far more dependable and promising navigation innovations include thought to be GPS or standalone cellular systems [11][12]. Although GPS is significant because to its widespread accessibility, it is impracticable to place it inside every sensor due to its high cost and energy consumption. As a consequence, in order to conserve power and resources, just a tiny subset of nodes, referred to as the anchor as well as beacon nodes, carry GPS units. With one of the several localization techniques, the remaining nodes learn their positions. Determination of reception phase of the antennas, the AoA technique compares the beam shapes and signal intensity. To perform this comparison, this tactic needs both directional and omni-directional antennas. A precise synchronisation is necessary for either the round or one-way propagation period to be provided by distance-related measurements. The signals from various receivers are measured, and the measurement is then used in some calculations, like cross-correlation, to determine the synchronized approximation of the transmitter provided by the TDoA. Without taking angle into account, the RSSI determines its output intensity in relation to the transmitter and receiver's range. Essentially, there are an endless amount of disturbances intermingled with the channel power, but RSSI measures energy on the channel over time.

A biassed location estimation vector had to be created for Yu to introduce an AoA-ToA transmission hypothesis in 2007[11]. Khan et al. [26] far nearly presented an enhanced and objective rendition in which they carried out a simulation to test whether the performance of the linear least squares estimator might be reliably predicted using the theoretical mean square error.An approximation pertaining towards the challenge of quasi target localization in a 3-D space serves as the foundation for the LLS and its optimisation, a different method Yu proposed in 2007 [27]. Khan et al. [28] presented a similar strategy two years later, with the primary exception being that the masses included the path-loss coefficient was utilized in their approach, and treated as an undetermined figure. Jiang et al. [14] established a slant-based localization model where a number of RSSI lights were used to establish the inclinations where technique utilizing biometrics as a foundation cease to deliver adequately when the network's circumstance altered. While Tomic et al. additionally established AoA.[29], the localization approach employed the RSSI. Lacking a description of the dissemination state where the Predicting AoA was regarded crucial, the nodes in these span approaches were located by the AoA. If distances are known, A methodology to evaluate the position inclinations was put out by Alkhatib et al. [30]. This method is based on the polar system's (R &) ability to localize nodes (see Fig. 2). The node spacing are contrasted (R) from unidentified and recognised nodes, the angles were estimated.

The localization approach was divided into centralised and dispersed techniques by Paul and Sato [31]. According to them, For WSN apps, the dispersed strategy was more commonly employed. We also noted that despite free ranging methedologies were a greater expense in terms of node hardware, ranging methedology were thought to be quite precise. A mathematical formulation that is variable for WSN localization by Liu and Liu [32] was put forth using measurements of the separations between M nodes in M-dimensional space. To determine the radial error, nodes with comparative weights were stored in a matrix.

3. Analysis of Available Techniques

Precise localization of sensors is among the very important criteria inside this plurality of data networks. Unfortunately, every possible approach to achieve this requires extra technology, like GPS units, sonar or thermal sensors, beams, or indeed directed radar. Each of these options raises the node's material requirements, requiring more power and spending more money on faster memory and CPUs. However, additional tests and experiments revealed that RSSI was still not dependable within circumstances where connections encountered problems like high absorbance or uncertain coefficients of path-loss. Numerous investigations have discovered that RSSI is indeed a node positioning indicator. According to a study by Cama-Pinto [33], RSSI can only be used as an indicator for distances under 40 metres; for distances between 40 and 100 metres, readings can vary between 60 and 65 dBm, which makes it difficult to use as an indicator. In actuality, these values weren't constants because almost everything had an impact on the outcomes, including the environment outside and inside, the weather (dry or humid), oShifting elements and obstructions. As a result, RSSI analysis was deemed unreliable or incorrect, or they may produce false findings in few specific softwares. A ToF and ToA are very popular approaches for localization in WSNs and various types of telecommunication systems. The following factors, however, limit the technique's applicability to theoretical studies of short-range communication:

$$Speed = \frac{Distance}{Time}$$

to calculate the required time for 1 meter

$$T = 1\frac{meter}{3} * 10^8 = 3.33\, nsec$$

One meter's difference was measured in 3.33 nsec, and ten meters' difference in 33.3 nsec. In order to reliably predict the ToF or ToA by 1meter and 30 MHz over a span of ten meters accuracy, It requires a CPU with a speed of at least 300 MHz (1/3.33 nsec). According to sensor network applications, nodes are spread out between each other at a range of 1 to 100 metres in the majority of cases where the sensors' detecting range only extends to a few metres or up to 700 nodes when utilizing an Xbee in the form specified in the Guide for Waspmote [34]. Arduino, Waspmote, or sensor nodes found on store shelves in general are not able to deliver such accuracy. In Figure 3,

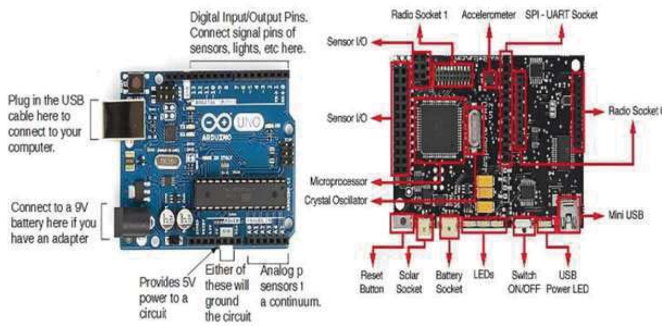


Fig. 3. Arduino and Waspmote.

Waspmote and Arduino both make use of ATMEGA processors. The ATMEGA 1281 featuring a 14.74MHz bandwidth is utilized by Waspmote, which translates to a processor's ability can generate 14,740,000 pulses each second to a clock. Datasheet for something like the ATMEGA1281 [35] indicates that the CPU comprises numerous timers that are utilized for various functions (see Fig. 4). The same source input, known as "AVR Clock," the first instance of this was delivered by a frequency-controlled crystal oscillator of 14.74 MHz, is used by all timers. These timers are used to time when the transmitter and receiver turn on and off. besides other timed activities. These timers have the ability to measure TOF, TOA, etc.

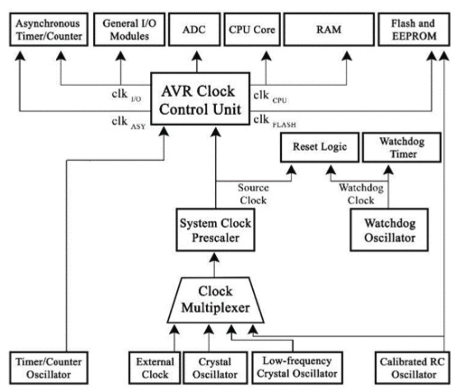


Fig. 4. Clock Distribution. [31]

The duration of a clock pulse is instr.484260515 nsec (1/f). Waspmote's minimum measurement duration is consequently approximately 67.5 nsec. Thus, a = 20.27 m when a Means (67.5 nsec denotes the duration of a single pulse) / (3.33 nsec TOF for 1 meter).Therefore, a clock = 20.27 meters. This technique was tested using Waspmote, but it was discovered that the component became unreliable for measuring pulses of the clock.A total of six devices were present in ATMEGA, but not one of them were capable of accurately counting a clock's phases [32] because the actual clock might drift with each trial and because they couldn't distinguish between processing-related pulses and unimportant pulses while counting pulses at such a fast speed. Without utilizing one of the prescaling choices of 64 and above, these timers provide unreliable counting (see Table 1). To put it another way, every count increases by sixty-four instr (or, more specifically, each counter value is increased by sixty-four instr) [36]. As a result, the TOF needed for 1297.3 meters is 64*67.5 nsec, or 4320 nsec. .Xbee, Zigbee, and other short-term protocols for communication for detector equipment are in fact exempt from this. However, it is limited to being used to Satellite or satellite data, which require millisecond-level timing precision and whose signals travel hundreds of thousands of miles. Waspmote can make precise calculations for these signals.The ATMEGA328p microcontroller used by Arduino operates at a frequency of around 20 MHZ, making one clock cycle equal to 50 nanoseconds or a 15-meter TOF. This problem affects Waspmote and Arduino as well, making it challenging to tally pulses of clock at this rate.

4. Methodical Study

Communication-based localization is regarded as a widespread technique. Further research revealed that it was the sole method that could be used on sensor nodes lacking any other parts, like GPS or ultrasonic sensors.We demonstrated an innovative approach using different Waspmote electrical transmission settings to identify the transmission frequencies relative to source networks. Several rounds of this procedure were used to attain improved accuracy. Twenty Waspmote nodes and a way of collecting networks were used in an experimental measurement to determine the proposed method's accuracy. (a Spanish product from Libelium) [33]. (see Fig. 5).

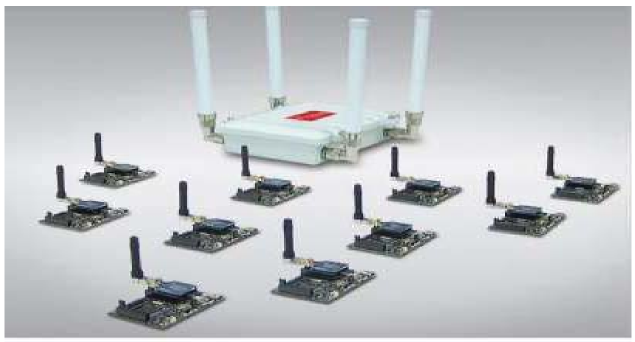


Fig 5. Waspmote and Meshlium. [37]

Waspmote nodes with an XBeePRO-S2 emitter that supports 802.15.4. According to Libleium [37], its transmission range can reach 700 metres.The first among five data can be chosen for the XBee power transmission. We manually measured the communication distance for the values (0, 1, 2, and 3) and presented the results (see Table 2).The experiment used a total of nine nodes as references, with data on their positions (X, Y) kept in memory. The experiment was conducted in a woodland environment (see Fig. 6). (reference nodes). All other 11 nodes were deployed at random and used specially designed frames in conjunction with the Waspmote IDE for locating one another through contacting each other nearby.

Fig. 6. The localization measurement in the forest.

During initial configuration, a variety of referencing sites must be manually loaded. These remaining nodes are randomly distributed (see Fig. 7).

The following is a summary of the suggested technique rounds:

Phase 1

- Because the power levels for every node were all configured with 0 = 10 dBm ways of collecting[xbee802.setPowerLevel(0)], the initial range for transmission and reception was 100 m.
- Reference nodes begin broadcasting messages with the node ID and their coordinates as (X, Y). 5Every single unfamiliar network attempts to get as many notifications as it's able to receive from in-range nodes that serve as references in order to determine every possible position within the cross-section of every one of the in-range standard networks.

- Two reference nodes must be within range at the very least (see Fig. 9).Alternatively, hold off until more accessible points of reference are available shortly.
- Location of Phase 1 is conceivable [ ][ ] = $[x1\ y1\ \ldots\ xn1\ yn1]$
- The Phase1 results include all potential sites and the average of all potential nodes as recommended locations for this phase.

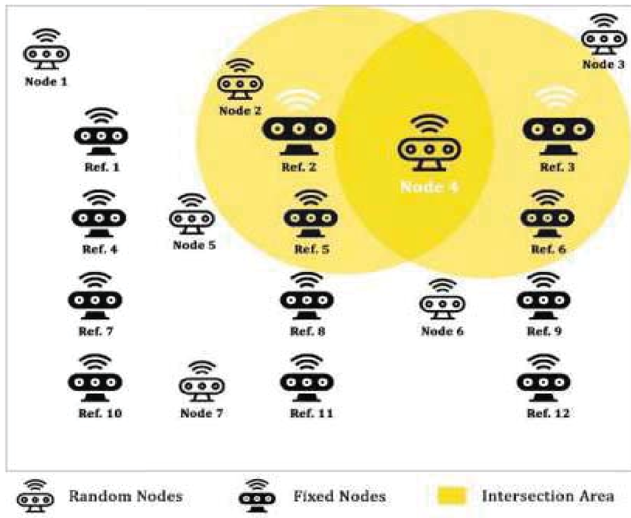


Fig. 7. Potential sites inside the chosen.

Phase 2

- During this round, a position would be reversed, with an ambiguous site setting their transmitting power to 1 before broadcasting to nearby reference nodes and waiting for their responses.
- Power transmission was set to 1 by the ambiguous sites (12 dBm). Up to 130 meters is the broadcast and receiving range thanks to [xbee802.setPowerLevel(1)].
- Unidentified nodes begin broadcasting communications that include their IDs and Mac addresses. utilizing respective (X, Y), as well as Identity data, all reference sites that have received a broadcast signal will respond to the unknown sites.
- Ambiguous sites will get these signals as well as only utilize these new entry sites at one radius somewhere between 100 to 130 meters [(accepted references) - (Phase 1 connections)].

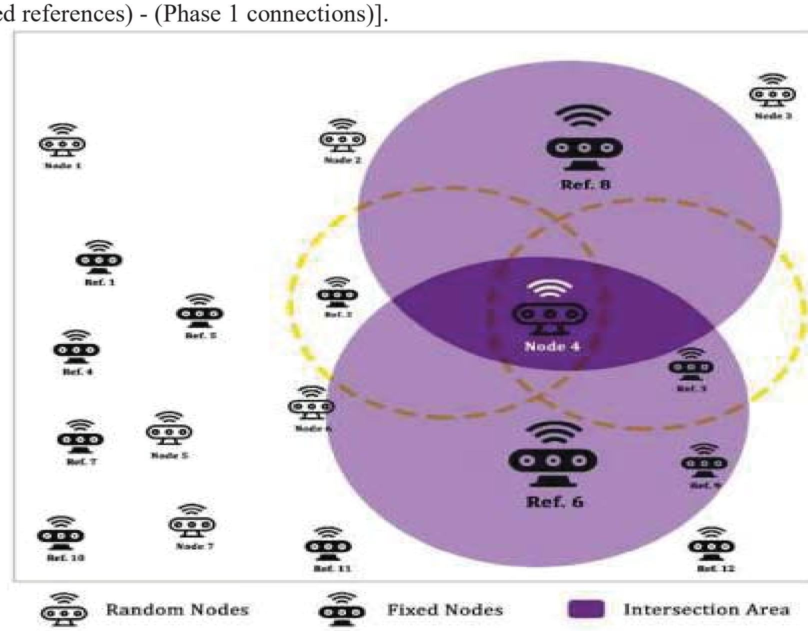


Fig. 8.The Suggested Destinations in the Phase 1 and Phase 2 interconnection of the Chosen References Node Range.

- This quantity of potential Destinations determined by Phase 1 will be reduced using the traditional distance-between-two-points algorithm, and the distance between each potential site determined by Phase 1 as well as all sources determined by Phase 2 shall be computed.

- It will be stored inside a potential destination of Phase 2[][] = [x2 y2... xn2 yn2] if (R1 = 100) Square root [(xround1 x reference phase 2)2 + (y phase 1 - y reference phase 2)2] = (R2 = 130). If not, it will be deleted. (Observe Figure 10).
- This round's locations will be used to determine the mean of the possible sites.

Phase 3

- Up to 180 meters of transmission and reception are possible with the energy transfer set by ambiguous site to 2 = 14 dBm [xbee802.setPowerLevel(2)].
- Unidentified nodes begin broadcasting messages with their ID and Mac address.
- Referencing sites will respond to unidentified sites with their (X, Y), and Identity information if they receive the broadcast signals.
- At a distance of between 130 and 180 meters, Ambiguous sites are only going to utilize new entry sites and get the information. As shown in the image, [(received reference) - (Round 2 references) - (Phase 1 references)].
- Each possible location from Phase 2 and each reference from Phase 3 will be measured in relation to one another. This will reduce the amount of Phase 2 origins that are feasible.
- The location will be saved in a probable location of Phase 3 [][] = [x3 y3... xn3 yn3] if (R2 = 130) Sqrt [(x phase2 x reference phase 3)2 + (y phase 2 - y reference phase 3)2] = (R3 = 180). If not, it will be deleted. (Refer to Figure 11).
- This round's destinations will be used to determine the mean of the possible sites.

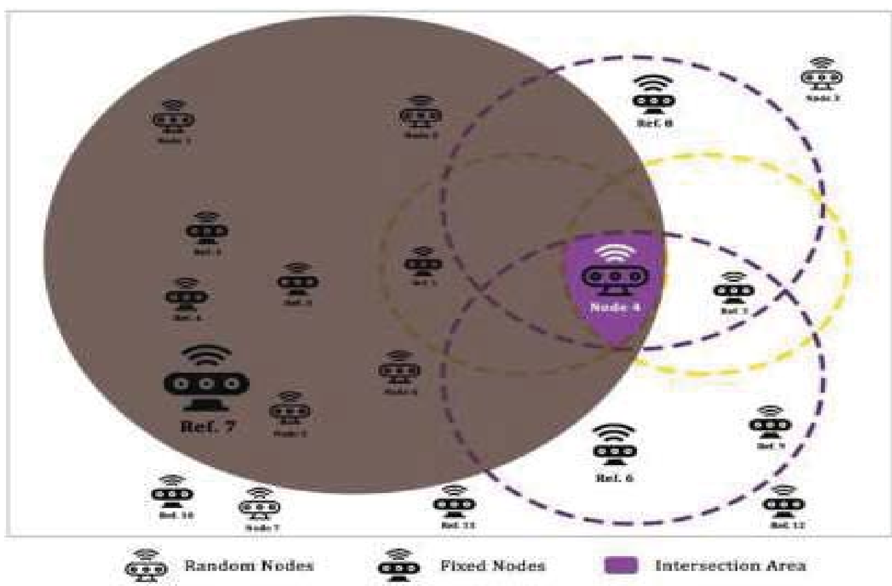


Fig. 9. The Suggested Destinations in the Phase 1, Phase 2 and Phase 3 interconnection of the Chosen References Node Range.

Phase 4

- Ambiguous sites set the energy transmission to 3 = 16 dBm [xbee802.setPowerLevel(3)], which allows for up to 260 meters of transmission and reception.
- Ambiguous sites begin transmitting broadcast messages containing their Mac address and Identity.
- As soon as you get broadcast signals, any Reference node will respond to any unknown nodes by providing their (X, Y), as well asIdentitydata.
- At a distance of between 180 and 260 meters, Ambiguous sites wouldgetthesesignals and solely employ the fresh reference sites. Phase 3 references, Phase 2 references, and (received reference) (Phase 1 references)
- Each prospective Phase 3 site's distance from every reference coming from phase 4 will be calculated. The distance formula would be applied to cut down on the number of Phase 3's potential places.
- If (R 3 = 180) Sqrt [(x phase 3 x reference phase 4) 2 + (y phase 3 - y reference round 4) 2] = (R 4 = 260),then a prospective destination will save the address of Phase 4 [] [] = [x 4 y 4... xn 4 yn 4]. It will be erased if not.

- The node is now prepared to shift its role and assist other unallocated nodes in localizing themselves by being suggested as the average of potential nodes' estimated locations. The method is summarized in the flowchart below. (See Figure 12).

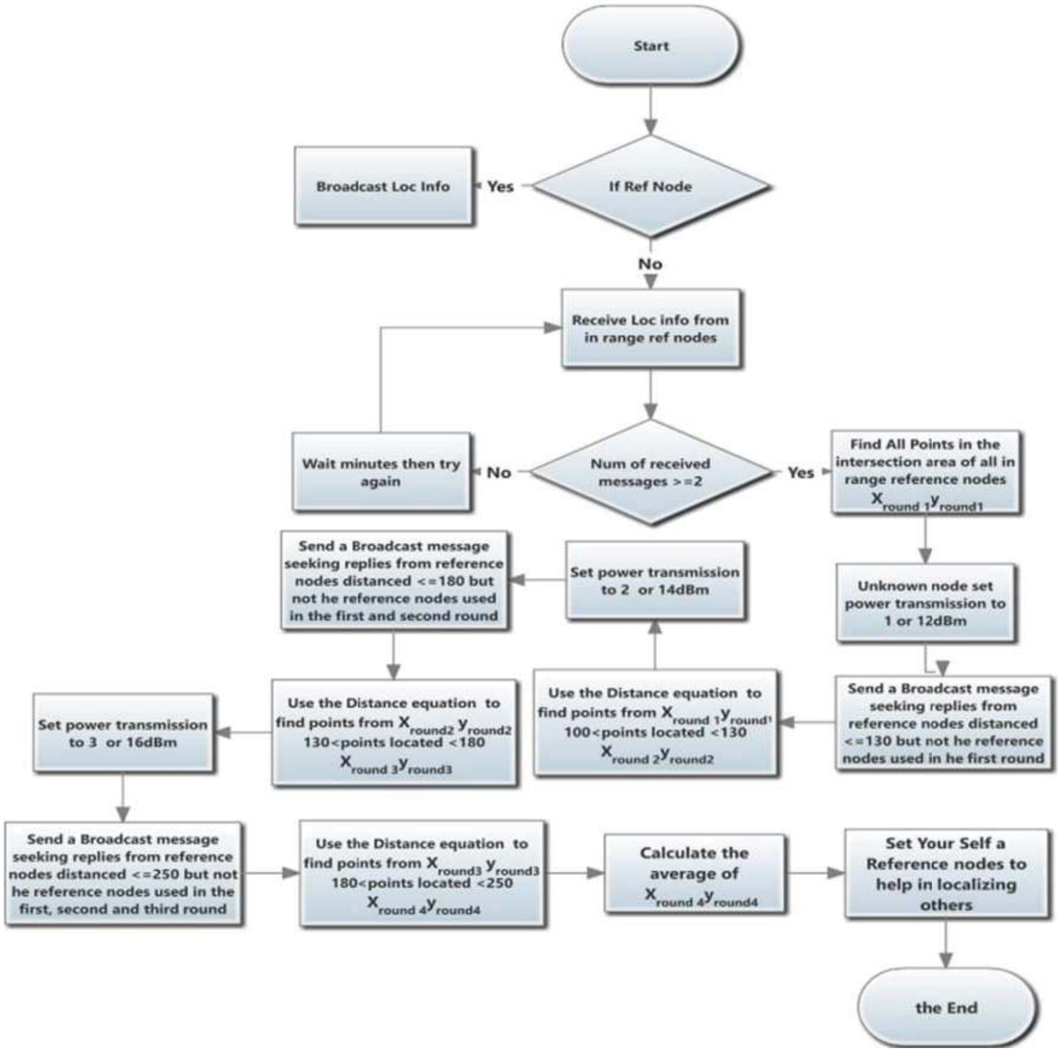


Fig. 10. Flow Chart of the technique.

1. **Start**
2. **Ref Nodes:** Set power to first level.
3. **Ref Node:** Send broadcast message.
4. **Receiving Xnode:** Calculate intersection points
5. **While** power <= max level
   - **Receiving Xnode:** Set power to the next level.
   - **Receiving Xnode:** Send a broadcast message.
   - **Receiving Ref Nodes:** Set power to the next level.
   - **Receiving Ref Nodes:** Reply with a unicast.
   - **Receiving Xnode:** Find the new intersection points.
6. **End While**
7. **End.**

Fig. 11. Pseudo Code.

5. Results and Discussion

We've conducted tests in a woodland setting. The communication range was manually measured within a 400 × 400 m region where the nodes were dispersed in order to determine whether XBee power transfer can be configured to 0, 1, 2, or 3. The positioning of the 20 sites. (Observe Fig. 14).

Four sets of experiments were conducted to assess how the suggested strategy affected localization accuracy. Each round's outcomes are described (see Table 3).

Table 3 Real nodes and their measured locations after four iterations.

Node Type	Node ID	Actual X	Actual Y	Round 1			Round 2			Round 3			Round 4		
				X	Y	# of Possible Locations	X	Y	# of Possible Locations	X	Y	# of Possible Locations	X	Y	# of Possible Locations
REF	1	100	100	-	-	-	-	-	-	-	-	-	-	-	-
REF	2	200	100	-	-	-	-	-	-	-	-	-	-	-	-
REF	3	300	100	-	-	-	-	-	-	-	-	-	-	-	-
REF	4	100	200	-	-	-	-	-	-	-	-	-	-	-	-
REF	5	200	200	-	-	-	-	-	-	-	-	-	-	-	-
REF	6	300	200	-	-	-	-	-	-	-	-	-	-	-	-
REF	7	100	300	-	-	-	-	-	-	-	-	-	-	-	-
REF	8	200	300	-	-	-	-	-	-	-	-	-	-	-	-
REF	9	300	300	-	-	-	-	-	-	-	-	-	-	-	-
Random	10	130	140	150	150	37	*	*	*	153	153	20	130	140	1
Random	11	190	150	200	150	127	200	150	9	195	155	2	190	150	1
Random	12	260	140	250	150	37	*	*	*	246	153	18	246	153	18
Random	13	120	270	142	257	48	120	280	10	113	266	3	113	286	3
Random	14	270	270	250	250	37	*	*	*	247	247	23	270	270	1
Random	15	334	147	300	150	127	*	*	*	346	152	21	340	150	15
Random	16	315	239	300	250	127	298	252	30	316	233	3	*	*	*
Random	17	145	62	150	100	127	*	*	*	152	53	21	150	60	15
Random	18	194	74	150	100	127	182	91	7	186	88	5	186	88	5
Random	19	229	347	250	300	127	*	*	*	247	346	21	250	340	15
Random	20	267	238	252	257	48	280	280	10	286	266	3	286	266	3

According to the results (see Table 4), the node results of potential sites were decreased in each cycle, which also decreased the likelihood that the expected location would be incorrect.

We observed that errors in rounds 1 and 2 ranged from 10 to 51 metres, round 3 from 7 to 26 metres, and round 4 from 0 to 22 metres. It's important to note that not all nodes locate themselves precisely and without making any mistakes. (View Figures 15 and 16). See table 5 to compare our findings to those of other approaches.

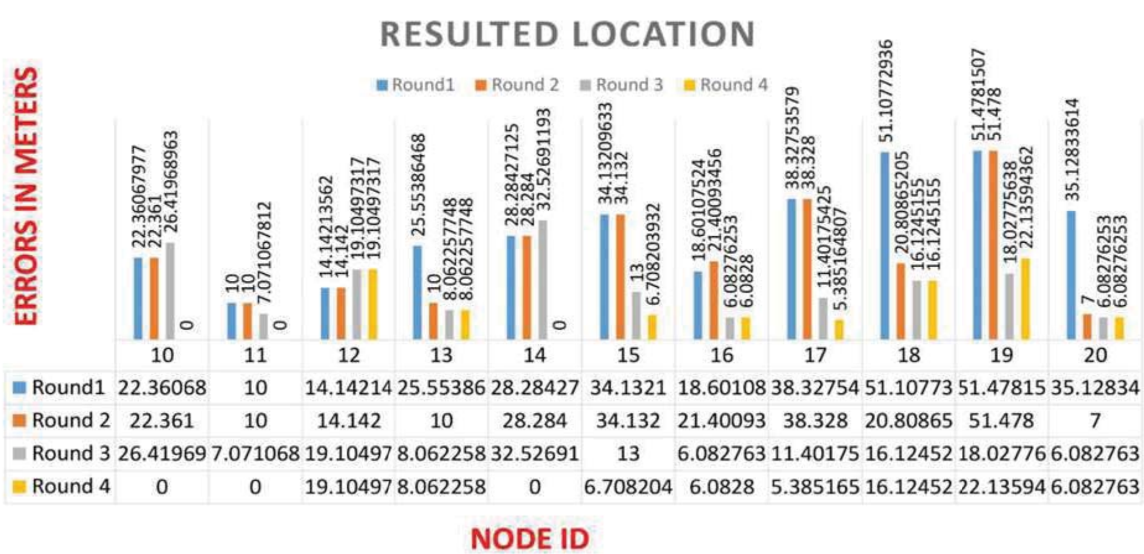


Fig. 12. The typical meters-per-round inaccuracy.

Fig. 13. The increase in the typical meter inaccuracy for each round.

6. Conclusions

Accurate localisation is one of the most crucial elements in the majority of WSN applications. Trilateration, Triangulation, AOA ,TDOA, TOF, GPS, TOA and TOF are common localization methods. The following issues with these methods could arise: 1. Both the cost and node life duration of GPS are excessively high. 2. RSSI is used by the majority of ranging techniques to determine distances. In actual situations, the signal strength does not always decrease as the range between transmitter and receiver increases, though. These RSSI readings can change irregularly over time, even when there is a consistent distance between the devices. 3. The TOA and TOF approaches cannot be used because WSN nodes powered by the only exact timing that Waspmote and Arduino CPUs can provide is millisecond resolution. Only one method that works in short-range communication networks without additional node localization components is the communication ranging technique. The range technique was employed several times in the suggested method to find nodes. By leveraging products that were already on the market, this method improved accuracy to the point where it occasionally neared 100%. Additionally, with each round, the number of potential locations shrank. The suggested method can be applied to external uses. Let’s select a woodland region as the location for the trail, which has a variety of challenges. The method has the benefits of price cutting as well asincreasing network lifespan even in environments with challenging radio propagation and in inhospitable locations. Moreover, it functions with lots of nodes.

References

1. Khan, R., Pathan, A., 2018. The state-of-the-art wireless body area sensor networks: A survey, International Journal of Distributed Sensor Networks, 14, (4).
2. Darwish, A., Hassanien. A., 2011. Wearable and implantable wireless sensor network solutions for healthcare monitoring, Sensors, 11, (6), p. 5561-95.
3. Idwan, S., Zubairi J.A. and Mahmood, I., 2016. Smart Solutions for Smart Cities: Using Wireless Sensor Network for Smart Dumpster Management’, International Conference on Collaboration Technologies and Systems (CTS), Orlando, FL, p. 493-497.
4. Guerrero-Ibáñez, J., Zeadally, S., Contreras-Castillo, J., 2018. Sensor Technologies for Intelligent Transportation Systems, Sensors, 18, (4), p.1212.
5. Hilmani, A., Maizate, A., Hassouni, L., 2018. Designing and Managing a Smart Parking System Using Wireless Sensor Networks, J. Sens. Actuator Netw.
6. Alkhatib. A., Hnaif. A. and Kanan. T., 2019. Proposed simple system for Road Traffic Counting, International Journal of Sensors, Wireless Communications and Control, 9, (2), p. 269-277.
7. Ez-Zaidi, A., Rakrak, S., 2016. A Comparative Study of Target Tracking Approaches in Wireless Sensor Networks, Journal of Sensors.
8. Alkhatib, A. A., 2014. A Review on Forest Fire Detection Techniques’, International Journal of Distributed Sensor Networks.
9. Alrajeh, N. A., Bashir, M., Shams, B., 2013. Localization Techniques in Wireless Sensor Networks, International Journal of Distributed Sensor Networks.
10. Cheng, L., Wu, C., Zhang, Y., Wu, H., Li, M., Maple, C., 2012. A Survey of Localization in Wireless Sensor Network, International Journal of Distributed Sensor Networks.
11. Vossiek, M., Wiebking, L., Gulden, P., Wieghardt, J., Hoffmann, C., Heide, P., 2012. Wireless local positioning, IEEE Microwave Mag., 4, p.77–86.
12. Jurdak, R., Corke, P., Cotillon, A., Dharman, D., Crossman, C., Salagnac, C., 2013. Energy-efficient localization: GPS duty cycling with radio ranging’, ACM Trans Sens Netw, 9, (2).

13. Nguyen, L. N., Vy, T., D., Shin, Y., 2019. An Efficient Hybrid RSS-AoA Localization for 3D Wireless Sensor Networks, Sensors, 19, (9), p.2121.
14. Jiang, J., Lin, C., Lin, F., Huang, F. ALRD, 2012. AoA localization with RSSI differences of directional antennas for wireless sensor networks', International conference on information society (i-Society), p. 304–309.
15. Luo, J., Zhang, X., Wang, Z., 2013. A new passive source localization method using AOA-GROA-TDOA in wireless sensor array networks and its Crame´r–Rao bound analysis, IEEE international conference on acoustics, speech and signal processing (ICASSP), p. 4031–4035.
16. Pita, R., Utrilla, R., Rodriguez-Zurrunero, R. and Araujo, A., 2019. Experimental Evaluation of an RSSI-Based Localization Algorithm on IoT End-Devices, Sensors, 19, (18), p. 3931.
17. Jun L., 2014. Range error correction in RSSI-based wireless sensor node localization', IEEE international conference on mechatronics and automation (ICMA), p. 379–383.
18. Sahu, P.K., Wu, E.H.-K., Sahoo, J., 2013. DuRT: Dual RSSI Trend Based Localization for Wireless Sensor Networks, Sensors Journal, IEEE, 13, (8), p. 3115- 3123.
19. Mukhopadhyay, B., Sarangi, S., Kar, S., 2015. Performance evaluation of localization techniques in wireless sensor networks using RSSI and LQI, in Communications (NCC)', Twenty First National Conference, p.1-6.

Real-Time Attendance System Using Face Recognition

Pavithra[a,*], Sakthi Priya[b], Swathy[c], K L Vaishnavi[d], Indira[e]

[a*,b,c,d,e] *Department of Computer Science Engineering, Amrita School of Computing, Amrita Vishwa Vidyapeetham, Chennai – 601103, India*

Abstract

Attendance systems have evolved significantly over the years, and the traditional method of marking attendance using pen and paper has become obsolete. With advancements in technology, a more reliable and efficient attendance system can be developed using real-time face recognition techniques. This paper presents a real-time smart attendance system that uses live video capturing and face recognition based on the Histogram of Oriented Gradients (HOG) feature extraction technique and the OpenCV library. Here OpenCV is built on CNN. The proposed system can accurately and efficiently detect and recognize faces in a live video stream, allowing for real-time attendance tracking and monitoring. The system uses a HOG-based face detector to locate faces in the video stream, and then applies a face recognition algorithm based on OpenCV to recognize the faces and match them against a pre-existing data of registered students or employees. Attendance will be automatically displayed in a web application that can be accessed only by authorized person. Experimental results demonstrate that the proposed system achieves real-time performance and high accuracy, with an average recognition rate of 98%.

Keywords: OpenCV, Histogram of Oriented Gradients (HOG), Face Recognition, Haar cascade.

1. Introduction

Face recognition is currently a research field where several innovative techniques have been discovered for accurate and effective face identification. At many institutions, the conventional technique of marking travel times to work might be a tiresome process. Schools have an important duty to record attendance by identifying students, which might take up to five minutes for a complete session. This can be a lengthy process. As a result, various institutions started using a wide range of additional recording techniques, including fingerprint recognition and frequency identification (RFID). These web-based methods can, however, be time-consuming and disruptive. Attendance systems are used in various educational institutions and organizations to record the presence of employees or students. Traditionally, attendance was marked manually, but this method is time-consuming and prone to errors. Therefore, automated attendance systems have been developed to increase efficiency and accuracy.

1.1. Problem with Traditional Method

Automatic attendance systems based on RFID (Radio Frequency Identification) employ radio waves to identify and track people, typically using RFID tags placed in ID cards or badges. Although these methods can be useful and effective, there are a few potential issues that should be taken into account:

- Interference: Other electrical equipment or materials, such as metal, can interfere with RFID signals and cause disruptions, which might result in inaccurate attendance records.
- Proxy attendance is when someone falsely records their presence on behalf of someone else who isn't there. RFID based automatic attendance systems are susceptible to this kind of fraud and are not impervious to proxy attendance. For instance, even if the owner of the RFID ID card or badge is not there, they could still register attendance on their behalf if they have access to it.

1.2. Proposed Solutions

In order to avoid the above mentioned shortcomings, biometric based attendance systems can be used. The two common biometric identification technology are fingerprint recognition and face recognition. There are many fingerprint based attendance systems in use. But the major issue with fingerprint based system is that the accuracy is low and the performance depends on the sensors used and the environmental conditions (error rates are high during extreme winter or summer). The False Acceptance Rate (FAR) and False Rejection Rate (FRR) are high in fingerprint based systems (total error rate is 1% approximately). To

*Pavithra G
E-mail address: pavig026@gmail.com

overcome the above challenges faced by fingerprint based attendance systems, face recognition is adopted. Using face recognition, we can reduce the rate of error and thus build an accurate system. Also, face recognition systems are contact-less thereby limiting the spread of infections or illness. The system is highly secure compared to fingerprint based systems thereby preventing incidents like forgery or buddy punching.

2. Related Works

In this paper [1] Student Attendance System using Face Recognition proposed by S. Dev, they have used the images captured are ameliorated using Generative Adversarial Networks to retain texture information. For face detection, Haar classifiers are used and 68 landmarks of each face is taken into account. For feature extraction, Gabor filters are used. AdaBoost is used for removing redundant features. For face recognition, KNN, SVM and CNN are used and results are compared. Its advantages include the accuracy obtained using KNN is high. The system has low computational complexity and high precision. The result is invariant to different poses, lighting and appearance changes like beard specs etc. This system is cost efficient as well as needs less manual work. while its disadvantage is SVM is giving poor output. The time taken by the system is comparatively high. The system does not handle different facial expressions in a good manner.

In this paper [2] Automated Smart Attendance System using Face recognition proposed by K. Preethi, they have used Local Binary Pattern (LBP) Histogram. Its advantage actually include Records attendance along with time stamp Have reliable outcomes for pose variance, and illumination. Takes less time to process the whole image and its disadvantages include that it can only recognize one face at a time

In this paper [3] Design of Intelligent Classroom Attendance System Based on Face Recognition proposed by W. Zeng, they have used used deep learning-related ideas to improve the AlexNet convolutional neural network with the use of the WebFace data set to improve the network training and testing. Its advantages are efficient, stable, reduced attendance costs and its disadvantage is time-consuming.

In this paper [4] Student Attendance System in the Classroom Using the Face Recognition Technique proposed by S. Lukas, techniques used were HFR has been widely used in many applications with a Combination of Discrete Wavelet Transforms (DWT) and Discrete Cosine Transform (DCT) to extract the features of the student's face which are followed by applying Radial Basis Function (RBF) for classifying the facial objects based on the known face recognition techniques in its endeavor to develop a specific computer application which can be used for recognizing any enrolled student feature extraction is involved with a purpose to extract features from any student's facial image that is required to come in a uniform size, in this case, 64x64 pixels. Its advantages are that it can be used for user authentication, recognition process becomes more robust. But it has one disadvantage Accuracy also varies when the image is blurred or dark

In this paper [5] Face Recognition System Face Recognition System proposed by Shivam Singh, they have used used KLT Algorithm, Viola-Jones Algorithm face detection which detect human face using Haar cascade classifier They apply a model combining to match the geometric characteristics of the human face. It is machine learning based approach where a cascade function is trained from images. It is used to detect objects in other images. Its advantage is that its automatically identifying a person from a still image or video frame reliable, secure and fast. It requires improvement in different lighting conditions which happens to be its disadvantage.

In this paper [6] an efficient automated attendance management system based on eigen face recognition proposed by E.Rekha, they have used PCA technique and it's advantages is to create a safe environment, it automatically update and save the attendance. Its disadvantage is that it time consuming and manipulation of attendance.

In this paper [7] Face Recognition Systems Under Morphing Attacks they have used robust algorithms proposed by U.Scherhag, printing and scanning technology. Its advantages are the manual annotation of images is very accurate automated detection of landmarks, reliable detection of morphed face images. Time consuming is a disadvantage.

In this paper [8] Integrating Conventional and Inverse Representation for Face Recognition research paper proposed by Y. Xu. They used conventional and inverse representation-based linear regression classification methods is used. Its advantage is that Face detection has proved successful when using the symmetry of the face, achieves extremely high precision and is fairly noise resistant. Its disadvantage is that contains large memory and is time taking.

In this paper [9] Face Recognition Based Attendance System proposed by Mekala, the technique is that the face recognition process is performed by using the Cognitive Face API which follows the PCA (Principal Component Analysis) algorithm. A collection of 25 photos taken from various angles is made for each student. Using the cognitive face API, features are extracted, and a database is created. The benefit is that it makes an effort to prevent human error and serves as a high accuracy substitute for recording student attendance. Its drawback is that it has slow speed and insufficient storage space for the data of each student.

In this paper [10] PCA based Facial Recognition for Attendance System proposed by T. A. Kiran, few techniques like human face recognition (HFR), Optimal character recognition, Principle component analysis method are used. It is time saving and has No false acceptance are its advantages. Its disadvantage is that it takes more storage and more time to recognize.

In this paper [11] Attendance system based on face recognition system using CNN-PCA method and real-time camera proposed by Winarno .They used hybrid feature extraction method using CNN-PCA. Its advantage is that it is effective and accurate in recognizing human faces in real time.

In this paper [12] High performance and efficient real-time face detector on central processing unit based on convolutional neural network proposed by Putro. They used Central Processing Unit based Convolutional Neural Network. Its advantage is that it requires less computation cost. The disadvantage is that it requires more computation power.

3. Methodology

The proposed system uses a HOG-based face detector to locate faces in the video stream, and then applies a face recognition algorithm based on OpenCV face recognition method to recognize the faces and match them against a pre-existing database of registered students.

3.1. Architecture

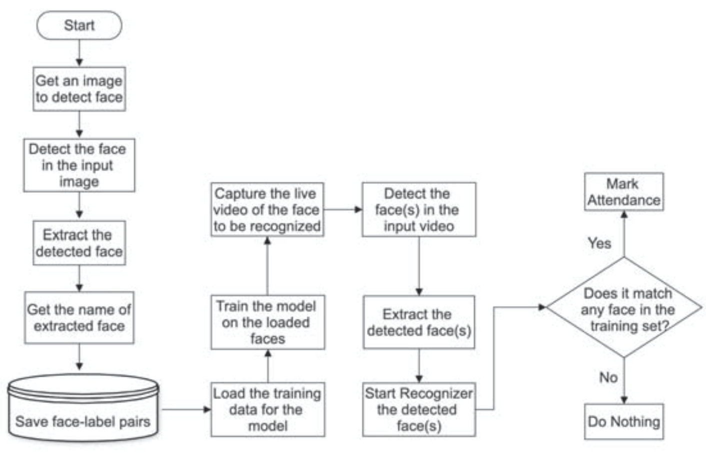


Fig. 1. System Architecture.

In the Fig.1, we give input as an image to detect the face and from that we extract features. Once we extract, we get the name for that individual in the image and save the face-label in the database. Then the training phase begins where we start to load the training data and train the model. We then proceed to the live stream video from which we need to detect the face and extract features from it. Once detection of face is completed, we move to the face recognition where the model will tell us whether the face recognized from live stream video matches in the training set. If it matches it will mark attendance , else it will do nothing.

3.2. Histogram of Oriented Gradients

A prominent feature extraction technique for object detection is Histogram of Oriented Gradients (HOG). HOG works by extracting gradient information from images and using it to represent the shape and look of an item. HOG is especially good at detecting objects with well-defined edges and texture, making it a good feature extraction technique for a real-time face recognition-based attendance system.

The HOG feature extraction process involves the following steps:

- Image Preprocessing: The input image is Preprocessing to remove noise and enhance contrast.

- Gradient Calculation: The image gradient is calculated using derivative filters such as Sobel, Prewitt or Scharr operators. The gradient is a measure of the rate of change of intensity at each pixel in the image.
- Orientation Binning: The image is divided into small cells, and the gradient orientation and magnitude of each pixel in the cell are computed. The orientations are then quantized into a fixed number of bins (e.g., 9 bins), and the magnitude is used to weight the contributions of each pixel.
- Histogram Calculation: A histogram of the orientations within each cell is computed by summing the weighted magnitudes of the gradients. The histograms are then concatenated to form a feature vector that represents the shape and texture of the object in the image.
- Normalization: The feature vector is normalized to reduce the effects of illumination variations and to make the feature vector more robust to changes in scale.

3.3. Open Source Computer Vision Library

For a variety of image and video processing applications, OpenCV (Open Source Computer Vision Library) is a free machine learning software library. Due to its capability to do real-time face detection and recognition, it is frequently utilized for face recognition applications. For face identification, OpenCV offers a number of pre-trained classifiers, such as Haar Cascades, Local Binary Patterns (LBP), and Histogram of Oriented Gradients (HOG). These classifiers operate by examining the image and locating the facial traits that are distinctive to faces. OpenCV provides solutions for face recognition, feature extraction, and facial landmark identification once a face has been identified. Moreover, OpenCV offers a user-friendly interface with tools for developing and managing face databases, recording live video streams, and showing identification results for face recognition applications.

3.4 Face Detection

Face detection is necessary because the image captured by the camera is fed to the system, and the face detection algorithm is applied to determine the presence of human faces in that image. Several image processing algorithms are utilized for detecting faces in images and the location of each of the identified faces. We used the HOG method to detect human faces in the provided image.

3.5 Face Positioning

A human face comprises 68 locations. This step's primary purpose is to detect particular points on faces and position the image. Face recognition algorithm is used to identify face points and the location of the face as precisely as possible without blurring the picture.

3.6 Face Encoding

The next stage is to extract the distinctive facial characteristic that can be utilized for recognizing each image after faces have been found in the input image. Whenever we obtain face localization, facial points are accurately retrieved for each input picture and stored in data in a folder for recognizing faces.

3.7 Face Matching

The method of facial recognition concludes at this phase. Deep learning, is one of the finest learning methods and generates feature vectors with real value. Our system creates ratification for each face after validating it. The face enc.py procedure is being used internally to estimate the faces in each image and the whole dataset. If the actual picture matches, it will move to attendance marking.

4. Result

In our project, we are not using any existing or available image data set in this instance. By capturing pictures of our classmates in various poses, angles and storing them as distinct classes, we created our own data set. A phone is used to take photos for the data set. Confusion Matrix determines the proportion of accurate and inaccurate estimates, which is then summed with the number of count values and their breakdown per class. It may be employed to measure things such as recall, precision, and accuracy. The sklearn metrics module provides a confusion matrix. The confusion matrix containing TP, FP, FN, and TN is shown in Table. Our model has an accuracy of 98%.

Table 1. Evaluation Metrics.

Dataset	TP	TN	FP	FN	Se.	Sp.	Acc.
IMPA-FACE3D	113	0	0	0	1.0	0.0	1.0
Work Dataset	43	2	1	1	0.9770	0.6667	0.9575
Total	156	2	1	1	0.9936	0.6666	0.9857

(a) (b) (c)

Fig. 2. Training Data (a) Face Recognized as Unknown Before Training and (b) Face Recognized as Per Their Name After Training and (c) Face Recognized as Per Their Name in Different Pose.

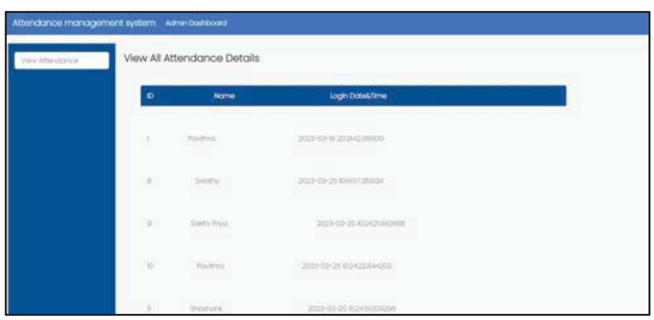


Fig. 3. Attendance Displayed In Web Application

5. Conclusion and Future Work

In this paper, we proposed a real-time face recognition based attendance system using HOG for face detection and OpenCV for face recognition. The proposed system achieved high accuracy of 98% and speed, making it suitable for practical applications. The proposed system can be further optimized by using more sophisticated feature extraction techniques and classifiers. The solution in our model is cost effective and efficient when compared to other biometric solutions. The data acquired for real time face recognition is large the cost and time are saved in this model. Since the whole process is automated with less intervention of humans, so there is no requirement of additional labour to perform the work manually. This model can be developed without the need for any specialized hardware. To develop real-time attendance system using face recognition a laptop with camera is sufficient. Future work can also explore the use of deep learning techniques for face recognition in real-time attendance systems.

References

1. Dev, S. and Patnaik, T., 2020. Student attendance system using face recognition. In 2020 international conference on smart electronics and communication (ICOSEC), IEEE, p. 90-96.
2. Preethi, K. and Vodithala, S., 2021. Automated smart attendance system using face recognition. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS), IEEE, p. 1552-1555.
3. Zeng, W., Meng, Q. and Li, R., 2019. Design of intelligent classroom attendance system based on face recognition. In 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE, p.611-615.
4. Lukas, S., Mitra, A.R., Desanti, R.I. and Krisnadi, D., 2016. Student attendance system in classroom using face recognition technique. In 2016 International Conference on Information and Communication Technology Convergence (ICTC), IEEE, p. 1032-1035.
5. Singh, S. and Jasmine, S.G., 2019. Face recognition system. International Journal of Engineering Research Technology (IJERT), 8(5).
6. Rekha, E. and Ramaprasad, P., 2017. An efficient automated attendance management system based on Eigen Face recognition, 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence, IEEE, p. 605-608.
7. Scherhag, U., Rathgeb, C., Merkle, J., Breithaupt, R. and Busch, C., 2019. Face recognition systems under morphing attacks: A survey. *IEEE Access*, *7*, p.23012-23026.
8. Xu, Y., Li, X., Yang, J., Lai, Z. and Zhang, D., 2013. Integrating conventional and inverse representation for face recognition. IEEE transactions on cybernetics, 44(10), p.1738-1746.
9. Mekala, V., Vinod, V.M., Manimegalai, M. and Nandhini, K., 2019. Face recognition based attendance system. International Journal of Innovative Technology and Exploring Engineering, 8(12), p.520-525.
10. Kiran, T.A., Reddy, N.D.K., Ninan, A.I., Krishnan, P., Aravindhar, D.J. and Geetha, A., 2020, September. PCA based Facial Recognition for Attendance System, International Conference on Smart Electronics and Communication (ICOSEC), IEEE, p. 248-252.
11. Winarno, E., Al Amin, I.H., Februariyanti, H., Adi, P.W., Hadikurniawati, W. and Anwar, M.T., 2019, December. Attendance system based on face recognition system using cnn-pca method and real-time camera, International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), IEEE, p. 301-304.
12. Putro, M.D., Kurnianggoro, L. and Jo, K.H., 2020. High performance and efficient real-time face detector on central processing unit based on convolutional neural network. IEEE Transactions on Industrial Informatics, 17(7), p.4449-4457.

Multimodal Sentiment Analysis: Exploiting Other Modalities for Improved Results

Vatsal Kansara[a*], Madhuri Chopade[b]

[a*]*PG Student, Computer Engineering Department, Gandhinagar Institute of Technology, Gandhinagar 382721,Gujarat, India*
[b]*Assistant Professor, Computer Engineering Department, Gandhinagar Institute of Technology, Gandhinagar 382721,Gujarat, India*

Abstract

An emerging field of NLP multimodal sentiment analysis (MSA) based on synthesized embeddings extracted from textual, visual, and acoustic sources predicts a speaker's sentiment tendencies based on both textual and acoustic information. These embeddings are generated by combining input unimodal raw data to produce a richer multimodal feature representation. By employing Transformer-based models, lexicon-based features perform better than other modalities due to their pre-training on large corpora of text. In spite of their powerful performance, learning a new self-supervised learning Transformer on a new modality is often not possible due to an insufficient amount of data, which is the case in language learning using multiple modalities. Bidirectional Encoder Representations from Transformer (BERT) is an efficient pre-trained language model which provides state-of-the-art results on question answering and natural language inference and many others. However, most of the work in sentiment analysis is worked based on textual data, how to learn better representations with multimodal information is still an area that needs to be explored. There are multiple modalities that we can exploit to get a better sentiment analysis of a person’s behavior. Here our approach is to leverage the audio-based features combined with lexical features to get an efficient and powerful model for sentiment analysis trained on the CMU-MOSI dataset.

Keywords: Multimodal Sentiment Analysis, CMU-MOSI Dataset, Transformers.

1. Introduction

Human communication does not limit itself to words. It includes acoustic annotations, body movements, facial expressions, and even body language. With the proliferation of social media platforms such as Facebook, WhatsApp, and YouTube, people produce large volumes of multimodal data with rich sentiment information every day. In addition to its role in human-computer interaction, sentiment analysis plays a vital role in artificial intelligence development and has been widely used in a variety of applications, including automatic driving, human-machine communication, and much more [1]. Text-based features which basically express sentiment through words, phrases, and relations generally outperform other modalities [2]. In recent times text sentiment analysis has achieved a lot of attention. Using pre-trained word vectors as inputs, TextCNN [3] is able to give state-of-art results on four out of seven sentiment classification tasks.

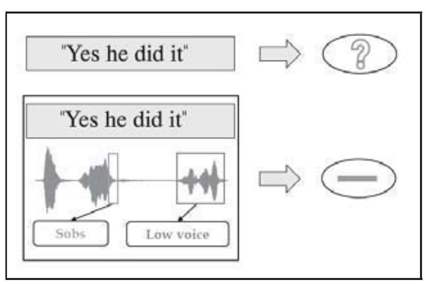


Fig. 1. Example of MSA. [16]

However, in some cases, it becomes difficult to identify the correct emotion of a person when there is only lexicon-based features are available. Many times, text modality also coexists with audio modality. Audio modality includes loudness, pitch, energy, vocal effort, and other frequency-related characteristics that convey sentiment information [4]. A combination of text and audio channels can give a more comprehensive picture of information and capture more emotions. In Figure - 1, we can see an example

*Vatsal Rakesh Kansara
E-mail address: 210120702006@git.org.in

of intermodally of text and audio. The sentence “Yes he did it” is ambiguous as it can address a variety of emotions in different situations. It's challenging to determine the sentiment of that kind of sentence just by taking those series of words as inputs. After introducing the speaker’s audio-related information, we can clearly identify that the intent of the speaker is negative in this scenario. Multimodal Sentiment Analysis (MSA) has been gaining increased popularity in recent years as a way to compensate for the disadvantage of a single modality of affective computing.

Recent large language models use Self-Supervised Learning and Transformer based methods utilizing large number of datapoints for pre-training in textual, vision and multimodal contexts. Collection of large volumes of data in multimodal context is not always easy. Recently, Bidirectional Encoder Representations from Transformer has presented state-of-the-art results on Eleven NLP tasks. The contextual representations generated by BERT are conditioned on both left and right context in each layer. So, the inference basically describes context content [5]. We explore approach to generalize pre-trained language model which exposes audio modality for better and efficient sentiment analysis.

2. Theory Background

An important subset of Natural Language Processing (NLP) is sentiment analysis, which focuses on identifying qualitative characteristics in data such as sentiment, opinion, thoughts, and behavioral intent. Many approaches [34] have been proposed to achieve state-of-the-art results in sentiment analysis. Recent advancements in supervised learning approaches have led to the growth of the field of Multimodal Sentiment Analysis.

2.1. Approaches

Based on the given feature, categorizing the polarity into Positive, Negative, or Neutral is important task. The approaches towards sentiment analysis are divided in to following categories.

2.1.1. Lexicon Based

A lexicon-based approach is a traditional approach for sentiment analysis that uses a set of manually framed rules to classify the input into positive or negative opinions [26]. Words that elicit positive or negative feelings in humans are scanned through documents using lexicon-based methods. The method involves different Natural Language Processing techniques like Lexicons, Stemming, Part of Speech, Tokenization, etc. This approach counts the number of words representing various emotions in the input, then classifies them accordingly. Sentiment analysis shows to be extremely dependent on the domain of interest [27]. For example, analyzing specific product reviews can yield very different results compared to analyzing social media data due to different forms of language used. This is one of the issues and adjusting the model based on that can be a time-consuming process. However, lexicon-based methods do not require training data which can be beneficiary for small datasets.

There are two lexicon-based approaches:

- Dictionary-based
- Corpus-based

In the Dictionary-based approach a small set of sentiment words are prepared and then the process iteratively expands the lexicon of sentiment words with synonyms and antonyms from existing dictionaries. Domain-specific lexicons can be created using corpus-based lexicons. The approach starts with a list of general-purpose sentiment words and discovers other sentiment words from a domain corpus based on co-occurring word patterns [29].

2.1.2. Machine Learning Based

Machine learning-based approaches for sentiment analysis can be divided as below:

- Unsupervised Learning
- Semi-supervised Learning
- Supervised Learning.

Unsupervised Learning algorithms mostly work with unlabeled type data and with the help of algorithms, different patterns/structures are discovered in data. One of the advantages of this kind of method is that we can use large datasets without investing more effort in human supervision and labeling.

Supervised Learning is one of the most widely known machine learning methods. In this approach, the datasets are pre-labeled with classes or ground truth. In this method, a model is trained with labeled source data, which can be used in making predictions using new unlabeled input data. In most cases, supervised learning often outperforms unsupervised and semi-supervised learning approaches, but the dependency on labeled training data can require lots of human effort and is therefore sometimes inefficient [26]. Some examples of this approach are Support Vector Machine (SVM), Neural Networks, Bayesian Networks, and Naïve Bayes.
Semi-Supervised Learning can work with both labeled and unlabeled datasets where unlabeled data are complemented with the labeled examples for training the model. Compared to supervised learning this method requires less effort in labeling and yields decent accuracy results.

2.1.3. Hybrid approach

Hybrid approaches mainly aim to extend machine learning models with lexicon-based knowledge [11]. Using both lexicon and machine learning-based features, the goal is to yield optimal results by combining both methods [12]. For example, Minaee et al. [13] developed an ensemble model using LSTM and CNN algorithms and demonstrated that this ensemble model provides better performance than the individual models.

A study by Pooja Mehta and Dr. Sharnil [14] shows that machine learning methods, such as SVM, Naive Bayes, and neural networks have the highest accuracy and can be considered as the baseline learning methods as well as in some cases lexicon-based methods are very effective. Hence our study on sentiment analysis will mostly be covering machine learning aspect.

2.2. Levels of Sentiment Analysis

Three different levels are considered while doing sentiment analysis namely the document level, the sentence level, and the entity or aspect level.

2.2.1. Document Level

The document-level analysis considers the whole text document as a unit of analysis [21]. This task assumes that only one opinion holder is involved in the creation of the document. It should be noted that document analysis has its own issues, such as the fact that multiple mixed opinions are sometimes expressed in the same document through implicit language [22].

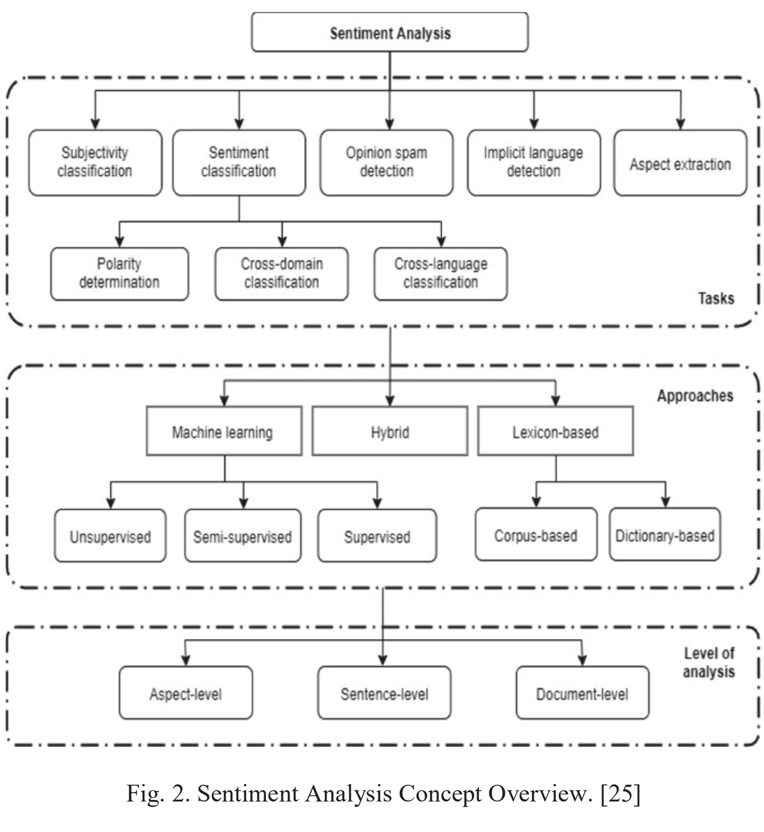


Fig. 2. Sentiment Analysis Concept Overview. [25]

2.2.2. Sentence Level

In sentence-level analysis, specific sentences in text are considered and categorized into different categories based on their sentiments. Basically, the method analyses individual sentences in a document to detect whether the sentence contains facts or emotions, and opinions. A sentence is neutral when it does not imply any opinion. In case a sentence is neutral, it’s termed an objective sentence, which is a sentence that reveals facts, as opposed to subjective sentences which present opinions and subjective views. Subjectivity and objectivity classification are the main advantages of sentence analysis [22].

2.2.3. Aspect Level

Aspect level sentiment analysis also known as entity-level or feature-level sentiment analysis refers to analyzing sentiments about specific entities and their aspects in a text document, not merely the overall sentiment of the document. The output of this Aspect level more detailed analysis as compared to both document and sentence-level analysis. It’s possible for opinion holders to have divergent opinions about specific aspects of an entity, despite the general sentiment of the document which can be either positive or negative [22]. To measure aspect-level opinion, aspects of the entity need to be identified [24]. stated that aspect-based sentiment analysis is beneficial to the business manager because customer opinions are extracted in a transparent way.

2.3. Deep Learning Based Methods

Figure – 3 Demonstrates recent advancements in sentiment analysis space and it's evident that Language models are one of the latest and biggest advancements in NLP and sentiment analysis space. One sub-branch of machine learning is Deep Learning which consists of approaches like CNN, RNN, LSTM[30], etc.

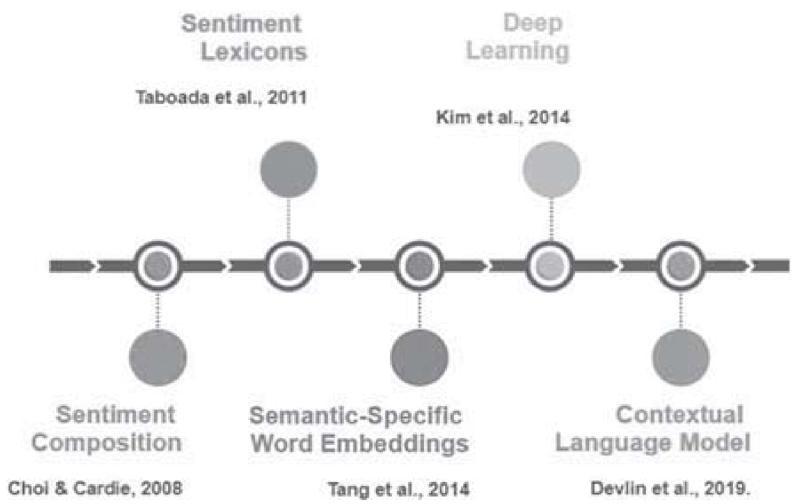


Fig. 3. Milestones of Sentiment Analysis. [25]

The latest advanced Language Models are built on top of many deep learning-based concepts like RNN, LSTM, Transformers, and Word Embeddings. Many Language models like BERT[33], GPT-2[32], and GPT-3[31] provide state-of-the-art performance on sentiment analysis data. To train them we just have to leverage a pre-trained model with a fine-tunning layer added on top of that for general NLP tasks.

2.4. Multimodal Sentiment Analysis

As an extension to traditional text-based sentiment analysis, multimodal sentiment analysis includes features such as text, speech as well as visuals. For multimodal sentiment analysis, a variety of two modality combinations image+text, speech+text, and speech+image combinations can be used which are called bimodal systems or some proposed models use a combination of all three of them. The use of visual features can describe something more effective than a long list of written words. Audio data can also provide important indicators such as pauses, the larger number of pauses indicates natural sentiment. The image below illustrates different phases of sentiment analysis based on multimodal fusion combining audio-visual features.

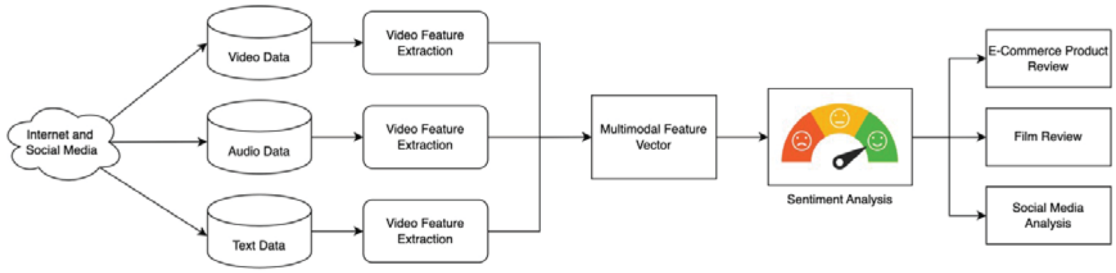


Fig. 4. Multimodal Sentiment Analysis Process.

3. Literature Survey

3.1. Contextual Inter-modal Attention for Multi-modal Sentiment Analysis

An important challenge associated with multimodal sentiment analysis is combining text, visual and acoustic inputs effectively. RNN-based approach was taken by Deepanway Ghosaly, Md Shad Akhtary, Dushyant Chauhany et al. [20] for a multi-modal attention framework that leverages the contextual information for utterance-level sentiment prediction. The approach taken applies attention to multi-modal multi-utterance representations and tries to learn the features. This approach was able to mark an F1 score of 82.31.

3.2. Multimodal Transformer for Unaligned Multimodal Language Sequences

Yao-Hung Hubert Tsai, Shaojie Bai, Paul Pu Liang et al.[19] explores the challenges like inherent data non-alignment due to variable sampling rates for multimodal sequences and long-range dependencies between elements across modalities. Multimodal Transformer (MulT) introduced here generically address these issues without explicitly aligning the data. Directional pairwise cross-modal attention attends to the interaction between multimodal sequences across distinct time steps and latency adapts streams across modalities. The model is evaluated with IEMOCAP, CMU-MOSI, and CMU-MOSEI datasets with an F1 score of 81.6 on the CMU-MOSEI dataset.

3.3. Gated Mechanism For Attention Based Multimodal Sentiment Analysis

In recent times key to achieving state-of-art performance on downstream NLP tasks is to fine-tune the trained contextual language model on task-specific datasets. For tasks related to lexical data, it's straightforward while, it’s not trivial for multimodal language. Challenges like cross-modal sentiment analysis, learning long-term dependencies, and Fusion of unimodal and cross-modal cues multimodal sentiment analysis were addressed by Ayush Kumar and Jithendra Vepa [18]. The aim of their approach is to learn the interaction between different modalities controlled by learnable gates. By this approach, F1 score of 81.17 was achieved also thy discuss issues related to audio quality leading to lower speech recognition accuracy eventually affecting sentiment score.

3.4. VAE-Based Adversarial Multimodal Domain Transfer for Video-Level Sentiment Analysis

Yanan Wang et al. [17] introduced VAE-based adversarial multimodal domain transfer (VAE-AMDT) and jointly trained it with a multi-attention module to reduce the distance difference between unimodal representations. For visual, linguistic, and acoustic representations to follow common distributions they perform variational autoencoder (VAR) and then transfer their all-unimodal representations to joint embedding space using adversarial space. They jointly train VAE-AMDT and multi-attention modules which consist of self-attention, cross-attention, and triple-attention components reaching to F1 score of 84.3 on the MOSI dataset. Also, Yanan Wang et al. [17] discussed future steps to improve their approach by using more powerful encoders to extract unimodal representations such as face identification methods for emotion.

3.5. CM-BERT: Cross-Modal BERT for Text-Audio Sentiment Analysis

Similar to Ayush Kumar and Jithendra Vepa [18] In Kaicheng Yang et al.[16] proposed Cross-Modal BERT (CM-BERT), which relies on the interaction of text and audio modality to fine-tune the pre-trained BERT model. The input of the model consists of text features which are basically output encoding of BERT's last layer and Audio features got by some pre-processing such as word-level alignment. Using text and audio modality together, the introduced multimodal attention dynamically adjusts word weight. They evaluated the modal on CMU-MOSI and CMU-MOSEI datasets achieving an F1 score of 84.5.

3.6. Integrating Multimodal Information in Large Pretrained Transformers

Wasifur Rahman et al. [15] designed a framework that allows BERT and XL-Net core structures to remain intact and only attaches a carefully designed Multimodal Adaption Gate to the models. They basically use conditional attention on nonverbal behaviors, the gate essentially maps the informative visual and acoustic factors to a vector with trajectory magnitude, and during fine-tuning this adaption vector modifies the internal state of BERT and XL-Net, allowing the models to seamlessly adapt to nonverbal inputs. By using this model achieved powerful performance (MAG-BERT - F1 Score 86.00) on CMU-MOSI (Multimodal Opinion Sentiment Index) dataset.

4. Challenges and Issues

As per the literature review few of the challenges and issues which are there in the multimodal analysis space can be found below.

4.1. Extend and test approaches to real-world datasets and scenarios

The effectiveness of multimodal sentiment analysis approaches needs to be tested on a variety of real-world datasets and scenarios to ensure that they are robust and reliable in practical applications.

4.2. Handling implicit sentiments like sarcasm

Identifying implicit sentiments, such as sarcasm or irony, is a significant challenge in multimodal sentiment analysis since these sentiments are often conveyed through indirect or ambiguous language.

4.3. Improving model accuracy and efficiency

Multimodal sentiment analysis models often require large amounts of data and computational resources, which can make them slow and computationally expensive. There is a need to improve the accuracy and efficiency of these models, while also reducing their computational requirements.

4.4. Handling different representations of different modalities

Different modalities, such as text, image, and audio, have different representations, and integrating them in a meaningful way requires sophisticated algorithms that can extract relevant features from each modality.

4.5. Limited availability of data for non-text modalities:

There is often a scarcity of labeled data for modalities other than text, such as images and audio, which makes it challenging to develop robust multimodal sentiment analysis models.

4.6. Generalizability of the models

Multimodal sentiment analysis models may perform well on specific datasets or in certain scenarios, but their generalizability to new datasets or scenarios is often limited. There is a need to develop models that can generalize well to a wide range of datasets and scenarios.

5. Problem Statement, Aims and Objectives

With great advancement in the digital world and communication technology, as the increasing popularity of various portable devices large amount of data is being uploaded as audio, video, and text. Consider an example of consumers of the products who record their reviews on a product on social media platforms like YouTube, and Facebook to inform a larger population about their views and opinions. This huge amount of data being generated every day and the requirement of organizations to make a quick decision about products and services from feedback requires an accurate and automated solution that can give insights about opinions and emotions about the individuals.

The primary advantage of analyzing audio and video over simple text analysis, for detecting emotions and sentiment, is the surplus of tonal and behavioral cues, and as the research on similar subjects shows that even just including audio modalities with text modalities can give us a significant boost to the sentiment analysis accuracy.

Recently many approaches have been proposed for multimodal sentiment analysis giving state-of-the-art results. However as discussed major issues and challenges remains there like the impact of each modality across datasets, the generalization ability of multimodal sentiment classifier, improving accuracy for larger databases, and handling different representations from different modalities.

The main goal here would be to make a powerful and efficient NLP modal that can do Multimodal Sentiment Analysis on real-world streaming data which then can be used in many applications of NLP such as Conversational AI, human-computer interactions, etc.

Specifically primary focus is to evolve current solution with below Aims and Objective,

1. To Improve accuracy
2. To lookout for solution with real time sentiment detections

6. Proposed Work

Language models built based on transformers provide a powerful way of learning for NLP tasks. Since lexical features are trained on the large text of corpora, they usually outperform other modalities in this domain. However, as discussed introducing multimodal data can provide us with more cues about a person's opinion which can help us to get better at tasks [2].

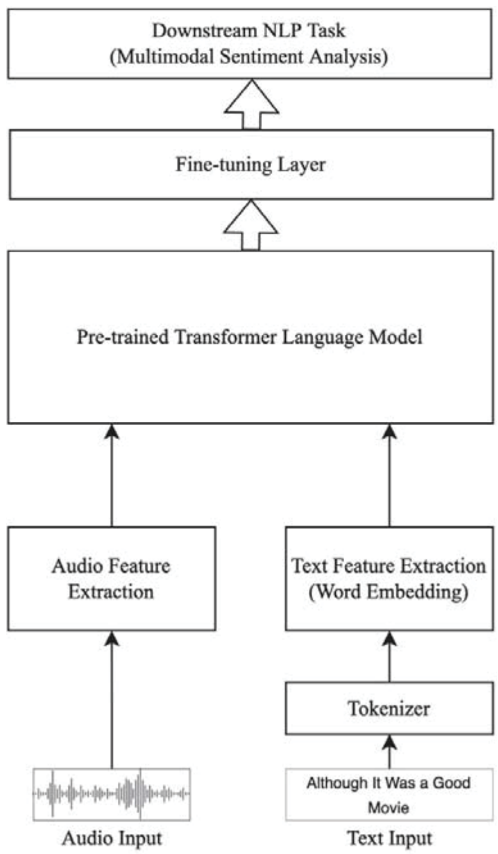


Fig. 5. Proposed Model.

Our focus would be more on working with audio modality in addition to lexicon features as working on raw speech data is less computationally intensive and also provides a more informative advantage in multimodal language settings [2].

As with the major language models approach, we plan to train the model in two phases. 1. Train a Pre-Trained Language Model with audio features on training tasks (Ex. Masked Language Model) for a smaller number of epochs. 2. Train the model on an actual dataset for sentiment analysis with a finetuning layer added on top of the trained model.

The main challenge here would be to make an audio feature extraction module that we can add as a prefix to an already trained model on large lexical data. Providing a small number of epochs in the first step with prefixed input will help the model to adapt prefixed audio features. Then in the second step, we can train the model on the actual database with the NLP task at hand i.e. multimodal sentiment analysis. The proposed approach will be more efficient in terms of training as compared to current systems which use different branches for preprocessing audio and then generating features and combining output before fine-tuning layer.

OpenAI Whisper APIs will be used to get a real-time and accurate transcription of the user speech which will serve as inputs to the multimodal sentiment analysis model. The overall flow for the prediction can be found in the below image.

We purpose to train such model on multimodal datasets available like CMU-MOSI, CMU-MOSEI, IEMO-CAP, etc., and to test on real-world scenarios to make an efficient and powerful model.

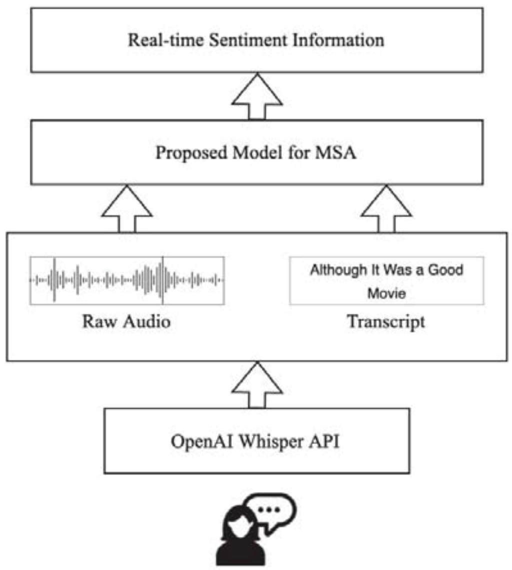


Fig. 6. Prediction Flow.

7. Conclusion

Combining other modalities with lexical data and harnessing the power of the already trained model with transformers on large text corpora can give us a significant boost in performance for NLP tasks at hand specifically sentiment analysis. The proposed model is efficient in terms of leveraging already known features from the lexical data model compared to other approaches that require training a new cross-model transformer model which we can extend to larger data sets for better generalizability.

8. Future work

In future research, there is a need to develop strategies to handle implicit sentiments, including sarcasm, as they can significantly impact the sentiment analysis results. Furthermore, to enhance the practical applicability of the models, it is essential to focus on improving their generalizability across different domains. One potential approach could be to incorporate visual data, such as images and videos, to better capture contextual information and improve the overall accuracy of the sentiment analysis models.

References

1. Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. 2018. Multimodal machine learning: A survey and taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence.
2. Erik Cambria, Devamanyu Hazarika, Soujanya Poria, Amir Hussain, and R. B. V. Subramanyam. 2017. Benchmarking Multimodal Sentiment Analysis.
3. Yoon Kim. 2014. Convolutional Neural Networks for Sentence Classification.
4. Runnan Li, Zhiyong Wu, Jia Jia, Yaohua Bu, Sheng Zhao, and Helen Meng. 2019.Towards discriminative representation learning for speech emotion recognition. In Proceedings of the 28th International Joint Conference on Artificial Intelligence (IJCAI), P.5060–5066.

5. Xinlong Li, Xingyu Fu, Guangluan Xu, Yang Yang, Jiuniu Wang, Li Jin, Qing Liu, and Tianyuan Xiang, 2020. Enhancing BERT Representation With Context-Aware Embedding for Aspect-Based Sentiment Analysis.
6. Jiahong Yuan and Mark Liberman, 2008. Speaker identification on the SCOTUS corpus.
7. Gilles Degottex, John Kane, Thomas Drugman, Tuomo Raitio, and Stefan Scherer, 2014. COVAREP A collaborative voice analysis repository for speech technologies.
8. Choi Y, Cardie C., 2008. Learning with compositional semantics as structural inference for subsentential sentiment analysis.
9. Duyu Tang, Furu Wei, Nan Yang, Ming Zhou, Ting Liu, and Qin Bing., 2014. Learning Sentiment-Specific Word Embedding for Twitter Sentiment Classification.
10. Devlin J, Chang MW, Lee K, Toutanova K., 2019. BERT: Pre-training of deep bidirectional transformers for language understanding.
11. Behera RN, Manan R, Dash S., 2016. Ensemble based hybrid machine learning approach for sentiment classification-a review.
12. Ahmad M, Aftab S, Ali I, Hameed N., 2017. Hybrid tools and techniques for sentiment analysis: a review.
13. Minaee S, Azimi E, Abdolrashidi A., 2019. Deep-sentiment: sentiment analysis using ensemble of cnn and bi-lstm models.
14. Pooja Mehta, Dr. Sharnil Pandya., 2020. A Review On Sentiment Analysis Methodologies, Practices And Applications.
15. Wasifur Rahman, Md. Kamrul Hasan, Sangwu Lee, Amir Zadeh, Chengfeng Mao, Louis-Philippe Morency, Ehsan Hoque. 2020. Integrating Multimodal Information in Large Pretrained Transformers.
16. Kaicheng Yang, Hua Xu, Kai Gao., 2020. CM-BERT: Cross-Modal BERT for Text-Audio Sentiment Analysis
17. Yanan Wang, Jianming Wu, Kazuaki Furumai, Shinya Wada, Satoshi Kurihara. 2022. VAE-Based Adversarial Multimodal Domain Transfer for Video-Level Sentiment Analysis.
18. Ayush Kumar and Jithendra Vepa., 2020. Gated Mechanism for Attention Based Multimodal Sentiment Analysis.
19. Yao-Hung Hubert Tsai, Shaojie Bai, Paul Pu Liang, J. Zico Kolter, Louis-Philippe Morency, Ruslan Salakhutdinov., 2019. Multimodal Transformer for Unaligned Multimodal Language Sequences
20. Deepanway Ghosaly, Md Shad Akhtary, Dushyant Chauhany, Soujanya Poria, Asif Ekbaly, Pushpak Bhattacharyya, 2018. Contextual Inter-modal Attention for Multi-modal Sentiment Analysis.
21. Wang G, Sun J, Ma J, Xu K, Gu J., 2014. Sentiment classification: the contribution of ensemble learning. Decis Support Syst 57, p.77–93.
22. Kumar A, Sebastian TM., 2012. Sentiment analysis: a perspective on its past, present and future. Int J Intell Syst Appl 4(10), p.1–14.
23. Taboada M, Brooke J, Tofiloski M, Voll K, Stede M., 2011. Lexicon-based methods for sentiment analysis.
24. Valdivia A, Luzón MV, Herrera F., 2017. Sentiment analysis in trip advisor.
25. Ligthart, A., Catal, C. & Tekinerdogan, B. Systematic reviews in sentiment analysis: a tertiary study.
26. Hemmatian F, Sohrabi MK., 2017. A survey on classification techniques for opinion mining and sentiment analysis.
27. Vinodhini G, Chandrasekaran RM., 2012. Sentiment analysis and opinion mining: a survey.
28. Shayaa S, Jaafar NI, Bahri S, Sulaiman A, Seuk Wai P, Wai Chung Y, Piprani AZ, Al-Garadi MA., 2018. Sentiment analysis of big data: Methods, applications, and open challenges.
29. Mite-Baidal K, Delgado-Vera C, Solís-Avilés E, Espinoza AH, Ortiz-Zambrano J, Varela-Tapia E., 2018. Sentiment analysis in education domain: a systematic literature review.
30. Ralf C. Staudemeyer, Eric Rothstein., 2019. Morris Understanding LSTM - a tutorial into Long Short-Term Memory Recurrent Neural Networks.
31. Tom B. Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child. Aditya Ramesh, Daniel M. Ziegler, Jeffrey Wu, Clemens Winter, Christopher Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, Dario Amodei., 2020. Language Models are Few-Shot Learners.
32. Alec Radford, Jeffrey Wu, Rewon Child, David Luan, Dario Amodei, Ilya Sutskever., 2019. Language Models are Unsupervised Multitask Learners.
33. Jacob Devlin, Ming-Wei Chang, Kenton Lee, Kristina Toutanova. 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.
34. Shah, A., Chopade, M., Patel, P., & Patel, P. 2022. Survey: Emotion Recognition from Text Using Different Approaches. In Futuristic Trends in Networks and Computing Technologies: Select Proceedings of Fourth International Conference on FTNCT 2021, Singapore: Springer Nature Singapore, p. 433-445.

Acknowledgement

I would like to thank my esteemed research guide Prof. Madhuri Chopade for her invaluable supervision, support, and tutelage during the thesis work providing all the necessary information and discussion about my thesis progress with the devotion of her valuable time. Her immense knowledge and plentiful experience have encouraged me all the time in my academic research and daily life. Additionally, this endeavor would not have been possible without the generous support from our Head of the Computer

Department Dr. Nimisha Patel, who provided valuable feedback toward improving my work.

I would also like to extend my sincere gratitude to Dr. H N Shah Director of Gandhinagar Institute of Technology, Gandhinagar, for his very humble support, motivation, continuous encouragement for innovation with new ideas, and for providing a wonderfully positive environment for research work.

I am thankful for and fortunate enough to get constant encouragement, support, and guidance from all Faculty members of the Computer Engineering Department which helped me in successfully completing my thesis work. At last, I hereby take this opportunity to thank everyone who directly or indirectly helped me in preparing this thesis work.

DL Based Approach for Diagnosis of Retinal Diseases from OCT Images

Sanayya[a*], Parul Aggarwal[b], Mohd Abdul Ahad[c]

[a*]*Jamia Hamdard, Mehrauli-Badarpur Rd, Hamdard Nagar, New Delhi-110062, India*

Abstract

Artificial intelligence has the potential to change the way diseases are diagnosed, classified, and identified. Models are trained on hundreds of thousands of images to detect and classify various types of diseases. It can serve as an additional tool to aid medical professionals, radiologists in the detection and triaging process. Early triage and identification of retinal diseases can lead to early intervention and prevent any extreme outcome for the patients. This study proposes a deep-learning approach for detecting human eye diseases such as choroidal neovascularization (CNV), diabetic macular edoema (DME), and drusen from optical tomography scans. In our research, we have used various deep learning architectures ranging from custom CNN models to customized transfer learning models such as VGG16, ResNet50, DenseNet121. These models were able to identify retinal deformities in the human eye with great accuracy. Our approach makes use of retinal OCT images that have been pre-processed and processed by resizing, orienting, and colour corrections. In our experimental testing, the model performed exceptionally well, with the best model having a classification accuracy of 0.9845 as compared to manual ophthalmological diagnosis.

Keywords: Deep learning; Optical coherence tomography (OCT); Convolutional Neural Network (CNN); Transfer Learning.

1. Introduction

Detection and identification of retinal deformities or diseases requires a precise and detailed observation of the retinal layers. Ophthalmologists use specialized medical equipment for the observation of the human eye.AI and CNN based DL techniques can play a pivotal role in the early, efficient, and effective detection and treatment of retinal deformities allowing professionals to focus on the mitigation of the diseases. The technique necessitates precise, proper ocular layer identification and extraction, all of which can be achieved with great accuracy with a well-trained model. Our proposed framework makes use of CNN based models to classify and identify an ocular condition. The retina is the portion of the eye responsible for transmitting light and images to the brain. A human eye is said to have normal eyesight when light focuses on the retina. An eye with a normal eyesight can see objects from both close and long distances. Infection of the retinal layer can result in vision loss, myopia, and macular degeneration. Some of the most well-known retinal diseases are choroidal neovascularization (CNV), drusen, diabetic retinopathy, and diabetic macular edoema (DME).

Optical coherence tomography (OCT) can look at the structures of the retina in real time [1]. It is easy to use, doesn't use ionising radiation, and has a high resolution, so it is often used in diagnostic ophthalmology [2]. Every year, approximately 30 million OCT scans are performed, and the analysis and interpretation of these images takes a significant amount of time [3]. Using projected light beams, this method generates (tomographic) sectional images of the object under investigation with high resolution in the depth dimension. The thickness of the retinal layers can be measured to aid in disease detection. The two types of optical coherence tomography are known as time-domain (TD) and spectral domain (SD). A two-dimensional scan of an inner structure sample can be generated with the help of the TD-OCT. It is estimated that the SD-OCT method is fifty times more efficient than the more conventional TD-OCT method. In addition, SD-OCT has an ultra-high resolution and is one hundred times faster than OCT. When compared to TD-OCT equipment, the scan produced by an SD-OCT scanner is of superior precision and quality.

Figure 1 demonstrates the OCT scans of the normal retina, as well as CNV, Drusen, and DME. The formation of new blood vessels from the choroid into the sub-retinal pigment epithelium (sub-RPE) or sub-retinal space via a rupture in the bruch membrane is known as choroidal neovascularization (CNV). CNV is one of the main reasons why people lose their sight, while DME is a side effect of diabetes. DME can happen to people with either type 1 or type 2 diabetes. DME happens when the macula of the eye starts to get too full of fluid. The macula oversees our ability to focus and see fine details. It is inserted into the retina, the blood vessel-filled lining at the back of the eye. Excess fluid accumulation in the macula can cause vision deformities.On the

*Sanayya
E-mail address: sanayya1998@gmail.com

other hand, Drusen are yellow deposits that develop behind the retina. Drusen are composed of fatty acids and proteins. Drusen that are larger in size can lead to advanced AMD, thereby resulting in vision loss. When the retina is functioning normally, light is focused directly on it rather than in front or behind it leading to normal vision.

Figure 2 provides a visual representation of the proposed method for identifying ocular degeneration utilising retinal OCT images. The framework utilizes retinal OCT images that have been pre-processed and processed by resizing, orienting, and colour corrections. To recognise the four ocular disorders, the dataset is analysed with two distinct convolution neural network (CNN) models (three and seven layers), as well as three distinct transfer learning models (custom VGG16 model, custom ResNet50 model, and custom DenseNet121 model).The primary goal is to assist patients and ophthalmologists in the process of making an automated and timely diagnosis. It can also help in improving the analytical performance by increasing diagnostic accuracy and supporting ophthalmologists in becoming more skilled at making speedy and time-saving diagnoses, which will be of great benefit to patients.

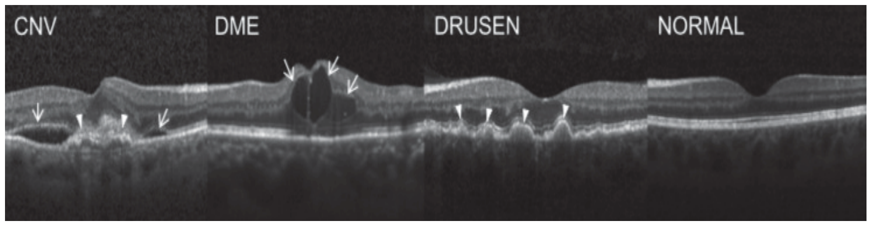


Fig. 1. Representative OCT images [4]

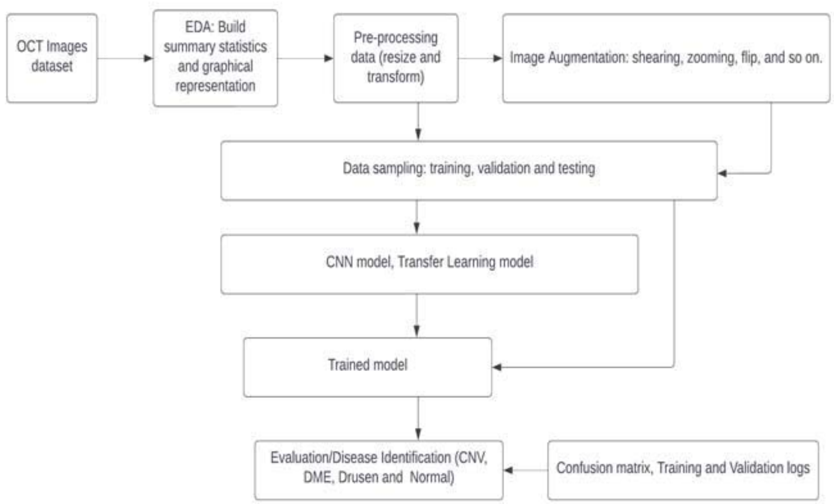


Fig. 2. Framework diagram for DL based approach for retinal disease detection.

2. Literature Review

Medical image analyses have made several breakthroughs thanks to deep learning algorithms. OCT scans of the human eye can be used to capture and assess a variety of ocular disorders, including DME, drusen, and CNV. The discovery of OCT in the 1990s marked a significant improvement in ocular imaging, as it is a non-invasive, radiation-free examination method capable of revealing the retinal layers with pinpoint accuracy.

According to Tălu et al. [5], OCT is a high-resolution imaging technology that is divided into two types: SD-OCT and TD-OCT. SD-OCT images provide a high-resolution cross-sectional and volumetric view of the retina. TD-OCT gives a two-dimensional picture of the retina's inner anatomy. OCT was found to be an excellent method for analysing, monitoring, and evaluating the various stages of AMD, according to the study. Moreover, drusen can be evaluated according to a range of structural features.

Karri et al. [6] used retinal OCT images and the inception network to present a transfer learning method for identifying retinal diseases. The dataset included OCT images of patients with dry AMD, DME, and healthy participants.The researchers discovered that, when compared to traditional learning approaches, the fine-tuned CNN was more effective at detecting anomalies. The accuracy of normal, AMD, and DME prediction was 99%, 89%, and 86%, respectively.

To detect dry AMD and DME with OCT images, Srinivasan et al. [7] described a classification strategy using support vector machine (SVM) classifiers and Histogram of Oriented Gradients (HOG) descriptors.The proposed technique did not segment the inner retinal layers. 45 volumetric pictures were included in the SD-OCT datasets, which were equally divided between the three classes- normal,AMD, and DME. This approach detects 100% of AMD cases, 100% of DME patients, and 86.67% of normal cases, giving it the best specificity and sensitivity.

Alsaih et al. [8] showed a way to automatically classify large amounts of SD-OCT imaging data to find DME. Their process was made up of four steps: pre-processing, detecting features, representing features, and classifying. The sensitivity and specificity for the SVM and principal component analysis were both 87.5%. The optimum classification of diseases required the use of LBP-ri vectors.

A DL-based method for identifying various retinal fluid types in the context of various macular disorders was put forth by Schlegl et al. [9] utilising OCT images.One thousand and two hundred OCT scans were included in their dataset, with patients' ages, genders, and prevalence of SMD, DME, and RVO (retina vein occlusion) all represented roughly proportionally. This fully automated approach was created to quantify and identify intra-retinal cystoid and sub-retinal fluid with a mean accuracy of 0.94, precision of 0.91, and recall value of 0.94.

Bhatt et al. [10] discussed the most popular DL models, their architectures, the benefits, and drawbacks of each, as well as potential applications for medical diagnosis and healthcare systems in their most recent work on CNN and image processing. Kim [11] created a strategy that combines a deep learning approach for classifying emotions from facial expressions with an image super-resolution technique. Nie et al. [12] presented a description of convolutional deep learning models for the retrieval of 3D objects. Zhao et al. [13] classified the various methods currently used for blood vessel segmentation into two main types: rule-based and machine-learning-based.In their study, Rajalingam et al. [14] presented a method for the merging of images that enables the MRI-CT-PET medical pictures to be viewed and evaluated with greater clarity. In a recent study by Xi et al. [15], the authors talk about how multiscale CNNs can be used to separate CNV from OCT data.

We discovered numerous substantial problems in the OCT images statistics during our approach. The influence of image augmentation over feeding raw images has not been examined by the researchers. A larger number of epochs has not been used to compile the model. We will apply much deeper architectures to help fill in these gaps.

3. Methodology

3.1 Dataset description:

We used publicly available dataset from Kaggle for the purpose of training and evaluation. The dataset consisted of several retinal OCT scans of four classes namely NORMAL, CNV, DME and DRUSEN. The images for each category were then divided into train, test and validation folders.

There are 84,495 JPEG OCT images and four classes (NORMAL, CNV, DME, DRUSEN). Each category's images are labelled as (disease)- (randomly assigned patient ID) - (image number by this patient).

3.2 Exploratory data analysis

EDA is one of the most important facets of this problem. To understand the data better, we built summary statistics for the dataset and constructed multiple graphical representations. The dataset is divided into three folders: train, test, and validation. Each folder has a subfolder for each type of image (NORMAL, CNV, DME, DRUSEN). We observed that the training dataset was slightly imbalanced as the classes DME and DRUSEN had a very small number of images.We also observed that images belonging to each class had different shapes meaning that each class of the OCT scans has some uniquely distinguishable pattern in their OCT scans.

3.3 Image augmentation

Image augmentation is a technique for creating fresh synthetic samples using basic image alteration techniques. As our dataset did not have many images, we supplemented this by generating augmented images from the existing images in the dataset. We applied various image transformation operations on the images in the dataset such as rotation, shearing, translation, zooming to create slightly altered versions of the images. The training set had 66788 images; validation dataset consisted of 16696 images whereas the test set contained 968 images. The images were then shuffled to ensure that there is no bias during training. These images were then inputted in batches of 128. Loading of images in batch becomes essential when there are many images and processing the whole dataset at once can lead to system memory getting full quickly.

Table 1. Related works.

References	Approach	Description	Performance
Feng Li et al. [6]	Ensembling Multi-ResNet 50	The majority of the improved ResNet50 is made up of convolutional, pooling, and fully connected layers.	accuracy is 97.9%, sensitivity is 96.8%, specificity is 99.4%, AUC is 0.998, and kappa is 0.969.
Kermany et al. [4]	Architecture for Inception V3	Each of the 11 inception modules that make up the network has a convolutional layer, an activation layer, a pooling layer, and a batch normalisation layer.	Accuracy is 96.6%, sensitivity is 97.8%, specificity is 97.3%, and AUC is 0.999.
Hussain et al. [17]	Random forest method	The test object's final class is decided via voting on sets of decision trees constructed using a subset of the training set.	Average accuracy is 96%, average sensitivity is 94%, average specificity is 85%, and mean AUC is 0.99.
Alsaih et al. [8]	SVM in linear	In order to classify the data, the algorithm generates either a line or a hyperplane.	Sensitivity and specificity are both 87.5%.
Srinivasan et al. [7]	SVM classifiers and HOG descriptors	This algorithm partitions the image into cells and measures the magnitude and orientation of spatial gradients within each cell to characterise the shape of nearby objects.	The accuracy of AMD, DME, and normal were each 100%, 100%, and 86.67% respectively.
Karri et al. [9]	Inception network	This method optimises a pre-trained CNN for prediction and recognises important responses during prediction to learn filter features.	Mean accuracy rates were 99%, 89%, and 86%.

We followed the below approach for data augmentation:

1. Reading files from various directories.
2. Resizing of images to 224 X 224 pixels.
3. Scaling the pixel values from 0-255 to the recommended range for neural network models i.e., 0-1.
4. Generating new but slightly modified versions of images in the training dataset by applying image transformation operations such as rotating, shearing, and zooming.
5. Conversion of images to a tensor data type to make them compatible with the model.

3.4 Deep learning models

In our approach, we compared the results of two custom built CNN-based model architectures and three different transfer learnings models on the dataset. CNN-based architecture aligned perfectly with our approach as our dataset consisted of many images. When it comes to computer vision challenges like image categorization, CNN models perform well and produce reliable results. The most significant advantage of transfer learning over training a model from scratch are resource savings and increased efficiency while training new models.

We used the following model architectures for our approach:

1. CNN Model with 3 Layers: This model had three layers, an input layer consisting of three channels withReLU (rectifier linear unit) activation function. Results from the first three CNN layers were fed into the max pool layer. We used a 3 x 3 kernel size. Then, we figured out the log-softmax probability and used it in the calculations that followed. We also used dropouts, and to avoid overfitting, we kept the probability value at 0.2.
2. CNN Model with 7 Layers: This model contained seven layers, including an input layer with three input channels that were all activated with ReLU (rectifier linear unit). The max pool layer received the output from the first, second, third, fourth, fifth, sixth, and seventh CNN layers. The kernel size utilized to create this image is 3 x 3. The log-softmax statistical likelihood was

calculated and utilized in subsequent calculations. To avoid overfitting, 0.2 probability dropout was utilized. Kernel size, dropout and the log-softmax probability were used in the same manner as in the previous 3-layer model.

3. VGG16 Transfer Learning [20]: VGG16 consists of 3x3 convolution layers with stride 1 and the same padding, as well as a 2x2 filter max pool layer with stride 2. Throughout the architecture, this configuration is maintained. It concluded with two fully connected layers and a softmax output layer.
4. ResNet50 Transfer Learning [21]: ResNet50 is a pre-trained convolutional neural network (CNN) that has been trained on a large dataset of images. Transfer learning involves taking a pre-trained network like ResNet50 and using it as a starting point to train a new model on a different dataset.With transfer learning, we can use the ResNet50 model as a feature extractor, where we remove the top layer(s) of the model and replace them with our own layers. This allows us to reuse the learned features from ResNet50 and train our own model to classify new images with less data and computing resources.
5. DenseNet121 Transfer Learning: The DenseNet (Dense Convolutional Network) architecture aims to make deep learning networks deeper while making them easier to train by making the connections between the layers shorter.Using DenseNet-121, we designed a feed-forward network in which each layer was directly connected to the next. It has four dense blocks, three transition levels, and 121 layers in total (117-conv, 3-transition, and1-classification). The original DenseNet study [22] explains that each convolution layer consists of (BN)-Relu-Conv processes. The Classification subnetwork has a 7x7 global average pooling, a 1000D fully connected layer, and softmax built in.

We used the following approach for model training:

1. To prevent any kind of bias, images are shuffled and then inputted with a set batch size and split ratio.
2. Outlining the model’s layers, activations, and input/output parameters.
3. Setting up the loss metric, optimizer, learning rate, and the number of epochs.
4. For every epoch:
   a) Using only a portion of the dataset for training.
   b) Setting up the optimizer, the input, and the labels.
   c) After receiving the image as input, the model will determine the training loss and use back-propagation to adjust the weights accordingly.
   d) After every kth iteration, the trained model is validated against the validation set. Afterwards, validation loss and accuracy is calculated.
5. Visualizing the outcomes.

4. Evaluation Metrics

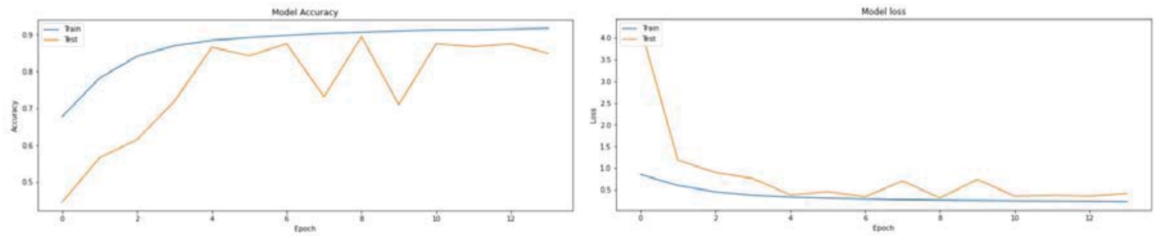
We used the following evaluation metrics to determine the accuracy of our CNN and Transfer Learning models in classifying retinal diseases:

- Accuracy: It measures how well the classifier can put the input into one of its categories. It is a measurement of how accurately a classifier can assign the correct label to an unseen input.

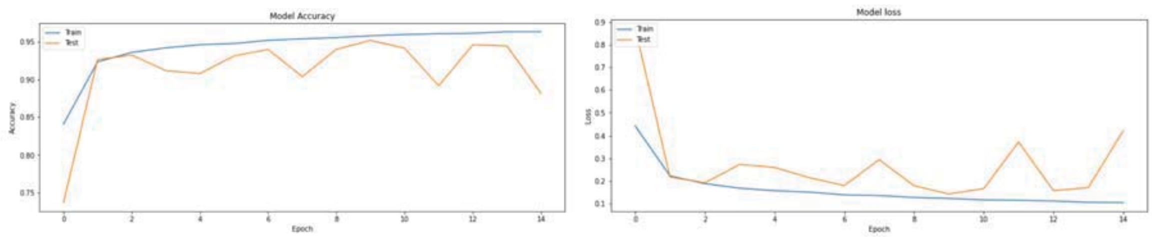
- Confusion Matrix: A confusion matrix is a table that is commonly used in machine learning and statistics to evaluate the performance of a classification model. The matrix provides a summary of the predictions made by the model by comparing them to the actual values. The confusion matrix is a square matrix that displays the number of true positive, false positive, true negative, and false negative predictions made by the model.True positive (TP) refers to the cases where the model predicted the positive class correctly. False positive (FP) refers to the cases where the model predicted the positive class incorrectly. True negative (TN) refers to the cases where the model predicted the negative class correctly. False negative (FN) refers to the cases where the model predicted the negative class incorrectly.

To accomplish the goals of this study, we trained CNN architectures of three and seven layers, each consisting of three and seven layers, as well as three distinct transfer learning models (custom VGG16 model, custom ResNet50 model, and custom DenseNet121 model).

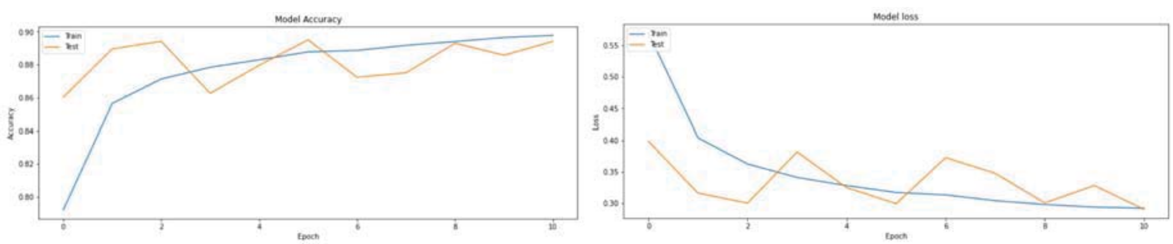
Figure 3 depicts the validation and training accuracy, and Figure 4 shows the confusion matrices for each model. The purpose of this comparison is to determine which of the models is the most suitable and effective for our dataset by contrasting the performance metrics and visualisation outputs of each of the models. Using these metrics and visualizations, we can determine the best performing model.



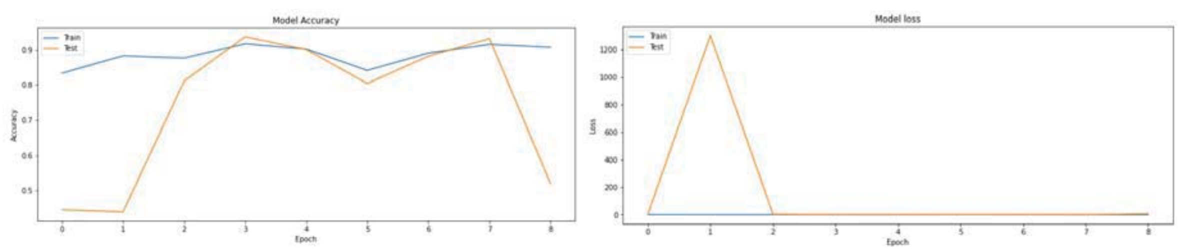
(a)



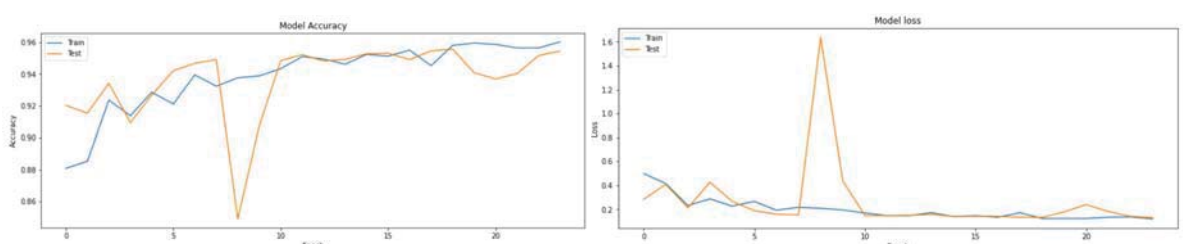
(b)



(c)

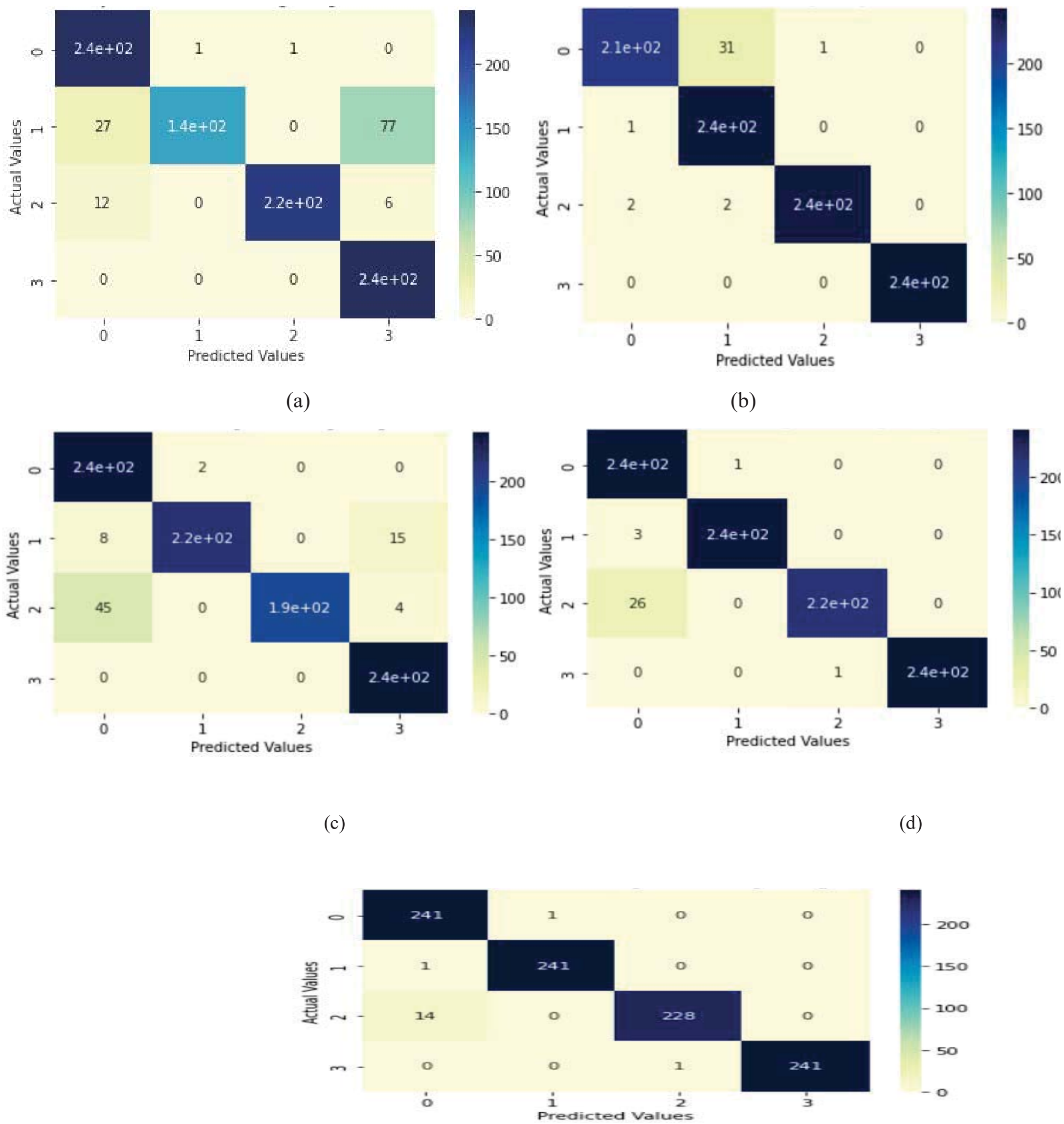


(d)



(e)

Fig. 3. Training and Validation Accuracy for (a) 3 Layer Model, (b) 7 Layer Model, (c) VGG16 Model, (d) ResNet50 Model and (e) DenseNet121 Model.



(a) (b) (c) (d) (e)

Fig. 4. Confusion matrices for (a) 3 Layer CNN Model, (b) 7 Layer CNN Model, (c) VGG16 transfer learning model, (d) ResNet50 transfer learning model and (e) DenseNet121 transfer learning model.

Result

To train the system, three distinct transfer learning models and two distinct (3- and 7-layer) convolutional neural network (CNN) models are used. Custom VGG16 models, custom ResNet50 models, and custom DenseNet121 models were used as transfer learning models. Figure 3 depicts the training and validation accuracy of each CNN model, while Figure 4 depicts the confusion matrix. We intend to find the best model for our dataset by comparing performance metrics and visualisations of all CNN layers. The confusion matrix, loss functions, and accuracy are three of the most important metrics for evaluating parameterised models.

- In case of 3-layer and 7-layer CNN models, total number of trainable parameters were 56,900 and 584,612 respectively whereas in case of VGG16, ResNet50, and DenseNet121 Transfer Learning model total number of trainable parameters were 6,458,324, 23,533,316 and 6,948,548 respectively.
- The three-layer CNN model had a loss of 0.2361 on the training set, 0.3140 on the validation set, and a test loss of 0.1821. We achieved an accuracy of 0.9442 on the test dataset using the 3-layer custom CNN Model.

- In terms of memory efficiency, the CNN model with seven layers is preferable. The seven-layer CNN models had an overall accuracy of 0.9638.
- On the other hand, custom VGG16 Transfer Learning Model has a test loss of 0.1956 and test accuracy of 0.9287.
- We observed that the ResNet50 Transfer Learning model fared quite well on the test dataset. Training the ResNet50 model with thousands of layers, we observed minimal change in the training error percentage. It gave an accuracy of 0.9618 which is comparable to the CNN model with 7 layers.
- Alternatively, the DenseNet121 Transfer Learning model had a minimal training loss of 0.1198, validation loss of 0.1327, and test loss of 0.1551. The Custom DenseNet121 Transfer Learning Model gave tremendous results, with an accuracy of 0.9845 on the test dataset.
- Both the CNN models took about the same amount of time to train, but the Transfer Learning models took about twice as long.

The findings of our investigation are presented in Figure 5 and Figure 6.

Model Name	Train Accuracy	Validation Accuracy	Test Accuracy
3CNN	0.9172	0.8944	0.9442
7CNN	0.9636	0.9521	0.9638
VGG16 Transfer Learning	0.8979	0.8951	0.9287
ResNet50 Transfer Learning	0.9171	0.9366	0.9618
DenseNet121 Transfer Learning	0.9479	0.9342	0.9845

Fig. 5. Comparative Chart.

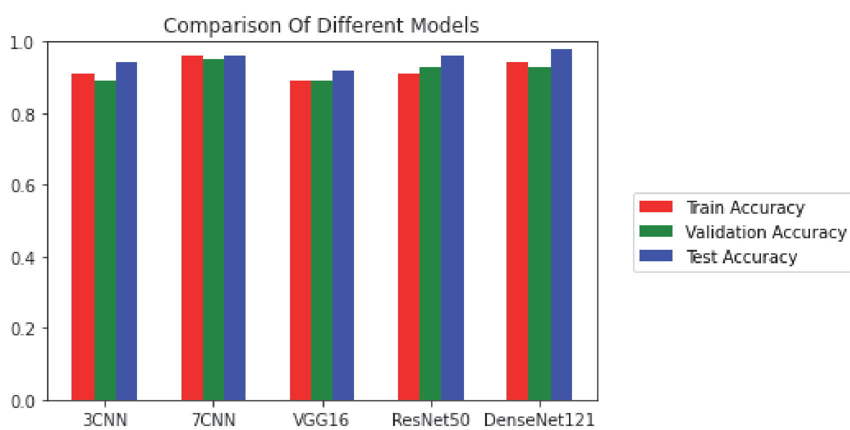


Fig. 6. Comparison of Different Models.

Conclusion and Future Scope

This research shows how deep learning algorithms may be used to efficiently classify and recognise ocular diseased characteristics in OCT images for CNV, Drusen, and DME. We used retinal OCT images that have been pre-processed and augmented by resizing, orienting, and colour corrections. We utilized two different CNN models (three and seven layers) and three different transfer learning methods (custom VGG16 model, custom ResNet50 model, and custom DenseNet121 model). All the models were able to differentiate and correctly classify an unseen OCT retinal scan as either CNV, DME, Drusen, or Normal with a high degree of accuracy. Additionally, the amount of time required for detection and epochs durations were nominal. We analysed and compared the accuracy, loss, and error metrics as well as the confusion matrices for each model. The Custom DenseNet121 Transfer Learning Model (with Image Augmentation) gives tremendous results, with an accuracy of 98% on the test dataset.

This framework can prove to be of great use to medical professionals and ophthalmologists as it could take off some of the burden from their shoulders and allow them to focus more on the treatment of the patients. This tool can certainly not replace trained professionals but can rather assist and aid in the diagnosis of retinal diseases. In areas where access to specialist doctors is scarce, patients can get a scan and General Physicians can submit those scans to the model for evaluation and based on those results, further treatment could be planned.

On the other hand, it cannot serve as the final diagnostic solution but should rather work in concert and tandem with trained professionals. In healthcare related models, a false negative can be extremely dangerous and there is still scope for further improvement in the model performance by utilizing deeper architectures with a diverse and vast dataset.

References

1. M. E. Velthoven, D. J. Faber, F. D. Verbraak, T. G. van Leeuwen, M. D. de Smet, 2007. Recent developments in optical coherence tomography for imaging the retina, Prog. Retinal Eye Res. 26(1), p. 57–77.
2. A. Lang, A. Carass, M. Hauser, E. S. Sotirchos, P. A. Calabresi, H. S. Ying, J. L. Prince, 2013. Retinal layer segmentation of macular OCT images using boundary classification, Biomed. Opt. Express 4(7), p. 1133–1152.
3. E. A. Swanson and J. G. Fujimoto, 2017. The ecosystem that powered the translation of OCT from fundamental approach to clinical and commercial impact," Biomed. Opt. Express 8(3), p. 1638–1664.
4. Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C., Liang, H., Baxter, S.L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M.K., Pei, J., Ting, M., Zhu, J., Li, C., Hewett, S., Dong, J., Ziyar, I., Shi, A., et al., 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. Cell 172(5), p. 1122-1131.
5. Ţălu, S.-D., Ţălu, Ş, 2012. Use of OCT imaging in the diagnosis and monitoring of age-related macular degeneration, age related macular degeneration—the recent advances in basic approach and clinical care, Gui-Shuang Ying. IntechOpen.
6. Karri, S.P., Chakraborty, D., Chatterjee, J., 2017. Transfer learning based classifcation of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. Biomed. Opt. Express 8(2), p. 579–592.
7. Srinivasan, P.P., Kim, L.A., Mettu, P.S., Cousins, S.W., Comer, G.M., Izatt, J.A., Farsiu, S., 2014. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. Biomed. Opt. Express 5(10), p. 3568–3577.
8. Alsaih, K., Lemaitre, G., Rastgoo, M., Massich, J., Sidibé, D., Meriaudeau, F., 2017. Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. Biomed Eng Online. 16(1), 68.
9. Schlegl, T., Waldstein, S.M., Bogunovic, H., Endstraßer, F., Sadeghipour, A., Philip, A.M., Podkowinski, D., Gerendas, B.S., Langs, G., Schmidt-Erfurth, U., 2018. Fully automated detection and quantifcation of macular fuid in OCT using deep learning. Ophthalmology 125(4), p. 549–558.
10. Bhatt, C., Kumar, I., Vijayakumar, V., et al., 2020. The state of the art of deep learning models in medical science and their challenges. Multimedia Syst.
11. Kim, P.W., 2020. Image super-resolution model using an improved deep learning-based facial expression analysis. Multimedia Syst.
12. Nie, W., Cao, Q., 2017. Convolutional deep learning for 3D object retrieval. Multimedia Syst. 23, p. 325–332.
13. Zhao, F., Chen, Y., Hou, Y., et al., 2019. Segmentation of blood vessels using rule-based and machine-learning-based methods: a review. Multimedia Syst. 25, p. 109–118.
14. Rajalingam, B., Al-Turjman, F., Santhoshkumar, R., et al., 2020. Intelligent multimodal medical image fusion with deep guided filtering. Multimedia Syst.
15. Xi, X., Meng, X., Yang, L., et al., 2019. Automated segmentation of choroidal neovascularization in optical coherence tomography images using multi-scale convolutional neural networks with structure prior. Multimedia Syst. 25, p. 95–102.
16. Yang, X., et al., 2016. Deep relative attributes. IEEE Trans. Multimedia 18(9), p. 1832–1842.
17. Hossain, M.S., Muhammad, G., Alamri, A., 2019. Smart healthcare monitoring: a voice pathology detection paradigm for smart cities. Multimedia Syst. 25(5), p. 565–575.
18. Hossain, M.S., Amin, S.U., Muhammad, G., Sulaiman, M., 2019. Applying deep learning for epilepsy seizure detection and brain mapping visualization. In: ACM Trans. Multimedia Comput. Commun. Appl. (ACM TOMM), vol. 15(1s).
19. Schmidt-Erfurth, U., Klimscha, S., Waldstein, S.M., Bogunović, H., 2017. A view of the current and future role of optical coherence tomography in the management of age-related macular degeneration. Eye (Lond.) 31(1), p. 26–44.
20. K. Simonyan and A. Zisserman, 2014. Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556.
21. K. He, X. Zhang, S. Ren, and J. Sun, 2016. Deep residual learning for image recognition," In Proceedings of the IEEE conference on computer vision and pattern recognition, p. 770–778.
22. Huang G, Liu Z, Van Der Maaten L, Weinberger, K. Q., 2017. Densely connected convolutional networks[C]// Proceedings of the IEEE conference on computer vision and pattern recognition, p.4700–4708.

A Review on Deep-Learning Approach to Copy-Move Forgery Detection and Source-Target Disambiguation

Krishna Alkeshbhai Brahmbhatt[a]*, Jahnavi Patel[b]

[a]*P.G Scholar, Information Technology Department, L. D. College of Engineering, 380015, India*
[b]*Asst. Professor, Information Technology Department, L. D. College of Engineering, 380015, India*

Abstract

The purpose of the research is to review available copy-move forgery detection approaches and to find an optimal Copy-Move forgery detection algorithm with accurate Source-Target Disambiguation which might also work for n-n Source-Target disambiguation and detection. Hence various approaches are studied out of which the two stage approach with backbone architecture based on VGG16 and Proposal Glue gives best performance for forgery detection and Source-Target disambiguation is performed best by Multibranch CNN including 4-TwinsNet which constitute of four 50-ResNet and Siamese Net. Further a more robust strategy for multi-target detection with Source-Target disambiguation that works on all cases even with the addition of the manipulation attacks like rotation, image blurring, JPEG compression etc.

Keywords: copy-move; forgery; image manipulation; deep learning;, image processing.

1. Introduction

In this digital era, the use of social media and hence use of image manipulation and editing tools have taken over the internet and world. To come across a digitally manipulated and forged picture has become such a common occurrence that 'seeing is believing 'adage does not hold anymore. Moreover, with such wide availability of tools and methods in the market for image forgery that it is very easy to forge an image without any identification through naked eye that the image is manipulated or any trace of being tampered. These tempered images bring a bucket load of problems along with them. Images are often considered trustworthy sources for information and investigation even in the cases of crime Investigation and news media. Various other fields that are affected by forged images are real estate, politics, insurance claims, public hysteria etc.

A digital image can be forged with various manipulations and Copy-move forgery is one such manipulation technique in which a section of image is copied and pasted in the same image widely; it may also include a section of photo being removed from image. In recent years copy move forgery has been one of most studied techniques of image manipulation and the reason for that is It is the easiest photo manipulation technique to perform but it is also one of the toughest techniques to prove that it has been performed, which results in copy move forgery being the most used image manipulation techniques. The purpose of copy-move forgery is generally to hide a particular part of the image using another section of same image or to add extra information to the image in order to contort the original meaning of the image. Hence the detection of copy-move forged image is an important research area in recent years in order to check the authenticity of the content we see online or in magazines which has the power to create a false public opinion or even mass hysteria or May also fuel unnecessary false rumor mill.

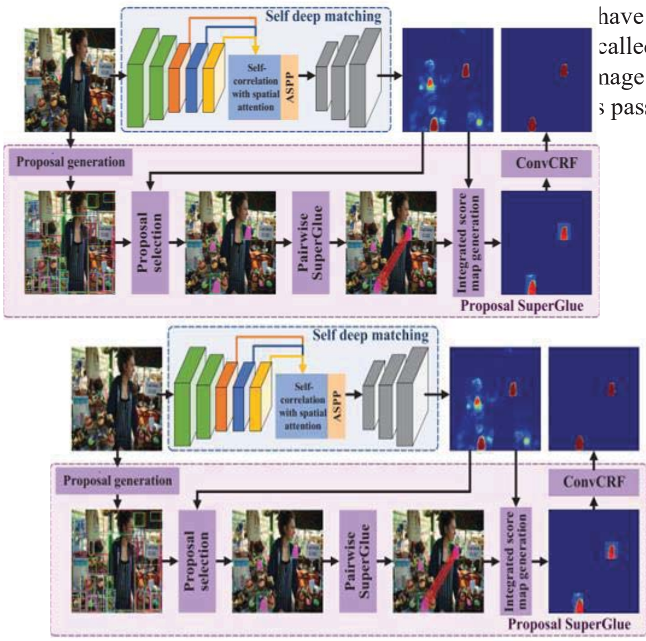
In the recent years many techniques have been researched for copy move forgery detection but a high accuracy detection approach with optimal source-target disambiguation is still a work in progress. Hence it has been of interest for researchers to dabble in. There are mainly two approaches to copy move forgery detection: keypoint based approach and block-based approach. In key point matching approach, the extracted keypoints are matched and in block-based approach the image is divided into blocks, and they are matched. In this paper some of the methods for copy move forgery detection are reviewed in order to find the most optimal approach.

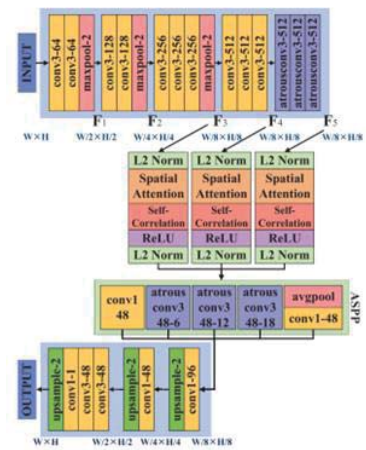
2. Literature Review

In [2] paper a two-stage Framework for copy-move forgery is used which includes end to end deep matching and customized key point matching algorithms are used in the pipeline. The copy move forgery detection is performed in two stages; the first stage includes a backbone deep matching network which gives backbone score maps that show the forged regions. The backbone network architecture includes feature extractor for which VGG16 is used but the pooling layer is removed from 4th and 5th convolution block and the 5th block is changed to atrous convolution: it is preferable for the resolution of convolution features as

* Krishna Alkeshbhai Brahmbhatt
E-mail address: krishnabrahmbhatt1503@gmail.com

well as field of view. This generates three large feature maps that are passed to skip matching. The feature maps have low level features with spatial and texture information and high-level features with semantic information both used to find visual similarities. These feature maps are passed to a self-correlation module with spatial attention to generate correlation maps. This process of generating correlation maps from feature maps is known as skip matching. Spatial attention and correlation module try to find correlation between every two features. Correlation maps are passed to Atrous spatial pyramid pooling (ASPP), It contains three parallel atrous convolution layers with 6, 12 and 18 atrous rates average pooling and convolution layer with 1*1 filter. ASPP generates 5 48-channel feature maps which are concatenated and fed to up sampling and convolution layers. The generated
have false alarm regions as well as in complete regions hence the first
called Proposal Superglue. Firstly, Proposal Generation is performed it
nage: bounding boxes are the suspicious region present in every corner
s passed as an input to Proposal Selection.





(a) (b)

Fig.1 (a) Proposed Methodology in the paper [2] (b) Architecture of Backbone Architecture [2]

Proposal selection takes output of proposal generation as well as the backbones Score map as input. Proposal selection rejects the proposals with less average scores, selects proposals with high IoU and merge proposals when intersection rates are high. Thresholds are fixed for this average threshold at 0.4, IoU threshold at 0.5 and intersection rates threshold at 0.8. Hence superglue provides high-quality proposals. High quality proposals are further passed to Super Point, It finds interest points and conducts key point matching. Points are matched through super glue using local features. Matching scores of the matching points pixel using a super pixel algorithm SEEDS. This pixel level score map and also integrated score map. The generated score maps are not refined, They have holes and false alarm regions, so in order to refine contour and remove regions with less probability using Convolution Conditional random field (CRF).

Further ablation study of backbone network is performed through step-by-step analysis till self-DM using a synthetic testing dataset with 120,000 images. Different state of the art methods are compared and also the given architecture in combination with 2 deeper networks; ResNet50 and ResNet101 and 3 light-weight networks: MobileNetV2, MobileNetV3, ShuffleNetV2 are tested out of these ResNet50 and MobileNetV3 give best performance. Hence, they were considered for further. They are compared using synthetic dataset. State of the art methods are compared using four data sets: Synthetic testing data set, CoMoFoD (with no attack), CASIA CMFD and MICC-F220.The F1 score for datasets is CASIA-0.7943(detected) and 0.5172(overall) CMFD-0.4782(pixel) and 0.7732(image-level) MICC-F220-0. 8559.The proposed method performs the best in almost all scenarios as shown in the figures.

In [3] methodology for source target disambiguation in copy move forgery detection is proposed. The given methodology works on the simplest scenario of single source single target copy move forgery. Moreover, it works best in the cases where the source and target regions are not overlapping each other and, in the cases, where a single region is returned instead of two different source target regions segmentation algorithms need to be used to distinguish between source and target. Generally, Copy-Move forgeries are performed through geometric transformation of source to target where a matrix represents relation between the two. Further interpolation processes are also performed in order to fit it into a 2D pixel grid resulting in non-invertible transformation. The copy-move forgery is often to followed with post processing such as landing of the borders of the target region in order to make it less obvious. Moreover, the post processing could also be applied globally so if it affects both the source and target. Facts that these post processing leaves are known as boundary artifacts. Proposed methodology exploits the non-invertible property due to interpolation and the boundary artifacts at the target regions for source target disambiguation.

The algorithm takes the localization mask of the forged image as input. The process perform source target disambiguation using interpolation artifacts works on the hypothesis formed that if the Copy move forgery process through geometric transformation is performed in the forward direction from source to target it gives the similar results to the image we already have at hand but on the other hand if the process is performed in the backward direction that is if the target is moved to source then the resulting image would be significantly different than the image we have at hand as there is no interpolation on the source. Hence it would be easy to identify which way is the forward way and which is the backward resulting in easily identifying source and target. For the implementation purpose we have used by bilinear kernel for interpolation. Sometimes only the interpolation artifacts are not enough in order to recognize source and target from each other. In such cases the second branch that is the branch that works on identifying the boundary artifacts is useful, moreover there is also a chance that the interpolation is very weak or not present at all due to either intense post processing process or rigid translation. The boundary artifact analysis works on the hypothesis that the target region would always have boundary artifacts in order to hide it or to make it less obvious, but source region would have no such arteffects hence the presence of this artifact is used in order for this ambiguation between source and target.

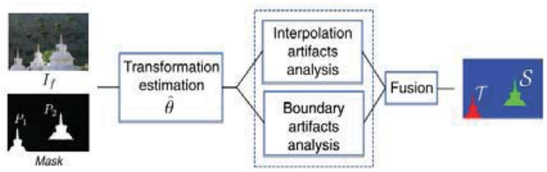


Fig.2 Scheme of the proposed CM disambiguation system. [3]

For the purpose of interpolation artifact analysis and boundary artifact analysis multibranch CNN networks are proposed: A parallel for branch module 4-Twins net which is used for interpolation artifact analysis and Siamese network for boundary artifact analysis. the output of both the networks are merged in a score level fusion module. Proposed architecture as a whole is referred to as DisTool. Architecture of the proposed 4 twins' network is shown in the figure. Before implementing the two networks, preliminary steps are carried out in order to identify the focus of attention (FOA), while the preliminary steps are similar for both the networks in the Siamese network the focus of attention is divided into four parts in order to get better results. Scores are generated based on a weighted sum the weights are assigned based on the reliability of both the network in specific scenarios. The four twins' network is more accurate in case of a more rigid translation while Siamese networks work better in the case of less rigid translation.

For evaluation purposes the four data sets are considered which includes a synthetically made dataset which further has three subsets in it which contains images with only rotation or only resizing or rigid translation. The methodology is also tested on USCISI, CASIA and GRIP data sets. The results were for Synthetic dataset SYN-T(28000 Images) for SYN-T-Rigid-97.00 accuracy, SYN-T-Rotation-97.90 accuracy and SYN-T-Resize-97.60 accuracy. for other already available datasets also results were comparatively better than previous approaches.

In [4] paper the proposed methodology is developed based on Busternet; busternet is a a dual branch parallel neural network with branches ManiDet and SimiDet -ManiDet is used in order to detect the manipulated regions and SimiDet is used to detect the similar copy move forged region. Hence with the help of results from both of them we can correctly disambiguate between source and target regions, but it has often been observed that the results from the internet are not much accurate because of the parallel branching structure hence in this paper a serial structure for the disambiguation is proposed.

The serial network can be divided into two sub parts: a Copy-Move forgery detection network (CMFDNet) and Source-Target Region Distinguishment Network (STRDNet). In the case of CMFD net the changes done to SimiDet for better performance as Simi date uses VGG16 and hence generates lower resolution features; include double level self-correlation and Atrous convolution to generate high resolution features. Also, the 4th pooling layer is removed to increase the resolution as each pooling layer divides the resolution by half. Also include double level self-correlation for finding relations between each two features and Atrous convolution to resolves the reduction in field of view filters due to pooling layer removal. The atrous convolution is only used for the fourth layer and not the rest in order to not increase the cost of GPU and balance in the batch size so as to keep training efficient. Self-Correlation module is followed by percentile pooling to remove irrelevant information from the score maps. The self-correlation module is followed by Mask Decoder module ASPP in order to take care of the multiscale features. In this module at the end a standard convolution and a softmax layer is used to find the final detection map which has both the source and target region.

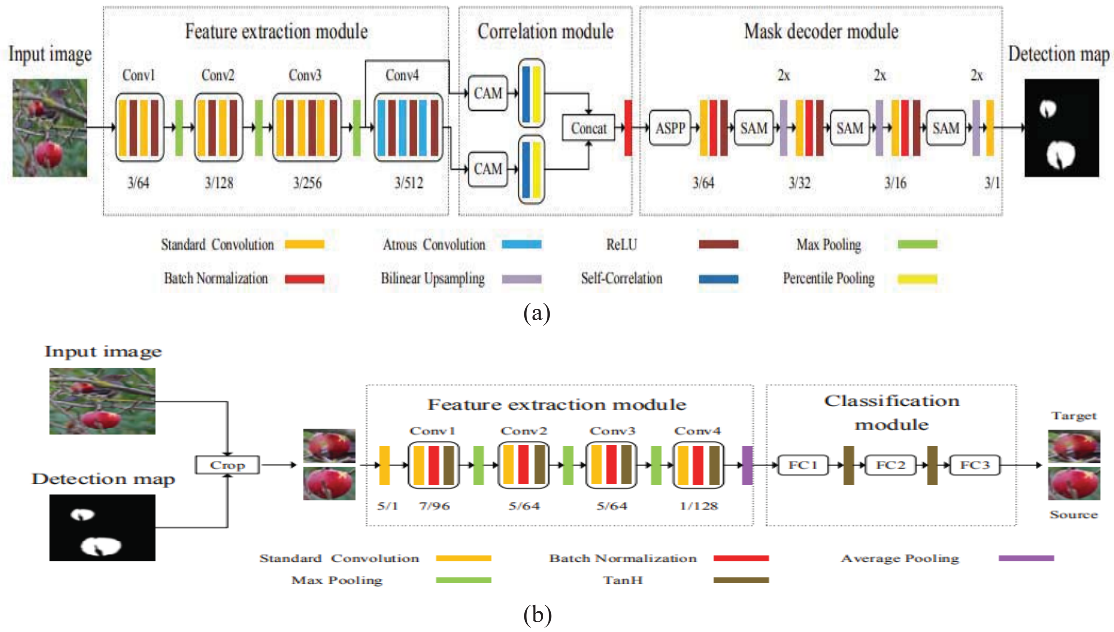


Fig.3. (a) Copy-Move Similarity Detection Network (b) Source-Target Distinguishment Network [4]

This detection map generated from the CMFD module Is used in order to crop the source and target from the actual input image. The way it is cropped is such that the smallest area containing the source and target and 15 pixels surrounding it. Cropped parts are passed to a feature extraction module containing for convolution groups first 3 followed by a max pooling and the last followed by an average pooling. Once the features are extracted there passed to a cost classification module having three fully connected layers with tanh activation function that classify is between tapered region and untampered region. Tampered region would be the target and the untampered would be the source.

For evaluation purposes a new synthetic data set (100100 images) was created and was used 8:1:1 for training testing and validation respectively. Evaluation was also done for CASIA v2.0(12614 images), CoMoFoD (200 images) and COVERAGE (200 images) data set for different state of the art methods. For data set are as follows CASIA v2.0-0.538, CoMoFoD-0.511 and COVERAGE-0.677 and for synthetic dataset-0.692 and for krocknecker convolution-0. 679.Evaluation in the case of the images under 6 well known attacks used to hide the forgery for different parameters was also considered. Comparison with the Busternet which was the base for the proposed methodology was also performed based on the correctly distinguished images and correctly classified images are synthetic data set(5394 images), CASIA v2.0(278 images), CoMoFoD(67 images) and COVERAGE(48 images).It was also tested for various attacks such as JPEG compression (JC), image blur (IB), Gaussian noise addition (GNA), color reduction (CR), contrast adjustments (CA), and brightness change (BC) with the proposed architecture performing better on all of the cases.

In [5] paper deep learning CNN model that is scale invariant to overcome. The inefficiency in performance due to scaling or rotation of manipulated objects. The method works with multi-scaled images this helps in feature extraction on multiple levels and hence gives an advantage of robustness in scaling. the architecture proposed in the paper can be divided into three parts: encoder phase, Decoder phase and classification phase. In encoder the input images of 256*256 dimensions are scaled multiple times up to 16*16 dimensions and all the 5 levels are taken as input. These images Convolution layer batch normalization and ReLu and down sampled using max pooling but max pooling is not good for segmentation and hence these max pooling is concatenated with activated feature space, continued till lowest scale input dimension. Convolution layer has padding same and stride 1. For visualization and segmentation corresponding to each max pooling up sampling is done. It is followed by convolution layer batch normalization and ReLu.The activated feature space is concatenated with the out of first up sampled layer. classified requires single Depth features space for training hence a 1*1 convolution with stride one and padding same Followed by sigmoid activation function So the image pixels are divided into two parts black and white for authentic and forged pixels. to evaluate the classifiers prediction a loss function is also used lower the value better the prediction; binary cross entropy loss function is used.

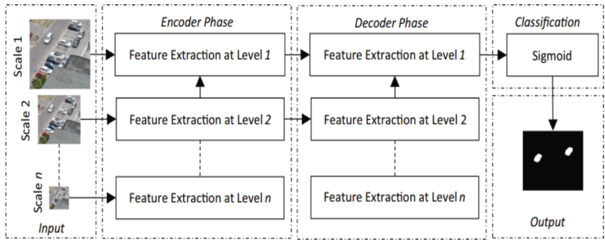


Fig.4. Architecture of proposed model for copy-move forgery detection using deep learning CNN model [5]

Evaluation purposes two data sets:CoMoFoD and CMFD are considered: for 3 kernel size, 6 post processing attacks With various parameters and comparison with various state of art methods. The ratio for training validation and testing is 70,20,10 respectively.

Table 1. Average test result using performance measures precision, recall, accuracy, TNR, FNR, F1-score and MCC value on different datasets [5]

Dataset	P	R	A	TNR	FNR	F1	MCC
CMFD	0.9892	0.9982	0.9878	0.7764	0.0018	0.9936	0.8329
CoMOoFoD	0.9863	0.9962	0.9839	0.8247	0.0038	0.9909	0.8578

Table 2. Average test result using performance measures precision, recall, accuracy, TNR, FNR, F1-score and MCC value on CoMoFoD dataset on different post-processing operations [5]

These abbreviations stand for—F: Only translation without post-processing, BC1-BC3: Brightness change, CA1-CA3: Contrast Adjustment, CR1-CR3: Color Reduction, IB1-IB3: Image blurring, JC1-JC9: JPEG Compression and NA1-NA3: Noise addition

Post-processing	P	R	A	TNR	FNR	F1	MCC
F	0.9863	0.9961	0.9838	0.8215	0.0039	0.9909	0.8558
BC1	0.9862	0.9961	0.9837	0.8186	0.0039	0.9908	0.8537
BC2	0.9860	0.9961	0.9835	0.8152	0.0039	0.9907	0.8510
BC3	0.9852	0.9961	0.9827	0.8029	0.0039	0.9903	0.8419
CA1	0.9864	0.9961	0.9840	0.8261	0.0039	0.9910	0.8589
CA2	0.9865	0.9962	0.9841	0.8293	0.0038	0.9910	0.8613
CA3	0.9858	0.9963	0.9835	0.8279	0.0037	0.9907	0.8607
CR1	0.9863	0.9961	0.9838	0.8213	0.0039	0.9909	0.8556
CR2	0.9863	0.9961	0.9838	0.8212	0.0039	0.9909	0.8556
CR3	0.9863	0.9961	0.9838	0.8214	0.0039	0.9909	0.8556
IB1	0.9861	0.9964	0.9838	0.8221	0.0036	0.9909	0.8575
IB2	0.9862	0.9964	0.9840	0.8217	0.0036	0.9910	0.8579
IB3	0.9862	0.9964	0.9841	0.8195	0.0036	0.9910	0.8560
JC1	0.9862	0.9962	0.9838	0.8277	0.0038	0.9909	0.8595
JC2	0.9860	0.9962	0.9835	0.8268	0.0038	0.9907	0.8573
JC3	0.9864	0.9963	0.9841	0.8287	0.0037	0.9910	0.8610
JC4	0.9867	0.9962	0.9843	0.8300	0.0038	0.9911	0.8611
JC5	0.9865	0.9961	0.9840	0.8298	0.0039	0.9910	0.8604
JC6	0.9865	0.9962	0.9841	0.8281	0.0038	0.9910	0.8602
JC7	0.9863	0.9961	0.9837	0.8261	0.0039	0.9909	0.8573
JC8	0.9863	0.9961	0.9837	0.8245	0.0039	0.9909	0.8570
JC9	0.9871	0.9961	0.9846	0.8456	0.0039	0.9913	0.8710
NA1	0.9869	0.9962	0.9846	0.8269	0.0038	0.9913	0.8600
NA2	0.9868	0.9961	0.9843	0.8278	0.0039	0.9911	0.8598
NA3	0.9868	0.9961	0.9843	0.8273	0.0039	0.9912	0.8593

In [6] paper a key point matching-based algorithm is used for copy move forgery detection. The proposed methodology in the paper consists of four steps: preprocessing, key point calculation, key point matching and morphological processing and LCS.Proposed methodology focuses on features from textured regions as well as a smooth region so in order to perform it properly different feature extraction is used for both scenarios. So different preprocessing needs to be performed for both approaches. For preprocessing of textured areas include removing unsharp edges and noise. The RGB image is converted to grayscale image. Moreover, unsharp masking technique is used to remove blurring and get a sharp image. For pre-processing of smooth regions, the wiener filter is used for adaptive noise removal and also mention that blurring is used for removing noise. Local neighborhood is used for local mean and variance for pixel.

For feature extraction phase, for textured region feature extraction Features from accelerated segment test (FAST) and Binary robust independent elementary features (BRIEF) descriptors are used and for smooth region feature extraction SIFT descriptors are used. In case of FAST Bresson Ham circle of radius 3 with 16 pixels along its circumference is used to find the neighbourhood and pixels below above right and left of this neighbourhood are checked. If 3 pixels out of this 4 are darker than it is checked further if 12 pixels are darker out of the 16 It is considered. Then BRIEF is used for binary feature descriptors finding.256 dimensions descriptors are generated after smoothing. For SIFT different smoothing version and difference of gaussian (DOG), also taylor series on scale space is applied for transformation invariance.

For feature matching a generalization of Lowe's matching technique named g2NN is used.it can also detect multiple copies. It is ratio of di and di+1 where d stands for Euclidean distance. A threshold is decided for ratio for matching. If the matched points are not too close to each other than morphological processing is applied. If normalized area is 40 percent of max the considered otherwise, they are considered outliers. Further localization is improved. Noise is removed through median filtering. super pixel segmentation is performed through linear spectral clustering (LSC). Segmentation is considered meaningful only if SSIM is higher than 0.6.Evaluation is performed on MICC-F220 with CPU time.

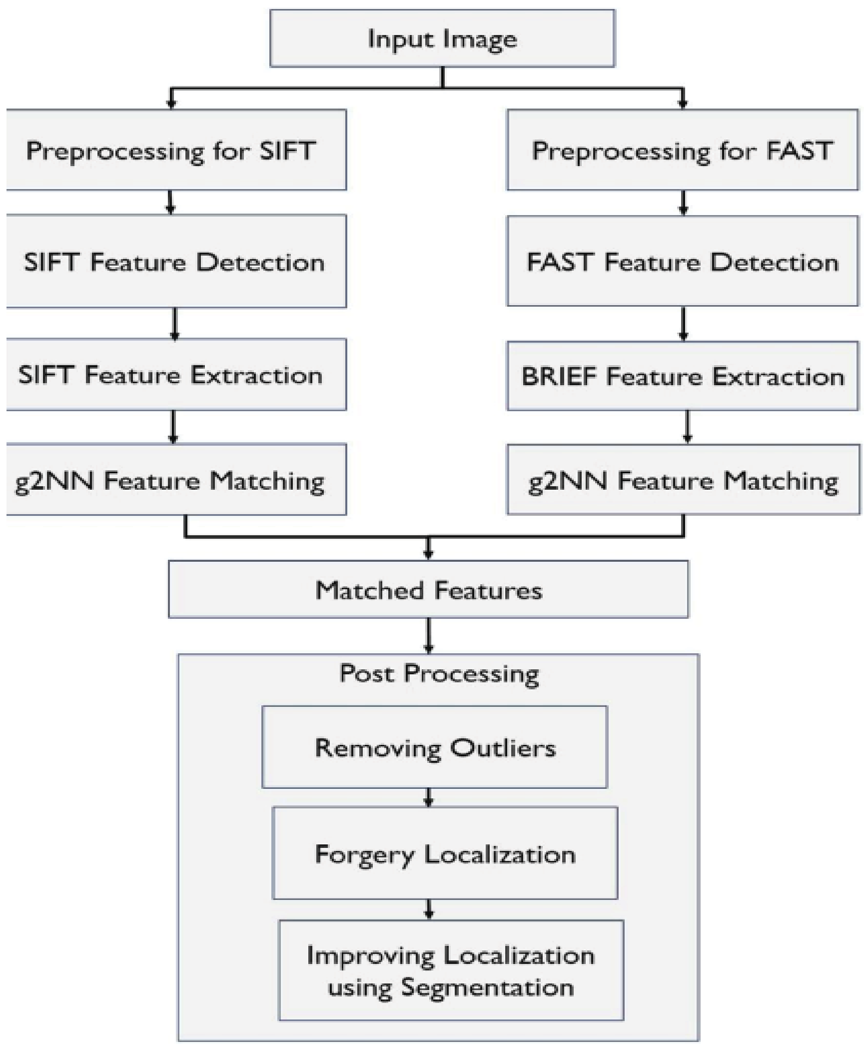


Fig 5. (a) Block diagram of proposed system [6]

Table 3. Performance measurement for 6 selected images [6]

Images	Pun et al. [18]	Ryu et al. [23]	Cao et al. [6]	Proposed
F-Measure				
Image 1	0.5052	0.1153	0.077	0.9626
Image 2	0.8244	0.9143	0.3767	0.9663
Image 3	0.6414	0.0044	0.0326	0.8567
Image 4	0.3241	0.0604	0.0444	0.725
Image 5	0.8317	0.0051	0.0397	0.9508
Image 6	0.6104	0.2646	0.0372	0.94864
Precision				
Image 1	1	0.0612	0.04	0.9476
Image 2	0.9966	0.8428	0.2325	0.9837
Image 3	1	0.0024	0.0173	0.9967
Image 4	1	0.082	0.0281	0.9853
Image 5	0.998	0.0026	0.0.0452	0.9164
Image 6	1	0.2239	0.0462	0.9023
Recall				
Image 1	0.3379	1	1	0.978
Image 2	0.7029	0.9991	0.9927	0.9495
Image 3	0.4721	0.024	0.2821	0.7513
Image 4	0.1934	0.0478	0.1067	0.5735
Image 5	0.7129	0.1582	0.0354	0.9878
Image 6	0.4393	0.3234	0.0325	1
CPU-time (in seconds)				
Image 1	30.335	33.339	97.148	5.622
Image 2	19.300	30.506	46.794	6.367
Image 3	15.613	33.590	63.938	8.4039
Image 4	16.336	30.845	35.470	7.236
Image 5	73.439	33.335	50.346	10.345
Image 6	61.478	44.672	79.345	11.457

3. Summarization

Sr no.	Title of paper	Publication details	Proposed Method	Tools and technology	Datasets	Research possibility
1. [2]	Two-Stage Copy-Move Forgery Detection With Self Deep Matching and Proposal SuperGlue .	Published in IEEE Transactions on Image Processing, Volume 31, 2022	A two stage approach with backbone architecture with deep matching and superglue with convCRF.	VGG16,Atrous convolution, Spatial attention, Self correlation, split matching, SuperPoint, superglue, convCRF	MICC-F220 CoMoFoD CASIA SYNTHETIC	The backbone architecture can be worked on and further multi target CMFD.
2. [3]	Copy Move Source-Target Disambiguation Through Multi-Branch CNNs	Published on IEEE Transactions on Information Forensics and Security, Volume 16,2022,	Geometric transformation followed by 4-twinsNet and Siamese network.	Geometric transformation, 4-Twins Net, Siamese Net, FOA	USCISI Grip CASIA SYN-Ts	Neuro- fuzzy network with back propagation can be used, multi-target detection, robustness by forger perspective.

3. [4]	A serial image copy-move forgery localization scheme with source/target distinguishment	Published on IEEE Transactions on Multimedia, Volume 23, 2021	A BusterNet with Atrous convolution and correlation and mask decoder with classification module.	BusterNet, Atrous Convolution, Mask Decoder, Classification module, Self Correlation.	Synthetic CASIA V2.0 CoMoFoD COVERAGE	Better source target disambiguation and robustness checking using various attacks.
4. [5]	Detection of Copy-Move Forgery in Digital Image Using Multi-scale, Multi-stage Deep Learning Model	Published on Springer in Neural Processing Letters, Volume 54, 2022	A encoder module with a decoder module with classification with sigmoid function.	Encoder, Decoder, sigmoid function	CMFD CoMoFoD	a segmentation using dictionary learning can be used with the proposed deep learning technique for a better result.
5. [4]	CNN-Transformer Based Generative Adversarial Network for Copy-Move Source/Target Distinguishment	CNN-Transformer Based Generative Adversarial Network for Copy-Move Source/Target Distinguishment	A generator and discriminator based deep learning approach for source target disambiguation	Generator, Discriminator, transformer	USCISI, CoMoFoD, CASIA2	The GAN can be improved and is supported by multi-branch CNN.
6. [5]	Disentangling copy-moved source and target areas	Published on Applied Soft Computing Volume 109, 107536, Elsevier on September 2021	An image component statistical deviation using GMM and log likelihood and histogram.	GMM, Empirical histogram, likelihood	CASIA2	without statistical modeling GMM but directly with the pixel values more advanced distance measures such as Wasserstein or MMD (Maximum Mean Discrepancy
7. [6]	FAST, BRIEF and SIFT based image copy-move forgery detection technique	Published on Springer in Multimedia Tools and Applications , 2022	SURF and FAST with BRIEF with pre-processing with G2nn matching and post processing.	SURF, FAST, BRIEF, G2NN , LCS	Synthetic	Multi-target detection and Source-Target disambiguation can also be performed.
8. [7]	A novel copy-move forgery detection algorithm via two-stage filtering	Digital Signal Processing, Elsevier, Volume 113, June 2021, 103032	A key point detection algorithm followed by key point Matching; Delaunay triangulation algorithm and 2 stage filtering scheme Grid-Based Filter and the Clustering-Based Filter	Delaunay triangulation, Grid-Based Filter and the Clustering-Based Filter	IMD CMHD CoMoFoD	Improving clustering based filter by relaxing angular constraint of group building and merging smaller groups nearer to each other.

9. [8]	AR-Net: Adaptive Attention and Residual Refinement Network for Copy-Move Forgery Detection	IEEE Transactions on Industrial Informatics, Volume 16, issue 10, 2020	AR-NET with channel and position attention with Adaptive Attention mechanism and ASPP and Residual Refinement module.	AR-NET, ASPP, Residual Refinement module, Adaptive attention mechanism	CASIAII	A multi-modes architecture for more information as opposed to single stream architecture of now
10. [9]	Copy Move Forgery Detection based on double matching	Journal of Visual Communicatio n and Image Representation , Elsevier, Volume 76, April 2021, 103057	Key-point detection with double matching first LIOP generated key points with Delaney triangle second adding triangles iteratively and Localization by DBSCAN.	LIOP, Atrous Convolution, Delauney triangulation, DBSCAN	MICC-F220	DBSCAN can be updated atrous convolution and self correlation for better feature generation.
11. [10]	Single and Multiple Copy–Move Forgery Detection and Localization in Digital Images Based on the Sparsely EncodedDistincti ve Features	Arabian Journal for Science and Engineering, Springer, Volume 45,April 2020	Uses SURF with BRISK descriptors and matching with 2NN and hamming distance followed by DBSCAN and noise clustering.	SURF, BRISK, 2NN, hamming distance, DBSCAN	MICC-F220 MICC-F2000 CoMoFoD	The technique needs to increase robustness from available attacks.
12. [11]	Copy-Move Forgery Detection Exploiting Statistical Image Features	Digital Image Forensics Journal, Springer, chapter 4, 2019	Using Mean and variance similarities between blocks made by Discrete wavelet transform.	Statistical transformatio n, DWT, Euclidean distance	USCISI + CoMoFoD SELF FORGED for testing	For block based methods a combination of DCT and FFT is used and can be combined with

4. Conclusion

In this paper recent advances in the field of copy move forgery have been studied in order to get an idea to find an optimal approach for forgery detection as well as source target disambiguation which is the process of recognising source from target.It was observed that backbone architecture with proposal superglue gives the best performance for detection of copy move forgery while a multi-branch CNN approach gives best results for source target disambiguation.The use of image processing based method is a step towards a accurate detection of manipulated images and helps in detecting forgeries that are not easily detected by naked human eye.

References

1. “COPY-MOVE-FORGERY-EXAMPLE.JPG”.
2. Y. Liu, C. Xia, X. Zhu and S. Xu, "Two-Stage Copy-Move Forgery Detection With Self Deep Matching and Proposal SuperGlue," in IEEE Transactions on Image Processing, vol. 31, pp. 541-555, 2022, doi: 10.1109/TIP.2021.3132828.
3. M. Barni, Q. -T. Phan and B. Tondi, "Copy Move Source-Target Disambiguation Through Multi-Branch CNNs," in IEEE Transactions on Information Forensics and Security, vol. 16, pp. 1825-1840, 2021, doi:10.1109/TIFS.2020.3045903.
4. B. Chen, W. Tan, G. Coatrieux, Y. Zheng and Y. -Q. Shi, "A Serial Image Copy-Move Forgery Localization Scheme With Source/Target Distinguishment," in IEEE Transactions on Multimedia, vol. 23, pp. 3506-3517, 2021, doi:10.1109/TMM.2020.3026868.
5. Jaiswal, A.K., Srivastava, R. Detection of Copy-Move Forgery in Digital Image Using Multi-scale, Multi-stage Deep

Learning Model. Neural Process Lett 54, 75–100 (2022). https://doi.org/10.1007/s11063-021-10620-9

6. Fatima, B., Ghafoor, A., Ali, S.S. et al. FAST, BRIEF and SIFT based image copy-move forgery detection technique. Multimed Tools Appl (2022). https://doi.org/10.1007/s11042-022-12915-y
7. Jixiang Yang, Zhiyao Liang, Yanfen Gan, Junliu Zhong,A novel copy-move forgery detection algorithm via two-stage filtering,Digital Signal Processing,Volume 113,2021,103032,ISSN 1051 2004, https://doi.org/10.1016/j.dsp.2021.103032.
8. Y. Zhu, C. Chen, G. Yan, Y. Guo and Y. Dong, "AR-Net: Adaptive Attention and Residual Refinement Network for Copy-Move Forgery Detection," in IEEE Transactions on Industrial Informatics, vol. 16, no. 10, pp. 6714-6723, Oct. 2020, doi: 10.1109/TII.2020.2982705.
9. Qiyue Lyu, Junwei Luo, Ke Liu, Xiaolin Yin, Jiarui Liu, Wei Lu,Copy Move Forgery Detection based on double matching,Journal of Visual Communication and Image Representation,Volume 76,2021,103057,ISSN 1047-3203, https:// doi.org/10.1016/j.jvcir.2021.103057
10.]Bilal, M., Habib, H.A., Mehmood, Z. et al. Single and Multiple Copy–Move Forgery Detection and Localization in Digital Images Based on the Sparsely Encoded Distinctive Features and DBSCAN Clustering. Arab J Sci Eng 45, 2975–2992 (2020). https://doi.org/10.1007/s13369-019-04238-2
11. Roy, A., Dixit, R., Naskar, R., Chakraborty, R.S. (2020). Copy-Move Forgery Detection Exploiting Statistical Image Features. In: Digital Image Forensics. Studies in Computational Intelligence, vol 755. Springer, Singapore. https:// doi.org/10.1007/978-981-10-7644-2_4
12. Y. Zhang *et al.*, "CNN-Transformer Based Generative Adversarial Network for Copy-Move Source/Target Distinguishment," in *IEEE Transactions on Circuits and Systems for Video Technology*, 2022, doi: 10.1109/TCSVT.2022.3220630.
13. Ludovic Darmet, Kai Wang, François Cayre,Disentangling copy-moved source and target areas,Applied Soft Computing,Volume 109,2021, 107536,ISSN 1568-4946, https://doi.org/10.1016/j.asoc.2021.107536.

Automated detection of Alzheimer's using Transfer learning models based on the Motion sensor Data

Junaid Gouher[a]*, Sushmita Upadhyay[b], Sanjaya Shankar Tripathy[c], Priyank Saxena[d]

[abcd]*Department of Electronics and Communication Engineering,*
Birla Institute of Technology, Mesra, Ranchi, Jharkhand, 835215, India

Abstract

Alzheimer's is an irremediable brain artifact affecting older people. The disease can't be cured and gradually impacts memory, behavior, thinking, and social skills. Timely disease detection can help control its progression and provide proper patient care. The non-invasive detection of the disease can be performed by imaging, camera, and sensor-based techniques. The imaging-based approach is costly, time-consuming, and requires expert supervision, while the camera-based approach's accuracy depends on the camera's correct position. The proposed work uses the sensory data approach by detecting the Freezing of Gait (FoG) event using motion sensors attached to the patient's body. The proposed work compares the classification accuracy of the Artificial Neural Network (ANN) and Convolution Neural Network (CNN) models based on the sensor data. The models are trained on the Daphnet dataset, and both models detect the disease identically.

Keywords: Alzheimer's disease; ANN; CNN; Daphnet dataset; Motion sensor.

1. Introduction

Alzheimer's disease (AD) is a massive cause of dementia and is predicted to impact more than 75 million individuals worldwide by 2030. The past decade has witnessed a rise in mental health problems, and there is a surge due to the recent covid-19 pandemic. Additionally, it ranks as the fourth leading cause of death [1]. AD is a chronic neurodegenerative disease leading to the decay of tissues and nerve cells throughout the brain, impairing cognition and memory in the patient. Amyloid protein, which is highly deposited in brain cells and obstructs signal propagation by increasing brain cell death rate, is the primary cause of AD [2]. There are three phases of AD progression: early, modest, and advanced cognitive levels. Patients in the subtle stage of cognitive development react violently, while those in their late stages possibly lead to fatal issues [3]. Based on the data available on AD, it is observed that 90% of patients having AD are diagnosed only in the later stages, reducing the chances of disease control and benefit from traditional treatment.

The patients diagnosed in the early stage can be benefited from the different assistance measures like corrective, reinforcing, and supportive helping them in social cognition and cognitive stimulation [4]. The remedial assistance guides the patients to correct their actions, while the reinforcement assistance can help delay the disease's effect. The supportive assistance helps the patient in their daily life activities. Therefore, timely disease detection is crucial for patient care and has attracted much attention to non-invasive detection techniques. The non-invasive detection identifies the brain disorder or the Freezing of Gait (FoG) events. Firstly, the data extraction is performed by any of the three types, i.e., imaging, camera, and sensor-based techniques, followed by the prediction. The prediction can be automatic or manual. Manual approaches are limited due to being time-consuming and requiring expertise.

Much research has been contributed to imaging-based approaches because of the availability of online datasets [1,4-11]. The imaging approaches use Magnetic Resonance Imaging (MRI) or Positron Emission Tomography (PET) scans. The application of Functional and Structural MRI techniques is also reported in the literature for AD detection. FMRI uses cerebral blood flow to diagnose AD to image brain activity, while structural MRI uses structural integrity of the affected brain regions [11]. The major disadvantage of these imaging approaches is that they must be done under expert supervision, and the readings must be understood well to make the prediction efficiently. Moreover, scanning is time-consuming and expensive.

The camera-based approaches have also contributed significantly to the literature [12-15]. The different types of cameras, i.e., Kinect, RGB, depth cameras, etc., have been used in these approaches. However, the major issue with these approaches is placing the camera in the correct position to get a clear picture/video. Also, such cameras with sensing devices are pretty expensive and ultimately increase the cost of the product.

The sensor-based approaches have gained widespread popularity in recent times. Many Internet of Things (IoT) based wearable products, i.e., wristbands, are available for monitoring and recording different health-related parameters. The sensor-based

* Junaid Gouher
E-mail address: mtec10010.21@bitmesra.ac.in

approaches detect the Freezing of Gait (FoG) events [16-19]. The FoG is a motor activity in ambulation controlled by the brain. The patients affected by AD often feel like their feet are stuck to the floor during the walk.

Sensor-based approaches can be differentiated based on the positioning of the sensors. Either the sensor can be attached to the body or installed in the patient pathway. In [16-18], the authors installed sensors along the patient's pathway and recorded the pressure the patient exerts on the floor and the distance between the two legs, which determined the detection of an FoG event. In [3,19], wearable sensors were attached to the body to predict FoG events.

The data collected using the abovementioned approaches is further analyzed for prediction purposes. Automatic prediction employing machine learning (ML) and deep learning (DL) classifiers has demonstrated exemplary performance in disease prediction [20-23]. Chitra Devi et al.'s [20] focus was on optimizing brain subregion analysis for Alzheimer's disease detection. The research showed that the hippocampus biomarker is essential for analyzing the disease. In [21], the authors used a fuzzy neural classifier to classify diseases using IoT and cloud-based prediction mechanisms.

A machine learning-based technique to distinguish between Alzheimer's and mild cognitive patients for healthy elderly cases was suggested by Long et al. [22]. In [12], the trial on twenty-four women's skeletal data was analyzed having equally healthy and affected women. According to the findings, feature extraction using support vector machines produced good categorization accuracy. A multi-template learning method was put forth by Liu et al. [23] for the automated diagnosis and staging of Alzheimer's disease. A feature selection algorithm models the connection between the template and individuals.

Thus, an inference can be drawn from the above that various machine-learning model for the detection of Alzheimer's disease has been approved for the detection of Alzheimer's disease. Still, none of the approaches focused on finding the best model for detecting the disease. The proposed work focuses on finding the best model for detection.

The contribution of the work is as follows:

- The proposed work uses wearable motion sensors (accelerometers) placed at the body's trunk, thigh, and ankle areas for data collection.
- The proposed work compares the performance of two deep learning networks, ANN and CNN, trained on the daphnet dataset.
- The dataset is utilized at 80% for training, 10% for validation, and 10% for testing.
- Both CNN and ANN perform equally well in disease prediction. ANN achieves an accuracy of 89.6 %, while CNN achieves an accuracy of 89.39%.

The subsequent work is divided as follows: the materials and methods used are discussed in section 2. Section 3 discusses the experimental results and discussion, followed by Section 4. The conclusion of this work is explained in Section 5.

2. Materials and Method

2.1 Data Set

The proposed work makes use of the Daphnet Dataset [18], which includes three different types of tasks: walking in a straight line, walking while making many turns, and lastly, a more realistic everyday activity where users enter various rooms while getting coffee, opening doors, etc. The distribution of data for training testing and validation is shown in Figure 1.

2.2 Data Preprocessing

The Daphnet dataset has three outputs: 0, 1, and 2; while 1 and 2 correspond to the no freezing and freezing of gait, 0 corresponds to the results where the subject performed tasks unrelated to the experiment. Hence these have to be removed from the dataset and were removed before feeding to the model.

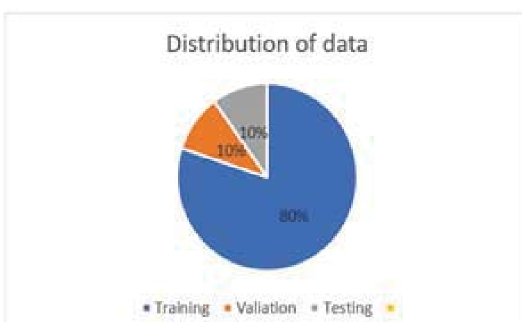


Fig 1. Data distribution

The above figure gives the percentage distribution of the data to be fed to the model. It shows that out of the total data, 80% percent is used for training,10% for validation, and 10% for testing.

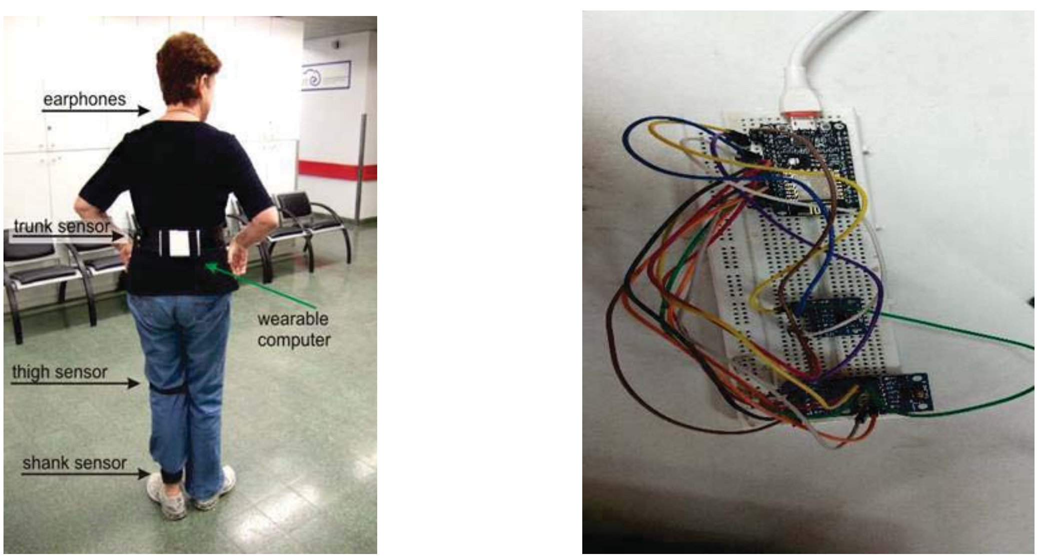
2.3 Process Flow

Figure 2 shows the overall process of the proposed work. The first step is the initialization of the sensors. We are using accelerometer sensors to predict acceleration in three directions. The next step involves the movement tracking of the patient under observation. The next step is data gathering. Micro-python and Python are used as software tools for data gathering. ANN and CNN models classify freezing and non-freezing gait events for disease detection. The same procedure is repeated for another patient.

Fig. 2. The process flow diagram.

2.3 Movement Tracking

Movement Tracking is done by accelerometer sensors ADXL345 attached to the patient's hip, thigh, and ankle, giving the movement in the x, y, and z direction. Fig 3(a) shows the sensor placement in our work. The three sensors are placed on the trunk, thigh, and ankle. The movement is recorded by Node-MCU (ESP8266) and stored as a .csv file in Python. The accelerometer ADXL345 is a capacitive effect type with a fixed and movable plate. The mass shifts as the acceleration is imparted to the sensor, affecting the spacing between the electrodes and capacitors. As a result, the capacitor's capacitance is altered. The acceleration value is computed based on the capacitance after measuring the shift in spacing between the plates. The Hardware setup using motion sensor and ESP8266 is designed in our lab for data collection, shown in figure 3(b).



(a) (b)

Fig 3. Illustration of sensors (a) Placement of the motion sensors [18] and (b) Hardware setup designed for data collection

2.5 Data logging

The data is logged using ESP8266 Node-MCU and stored as a .csv file in Python. A sample of the data collected using the hardware designed in the lab is shown in Figure 4.

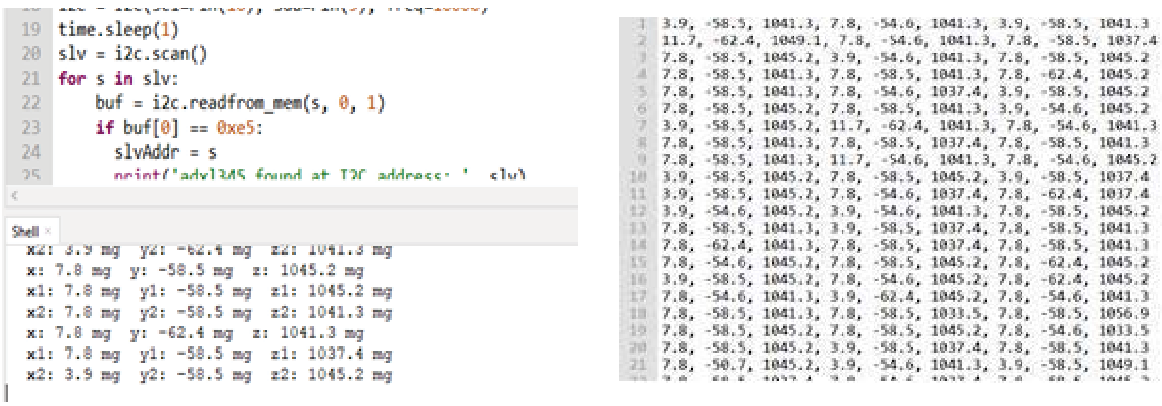


Fig 4: Sample of Data logging.

2.6 ANN Architecture

The ANN architecture implemented in this work uses two dense layers as the hidden layers and one dense layer for the input and one for the output. A dropout of 0.2 is applied after each layer. A dropout of 0.2 means that 80% of data is passed on to the next layer. For the input and output layers, tanh is used, and ReLu for the hidden layers as the activation function. The model is sequential because the data is sequential Time Series data. Two hidden layers give the best accuracy for the model on increasing the number of hidden layers, the model overfits, and accuracy is compromised. Figure 5 depicts the architecture.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 dense (Dense)               (None, 64)                640

 dropout (Dropout)           (None, 64)                0

 dense_1 (Dense)             (None, 64)                4160

 dropout_1 (Dropout)         (None, 64)                0

 dense_2 (Dense)             (None, 64)                4160

 dropout_2 (Dropout)         (None, 64)                0

 dense_3 (Dense)             (None, 1)                 65

=================================================================
Total params: 9,025
Trainable params: 9,025
Non-trainable params: 0
_________________________________________________________________
```

Fig. 5 ANN Model parameters.

2.7 CNN Architecture

The proposed network consists of two 1D convolutional layers, a 1D max pooling layer, and two 1D convolutional layers, and a 1D max pooling layer. These layers are followed by a flattened layer and a dense layer, and the output layer is also a dense layer. Except for the output layer, which uses SoftMax as the activation function, all the levels use tanh as the activation function. Figure 6 depicts the architecture.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv1d (Conv1D)             (None, 126, 128)          3584

 conv1d_1 (Conv1D)           (None, 124, 128)          49280

 max_pooling1d (MaxPooling1D  (None, 62, 128)          0
 )

 conv1d_2 (Conv1D)           (None, 60, 64)            24640

 conv1d_3 (Conv1D)           (None, 58, 64)            12352

 max_pooling1d_1 (MaxPooling  (None, 29, 64)           0
 1D)

 conv1d_4 (Conv1D)           (None, 27, 32)            6176

 conv1d_5 (Conv1D)           (None, 25, 32)            3104

 max_pooling1d_2 (MaxPooling  (None, 12, 32)           0
 1D)

 flatten (Flatten)           (None, 384)               0

 dense (Dense)               (None, 128)               49280

 dense_1 (Dense)             (None, 2)                 258

=================================================================
Total params: 148,674
Trainable params: 148,674
Non-trainable params: 0
```

Fig. 6 CNN Model parameters

3. Results & Discussion

The data set used, the experimental framework, the evaluation metrics used, and the observed outcomes are all discussed in this section. The networks are implemented in Python 3.7 using Keras with TensorFlow as the backend. The entire experiment is performed on Google Colab's (Google's cloud computing platform) Tesla P100-PCIE GPU. Figures 7 (a) and (b) show the confusion matrix for ANN and CNN, respectively. The accuracy, precision, and recall of the models depicted in Table 1 were calculated using the formulae below, which are defined by equations 1-3. This evaluation was based on the true positives (true(+ve)), true negatives (true(-ve)), false positives (false(+ve)), and false negatives (false(-ve)) categories.

3.1 Evaluation Metrics

3.1.1 Accuracy

This is equivalent to the percentage of predictions the model accurately classified.

$$\text{Accuracy (ACC)} = TP + TN / (TP + FP + TN + FN) \qquad (1)$$

3.1.2 Precision

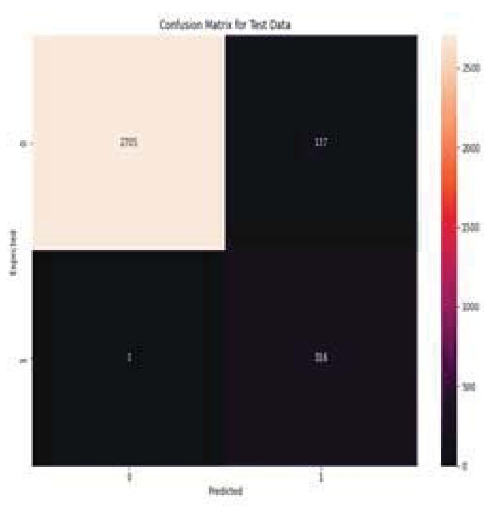
Precision, also referred to as positive predictive value, is the percentage of pertinent retrieved instances.

$$\text{Precision} = TP / (TP + FP) \qquad (2)$$

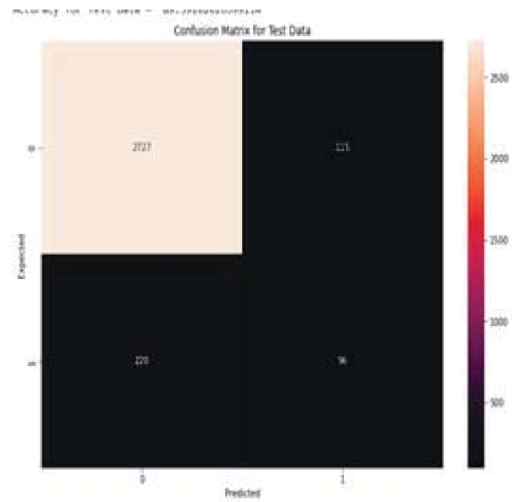
3.1.3 Recall

The recall is the percentage of all relevant instances that were successfully detected, also known as the Sensitivity (SEN), hit rate, or true positive rate (TPR).

$$\text{Recall} = TP / (TP + FN) \qquad (3)$$



(a)



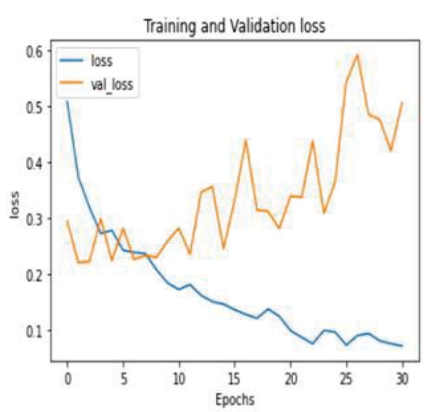
(b)

Fig. 7 Illustration of a confusion matrix for (a) ANN (b) CNN

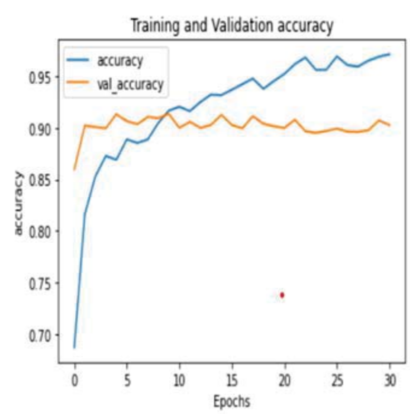
Table 1. Model Classification Evaluation Outcome.

Evaluation parameters	**ANN**	**CNN**
TP	2705	2727
TN	316	96
FP	137	115
FN	0	220
Accuracy	89.6%	89.39%
Precision	90.25%	89.39%
Recall	90.25%	89.39%

As we can see from Table 1, the true positive and the true negative values are better for the ANN model than the CNN model, and the precision and recall percentage is also better for the ANN model. The training and validation accuracy and loss graphs for both models are shown in Figures 8 and 9, respectively.

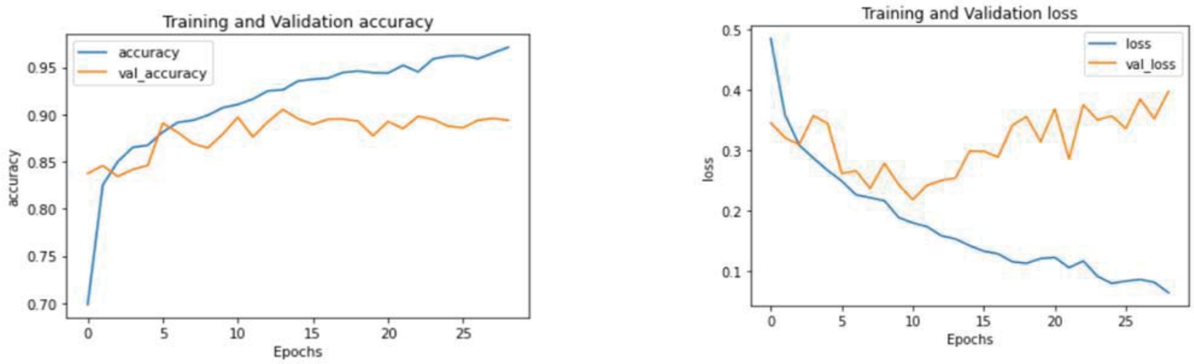


(a)



(b)

Fig. 8. Illustration for ANN model (a) Training and Validation loss (b) Training and Validation accuracy



(a) (b)

Fig. 9 Illustration for CNN model (a) Training and Validation loss (b) Training and Validation accuracy

4. Conclusion

This work mainly aimed to classify freezing and non-freezing of the gait based on the motion sensor data. Since the dataset used in this work is imbalanced, the accuracy can't be used to judge the model's effectiveness. Nevertheless, the precision and recall values will be more critical in determining the best model for the prediction. Hence, regarding precision and recall, it can be concluded that ANN outperforms CNN for the prediction of Alzheimer's disease. The major limitation of this work is the imbalanced dataset with many outliers. The future extension of this work is to improvise the models to have low inference time for real-time testing and the creation of our dataset.

DISCLOSURE STATEMENT

There is no conflict of interest.

FUNDING SUPPORT

IDEAS, Technology Innovation Hub at ISI, Kolkatta under National Mission on Cyber-Physical Systems, has funded this research.

References

1. Marwa, E.G., Moustafa, H.E.D., Khalifa, F., Khater, H. and AbdElhalim, E., 2023. An MRI-based deep learning approach for accurate detection of Alzheimer's disease. Alexandria Engineering Journal, 63, pp.211-221.
2. Pallawi, S. and Singh, D.K., 2023. Study of Alzheimer's disease brain impairment and methods for its early diagnosis: a comprehensive survey. International Journal of Multimedia Information Retrieval, 12(1), p.7.
3. Singh, N., Soni, N. and Kapoor, A., 2022. Automated detection of Alzheimer disease using MRI images and deep neural networks-A review. arXiv preprint arXiv:2209.11282.
4. Sharma, S., Dudeja, R.K., Aujla, G.S., Bali, R.S. and Kumar, N., 2020. DeTrAs: deep learning-based healthcare framework for IoT-based assistance of Alzheimer patients. *Neural Computing and Applications*, pp.1-13.
5. Bandyopadhyay, A., Ghosh, S., Bose, M., Singh, A., Othmani, A. and Santosh, K.C., 2023, January. Alzheimer's Disease Detection Using Ensemble Learning and Artificial Neural Networks. In Recent Trends in Image Processing and Pattern Recognition: 5th International Conference, RTIP2R 2022, Kingsville, TX, USA, December 1-2, 2022, Revised Selected Papers (pp. 12-21). Cham: Springer Nature Switzerland.
6. Sisodia, P.S., Ameta, G.K., Kumar, Y. and Chaplot, N., 2023. A Review of Deep Transfer Learning Approaches for Class-Wise Prediction of Alzheimer’s Disease Using MRI Images. Archives of Computational Methods in Engineering, pp.1-21.
7. Prasath, T. and Sumathi, V., 2023. Identification of Alzheimer's Disease by Imaging: A Comprehensive Review. International Journal of Environmental Research and Public Health, 20(2), p.1273.

8. Lin, H., Jiang, J., Li, Z., Sheng, C., Du, W., Li, X. and Han, Y., 2023. Identification of subjective cognitive decline due to Alzheimer's disease using multimodal MRI combining with machine learning. Cerebral Cortex, 33(3), pp.557-566.
9. Dhinagar, N.J., Thomopoulos, S.I., Rajagopalan, P., Stripelis, D., Ambite, J.L., Ver Steeg, G. and Thompson, P.M., 2022. Evaluation of transfer learning methods for detecting Alzheimer's disease with brain MRI. *bioRxiv*, pp.2022-08.
10. Puente-Castro, A., Fernandez-Blanco, E., Pazos, A. and Munteanu, C.R., 2020. Automatic assessment of Alzheimer's disease diagnosis based on deep learning techniques. *Computers in biology and medicine*, *120*, p.103764
11. Matthews, P.M. and Jezzard, P., 2004. Functional magnetic resonance imaging. Journal of Neurology, Neurosurgery & Psychiatry, 75(1), pp.6-12.
12. Seifallahi M, Soltanizadeh H, Hassani Mehraban A, Fa Khamseh ,2019. Detection of Alzheimer's disease in elder people using gait analysis and kinect camera. J Health Biomed Inform 6(3):178–196
13. Cao, Z., Simon, T., Wei, S.E. and Sheikh, Y., 2017. Realtime multi-person 2d pose estimation using part affinity fields. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7291-7299).
14. Moreno-Noguer, F., 2017. 3d human pose estimation from a single image via distance matrix regression. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2823-2832).
15. G. Palvakos, X Zhou, K.G. Derpanis, K. Daniilidis.,2017, Coarde to fine volumetric prediction for single 3D human pose in: : 2017 IEEE Conference on Computer Vision and Pattern Recognition, IEEE 2017, pp-1263-1272
16. Vora J, Tanwar S, Tyagi S, Kumar N, Rodrigues JJ (2017) Faal: fog computing-based patient monitoring system for ambient assisted living. In: IEEE 19th international conference on e-health networking, applications and services (Healthcom), vol 2017. IEEE, pp 1–6
17. Prakash, C., Kumar, R. and Mittal, N., 2018. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artificial Intelligence Review*, *49*, pp.1-40.
18. Bachlin, M., Plotnik, M., Roggen, D., Maidan, I., Hausdorff, J.M., Giladi, N. and Troster, G., 2009. Wearable assistant for Parkinson's disease patients with the freezing of gait symptom. *IEEE Transactions on Information Technology in Biomedicine*, *14*(2), pp.436-446
19. Casolla, G., Cuomo, S., Di Cola, V.S. and Piccialli, F., 2019. Exploring unsupervised learning techniques for the Internet of Things. *IEEE Transactions on Industrial Informatics*, *16*(4), pp.2621-2628.
20. Chitradevi, D. and Prabha, S., 2020. Analysis of brain sub regions using optimization techniques and deep learning method in Alzheimer disease. *Applied Soft Computing*, *86*, p.105857.
21. Kumar, P.M., Lokesh, S., Varatharajan, R., Babu, G.C. and Parthasarathy, P., 2018. Cloud and IoT based disease prediction and diagnosis system for healthcare using Fuzzy neural classifier. *Future Generation Computer Systems*, *86*, pp.527-534.
22. Long, X., Chen, L., Jiang, C., Zhang, L. and Alzheimer's Disease Neuroimaging Initiative, 2017. Prediction and classification of Alzheimer disease based on quantification of MRI deformation. *PloS one*, *12*(3), p.e0173372.
23. Liu, M., Zhang, D. and Shen, D., 2016. Relationship induced multi-template learning for diagnosis of Alzheimer's disease and mild cognitive impairment. *IEEE transactions on medical imaging*, *35*(6), pp.1463-1474.

IoT Based Soil Health Monitoring System Using Soil Moisture Sensor and pH Sensor

Mrinalini Barik[a]*, Sitanshu Sekhar Sahu[b], Rupesh Kumar Sinha[c]

[abc] *Department of Electronics & Communication, Birla Institute of Technology, Mesra, Ranchi, Jharkhand*

Abstract

Soil management is a fundamental component of land management which emphases on the diverse soil forms and features that can boost the soil production capabilities. This paper proposes a process to apply an Arduino based soil monitoring framework. This framework uses two major ecological features like moisture and pH value of soil using soil moisture sensor and pH sensor probe. Data is managed by ESP32 which may be connected to Thing speak by the concept of Internet of Things (IoT). This permits the client to control water system pumps from far distances through the Thing speak App, which would assist farmers with greatest yielding and quality harvests.

Keywords: Internet of Things, ESP32, Soil moisture, pH value, Thingspeak.

1. Introduction

In agriculture, rich and healthy soil with proper nutrients can significantly contribute to a successful farming season. Soil offers the primary help to plant. Soil has a wide range of varying molecular and physical characteristics. There are many factors such as leaching and microbial activity that can make an entirely different soil type. Good soil structure includes proper management of soil. It is necessary to monitor physical and chemical features of soil, considering all the factors. Soil acts as a pantry for plants and provide human nutrition. Managing soil is important as carelessness can hamper the growth of plants, resilient to droughts and affect life of farmer and rest of the population. Proper soil management protects and enhances soil performance. It also prevents pollution and improves crop quality. Therefore, it is necessary to manage land more effectively by analysing the soil. Nowadays, smart systems are introduced for various problems in real life. IoT, more up to date innovation is a continuous thought which manages the cooperation among objects and collects data online. It is one of the essential components of smart system.

The climatic condition is also an essential reason for crop growth [1]. The combination of IoT and modern technologies can help in increasing production and quality of soil [2]. Proper water system which will conserve water and as water is essential for plant growth [3] [4]. A regular water system that can be controlled so that the soil moisture condition is balanced according to the requirement of the plant. [5] The study was aimed to develop an automatic watering process with a unit to control the motor using MQTT protocol for communication. [6], To make a less cost and automatic system to control soil moisture in rural farms using Arduino UNO. IoT enabled sensor with neural network and predict seasonal rainfall in North Karnataka. [7], Design of less price IoT structure for analysing soil status, temperature and humidity using wireless sensor network. [8] This study associates the environmental parameter like humidity, temperature with IoT based controlling using Arduino and Raspberry pi. [9], This project includes different sensors positioned over the plant and transfer the required parametric values. [10], The study is about maintaining fish in aquarium by automated system using pH, ultrasonic and servo data receiver. [11], Study suggesting high correctness crop control by data gathering and computerized farming method using ADS1115 I2C protocol. [12], This study includes smart agriculture technique like GPS remote controlling, parameter sensing and leaf dampness using IoT. [13], The study implements a wireless sensor network with different sensor using ESP-WiFi-MESH protocol to monitor small area. The parameter includes moisture of soil, temperature, humidity and sunlight intensity. [14], proposed automated system monitoring the turmeric farm by wireless sensor network technology. [15], The research proposed a smart agriculture, with sensor-based monitoring in real time with less human interaction.

The proposed study for IoT created soil health monitoring structure with two different sensors consists of a brief introduction about smart soil management, then the materials and methodology are discussed at last comes the result for the posed work.

* Mrinalini Barik
E-mail address: mrinalinibarik99@gmail.com

2. Smart Soil Management

Most of the ground is covered in loose surface material called soil. It offers actual support to plants and is additionally their wellspring of water and supplements in agribusiness. Soil management is a process to protect soil resources and make soil and crops less susceptible to pests, diseases, and weeds.

Smart soil management is an effective and efficient way of knowing the different properties of soil. It informs the farmers with real time data and monitors the soil. Checking of soil parameter is one of the main issues in horticultural practices. The creation and support of yields are straightforwardly impacted by soil boundaries. Hence, approaching a smart soil management system can help farmers to expand their yield at least expense. In the smart soil management model, we use different types of sensors for informing the real time data and the use of sprinklers to chill off the temperature of the yields. Our methodology is to make this framework open for far distance so farmers/clients can get the data and control the field 24╳7. The setup is controlled by Arduino, the information is transferred to the Wi-Fi unit in Node MCU like, ESP32.

3. Materials and Methods

The main part of the system is ESP32 board, which is a microcontroller with integrated Wi-Fi and dual mode Bluetooth network that makes it possible to create an Internet of Things (IoT) application. In this work, the process is separated into two parts. The first one is to develop the hardware setup with the sensors and setup for monitoring the soil moisture at the same time. Secondly, the transfer and storing of data to cloud. The procedure of the complete system is given below in Fig.1.

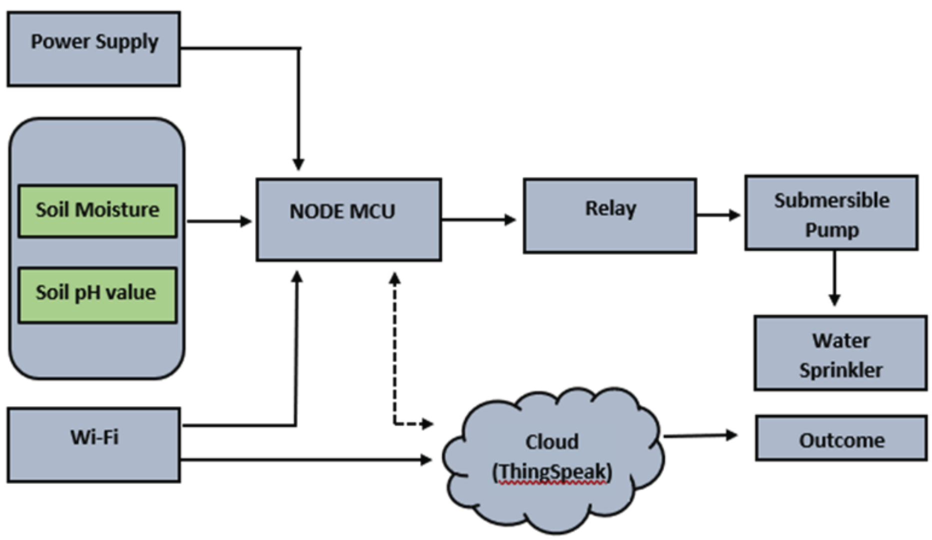


Fig.1. System workflow diagram

3.1 Hardware Specification

Here, different elements were utilized for developing the entire model. The components were put together to perform their respective tasks. The microcontroller is used as system input and the two sensors as output for the sensor readings. The subsequent components are used in this proposed system.

3.1.1 ESP32 (Wi-Fi module)

The controller named Node MCU ESP32 board, fig.2 is a chip based on Espressif product which is scalable, adaptive, and compact board with inbuilt Wi-Fi and Bluetooth capability to communicate. An ESP32 board module is utilized as the fundamental module for collecting sensor data. It can develop a program and as well as control movements of every kind in the proposed framework. Hence, the microcontroller was able to communicate with sensors and other devices.

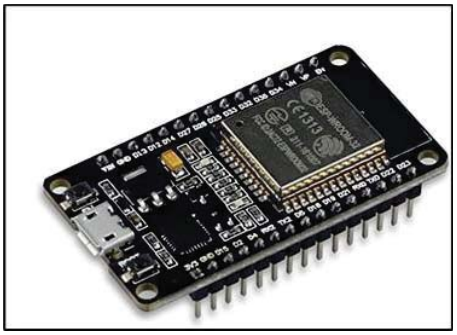


Fig.2. ESP32 microcontroller board

3.1.2 Soil Moisture Sensor

The capacitive soil moisture sensor, sensor which measures moisture level of soil by capacitive sensing. Capacitance measures dielectric permittivity of the surrounding area. It does not measure soil moisture directly, instead it measures the ions that are dissolved in the moisture. While measuring, the current cannot go through the dry soil as the water content is very low and when the soil is wet, the current can undoubtedly pass. To determine if soil is wet or dry enough, the moisture threshold number can be changed. Fig.3 shows soil moisture sensor. The moisture sensor and relay actuator were integrated into the module, which is a submersible pump of 3 – 6 V to regulate and monitor the water flow and water content in intended agricultural land as shown in Fig.4.

Fig.3 Capacitive soil moisture sensor

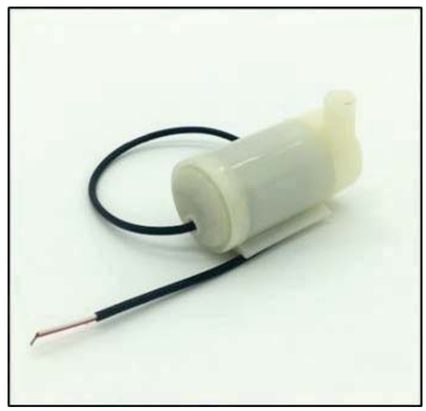


Fig.4 Submersible motor pump

3.1.3 pH Sensor

The second sensor shown in fig.5, which is used in this model is E-201-C composite electrode pH sensor probe with BNC connector. This pH sensor measures the pH value of soil, so that the chemical nature of the soil can be known in the form of acid or base. It can be interfaced directly by any microcontroller.

Fig.5 E-201-C pH sensor probe

3.2 Software Specification

In the software module, Arduino IDE 1.8.19, the open-source Arduino software where interfacing is quite simple to use. The communication process was analysed and visualized by Thingspeak, an IoT platform that allows to aggregate and analyse live information streams in the cloud. Thingspeak allows sensors, instrument to lead data to the cloud and stored in channel. It is the best open source IoT platform where sensor can continuously send data to cloud platform and the data are interpreted in form of graph. It can measure and report simple as well as complex data.

3.3 System Design

The work process of this proposed model works using ESP 32 developer kit. This device is powered using an external power supply. The two sensors which are going to give live data are soil moisture sensor and pH value sensor. The working begins when microcontroller gets power and the sensors are connected to it. ESP32 start analysing the data and keeps on displaying on arduino IDE. The Wi-Fi module which is inbuild in ESP32 helps in transferring measurement reading to IoT cloud platform. Coming to the sprinkler system we connected a relay actuator with a motor which pumps water according to the current requirement of the field, which depends on the water content. This workflow for measurement of soil moisture and the condition to control the pump is shown in Fig 6.

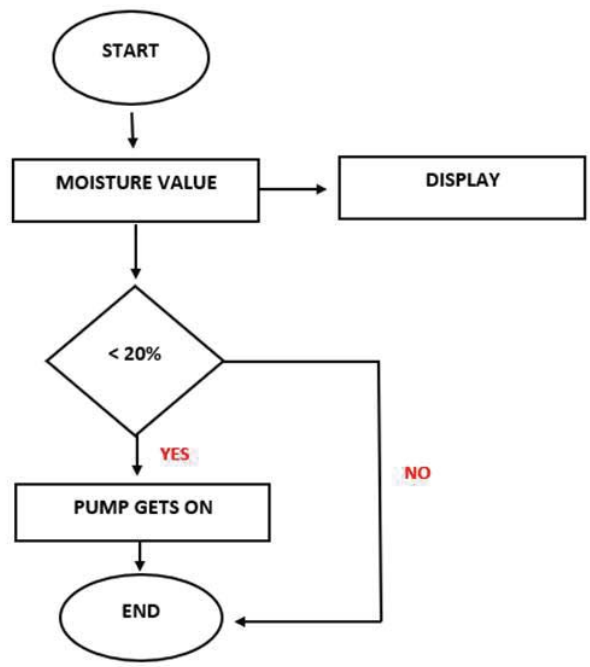


Fig.6. Pump control process

The procedure flow for configuring the water pump follows by first sensing moisture value of the soil. Referring to readings the point rate is set as 20%, the pump starts operating according to the condition of the land.

The observed data of the moisture measurement and pH value was collected by a broker where all the information were gathered. Further, the data were transferred from device over the internet to cloud platform where the data are analysed and visualized in the form of graphs with real time updated value.

4. Results

The working of the projected Arduino based smart soil management with ESP32 Wi-Fi module was tested with a real time scenario. The sensor data were received by implementing the whole module as shown in Fig.7. The pump control process is attached with the module as shown in Fig.8 which is to provide a sprinkler system according to the moisture level of the soil. The real time data was collected from the soils of Mesra, Ranchi, Jharkhand.

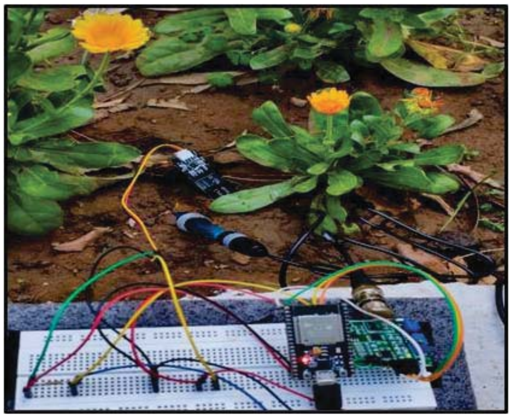


Fig.7 System model

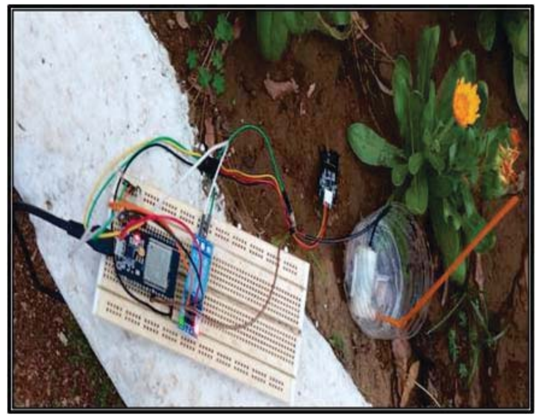


Fig.8 Pump control model

The main component in the proposed model is ESP32 Wi-Fi unit that permits to transfer and receive data via the internet. It was used in a local area network (LAN) approach for the transferring of data to the cloud. After the setup was completed, we started taking the reading for the moisture sensor and pH sensor at the same time interval for two months interval. First reading was taken in the month of November to December and second as January to February. The data collected for sample soil can be seen in Table.1 and Table.2 below.

Table 1. Sample Collected for NOVEMBER – DECEMBER (2022)

TIME STAMP	SOIL MOISTURE	SOIL pH VALUE
16:40:54	18.00%	7.9
16:40:55	18.20%	7.8
16:40:56	18.20%	7.9
16:40:57	19.40%	7.9
16:40:58	19.40%	7.5
16:40:59	19.70%	6.9
16:41:00	19.75%	7
16:41:01	19.75%	7.2
16:41:02	19.85%	7.5
16:41:03	19.88%	6.6
16:41:04	20.00%	7.2
16:41:05	20.00%	7.5
16:41:06	19.55%	7.5
16:41:07	19.89%	7.8

Table 2. Sample Collected for JANUARY – FEBRUARY (2023)

TIME STAMP	SOIL MOISTURE	SOIL pH VALUE
16:40:54	12.00%	7.9
16:40:55	12.20%	7.8
16:40:56	12.20%	7.9
16:40:57	12.40%	7.9
16:40:58	15.40%	7.5
16:40:59	15.70%	6.9
16:41:00	15.75%	7
16:41:01	15.75%	7.2
16:41:02	16.85%	7.5
16:41:03	16.88%	6.6
16:41:04	17.00%	7.2
16:41:05	17.00%	7.5
16:41:06	18.55%	7.5
16:41:07	18.89%	7.8

From the Table.1, it is states the samples collected from moisture sensor and pH sensor at same timestamp for sample soil. The soil moisture ranged from 18.5% to 20% and soil pH ranged from 7.5 to 7.9. Table.2, shows that the moisture ranged from 12% to 18% and Ph value is same as the soil and patch area was same. The moisture value can vary season to season and is also depended on temperature. The pH of the soil remains the same but if the user wants to grow.
crops that need acidic nature soil, then the user can add different fertilizers to change the pH of the soil. After the collection of data, the readings are transferred from the Arduino database to Thingspeak every minute with a common Wi-Fi network. The continuous data in the form of graphs is shown in the main page of the Thingspeak with two different fields as shown in Fig. 9 and Fig. 10 to make it simple for clients to see the historical backdrop of past information.
This working structure was able to gather information about the soil state with two different parameters. Considering Jharkhand state, Ranchi region as the samples is collected here. Different crops and plants which are commonly grown in this region having specific range for the considered parameter is mentioned as shown in table. 3. The measured value is aligned to the range mentioned in the table.3.

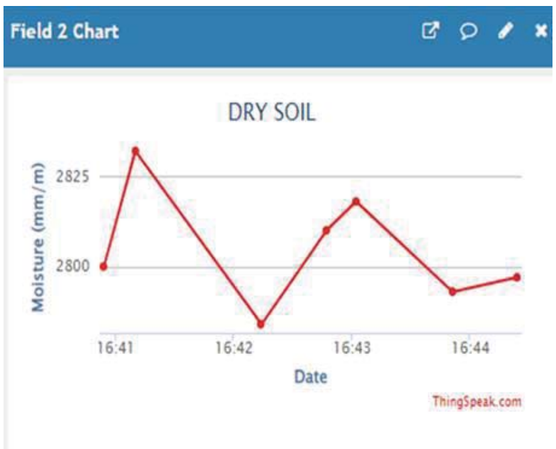


Fig.9 Graph for soil moisture

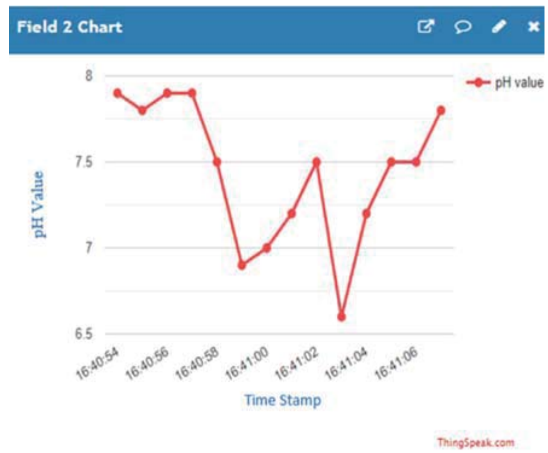


Fig.10 Graph for pH value of the soil

Table 3. Common crops grown in Ranchi region

CROPS	SOIL MOISTURE	SOIL pH VALUE
POTATO	60-70%	5.5 – 7.0
TOMATO	<50%	6.0 – 7.0
FRENCH BEANS	50 - 70%	5.5 - 6.0
PUMPKIN	30 - 50%	6.0 - 6.8
LADYFINGER	50 - 60%	6.5 - 7.0
WATERMELON	40 - 60%	6.2 - 6.8
RICE	60 - 80%	6.0
HIBISCUS	30 - 40%	5.5 - 6.5
MARIGOLD	35 - 50%	6.2 - 6.5

5. Conclusion

In the proposed methodology, a module has been explained which can monitor the moisture level and create a sprinkler structure accordingly with less human interference. The framework effectively gave the data in regards to the state of soil, particularly moisture and pH and delivered it to the cloud system. The IoT based soil health monitoring system using two different sensors was successfully designed for any season condition. Moisture level is predicted by knowing the water content of soil. The pH value validated the soil in two conditions, acidic or alkaline. The two parameters in the proposed system are equally important for soil and the growth of a good quality crop. For future development, the watering system can be changed in a more proficient and enlarge way with proper sprinkler system, alert and inform the user through web interface. The system can be modified by using other different sensors.

References

1. J. E. Tamil Malar and M. Vaishnavi, 2022 "IoT based Smart Irrigation System using ESP32," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India.
2. A. Chakraborty, M. Islam, A. Dhar and M. S. Hossain, 2022 "IoT Based Greenhouse Environment Monitoring and Smart Irrigation System for Precision Farming Technology," 2022 International Conference on Innovations in Science, Engineering and Technology (ICISET), Chittagong, Bangladesh.
3. Pushkar Singh, Sanghamitra Saikia," 2014 Arduino-Based Smart Irrigation Using Water Flow Sensor, Soil Moisture Sensor, Temperature Sensor and ESP8266 WiFi Module" IEEE International Conference on Control System, Computing and Engineering (ICCSCE)
4. Vaishali S, Suraj S, Vignesh G, Dhivya S and Udhayakumar S," Mobile Integrated Smart Irrigation Management and Monitoring System Using IOT" International Conference on Communication and Signal Processing, April 6-8, 2017, India.
5. I Komang Agus Ady Aryanto, Roy Rudolf Huizen, Kadek Yota Ernanda Aryanto. "Design of Soil Humidity Monitoring System Using the Internet of Things Concept and MQTT", 2020 International Conference on Smart Technology and Applications (ICoSTA), 2020
6. Suhas Athani, CH Tejeshwar, Mayur M Patil," 2017 Soil moisture monitoring using IoT enabled Arduino sensors with neural networks for improving soil management for farmers and predict seasonal rainfall for planning future harvest in North Karnataka - India" IEEE
7. Madhura U K, Akshay P, Akshay J Bhattad, Nagaraja G S, 2017" Soil Quality Management using Wireless Sensor Network" 2nd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions.
8. Lalbihari Barik, 2019 "IoT based Temperature and Humidity Controlling using Arduino and Raspberry Pi" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 10, No. 9.
9. Nikesh Gondchwar, R. S. Kawitkar, "IOT based smart agriculture," International journal of Advanced research in computer and Communication Engineering (IJARCCE), vol. 5, no.6, Jun. 2016.
10. Riko Reynaldi, Dani Hamdani "Design and Implementation of an Aquarium Automation System using a NodeMCU", International Journal of Industrial Research and Applied Engineering (JIRAE), Vol. 5, No. 1, April 2020.
11. V.D.Bachuwar, A.D.Shligram, L.P.Deshmukh, 2018 "Monitoring the soil parameters using IoT and Android based application for smart agriculture" AIP Conference Proceedings 1989.
12. N. Suma, Sandra Rhea Samson, S. Saranya, G. Shanmugapriya and R. Subhashri, "IoT Based Smart Agriculture Monitoring System", *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 5, no. 2, Feb 2017.
13. A. G. Q. Aquino, A. H. Ballado and A. V. Bautista, "Implementing a Wireless Sensor Network with Multiple Arduino-Based Farming Multi-Sensor Tool to Monitor a Small Farm Area Using ESP32 Microcontroller Board," *2021 IEEE 13th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)*, Manila, Philippines, 2021
14. M. P. M and V. J, "Turmeric Farm Monitoring System using Wireless Sensor Network with ESP32 Module," *2021 Smart Technologies, Communication and Robotics (STCR)*, Sathyamangalam, India, 2021
15. T. S. Gunawan, N. N. Kamarudin, M. Kartiwi and M. R. Effendi, "Automatic Watering System for Smart Agriculture using ESP32 Platform," *2022 IEEE 8th International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA)*, Melaka, Malaysia, 2022

Potato Leaf Disease Classification Using Image Processing and Machine Learning

Priyanshu Tripathi[a], Mihir Daiv[b], Gourav Gupta[c], Saubhagya Monal[d], Dr. Sandip Langade[e]*

[a,b,c,d]*Student, MIT Academy of Engineering, Alandi, Pune-412105, India*
[e]*Assistant Professor, MIT Academy of Engineering, Alandi, Pune-412105, India.*

Abstract

Food is important for every living being, and so its quality and safety measures are also critical for the entire food chain. Food safety and quality assurance are complicated processes that require consideration of all phases of food processing, including cultivation, harvesting, storage, and consumption. These procedures, on the other hand, are frequently time-consuming. The present work focuses on the use of image processing and artificial intelligence in the sector of agriculture, with a focus on quality analysis and grading. Sorting fruits by grade is a time-consuming and labour-intensive operation, which makes it difficult for growers with limited resources. Automation comes to our rescue in order to design a system that can address these issues. The use of automation on fruit presents some problems, such as preserving fruit quality and safety while being processed by such methods. As a result, this research proposes a non-destructive approach that does not injure the fruit. The proposed methodology is based on image processing and deep learning techniques, which are the most appropriate and effective solutions for the above-mentioned issues. Edge detection, image contours, and artificial neural networks are some of the methodologies discussed in this study for implementing the system. Different kinds of potato plant diseases, including early blight, late blight, septoria blight, etc., will affect the plants and show their disorder in the leaves. If these outbreaks are discovered at the primary stage and appropriate intervention is taken, the farmer would not be at risk of suffering significant financial losses. The suggested model would effectively identify and diagnose potato leaf stand illnesses using image processing techniques. The CNN model is utilised in this study to identify the disease from photos of the potato leaf because CNN is used for image classification and produces the best results.

Keywords: Machine Learning; Image Processing; Convolutional Neural Network (CNN); Agriculture; Potato; etc.

1. Introduction

Fruit classification and recognition may be a subject involving a large variety. At present, the research on fruit image classification and recognition in the main focuses on fruit quality classification, maturity recognition, defect detection, and mechanism choosing, and typically aims at a definite sort of fruit, whereas there's less analysis on multi-sort fruit classification and recognition in China [1]. The classification and recognition of various fruits has a wide range of practical applications. As an example, in supermarkets in developed countries, folks use multi-class fruit image recognition to aid fruit purchase; within the line, the classification and recognition of multiple types of fruits may also scale back manual errors and improve production potency [2]. Additionally, the classification and recognition of fruit pictures also have sure analytical significance within the fields of intelligent agriculture and digital health care [3]. Within the facet of intelligent agriculture, the fine management of fruit trees and automatic fruit choosing are accomplished through the identification of fruits; within the field of digital drugs, the popularity of fruits is employed to investigate the kinds of fruits so as to acquire the nutrients contained in them and facilitate patients in formulating a reasonable dietary combination during the later stages of recovery [4]. The leaves are a vital element of the plant, and assessing agricultural crops is critical. The texture and color of the leaf are the most noticeable visual properties [5]. As a result, identifying leaf diseases is critical for assessing agricultural goods, increasing market value, and meeting quality requirements. It is also critical to notice and take further actions to prevent illness transmission. It will be far too sluggish. When physical techniques are employed to classify and categorize the approach, we also need specialists to guarantee that it doesn't go wrong and who are less commonly available [6]. Also, the motivation comes from automating the agricultural field using the recent advancements in technologies of AI, ML, Deep Learning and Data Science. All this will simultaneously result in reducing the manpower needed for the purpose.

* Priyanshu Tripathi
E-mail address: tripathipt2001@gmail.com

2. Literature Survey

Narayana [1]: The author covers the area of describing the improvements of AI research for focusing the Deep Neural Network. The Study also details about the various uses of AI model into recognizing medical images, which will help in medical services. It also entails about the impacts of the AI system in changing the building strategy of the image processing. Khan & Habsi [2]: In this paper, the author covers Computer Vision can be used in detection of images, classification of faults in railroad ties using images, interpreting remote sensing data for geographical information systems, etc. Wiley & Lucas [3]: Creating models and feature extraction is done using Computer vision and Image Processing is all about computational transformations for images is what this paper suggests. Muresan & Oltean [4]: In this paper, the author has proposed a dataset of fruit images and used it to build a model for training of identification of different fruits using Deep Learning. Rokunuzzaman & Hemantha et al. [5]: Employs a machine vision technique to classify the tomatoes with flaws. The blossom end rot and tomato cracks are the two main flaws covered by the author. The neural network approach resulted in an accuracy of 87.5%. Additionally, the author contrasted rule based and neural techniques. Seng & Mirisaee et al. [6]: To identify the fruit photos based on their feature value, Woo Chaw Seng et al presents a neighbors classification technique. According to the author, this strategy improves accuracy by up to 90%.

3. Proposed Approach

So basically, what we are trying to do here is predict the type of disease that will affect the potato leaf. There are three classes in which we have classified our data: the early blight, the late blight, and the healthy one. So, the user here will give the image as an input, and the model will be predicting the type of disease with which the leaf has become infected or whether the leaf is healthy and safe. So, for this purpose, we analyzed the data, and we went through the different articles and research papers, and we came across using CNN as the technique for this project, which was found to be a very optimal method. Several different ways to approach the problem have used the CNN technique.

The first deep learning model in the field of computer vision was developed by the Japanese scholar Fukushima. A convolutional neural network can be used for a biological vision system. The CNN can be compared to the BP neural network, which uses the concept of forward propagation to calculate the output price and uses back propagation to adjust the load and biases. The architecture of CNN contains five sections: the input layer, the convolution layer, the down sampling layer, the fully connected layer, and the output layer. Each convolution layer consists of multiple feature extraction maps. Every feature matrix contains several freelance neurons, and every neighboring cell is regionally connected with the neurons within the input (feature map) of the higher layer.

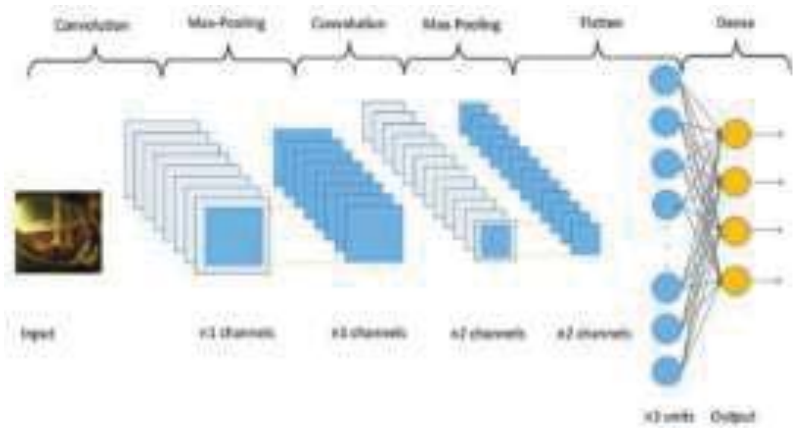


Fig 1. CNN Architecture

Due to a predetermined relationship between entirely unlike native options, once a feature is extracted, its position in relation to alternative alternatives is predetermined rather than its true position being known. The input picture (feature map) from the previous layer is convoluted with various filters to produce completely distinct feature maps within the convolution layer. The weights of each neuron inside a feature map are equal since each one represents a feature. The lower sampling layer, represented by dense layers, decreases the dimensions of the images of the feature map in the network by performing basic scaling on different feature maps within the convolution layer. To mix the choices and ultimately identify and classify them, a full affiliation layer may be linked after convolution and down sampling before the network's output layer. The results are provided at the output layer as a result of the combination of feature extraction and pattern classification, which yields the desired parameters through iterative improvement. The number of parameters that must be trained within the network is drastically decreased using thin associations and weight sharing, which leads to an improvement in generalization performance.

4. Methodology

The basic steps of plant disease detection are shown below:

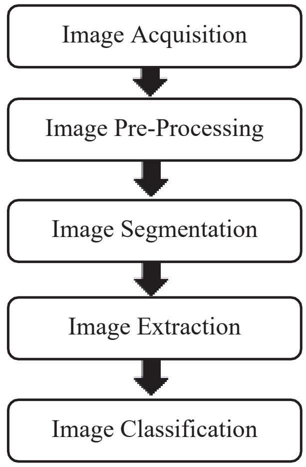


Fig. 2. Methodology

4.1 *Image Acquisition:* The first step in the process is to gather all images of diseased or infected plant leaves using a camera or a supply that records all such data. The first image is sorted using RGB (Red, Green, and Blue). In this instance, an accompanying RGB leaf image is used in the color transformation process, and a device-independent color-space transition is constructed for it.

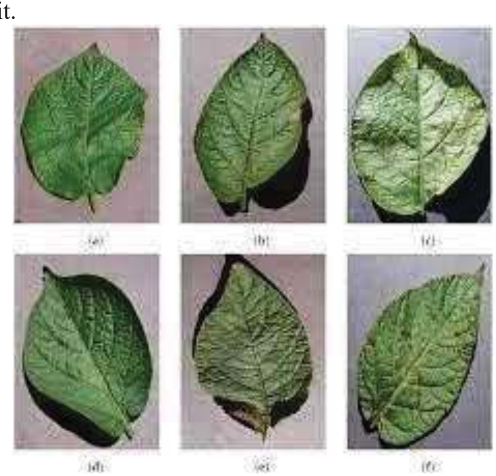


Fig 3. Image Acquisition

4.2 *Image Pre-processing:* The plant leaf that is infected produces noise in the image. Presence of dust on the leaf, sand, and other things could also make this noise. Getting rid of the yelling information from the image is crucial to achieving greater accuracy. The removal of yelling image pictures then uses pre-processing techniques. There are several different preprocessing methods available. To capture the attention of the world, an image is flipped, clipped, or cropped, to make it more interesting. Using the smoothing filter, the idea is made more uniform. It is distributed to improve the image in order to increase contrast. It uses a variety of techniques, each of which has an impact on the outcome.

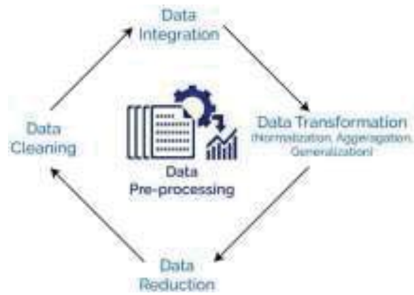


Fig. 4. Data Preprocessing

4.3 *Image Segmentation:* Segmentation refers to dividing an image into distinct parts that are similar in nature or appearance. Numerous methods, including k-means clustering, transforming an RGB image into a HIS model, etc., can be used to segment data. The clustering of K-means is used to divide objects into a number of categories K depending on the set of options we provide. By decreasing the addition of the squares of the space between the item and the cluster within the query, objects are classified.

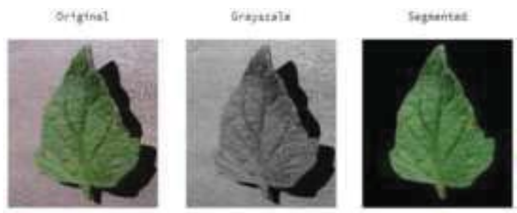


Fig. 5. Image Segmentation

4.4 *Feature Extraction:* Feature extraction plays a very important role in the identification of any kind of disease. Many image processing applications use the extraction operation. The major characteristics that we will use for identifying the plant diseases are color, texture, form, edges, etc. It has been discovered that morphological outcomes outperform the alternative possibilities. Texture refers to the coloration, abrasiveness, and hardness of an image. With this process, the traits of contaminated vegetable leaves are gathered.

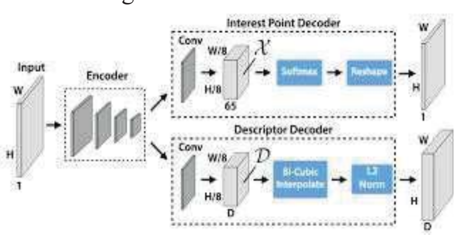


Fig. 6. Feature Extraction

4.5 *Classification:* Once the features have been extracted, a classifier is developed and used to categorize the difference in the features. Plant diseases are categorized and described based on the health or illness of their leaves. Based on the type of disease it has been affected by unhealthy leaf or a healthy leaf is classified. Unhealthy leaf is any hierarchic based upon the sort of sickness with which it has been infected.

Fig. 7. Classification

In this proposed model, the dataset which was used was named 'Plant Village.' It consisted of a total of 2152 images of potato leaves which comprises of three different types of textures of leaves such as,

- 1000 images consisted of Early-Blight Disease.
- 1000 images consisted of Late-Blight Disease.
- 152 images of Healthy Potato Leaves.

The images needed to be adjusted to 256 × 256 size, and then randomly resized and rescaled for the purpose of increasing diversity of samples. These images were sent as input to the network operation. Dataset is divided into three major parts i.e.; Training part comprises of 80% of Dataset and the Testing and Validation comprises of 10% of Dataset each. After the model building, we have used the optimizer "Adam" and the loss function used is "Categorical Crossentropy" because we are doing multi-class classification of images and we compile the model based on accuracy.

Fig. 8. Validation

5. Conclusion

The following conclusions were achieved:

- The advantages of the farmers and, consequently, the farming industry, were taken into consideration when designing a planned system.
- The created system can identify the plant illness and offering a solution that can be used to treat it.
- We will handle it with the correct data of the unwellness and therefore the cure to enhance the plant's health.
- The model is trained with additional information that may cowl all the diseases that are gifted, and it ought to be ready to counsel the precautions.
- A well calculated study was being done on the selected data. It was ell studied and accordingly the model was built up.
- This may scale back a lot of the farmer's work and can bring automation within the agricultural trade.

References

1. Sathya Narayana PSV 2019, "AI in Image Analytics" International Journal ofComputational Intelligence Research ISSN 0973-1873 Volume 15, Number 2.
2. Asharul Islam Khan, Salim Ai-Habsi 2020, "Machine Learning in Computer Vision" International Conference on Computational Intelligence and DataScience Published by Elsevier B.V.
3. Victor Wiley, Thomas Lucas 2018, "Computer Vision and Image Processing" International Journal of Artificial Intelligence Research 2, No. 1, June.
4. Horea MURESAN Mihai OLTEAN 2018, "Fruit recognition from images using deep learning" Acta Univ. Sapientiae, Informatica 10, 1 26–42, DOI: 10.2478/ausi-2018-0002.
5. Rokunuzzaman, Md & Jayasuriya, Hemantha. 2013, Development of a low-cost machine vision system for sorting of tomatoes. Agricultural EngineeringInternational: CIGR Journal, 15. 173-180.
6. Seng, Woo & Mirisaee, Seyed. 2009, A new method for the fruit recognition system. Proceedings of the 2009 InternationalConference on Electrical Engineering and Informatics, ICEEI. 1. 130 - 134. 10.1109/ICEEI.2009.5254804.
7. Femling, Frida & Olsson, Adam & Alonso-Fernandez, Fernando. 2018, "Fruit and Vegetable Identification Using Machine Learning for Retail Applications".
8. Kishan Das Menon, H. et al. 2020, "Digital grading and sorting of fruits." MaterialsToday: Proceedings.
9. Mr. Sumit S. Telang, Prof.S.M. Shirsath, 2017, ''Fruit Quality Management using Image Processing", International Conference on Ideas, Impact andInnovation in Mechanical Engineering Volume 5, Issue 6.
10. Anuja Bhargava, Atul Bansal, 2021, "Fruits and vegetables quality evaluation using computer vision: A review", Journal of King Saud University - Computer and Information Sciences, Volume 33, Issue3,Pages 243-257,ISSN 1319-1578,
11. Ismail, Nazrul & Malik, Owais, 2021, "Real-time Visual Inspection System for GradingFruits using Computer Vision and DeepLearning Techniques". InformationProcessing in Agriculture, DOI: 10.1016/j.inpa.2021.01.005.

12. Chen, Jieling & Zhang, Min & Xu, Bao-guo & Sun, Jincai & Mujumdar, Arun. 2021, "Artificial intelligence assisted technologies for controlling the drying of fruits and vegetables using physical fields: A review". Trends in Food Science & Technology. DOI: 105. 251-260. 10.1016/j.tifs.2020.08.015.

Artificial Intelligence Enabled Automated Basketball Score Point Detection System

Varun S V[a*], Muthaiah U[b], Shankar V P[c], Roshan R[d]

[a,b,c,d]*Department of Computer Science and Engineering, Amrita School of Engineering, Chennai, India*

Abstract

Basketball is one of the most popular games in the world, and the industries that are tied to it have also greatly benefited the economy. Basketball-related applications of artificial intelligence (AI) technology have received a lot of interest recently. Most of the research have demonstrated that artificial intelligence (AI) technology can raise basketball players' training standards, assist coaches in developing effective game plans, reduce sports-related injuries, and enhance game enjoyment. This article proposes an AI-based automated basketball scoring system that utilizes computer vision and deep learning techniques to accurately track and record the scores of players in real-time. The proposed system utilizes pose detection, basketball, and hoop detection, and tracking methods to detect scores during the game. The algorithm then interprets the data and updates the scoreboards and game statistics in real-time. The article provides the system architecture's specifics as well as the findings of experimental testing carried out to confirm the system's precision and dependability.

Keywords: AI, frame processing; detection; tracking; deep learning; Basketball; Scoreboard

1. Introduction

The use of technology and artificial intelligence (AI) in sports has become increasingly popular in recent years, particularly in sports analytics. However, basketball heavily relies on sensors, which involves a huge cost for deployment setup. To address this, we are introducing a profitable venture - an automatic scoring system for basketball leagues and tournaments. In the National Basketball Association, discussions and voting are often held to decide the scoring factors in basketball shooting scenarios. Referees are critical in deciding whether a shot made is a 2-pointer or a 3-pointer. However, human errors tend to occur during pressure situations, despite referees being trained to handle such situations efficiently. Additionally, a committee is formed to analyse each shot to determine the scoring point from the origin of the shot. Six members take part in these activities to ensure the score is provided correctly, making the system completely manual. Hence, automating such a manual system was an interesting concept to work on. This project is the first fully automated basketball scoring system that uses past frame processing, tracing the ball's path from the moment it entered the hoop back to the player who released it from their hand. This project is the first step towards making the basketball scoring system fully automated, without any interference from humans.

Currently, the official NBA basketball is scored using sensors and a panel of members (former players) from other teams who view different angles of the same shot. Based on the panel's review, points are awarded to the team. We went through all combinations to make the scoring system cheaper and more efficient. We explored three different approaches: hardware, software, and a mix of both. We found that the most efficient method was a mix of software and hardware, but the setup cost was high. Therefore, for the best money-to-performance ratio, we determined that the software-based solution was the best option. Our project aims to automate the scoreboard, reducing the expenditure on VAR (Video Assisted Referee) and using our proposed method for points. We suggest using VAR only for verifying fouls. Our software approach uses OpenCV and deep learning techniques to track players and the score of each team. Overall, the proposed AI-based automated basketball scoring system represents a significant contribution to the field of sports analytics and has the potential to improve the experience of watching and analysing basketball games. The proposed system can also be extended to other sports and real-time monitoring scenarios.

2. Literature Survey and Findings

Automated stat monitoring, referee rule verification, and video annotation are popular in basketball today, with player tracking being the main feature. This paper presents strategies for player detection, team correlation, player monitoring, and arena to top-down view homographs [1]. It provides dependable detection and tracking in optimal conditions and can help teams learn about their opponents' plays, formations, and tactics. Real-time picture analysis could eliminate referees' human errors and improve rule verification. The paper offers a MATLAB software and procedures for tracking each player and projecting each location on a top-

* Muthaiah U
E-mail address: u_muthaiah@ch.amrita.edu

down homographic view of the basketball court, with satisfying player identification accuracy and prominent position monitoring accuracy. However, the present method may encounter issues like player track shifting, overlapping, or dropping.

In this research [2], they proposed an image-based Basketball Scoring Detection (BSD) system that combines object tracking with object detection. The You Only Look Once (YOLO) model is used to identify and locate the basketball hoop in real-time, and motion detection based on method called as frame difference which is used to identify any item motion around the hoop. The suggested BSD technique has satisfactory real-time accuracy, and tests on real-world basketball court videos show its precision. The technique has been implemented at several basketball courts in Beijing for intelligent basketball assessment devices. Automatic basketball scoring detection can simplify highlight clip editing, posting, storage, and distribution.

A technique to rebuild the course of the basketball in 3D space from a 2D plane [3]. They talk about camera calibration first, then they get projection matrix that is affine in nature. In this research, a morbid solution is found to the problem of estimating of the 3d information from the 2D values using the continuous points that make up to the basketball trajectory. The basketball's motion is analyzed and mapped to 3-D plane and basketball court scene are generated in the end.

A method for accurately analyzing the basketball trajectory is proposed by utilizing an image segmentation algorithm for basketball object extraction [4]. The algorithm models the basketball game's video frame to extract the basketball object for precise tracking in a video clip retrieval system. The algorithm uses the variance on the image block which makes them distinct, thus leading to form a training feature vector which achieves a segmentation accuracy rate of over 95.2%. The study concludes that this method is effective in accurately segmenting and extracting basketball objects in basketball videos.

In [5], the authors proposed a camera-based method for detecting and tracking basketball players during games, which accurately maps their locations onto a basketball court image. The method outperforms CNNs (convolutional neural networks) with fewer parameters and achieves high accuracy in player identification and tracking across various datasets. It also generates a heat map of players' locations based on team colors, surpassing other models like ICF, Faster RCNN, and SSD512. The technique is also tested on a pedestrian recognition dataset and shows competitive performance. The paper concludes that the proposed method is lightweight, efficient, and resilient compared to other advanced detection methods.

A method to predict the ownership of basketball using monocular view while only using the trajectories of a player. The study [6] rearranges using various data such as unary, pairwise and ball evidence to obtain a concise feature that can accurately predict the basketball owner by learning the patterns of multiple factors like geographical and momentum of players. The paper primarily focuses on basketball, using data obtained through tracking and applying permutations on the data by finding the set of permutation matrices to reduce up to maximum capacity of the total variance to entropy of the data. The approach also incorporates contextual and tactic features and integrates image-based detectors to enhance performance. The method demonstrates high accuracy and proposes to be improved by using sources such as a vision-based ball detector.

The paper [7] proposes a method for recognizing basketball scoring in AER sequences. The novel approach involves designing a model using spiking neural networks to retrieve active features and distinguish them on various stream events. There were problems on balancing of the dataset, which led them to use an imbalanced Tempotron algorithm. AER sequence dataset is used, segmentation into streams of events with both polarising samples. The proposed method outperforms existing models in performance. The dataset is collected using a DVS camera, which records dynamic events and consumes less memory. The authors aim to apply this framework to more time series datasets, combine more pruning methods, and extend their applications to neuromorphic chips for lower cost consumption.

A system for analysing long-shots in basketball games is proposed [8]. It uses methods like trajectory-based ball tracking from input video sequences to detect and track the basketball's throwing angle of shoot and velocity of flight path, enabling accurate shot analysis. The system generates candidate trajectories to reduce errors caused by occlusion, ball and camera motion, and other moving non-target in the linear perspective planes. The method also uses a segmentation method for identifying moving targets in the foreground, and feature-based pruning to reduce computational complexity. The system's effectiveness is demonstrated in dynamic background scenes with varying illumination conditions, and it is low cost as it does not need complex hardware like high-end cameras.

[9] was published prior to the [8] which had major improvements. [9] presents an algorithm that estimates the location of a ball using trajectory information. It detects objects in the shape of a sphere and generates a set of features for every frame, plotting a set of trajectories in which the candidates are estimated to the actual ball trajectory. It also introduces an effective detection method of the background plane that uses the value obtained from calculation of the median of background pixels to distinguish the moving objects in the foreground, even in the presence of occlusion and other objects in the scene. The algorithm can locate the ball even when it is occluded or overlapped with other objects in the frame, and it is straightforward to implement without requiring intricate calculations to determine the ball projection angle and speed. Simulation results demonstrated that the model was efficient enough to detect the presence of most moving objects in the background.

professional sports players and coaches are interested in analyzing high-level semantics from a different perspective. An advanced technique for analyzing basketball matches in broadcast videos by calibrating court-based cameras is proposed [10]. This new approach utilizes a quadrangle candidate generation algorithm and refined model fitting score to achieve precise calibration results. Player trajectories are then extracted using a Cam Shift-based tracking method and mapped to real-world court coordinates based on the calibration. This resulting player position and trajectory information is particularly useful for professional-oriented applications. The paper also showcases three practical applications to demonstrate the system's applicability, and experimental results show the effectiveness and robustness of the proposed calibration and tracking algorithms.

In paper [11,] they suggest that repeatable practice and appropriate feedback are crucial for individual improvement in basketball. To aid in this process, a basketball tracking system can recognize and track the ball during practice sessions. The collected data can then be analyzed to provide insights on missed and successful shots, progression under practice, and personalized analysis report for improvement. This system enables athletes to focus on their technique and practice, without having to worry about data collection and analysis.

The paper [12] proposes a method to automatically detect handball players performing specific techniques in training footage, which is a challenging task due to cluttered scenes, poor illumination, and numerous players of varied sizes and distances from the camera. To detect active players, the proposed method combines the YOLO object detector with two activity measures based on spatial-temporal interest points (STIPs) and optical flow (OF). One advantage of this method is that it can be directly applied without additional model learning, but manual threshold adjustments are needed. The accuracy of the activity measures depends heavily on precise player detection, as motion-related features are only considered within the bounding box of a detected object.

A novel method for simulating patch motion in single target tracking is proposed that is easily able to adapt to many trackers which are part-based [13]. Adaptive Local Movement Modelling (ALMM) method is suggested which can easily adapt to many part-based trackers and simulate the spatial distribution of image patch reliability and the tracked object. A model has been constructed using Gaussian Mixture Model (GMM) to patch mobility with object's gravity center, and is coupled with a basic tracker in a framework that is boosted to create a new integrated classifier for tracking, allowing for anomaly detection in patch motion. The method has configurable options for adjusting to changes in the object and number of patches used. Experimental results on common datasets demonstrate significant improvements in tracking performance over modern trackers.

In this paper [14], a detailed analysis of various geometric and semantic visual features for effectively detecting and tracking basketball players. To demonstrate that a well-defined tracking system can be created utilizing Deep Learning features without the inclusion of features like proximity, color similarity, or camera stabilization methods, a study is being done. The suggested tracker contains two steps: detection, which estimates players' poses using a pre-trained deep learning model, and tracking, which makes use of pose and semantic data from the result of a convolutional layer in a VGG network. For a basketball dataset with over 10,000 cases, the tracker's performance is assessed in terms of MOTA.

The authors of [15] introduce a real-time multi-person recognition and monitoring system that finds and follows players in video clips using YOLOv2 and SORT. The framework was tested on a subset of the NCAA Basketball Dataset, and a two-layer LSTM was trained for action recognition. They introduce a real-time framework that detects and tracks basketball players using YOLOv2 and SORT. The performance was evaluated using the PASCAL VOC criteria, and accurate results were obtained.

While the papers dealt with ball detection [8][9], object tracking & segmentation [4][13], player detection & tracking [1][5][6][10][15], hoop detection [7] and shot detection [2][3][7][11]. While most of the papers used various technologies to execute the previously mentioned models and attain results and accuracy scores. Most shot detection models used static camera. We wanted our model to be versatile and work in any mode, static or moving. We found that there were no articles which dealt with the identification of the score of the shot made. Also, Complete Automation till score detection was not employed. We referenced previous researches on player detection, pose detection, hoop detection and shot detection to integrate all those and thus also leading us to propose a novel approach to detect the score of the shot made by the player regardless of the camera movement.

3. Dataset

A large set of basketball dataset was taken from ImageNet which is an image database open for public use worldwide. They were manually labelled and split into training, testing, and validation sets in a 70:20:10 ratio. These images were recorded from moving cameras during live matches and used as training data for creating a model capable of detecting nets and basketballs in real-time. The model's parameters were adjusted through an iterative process using an optimization algorithm to minimize the difference between the predicted outputs and true labels. A labelling tool was used to label all features in the frames, and their coordinates were stored in a JSON file. The model's performance was tested using the validation dataset, and parameters were fine-tuned to reduce overfitting. The trained model was then used to make predictions on new basketball videos in the proposed methodology. The raw footage was trained to form an input model for the proposed model.

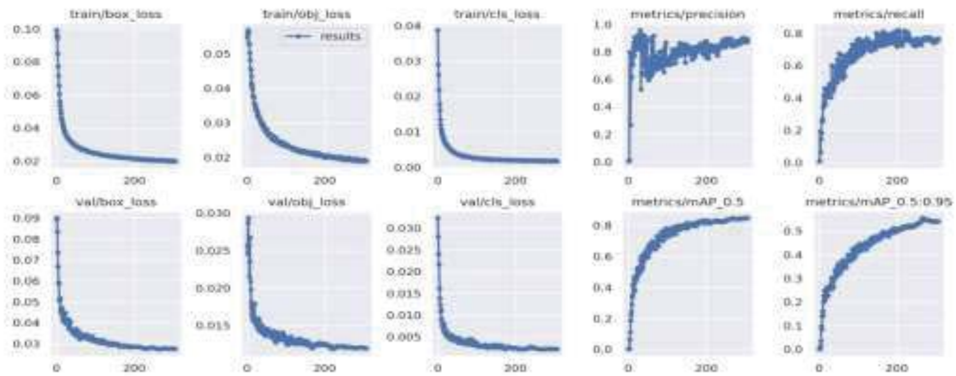


Fig. 1. Graphs depicting training loss for labelled data.

4. Proposed Model

The proposed method is a multi-step approach to analysing a video. It involves object detection and tracking algorithm, to detect and track objects in the video. Shot detection is performed by calculating the intersection area of consecutive frames. Past frame processing is used to reverse the frames, enabling the algorithm to detect the shooter. Then the next step involves person and pose detection of the shooter. Euclidean distance is calculated between the shooter and the target with a score is generated based on the frames of the shot.

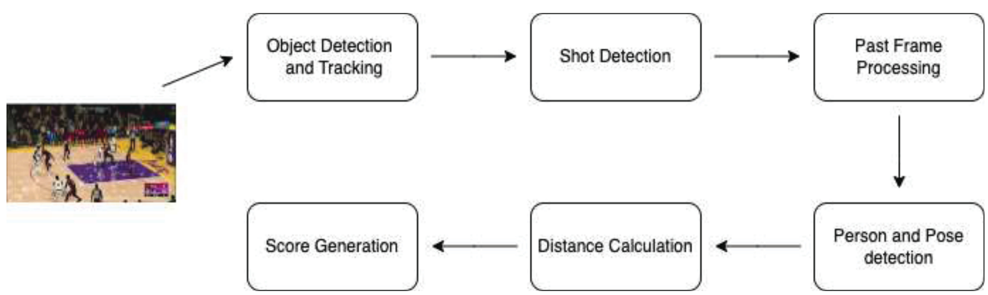


Fig. 2. Model Architecture

4.1. Object Detection and Tracking

The first step in the score detection algorithm for basketball is to detect and track objects like basketball and hoop. To detect and track basketball and hoop in the score detection algorithm, YOLOv4 (You Only Look Once Version 4) is used. This real-time object detection system detects and labels multiple objects like hoop, basketball in an image or video frame. Once trained, the model is used to detect and track the objects in the video by analysing each frame, outputting their location as bounding boxes, and updating their position across consecutive frames. This information, combined with pose detection data, helps determine the player's position relative to the objects. Thus, it allows us to accurately detect and track these objects in the video, which is shown in Figure 3.

4.2. Shot Detection

To detect successful shots in basketball, overlapping bounding boxes of the basketball and hoop indicate that the ball has entered the goal. The overlap is determined by comparing the coordinates of the two bounding boxes. The system uses a function that computes the intersection area between two bounding boxes, which takes two parameters: boxA and boxB. The function calculates the (x, y)-coordinates of the intersection rectangle and then the intersection area as the absolute value of the product of the width and height of the intersection rectangle. The function calculates only the intersection area and not the union or IOU (Intersection Over Union) between the two bounding boxes. The shot detection can be seen in Figure 4.

$$interArea = \left|\left(max\left(\left(B_x - A_x, 0\right)\right) \times max\left(\left(B_y - A_y\right), 0\right)\right)\right| \quad (1)$$

where (Ax, Ay) and (Bx, By) are 2D coordinates of bounding boxes present in hoop and basketball in Eq. (1), respectively.

4.3. Past Frame Processing

After shot detection, the next step is to start past frame processing. To determine the trajectory and success of a shot in basketball, the system uses YOLOv4 to detect the basketball and hoop in the current frame, followed by processing of past frames. During this step, the frames leading up to the shot are stored in a buffer memory, and when a shot is detected, the processing of buffer frames begins in reverse order until the ball crosses the hoop. This step helps determine the ball's trajectory, shot release point, and player position during the shot. To optimize the PyTorch model and enable real-time inference on the video frames, the model can be converted to TensorRT format using ONNX. The reversed frames are then rendered using OpenCV. By analysing frames in reverse order, the system can determine shot success and type, providing more accurate and reliable scoring. This helps to ensure that the system accurately detects and tracks the ball's trajectory, allowing for more accurate and reliable shot detection and scoring.

4.4. Person and Pose Detection

To accurately identify the player holding the ball and detect their pose, we use a combination of person detection and pose detection models. Firstly, YOLOv7 is used for person detection to identify the player shooting the ball. It returns bounding box coordinates of the shooter, which the pose detection model uses to detect their pose. Detecting the body position and orientation helps determine the shooting motion and trajectory of the ball, and whether it passes through the hoop. The coordinates from YOLOv7 are used as input to accurately detect the player's pose, enhancing the accuracy of the score detection algorithm. After person detection, the pose of the player holding the ball is detected to determine their body position and orientation. MMPose is a pose detection model that is used here to detect key points of the shooter's body and estimates the position and orientation of each joint in 3D space using CNNs and graphical models. It is trained on a large dataset of annotated images and videos of basketball players, enabling real-time detection of complex poses. Once the player's pose is detected, it is used to determine the shooting motion and trajectory of the ball. This step is critical to the score detection algorithm and helps determine the shooter and appropriate score for the shot made.

4.5. Distance Calculation

The next step is to calculate the distance between the ball and the palm of the hand of the player who shoots. To calculate it, Euclidean formula is used here. This determines if the shot is within scoring range. The ball and palm positions are obtained from a pose detection model, with the ball represented by 2D coordinates (x,y). The palm's position is determined by detecting key points of the hand and calculating the Euclidean distance between them. The distance between the ball and the palm is determined by applying the Euclidean formula, and if it falls within a certain range, the shot is valid, and the appropriate score is awarded. Accurately determining the shot's scoring range relies on this crucial distance calculation step.

$$distance = \sqrt{\left(\left(x_2 - x_1\right)^2 + \left(y_2 - y_1\right)^2\right)} \quad (2)$$

where (x1, y1) and (x2, y2) are the 2D coordinates of the ball and the palm of the hand in Eq. (2), respectively.

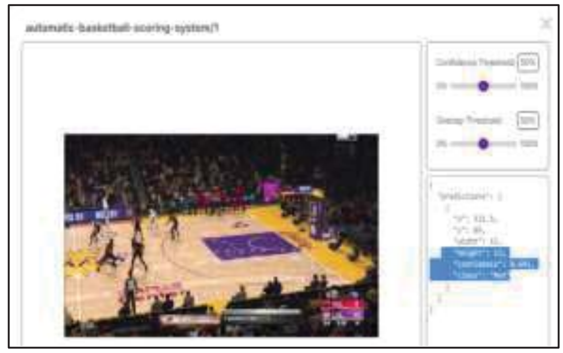
4.6. Score Generation

The shooter's distance from the ball is calculated and the frames captured during the shot are analysed from the buffer memory. The number of frames in the *shot_n* folder is used to determine if it is a two-pointer or a three-pointer in basketball. A three-pointer is scored when the ball is shot beyond the three-point line and goes through the hoop, which takes longer to travel, resulting in more frames captured. To determine the score, the number of frames is compared to a predetermined value of 25, which corresponds to the minimum time of flight required for a three-pointer. If the number of frames exceeds 25, it is a three-pointer, and if it is less, it is a two-pointer. The corresponding score is then allotted to the team, determining the scores earned by both teams, as shown in Figure 6.

5. Experimental Results

The proposed system utilizes Visual Studio Code for the Python IDE and various libraries such as TensorFlow, Fast RCNN, OpenCV, Yolov4, Yolov7, and mmpose tool. It shows promising results in generating basketball scores automatically by accurately detecting and tracking the basketball, hoop, and players in real-time. The system generates output pictures with bounding boxes around the detected objects, enabling easy identification of system components with average accuracy of 78.74%.

(a)



(b)

Fig. 3. (a) Basketball and (b) Hoop detection

The shot detection algorithm was able to accurately identify when a shot was taken, with average accuracy of 83.75%. This ensured that the system only generated scores when a shot was successful, minimizing false positives. The person and pose detection algorithms were also effective in identifying and tracking the players, with good accuracy rates. The score generation module of the system was able to generate scores for each successful shot made, based on the detected objects and their respective positions. This provided a convenient and automated way of keeping score, which can be useful for coaches, broadcasters, and fans.

Fig. 4. Shot Detection

Fig. 5. Person and Pose detection

Fig 6. Shot generation

The average accuracy of the pose detection algorithm was 78.33%, which enabled us to easily identify the shooter, generating close to an average of 85.52% accuracy for all the shots made in the video. The graph depicted in figure 7 shows the accuracy of our model in detecting and classifying objects over different set of input video lengths, the accuracy was determined by the results mentioned in the table 1. The graph depicted in figure 8 shows the number of actual shots being made and the shots counted over different matches along with its accuracy which was determined by the results mentioned in the table 2. The graph depicted in figure 9 shows the accuracy of identifying the shooter over different matches, which was calculated from results mentioned in the table 3. The graph depicted in figure 10 shows the accuracy in tracking the correct frame of the shooter shooting the ball, which was calculated from the results mentioned in the table 4. The graph depicted in figure 11 shows the accuracy of scores provided to the teams, this was calculated from the cumulative results present in the table 5.

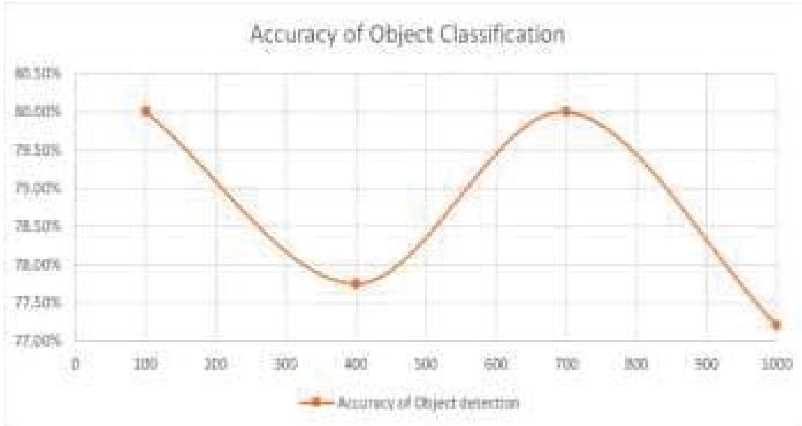
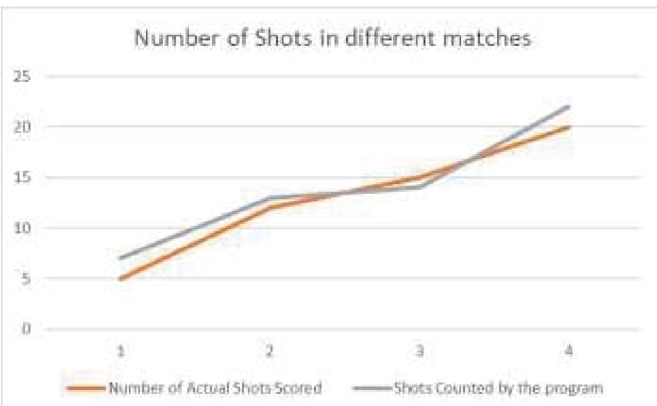
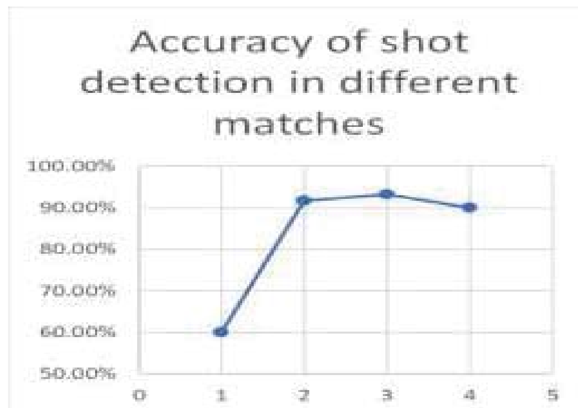


Fig. 7. Accuracy of the Object Classification



(a)



(b)

Fig. 8. (a) Accuracy of detecting and (b) counting of the shot.

Fig. 9. Accuracy of Shooter Identification

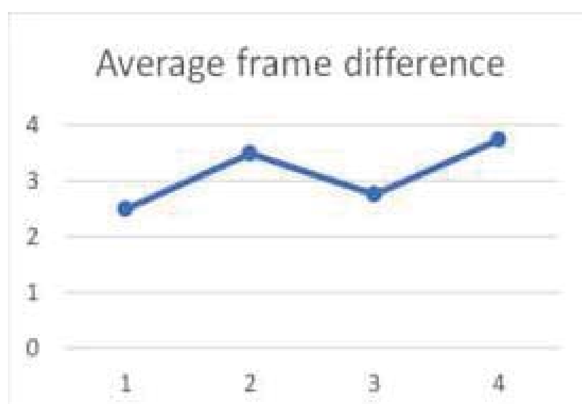


Fig. 10. Accuracy of correct shoot frame

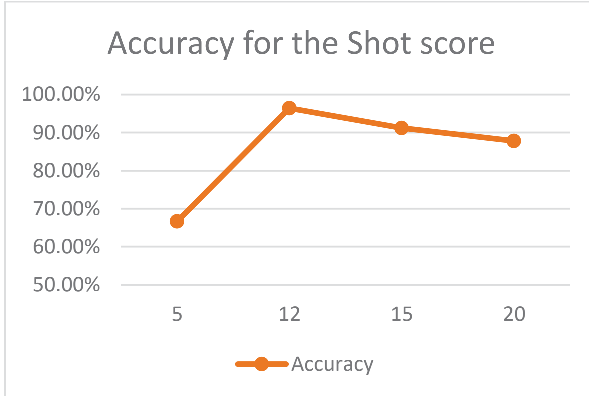


Fig. 11. Accuracy of the scores assigned for each shots.

Table 1. Accuracy of the Object Classification

Number of frames	Total Number of classes	Identified Classes	Number of frames object detection failed	Accuracy of Object detection
100	3	3	20	80.00%
400	3	3	89	77.75%
700	3	3	140	80.00%
1000	3	3	228	77.20%

Table 2. Accuracy of Detecting shot

Match Number	Number of Actual Shots Scored	Shots Counted by the program	Accuracy
1	5	7	60.00%
2	12	13	91.67%
3	15	14	93.33%
4	20	22	90.00%

Table 3. Accuracy of Identifying the correct shoot frame

Shots made	Average frame difference
5	2.5
12	3.5
15	2.75
20	3.75

Table 4. Accuracy of Identifying the Shooter

Match Number	Number of Shots made	Number of shoots identified	Number of correct shooter identification	Shooter Identification Accuracy
1	5	7	5	60.00%
2	12	13	11	83.33%
3	15	14	11	80.00%
4	20	22	20	90.00%

Table 5. Accuracy for the scores assigned for each shots.

Number of Shots Scored	Number of 2 Pointers	Number of 3 pointers	Number of shots counted	Counted 2 pointer	Counted 3 pointer	Actual Score	Counted Score	Accuracy
5	3	2	7	5	2	12	16	66.67%
12	8	4	13	10	3	28	29	96.43%
15	11	4	14	11	3	34	31	91.18%
20	19	1	22	20	2	41	46	87.80%

6. Conclusions and Future Work

The proposed method tackles some of the key issues with basketball scoring systems, including false detections, inaccurate player placement, and the ability to discern between in-bounds and out-of-bounds occurrences. The experimental findings demonstrate how well the suggested system tracks player positions and detects scores with exceptional accuracy and efficiency. The accuracy is 85.52% for detecting correct scores, The proposed method can also assist coaches and teams in better understanding their performance, making data-driven decisions, and enhancing their overall strategy. Future work on this project involves optimizing our model to execute the frames on real-time with no delay. Our project inputs video as input as of now, we are aiming to deploy the model on broadcasting footages which required us to process the frames through threads for faster analysing on real-time.

References

1. Parsons, Scott, and Jason Rogers. (2014). “Basketball Player Tracking and Automated Analysis.”
2. Xu-Bo Fu, Shao-Long Yue, De-Yun Pan. Camera-based Basketball Scoring Detection Using Convolutional Neural Network. International Journal of Automation and Computing, vol. 18, no. 2, pp.266-276, 2021. https://doi.org/10.1007/s11633-020-1259-7
3. Wu Lifang, Shen Wei, Han Xiuli, Meng Xianglong and Liu Chao, "A method of the basketball trajectory reconstruction," *2008 9th International Conference on Signal Processing*, Beijing, 2008, pp. 1360-1363, doi: 10.1109/ICOSP.2008.4697384.
4. Huachen Zhu, Long Liu. (2022). Basketball Object Extraction Method Based on Image Segmentation Algorithm. *Security and Communication Networks, 2022, 3021682.*
5. P. K. Santhosh, B. (2019). An Automated Player Detection and Tracking in Basketball Game. *Computers, Materials & Continua, 58(3), 625–639.*
6. X. Wei, L. Sha, P. Lucey, P. Carr, S. Sridharan and I. Matthews, "Predicting Ball Ownership in Basketball from a Monocular View Using Only Player Trajectories," *2015 IEEE International Conference on Computer Vision Workshop (ICCVW)*, Santiago, Chile, 2015, pp. 780-787, doi: 10.1109/ICCVW.2015.106.
7. J. Shen, Y. Zhao, J. K. Liu, and Y. Wang, "Recognizing Scoring in Basketball Game from AER Sequence by Spiking Neural Networks," *2020 International Joint Conference on Neural Networks (IJCNN)*, Glasgow, UK, 2020, pp. 1-8, doi: 10.1109/IJCNN48605.2020.9207568.
8. B. Chakraborty and S. Meher, "A trajectory-based ball detection and tracking system with applications to shooting angle and velocity estimation in basketball videos," *2013 Annual IEEE India Conference (INDICON)*, Mumbai, India, 2013, pp. 1-6, doi: 10.1109/INDCON.2013.6725963.
9. B. Chakraborty and S. Meher, "Real-time position estimation and tracking of a basketball," *2012 IEEE International Conference on Signal Processing, Computing and Control*, Solan, India, 2012, pp. 1-6, doi: 10.1109/ISPCC.2012.6224370.
10. M. -C. Hu, M. -H. Chang, J. -L. Wu and L. Chi, "Robust Camera Calibration and Player Tracking in Broadcast Basketball Video," in *IEEE Transactions on Multimedia*, vol. 13, no. 2, pp. 266-279, April 2011, doi: 10.1109/TMM.2010.2100373.
11. Širmenis, Juozas; Lukoševičius, Mantas, “Tracking Basketball Shots – Preliminary Results,” in Proceedings of the 26th international conference on information society and university studies (IVUS 2021), Kaunas, Lithuania, April 23, 2021, CEUR-WS. 2021, vol. 2915, art. no. 21, p. 181-190. ISSN 1613-0073.
12. M. Ivasic-Kos, M. Pobar and J. Gonzàlez, "Active Player Detection in Handball Videos Using Optical Flow and STIPs Based Measures," *2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS)*, Gold Coast, QLD, Australia, 2019, pp. 1-8, doi: 10.1109/ICSPCS47537.2019.9008460.
13. B. Zhang, Z. Li, A. Perina, A. D. Bue and V. Murino, "Adaptive Local Movement Modelling for Object Tracking," *2015 IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, HI, USA, 2015, pp. 25-32, doi: 10.1109/WACV.2015.11.

(ICIEV-ISCMHT), Himeji, Japan, 2017, pp. 1-6, doi: 10.1109/ICIEV.2017.8338571.

14. Arbués-Sangüesa, A., Ballester, C., & Haro, G. (2019). Single-Camera Basketball Tracker through Pose and Semantic Feature Fusion.
15. Acuna, D. (2017). Towards Real-Time Detection and Tracking of Basketball Players using Deep Neural Networks.

Key Frame Extraction from Video Sequence: On-The-Fly A Comparison and Analysis

Swati J. Patel[a,b], Mehul C. Parikh[b]

[a] Information Technology Department, L.D.College of Engineering, Ahmedabad, Gujarat, India
[b] Gujarat Technological University, Ahmedabad, Gujarat, India

Abstract

As a result of the digital revolution, media on the Internet has completely taken over the world. Digital video may now be made available to everyone because of the growing use of internet services, declining prices for digital storage devices, and the advent of 4-G technology. An extensive collection of videos is constantly growing, and it always takes time to analyze such a significant amount of data. Several still images called frames make up the video sequence. A video has a lot of information, and because of this, the frames often contain extraneous and identical data that is pointless if the film's content is a concern. The efficient processing of video content requires a practical and informative presentation. It is crucial to select pertinent and informative content from videos automatically. By removing replications and extracting important frames from the video, keyframe extraction is considered appropriate for thorough video analysis. A key frame is a representative frame that includes the facts of the video collection, representing vital information from the video. It is not only helpful to recognize the whole video but also can reduce the processing time, computational costs, and storage requirements of each video sequence in various applications. Extraction of these frames is one of the essential tasks in video processing. This paper presents different methods of key-frame extraction proposed in the past.

Keywords: Key-frame; Video Processing; Video Key Frame; Key Frame Extraction

1. Introduction

With the development of video recording devices such as smartphones, portable cameras, surveillance equipment, and others, video capturing, sharing, and creating becomes a straightforward process. The amount of video data has been explosively increasing. Due to the tremendous use of digital media over the Internet in information, education, entertainment, business, and surveillance, video processing has become a popular research topic in image processing [1]. It is not advisable to process a whole video sequence as a video composed of many frames at a frame rate of at least 24 frames per second (fps) for high-definition video. So it would be best if the methods could extract important frames in a video sequence that would be sufficient to represent the video and could be used to recognize the whole video sequence [2]. Key-frames provide a quicker view of video content and help reduce the computational complexity for various video analysis and retrieval applications. The video is recreated using extracted Keyframes. [3]. Keyframes are the fundamental building blocks for multiple tasks, including video browsing, summarization, searching, understanding, and chapter titles in DVDs. They are also used in numerous applications, including surveillance, medical, underwater, web browsing videos, sports and news programs, indoor and outdoor videos, and surveillance.

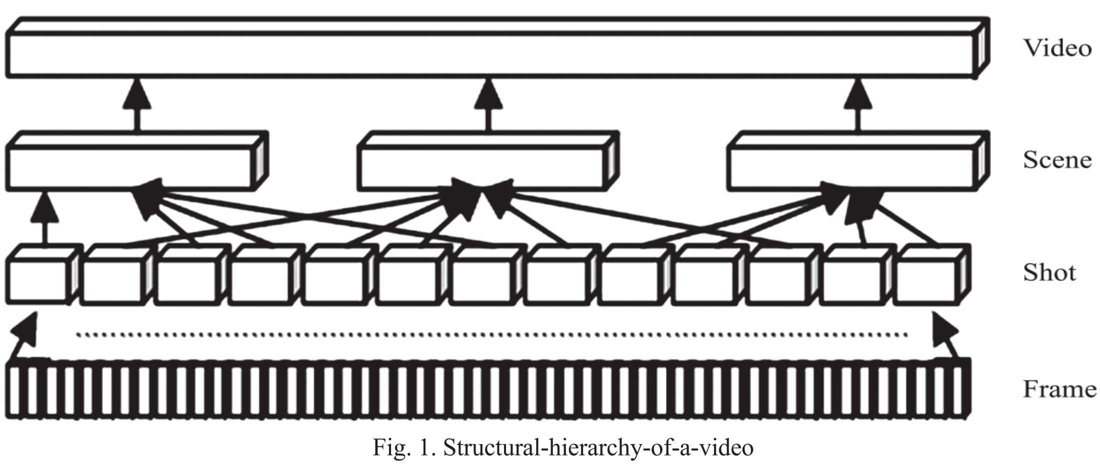


Fig. 1. Structural-hierarchy-of-a-video

Video is an immense volume of data objects containing highly redundant and insignificant information. As shown in Figure 1, the video has a complex structure consisting of scenes, shots, and frames [5]. A shot is a consecutive, adjacent sequence of frames captured by a single camera in continuous action. The key Frame is the video part representing a visual summary and meaningful information about the sequence [6]. The key Frame must contain the high-priority entities and events of the video and be free of repetition and redundancy [7]. Video processing is essential in many applications, including watermarking, various scene segmentation, shot boundary detections from that scenes, and key-frame extraction from that shots. Key-frame is either a frame or a set of frames that represent the entire content of the video clip. It refers to the image frame in the video sequence, which is representative and can reflect the summary of video content. It must contain most of the salient features of the represented video clip [3]. The concept of key-frame extraction focuses on the most specific part of a video sequence and is selected such that the video can be reproduced using the key-frames. Depending on the content complexity of the shot, one or more key-frames can be extracted from one single shot. A shot is defined as an unbroken sequence of frames recorded from a single camera, which forms the building blocks of video. In video data that contains multiple shots, it is necessary to identify individual shots for key-frame extraction [8]. Selecting key-frames from a video is a ranking process of unique frames regarding their representativeness to the video [3]. One can express the main content of video data clearly and reduce the amount of memory needed for video data processing and complexity by using the keyframe. Three properties must be considered when selecting key-frames: continuity, priority, and repetition. Continuity means that the video must be as uninterrupted as possible. Priority means that particular objects or events may be more critical than others based on a given application, and thus the key-frame must contain high-priority items. It is a highly task-dependent property. Repetition means that it is essential not to represent the same events repeatedly. It is often challenging to successfully incorporate these semantic properties [9].

2. Classification of Key Frame extraction method

2.1. Uniform Sampling Method

Uniform sampling is the most common method to extract the key Frame. In this method, every kth Frame from the video sequence is extracted, where the predefined value of k is decided from the length of the video. If the video sequence is more, then the value of k is large, else is kept small. On average, 5% to 15% from the original video, the total number of key-frames should be extracted. This concept is straightforward and does not have semantic relevance [10] [7]. As it is based on the predefined fixed value, these approaches are not content-based and do not consider the dynamics of the visual content, and selected frames are often unstable [11].

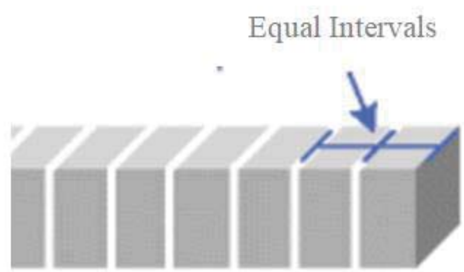


Fig. 2. Uniform sampling

2.2. Pixel Compare Method

In this method, every consecutive Frame is compared pixel-wise, and when the comparison difference crosses a given threshold value, the system identifies that Frame as a keyframe. This method is highly time-consuming and too sensitive to the motion of objects in Frame [12].

2.3. Image Histogram Method

Image histogram gives us the number of pixels for specific brightness values rated from 0 to 256. It contains essential information about an image and hence can be utilized to extract the keyframes. In this method, the histogram of each Frame is calculated and, based on the difference between the two frames, can decide the dissimilarity between them. If the histogram of two consecutive frames is either 50% or more dissimilar, then we can extract that Frame as a keyframe [7].

2.4. Scale-Invariant Feature Transform

Scale-invariant feature transform (SIFT) is used for feature detection to detect and describe the local features in an image. It is the most important method used for local features in computer vision applications. SIFT feature descriptor is invariant to uniform scaling, orientation, illumination changes, translation, and rotation and partially invariant to affine distortion. So we can use SIFT features for key frame extraction. Important locations are first defined using a scale space of smoothed and resized images and applying the difference of Gaussian functions on these images to find the maximum and minimum responses. Non-maxima suppression is performed, and putative matches are discarded to ensure a collection of highly interesting and distinct collection of key points. A histogram of oriented gradients is performed by dividing the image into patches to find the dominant orientation of the localized key points. These key points are extracted as local features [7 13]

2.5. Cluster-Based Method

Clustering is a popular technique for the keyframe extraction method. Clustering algorithms can automatically classify video data according to their similarity. In this method, key frame clusters are created using the data points and various features of video sequences. The set of keyframes is created with frames that have the closest distance from the center of the cluster. The advantage of this method is that it covers the global characteristics of the scene. Still, it requires a high computational cost for cluster generation and feature extraction from the scene [12]. The main drawback of these methods is that, depending on the number of clusters, keyframes can be either redundant or fail to represent the content of the whole shot efficiently.

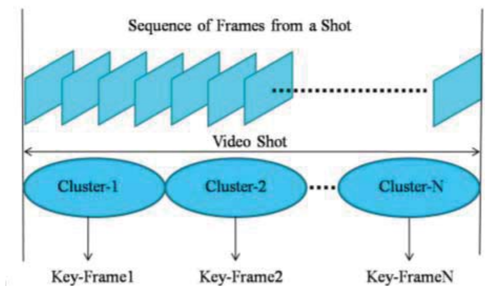


Fig. 3. Cluster-based [14]

2.6. Shot-Based Method

One technique for identifying the significant changes in the video's content is shot boundary detection. The keyframe extraction is done by extracting a keyframe per shot. The number of keyframes used to abstract a shot should be compliant with visual content complexity within the shot, and the placement of keyframes should represent the most salient visual content. The shots in the video are divided into sub-shots. For each sub-shot, entropy is calculated, and the extraction of the key Frame in each shot is based on the maximum entropy value of each shot. However, this method has the drawback of not including the content complexity and is also not appropriate and accurate for a video which is having a big shot [6].

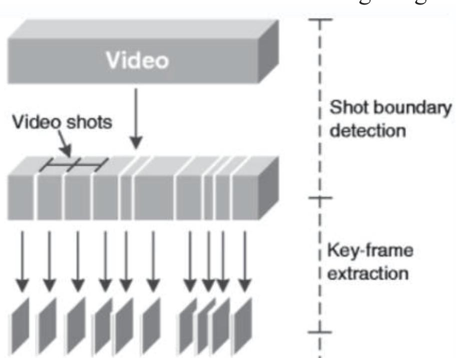


Fig. 4. Shot based

2.7. Content- Analysis-Based Method

In this method, keyframes are selected based on the color, texture, and other valuable visual information of each Frame. All the frames of the video in which this information is changing significantly are considered as the keyframes. First of all, the first Frame is selected as a new frame, and the next subsequent frames are compared with this reference frame. The kth Frame becomes a new reference frame if the distance between kth and the reference frame exceeds the predefined threshold value. This method selects the keyframe based on the degree of change in the content of the Frame. It is very insensitive to camera movement and hence produces unstable and very poor efficient keyframes [12]

2.8. Motion-Based Method

This motion-based method first segments an input video clip into homogeneous parts based on major types of camera motion, e.g., pan, zoom, pause, and steady, and dedicated rules are used to extract keyframes from each segment. Movement in the video shots can easily be detected or analyzed by analyzing the optical flow of the video sequence. In this method, the local minimum in the movement is considered the keyframe. One of the drawbacks of this method is its low robustness, as this method depends on the local information and does not count the global information for the keyframe extraction [15].

2.9. Sparse Representation-Based Method

In this sparse representation-based method, video frames are projected to a low dimensions feature space using a random projection matrix, and sparse representation is exploited in the random feature space to analyze the Spatiotemporal information of the video data and generate keyframes [15]. This approach does not require shot(s) detection, segmentation, or semantic understanding and is computationally efficient.

3. Comparison of Key Frame extraction method

Table 1. compares keyframe extraction methods based on their characteristics, merits, and demerits.

Method	Characteristics	Merits	Demerits
Uniform Sampling	Most common method	Straightforward method	Not content-based and Selected frames are often unstable
Pixel Compare	Pixel-wise comparison	Easy to evaluate	Time-consuming
Image Histogram	Similarity measure between keyframe	High-level segmentation	Don't consider the local similarities
Scale-invariant feature transform	Describe the local features in an image	Most prominent local feature	-
Cluster-based	Clustering similar frames/shots	Covers global characteristics of the scene	High computational cost (Takes 10 times the video length)
Shot based	keyframe in each shot is based on the maximum entropy value of each shot	-	Not appropriate for a big shot
Content-analysis based	Keyframes extraction based on the degree of change in the content of the Frame	Maintain good segmentation results	Insensitive to camera movement.
Motion-based	Adopts advantage of the digital capture device.	Reduce the spatiotemporal effects	High-quality video expected

4. Prior Art of Key frame extraction method

The earlier studies done in the key frame extraction field are briefly summarized in this section.

The first clustering-based keyframe extraction algorithm was published in 1998 by Zhuang et al [16]. The number and size of the clusters are used to determine which keyframes to use. A cluster is formed by comparable visual frames, and the cluster's visual content may include colour, texture, and shape. This method is more efficient, quick to compute, and simple to apply to online processing. This method was tried on two films: a romantic comedy(movie 1) and an action movie (movie 2). The second movie has more keyframes than the first.

A frame from the collection that is frequently distinct from its succeeding neighbor is chosen as the key Frame. The fuzzy C-means clustering algorithm is used to group the visually comparable frames into one group. After clustering, frames with change ratios—a metric for content variation—that is higher than the average value of the cluster are handled as Keyframes. This technique was tested using available video datasets, football footage, and sports videos on YouTube [17].

By fusing the key aspects of the video, Jiaxin Wu et almethod .'s [18] allows comparable frames to cluster together. Pre-sampling is used in the first stage to decrease the video frame's redundancy and generate candidate frames. In order to represent the visual content of candidate frames, the BoW (Bag of Words) model is used. Lastly, the VRHDPS (Video Representation based High-

Density Peaks Search) clustering technique groups candidate frame data into groups. As a keyframe, the central value of each cluster is compiled.

Keyframe extraction requires two phases, according to Besiris et al. [19]. The creation of the MST (Minimum Spanning Tree) graph, in which each node is connected to a single frame of the shot, comes first. The keyframes are extracted in the second stage using the maxim speed approach. The number of selected keyframes is controlled by an adaptively set threshold.

On the basis of spatial and temporal color distribution, Zhonghua et al [20].'s research focused on video keyframe extraction. First, a frame is built during the video shot that takes into account the spatial and temporal distribution of the pixels. The shot calculates the weighted separation between each Frame's color histogram. As keyframes, they choose the frames that are closest to the distance curve's peaks.

According to Spyrou et al. [21], keyframes are taken from video clips based on their semantic context. Keyframe regions are used to extract the color and texture features. Each Frame's local region thesaurus is created using a hierarchical clustering method. Each photo contained a local extraction of the visual thesaurus.

Each video frame is assigned a collection of features. Semantic features and frame-based features are among the features. Semantic characteristics pinpoint the likelihood of the Frame's semantic concepts. Each Frame in each segment of the video is connected to at least one of the semantic attributes. For each collection of frames, a score is generated based on the semantic value. Lastly, the score value is used to choose the representative frame [22].

Keyframe extraction was created by Ling Shao et al. [23] based on intra-frame and inter-frame motion histogram analysis. Keyframes are extracted from the frames that contain complicated motion and are more significant than their neighboring frames. It includes more of the video's actions and activities. Finding peaks in the entropy curve that is produced using the motion histograms in each video frame is the first step in initializing the keyframes. The peaked entropies are weighted using inter-frame saliency, which employs histogram intersection and results in the creation of final keyframes. The foreground objects' motion complexity maxima and the variance in motion between subsequent frames are used by this approach to extract keyframes.

The keyframe extraction approach that was performed hierarchically to produce a keyframe with a tree-structured was discussed by Hyun Sung Chang et al. in their study [24]. There are a lot fewer frame comparisons as a result. It creates the video's multilevel abstract. By utilizing the depth-first search technique with pruning, it offers an effective content-based retrieval.

Keyframe extraction and object segmentation are concurrently built by a unified feature space, according to Xiaomu Song and his colleagues [25]. The keyframe selection is articulated as a feature selection inside the context of the Gaussian Mixture Model (GMM) for object segmentation. In this case, keyframes are extracted using two divergence criteria. Maximizing pairwise interclass divergence between GMM components is one strategy. The next step is to maximize the marginal divergence, which determines how the mean density varies within frames. With this method, the representative keyframes for object segmentation are extracted. Combining keyframes and objects allows for the performance of this content-based video analysis. This scheme demonstrates an integrated content-based video analysis that offers a novel and adaptable functionalization of frames and objects.

The entropy difference approach was investigated by Markos Mentzelopoulos et al., [26] in an effort to segment spatial frames. The entropy that the dominating item possesses can be used to extract the keyframe. When the object can be distinguished from the backdrop, this work produces good results. Yet, when transient changes like flashes happen, performance suffers.

Keyframe attributes like texture, edge, and motion were used to analyze the content-based video indexing and retrieval. Keyframes were retrieved using clustering techniques based on K-means. Comparing this method's effectiveness to the Volume Local Binary Pattern (VLBP)[27].

To modify the fundamental properties of human motion capture data, Joint Kernel Sparse Representation was developed to capture human motion data for keyframe extraction. No matter how motions are captured, this method models the sparseness and Riemannian manifold structure of human motion, which are two crucial characteristics of motion data. The internal structure of the motion capture data can be obtained by joint representation. Moreover, the triangle restriction ensures that keyframe extraction is valid locally, particularly for periodic motion sequences. As compared to other state-of-the-art methods, the experimental results are favorable [28].

The video frames are projected to a low dimensional random feature space, and keyframes are recovered based on sparse representation from creating consumer films. The author used the notion of sparse signal representation to evaluate the spatial and temporal information of the video and produce keyframes. Shot detection, segmentation, or semantic comprehension are not necessary with this technique [15].

The keyframe selection process takes into account local features based on a key point-based architecture. Based on the two obvious parameters of coverage and redundancy, the appropriate keyframes are chosen. One of the promising keyframe extraction techniques is this one [29].

In order to extract keyframes, Badre et al. [30] described the Haar wavelet transform with different levels and thepade's sorted pentnary block truncation coding. The Alias Canberra distance, Sorensen distance, Wavehedge distance, Euclidean distance, and mean square error similarity measurements are used to measure variety among successive frames.

5. Summary and Conclusion

The key frame extraction procedure eliminates the majority of the unnecessary frames from the video and is regarded as a fundamental unit in the structural analysis of the video. An accurate representation of the complete shot is given to the user. It's quite important in a lot of different areas, including video summarization, content-based video indexing and retrieval, video searching, video compression, and many more. This paper provides a thorough analysis of the methods used to locate the Key Frames, their benefits and drawbacks, and the challenges that a user encounters when attempting to extract the Key Frame. Although there are no standard metrics for evaluating Key Frames extraction methods, these approaches should be highly effective, reliable, and computationally simple, and the extracted Key Frames must be as small as feasible and reflective of the full video's sequence frame. Depending on the use for which it is intended, a particular way may be best. An advanced strategy for key frame extraction is the cluster-based approach.

References

1. Ma M, Mei S, Wan S, Hou J, Wang Z, Feng DD. Video summarization via block sparse dictionary selection. *Neurocomputing*. 2020;378:197-209. doi:10.1016/j.neucom.2019.07.108
2. Elahi GMEM, Yang YH. Online learnable keyframe extraction in videos and its application with semantic word vector in action recognition. *arXiv*. 2020.
3. Mei S, Guan G, Wang Z, Wan S, He M, Feng DD. Author ' s Accepted Manuscript reconstruction Video Summarization via Minimum Sparse Reconstruction. *Pattern Recognit*. 2014. doi:10.1016/j.patcog.2014.08.002
4. Asha Paul MK, Kavitha J, Jansi Rani PA. Key-Frame Extraction Techniques: A Review. *Recent Patents Comput Sci*. 2018;11(1):3-16. doi:10.2174/2213275911666180719111118
5. Ali IH, Al-Fatlawi TT. Key Frame Extraction Methods. *Int J Pure Appl Math*. 2018;119(10):485-490.
6. Gawande U, Hajari K, Golhar Y. Deep Learning Approach to Key Frame Detection in Human Action Videos. *Recent Trends Comput Intell*. 2020:1-16. doi:10.5772/intechopen.91188
7. Jadon S, Jasim M. Unsupervised video summarization framework using keyframe extraction and video skimming. *2020 IEEE 5th Int Conf Comput Commun Autom ICCCA 2020*. 2020:140-145. doi:10.1109/ICCCA49541.2020.9250764
8. Sheena C V., Narayanan NK. Key-frame Extraction by Analysis of Histograms of Video Frames Using Statistical Methods. *Procedia Comput Sci*. 2015;70:36-40. doi:10.1016/j.procs.2015.10.021
9. Gianluigi C, Raimondo S. An innovative algorithm for key frame extraction in video summarization. *J Real-Time Image Process*. 2006;1(1):69-88. doi:10.1007/s11554-006-0001-1
10. Jeong D ju, Yoo HJ, Cho NI. A static video summarization method based on the sparse coding of features and representativeness of frames. *Eurasip J Image Video Process*. 2016;2017(1):1-14. doi:10.1186/s13640-016-0122-9
11. Barhoumi W, Zagrouba E. On-the-fly Extraction of Key Frames for Efficient Video Summarization. *AASRI Procedia*. 2013;4:78-84. doi:10.1016/j.aasri.2013.10.013
12. Kavita Sahu MSV. Key Frame Extraction From Video Sequence : A Survey. *Int Res J Eng Technol*. 2017;04(05):1346-1350. https://www.irjet.net/archives/V4/i5/IRJET-V4I5404.pdf.
13. Jin H, Yu Y, Li Y, Xiao Z. Network video summarization based on key frame extraction via superpixel segmentation. *Trans Emerg Telecommun Technol*. 2020;(February):1-11. doi:10.1002/ett.3940
14. Janwe NJ. Video Key-Frame Extraction using Unsupervised Clustering and Mutual Comparison. 2016;(10):73-84.
15. Kumar M, Loui AC. Key frame extraction from consumer videos using sparse representation. *Proc - Int Conf Image Process ICIP*. 2011;(1):2437-2440. doi:10.1109/ICIP.2011.6116136
16. Zhuangt Y, Rui Y, Huang TS, Mehrotra S. ADAPTIVE KEY FRAME EXTRACTION USING UNSUPERVISED CLUSTERING. 1998;(94):866-870.
17. Angadi S, Naik V. Entropy based fuzzy C means clustering and key frame extraction for sports video summarization. *Proc - 2014 5th Int Conf Signal Image Process ICSIP 2014*. 2014:271-279. doi:10.1109/ICSIP.2014.49
18. Wu J, Zhong S, Jiang J. A novel clustering method for static video summarization. *Multimed Tools Appl*. 2016.

doi:10.1007/s11042-016-3569-x

19. Besiris D. Key frame extraction in video sequences : a vantage points approach. 2007:0-3.
20. Chen H. Video Key Frame Extraction Based on Spatial-temporal Color Distribution. 2008:196-199. doi:10.1109/IIH-MSP.2008.245
21. Spyrou E, Avrithis Y. Keyframe Extraction using Local Visual Semantics in the form of a Region Thesaurus. 2007:98-103. doi:10.1109/SMAP.2007.39
22. R JVCI, Lai J, Yi Y. Key frame extraction based on visual attention model. *J Vis Commun Image Represent*. 2012;23(1):114-125. doi:10.1016/j.jvcir.2011.08.005
23. Shao L, Ji L. Motion Histogram Analysis Based Key Frame Extraction for Human Action / Activity Representation. 2009. doi:10.1109/CRV.2009.36
24. Chang HS, Sull S, Lee SU, Member S. for Content-Based Retrieval. 1999;9(8):1269-1279.
25. Song X, Fan G. Joint Key-Frame Extraction and Object Segmentation for Content-Based Video Analysis. 2006;16(7):904-914.
26. Mentzelopoulos M, Psarrou A. Key-Frame Extraction Algorithm using Entropy Difference. 2004:39-45.
27. Ravinder M, Venugopal T. Content-Based Video Indexing and Retrieval using Key frames Texture , Edge and Motion Features. 2016;6(2):672-676.
28. Xia G, Sun H, Niu X, Zhang G, Feng L. Keyframe Extraction for Human Motion Capture Data Based on Joint Kernel Sparse. 2016;2(c). doi:10.1109/TIE.2016.2610946
29. Guan G, Wang Z, Lu S, Deng J Da, Feng DD. Transactions Letters. 2013;23(4):729-734.
30. Badre SR, Coding ABT. Summarization with Key Frame Extraction using Thepade ’ s Sorted n-ary Block Truncation Coding Applied on Haar Wavelet of Video Frame. 2016:332-336.

IoT Enabled Things Finder Device using Node MCU

Dr. Ankit Bhavsar[a*], Dr. Poonam Dang[b], Dr. Jyoti Dubey[c]

[a*bc]*Assistant Professor, FCAIT, GLS UNIVERSITY, Ahmedabad 380006, India*

Abstract

Frequently, we face the challenge of recalling the location where we have kept essential items such as official documents, laptops, keys, briefcases and many more. To address this problem, we propose a smart IoT-enabled device that can locate lost items. The prototype model utilizes Node MCU ESP8266, an IoT-enabled Wi-Fi platform, and includes an LED along with a buzzer as indicating devices. The proposed prototype architecture is designed to assist individuals in finding their misplaced valuable things by identifying their last known location. This feature enables users to locate their lost belongings quickly, with just a click using the Blynk application on their smartphone.

Keywords: Node MCU, IoT, GPS Module, Microcontroller

1. Introduction

In today's world, the average person spends 10 to 12 hours a day at work, with a workload of 8 to 10 hours per day, also an additional 2 to 4 hours of travel time. With just 24 hours in a day, a person spends a significant portion of their time at work, leaving little time for other responsibilities such as family, social activities, child education, and caring for aging parents. This lifestyle often leads to stress, which can result in health issues such as lack of sleep and short-term memory loss. Other factors such as alcoholism and depression also contribute to memory loss, with recent research indicating that stress can cause a person to forget 56% of things within an hour, 66% after a day, and 75% after six days [1]. This forgetfulness can cause significant stress and problems, such as misplacing important items and spending valuable time searching for them.

In this context, the Internet of Things (IoT) has emerged as a promising technology that uses sensors connected to the internet to monitor events in real-time. The IoT system is fully automated and requires minimal human intervention. The Node MCU is a low-cost open-source hardware and software environment used for the IoT platform, while GPS sensors provide accurate location data.

This paper addresses one of the common problems faced by individuals in their daily lives - forgetting the location of important items such as car keys, home keys, and laptops. As individuals go through their busy daily schedules, they may forget the location of these items, resulting in lost time and added stress. The proposed solution is an IoT-based Things Finder device that individuals can attach to valuable items. When a person forgets the location of an item, the IoT-based device informs them of the item's location, thanks to the connectivity provided by the internet and the GPS sensor's accurate location data.

2. Internet Of Things (IoT)

The Internet of Things, IoT is an Ecosystem of interconnected computing devices, digital and mechanical machines, objects, animals or human beings with unique identifiers and with the ability to transfer the data across the network without human to computer or human to human interaction. An IoT Ecosystem consists of internet enabled devices that use embedded systems like processors, sensors and communication hardware to collect, process and send the data across the networks. The IoT Ecosystem collects the data based on particular events that occurred in the real environment. The IoT devices collected data from the environment and sent it to the IoT gateway where data is analyzed. The result of analysis transfers to the user interface like mobile applications, business applications or any other backend system. IoT devices do the most of the task without human intervention.

Fig.1 diagram shows the data flow within IoT Ecosystem. IoT Ecosystem has three major components: IoT Device, IoT Gateway and Application. Here an IoT device has a kind of sensor that monitors the environment and reacts on particular events. Sensors

*Dr. Ankit R Bhavsar
E-mail address: ankit@glsica.org

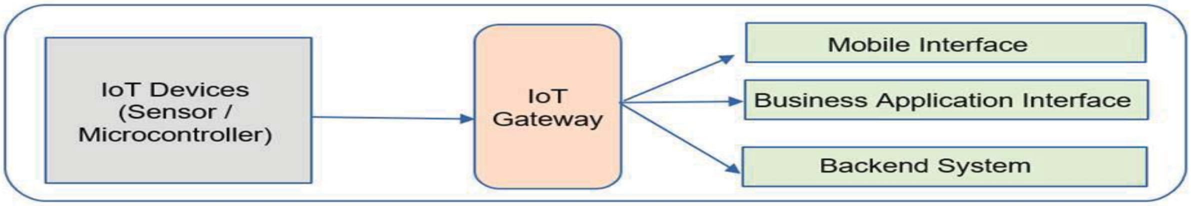


Fig 1. IoT Eco System.

capture the data and transmit to the IoT Gateway. IoT Gateway acts as an intermediate between IoT Devices and Application. IoT Gateway made from the combination of data processing and computing environment. Applications provide an interface to end users to manage the IoT Ecosystem.

IoT makes the life of people very easy. The IoT enabled device provides a facility to monitor the concerned environment from the remote location and take necessary action as and when required. It offers various applications like smart home, smart city, smart framing, smart grids, health monitoring etc. Eventually IoT is the result of people's curiosity to lead a convenient and connected lifestyle with reducing work. That's why we propose IoT enable things finder by this research paper. The research paper is focused on IoT based solutions for the things finding process.

3. Related Work

IoT or Node MCU has been used in various applications related to monitoring, controlling and tracking of objects. [4],[5],[7],[8],[10],[11],[13],[14],[15],[17],[19],[20] Some of the applications are as mentioned below:

Sejal Bagde et. al [2] has proposed an IoT based Smart Switch. They designed and constructed an economical Wi-Fi based automation system for smart home and industrial usage prototype by using ESP8266 Wi-Fi module and smart switch. The proposed research work has developed an Android application with a unit consisting of ESP8266 Wi-Fi module, relay, logic level converter module, capacitive touch sensor module and also a Wi-Fi technology has been used to control the switches.

Kesevan et.al [3] proposed and integrated IoT-based technologies to create an interactive cultivation sensing system to assist farmers in growing their crops in optimum condition and also overcome the problem of the labor shortage. An interactive cultivation sensing system consisting of IoT-based technologies is designed and realized to ensure the continuous growth of crops in optimum conditions daily. With this, progress will be made in determining the efficient cultivation conditions for machine learning, and in finding solutions to future problems of agriculture.

P Siddarth et.al [6] has proposed the IoT based wheelchair fall detection system. They used the accelerometer sensors, Node MCU and RFID technology operating through indoor and outdoor tracking using the embedded system with the thresholds.

Dileep Reddy et.al [9] developed a system to fully automate the petrol bunks with the help of various electronic devices, components, and circuits. Mainly this project is featured on the microcontroller and Radio Frequency Identification (RFID)/Wi-Fi card in which the microcontroller acts as an active device while RFID/Wi-Fi card acts as a passive device. Automatic petrol pump provides the feature of instant recharge. It gives accurate information about selling and control over any adulteration.

Ajay Prakash et.al [12] developed an automated irrigation system that is specifically meant to be used for crops. The system includes a soil moisture sensor which is used to monitor the condition of the soil. The proposed irrigation system is based on the Internet of Things (IoT) technology. The main objective of this model is to monitor contaminated water and protect the field from the light emitted by AC loads and light pillars that fall on LDR. The system employs soil moisture sensors and LDR to measure the current value of the physical parameters. The data collected by the sensors is then transmitted to the microcontroller which displays it on a screen and also sends it wirelessly.

Lukito Hasta Pratopo et.al [16] developed the system to monitor the Temperature and pH level of the Cocoa Beans. They used various components such as Node MCU ESP32, DS18B20 temperature sensor, and SKU SEN0161 pH sensor. The proposed system combined with communication media between smartphones and sensor detection devices via internet, a temperature and pH monitoring system has been successfully created in the cocoa bean fermentation process. Based on the results of the performance test, the monitoring system succeeded in reading the temperature and pH during the fermentation process.

Mardianus et.al [18] developed the prototype smart security tools on doors using RFID with a Node MCU-based telegram monitor in the process of opening the door. They developed a system with the aim to provide security and comfort and make work easier to make it faster, more effective and efficient. The prototype was built using the C programming language with the Arduino IDE (Integrated Development Environment) application, and the system in the form of WEB using the PHP Native programming language using the Sublime Text 3 application and using the MySQL database and using the Telegram application as a medium for receiving notifications in the form of messages to be received by head of the Computer Laboratory. Based on the results of testing, the smart security prototype on the door using RFID can function properly, where each device can function properly and notifications are also sent according to the process of opening the door.

3. Proposed Architecture for IoT Based Thing Finder Device

Many times, we faced the problem of not remembering where we have kept our important things such as: important official documents, briefcase, Laptop, home keys, car keys etc. To provide a reliable solution to cope up with such problems the proposed architecture is based on a smart IoT enabled things finder device. The proposed prototype model is based on IoT, consisting of Node MCU ESP8266 an IoT WI-FI enabled platform. As an indicating device the proposed prototype involves the use of an LED and a buzzer. The proposed architecture will also be able to provide the current location of the lost valuable things using the Blynk IoT mobile application. The proposed architecture involves the interfacing of the Buzzer, LED and GPS module with the Node MCU. Node MCU can be programmed for the required application of the prototype with Arduino IDE. Node MCU needs to be connected with wi-fi with which other components interfaced with Node MCU can be controlled with Blynk application. Using the proposed prototype, the real time monitoring of the precious stuff can be made possible from anywhere at any time. If proper internet connectivity is provided to the node MCU then it can send real time feedback to the Blynk IoT server.

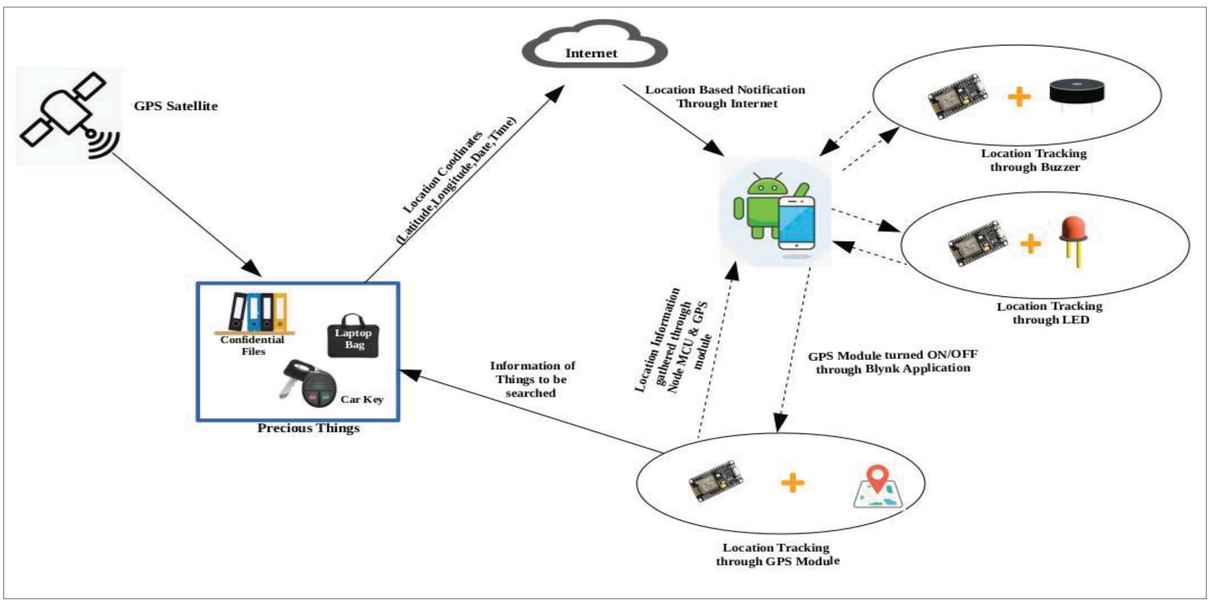


Fig. 2. IoT Based Things Finder Architecture.

The real time detection of the lost stuff can be made in three ways: In the first scenario using the Blynk IoT application we can make LED turn ON interfaced with node MCU which is in contact with the precious device. With the blinking of LED attached with node MCU the stuff can be detected. The second scenario is based on detecting the lost stuff but instead of LED we want to raise a buzzer sound from the place where exactly the stuff is. To do so we can operate a module for the buzzer to turn ON to indicate its position. The third scenario arises when the lost stuff is completely out of our reach and can't be detected with the help of a buzzer and LED module. In such cases the live location of the lost stuff can be made using a GPS module. The GPS module is interfaced with node MCU and displays the GPS data such as coordinates of latitude, longitude date and time on a local web server.

4. Prototype Implementation

The major components of the proposed prototype consist of Node MCU, GPS module, Buzzer and an LED. Interfacing of the components such as GPS module, Buzzer and Led is done with Node MCU. For real time indication of the precious stuff Blynk IoT app is used. Connection establishment is done between Node MCU and Blynk IoT application. Fig. 3 shows the implementation of the things finder system. In the proposed prototype first of all interfacing of Node MCU with GPS module, buzzer and LED is done. The VCC, GND, Rx and Tx pins of the GPS module are connected with 3.3V, GND, D1 and D0 pins of

Node MCU. LED and a Piezo-Buzzer is connected to D5 and D6 pins of Node MCU. Buzzer, Led and GPS module will work as indicating factors in identifying the lost precious stuff such as: laptop bag, car keys and confidential files etc. The coding for the proposed prototype is to be done in Arduino IDE. Blynk IoT application needs to be installed on an android phone for real time tracking of the object. Blynk IoT is set up on an android phone by creating a Blynk account. Once the account is created the authentication code is sent on the registered email id. To establish connection between the Blynk IoT app and Arduino IDE authentication key received on registered mail id is added in the code. Also, Wi-Fi id and password also needs to be added in the code to have seamless internet connection mandatory for working of Node MCU and Blynk IoT application. Once the whole setup is completed as per the mentioned scenarios the lost stuff can be detected either by using GPS module, Buzzer or an LED light depending upon the vicinity of the object.

The proposed prototype includes a Node MCU, GPS module, buzzer, and LED as its main components. The components are interfaced with Node MCU to work together. Real-time tracking of the object is achieved through the Blynk IoT app. The connection between Node MCU and Blynk IoT application is established first. The GPS module, buzzer, and LED are then interfaced with Node MCU. The GPS module's VCC, GND, Rx, and Tx pins are connected to the 3.3V, GND, D1, and D0 pins of Node MCU. The LED and Piezo buzzer are connected to the D5 and D6 pins of Node MCU.

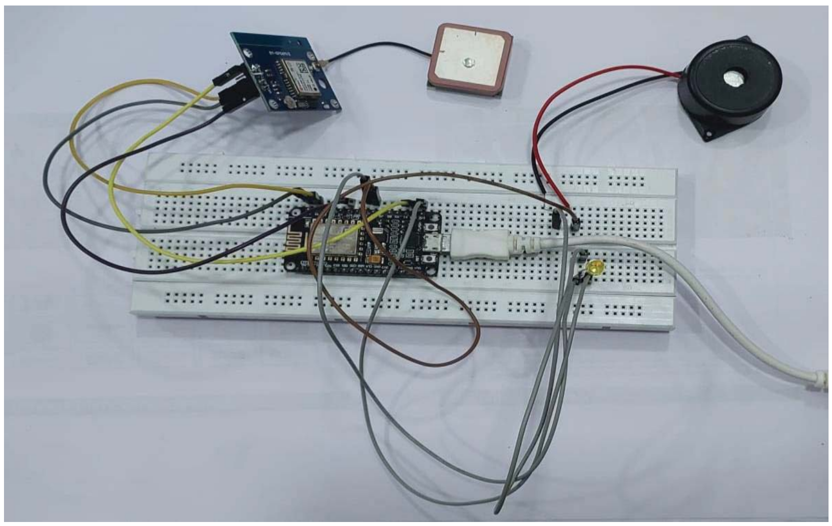


Fig 3. Things Finder Implementation.

The buzzer, LED, and GPS module work together to indicate the presence of lost items such as a laptop bag, car keys, confidential files, etc. The coding for the prototype is done in the Arduino IDE. The Blynk IoT app is installed on an android phone for real-time tracking of the object. To set up Blynk IoT on an android phone, a Blynk account is created, and the authentication code is sent to the registered email ID. To establish a connection between the Blynk IoT app and Arduino IDE, the authentication key received on the registered email ID is added to the code. The Wi-Fi ID and password are also added to the code for seamless internet connection, which is mandatory for the Node MCU and Blynk IoT application to work together. Once the setup is completed according to the mentioned scenarios, the lost items can be detected using the GPS module, buzzer, or LED light depending on the visibility of the object.

5. Conclusion

The proposed architecture aims to enhance the conventional process of locating objects within or beyond defined boundaries by utilizing an IoT-enabled Node MCU-based remote tracking and notification system for individuals. With this system, users can obtain real-time GPS location data of the object via a mobile application. The buzzer-based location tracking feature assists in finding objects within enclosed areas, while the GPS module-based location tracking feature aids in locating objects in large, unbounded open spaces in real-time.

References

1. Murre JMJ, Dros J., 2015. Replication and analysis of Ebbinghaus' forgetting curve. Chialvo DR, ed. *PLoS ONE*. 10(7): e0120644. doi: 10.1371/journal.pone.0120644.

2. Sejal Bagde, Pratiksha Ambade, Manasvi Batho, Piyush Duragkar, Prathmesha Dahikar and Avinash Ikhar, 2021. Internet of Thing (IoT) Based Smart Switch", Journal of ISMAC, Vol. 03/No.02, 2021, p. 149-162
3. Kesevan Veloo, Hayate Kojima, Shogo Takata, Masashi Nakamura, Hironori Nakajo, 2019. Interactive Cultivation System for the Future IoT Based Agriculture, Seventh International Symposium on Computing and Networking Workshops (CANDARW), DOI 10.1109/CANDARW.2019.00059
4. Junhu Ruan, Hua Jiang, Chunsheng Zhu, Xiangpei Hu, Yan Shi, Tianjun Liu, Weizhen Rao, Felix Tung Sun Chan, 2019. Agriculture IoT: Emerging Trends, Cooperation Networks, and Outlook, IEEE Wireless Communications, 10.1109/MWC.001.1900096
5. Prof. P.G. Patil, Prasad Javharkar, Manoj Dashpute, Akshay Damale, Sanju Kheriya, 2018. IT BASED NAVIGATION ROBOT", Journal of Analysis and Computation (JAC), Volume XI, Issue II, ISSN 0973-2861.
6. P Siddarth, 2021. IoT BASED WHEELCHAIR FALL DETECTION", https://ssrn.com/abstract=3918356, http://dx.doi.org/10.2139/ssrn.3918356.
7. George Oguntala, Raed Abd-Alhameed, Stephen Jones, James Noras, Mohammad Patwary, Jonathan Rodriguez, 2018. Indoor Location Identification Technologies for Real-Time IoT-based Applications: An Inclusive Survey, Elsevier.
8. BAKYALAKSHMI V, ANITHAMARY M, BHARATHY R, JHANANI SHREE U, 2021. SYSTEMATIZED WAREHOUSE BASED ON IoT, Iraqi Academics Syndicate International Conference for Pure and Applied Sciences adn Journal of Physics: Conference Series, doi:10.1088/1742-6596/1818/1/012226.
9. Dileep Reddy Bolla, Jijesh J J, Satya Srikanth Palle, Mahaveer Penna, Keshavamurthy, Shivashankar, 2020. An IoT Based Smart E-Fuel Stations Using ESP-32, 5 th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT).
10. Fatema Tuz Zohora, Md. Rezwanur Rahman Khan, Md. Fazla Rabbi Bhuiyan, Amit Kumar Das, 2017. Enhancing the Capabilities of IoT Based Fog and Cloud Infrastructures for Time Sensitive Events, International Conference on Electrical Engineering and Computer Science (ICECOS).
11. Bhisham Sharma, Mohammad S. Obaidat, 2019. Comparative analysis of IoT based products, technology and integration of IoT with cloud computing, IET Journals, ISSN 2047-4954, doi: 10.1049/iet-net.2019.0180.
12. Ajay Prakash, Rahul Kumar, Prabhat Kumar Singh, Ankit Singh, Asst. Prof. Shahab Ahmed, Asst. Prof. Abhishek Jain, 2021. A Review on IoT Based Water Irrigation And Farm Protection using Arduino, International Journal of Scientific Research & Engineering Trends,Volume 7, ISSN: 2395-566X.
13. P. Visalakshi, D. V. Sai Teja, K. Muni Kiran, 2019. IoT based ATM monitoring using cloud data with sensors, International Journal of Advance Research, Ideas and Innovations in Technology, Volume 5, Issue 2, p. 1325-1327, ISSN: 2454-132X.
14. C. Shobana Nageswari, C.G. Sangeetha, V.B. Yogambigai, 2018. IoT based Smart Mine Monitoring System, International Journal of Electronics, Electrical and Computational System (IJEECS), Volume 7, Issue 4, p. 690-695, ISSN 2348-117X.
15. MD. Imtiaz Malek, Refah Nanjiba, Zannatun Nayeem, 2020. Real-Time IoT Based Urban Street Water-Logging Monitoring System Using Google Maps, 2nd International Conference on Image Processing and Machine Vision, Association for Computing Machinery,New York, ISBN: 978-1-4503-8841-2, https://doi.org/10.1145/3421558.3421574.
16. Lukito Hasta Pratopo, Ahmad Thoriq, Eko Heri Purwanto, Daffa Afian Wiradwinanda, 2022. Temperature and pH Monitoring System Design in the Fermentation of Cocoa Beans Based on Android, PELITA PERKEBUNAN, Volume 38, Number 1, ISSN: 0215-0212 / e-ISSN: 2406-9574, DOI: 10.22302/iccri.jur.pelitaperkebunan.v38i1.494.
17. Viraj Patil, Anup Randive, Mr. Deepak Gaikwad, 2021. SMART CAR PARKING SYSTEM", Open Access International Journal of Science & Engineering, Volume 6, Special Issue- AIMTMREF, ISSN (Online) 2456-3293, DOI 10.51397/OAIJSE06.2021.0041.
18. Mardianus, Andi Yusika Rangan, Salmon, Prototype Smart Security on Doors using RFID with Telegram Monitor NodeMCU Based, TEPIAN Vol. 2 No. 1, ISSN 2721-5350 e-ISSN 2721-5369, doi.org/10.51967/tepian. v2i1.293.
19. Rifqi Firmansyah, Muhamad Yusuf, Pressa P. Surya Saputra, IoT Based Temperature Control System Using Node MCU ESP 8266, International Joint Conference on Science and Engineering (IJCSE), Advances in Engineering Research, volume 196, ATLANTIS PRESS.
20. Satien Janpla, Nisanart Tachpetpaiboon, Chaiwat Jewpanich, 2019. Development of Automatic Home-Based Fish Farming Using the Internet of Things, International Journal of Recent Technology and Engineering (IJRTE), Volume-8 Issue-2, ISSN: 2277-3878, DOI: 10.35940/ijrte. B2677.078219.

Market Basket Analysis and Design of Chatbot for Recommendations to Startups in Tamil Region

Shruthee B[a], Shekinah Olive[a], Dr K Sundarakantham[a]

[a]*Department of Computer Science and Engineering, Thiagarajar College of Engineering, Tamil Nadu, 625015, India*

Abstract

Decision-makers of an organization can use Association Rule Mining (ARM), which is a data mining technique, to extract hidden information from databases and increase their overall profit. The Apriori algorithm is a traditional or classical method for association mining rule, but since it scans the entire database to generate the candidate item sets it takes a long period of time. The Apriori algorithm is being used on a Tamil dataset that is a dataset of grocery items which are needed in the Tamil household. This customized dataset is created by distributing a survey form to friends and families who come from Tamil backgrounds. The association rules that are formed need to cross the threshold value: minimum support:0.0045, minimum confidence: 0.2, and minimum lift: 3. For our database Apriori algorithm works better when compared to the FP-Growth algorithm. Finally, the acquired association rules are made available to the vendors via a chatbot.

Keywords: Apriori Algorithm; Association rule; Support; Lift; confidence; chatbot.

1. Introduction

Data mining is a popular method to obtain information from a large database. It is widely used in real-time data like textile showroom or supermarket data. Association rule mining is one of the popular techniques used in data mining that can generate association rules among the items in the dataset.

Association rules [4] involve the identification of unique patterns among the items in the given dataset and extract different association rules from the database to measure support and confidence; and is one of the basic Market Basket Analysis (MBA) that gives information to increase business sales [9] and also make purchases for customers easier.

1.1. Related Works

R Suganya et al [1] presented a paper on frequent pattern mining using Apriori-based algorithms. The provided techniques for finding association rules in huge databases are not only effective but also quick. This method's crucial contribution is how much the I/O overhead associated with the Apriori algorithm is reduced. These techniques will be helpful in several actual boolean data storage database mining applications. These methods now only operate with boolean datasets, thus additional work is required to make them relevant to all types of datasets.

Wiwit Pura Nurmayanti et al [2] presented a paper on market basket analysis using an apriori algorithm on outdoor product sales data mainly focused on Indonesia. The Apriori algorithm provides transaction patterns in the sale of outdoor goods as the result. The paper uses a minimum support value of 0.296, a threshold confidence value of 0.774, and a threshold lift value of 1.49. It shows that consumer purchase patterns regarding the purchase of portable stove items have a potential to buy portable gas items. FP-Growth algorithm is used to find transaction patterns in the sale of outdoor goods in this paper.

Yusuf Kurina et al [3] expressed their idea of market basket analysis for knowing the sales pattern at the O! Fish restaurant using an Apriori algorithm. O! Fish restaurants may make more effective advertising tactics by mentioning goods (menus) that are frequently bought together using information about sales patterns. Data mining approaches utilizing Apriori algorithms may be coupled with knowledge of the goods that customers often buy together to determine their buying habits. To create association rules, an Apriori algorithm is utilized. By referring to a combination of goods that are frequently purchased concurrently, O! Fish restaurants may utilize information about the association's regulations in consumer menu purchases to develop more possible promotional techniques to increase sales.

Manpreet Kaur et al [4] published a paper on market basket analysis to identify the changing trends of market data using association rule mining Marketing, bioinformatics, education, nuclear science, and other industries can all benefit from using Market Basket Analysis, also known as Association Rule Learning or Affinity Analysis. The primary objective of an MBA is to help retailers gain knowledge about consumer purchasing patterns so that they may make sound decisions to help their business. For doing market basket analysis, a variety of algorithms are available. The current algorithms can only work with static data so, the changes that take place over time are not included. This paper discusses association rule mining and provides a new algorithm which may help to examine customer behavior and assists in increasing sales.

Luis Cavique [5] designed a method on a scalable algorithm for market basket analysis. Market basket analysis is a useful technique for helping retailers connect different item sets with each other to help them flourish their business. Finding big baskets is crucial in any industry that deals with a huge number of things, but it's especially important in retail. Even though certain methods can locate big item sets, they may not be computationally efficient. This work aims to offer a big itemset pattern discovery technique for market basket analysis. The market basket aim is converted into a maximum-weighted clique aim in order to acquire the condensed data, which are then employed in this method.

Kwei Tang et al [6] published a paper on context-based market basket analysis in a multi-store environment. A novel method was provided to conduct market basket analysis across various stores and time periods. According to the user's application and demands, the user first establishes a time concept hierarchy and a location hierarchy. By merging the levels of the two hierarchies, a collection of contexts is methodically obtained. Then an effective approach for deriving association rules that satisfies all context-specific support and confidence requirements was created. This method enables a decision-maker to examine purchasing trends at various concepts like time and location for a combination of days and stores, a combination of quarters and states, or a combination of days and regions. The association rules are highly ordered in addition to being flexible since they are acquired from the time and place hierarchies. The algorithm may produce more detailed and substantial information compared to the store-chain rules and the conventional rules, according to a numerical evaluation of its performance.

Abhishek B Rao et al [7] published a paper on the application of market analysis in healthcare. In the modern day, data analysis is essential since it enables us to comprehend patterns via thoughtful exploration. The Apriori technique is extensively used by academics to locate frequently recurring objects in transactional databases, making it one of the primary strategies. In this essay, market basket analysis' applicability to the healthcare sector is discussed. The Apriori method is used in the current study to identify common illnesses that co-occur in a region. This could encourage locals to be more aware of common illnesses and to take every safety action at their disposal to protect their health.

Dr Sandeep A. Thorat et al [8] published a paper on computer–human interaction. The popular approach is through the use of chatbots, which are computer programs designed to make this process easy and engaging. Current artificial intelligence techniques often struggle to provide the most appropriate response to user queries, leading to the predominance of rule-based chatbot systems in industry. This paper presents a comprehensive study on the implementation of rule-based chatbot systems, including discussions on performance measurement parameters for such systems. Furthermore, a comparison is made between two of the most widely used rule-based chatbot implementation frameworks, Google Dialogflow and IBM Watson. Finally, the paper concludes by listing expectations for future chatbot systems.

Robert C. Blattberg et al [9] published a chapter on the examination of the items that customers commonly purchase together. This information is to determine which products should be promoted or cross-sold together. This term is derived from the shopping carts used by customers in supermarkets during their shopping trips. With the advent of the internet, there are now new opportunities for collecting and analyzing such data. This chapter provides an overview of the fundamental concepts of "confidence," "support," and "lift" in relation to market basket analysis, and explains how these concepts can be translated into practical metrics that can be expanded upon.

Jagdish Singh et al [10] published a paper to present the implementation of a rule-based enquiry chatbot that is designed exclusively for students of Asia Pacific University (APU). This chatbot is called ‘APU Admin Bot’. It can help students with a fast solution in a methodical way to resolve their questions and doubts instead of seeking help from the administrative staff. This chatbot makes use of a rule-based approach in the area of pattern recognition. It uses certain words or phrases and sometimes even actions to fetch and display the whole set of responses from the chatbot. It is built entirely from the Chatfuel platform. The chatbot is a messaging platform and is more reliant on a code-less platform.
Bulleted lists may be included and should look like this:

- First point
- Second point
- And so, on

1.2. Our Contributions

Our contributions to this paper include,

- Creation of customized regional language dataset.

- Creation of association rule for the Tamil language customized grocery dataset.
- Creation of a chatbot, to use the association rules generated for the retailer's purpose.

1.3. Overview of our technique

This work makes use of the Apriori algorithm. It is a data mining technique which is widely utilized to do market basket analysis.

The novelty of the proposed work is that it has a regional language dataset created from scratch and a chatbot which will give suggestions to the retailers based on the association rule; backed by the Apriori algorithm.

Market Basket Analysis (MBA) is a technique of determining what product is purchased when a specific product is purchased by a customer. This is determined by analyzing the customer's purchase history.

So, the purchase history of customers is collected in order to determine the products which are most likely to be purchased with a product.

The products purchased by the customers are referred to as itemsets. Market Basket Analysis is done on the customer purchase history to find relationships between the different item sets.

Since MBA is useful in finding the association or relationships among the items purchased by the customer, it can be made use of by retailers to help their business. Retailers can gather customer purchase data and analyze that data using Market Basket Analysis and extract useful information regarding the relationship between the different item sets.

This relationship between different itemsets is known as the association rule. Each rule is determined by its:

- Confidence
- Support
- Lift

Key differences between the proposed work and existing work,

(i) While there are a number of works available for English language datasets, works on regional language datasets are very low.

(ii) The proposed work applies the Apriori algorithm for low resource Tamil language, which is used by the chatbot application.

2. Proposed Methodology

We propose the use of the Apriori algorithm to find association rules for the regional language dataset. The overall architecture of the methodology that is put forward is shown in Fig 1. The following sections explain the methodology in detail.

2.1. Association Rule Mining

We have used association rule mining to perform market basket analysis. An approach to machine learning, based on rules is association rule mining. This helps us to find significant connections between various items as they occur together in the data collection.

Three core measures that are used in the association rule are Support, Confidence and Lift.

- Support
- Confidence
- Lift

Support:
The percentage of groups that include all the items stated in an association rule is known as the rule's support. This is calculated from all the groupings that were taken into consideration as the percentage value.

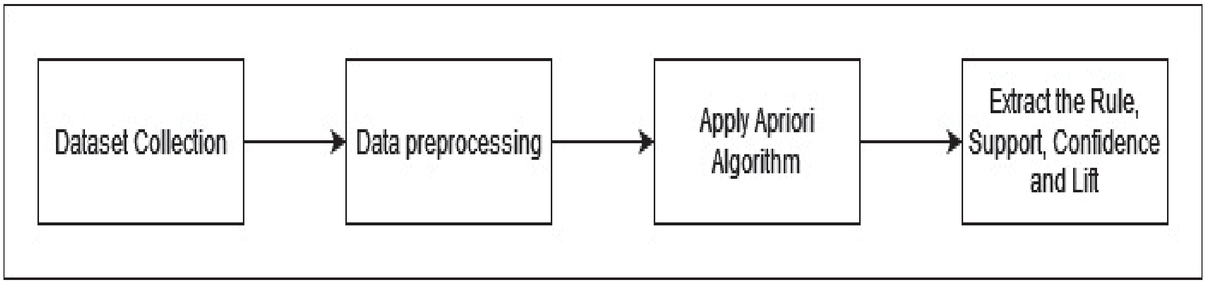


Fig. 1. Workflow of the Market Basket Analysis

Support (X ⇒ Y) = Count of X ⇒ Y/ Total number of Transactions

Confidence:

Confidence is defined as the percentage of how frequently the rule head appears across all the groups that also include the ruling body This rule's reliability is indicated by the confidence value.

Confidence (X ⇒Y) = Support (X ⇒Y) / Support (X)

The proposed methodology is as follows:

- Dataset collection: The data for the dataset is collected from local retail stores and acquaintances which makes it more reliable and accurate for real-time use.
- Data preprocessing: The collected data can have null values. These if left in the dataset will give unreliable results, therefore we removed them to only take the required text data into consideration.
- Apply Apriori algorithm: Apriori algorithm makes use of prior knowledge of frequent itemset properties, which helps to generate the association rules. The frequent itemsets are generated in a level-wise manner. The Apriori property that, all the nonempty subsets of a frequent itemset must also be frequent, is followed over here. The algorithm is applied with minimum support of 0.0045, minimum confidence of 0.2, a minimum lift of 3, minimum length of 2.
- Extraction of Rule, Support, Confidence and Lift: After applying the Apriori algorithm to the records of the dataset we obtained the association results. From which the rules and their necessary parameters can be extracted.

Table 1. An example of a table

An example of a column heading	Column A (*t*)	Column B (*T*)
And an entry	1	2
And another entry	3	4
And another entry	5	6

If table footnotes should be used, place footnotes to tables below the table body and indicate them with superscript lowercase letters. Be sparing in the use of tables and ensure that the data presented in tables do not duplicate results described elsewhere in the article.

2.2. Apriori Algorithm

The apriori algorithm is a popularly used algorithm in association rule learning. It is used to identify the items in a data set, and create itemsets out of it to create association rules.

An itemset has a low likelihood of occurring if:

- Minimum support threshold is P(I), where I is any itemset that isn't empty.
- Any subset in the itemset's value that is below the minimum support level is ignored.

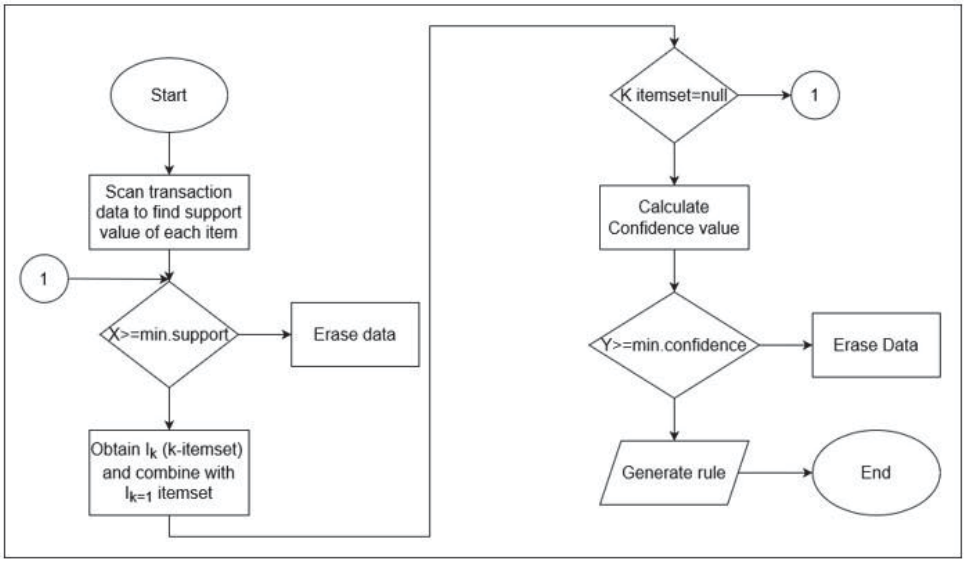


Fig. 2. Flowchart of Apriori algorithm

The Anti-monotone Property is the second feature. An illustration for it would be that if the chance of buying a burger is below the minimum support, the possibility of buying a burger and fries will be lower as well.

Steps in the apriori algorithm:

- Computing the support, also known as frequency, for each item individually.
- Set a threshold value for support.
- Selecting the frequent items.
- Find the frequent itemsets support.
- Repeat for the dataset.
- Construct the association rules and calculate their confidence.
- Check the list of rules.

The implementation for the dataset is as follows:

Algorithm 1 (Apriori algorithm for MBA to find association items)
1. Download Apriori on your environment
2. Import necessary packages and dataset creation
3. Give the custom dataset as input
4. Do data preprocessing and perform descriptive analysis
5. Set minimum support, confidence, and lift.
6. For each item in the association rule results:
6.1. Extract the rule, support, confidence and lift.
6.2. Print association rules, which is expected.

2.3. Chatbot

A chatbot is a computer program used to understand questions a user asks and automate the response process to simulate human conversation. The process of finding information is made easier through Chatbots by giving a direct reply to the requested queries. The queries can be of any format like text, audio, or both.

The chatbot we have developed is a simple chatbot which is text-based. It is pre-programmed to give replies to a limited set of inputs with answers that had been written when the chatbot was developed. It operates much similarly to the ones for an interactive FAQ chatbot. The chatbot does not give answers to untrained inputs but just gives an unavailable response.

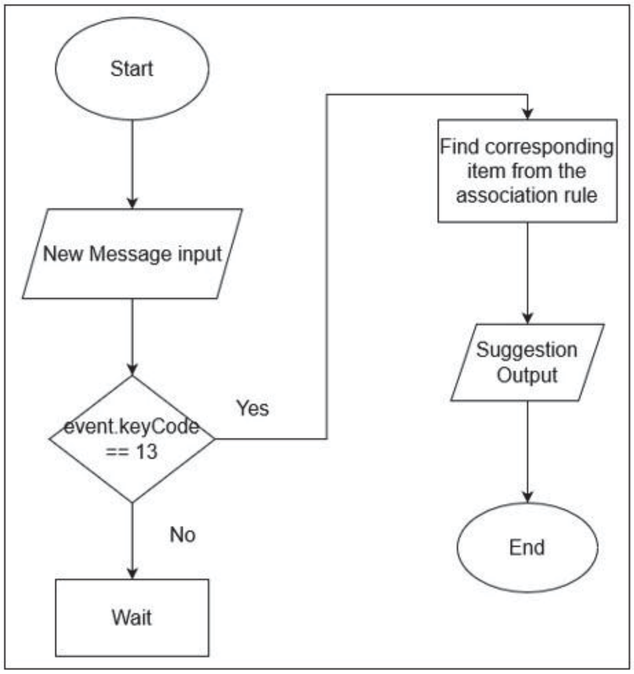


Fig. 3. Flowchart of Apriori algorithm

This chatbot can be used by people who would need a recommendation tool to give them suggestions regarding their grocery business, whether it be a new or existing business. A user would not need to rely on searching the internet to find the related items for their given inputs.

3. Experimental results and analysis

3.1. Dataset collection

The data set is created by gathering information from our acquaintances and nearby stores. The custom dataset contains transaction data of about 10000 different transactions, all in the Tamil language.

The dataset has a large variety of items and there are transactions of different lengths ranging from 1 item in a transaction to 20 items in a transaction.

Most of the items are checked for spelling mistakes. The dataset itself is random and is not ordered in any specific pattern.

Table 1. Parameters of the Apriori algorithm used

Parameters	Values
minSupport	0.0045
minConfidence	0.2
minLift	3
minLength	2

The dataset is stored in Excel with a variety of items in different transactions.

	A	B	C	D	E	F	G
7502	இறைச்சி	பால்	ரொட்டி				
7503	பால்	சோடா					
7504	இறைச்சி	சோடா	குளிர்களி				
7505	இறைச்சி	ரொட்டி					
7506	இறைச்சி	தயிர்					
7507	காய்கறிகள்						
7508	வெண்ணெய்	பால்	பழங்கள்	சர்க்கரை			
7509	பால்	சாக்கலேட்					
7510	காய்கறிகள்	சலவைத்தூள்					
7511	இறைச்சி	ரொட்டி					
7512	பழங்கள்	பால்					
7513	பழங்கள்	சோடா	தயிர்	காய்கறிகள்	முட்டை	சர்பத்	
7514	இறைச்சி	சர்பத்	இனிப்பு				
7515	மீன்	பானங்கள்					
7516	பால்	இறைச்சி	மெழுகுவர்த்திகள்	திராட்சை	மூலிகைகள்	காய்கறிகள்	தயிர்
7517	பாலாடைக்கட்டி	நீர்					
7518	பால்	மிட்டாய்					
7519	காய்கறிகள்						
7520	தயிர்	இறைச்சி					

Fig. 4. Customized Dataset Screenshot

3.2. Analysis of the Apriori Algorithm

The advantages of using the Apriori algorithm for extracting association rules:

- (i) One of the simplest and easily understandable algorithms is the Apriori algorithm.
- (ii) The derived associations are understandable and easy to explain to the client.
- (iii) There is no need to use labelled data because the algorithm is a fully unsupervised algorithm. So, it can be used in various circumstances as unlabeled data is more accessible.
- (iv) It is a comprehensive algorithm, so it detects each rule for a specified level of support and confidence.
- (v) Apriori algorithm is the most suitable for the work because it uses join functions to check all possible combinations that fits above the given threshold.
- (vi) Join and Prune steps are easy to implement on large itemsets in large databases.

Table 2. Comparative analysis of results

Paper	Description
Frequent Pattern Mining Using Apriori-Based Algorithm [1005D	They presented a paper on frequent pattern mining using apriori-based algorithms. The provided techniques for finding association rules in huge databases are not only effective but also quick.
Market Basket Analysis with Apriori Algorithm and Frequent Pattern Growth (FP-Growth) on Outdoor Product Sales Dat [2]	They presented a paper on market basket analysis using an apriori algorithm on outdoor product sales data mainly focused on Indonesia. The Apriori algorithm generates transaction patterns in the sale of outdoor goods. The level of strength of the rules at minimum support is 0.296, the confidence is 0.774, and the lift value is 1.49. This shows that the consumers who buy portable stove items also have the potential to buy portable gas items.
Proposed method	We applied the Apriori algorithm to the Tamil Language dataset, which mostly focuses on locally available groceries. It helped us to understand the purchase patterns of people living here.

The output obtained is as follows.

	LHS	RHS	Support	Confidence	Lift	Frequency
0	முழு கோதுமை பாஸ்தா	இடலை எண்ணெய்	0.007961	0.271493	4.142195	0.191057
1	பாஸ்தா	இறால்	0.005042	0.322034	4.528301	0.121003
2	பாஸ்தா	இறைச்சி	0.006103	0.389831	3.460721	0.146477
3	தக்காளி சட்னி	மாட்டிறைச்சி	0.005307	0.377358	3.859092	0.127372
4	மூலிகை & மிளகு	மாட்டிறைச்சி	0.015921	0.323450	3.307793	0.382115
5	ஸ்பாகெட்டி - அரிசி மாவு	இடலை எண்ணெய்	0.005042	0.201058	3.067562	0.121003
6	காய்கறிகள் - இடலை எண்ணெய்	பால்	0.005042	0.395833	3.034991	0.121003
7	காய்கறிகள் - ஸ்பாகெட்டி	இடலை எண்ணெய்	0.006369	0.200837	3.064184	0.152846
8	கோழிக் கறி - இடலை எண்ணெய்	பால்	0.004909	0.411111	3.152131	0.117819
9	சூப் - நீர்	இடலை எண்ணெய்	0.005174	0.222857	3.400150	0.124187
10	மாட்டிறைச்சி - பால்	இடலை எண்ணெய்	0.004909	0.224242	3.421286	0.117819

Fig. 5. Output of Apriori algorithm

Fig. 6. Output of Chatbot

Some of them with their obtained parameters are:

Table 3. Results obtained

Antecedents	Consequents	antecedent support	consequent support	Support	Confidence	Lift	Leverage
(உழுத்தம் பருப்பு, கோதுமை)	(கடலை பருப்பு, திணை)	0.182857	0.182857	0.182857	1.0	5.468750	0.149420
(கடலை பருப்பு, திணை)	(உழுத்தம் பருப்பு, கோதுமை)	0.182857	0.182857	0.182857	1.0	5.468750	0.149420
(உழுத்தம் பருப்பு, கோதுமை)	(திணை, பயத்தம் பருப்பு)	0.182857	0.182857	0.182857	1.0	5.468750	0.149420
(திணை, பயத்தம் பருப்பு)	(உழுத்தம் பருப்பு, கோதுமை)	0.182857	0.182857	0.182857	1.0	5.468750	0.149420
(உளுந்து, திணை)	(உழுத்தம் பருப்பு, கோதுமை)	0.182857	0.182857	0.182857	1.0	5.468750	0.149420

4. Conclusion

We applied the Apriori algorithm to the Tamil Language dataset, which mostly focuses on locally available groceries. It helped us to understand the purchase patterns of people living here. It can be used to assist retail stores with their stocking schedule and stock maintenance. With the help of this study, the stores can make offers and discounts to increase their sales. This can also be used by the producers to understand customer demand and purchase patterns. We obtained 22 association rules for the dataset.

References

1. R. Suganya, R. Tamil Selvi, 2014. Frequent Pattern Mining Using Apriori Based Algorithm; IJERT.
2. Wiwit Pura Nurmayanti, Hanipar Mahyulis Sastriana, Abdul Rahim, Muhammad Gaz-ali. Ristu Haiban Hirzi, Zuhut Ramdani, Muhammad Malthuf, 2021. Market Basket Analysis with Apriori Algorithm and Frequent Pattern Growth (Fp-Growth) on Outdoor Product Sales Data; International Journal of Educational Research and Social Sciences.
3. Yusuf Kurnia, Yohanes Isharianto, Yo Ceng Giap, Aditiya Hermawan & Riki, 2019. Study of application of data mining market basket analysis for knowing sales pattern (association of items) at the O! Fish restaurant using apriori algorithm; Journal of Physica: Conference Series.
4. Manpreet Kaur & Shivani Kang; 2016. Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining; Procedia Computer Science 85, p. 78-85.
5. Luis Cavique; A scalable algorithm for the market basket analysis, 2007. Journal of Retailing and Consumer Services 14, p. 400-407.
6. Kwei Tang, Yen-Liang Chen, Hsiao-Wei Hu, 2008. Context-based market basket analysis in a multiple-store environment; Decision Support Systems 45, p. 150-163.
7. Abishek B. Rao, Jammula Surya Kiran, Poornalatha G, 2021. Application of market–basket analysis on healthcare; International Journal of System Assurance Engineering and Management.
8. Dr Sandeep A. Thorat, Vishakha D. Jadhav, 2020. A Review on Implementation Issues of Rule-based Chatbot Systems; International Conference on Innovative Computing and Communication.
9. Robert C. Blattberg, Byung-Do Kim, Scott A. Neslin, 2008. Market Basket Analysis; Database Marketing; Springer; p. 339–351.
10. Jagdish Singh, Minnu Helen Joesph and Khurshid Begum Abdul Jabbar, 2019. Rule-based chabot for student enquiries; International conference on computer vision and machine learning.

Violent Interaction Detection in Video Source Using Deep Learning

Priyank Dave[a*] Nisha Khurana[b]

[a*]*PG Scholar, Computer Engineering Department, Gandhinagar Institute of Technology*

[b]*Professor,* Information Technology *Department,* Gandhinagar Institute of Technology

[a]*210120702001@git.org.in,* [b]*nisha.khurana@git.org.in*

Abstract

Violent activity detection is very important in some video analysis eventualities like railway and metro stations, prisons or alternative public places. With the rise of video data set to watch the demands of such a system, that acknowledges the violence and suspicious events has been increased automatically. Violence action detection has become a vigorous research area of computer vision and video processing to attract new researchers. During this work, we tend to discuss developing a method for the automated analysis of violence activity through video supply. Machine learning techniques and tools are used to detect sequences of violence interaction in videos.

Keywords: Violence Detection, Action and Activity Recognition, Anomaly Detection, Machine Learning for VD.

1. Introduction

The majority of the systems require manual human inspection, for identifying violent interaction scenarios in video. Which is both ineffective and practically impossible. Having such a practical system that can automatically monitor videos and identify the violent behavior of humans will be of immense help and assistance to the law-and-order establishment. In recent years, machine learning and computer vision techniques have made it possible to recognize human action from video. Automated video analysis techniques are utilized for a wide range of purposes, including: indexing and finding videos, Summary of a video [1], Recognizing action and activity [2], Classification of videos [3]

Due to its numerous applications in the fields of medicine, security, sports, and entertainment, among many others, action and activity detection has attracted a lot of study interest in the field of video categorization.

Action recognition is a technology that can identify human actions. Based on the complexity of the acts and the number of bodily parts involved, human activities are divided into four types.

The four types are gestures, actions, interactions, and group activities. Whereas violence detection is one of the broad classes that aim on identifying violent and injurious event patterns in an input video. Video classification can be used to identify specific actions from video. Sequences of images make up videos. Images taken from video frames can be used to extract features, and with the help of these features, predictions can be made.

As we can see in Fig. 1.1, Video classification can be used in applications like, activity recognition, violence detection, anomaly recognition.

*Priyank Dhirenkumar Dave

E-mail address: 210120702001@git.org.in

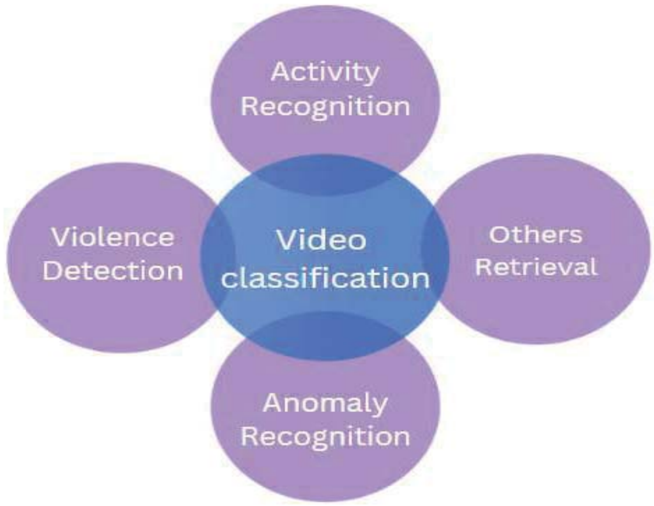


Fig 1.1 Video classification applications.

Fig. 1.2 shows clips extracted from video and makes predictions of activity in frame. As we can see in the image both are examples of running activity.

2. Background:

The community that studies action recognition has generally concentrated on relatively simple activities like clapping, walking, jogging, etc. Comparatively less research has been done on the recognition of specific events like fight behavior and violence detection. Such a feature might come in very handy for applications like video monitoring in jails, public spaces, psychiatric facilities, or even embedded in camera phones and social media.

Violence Detection is the selection or extraction of anomalous patterns from a typical surveillance video that may last for a very brief or an extended period of time. By starting countermeasures such as alerting the closest relevant departments for reaction activities, its early identification aids in either prevention or decreasing the residual harm in terms of human lives and their properties. Violence is any abnormally cruel human action, including hitting, harming, destroying, and wounding others.

Many real-world scenarios can be used with video data. When compared to the typical pattern of ongoing video surveillance, it is obvious that violent incidents are incredibly uncommon. As a result, using human resources to watch video streams might not be viable because it would take a lot of training to spot odd patterns. A video data system's primary objective is to identify and report any abnormal behavior that deviates from the typical pattern of activity.

Since decades, Violence Detection has been used to analyze video data utilizing soft computing methods that were initially based on conventional image processing methods. Early Violence Detection literature took into account a variety of factors for decision-making, including human motion acceleration, appearance, and motion flow, among many others Video data preparation, feature extraction, and segmentation into violent or nonviolent segments are all common processes in baseline research. Segmenting videos and cleaning up data are both part of the preprocessing stage. The term "features extraction" refers to the processing of individual-level (spatial) frames utilizing specific feature extractors including motion, speed, and optical flow.

3. Literature Survey

Violence detection methods can be categorized in two methods:

1. Violence detection using machine learning methods.
2. Violence detection using deep learning methods.

Machine learning methods have gotten more attention from recent decades, but in recent years deep learning approaches are popping up in the attention of researchers.

3.1 - Violence detection using machine learning methods

Penet et al. [6] investigated the different Bayesian network learning algorithms using temporal and multimodal information for a violent shot detection system.overall results gave a false alarm rate of 50% and 3% missed-detection. Nievas et al. [7] proposed to use the bag-of-words (BoW) approach and two motion descriptors, space–time interest points (STIP) and motion scale invariant feature transform (MoSIFT) for fight detection. Detect fights with 90% accuracy. A violent flow (VF) variation for violence detection based on the combination of SVM and Horn–Schunck optical flow algorithm was proposed by Arceda et al. [8]. Das et al. [9] classifier achieves 86%

accuracy.

3.2 - Violence Detection Based on Deep Learning Techniques

Ding et al. [10] used 3D convolution with back propagation strategy for violence detection. Xia et al. [11] used Bi-channels CNN with SVM are used for violence detection. Mu et al. [12] CNN are used to detect violent based on acoustic information. Meng et al. [13] Integrating frame of trajectory and Deep CNN to detecting human violent behavior in videos.

3.3 - Data set availability

Hence data is related to violence, so more dataset is not made publicly available but there are various datasets available like

1. RWF-2000 dataset for violence detection by Cheng et al. [4]
2. Movie Fights dataset by Nievas et al. [5]
3. Hockey Fights dataset by Nievas et al. [5]

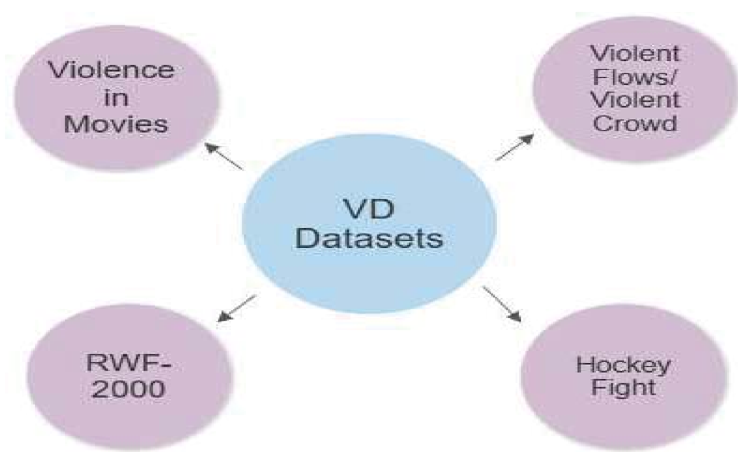


Fig. 3.1. Overview of VD Datasets.

4. Proposed Model

To classify violent or non-violent actions, our model must be able to anticipate sequences in consecutive frames, such as a pattern in the subjects' movement or a degree of their motion, and so on. This is not possible if only the spatial features (features that pertain to a specific frame) of the frames are considered. When detecting sequences in frames, temporal or time-related features must also be examined. The temporal features can be processed in either the forward or reverse sequence. Our model processes temporal characteristics in both directions in addition to spatial features, allowing the model to become more accurate while consuming less computational time.

The proposed model aims to keep performance similar to cutting-edge violence detection models while reducing the computational complexity.

Three steps primarily make up the proposed algorithm:

1. Spatial feature extractor
2. Temporal feature extractor
3. Classifier

The videos' frames have been extracted. The captured frames are reshaped to 64 x 64 pixels (denoted as x y). The training data is a Numpy array, with each row indicating a sequence or pattern in video. A sequence could include a degree of movement and movements, such as whether an arm movement is a punch or a handshake, and so on. A series can be extracted with as few as two frames. However, we extracted temporal features (time-related features) using 16 sequential frames.

When the model receives a video frame, the first step is spatial feature extraction using a network. This network employs MobileNet V2 as an encoder used to perform sequential-time-distributed static single frame spatial feature extraction. As a classifier for the model, we selected MobileNet V2, a simple state-of-the-art classifier for spatial feature extraction. MobileNet V2 gets great accuracy while using a much smaller network size. Then it is passed to BiLSTM for temporal feature extraction. Following fig shows the structure of the model.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 time_distributed (TimeDistr  (None, 16, 2, 2, 1280)   2257984
 ibuted)

 dropout (Dropout)           (None, 16, 2, 2, 1280)    0

 time_distributed_1 (TimeDis  (None, 16, 5120)         0
 tributed)

 bidirectional (Bidirectiona  (None, 64)               1319168
 l)

 dropout_1 (Dropout)         (None, 64)                0

 dense (Dense)               (None, 256)               16640

 dropout_2 (Dropout)         (None, 256)               0

 dense_1 (Dense)             (None, 128)               32896

 dropout_3 (Dropout)         (None, 128)               0

 dense_2 (Dense)             (None, 64)                8256

 dropout_4 (Dropout)         (None, 64)                0

 dense_3 (Dense)             (None, 32)                2080

 dropout_5 (Dropout)         (None, 32)                0

 dense_4 (Dense)             (None, 2)                 66

=================================================================
Total params: 3,637,090
Trainable params: 3,060,642
Non-trainable params: 576,448
_________________________________________________________________
```

Fig. 4.1. Structure of the model.

5. Results

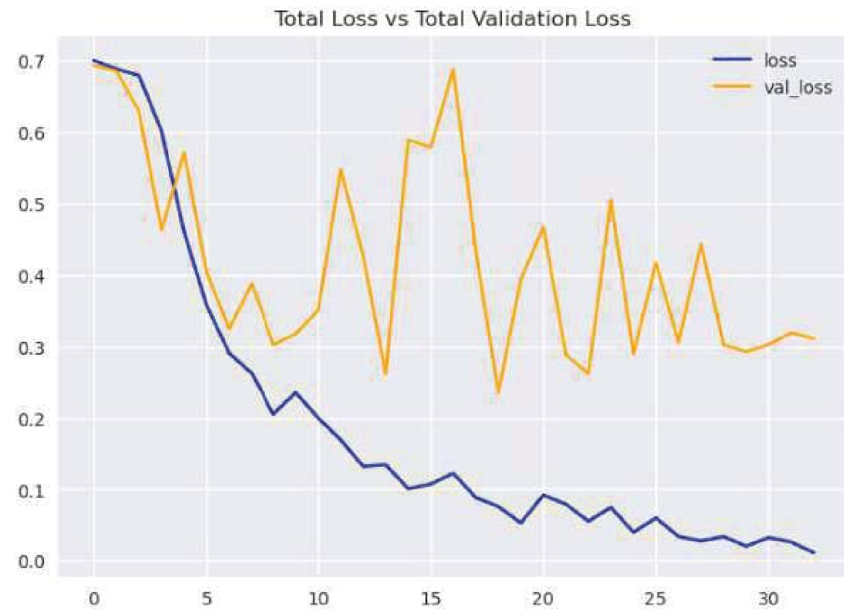


Fig. 5.1. Total loss vs Total Validation Loss.

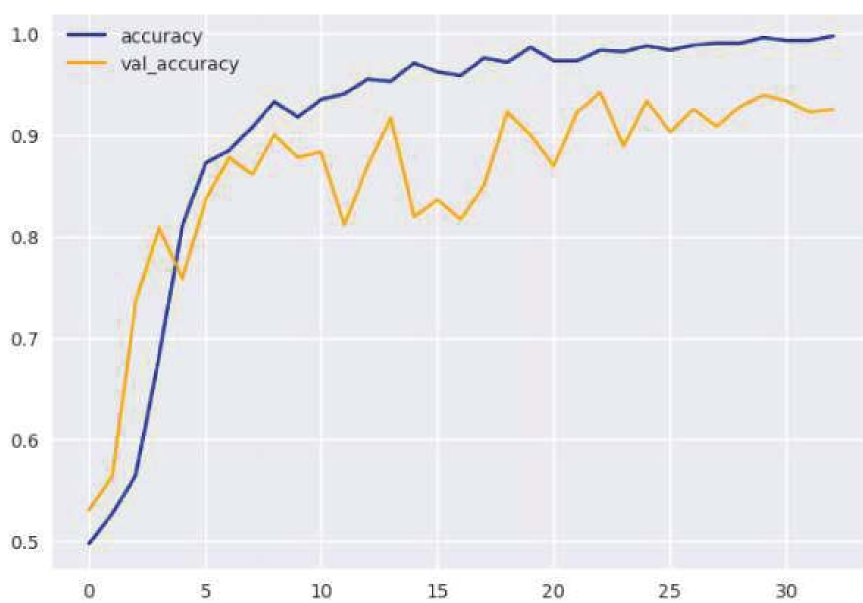


Fig. 5.2. Total Accuracy vs Total Validation Accuracy.

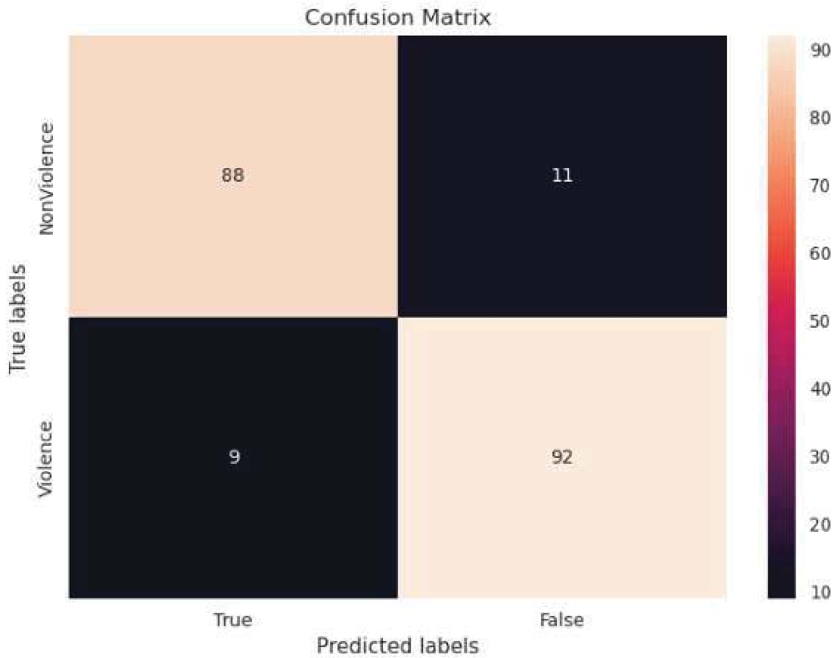


Fig. 5.3. Confusion Matrix.

Classification Report is :

	precision	recall	f1-score	support
0	0.91	0.89	0.90	99
1	0.89	0.91	0.90	101
accuracy			0.90	200
macro avg	0.90	0.90	0.90	200
weighted avg	0.90	0.90	0.90	200

Fig. 5.4. Classification Report.

6. Conclusion

The suggested model extracts spatial features using MobileNet V2 as an encoder, then extracts temporal features using Bi-LSTM followed by a classifier. There are 3,637,090 parameters in the model. The model's architecture makes it computationally quick and light. Experiments using a complicated dataset of real security camera footage based on RWF2000 revealed average accuracy of 0.90.

Although our suggested model performed satisfactorily, it still has to be further verified using more common datasets where it is difficult to identify one to many or many to many violent acts, including the use of weapons. Again, by analyzing the past sequence of events for an individual or a group of people, the detection model can be extended to a prevention model. We will expand this work in the future to address the aforementioned challenges in detecting violent and non-violent activities.

References

1. Wu, J., Zhong, S.h., Liu, Y., 2020. Dynamic graph convolutional network for multi-video summarization. Pattern Recognition 107, 107382.
2. Dang, L.M., Min, K., Wang, H., Piran, M.J., Lee, C.H., Moon, H., 2020. Sensor-based and vision-based human activity recognition: A comprehensive survey. Pattern Recognition 108, 107561.
3. Sasithradevi, A., Roomi, S.M.M., 2020. Video classification and retrieval through spatio-temporal radon features. Pattern Recognition 99, 107099.
4. Cheng, M.; Cai, K.; Li, M. RWF-2000, 2021. An open large scale video database for violence detection. In Proceedings of the 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 10–15, p. 4183–4190.
5. Bermejo Nievas, E.; Deniz Suarez, O.; Bueno García, G.; Sukthankar, R., 2011. Violence detection in video using computer vision techniques. In International Conference on Computer analysis of Images and Patterns; Springer: Berlin/Heidelberg, Germany, p. 332–339
6. Penet, C.; Demarty, C.H.; Gravier, G.; Gros, P., 2012. Multimodal information fusion and temporal integration for violence detection in movies. In Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Kyoto, Japan, p. 2393–2396.
7. Bermejo Nievas, E.; Deniz Suarez, O.; Bueno García, G.; Sukthankar, R., 2011. Violence detection in video using computer vision techniques. In International Conference on Computer analysis of Images and Patterns; Springer: Berlin/Heidelberg, Germany, p. 332–339.
8. Arceda, V.M.; Fabián, K.F.; Gutíerrez, J.C., 2016/ Real Time Violence Detection in Video; IET: Talca, Chile.
9. Das, S.; Sarker, A.; Mahmud, T., 2019. Violence detection from videos using hog features. In Proceedings of the 2019 4th International Conference on Electrical Information and Communication Technology (EICT), Khulna, Bangladesh, p. 1–5.
10. Ding, C., Fan, S., Zhu, M., Feng, W., Jia, B., 2014. Violence detection in video by using 3d convolutional neural networks. In: International Symposium on Visual Computing, p. 551–558. Springer.
11. Xia, Q., Zhang, P., Wang, J., Tian, M., Fei, C.,2018. Real time violence detection based on deep spatiotemporal features. In: Chinese Conference on Biometric Recognition, p. 157–165. Springer.
12. Mu, G., Cao, H., Jin, Q. 2016. Violent scene detection using convolutional neural networks and deep audio features. In: Chinese Conference on Pattern Recognition, p. 451–463. Springer.
13. Meng, Z., Yuan, J., Li, Z., 2017. Trajectory-pooled deep convolutional networks for violence detection in videos. In: International Conference on Computer Vision Systems, p. 437–447. Springer.
14. https://www.nytimes.com/2019/05/14/technology/facebook-live-violent-content.html
15. Serrano Gracia I, Deniz Suarez O, Bueno Garcia G, Kim TK., 2015. Fast fight detection. PLOS ONE 10(4):e0120448.
16. Zhou P, Ding Q, Luo H, Hou X., 2018. Violence detection in surveillance video using low-level features. PLOS ONE 13(10):e0203668.

17. Ribeiro PC, Audigier R, Pham QC. RIMOC, 2016. a feature to discriminate unstructured motions: application to violence detection for video-surveillance, 144, p. 121–143.
18. Yao C, Su X, Wang X, Kang X, Zhang J, Ren J., 2021. Motion direction inconsistency based fight detection for multiview surveillance videos, p. 1–11.
19. Febin IP, Jayasree K, Joy PT., 2020. Violence detection in videos for an intelligent surveillance system using MoBSIFT and movement filtering algorithm. 23(2):p. 611–623.

Various Stock Market Prediction Methods Using Data Science and Machine Learning

Spandan Joshi, Mr. Bakul Panchal*

Research Scholar, Information Technology Department, L. D. College of Engineering, Ahmedabad, Gujarat, India

**Assistant Professor, Information Technology Department, L. D. College of Engineering, Ahmedabad, Gujarat, India*

Abstract

The Stock Market trends are highly volatile in nature. The recent studies have shown that news and sentiments also have an impact on stock prices. Even without it, it is difficult to accurately predict the prices and crisis. Therefore, we use data science, machine learning algorithms and statistics in order to do the huge calculations and make correct predictions. This review proposes a study of 5 such methods involving various algorithms and techniques for the prediction of stock market prices. The methods in this review contain the prediction techniques for predicting stock prices, crisis, and also the impact of news and social media sentiment on the stock price.

Keywords: Stock Market, Prediction, Algorithm.

---

1. Introduction

As we know, the stock market is highly volatile. So it's really hard to predict trends/crisis, such things in the case of stock market. But with the recent advancements in computing and analytics, we have found many ways and algorithms to help us with the task of the prediction of the stock market. This review aims to present 5 such methods developed recently that help us in the prediction of stock market trends. This review comprises of techniques to predict stock prices, crisis, anomalies, effects of news and sentiments on stock prices. This review presents various stock market prediction methods developed using machine learning algorithms and data science that give better predictions and lesser error than the traditional methods.

2. Overview

A. "A Hybrid Approach of Bayesian Structural Time Series With LSTM to Identify the Influence of News Sentiment on Short-Term Forecasting of Stock Price" [1]
This paper presents a method created by combining LSTM (Long Short Term Memory) with the BST (Bayesian Structural Time Series) in order to predict the stock price for short term. This method also takes news and social media sentiment into account while making the prediction.
This method gives significantly less error percentage than the traditional models like ARMA (Auto Regressive Moving Average), VAR (Vector Auto Regression), ARIMA (Auto Regressive Integrated Moving Average) etc. [1]

B. "Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques With a Novelty Feature Engineering Scheme" [2]
This paper presents a method for deciding the investment strategy by combining the data collected from stock market, candle stick patterns, processing it with technical indicators and 8 triagram scheme using a neural network. This method works for prediction of stock prices. Also, this method uses multiple algorithms such as SVM (Support Vector Machine), LR (Linear Regression) etc. according to the situation so it gives better results than the traditional methods. [2]

C. "Novel Stock Crisis Prediction Technique—A Study on Indian Stock Market" [3]
This paper presents a method for prediction of stock market crisis by collecting the data from the National Stock Exchange (NSE) and then selecting features by using a hybrid feature selection algorithm that comprises of the RFE (Recursive Feature Elimination) and BFS (Boruta Feature Selection), an intersection of these two is performed and the final features are selected from the total 42 features. After this, the stock price bubble identification is done by using the RSI (Relative Strength Index) statistics. And finally, after all this data has

been collected, XGBoost (eXtreme Gradient Boosting) and DNN (Delayed Neural Network) Regression methods are used to predict stock crisis. [3]

D. "One Step Ahead: A Framework for Detecting Unexpected Incidents and Predicting the Stock Markets" [4]
This paper presents a method for prediction of stock prices by combining the news and sentiment data with the market data and predicting the behaviour of stock prices by using various machine learning algorithms like the decision tree, random forest etc. This method basically focuses on forecasting the stock prices around the days with news of terrorist attacks. [4]

E. "Forecasting Stock Market Indices Using Padding-Based Fourier Transform Denoising and Time Series Deep Learning Models" [5]
This paper presents a method for predicting the stock market prices by collecting the data from the stock market indices and the time series, denoising the data by using PFTD (Padding based Fourier Transform Denoising) method. Then the deep learning models are trained using methods like RNN (Recurrent Neural Network), LSTM (Long Short Term Memory), GRU (Gated Recurrent Unit). After this, the prediction is done. The prediction accuracy increases by more than 20% after applying the Padding based Fourier Transform.

3. Methodology

A. "A Hybrid Approach of Bayesian Structural Time Series With LSTM to Identify the Influence of News Sentiment on Short-Term Forecasting of Stock Price**"** [1]

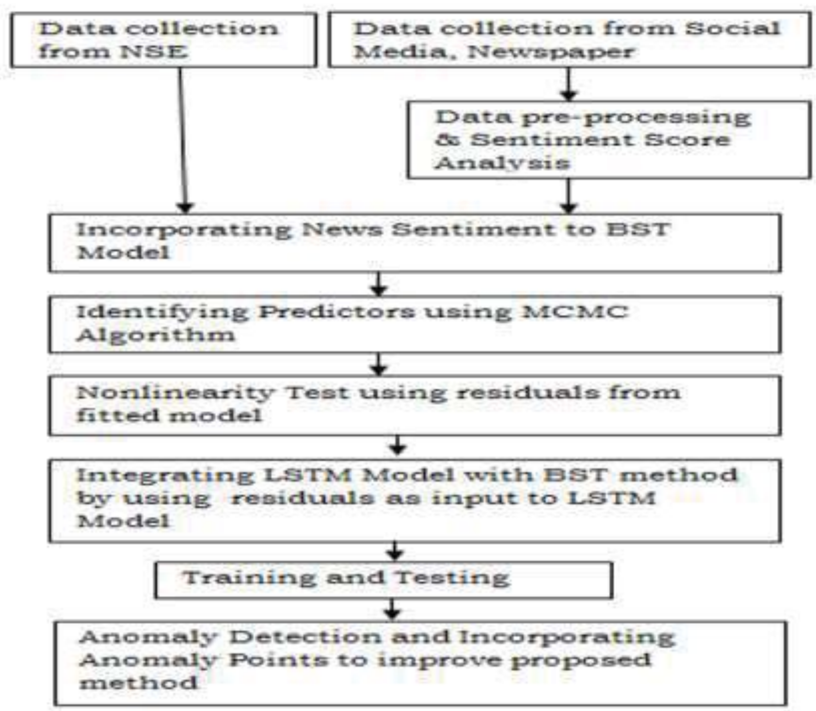


Figure 1. Proposed Workflow Model) [1]

The proposed workflow model of this method is shown in the above figure. The data is collected from the National Stock Exchange, news portals and social media. After this, the data pre-processing and sentiment score analysis is done on the data collected from the news portals and social media. After this, it is incorporated with the BST (Bayesian Structural Time Series). After this, predictors are identified using the MCMC (Markov Chain Monte Carlo) Algorithm. Then the non-linearity test is done from the fitted model. Then the LSTM (Long Short Term Memory) and BST (Bayesian Structural Time Series). Then the training and testing is performed. After this, the anomaly detection is done using the Isolation Forest (iTree) algorithm. This data is incorporated to improve the proposed method. [1]
Tools used: [1]

- R and Python [1]

- NLP and Stanford CoreNLP [1]
- SentiWordNet (It is a lexical resource used for identifying polarity of a statement (for example, polarity of a tweet). It can range from -1 to +1 indicating negative and positive polarities respectively.) [1]
- Slice Matrix-IO (used to detect the anomaly) [1]

Results: [1]

Mean Absolute Percentage Error(MAPE)	
Model	**MAPE**
ARIMA	14.97 %
VAR	14.30 %
Local Linear Trend Model without Sentiment Score (**Model-1**)	12.75 %
Local Linear Trend Model with Sentiment Score (**Model-2**)	10.8 %
Bayesian Nowcasting Model with Sentiment Score & Anomaly Points (**Model-3**)	**3.5 %**

Figure 2. Results of the proposed method [1]

As shown in the figure above, we can see that the Mean Absolute Percentage Error is really low compared to the traditional models ARIMA, VAR etc. The traditional models have this mean absolute percentage error around 13%, but the proposed model has only 3%, which is very less compared to that. Overall, this method is very efficient and a cut above the traditional methods.

B. "Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques With a Novelty Feature Engineering Scheme**"** [2]
This paper presents a method for deciding the investment strategy by combining the data collected from stock market, candle stick patterns, processing it with technical indicators and 8 triagram scheme using a neural network. [2]

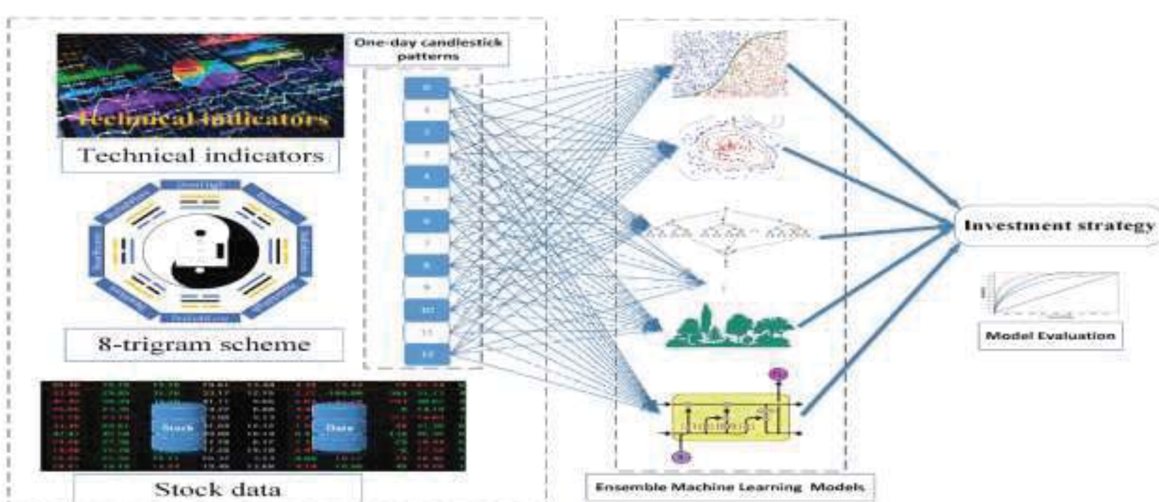


Figure 3. Proposed model [2]

The proposed model is as shown in the figure above. First, the data is collected from the stock market, the technical indicators are calculated, 8-trigram method is used to decide the price movements, then all this, combined with the candle stick patterns is given as an input for the proposed model. The model then decides which method to use and then gives an investment strategy. This method is really useful for stock price prediction.

Group of indicators	Technical indicators
Overlap indicators	Moving Average (MA), Exponential Moving Average (EMA), Double Exponential Moving Average (DEMA), Kaufman's Adaptative Moving Average (KAMA), Simple Moving Average (SMA), Parabolic SAR (SAR)
Momentum indicators	Average Directional Movement Index (ADX), Price Oscillator - Absolute (APO), Balance of Power (BOP), Commodity Channel Index (CCI), Moving Average Convergence/Divergence (MACD), Money Flow Index (MFI), Momentum (MOM), Relative Strength Index (RSI)
Volume indicators	Chaikin A/D Line (AD), Chaikin Oscillator (ADOSC), On Balance Volume (OBV)
Volatility indicators	True Range (TRANGE), Average True Range (ATR), Normalized Average True Range (NATR)

Figure 4. The Technical Indicators [2]

The above figure shows the Technical Indicators taken into account for the process. They are divided into four classes namely Overlap, Momentum, Volume and Volatility Indicators.

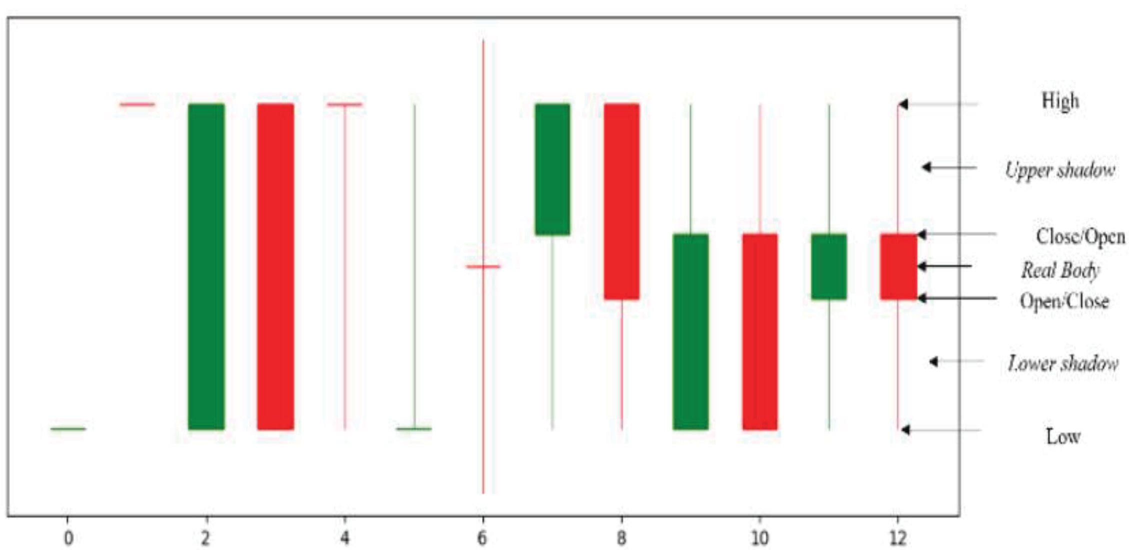
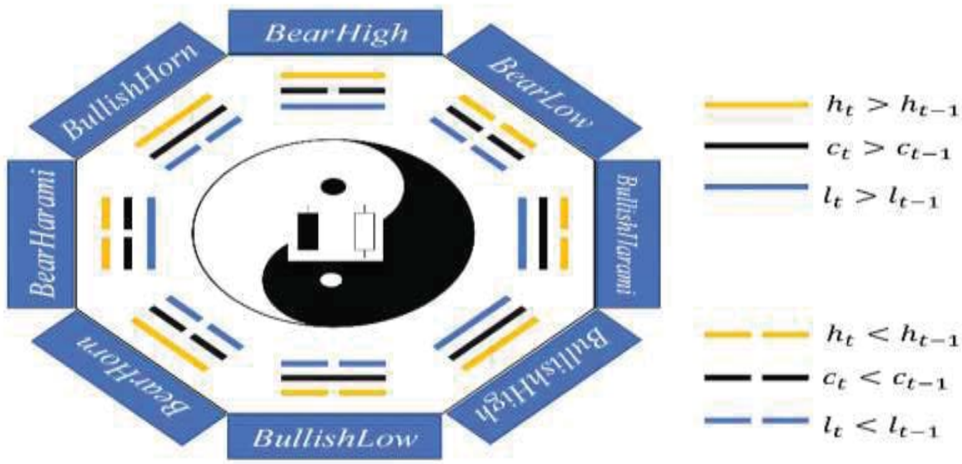


Figure 5. Sample of a Candle Stick Pattern [2]



Symbols	Description
h	h_t represents the highest price at time *t*, h_{t-1} represents the highest price at time *t-1*
c	c_t represents the closing price at time *t*, c_{t-1} represents the closing price at time *t-1*
l	l_t represents the lowest price at time *t*, l_{t-1} represents the lowest price at time *t-1*

Figure 6. The 8 Trigrams of Price Movements [2]

The above figure shows the 8-trigrams of price movements, along with its parameters and description of them.

Algorithm: Model Evaluation

Input: Patterns data which includes feature engineering data and different indicators

Output: *BestModel*, *F1 score*

```
Evaluation (features):
  foreach p in patterns: Generate p_data of p;
      LogisticRegression (p_data);
      GridSearchCV of KNN (p_data);
      GridSearchCV of SVM (p_data);
      GridSearchCV of RF (p_data);
      GridSearchCV of GBDT (p_data);
      LSTM (p_data);
      BestModel = MAX_F1 (LR, KNN, SVM, RF,
GBDT, LSTM);
      Save the best performance model BestModel,
and F1 score for pattern p
Output: List of best performance model, F1 score
for each pattern
```

Figure 7. The Model Evaluation Algorithm [2]

The figure 7 shows the main model evaluation algorithm. The model evaluation algorithm decides which algorithm to use in order to predict the stock prices out of Logistic Regression, KNN (K-Nearest Neighbours), SVM (Support Vector Machine), Random Forest, GBDT (Gradient Boosted Decision Tress) and LSTM (Long Short Term Memory). This model gives the best model as output along with the F1 scores.

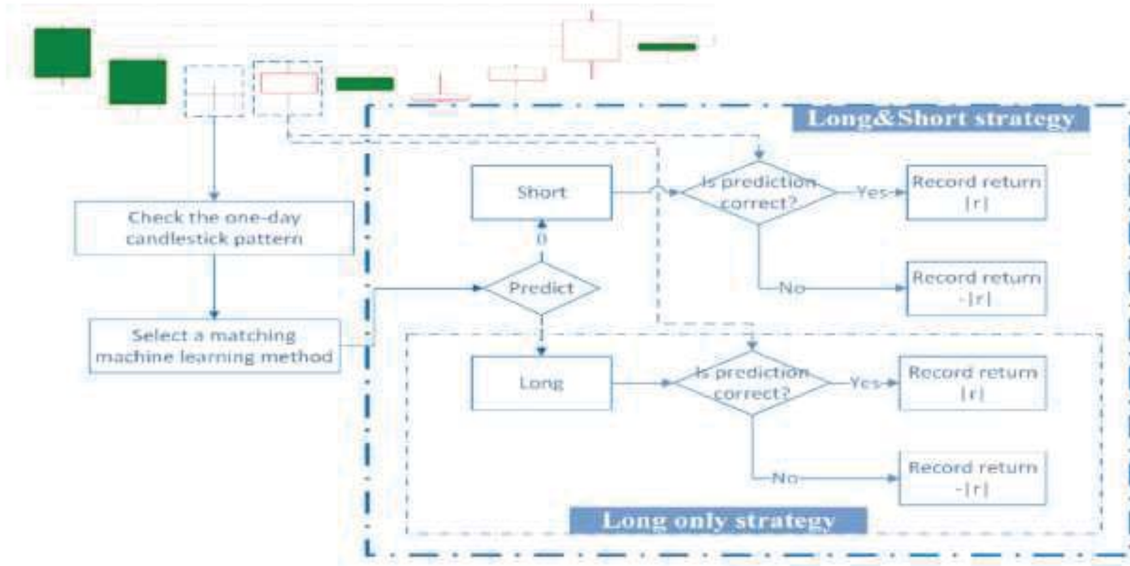


Figure 8. Prediction Flow Diagram [2]

The above figure shows the flow diagram of the proposed model. As we can see, the method has two parts, for the prediction of long and short term stock price respectively. Whether the predictions are correct or incorrect, the result is stored and incorporated in order to improve the existing model.

4. Results

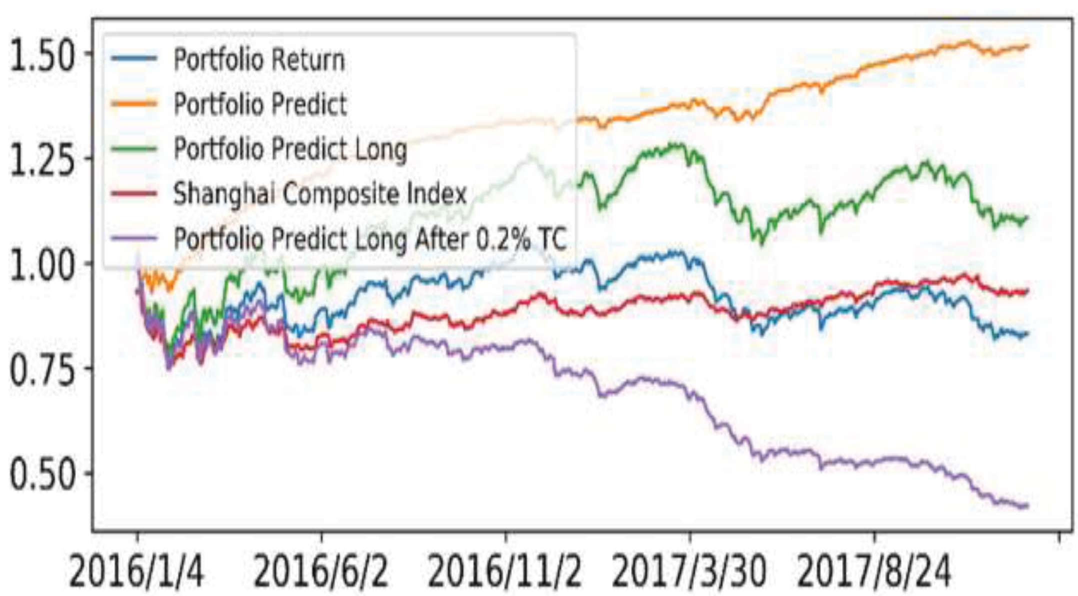


Figure 9. Portfolio Prediction [2]

Table 1. Results of prediction [2]

	Portfolio	**Predict1**	**Predict2**	**Predict3**	**Index**
Max Draw down	26.10%	9.70%	24.10%	59.04%	21.00%
Sharpe Ratio	-0.36	1.65	0.24	-1.76	-0.08
Sortino Ratio	-0.023	0.174	0.029	-0.139	-0.009

The above figures explain the results of prediction, the figure 9 show the graph of portfolio prediction and the table 1 shows the results of prediction. In table 1, Predict1 means the portfolio including the short and long. Predict2 means the portfolio only go long. Predict3 means the portfolio only go long after 0.2% transaction cost. The Index means Shanghai Composite Index.

Overall, this is a good method as it relies not only on one method, but selects one based upon the situation. And thus it has better chances of predicting more accurately than the traditional methods.

C. "Novel Stock Crisis Prediction Technique—A Study on Indian Stock Market**"** [3]

This paper presents a method for prediction of stock market crisis using a hybrid feature selection technique and XGBoost and DNN.

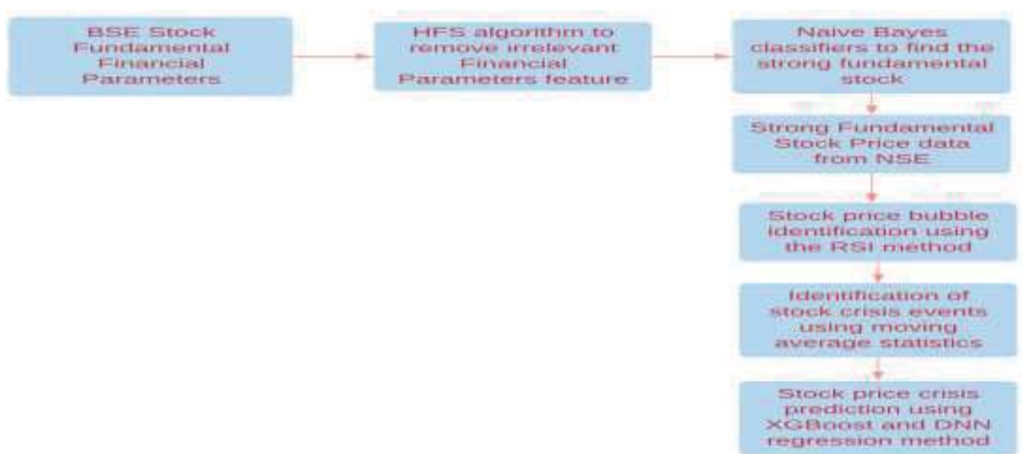


Figure 10. Proposed model [3]

The above figure shows the workflow of the proposed model. First, the data is collected from the stock market, then the HFS (Hybrid Feature Selection) algorithm is used to select the relevant stock features only. Then the Naive Bayes is used to predict the strong fundamental stock. Then the price data is collected from National Stock Exchange. Then the stock price bubble identification is done using the RSI (Relative Strength Index). Then the identification of stock crisis events is done using moving average statistics, then finally the stock price crisis is predicted using XGBoost and DNN.

Figure 11. Parameters List [3]

The above figure shows the list of all 42 parameters taken into account. Out of this list, a few parameters will be taken for the prediction of the stock price.

Figure 12. The Hybrid Feature Selection Algorithm [3]

The above figure shows the HFS (Hybrid Feature Selection) method that comprises of the RFE (Recursive Feature Elimination) and BFS (Boruta Feature Elimination) techniques. After this, the features in the intersection of the results of both are taken as the final features.

Stock Price Bubble Identification: [3]

Relative Strength Index (RSI) statistics are used to find the bubble in stock price. The RSI technical indicator value ranges from 0 to 100. The RSI values below 30 indicates that the stock price is oversold, and RSI values above 70 indicates the overbought levels. When the RSI indicator value reaches above 70, there is a high chance that stock price is falling. Most of the existing work RSI computed is based on 14 days, However, in this approach, they have considered 200 days in RSI to find the stock price bubble.

The reason for doing this is 14 days is used for intra-day trading and not for the long term.

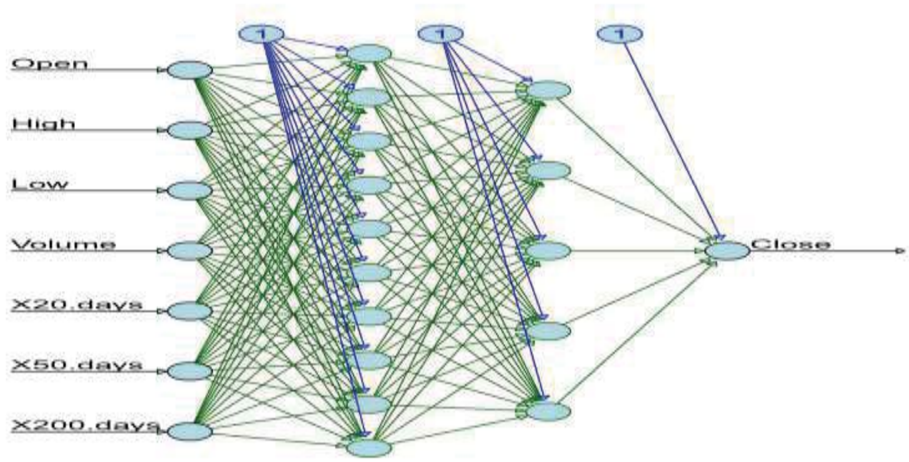
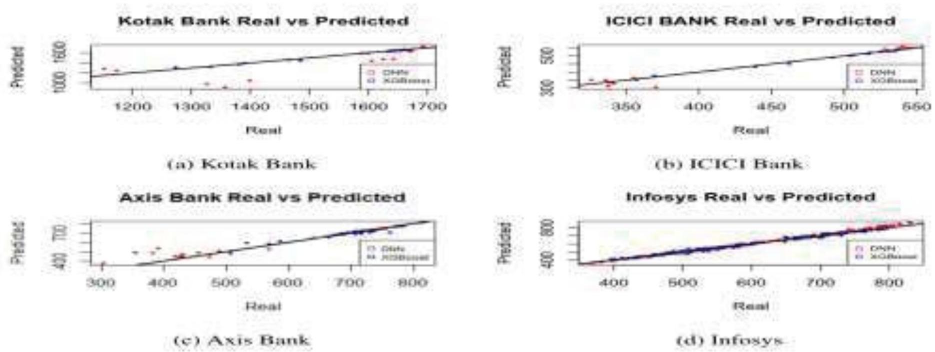


Figure 13. The Delayed Neural Network [3]

The above figure shows the structure of the Delayed Neural Network for the prediction of crisis.
Results: [3]



(
Figure 14. Results of XGBoost and DNN methods [3]

The above figure shows the results of the predictions using both, the XGBoost and DNN.
Overall, this is a good method as it takes only required features into account and finds strong stocks. And it predicts for both, the long and short.

D. "One Step Ahead: A Framework for Detecting Unexpected Incidents and Predicting the Stock Markets**"** [4]
This paper aims to predict the short term stock prices by taking into account the impact of news and social media, especially the terrorism news.

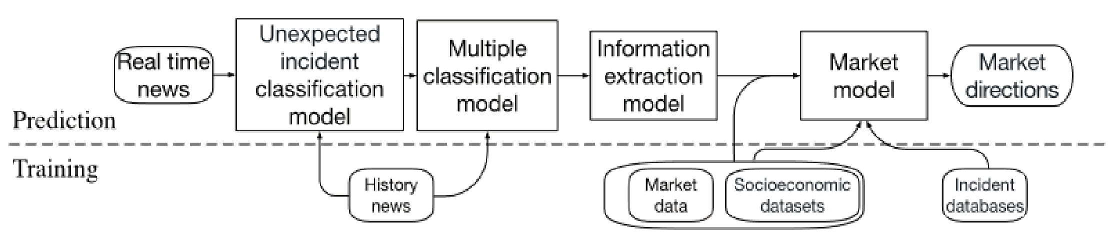


Figure 15. Basic Workflow of the proposed model [4]

The above figure shows the basic workflow of the proposed model. Firstly, the data is collected from the real time news, history news, stock market, incident databases and socioeconomic databases. Then all this data is given as input after performing the data cleaning. Then the model predicts based upon the data using algorithms like Decision Tree, Support Vector Machine, Logistic Regression and Random Forest.

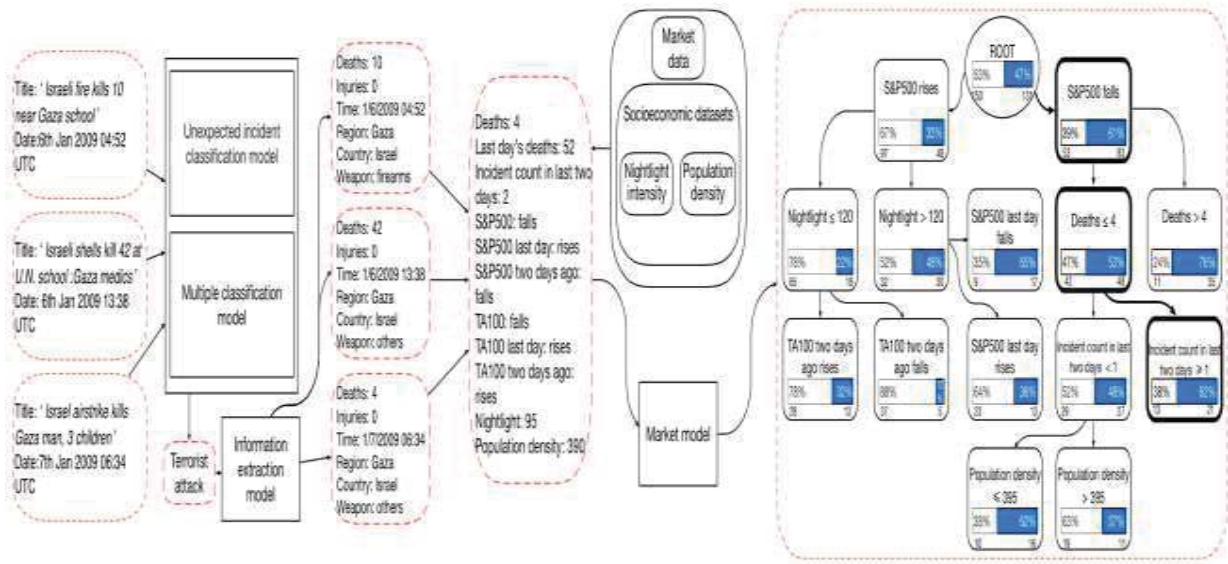


Figure 16. Sample prediction of the proposed model [4]

Data Collection: [4]
News Data is collected from various news portals such as Reuters, Twitter etc. [4]
Market Data is collected from Major Indices data such as IGBC (Columbia Stock Exchange General Index), TA100 (Tel Aviv Stock Exchange General Index), IGBM (Madrid Stock Exchange) etc. [4]
Terrorism Data is obtained from the GTD (Global Terrorism Database). [4]
Socio-Economic Data is obtained from NASA's global nightlight intensity data (used to measure economic developments for areas where incidents happen) and SEDAC's gridded global population density data (to estimate the number of people directly affected by incidents). [4]

Table 2. Results of prediction [4]

	Days with Terrorist Attacks			**Days without Terrorist Attacks**		
Features	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
Market Only	51.6%	47.0%	49.2%	59.5%	53.2%	56.2%
Full Feature	64.8%	53.0%	58.3%	64.4%	56.1%	60.0%

Workflow: [4]
The sample prediction of this proposed model is shown in the figure 16. As we can see, firstly, the news data is collected and it is put through an incident classification model and a multiple classification model. Then if the data is relevant, it is sent to the information extraction model. The information extraction model takes out the information required and eliminated the unnecessary parts of the data. After this, the information is combined with the data from the market and socioeconomic data.

After this, the combined data is provided as an input to the market model, which predicts the market ups and downs. As we can see in the figure, the output has two percentages, the percentage on the left correspond to the market ups and the right to market downs.

Results: [4]

The table 2 shows the results of the prediction. And as we can see, after combining the data, we get nearly 13% boost in prediction during the days with the terrorist attacks. Even in the days without terrorist attacks, we get a precision boost of nearly 5% after combining the data.

Table 3. Algorithm wise precision [4]

Algorithms	Precision	Recall	F1
Decision Tree	70.6%	57.6%	63.4%
Logistic Regression	61.2%	60.0%	60.6%
Random Forest	63.0%	54.4%	58.4%
Support Vector Machine	52.3%	50.0%	51.1%

The table 3 shows the algorithm wise precision for each of the algorithms used. As we can see, the decision tree algorithm has the highest precision of 70.6%.
Overall, this proposed model works fine and improves prediction accuracy for both, days with or without incidents. The model helps a lot in days with incidents, especially terrorism incidents.

E. "Forecasting Stock Market Indices Using Padding-Based Fourier Transform Denoising and Time Series Deep Learning Models**"** [5]

This model uses P-FTD (Padding based Fourier Transform Denoising) technique in order to remove noise (unnecessary part of the data) from the data and makes it more useful for prediction.

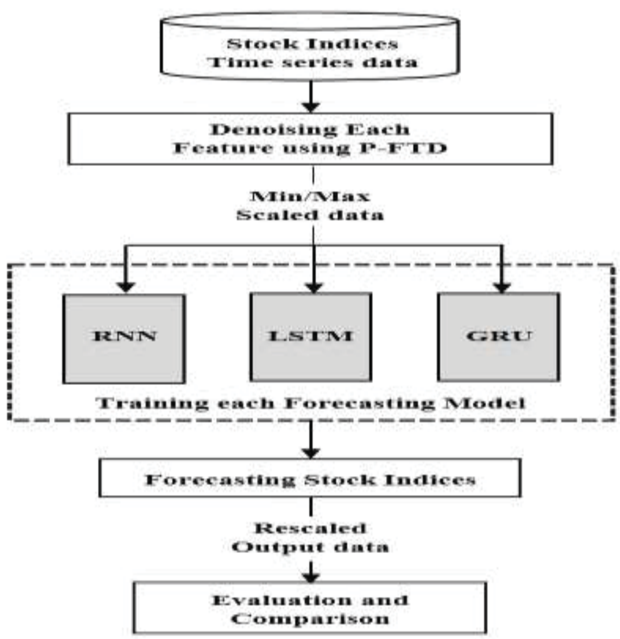


Figure 17. Proposed model's workflow [5]

The above figure shows the workflow of the model proposed. Firstly, the data is collected from the stock indices and the time series. Then each feature is denoised using the padding based fourier transform denoising technique. Then, the minimum or maximum scaled data is given as an input for the training of each forecasting model. The models used are RNN (Recurrent Neural Network), LSTM (Long Short Term Memory), GRU (Gated Recurrent Unit). After the training, the stock indices are forecasted. Then the output data is rescaled and the evaluations and comparisons are done.

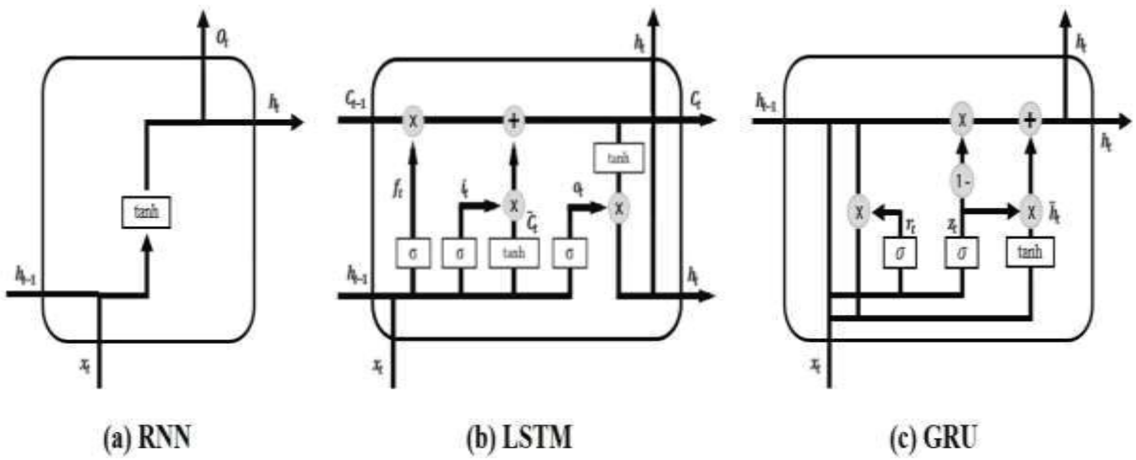


Figure 18. The Cell Diagrams of the models used [5]

The above figure shows the cell diagrams of the models used. It contains the paths and the functions used in them.

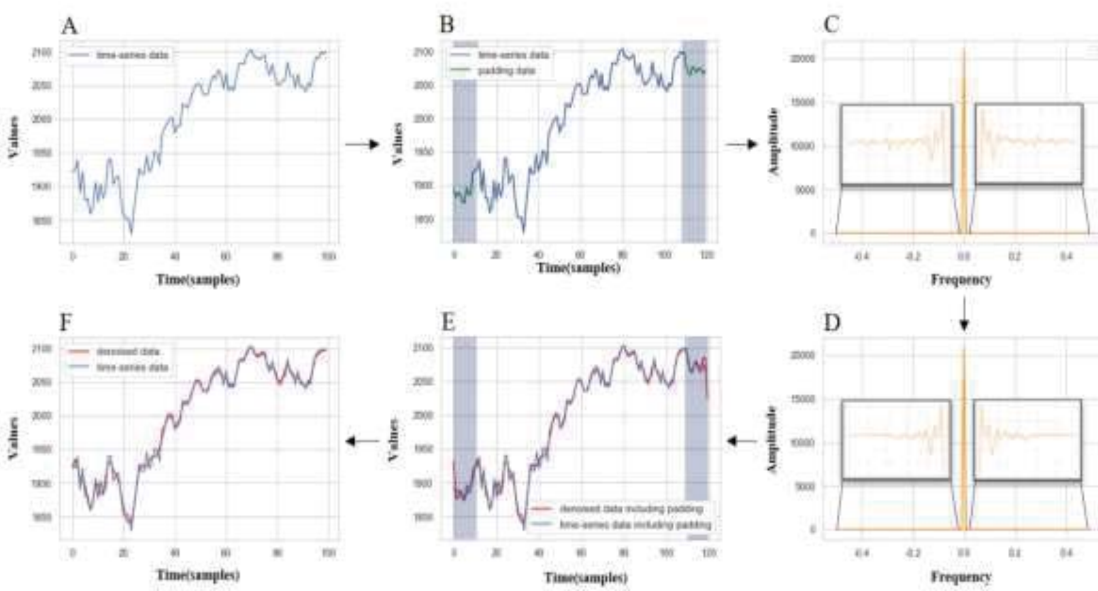


Figure 19. Illustration of the steps of P-FTD [5]

The figure 19 illustrates the steps of the padding based fourier transform denoising technique.

Figure 20. The sample of data before and after performing the P-FTD and the noise removed [5]

The above figure shows the sample of data before and after applying the padding based fourier transfom denoising technique. We can see that how much unnecessary data was removed after applying the padding based fourier transform denoising.

Data Used: [5]
The Standard & Poor 500 (S&P500) index (It is an index that tracks the performance of 500 large companies listed in the US Stock Exchange)
KOSPI (It is the Korea Composite Stock Price Index)
SSE (It is the Shanghai Stock Exchange)

Results

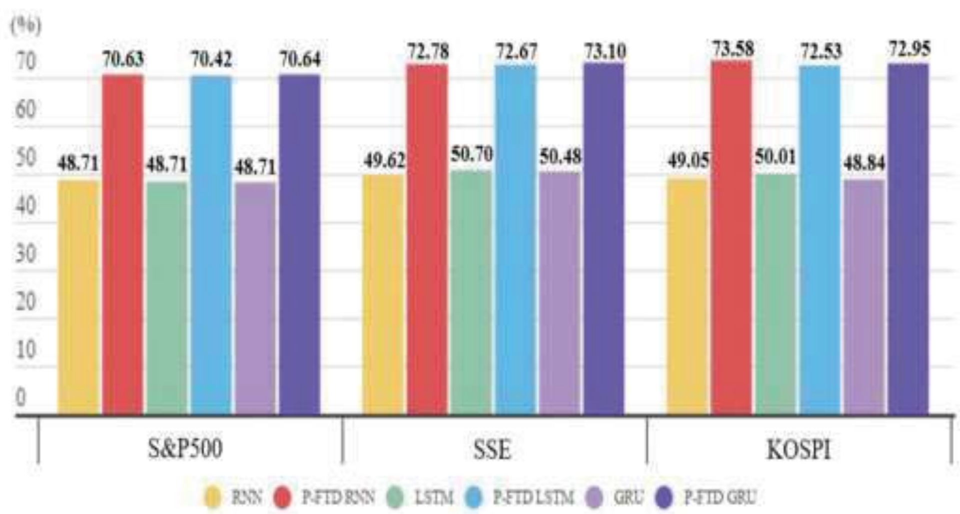


Figure 21. Prediction accuracy comparison before and after [5]

The above figure shows the prediction accuracies for each index and model before and after applying the padding based fourier transform denoising technique. As we can see, after applying the padding based fourier transform denoising, we get a boost of more than 20% for each model's prediction accuracy.
Overall, this model is really useful as it not only reduces the data to process, but also gives a really huge boost to the prediction accuracy.

Summary of Papers:

Sr No.	Name of the Paper	Publication Details	Proposed Conceptualization	Tools Used	Research Possibility
1	A Hybrid Approach of Bayesian Structural Time Series With LSTM to Identify the Influence of News Sentiment on short term forecasting of stock price [1]	IEEE Transactions on Social Computational Systems, 2021	A method created using combining BST with LSTM for better prediction of short term stock prices than the traditional models	R and Python, SentiWordNet, Slice Matrix I-O	The Algorithm can be further optimized in order to minimize the error
2	Stock Trend Prediction using Candlestick Charting and Ensemble Machine Learning Techniques with a Novel Feature Engineering Scheme [2]	IEEE Access Volume 9, 2021	A method involving candlestick patterns and machine learning techniques to predict stock market trends	Python and R	Further improvising can be done in order to gain more accuracy
3	Novel Stock Crisis Prediction Technique: A Study on Indian Stock Market [3]	IEEE Access Volume 9, 2021	A method that uses a Hybrid Feature Selection Technique and DNN and XGBOOST	Python	The method can be improvised in order to gain better performance
4	One Step Ahead: A Framework for Detecting Unexpected Incidents and Predicting the Stock Markets [4]	IEEE Access Volume 9, 2021	A model created to predict the short term effects on stocks due to incidents	Python	A generalized model can also be created for all the anomalies (not just terrorist attacks)
5	Forecasting Stock Market Indices using Padding Based Fourier Transform Denoising and Time Series Deep Learning Models [5]	IEEE Access Volume 9, 2021	A method that uses Padding based Fourier Transform in order to denoise the data and then makes prediction for stock market indices	Python	The accuracy can be boosted even further by improvising

5. Conclusion

We studied several methods for the prediction of stock market. All of these methods surpassed the traditional methods by a lot. But the stock market is highly volatile, so even these methods can't give us accurate results all the time. We can also see that news, rumours etc. also have a good amount of impact on stock price for a short time.

References

1. P. Ray, B. Ganguli and A. Chakrabarti, "A Hybrid Approach of Bayesian Structural Time Series With LSTM to Identify the Influence of News Sentiment on Short-Term Forecasting of Stock Price", in IEEE Transactions on Computational Social Systems, vol. 8, no. 5, pp. 1153-1162, Oct. 2021, doi: 10.1109/TCSS.2021.3073964.
2. Y. Lin, S. Liu, H. Yang and H. Wu, "Stock Trend Prediction Using Candlestick Charting and Ensemble Machine Learning Techniques With a Novelty Feature Engineering Scheme", in IEEE Access, vol. 9, pp. 101433-101446, 2021, doi: 10.1109/ACCESS.2021.3096825.
3. N. Naik and B. R. Mohan, "Novel Stock Crisis Prediction Technique—A Study on Indian Stock Market", in IEEE Access, vol. 9, pp. 86230-86242, 2021, doi: 10.1109/ACCESS.2021.3088999.
4. Z. Li, S. Lyu, H. Zhang and T. Jiang, "One Step Ahead: A Framework for Detecting Unexpected Incidents and Predicting the Stock Markets", in IEEE Access, vol. 9, pp. 30292-30305, 2021, doi: 10.1109/ACCESS.2021.3059283.
5. D. Song, A. M. Chung Baek and N. Kim, "Forecasting Stock Market Indices Using Padding-Based Fourier Transform Denoising and Time Series Deep Learning Models", in IEEE Access, vol. 9, pp. 83786-83796, 2021, doi: 10.1109/ACCESS.2021.3086537.
6. A. Harvey and S. Koopman, "Structural time series models," in StatsRef: Statistics Reference Online. Hoboken, NJ, USA: Wiley, 2014, doi: 10.1214/19-AOAS1265.
7. Y. Bao, C. Quan, L. Wang, and F. Ren, "The role of pre-processing in twitter sentiment analysis," in Proc. Int. Conf. Intell. Comput. Cham, Switzerland: Springer, 2014, pp. 615–624, doi: 10.1007/978-3-030-82469-3.
8. J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," J. Comput. Sci., vol. 2, no. 1, pp. 1–8, Mar. 2011, doi: 10.48550/arXiv.1010.3003.
9. I. K. Nti, A. F. Adekoya, and B. A. Weyori, ‘‘A systematic review of fundamental and technical analysis of stock market predictions,’’ Artif. Intell. Rev., vol. 53, pp. 1–51, Aug. 2019, doi: 10.1007/s10462-019-09754-z.
10. P. Yu and X. Yan, ‘‘Stock price prediction based on deep neural networks,’’ Neural Comput. Appl., vol. 32, no. 6, pp. 1609–1628, Mar. 2020, doi: 10.1007/s00521-019-04212-x.
11. A. Mundra, S. Mundra, V. K. Verma, and J. S. Srivastava, ‘‘A deep learning based hybrid framework for stock price prediction,’’ J. Intell. Fuzzy Syst., vol. 38, no. 5, pp. 5949–5956, May 2020.
12. B. J. Blair, S. H. Poon, and S. J. Taylor, ‘‘Forecasting S&P 100 volatility: The incremental information content of implied volatilities and high-frequency index returns,’’ in Handbook of Quantitative Finance and Risk Management, C. F. Lee, A. C. Lee, and J. Lee, Eds. Boston, MA, USA: Springer, 2010, doi: 10.1007/978-0-387-77117-5_88.
13. A. Safari and A. A. Ghavifekr, ‘‘International stock index prediction using artificial neural network (ANN) and Python programming,’’ in Proc. 7th Int. Conf. Control, Instrum. Autom. (ICCIA), Feb. 2021, pp. 1–7, doi: 10.1109/ICCIA52082.2021.9403580.
14. X. Li, H. Xie, R. Wang, Y. Cai, J. Cao, F. Wang, H. Min, and X. Deng, ‘‘Empirical analysis: Stock market prediction via extreme learning machine,’’ Neural Comput. Appl., vol. 27, no. 1, pp. 67–78, Jan. 2016, doi: 10.1007/s00521-014-1550-z
15. Y. Xu and S. B. Cohen, ‘‘Stock movement prediction from tweets and historical prices,’’ in Proc. 56th Annu. Meeting of Assoc. Comput. Linguistics (COLING), 2018, pp. 1970–1979, doi: 10.18653/v1/P18-1183.
16. Z. Hu, W. Liu, J. Bian, X. Liu, and T. Y. Liu, ‘‘Listening to chaotic whispers: A deep learning framework for news oriented stock trend prediction,’’ in Proc. 11th ACM Int. Conf. Web Search Data Mining (WSDM), 2018, pp. 261–269, doi: 10.1145/3159652.3159690.
17. H. Lee, M. Surdeanu, B. MacCartney, and D. Jurafsky, ‘‘On the importance of text analysis for stock price prediction,’’ in Proc. LREC, 2014, pp. 1170–1175, url: "http://www.lrec-conf.org/proceedings/lrec2014/pdf/1065_Paper.pdf".
18. S. M. Idrees, M. A. Alam, and P. Agarwal, ‘‘A prediction approach for stock market volatility based on time series data,’’ IEEE Access, vol. 7, pp. 17287–17298, 2019, doi: 10.1109/ACCESS.2019.2895252.
19. X. Zhang, S. Qu, J. Huang, B. Fang, and P. Yu, ‘‘Stock market prediction via multi-source multiple instance learning,’’ IEEE Access, vol. 6, pp. 50720–50728, 2018, doi: 10.1109/ACCESS.2018.2869735
20. M. Nabipour, P. Nayyeri, H. Jabani, S. Shahab, and A. Mosavi, ‘‘Predicting stock market trends using machine learning and deep learning algorithms via continuous and binary data; a comparative analysis,’’ IEEE Access, vol. 8, pp. 150199–150212, 2020, doi: 10.1109/ACCESS.2020.3015966
21. M. Nabipour, P. Nayyeri, H. Jabani, A. Mosavi, E. Salwana, and

S. Salwana, ‘‘Deep learning for stock market prediction,’’ Entropy, vol. 22, no. 8, p. 840, Jul. 2020, doi: 10.3390/e22080840

A Review on Convolution Neural Network based Ophthalmic Disease Detection using Fundus Images

Parth Bipinbhai Nirmal[a*], Bakul Panchal[b], Manoj Patel[c]

[a*]*P.G. Scholar, Information Technology Department, L.D. College of Engineering, 380015, India*
[b]*Asst. Professor, Information Technology Department, L.D. College of Engineering, 380015, India*
[c]*Asst. Professor, Information Technology Department, L.D. College of Engineering, 380015, India*

Abstract

It is estimated that approximately 2.2 billion individuals worldwide suffer from visual impairments, predominantly stemming from age-related ophthalmic diseases that may result in blindness if left unaddressed. The present investigation endeavors to detect such afflictions at an early stage by employing retinal fundus images acquired from online datasets. Preprocessed fundus images are analyzed utilizing a convolutional neural network (CNN) followed by classification of the CNN output. Various online datasets serve as training and testing data for the proposed model. Performance of the proposed model is evaluated vis-à-vis other optimized models based on metrics such as precision, accuracy, specificity, recall, and F1 score. Such an approach holds great promise for medical professionals tasked with managing ocular disorders.

Keywords: Self-supervised learning, retinal disease classification, ophthalmic diseases

1. Introduction

Early detection and timely treatment of ocular diseases are vital to prevent vision loss and improve quality of life. Age-related macular degeneration (AMD), diabetic retinopathy (DR), glaucoma, and cataracts are among the most common ophthalmic disorders [1] [2] [3] [4]. Diabetic retinopathy results from elevated blood glucose levels that cause damage to the retinal blood vessels. AMD is caused by macular injury, while glaucoma is the result of fluid buildup in front of the eye that raises pressure and damages the optic nerve [5]. Cataracts are characterized by clouding of the eye's lens. The projected number of DR cases worldwide may increase to 191.0 million by 2030 due to the rising number of diabetic patients [4] [5].

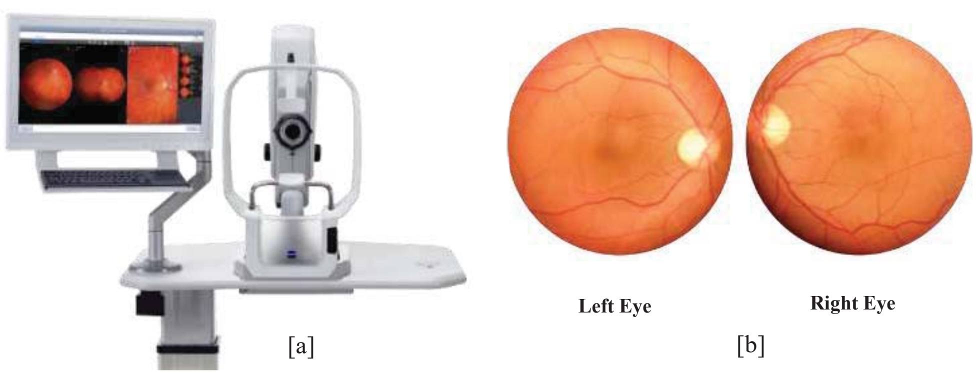


Fig. 1. [a] Clarus 700 camera [b] Color fundus photographs of left eye and right eye [1]

*Parth Nirmal
E-mail address: pnirmal.pn@gmail.com

Deep learning techniques have opened up various possibilities in clinical medicine and are being rapidly integrated for image analysis [3] [4]. Diagnostic tests using retinal imaging capture multiple fundus and retinal photos [5]. The most common fundus cameras used for this purpose are Clarus 700 [4] and CIRRUS 6000 [2]. These images can be analyzed to diagnose multiple ophthalmic disorders, including AMD, DR, cataracts, and glaucoma [5].

The field of ophthalmology has greatly benefited from the advancements in computer vision and machine learning [5]. In recent years, various studies have been conducted on developing optimized convolution neural network models for the detection of multiple eye diseases [5]. These studies aim to provide a reliable and efficient system for the automated detection of eye diseases, including age-related macular degeneration, diabetic retinopathy, and glaucoma, among others. One such study proposed a multi-class multi-label classification system based on transfer learning using convolutional neural networks [6] [7].

Another study focused on detecting age-related macular degeneration from fundus images by utilizing a multi-view instance discrimination method, which captures discriminative regions in the image. Furthermore, it employed Guided Grad-CAM for high-resolution discriminative visualizations and achieved 89.6% accuracy in detecting age-related macular degeneration [6] [7]. In addition, a study on rotation-oriented collaborative self-supervised learning showed promising results for retinal disease diagnosis by using rotation-based augmentation and multi-view feature fusion [1] [6] [7].

Another approach proposed a hybrid algorithm that enhances color retinal fundus images by utilizing a Wiener filter and Contrast Limited Adaptive Histogram Equalization (CLAHE) technique [8]. This study demonstrated significant improvements in image enhancement, noise reduction, and artifact elimination compared to the traditional CLAHE method. Furthermore, a study on multi-class multi-label ophthalmologic al disease detection used transfer learning based convolutional neural network approaches to classify eight categories of ocular fundus images with high accuracy, particularly in detecting cataract disease and hypertension [9] [10].

Overall, these studies highlight the potential of utilizing convolutional neural networks and machine learning for automated and accurate detection of various eye diseases. The development of such systems can significantly aid healthcare professionals in early diagnosis and timely treatment of eye diseases, ultimately leading to better patient outcomes. The use of deep learning techniques in ophthalmology has resulted in significant advancements in the field of medical image analysis. These five papers provide promising approaches for detecting various eye diseases using deep learning methods, achieving state-of-the-art performance on benchmark datasets. The proposed techniques have the potential to improve the accuracy and efficiency of medical image analysis, facilitating early diagnosis and treatment of various eye diseases.

In summary, these five research papers aim to improve the detection and diagnosis of eye diseases using various techniques in image processing and machine learning. The first paper proposes an optimized convolutional neural network for multiple eye disease detection with high accuracy. The second paper focuses on age-related macular degeneration detection using a deep learning-based approach for the AMD Detection from Fundus Images (ADAM) challenge. The third paper introduces a rotation-oriented collaborative self-supervised learning method for retinal disease diagnosis. The fourth paper implements transfer learning with a convolutional neural network for multi-class multi-label ophthalmological disease detection. Lastly, the fifth paper proposes a hybrid algorithm that combines a Wiener filter and CLAHE technique for enhancing color retinal fundus images. Overall, these papers demonstrate the potential for machine learning and image processing techniques to improve the accuracy and efficiency of eye disease diagnosis, potentially leading to earlier detection and better patient outcomes.

2. Literature Review

2.1. Paper 1: Optimized convolution neural network based multiple eye disease detection [1]

The aim of this study is to develop an automated approach for detecting multiple diseases in retinal fundus images using a combination of maximum entropy transformation, convolutional neural network (CNN), and multiclass support vector machine (SVM) classifier. The fundus image is first subjected to maximum entropy transformation to preserve the maximum amount of information. The pre-processed image is then fed into a CNN consisting of two convolution layers, a Max-pooling layer, and an activation function to extract features.

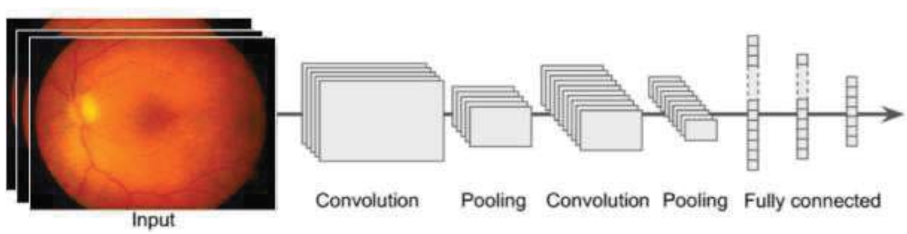
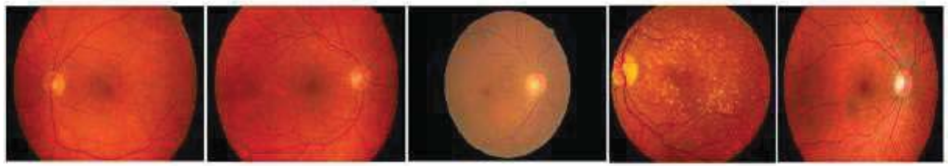


Fig. 2. Convolution Neural Network Architecture [1]

The hyperparameters of the CNN are optimized using the Flower Pollination Optimized Algorithm (FPOA) to improve accuracy. The extracted features are then fed to a multiclass SVM classifier to determine the type of the diseased image. The architecture of the proposed method is shown in Fig. 2, while the different types of diseases for classification are illustrated in Fig. 3.

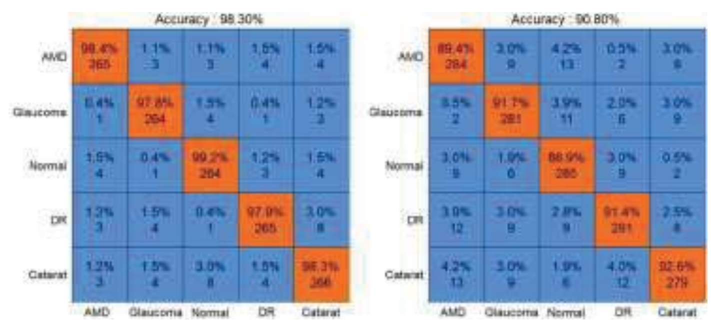


[a] [b] [c] [d] [e]

Fig. 3. (A) DR image (B) Glaucoma image (C) Cataract image (D) AMD image (E) Normal image [1]

The research utilized the Ocular Disease Intelligent Recognition (ODIR) database, which contains a large collection of color fundus images of both left and right eyes of 5000 patients, each accompanied by diagnostic keywords from medical professionals. Training progress plots of the convolutional neural network (CNN) were presented in Fig 4, which displays the training accuracies and losses on the Y-axis and iterations on the X-axis, demonstrating the effectiveness of the proposed approach.

Fig. 4. Training progress of CNN [1]



[a] [b]

Fig. 5. Confusion matrix of (a) Proposed method and (b) Simple CNN [1]

If table footnotes should be used, place footnotes to tables below the table body and indicate them with superscript lowercase letters. Be sparing in the use of tables and ensure that the data presented in tables do not duplicate results described elsewhere in the article.

2.2. ADAM Challenge: Detecting Age-Related Macular Degeneration from Fundus Images [2]

This study proposes an architecture for detecting age-related macular degeneration (AMD) from fundus images, which involves four tasks. The first task is the classification of AMD and non-AMD, followed by the detection and segmentation of the optic disc, localization of the fovea, and detection and segmentation of lesions. The challenge involves the segmentation of five types of lesions (drusen, exudate, hemorrhage, scar, and other); however, only three types of lesions are illustrated in Fig 6. The proposed architecture provides a comprehensive approach to detect AMD and identify the type of lesions present in the fundus images.

The study employed the ADAM dataset, comprising 1200 retinal fundus images stored in JPEG format with 8 bits per color channel, provided by Zhongshan Ophthalmic Center, Sun Yat-sen University, China. The fundus images were captured using a Zeiss Visucam 500 fundus camera with a resolution of 2124 × 2056 pixels (824 images) and a Canon CR-2 device with a resolution of 1444 × 1444 pixels (376 images).

To extract image features, the proposed framework employed 15 finding networks, as illustrated in Fig 7.

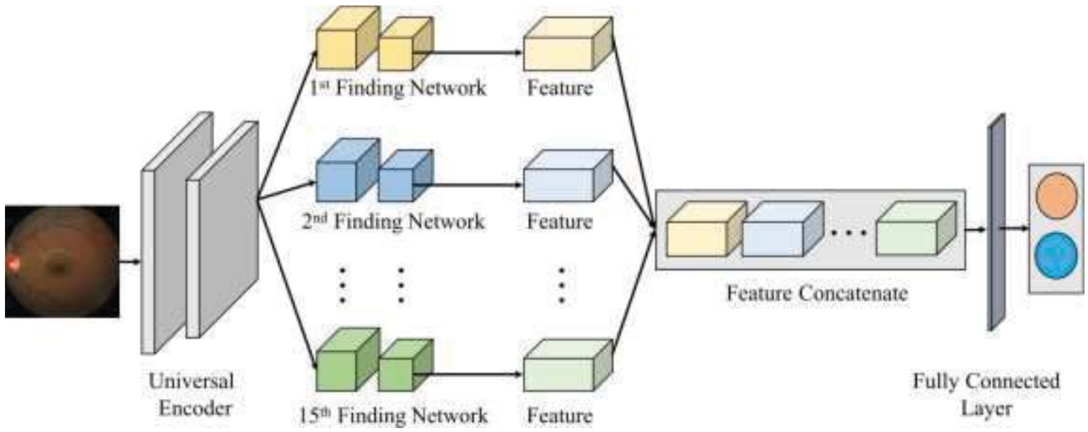


Fig. 7. ADAM proposed framework [2]

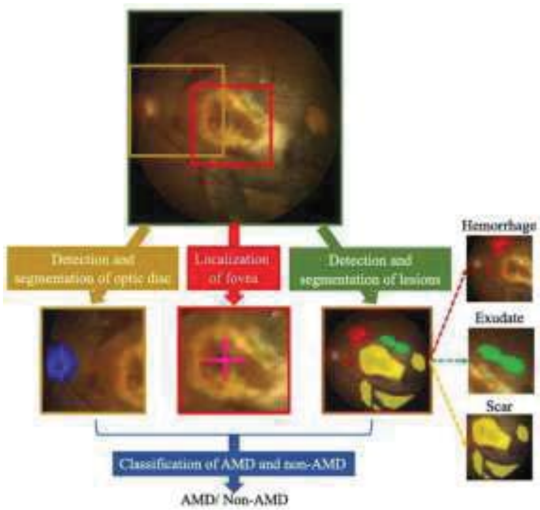


Fig. 6. The four tasks of the Proposed Methodology [2]

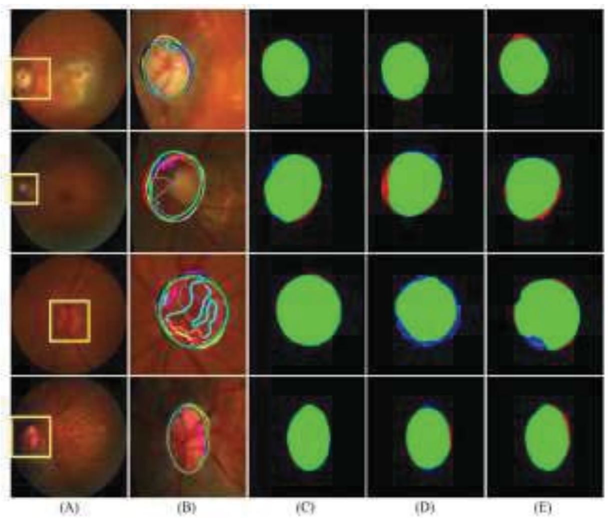


Fig. 8. (A) Original fundus image (B) Boundaries of optic disc (C - E) Segmentation results [2]

The results depicted in Figure 8 illustrate the qualitative examples of the optic disc segmentation process. The original fundus image, the boundaries of the segmented optic disc obtained by ten teams, and segmentation results of the top three teams compared with the ground truth are presented. The proposed framework, which involves segmentation, detects age-related macular degeneration with 97% accuracy.

2.3. Rotation-Oriented Collaborative Self-Supervised Learning for Retinal Disease Diagnosis [3]

The paper employs a multi-view instance discrimination method to learn feature representations that are similar to transformed versions of the input image and distinct from those of other images. The transformations involve rotation, scaling, cropping, and adjusting the image's brightness, contrast, and saturation. The proposed method captures more discriminative regions in the input image, as demonstrated in the sixth column of Fig 10. Additionally, the Guided Grad-CAM technique generates high-resolution discriminative visualizations that reveal salient structures such as vessels.

The visual similarity of the retrieved images with the test image is high, as evident from Fig 11. This feature can assist in assigning a correct class to the test image, as observed by the majority vote of the 5-nearest neighbors. The proposed framework achieved an accuracy of 89.6% in detecting age-related macular degeneration (AMD) and 98.5% accuracy in detecting pathologic myopia (PM), as demonstrated in Fig 11.

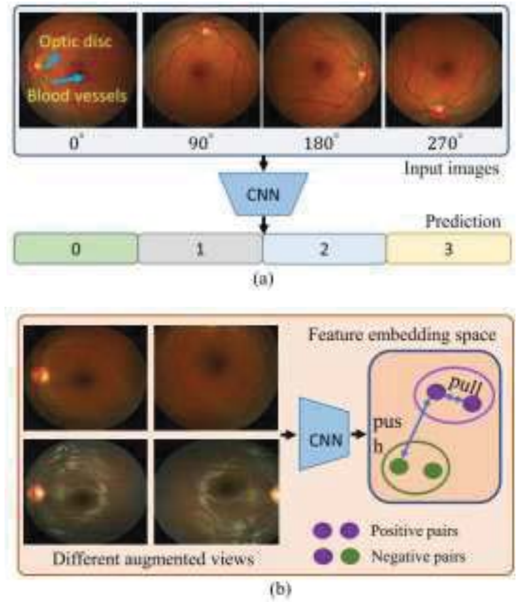


Fig. 9. (A) The rotation prediction task (B) Images generated from one patient image under different augmentations (positive pairs) should be similar in the embedding space, while images from different patients (negative pairs) should be dissimilar [3]

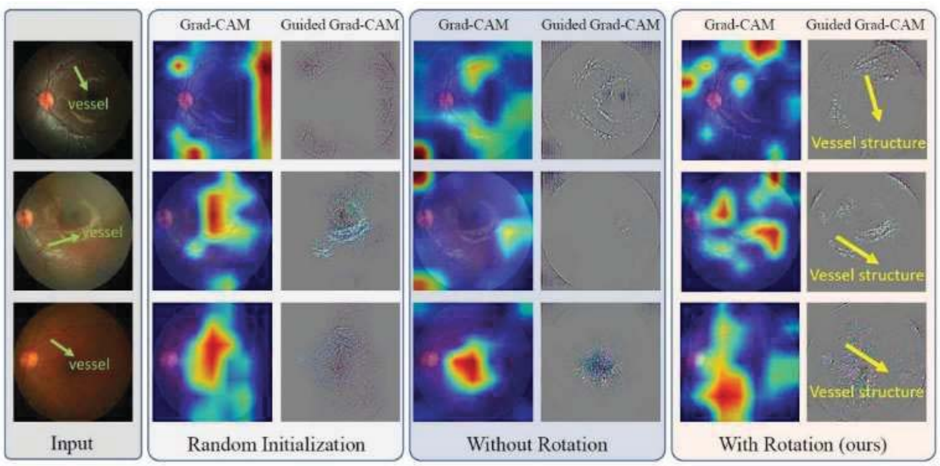


Fig. 10. The visualization of the features of three examples from the proposed models [3]

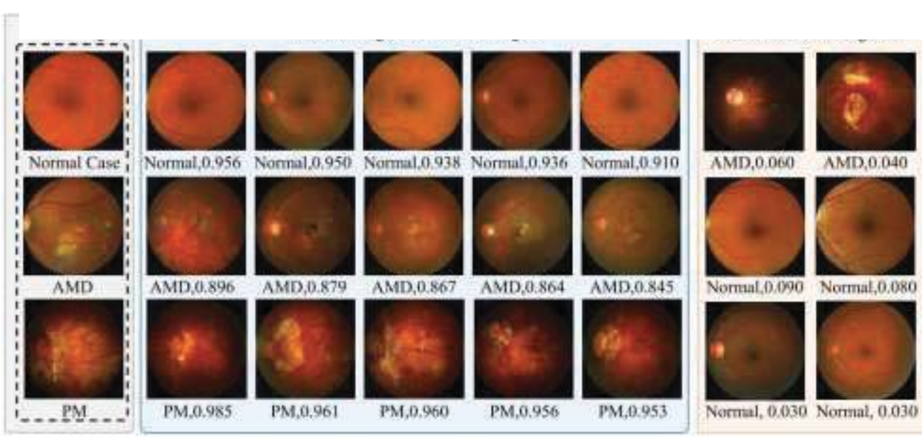


Fig. 11. The final result is obtained by a k-Nearest Neighbor [3]

2.4. Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network [4]

The study puts forth a new approach for multi-class multi-label classification of ophthalmological diseases using two transfer learning-based CNN models. The proposed models use pre-trained input CNN architectures for accurate classification of ocular diseases. The Ocular Disease Intelligent Recognition (ODIR) dataset is utilized to implement the proposed method. Figure 12 illustrates the two pre-trained input CNN architectures used for multi-class multi-label classification of ocular diseases.

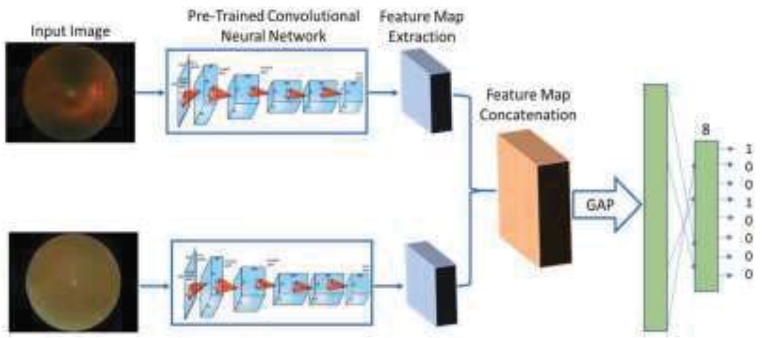


Fig. 12. Proposed two input Convolution Neural Network architecture [4]

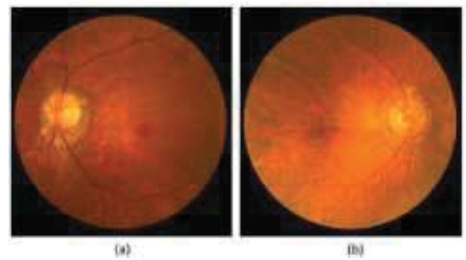


Fig. 13. (A) Normal (B) Having moderated retinopathy Fundus Images [4]

S.No.	Class	Training	Offsite testing	Onsite testing	Total
1.	Normal	1135	161	324	1620
2.	Diabetes	1131	162	323	1616
3.	Glaucoma	207	30	58	307
4.	Cataract	211	32	64	243
5.	AMD	171	25	47	295
6.	Hypertension	94	14	30	138
7.	Myopia	177	23	49	249
8.	Other abnormalities	944	134	268	1346

Table 1. Class distribution of eight categories of ocular images [4]

Class ↓/Metrics →	ACC
Normal	0.40
Glaucoma	0.54
Diabetic retinopathy	0.89
AMD	0.88
Hypertension	0.95
Cataract	0.97
Myopia	0.90
Other abnormalities	0.44

Table 2. Accuracy of various eight categories of ocular fundus images [4]

Table 1 presents the complete list of eight categories of ocular fundus images, including information on each category such as the number of images available in the dataset. On the other hand, Fig 13 illustrates the comparison between two fundus images: one fundus image depicts a normal case without any signs of disease, while the other shows an image of moderate Diabetic retinopathy disease. This comparison provides an opportunity to observe the differences between a healthy fundus image and an image affected by a disease. The accuracy of detecting all the eight diseases is shown in Table 2. The proposed method achieves high accuracy in detecting Cataract disease and Hypertension.

Artwork has no text along the side of it in the main body of the text. However, if two images fit next to each other, these may be placed next to each other to save space, see Fig 1.

2.5. A Hybrid Algorithm to Enhance Color Retinal Fundus Images Using a Wiener Filter and CLAHE [5]

The paper proposes the use of the CLAHE technique for improving the color fundus image and reducing noise. Applying the wiener filter pass before applying CLAHE results in improved detail in the green channel, reduced noise, and elimination of artificial boundaries associated with the CLAHE algorithm. The framework for enhancing the input image is shown in Fig 14, where the wiener filter is used to filter out noise from the corrupted image.

Fig 15 depicts the enhanced retinal color fundus images obtained by using the CLAHE enhancement technique with the red and blue channel. Furthermore, the proposed method is also shown in the figure, which includes implementing and applying the wiener filter for noise reduction first, followed by the application of CLAHE for further enhancement.

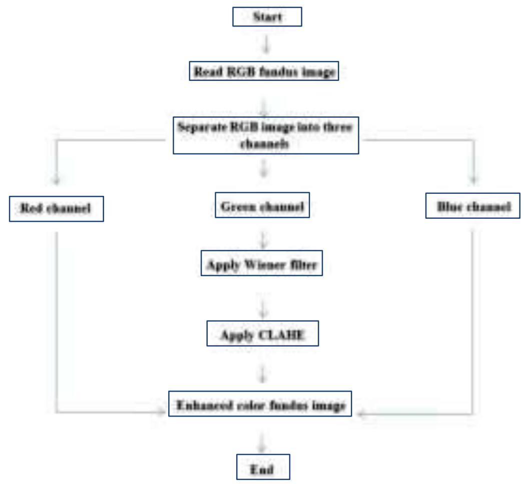


Fig. 14. Architecture of image enhancement process using CLAHE and Wiener Filter [5]

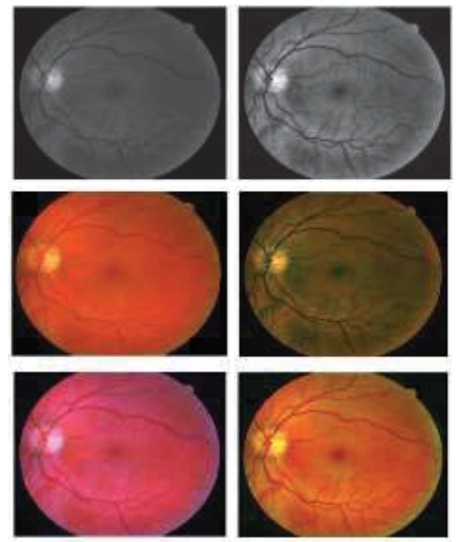


Fig. 15. Retinal colour fundus image of different condition: (A) Normal grey scale Image (B) grey scale CLAHE result Image (C) Original Image (D) CLAHE to red channel (E) CLAHE to blue channel (F) Proposed method image [5]

3. Summarization

Sr No.	Name of the Paper	Publication Detail	Proposed Conceptualization	Research Possibility	Tools and Technology Used
1.	Optimized convolution neural network based multiple eye disease detection	Published in Computers in Biology and Medicine 146 (2022) 105648	Identification of the age-related eye diseases at an early stage using retinal fundus images taken from online dataset and pre-processed using maximum entropy transformation.	More effective architecture can be used for future development in terms of performance	Flower pooling optimization algorithm, Ocular Disease Intelligent Recognition (ODIR) Dataset
2.	ADAM Challenge: Detecting Age- Related Macular Degeneration from Fundus Images	Published In IEEE Transactions on Medical Imaging, Vol. 41, No. 10, October 2022	Age-Related macular degeneration detection using Automatic Detection challenge on Age-related Macular degeneration(ADAM)	develop more robust and novel algorithms for AMD screening from fundus images	Zeiss Visucam 500 fundus camera, Canon CR-2 device
3	Rotation- Oriented Collaborative Self-Supervised Learning for Retinal Disease Diagnosis	Published IEEE Transactions on Medical Imaging, Vol. 40, No. 9, September 2021	Present a rotation-oriented collaborative method that explores rotation- related features. Which explore the invariant property used for retinal disease classification	Can work better with huge amount of unlabeled data	Ichallenge-AMD dataset, EyePACS dataset, ResNet,

4	Multi-class multi-label ophthalmologic al disease detection using transfer learning based convolutional neural network	Published in Elsevier Biomedical Signal Processing and Control 2020	Eight class diseases detection with using two pre-processed convolution neural network architecture	EfficientNet CNN architecture can give more accurate results	ODIR Dataset, Pre-Trained model
5	A Hybrid Algorithm to Enhance Colour Retinal Fundus Images Using a Wiener Filter and CLAHE	Published In IEEE Journal of Biomedical and Health Informatics, Vol. 25, No. 9, September 2021	Presented a Hybrid algorithm with CLAHE and Wiener filter for the enhancement of original fundus image	Resizing the image results reduction in time consumption	contrast limited adaptive histogram equalization (CLAHE), OIA-ODIR with 5000 patients data

4. Conclusion

Using convolution neural network along with pre-processed fundus images, early detection of various ophthalmic diseases such as AMD, cataract, glaucoma and DR was performed with the high accuracy rate as compare to other optimized CNN models. Optimization of the fundus image helps in increasing the efficiency and accuracy of the proposed model. In comparison with the other optimized networks, the proposed model provided better performance metrics in terms of precision, recall, specificity and F1 score. Proper annotation of the image can be implemented in future to get more accurate results. The study also suggests the enhancement of network architecture to improve the performance. The paper proposes the use of the CLAHE technique for improving the color fundus image and reducing noise. Applying the wiener filter pass before applying CLAHE results in improved detail in the green channel, reduced noise, and elimination of artificial boundaries associated with the CLAHE algorithm. The framework for enhancing the input image is shown in Fig 14, where the wiener filter is used to filter out noise from the corrupted image.

References

1. P. G. subin and P. Muthukannan, "Optimized convolution neural network based multiple eye disease detection," *Computers in Biology and Medicine,* vol. 146, p. 105648, 18 July 2022.
2. H. Fang, F. Li, H. Fu, X. Sun, X. Cao, F. Lin, J. Son, S. Kim, G. Quellec, S. Matta, S. M. Shankaranarayana, Y.-T. Chen, C.-H. Wang, N. A. Shah, C.-Y. Lee, C.-C. Hsu, H. Xie, B. Lei, U. Baid, S. Innani, K. Dang, W. Shi, R. Kamble, N. Singhal, C.-W. Wang, S.-C. Lo, J. I. Orlando, H. Bogunovic, X. Zhang and Y. Xu, "ADAM Challenge: Detecting Age-Related Macular Degeneration From Fundus Images," *IEEE Transactions on Medical Imaging,* vol. 41, p. 2828–2847, October 2022.
3. X. Li, X. Hu, X. Qi, L. Yu, W. Zhao, P.-A. Heng and L. Xing, "Rotation-Oriented Collaborative Self-Supervised Learning for Retinal Disease Diagnosis," *IEEE Transactions on Medical Imaging,* vol. 40, pp. 2284-2294, 2021.
4. N. Gour and P. Khanna, "Multi-class multi-label ophthalmological disease detection using transfer learning based convolutional neural network," *Biomedical Signal Processing and Control,* vol. 66, p. 102329, April 2021.
5. M. J. Alwazzan, M. A. Ismael and A. N. Ahmed, "A Hybrid Algorithm to Enhance Colour Retinal Fundus Images Using a Wiener Filter and CLAHE," *Journal of Digital Imaging,* April 2021.
6. T. Li, W. Bo, C. Hu, H. Kang, H. Liu, K. Wang and H. Fu, "Applications of deep learning in fundus images: A review," *Medical Image Analysis,* vol. 69, p. 101971, 2021.
7. Z. Shen, H. Fu, J. Shen and L. Shao, "Modeling and Enhancing Low-Quality Retinal Fundus Images," *IEEE Transactions on Medical Imaging,* vol. 40, p. 996–1006, March 2021.
8. W. Wang, X. Li, Z. Xu, W. Yu, J. Zhao, D. Ding and Y. Chen, "Learning Two-Stream CNN for Multi-Modal Age-Related Macular Degeneration Categorization," *IEEE Journal of Biomedical and Health Informatics,* vol. 26, p. 4111–4122, August 2022.
9. S. Sengupta, A. Singh, H. A. Leopold, T. Gulati and V. Lakshminarayanan, "Ophthalmic diagnosis using deep learning with fundus images – A critical review," *Artificial Intelligence in Medicine,* vol. 102, p. 101758, 2020.
10. K. Li, X. Qi, Y. Luo, Z. Yao, X. Zhou and M. Sun, "Accurate Retinal Vessel Segmentation in Color Fundus Images via Fully Attention-Based Networks," *IEEE Journal of Biomedical and Health Informatics,* vol. 25, p. 2071–2081,

June 2021.

11. B. J. Bhatkalkar, S. V. Nayak, S. V. Shenoy and R. V. Arjunan, "FundusPosNet: A Deep Learning Driven Heatmap Regression Model for the Joint Localization of Optic Disc and Fovea Centers in Color Fundus Images," *IEEE Access,* vol. 9, p. 159071–159080, 2021.
12. K. Aurangzeb, R. S. Alharthi, S. I. Haider and M. Alhussein, "An Efficient and Light Weight Deep Learning Model for Accurate Retinal Vessels Segmentation," *IEEE Access,* vol. 11, pp. 23107-23118, 2023.

Deep Learning Approach for Performance Evaluation of Pre-trained Network in Predicting Skin Disease

Shaikh Adiba[a*], Quadri Syed Adil[b], Mazhar Kazi[c], Syed Zebanaaz[d]

[a*]*Dr. G.Y. Pathrikar College of CS and IT, MGM Unversity, Chhatrapati Sambhajinagar(Auranagabad),431003, India.*
[b]*Kamla Nehru Polytechnic College of Pharmacy, Dr. Rafiq Zakaria Campus, Auranagabad, 431001, India*
[c]*Shikshan Maharshi Dnyandeo Mohekar Mahavidyalaya, Kalamb, Osmanabad, India*
[d]*Dr. Rafiq Zakaria Centre for High Learning & Adv. Research, Dr. Rafiq Zakaria Campus, Auranagabad, 431001, India*

Abstract

The Prediction of skin diseases becoming challenging task even for dermatologist working from many years because of too tiny contrast between surrounding skin and lesions, the usual resemblance between skin lesions. In order to diagnose and treat skin diseases, doctors currently use a biopsy procedure that is examined and carried out by them. A hybrid approach can avoid human judgment, producing positive results quickly. A thorough examination suggests that frameworks for recognizing various skin disorders may be built using deep learning techniques. To find skin illnesses, it is necessary to distinguish between skin and non-skin tissue. Computer-aided vision system helps eliminate to prognosis malignant skin lesions at the earliest time. With the advent in the deep learning includes CNN to have improved accuracy in prediction. To evaluate the performance and implement CNN for identifying skin disease, we have chosen the transfer learning with the popular architectures like Resnet50, Resnet101, AlexNet and Squeezenet.

Keywords: Deep Learning, Skin Disease, Performance, F1 Score, Accuracy, Precision, SVM, classifier, Feature extraction, Recall, Pre-trained network, Resnet-50, Resnet-101, Alexnet, Squeezenet.

1. Introduction

One of the most prevalent illnesses in the world is skin disease. Skin conditions come in many different forms, including squamous cell carcinoma (SCC), intraepithelial carcinoma, melanoma, and basal cell carcinoma (BCC) [1]. Research has shown that skin cancer, in particular, is the most prevalent cancer in the United States and that one-fifth of all Americans will develop skin cancer at some point in their lifetime [2, 3]. With a death rate of 1.62% among all skin cancers, melanoma is said to be the most lethal type [4].

The American Cancer Society predicts that there will be roughly 100,350 new instances of melanoma in the country in 2020 and that 6,850 individuals will pass away from the disease [5]. Contrarily, BCC is the most prevalent form of skin cancer, and although while it seldom results in death, it has a significant negative impact on healthcare services [6]. Thankfully, early skin cancer detection and treatment can increase five-year survival rates by about 14% [7].

However, correctly identifying a skin illness is difficult since it requires a number of visual cues, including the individual lesion morphology, the body site distribution, color, scaling, and arrangement of lesions. The diagnosing procedure can be difficult to understand when the different components are examined independently [8].

For melanoma, the ABCD rules, pattern analysis, Menzies method, and 7-Point Checklist are the four main clinical diagnosis techniques. Only doctors with plenty of experience can frequently use these techniques to make accurate diagnoses [9].

The gold standard for determining the existence of a skin illness is the histo-pathological analysis of a biopsy taken from a suspicious lesion. Numerous illustrations of various skin conditions are shown in Fig. 1. So, it would be advantageous to develop a technology that can automatically distinguish between skin cancer and non-cancer and differentiate between different types of skin cancer.

Fig. 1. Numerous illustrations of various skin conditions. The Dermofit Picture Library is where these pictures were found [10].

In this world of development, knowledge, progress, the advancements in research have reached such a high stage, where it won't be wrong to say that this is the "Computer World". Wherein, almost all processes are being taken up by the computer programs. So is the field of Artificial Intelligence being raised to an extent where diagnosing the disease at the initial stage won't require visiting a doctor. We can get to know about the disease by just placing their symptoms in the Expert System.
In this article, we tried to evaluate the performance of pre-trained networks using deep learning in predicting skin disease. Deep Learning sounds quite similar to Machine Learning. No doubt, both contribute to Artificial Intelligence as a whole, deep learning and machine learning do have differences.

2. Skin Disease

The human body's largest and thickest organ, the skin is made up of the epidermis, dermis, and hypodermis. The three primary functions of the skin are protection, sensation, and thermoregulation, and they all work together to offer a strong defense against environmental assault. The epidermis' top layer, the stratum corneum, is a layer of protective tissue that is optically neutral and varies in thickness. Keratinocytes in the stratum corneum create keratin, which is beneficial to the skin's ability to shield the body. The stratum corneum scatters light when it strikes the skin. Melanocytes are present in the epidermis' basal layer. In particular, melanocytes cause the skin to produce melanin, a pigment that gives skin its tan or brown hue. By producing more melanin, melanocytes serve as filters and shield the skin from damaging ultraviolet (UV) rays from the sun. The number of melanocytes determines how much UV radiation is absorbed. Melanoma, however, is brought on by melanocytes that develop abnormally. The middle layer of skin, known as the dermis, is made up of collagen fibers, sensors, receptors, blood vessels, and nerve endings. It gives the skin flexibility and vitality [32]. Nucleotide molecules make up Deoxyribonucleic Acid (DNA), which is made up of them. A phosphate, sugar, and nitrogen base are all components of a nucleotide. The genes are created from the sequence of nitrogen bases in the DNA. Cell development, division, multiplication, and death are all controlled by genes. Cell division and proliferation are regulated by oncogenes. Tumor suppressor genes are often referred to as protective genes. They typically do this by keeping track of how quickly cells divide into immature cells, repairing damaged DNA, and regulating when a cell dies. When the tumor suppressor genes mutate, a cell becomes uncontrollable and eventually forms a mass known as a tumor (cancer). The DNA can be harmed by UV radiation, which leads the melanocytes to manufacture melanin at an excessively high rate. A healthy quantity of UV radiation helps the skin produce vitamin D, but too much can lead to pigmented skin lesions [34]. Melanoma is the name given to the malignant tumor that develops when melanocytes proliferate abnormally [35]. Malignant melanoma (MM), squamous cell carcinoma, and basal cell carcinoma are the three main subtypes of skin cancer. The latter two, termed as keratinocyte carcinoma, in particular, are derived from basal and squamous keratinocytes (KC). In the United States, about 4.3 million cases of BCC and 1 million cases of SCC are identified each year, though it's probable that these figures are understated [36]. These are the most frequent skin malignancies that affect both men and women. Nevertheless, MM, an aggressive melanocyte malignancy is a less prevalent but significantly more lethal form of skin cancer. It frequently begins tiny and gradually grows and changes in color. Melanin's hue primarily depends on where it is located in the skin. Melanin found in the stratum corneum is what gives skin the color ebony. The top epidermis, papillary dermis, and reticular dermis are each shown to be light to dark brown, gray to gray-blue, and steel-blue, respectively. The excessive melanin deposit is present in the epidermis in cases of benign lesions. The most significant indication of melanoma that causes noticeable vicissitude in skin color is melanin present in the dermis. In addition to pale lesion regions with a significant blood supply at the periphery, thicker collagen fibers are one more indicator of melanoma. The pigmented lesion's shape, size, color, border, and symmetry are further gross morphologic characteristics. If the ocular approximation confirms a suspicion of skin cancer, a biopsy and histology are necessary to provide an exact diagnosis [37]. There are four main kinds of melanoma based on microscopic evaluations of the lesion: superficial spreading melanoma (SSM), nodular melanoma (NM), lentigo malignant melanoma (LMM), and acral lentiginous melanoma (ALM).

3. Skin disease diagnosis with deep learning

As deep learning has grown in popularity, it has been applied to a variety of dermatological problems. Here, we take a look at the research that has already been done using deep learning to diagnose skin diseases. From a machine learning viewpoint, we first offer a general overview of the data preparation and augmentation techniques used in deep learning, and then we present a literature evaluation of deep learning's applications in skin disease detection organized by task type.

3.1. Data preprocessing and augmentation

3.1.1. Data preprocessing

When using deep learning to diagnose skin diseases, data preparation is crucial. Since skin disease datasets (like ISIC, PH2, and AtlasDerm) contain images with widely varying resolutions, and deep networks typically accept inputs with certain square sizes (like 224 by 224 and 512 by 512), it is necessary to crop or resize these images before feeding them into deep learning networks. Images may be distorted or lose important details if they are resized or cropped directly to meet specific dimensions [13, 32]. Practical solutions to this problem include scaling images uniformly along their shortest side while preserving their aspect ratio. Before being fed into a deep learning network, images are usually normalized by subtracting the mean value and dividing by the standard deviation, both of which are calculated over the whole training subset. Since the lighting, skin tones, and perspectives of skin disease photos can vary widely across a dataset, it has been observed in previous publications [33, 32] that subtracting a uniform mean value does not properly normalize the illumination of individual photographs. Yu et al. [32] corrected this by removing the image's channel-wise mean intensity values from each individual image to normalize them. The experimental results presented in their research demonstrated that the performance of a deep network degrades when the mean pixel value is subtracted. Moreover, hair or other unrelated stuffs should be deleted from skin pictures using techniques such thresholding methods [34, 35], morphological approaches [36], and deep learning algorithms [22, 21, 22] for more accurate segmentation and classification.

3.1.2. Data augmentation

It is well-known that training a deep learning network requires a vast amount of data in order to prevent overfitting and produce optimal results. Unfortunately, large amounts of labeled training data are difficult to come by for many applications, such as skin disease diagnosis. Due to factors such as disease rarity, patient privacy, the need for labeling by medical specialists, and the high expense to gather medical data, medical image analysis typically works with limited data [37]. To address this challenge, researchers have devised techniques for "data augmentation," which involve artificially modifying the original data using the right techniques to enhance the quantity of accessible training data. It is possible to increase the quantity and quality of the available training data through the use of feasible data augmentation techniques. Deep learning systems can pick up more useful features like rotation and translation invariance with more training data. Geometric transformations (such as flipping, cropping, translating, and rotating), color space augmentations, kernel filters, mixing images, random erasing, feature space augmentations, adversarial training, generative adversarial networks, neural style transfer, and meta-learning are all common data augmentation techniques [37]. The 4,000 dermoscopy images used by Al-Masni et al. [38] were rotated by 0, 90, 180, and 270 degrees to provide additional training data. By doing so, we were able to lessen the effects of over-fitting and increase the deep networks' stability. The photos were rotated by 0, 90, and 180 degrees by Yu et al. [32], and then randomly translated (with a shift between -10 and -10) pixels. Experiments conducted on the ISIC skin dataset showed a considerable improvement after data augmentation was used. For a more in-depth discussion of data augmentation, the reader is directed to the work of Shorten et al. [37].

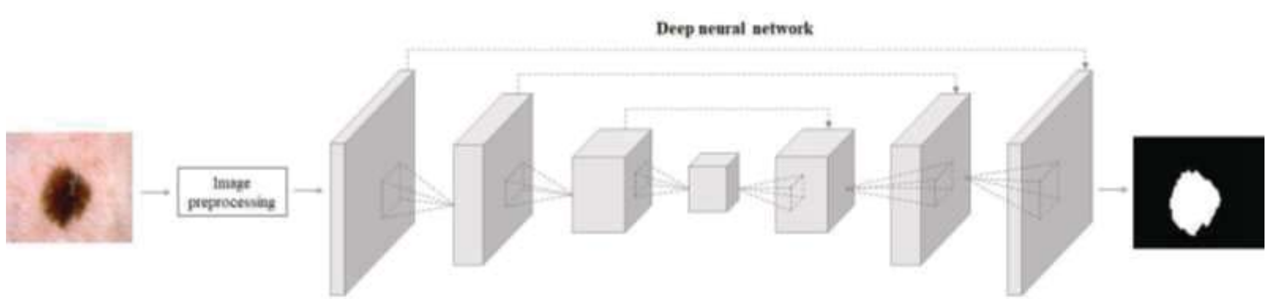


Fig. 2. The workflow of a typical skin disease segmentation task

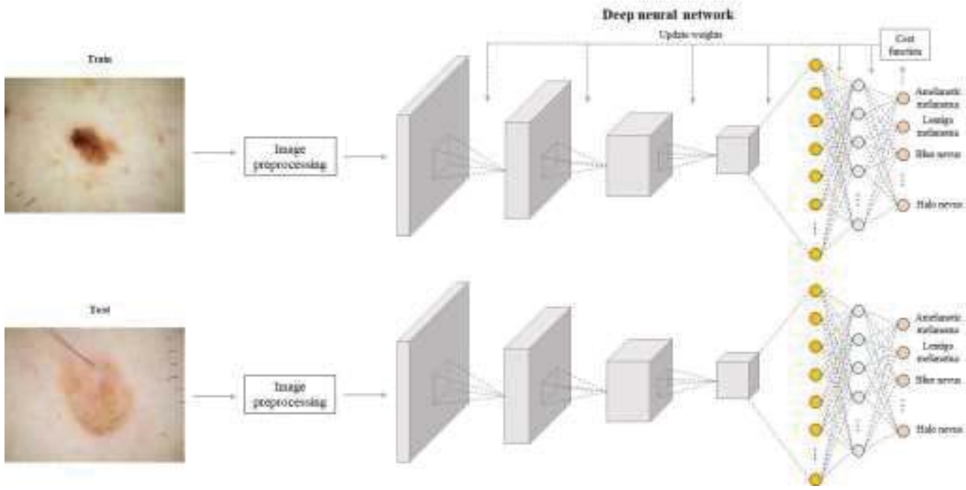


Fig. 3. The workflow for a typical skin disease classification task

4. Methodology:

4.1. Dataset:

The DermNet Skin Disease Atlas collected and labeled over 23,000 photos and made them accessible for free use as the DermNet dataset. The URLs for the remaining photographs looked to be broken, so limited images were downloadable. Nevertheless, the data had a few redundant sub-classes that were either blank or otherwise unimportant. From http://www.dermnet.com/dermatology-pictures-skin-disease-pictures, we gathered data consisting of images of 08 different skin illnesses. The table below shows type of skin illnesses and the number of samples of each type collected for our experiment.

Table 1. Type of Skin Illnesses and Data Collected

Skin Illness Type	Number of Samples Gathered
Acne	355
Actinic Keratosis	214
Basal Cell Carcinoma	444
Hives	121
Nail Infection	148
Psorasis	691
Rhus Dermatitis	115
Rosacea	196

There are altogether 2,284 samples in total, with a proportion of 80 % used for training and the rest 20% images used for testing. Images are in JPEG format, which uses three channels (RGB) to represent color. While there is some variation in image and category resolutions, none of this is really high-resolution data.
The block diagram in figure 4 shows the step by step process of predicting skin disease with Deep learning models

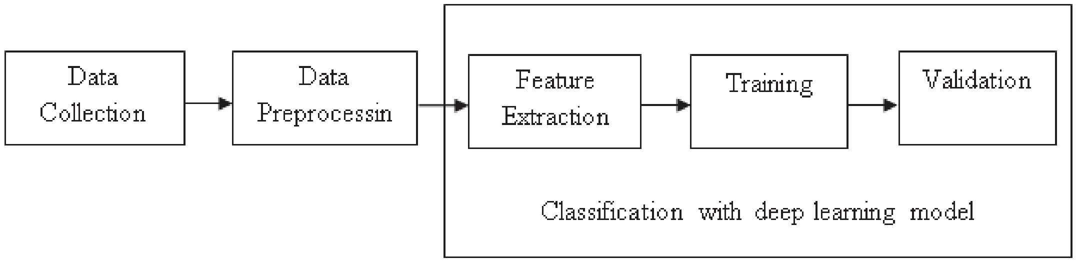


Fig. 4. Proposed model using deep learning algorithm

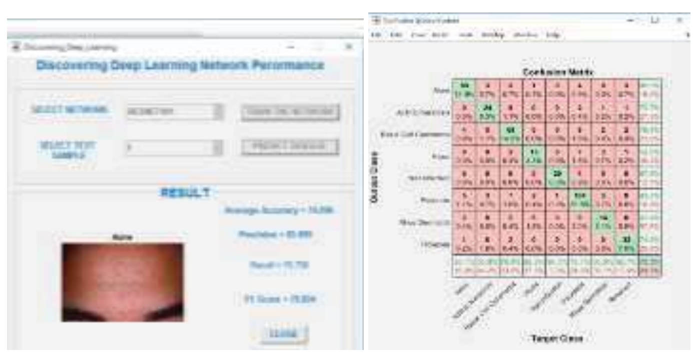
5. Results

The Pre-trained learning models namely ResNet50, ResNet101, AlexNet, SqueezeNet are considered here in this experiments. In order to evaluate the performance of Pre-trained model with deep learning approach, we calculated four different metrics viz accuracy, precision, recall, F1 score.

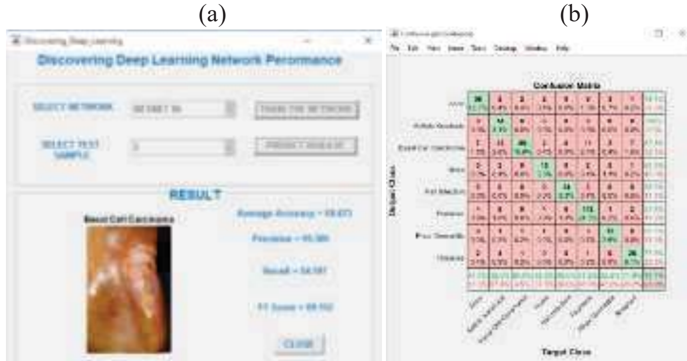
The most popular option for resolving challenging computer vision issues has been deep learning methods. Model correctness is the most common evaluation statistic used to determine an algorithm's performance. For a very long time, the primary parameter utilized to compare machine learning models was accuracy.

- How often a model correctly predicted throughout the entire dataset is determined by the accuracy statistic. This measure can only be trusted if the dataset is class-balanced, meaning that each class contains an equal amount of samples.
- How many of the model's "positive" predictions came true is a measure of precision.
- Recall: Recall measures how many of the positive class samples present in the dataset were correctly identified by the model.
- Accuracy, on the other hand, merely measures the proportion of times a model correctly predicted across the full dataset, which is still valid if the dataset is class-balanced.
- An alternative machine learning evaluation statistic called F1 score evaluates a model's predictive ability by focusing on its performance inside each class rather than its overall performance as is done by accuracy. A model's precision and recall scores are combined into one metric, the F1 score, which has led to its extensive use in recent research.

We used precision and recall as a criterion to assess the performance of these pre-trained models because our dataset is very unbalanced and of the multiclass type. Figure 5 (a) through (j) shows the results from the proposed system.



(a) (b)



(c) (d)

(e) (f)

(g) (h)

(i) (j)

Fig. 5. Screenshots and Confusion Matrix of the proposed system implemented in MATLAB

6. Conclusion

In the near future, initial disease diagnosis won't require a trip to the doctor thanks to advancements in artificial intelligence. By just entering the disease's symptoms into the Expert System, we can learn more about the condition. Deep learning has becoming more popular since it requires less data samples while still producing amazing results. This article illustrates how well deep learning-based pre-trained networks do at predicting skin diseases.

Thus, we may draw the conclusion that ResNet50 has demonstrated the most promising results in identifying 8 distinct skin diseases, needing less training samples and no transfer learning. After transfer learning, AlexNet and SqueezeNet may become more effective in our instance.

References

1. Srinivasu, P.N.; SivaSai, J.G.; Ijaz, M.F.; Bhoi, A.K.; Kim, W.; Kang, J.J. 2021. Classification of Skin Disease Using Deep Learning Neural Networks with MobileNet V2 and LSTM. *Sensors 21*, 2852. https://doi.org/10.3390/s21082852

2. Mostafiz Ahammed, Md. Al Mamun, Mohammad Shorif Uddin. 2022. A machine learning approach for skin disease detection and classification using image segmentation, Healthcare Analytics, Volume 2, 100122, ISSN 2772-4425, https://doi.org/10.1016/j.health.2022.100122.
3. Rifat Sadik, Anup Majumder, Al Amin Biswas, Bulbul Ahammad, Md. Mahfujur Rahman. 2023. An in-depth analysis of Convolutional Neural Network architectures with transfer learning for skin disease diagnosis, Healthcare Analytics, Volume 3, 100143, ISSN 2772-4425, https://doi.org/10.1016/j.health.2023.100143.
4. Wu H, Yin H, Chen H, Sun M, Liu X, Yu Y, Tang Y, Long H, Zhang B, Zhang J, Zhou Y, Li Y, Zhang G, Zhang P, Zhan Y, Liao J, Luo S, Xiao R, Su Y, Zhao J, Wang F, Zhang J, Zhang W, Zhang J, Lu Q. 2020. A deep learning, image based approach for automated diagnosis for inflammatory skin diseases. Ann Transl Med. 8(9):581. doi: 10.21037/atm.2020.04.39. PMID: 32566608; PMCID: PMC7290553.
5. V. Balaji, S. Suganthi, R. Rajadevi, V.K. Kumar, B.S. Balaji, S. Pandiyan. 2020. Skin disease detection and segmentation using dynamic graph cut algorithm and classification through Naive Bayes classifier, Measurement 107922.
6. American Cancer Society. 2020. Cancer facts & figures for hispanics/latinos 2018–2020.
7. R. Kasmi, K. Mokrani. 2016. Classification of malignant melanoma and benign skin lesions: implementation of automatic ABCD rule, IET Image Process. 10 (6) 448–455, http://dx.doi.org/10.1049/iet-ipr.2015.0385.
8. A.D. Mengistu, D.M. Alemayehu. 2015. Computer vision for skin cancer diagnosis and recognition using RBF and SOM, Int. J. Image Process. (IJIP) 9 (6) 311–319.
9. M. Pawar, D.K. Sharma. 2014. R. Giri, Multiclass skin disease classification using neural network, Int. J. Comput. Sci. Inform. Technol. Res. 2 (4) 189–193.
10. L.-s. Wei, Q. Gan, T. Ji, Skin disease recognition method based on image color and texture features, Comput. Math. Methods Med. 2018 (2018).
11. M.N. Islam, J. Gallardo-Alvarado, M. Abu, N.A. Salman, S.P. Rengan, S. Said. 2017. Skin disease recognition using texture analysis, in: 2017 IEEE 8th Control and System Graduate Research Colloquium, ICSGRC, pp. 144–148, http://dx.doi.org/10.1109/ICSGRC.2017.8070584.
12. A. Nawar, N.K. Sabuz, S.M.T. Siddiquee, M. Rabbani, A.A. Biswas, A. Majumder. 2021. Skin disease recognition: A machine vision based approach, in: 2021 7th International Conference on Advanced Computing and Communication Systems, vol. 1, ICACCS, pp. 1029–1034, http://dx.doi.org/10.1109/ICACCS51430.2021.9441980.
13. F. Curia. 2021. Features and explainable methods for cytokines analysis of Dry Eye Disease in HIV infected patients, Healthc. Anal. 1 100001.
14. V. Chang, V.R. Bhavani, A.Q. Xu, M. Hossain. 2022. An artificial intelligence model for heart disease detection using machine learning algorithms, Healthc. Anal. 2 100016.
15. S. Dev, H. Wang, C.S. Nwosu, N. Jain, B. Veeravalli, D. John. 2022. A predictive analytics approach for stroke prediction using machine learning and neural networks, Healthc. Anal. 2 100032, http://dx.doi.org/10.1016/j.health.2022.100032, URL https://www.sciencedirect.com/science/article/pii/S2772442522000090.
16. R. AlSaad, Q. Malluhi, I. Janahi, S. Boughorbel. 2022. Predicting emergency department utilization among children with asthma using deep learning models, Healthc. Anal. 2 100050, http://dx.doi.org/10.1016/j.health.2022.100050, URL https://www.sciencedirect.com/science/article/pii/S2772442522000181.
17. M. Ahammed, M.A. Mamun, M.S. Uddin. 2022. A machine learning approach for skin disease detection and classification using image segmentation, Healthc. Anal. 2 100122, http://dx.doi.org/10.1016/j.health.2022.100122, URL https://www.sciencedirect.com/science/article/pii/S2772442522000624.
18. S. Serte, A. Serener, F. Al-Turjman. 2020. Deep learning in medical imaging: A brief review, Trans. Emerg. Telecommun. Technol. e4080.
19. N.C. Thompson, K. Greenewald, K. Lee, G.F. Manso. 2020. The computational limits of deep learning. arXiv preprint arXiv:2007.05558.
20. H. Pan, Z. Pang, Y. Wang, Y. Wang, L. Chen. 2020. A new image recognition and classification method combining transfer learning algorithm and MobileNet model for welding defects, IEEE Access 8. 119951–119960.
21. W. Wang, Y. Li, T. Zou, X. Wang, J. You, Y. Luo. 2020. A novel image classification approach via dense-MobileNet models, Mob. Inf. Syst. 2020.
22. K. Sriporn, C.-F. Tsai, C.-E. Tsai, P. Wang. 2020. Analyzing lung disease using highly effective deep learning techniques, Healthcare 8 (2) 107.

23. T. Ghosh, M.M.-H.-Z. Abedin, S.M. Chowdhury, Z. Tasnim, T. Karim, S.S. Reza, S. Saika, M.A. Yousuf. 2020. Bangla handwritten character recognition using MobileNet V1 architecture, Bullet. Electr. Eng. Inform. 9 (6) 2547–2554.
24. T.M. Angona, A. Siamuzzaman Shaon, K.T.R. Niloy, T. Karim, Z. Tasnim, S. Reza, T.N. Mahbub. 2020. Automated bangla sign language translation system for alphabets by means of MobileNet, Telkomnika 18 (3).
25. M. Rahimzadeh, A. Attar. 2020. A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of xception and ResNet50V2, Inform. Med. Unlocked. 100360.
26. E. Ayan, H.M. Ünver. 2019. Diagnosis of pneumonia from chest X-Ray images using deep learning, in: 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science, EBBT, IEEE. pp. 1–5.
27. L. Yang, P. Yang, R. Ni, Y. Zhao. 2020. Xception-based general forensic method on small-size images, in: Advances in Intelligent Information Hiding and Multimedia Signal Processing, Springer. pp. 361–369.
28. C. Shi, R. Xia, L. Wang. 2020. A novel multi-branch channel expansion network for garbage image classification, IEEE Access 8. 154436–154452.
29. A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, H. Adam. 2017. MobileNets: Efficient convolutional neural networks for mobile vision applications. arXiv:1704.04861.
30. F. Chollet. 2017. Xception: Deep learning with depthwise separable convolutions, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 1251–1258.
31. D.S. Reddy, P. Rajalakshmi. 2019. A novel web application framework for ubiquitous classification of fatty liver using ultrasound images, in: 2019 IEEE 5th World Forum on Internet of Things (WF-IoT), IEEE. pp. 502–506.
32. R. Yasir, M.A. Rahman, N. Ahmed. 2014. Dermatological disease detection using image processing and artificial neural network, in: 8th International Conference on Electrical and Computer Engineering, IEEE. pp. 687–690.
33. J.S. Alarifi, M. Goyal, A.K. Davison, D. Dancey, R. Khan, M.H. Yap. 2017. Facial skin classification using convolutional neural networks, in: International Conference Image Analysis and Recognition, vol. 10317, Springer, Cham. pp. 479–485, http://dx.doi.org/10.1007/978-3-319-59876-5_53.
34. Y. Li, L. Shen. 2018. Skin lesion analysis towards melanoma detection using deep learning network, Sensors 18 (2) 556.
35. J. Rathod, V. Wazhmode, A. Sodha, P. Bhavathankar. 2018. Diagnosis of skin diseases using convolutional neural networks, in: 2018 Second International Conference on Electronics, Communication and Aerospace Technology, ICECA, IEEE, pp. 1048–1051.
36. M. Chen, P. Zhou, D. Wu, L. Hu, M.M. Hassan, A. Alamri. 2019. AI-skin: Skin disease recognition based on self-learning and wide data collection through a closed-loop framework, Inf. Fusion 54 1–9.
37. M.A.A. Milton. 2019. Automated skin lesion classification using ensemble of deep neural networks in isic 2018: Skin lesion analysis towards melanoma detection challenge. arXiv preprint arXiv:1901.10802.
38. H. Liao, 2015. A deep learning approach to universal skin disease classification.
39. T. Shanthi, R. Sabeenian, R. Anand. 2020. Automatic diagnosis of skin diseases using convolution neural network, Microprocess. Microsyst. 103074.

Acknowledgement

Shaikh Adiba is working as an Assistant Professor in the field of computer science and IT at Dr. G.Y. Pathrikar College of CS and IT. Her area of interest is Pattern Recognition. She came up with the idea of the paper, wrote the code in matlab and captured the experimental results.

Quadri Syed Adil is working as Assistant Professor in the field of Pharmacy at Kamla Nehru Polytechnic College. He has an experience of teaching in medicine and has a thorough knowledge of various types of skin diseases. He made his contribution in understanding skin diseases and forming dataset for experimental work.

Mazhar Kazi is working as Assistant Professor in the field of computer science at Shikshan Maharshi Dnyandeo Mohekar Mahavidyalaya, Kalamb. His area of interest is Neural Network. He contributed on literature review and framed the experiment into an article.

Syed Zebanaaz is working as an Assistant Professor in computer science at Dr. Rafiq Zakaria Centre of Higher Learning & Advanced Research. She contributed in article language correction through proofreading.

Container Based Virtualization Architecture for Performance Evaluation

Jayasudha Mudaliar[a*], Dr. Nimisha Patel[b]

[a*,b]*Gandhinagar Institute of Technology, Gandhinagar, Gujarat 382721, India*

Abstract

The provision of resources to users with the service and deployment models is supported by the usage of fundamental technologies like virtualization and containerization. The current virtualization landscape focuses on creating an abstraction layer over the computer hardware that enables the hardware components of a single computer to be divided into multiple virtual computers, also known as the virtual machines. Similar to that, containerization offers a solid foundation on which one or more containers can run under the same common operating system, despite the fact that both techniques are undergoing rapid improvement. Here, containerization lacks OS stability and platform compatibility while virtualization lacks in performance because of this, our aim is to make a combined form of virtualization and containerization, and to create an architecture that allows us to install containers on top of virtual machines. On the basis of this design, performance evaluation work is done. The deployments are carried out in a predetermined order using calculations based on the dynamic usage value, and the study considers the ideas involved in the creation of an improved model.

Keywords: Virtualization, Containerization, Mapping, Scaling, Computing.

1. Introduction

The concept for the development of hardware and software requires the use of new technologies, integrated tools, and distributed computer infrastructures. The usage of fundamental technologies like virtualization and containerization used for the service and deployment models provide resources to users. Virtualization in the simplest and easiest words is the abstraction of computer resources. Virtualization has the ability to allow single physical resource serve as a multiple virtual resource [1]. Virtualization concept focuses over the computer hardware that allows the hardware elements of a single computer to be divided into multiple virtual computers (virtual machines). A container is a standardized software component that wraps up code and all of its dependencies to ensure that an application will run efficiently and consistently in different computing environments. Containers provide a new way of virtualization which is lightweight and agile. This has helped in the popularity of containers. Adapting to containers for new applications or migrating to container architecture for existing applications have lots of benefits for the application life cycle management [2].

Based with these technologies implementation of both the concept is done. Here first virtualization process takes place further which the containerization process is included. The same process is carried out with the scenario of the validation, evaluation and solution. The implementation of the stated concept is carried out with the phases like clustering, selector and labelling, scaling and computation of scaling. To prepare an optimized system the work is implemented with the computation of the scaling where the scaling is computed with dynamic value. Hence, the system utilization is measured evaluated and then validation is done. However, basis on this parameter based evaluation is performed.

2. Theory Background

The most common use for a conceptual model of a virtual infrastructure is in cloud computing, which uses virtualization design. The act of developing and delivering a virtual version of something as opposed to a physical one is known as virtualization. A desktop, an operating system (OS), a server, a storage device, or network resources could all fit this description. The main purpose of VMs is to operate multiple operating systems at the same time, from the same piece of hardware. Without virtualization, operating multiple systems would require two separate physical units.

[*]Jayasudha Mudaliar

E-mail address: jayasudha.mudaliar@gmail.com

The main feature of virtualization is achieved through an isolation layer, built on top of the operating system level [5]. Containerization is the packaging together of software code with all it's necessary components like libraries, frameworks, and other dependencies so that they are isolated in their own container. Container use all available resources, and they minimize overhead.

- **Need of Virtualization:**

Virtualization enables the much more effective use of all underutilized resources. When compared to traditional systems, virtualization offers a more fault-tolerant environment because there are always other servers accessible to use in case a server fails for any reason. Due to the fact that virtualization may be done utilizing a variety of virtual machines, it also decreases the numerous server needs. As a result, it takes a lot of energy to run those systems and maintain the system. The need can be reduced and resource use can become effective with the use of virtual machines. In addition to all of that, this is good for the environment.

- **Need of Containerization:**

The creation of containers has solved a number of computing related problems, including resource management, scheduling, and packaging. The needs of application administration and packaging is also resolved. Containers can provide an effective solution for the better deployments. A container is a bundle that includes middleware, business logic, and application elements that are prepared for deployment. Containers are extremely safe to use and scalable. These are simpler to deploy. Container as an open-source platform that enables users and programmers to segregate application dependencies.

3. Proposed Work

The work focused in the terms of tools, models, methods and processes, in relation to virtualization, containerization and composition concept. The systematic mapping of the area like validation, evaluation and solution, with the base concepts of virtualization and containerization. The work is based on the composition of the concepts of the virtualization and containerization. The layers and modules in the architecture of the composition of process virtualization and containerization, which states the container based virtualization architecture, are shown in the diagram in Figure 3.0.1.

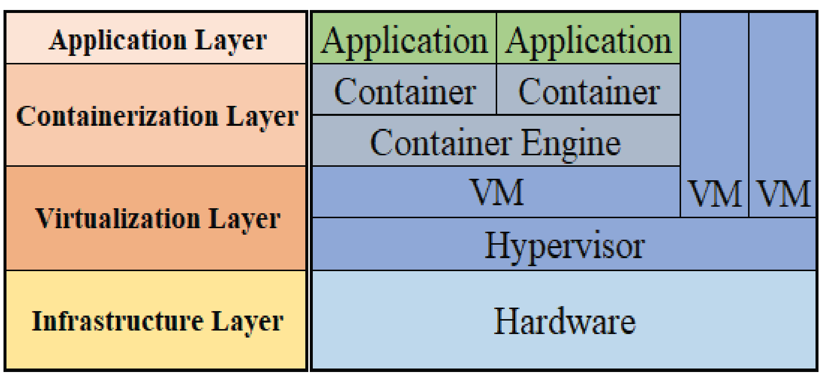


Fig. 3.0.1 Container Based Virtualization Architecture

The figure 3.0.1 illustrates the container-based virtualization approach, in which the virtual machine is constructed over the hardware by deploying the virtualization layer over the infrastructure layer. Following this is the deployment of the containerization layer over the virtualization layer, during which containers are deployed on top of virtual machines. The final phase involves performing application configurations with containers while using the application layer over the containerization layer.

3.1 Mapping of Containers

The structured mapping of the module is the main focus of the effort. Clustering the virtual machines together after creating the many virtual machine instances in order to deploy the containers. Grouping and scaling are carried out in order to get the optimum model ready. Figure 3.1.1 demonstrates that the model is built on an effective container mapping that is distributed across the instance.

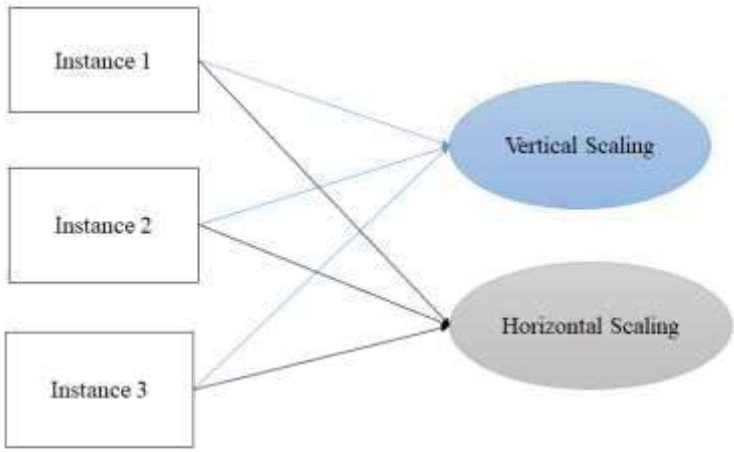


Fig. 3.1.1 Node Group and Scaling Model

The figure 3.1.1 shows that the model is based on a well-organized mapping of the containers, with containers efficiently dispersed over the instance. The types of scaling where both vertical and horizontal scaling are present are shown in figure 3.1.2.

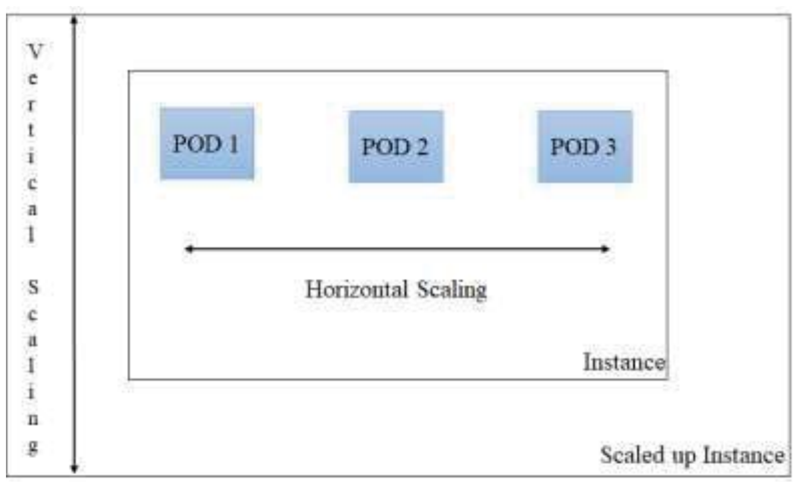


Fig. 3.1.2 Types of Scaling

The figure 3.1.2 states the types of scaling where both the vertical and horizontal scaling are included for the achieving the performance and model efficiency. Here the vertical scaling is done for the CPU utilization and horizontal scaling is done for the storage utilization.

3.2 Algorithm

Step 1: Setup Hardware and create VM Instance
Step 2: To deploy the pods create the cluster between the nods

```
Step 3: To deploy the pod check the nod group
if the configuration = Group1
    then deploy to the Instance 1
else if the configuration = Group2
    then deploy to the Instance 2
else if the configuration = Group3
    then deploy to the Instance 3
Step 4: After the deployment check the vertical scaling and horizontal scaling
```

Vertical scaling

```
if the CPU_utilization >= Utilization_threshold (F)
    then CPU_utlization increased
    do
        reset the pods and allocate new pods
    done
    pod = pod+n(pod)

else if CPU_utilization <= Utilization_threshold (F)
    then CPU_utilization decreased
    do
        reset the pods and deallocate the pod
    done
    pod = pod-n(pod)
```

Horizontal scaling

```
if the Memory_utilization >= Memory_threshold (M)
    then memory overloaded
    do
        increase the memory of the nods
    done
    m_node= node_memory + new_memory

else if Memory_utilization <= Memory_threshold (M)
    then memory underloaded
    do
        decrease the memory of the nods
    done
    m_node = node_memory - new_memory
```

Step 5: Perform the scaling i.e., Step 4 with the specified time interval and perform the pod downscale as scheduled time period.

Here the algorithm states the deployment. At first the setup of the hardware and create the virtual machine instances is performed. Then the clustering between the nods and the deployment of the pods are performed with this the node group is also performed to deploy the pods in an efficient manner.

3.3 Threshold Calculation

The threshold calculation is used as the foundation for the scaling. The threshold values are also used as the foundation for the scale up and scale down. Scale up is carried out when the threshold value is raised, and scale down is carried out when the threshold value is lowered. On the basis of the threshold values, scaling up and scaling down are conducted here. The probability value used to determine these threshold values is the average cumulative value.

On the basis of the iterations of the mean variables aligned with the system, cumulative value is estimated, equation 1 gives the resultant mean value; equation 2, the average cumulative value, is given by this equation.

$$A = \frac{\bar{a}}{n} \tag{1}$$

where,
n = number of mean variables;
a = value of variables;
A = mean value of variables
Equation 1. States the resultant mean value.

$$F = \sum_{A=0}^{N} \frac{A_N}{N} \tag{2}$$

where,
N = Number of mean variable;
F= Cumulative Variable

Equation 2. Depicts the Variable F which is the resultant value of the average cumulative value
The resultant F will be the obtained threshold.
The system's total iterations are used to determine the cumulative value. Our calculated cumulative value is based on the iterations. Both the memory and CPU use values are included in the iterations.

3.3.1 CPU utilization threshold

The CPU utilization threshold is calculated on the basis of the below given algorithm:

Utilization_threshold (F)
Step1: Get the average mean value of the cpu_utilization and store it to I.
Step 2: Similarly perform multiple iteration and store w.r.t In.
Step 3: Perform probability of cumulative value of i events and obtain F1
  Store the value of F1 in F.
  F will be the threshold value.
Step 4: Perform the step 1 to 3 and obtain the value of F2
Step 5: check
  if F2 > F
    then F = F2
  else
      F = F
Step 6: Stop
The Utilization_threshold is calculated as stated in the algorithm where the average mean value is stored in the I variable. Following to this we perform multiple iterations and store all those values with respect to the In. The F are obtained on the scheduled time which is the Utilization_threshold (F).

3.3.2 Memory utilization threshold

The Memory utilization threshold is calculated on the basis of the below given algorithm:

Memory_threshold (M)
Step1: Get the average mean value of the memory_utilization and store it to J.
Step 2: Similarly perform multiple iteration and store w.r.t Jn.
Step 3: Perform probability of cumulative value of j events and obtain M1
  Store the value of M1 in M.
  M will be the threshold.
Step 4: Perform the step 1 to 3 and obtain the value of M2

Step 5: check
    if M2 > M
        then M = M2
    else
        M = M
Step 6: Stop
The Memory_threshold is calculated as stated in the algorithm where the average mean value is stored in the J variable. The M are obtained on the scheduled time which is the Memory_threshold (M).

3.4 Implementation of Test Platform

The implementation for the topic "Container based Virtualization architecture" has been performed in a few stages. Below are the phases of the implementation:

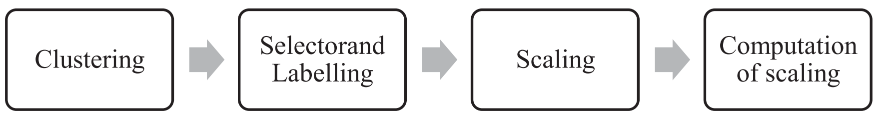


Fig. 3.4.1 Phases of the Implementation

The process that has been used throughout implementation is shown in figure 3.4.1, Phases of Implementation, where each phase is further detailed below.

3.4.1 Clustering Execution

In the cluster Execution the nodes and pods are stated where nodes refers to the virtual machines and pods refers to the containers. For the proposed model states the nodes and pods configured for the same.

3.4.2 Selector and Labelling

- Pods selectors on the basis of the node labelling.
- Node Labelling attributes:
    Hardware = Highspec & Lowspec
- Labelling the same with the nodes:
    kubectl label node worker-node-1 hardware=highspec
    kubectl label node worker-node-2 hardware=lowspec
- Pod Labelling attributes:
    Labels: dev, prod and type
    Node labelling w.r.t node attributes

3.4.3 Labelling Execution with Nodes

```
root@kubernetes-master:~/definations# kubectl get nodes --show-labels
NAME                STATUS   ROLES                  AGE     VERSION   LABELS
kubernetes-master   Ready    control-plane,master   5d23h   v1.23.4   beta.kuberne
tes.io/arch=amd64,beta.kubernetes.io/os=linux,kubernetes.io/arch=amd64,kubernetes.
io/hostname=kubernetes-master,kubernetes.io/os=linux,node-role.kubernetes.io/contr
ol-plane=,node-role.kubernetes.io/master=,node.kubernetes.io/exclude-from-external
-load-balancers=
kubernetes-node1    Ready    <none>                 5d23h   v1.23.4   beta.kuberne
tes.io/arch=amd64,beta.kubernetes.io/os=linux,hardware=highspec,kubernetes.io/arch
=amd64,kubernetes.io/hostname=kubernetes-node1,kubernetes.io/os=linux
kubernetes-node2    Ready    <none>                 5d23h   v1.23.4   beta.kuberne
tes.io/arch=amd64,beta.kubernetes.io/os=linux,hardware=lowspec,kubernetes.io/arch=
amd64,kubernetes.io/hostname=kubernetes-node2,kubernetes.io/os=linux
root@kubernetes-master:~/definations#
```

Fig. 3.4.4 Labelling the Nodes

The Figure 3.4.4 depicts about the node labelling, for the better management of the nodes, the labeling are performed where the worker nodes are labelled in which kubernetes node 1 is labelled as the **highspec** node and kubernetes node 2 is labelled as the **lowspec** node. With respect to this labelling the pods are deployed.

- **Labelling Execution with Pods using Node Selectors**

The Figure 3.4.5 states about the pod labelling where labelling of the pod is done with respect to the configurations.

```
root@kubernetes-master:~/definations# vim 1.pod1.yaml
root@kubernetes-master:~/definations# k apply -f 1.pod1.yaml
pod/nginx1 configured
root@kubernetes-master:~/definations# k get pods --show-labels
NAME     READY   STATUS    RESTARTS   AGE     LABELS
nginx    1/1     Running   0          2m12s   env=dev
nginx1   1/1     Running   0          45s     env=prod
nginx3   1/1     Running   0          3m52s   env=dev,type=frontend
root@kubernetes-master:~/definations#
```

Fig. 3.4.5 Labelling the Pods

The Figure 3.4.5 depicts about the pod labelling, the pods labelling is done with in the yaml file where the env is specified. The types of the labelling is used are dev, prod and type according to the case the engines are described these are nginx with label=dev, nginx1 with label=prod and nginx3 with label=dev and type=frontend. This labelling is done for the optimized management of the pods deployed over the nodes. The deployment of the pods over the nodes is specified where the deployment is done via the labels and selectors. Whereas,
nginx is deployed in the kubernetes-node1 with the node label=highspec and pod env = dev.
nginx1 is deployed in the kubernetes-node1 with the node label=highspec and pod env = prod.
nginx3 is deployed in the kubernetes-node2 with the node label=lowspec and pod env = dev with type=frontend.

3.4.4 Scaling

Scaling is done while taking into account the CPU and Memory Utilization. Using this scaling, the average cumulative frequency is used. Below are the parameters used for scaling.

Parameters:

Requests:

cpu: Fn
memory: Mn

- **Scaling Execution**

```
root@kubernetes-master:~/definations# k get pods --show-labels
NAME              READY   STATUS    RESTARTS      AGE   LABELS
myapp-rc-j4tgq    1/1     Running   0             83s   app=myweb,type=frontend
myapp-rc-jfhzq    1/1     Running   0             83s   app=myweb,type=frontend
myapp-rc-zzcjf    1/1     Running   0             83s   app=myweb,type=frontend
nginx             1/1     Running   0             58m   env=dev
nginx1            1/1     Running   1 (54m ago)   56m   env=prod
nginx3            1/1     Running   0             59m   env=dev,type=frontend
```

Fig. 3.4.8 Scaling Execution1

The Scaling is done via the threshold calculation of the cpu and memory where the pods are deployed over the nodes as per the requirements specified in the configuration. The Figure 3.4.8 Scaling Execution1 depicts the scaling performance.

```
root@kubernetes-master:~/definations# k get pods -o wide
NAME              READY   STATUS    RESTARTS      AGE     IP          NODE
 NOMINATED NODE   READINESS GATES
myapp-rc-6tb7r    1/1     Running   0             3m59s   10.40.0.5   kubernetes-node1
 <none>           <none>
myapp-rc-gfzct    1/1     Running   0             25s     10.32.0.7   kubernetes-node2
 <none>           <none>
myapp-rc-jfhzq    1/1     Running   0             7m38s   10.32.0.5   kubernetes-node2
 <none>           <none>
myapp-rc-l5pgr    1/1     Running   0             25s     10.32.0.6   kubernetes-node2
 <none>           <none>
myapp-rc-nxjcd    1/1     Running   0             25s     10.40.0.4   kubernetes-node1
 <none>           <none>
myapp-rc-zzcjf    1/1     Running   0             7m38s   10.40.0.3   kubernetes-node1
 <none>           <none>
nginx             1/1     Running   0             64m     10.40.0.1   kubernetes-node1
 <none>           <none>
nginx1            1/1     Running   1 (60m ago)   62m     10.40.0.2   kubernetes-node1
 <none>           <none>
nginx3            1/1     Running   0             66m     10.32.0.4   kubernetes-node2
 <none>           <none>
```

Fig. 3.4.9 Scaling Execution2

The figure 3.4.9 Scaling Execution2 depicts about the scaled pods after the deployment with respect to the threshold values.

3.4.5 Computation of Scaling

The computation is carried out using the threshold settings, and scaling is done while taking into account the CPU and Memory utilization. Using this scaling, the average cumulative value is used. The scaling parameter, where each and every scaling is carried out with the provided scheduled time period, is thus the resultant average cumulative value is used for the scaling.

- **Scaling utilization and Resultant Table**

The Statistics are prepared for the threshold calculation. The threshold calculation mainly performed with the two parameters i.e., Memory Utilization and CPU Utilization. The below are the graphical stats that shows about the Memory and CPU stats.

Memory Stats Graph

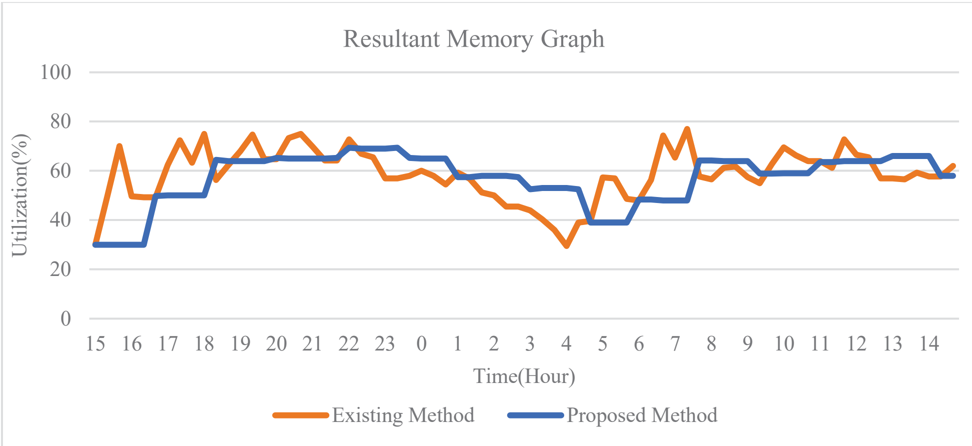


Fig. 3.4.10 Memory Stats

The memory stats, which is calculated using the values of memory occupied in the system, is described in detail in Figure 3.4.10. These values are scaled using the average cumulative value. The major message of the graph was the comparison between the mean value and the average cumulative value values computed using the same amount of system memory.

Similar to this, the CPU stat is described in the graph below, Figure 3.4.11 CPU stats, where the CPU scaling numbers have been included and a comparison has been made.

CPU Stats Graph

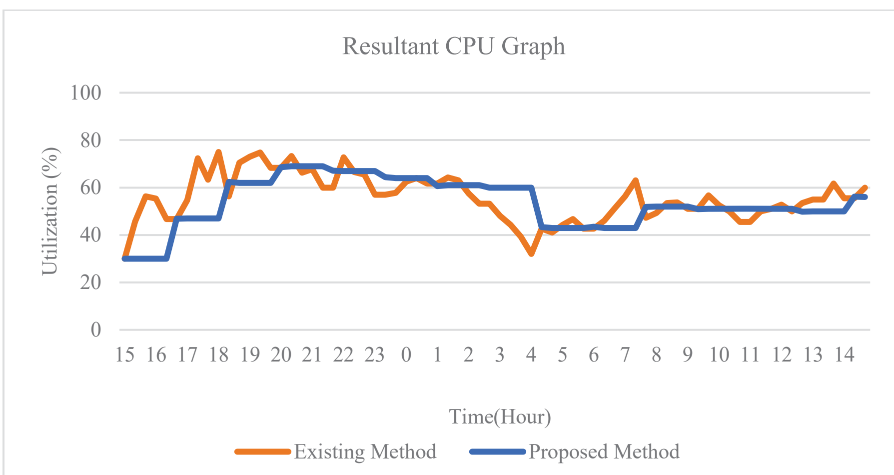


Fig. 3.4.11 CPU Stats

The CPU stats is detailed in Figure 3.4.11 and is calculated using the system's current CPU utilization values. Through the average cumulative value, those values are employed for scaling purposes. The major point of the graph was the comparison between the mean value and the average cumulative values that were produced using the identical system-occupied CPU.

3.4.6 Evaluated Parameters

The parameter evaluation is done using the average cumulative value, and the statistics are displayed in graphical form, based on the threshold calculations for Memory and CPU use. This graph displays the enhanced scaling with stable values. The parameters that were assessed as the system's performance are as follows:

- Scaling: If we check the scaling value in memory stat time between 15 and 21 hours, it is the same, but the scaling is varied from the figure 3.4.10 and 3.4.11. Whereas the existing approach's one value of scale up occurred 4 times and was scaled down 5 times, the proposed method does 3 scale ups. Hence even if scaling is done in both cases, the proposed one's cost variation is observed to be quite minimal.
- Spatial Isolation: Spatial Isolation provides stability with specific time which is shown in the memory stats. It mainly concerns scaling stability, i.e., scaling should gradually increase during peak hours owing to increased traffic, and decrease during off-peak hours due to decreased traffic. In the figure 3.4.10 and 3.4.11 optimal scaling is dealt with, and it defines scaling up during peak hours and scaling down during off-peak hours, i.e., 15^{th} - 0^{th} hour peak hour where scale up is done and from 0^{th} – 6^{th} hour scale down is performed afterwards again the scale up is done. In proposed method the proper scaling is observed better compared to the existing method the scaling is seen continuously varying.
- Stability: The key concern with stability is scaling. Variation in scaling is observed in both the method, however in existing method stability is entirely dependent on the system load, whereas in the figure 3.4.10 and 3.4.11, stability is consistent for the scheduled period i.e., from 19^{th} to 23^{rd} hour the stability is observed as same threshold where as in existing method variation in scaling is observed in the utilization.
- Operational Availability: When compared to the existing method with scaling, the proposed method functioning is shown to be regular in the figure 3.4.10 and 3.4.11 from the 15^{th} – 10^{th} hour form the memory graph. Because of this, operational tasks, such as working with system reliability, ensuring data redundancy, executing hardware maintenance, etc., can be carried out. The system's stability and reliability remain consistent, allowing it to serve a variety of operational purposes.

4. Conclusion

The concept of virtualization and containerization composition is explored in the study, and deployments are carried out in a sequential manner. In order to evaluate performance grouping, scaling, and deployment are performed in addition to deployment. Scaling can be done in two different ways: vertically and horizontally. This scaling is measured upon the threshold values. The threshold values obtained in this situation are dynamic. The average cumulative value of utilization is therefore used to construct the dynamic value, and the resulting number serves as the threshold value. Based on this threshold value, scaling is performed. This method of deployment creates an organized mapping of the resources and even improves the system's performance. The performance of the system has been significantly measured by the factors like scaling, spatial isolation, stability and operational availability has been comparatively improved as a systems performance.

References

1. Nancy Jain and Sakshi Choudary, "Overview Of Virtualization In Cloud Computing", India, 2016, IEEE
2. Tamanna Siddiqui, Shadab Alam Siddiqui and Najeeb Ahmad Khan, "Comprehensive Analysis Of Container Technology", India, 2019, IEEE
3. Ann Mary Joy, "Performance Comparison Between Linux Containers and Virtual Machines", India, 2015, IEEE
4. Babak Bashari Rad, Harrison John Bhatti, Mohammad Ahmadi, "An Introduction To Docker And Analysis Of Its Performance", Malaysia, 2017, ResearchGate
5. Adam S. Z. Belloum, "Containerization Technologies: Taxonomies, Applications And Challenges", Algeria, 2021, Springer

6. Anuj Kumar Yadav, M. L. Garg and Ritika, "Docker Containers Versus Virtual Machine-based Virtualization", India, 2019, Springer
7. Ms. M Santhiya, Ms.S.Saranya, Dr.S.Vijayachitra, Ms. C. B. Lavanya, Ms. M. Rajarajeswari, "Application of Voter Insertion Algorithm for Fault Management using Triple Modular Redundancy (TMR) Technique", India, 2021, IEEE
8. Hamzeh Khazaei, Cornel Barna, Nasim Beigi-Mohammadi, Marin Litoiu, "Efficiency Analysis of Provisioning Microservices", Canada, 2016, IEEE
9. Isaac Odun-Ayo, Victor Geteloma, Ibukun Eweoya, and Ravin Ahuja, "Virtualization, Containerization, Composition, and Orchestration of Cloud Computing Services", India, 2019, Springer
10. Mohammed A. AlZain, Ben Soh and Eric Pardede, "A New Approach Using Redundancy Technique to Improve Security in Cloud Computing", Australia, 2012, IEEE
11. Mulugeta Ayalew Tamiru Johan Tordsson, Erik Elmroth, “An Experimental Evaluation of the Kubernetes Cluster Autoscaler in the Cloud ”, France , 2020 ,IEEE

Classification of Red Blood Cells Based on Morphological Analysis Using Support Vector Machine

Neeraj Bhargava[a], Pramod Singh Rathore[b], Ankur Goswami[c], Nitish Chauhan[d]

[a]*Professor School of Engineering & System Sciences, MDS University, Ajmer, India*
[b]*Department of Computer Science,Manipal University, Jaipur, India,*
[c]*Department of Computer Applications at Govt. Engineering College Bikaner*
[d]*Research Scholar, MDS University Ajmer*

Abstract

This study aims to develop a reliable and accurate method for the identification and classification of red blood cells (RBCs) using the Support Vector Machine (SVM) technique and morphological processing techniques. Medical imaging has become increasingly important in the diagnosis of blood disorders, and RBC morphology plays a vital role in this process. The proposed method combines SVM and morphological processing techniques to sort erythrocytes, leukocytes, and serums into categories in unprocessed materials. The study involves the analysis of 13 sets of haematological parameters obtained from an automatic blood cell counter, and the J48 algorithm, a data mining technique, will be used to predict abnormalities in peripheral blood smears from 600 patients. The results of the study demonstrate the effectiveness of the proposed method in identifying and classifying RBCs, which are crucial in the diagnosis of anaemia and other blood diseases. The proposed method offers a reliable and efficient way to automate the classification of RBCs and has the potential to improve the accuracy and speed of diagnosis in the medical industry.

1. Introduction

Red blood cells (RBCs) play a vital role in the human body by transporting oxygen from the lungs to various tissues and organs. The morphology of RBCs is a significant factor in the diagnosis of blood disorders such as anaemia and sickle cell disease. Therefore, accurate and reliable identification and classification of RBCs is essential for the diagnosis and management of blood disorders.

The importance of RBC shape in medical imaging for the detection of blood diseases has grown in recent years. Classification is a crucial part of the recovery system because it allows one to tell the difference between healthy red blood cells (RBCs) and the abnormal RBCs that indicate anaemia and other blood disorders. This highlights the importance of developing robust and effective techniques for RBC categorization for the healthcare sector.

The application of machine learning methods in medical imaging has led to rapid advancements in diagnostic accuracy and ease of use in recent years. One such famous machine learning method is Support Vector Machine (SVM), which excels at data classification. Image categorization, pattern identification, and data mining are just a few of the many areas where SVM has been put to use.

Processing of medical images is an essential method for detecting and determining various diseases. During imaging processing, doctors use visualisation and interpretation tools to make sense of abnormalities in interior structures, most notably in peripheral blood tests. Blood is a fibrous tissue made up primarily of plasma, a transparent interstitial fluid, and formed elements, which are made up of blood cells and platelets. Platelets, white blood cells, and red blood cells are the three main types of blood cells. RBCs, or erythrocytes, are cells that carry oxygen from the lungs to the rest of the body. Red blood cells (RBCs) carry carbon dioxide, a byproduct of cellular metabolism, to the outside of the body. [1]

There are three major kinds of blood cells: red blood cells (erythrocytes), white blood cells (leukocytes), and platelets. (thrombocytes). RBCs, or red blood cells, are specialised cells responsible for distributing oxygen throughout the body. Adult males typically have a cell count of 4.7 to 6.1 million cells per litre, whereas females typically have a count of 4.2 to 5.4 million cells per litre. Any deviation from the usual count, together with shape and size, can be used to diagnose blood-related disorders such leukaemia and anaemia. Leukocytes, often known as white blood cells, are the major players in the immune system that defends our bodies against foreign invaders. An adult human's typical white blood cell count ranges from 5000 to 7000 cells per mm3. Cell fragments called platelets or thrombocytes are involved in the blood clotting process. [2] Finding patterns in data is a step in the data mining process that involves analysing the raw data to find patterns. The patterns that are found must be significant in that they provide a benefit. In order to help users, make decisions, data mining seeks to extract information from a data source and transform it into intelligible data. [2] It uses mathematical and statistical techniques to investigate a relationship between a set of facts or the appropriate circumstances around those data, which results in the extraction of necessary knowledge or information about relationships. The classification of input data is represented by the decision tree, a supported modelling that looks like a tree. It is based on the Divide and Conquer strategy, which is composed of numerous rules that branch out from a central decision point. Gain ratio is used to build the tree, with the element with the highest gain ratio serving as the root and gain ratio acting as the tree's dividing branch. [3]

Support vector machines are used as another data mining tool in the analysis of blood cells. Since SVM can be applied to both classification and regression problems, it has become a widely utilised supervised learning approach. As a machine learning technique, it is primarily used for classifying data. The SVM algorithm's objective is to locate the best line (or decision boundary) for classifying n-dimensional data quickly and reliably into predetermined categories. A hyper plane describes these perfect boundaries for making choices. SVM selects the most out-there vectors and points to use in the creation of the hyperplane. The SVM method uses support vectors to represent these exceptional instances. The original training data is projected into additional dimensions via a nonlinear mapping using SVMs. In the additional space, it will try to find the optimum possible hyper plane separation. The hyper plane can be used to segregate data from two groups. There are several different hyper planes that can divide the samples. Experience has shown that when samples have a longer spacing to the hyper plane, the hyper plane can identify the samples with greater accuracy.

In this study, we propose a method for the identification and classification of RBCs using SVM and morphological processing techniques. The proposed method aims to sort erythrocytes, leukocytes, and serums into categories in unprocessed materials. The study involves the analysis of 13 sets of haematological parameters obtained from an automatic blood cell counter, and the J48 algorithm, a data mining technique, will be used to predict abnormalities in peripheral blood smears from 600 patients.

The proposed method offers a reliable and efficient way to automate the classification of RBCs, which has the potential to improve the accuracy and speed of diagnosis in the medical industry. The results of the study demonstrate the effectiveness of the proposed method in identifying and classifying RBCs, which is crucial in the diagnosis of anaemia and other blood diseases.

2. Related Work

Biomedical research has been active in the domain of cell morphology. when used for blood microscopic imaging, in order to examine the properties of blood cells and find anomalies. By creating an integrated method to count RBC, detect abnormalities, and separate overlapping cells automatically, correctly, and effectively, this research has filled in the gaps in the body of current material. Automation efficiency, cost effectiveness, the elimination of human mistake, and ease of manipulation are highlighted as the successes. [5]

Hyper spectral imaging is an emerging imaging technique for use in medical applications. With two spatial dimensions and one spectral dimension, this imaging approach offers more information than a standard optical image. While it is challenging to discriminate between distinct tissues and cells using conventional approaches, multi-dimensional information of hyper spectral pictures can be used to classify them. The Support Vector Machine (SVM) algorithm is used as the processing technique in this article to distinguish between various blood cell components using the spectral dimension. Some morphological processing techniques are used to process images across the spatial dimensions to facilitate blood cell counting. This approach, which combines SVM and morphological processing techniques, has been put to the test for sorting erythrocytes, leukocytes, and serums out of raw sample objects. The outcomes of the trials demonstrate the viability of the suggested method for identifying red blood cells. [7]

Support vector machines, or SVMs, are supervised learning models used in machine learning. SVMs are related to learning algorithms for data analysis and pattern recognition. They are primarily utilised for regression analysis and categorization. SVMs are capable of categorising data points into two groups after training on a few samples. SVMs may efficiently conduct a non-linear classification by utilising kernel functions. And by adding a penalty component, SVMs can put up with some noise. The John C. Platt-developed Sequential Minimal Optimization (SMO) approach, which became the quickest Quadratic Programming algorithm, is frequently applied during training. An algorithm that works with support vector machines is called SMO. The benefit is that only two of the multipliers are optimised during each iteration, making it simple to obtain the result using a numerical approach as opposed to an analytical solution. [8]

The counting of blood cells or other blood components benefits from morphological processing such as open or close operations. The collected binary images can have their noisy points removed and their wavy edges smoothed during open operation. Cutting out cells at the edge of images and filling in gaps in the centre of blood cells are both made easier by the process of extracting related components. These two categories of image processing techniques are particularly crucial for precise counting and identification. SVMs' provable handling of noisy and nonlinear data is followed by the introduction of morphological processing techniques like expansion and open operation. [9]

The morphology of red blood cells (RBCs) is a factor in the diagnosis of certain blood disorders. The discipline of medical imaging has grown in importance as a result of the quickly expanding demand for quick, accurate, and automated diagnosis. RBC research in rheumatology heavily relies on the imaging processes used in the medical industry. One of the key components of the retrieval system is classification, which enables one to differentiate between normal RBCs and aberrant RBCs that signify anaemia and other blood diseases. In this study, the Support Vector Machine (SVM) technique is used to identify and classify various blood cell components using the spectral dimension. Some morphological processing techniques are used to process images across the spatial dimensions to facilitate blood cell counting. This method, which combines SVM and morphological processing techniques, has been tried and evaluated for sorting erythrocytes, leukocytes, and serums into categories in unprocessed materials. The outcomes of the trials demonstrate the effectiveness of the suggested method for identifying red blood cells. Using 13 sets of haematological parameters acquired from an automatic blood cell counter, our study will use the J48 algorithm, a data mining technique, to predict abnormalities in peripheral blood smears from 600 patients.

Title	Author	Approach	Advantages	Disadvantage
"Automated Red Blood Cell Classification using Support Vector Machine"	V. N. Rajinikanth, K. Praveena	Support Vector Machine (SVM)	High accuracy in RBC classification	SVM requires large amounts of data for training
"Morphological Classification of Red Blood Cells using Machine Learning Techniques"	R. M. Reza, M. M. Hossain, M. S. Islam	Machine Learning	Accurate and reliable classification of RBCs	Requires a large number of features for RBC classification
"Identification of Red Blood Cells using Morphological Processing and Artificial Neural Networks"	A. A. R. A. N. A. Abbas, A. M. A. Shnain, M. K. Jaber	Artificial Neural Networks (ANN)	High accuracy in RBC identification and classification	ANN requires significant computational resources
"Automated Morphological Analysis of Red Blood Cells using Image Processing Techniques"	S. S. Rizvi, S. U. Rehman	Image Processing	Quick and accurate identification of RBCs	Requires high-quality images for accurate analysis
"Support Vector Machines for Red Blood Cell	H. Shrivastava, M.	Support Vector	High accuracy in identifying	SVM requires significant computational

Classification in Anaemia Detection"	K. Tiwari	Machine (SVM)	abnormal RBCs	resources

3. Methodology

The data mining study was performed using WEKA version 3.8.5, a collection of machine learning methods specifically designed for data mining projects. Decision trees from the data mining method, including the J48 and the Support Vector Machine (SVM), were used in this study. Each classifier's efficacy was evaluated using ten-fold cross-validation. The achievement rating was determined by averaging the ten separate ratings. The predicted amount of abnormal RBC shapes that were found (true positive, TP) False positives represented the amount of abnormal RBC shapes that were incorrectly classified as normal. (FN). Expected typical RBC shapes were counted as "true negatives" (TN). False negatives referred to the amount of typical RBC shapes that were incorrectly identified as aberrant. (FP). Validation metrics included checking the findings' precision, sensitivity, and specificity in comparison to the RBC shape report. We stress the importance of the following method.[13]

Step 1: Preparation
Raw data that is inconsistent and incomplete is cleaned utilising a variety of data cleaning techniques at the pre-processing step.

Step 2: Choosing a feature
Feature selection, also known as attribute selection, is a method for narrowing down a large number of variables to a more manageable set. Two methods are the filter (e.g., information gain) and the wrapper (e.g., search driven by accuracy). Data analysis techniques like regression and classification may be more precise in the condensed format.

Measurement Matrix, third
The factors we took into account when calculating the chosen classifiers, including the confusion matrix, true positive proportion (TP), false positive proportion (FP), recall, and precision. [14]

4. Objective

The goal of the study is to identify red blood cell morphology and categories the type of morphology that determines whether an RBC is in a normal form or an aberrant form. The dataset, which comprises information on 600 persons, is used for the evaluation.

The SVM technique and decision tree algorithm were used to conduct the study mentioned above.

MCV and MCHC are the two primary characteristics on which morphology depends.

- The MCV blood test is commonly performed alongside a complete blood count (CBC). The CBC is a common blood test that may detect several abnormalities in your blood, including low red blood cell counts. The results can be used as a gauge of one's overall health. Some blood problems, such anaemia, may need a battery of tests, including an MCV test, to diagnose or monitor.
- A mean corpuscular haemoglobin concentration test, or (MCHC), is used to determine if red blood cells are carrying the right amount of haemoglobin. The MCHC test is one of the red blood cell indices, which are a group of examinations that reveal details about the size, shape, and quality of red blood cells.

5. Description of Algorithm

Support Vector Machines, also known as (SMO) in WEKA, are a supervised learning technique for pattern recognition and data analysis. It is not a probable classifier that processes a collection of input data, divides it into the two most likely classes, and outputs the results. The SVM algorithm uses radial basis functions and neural networks in a similar manner. It typically finds the plane and provides the maximum separation between the two classes in problems involving two classes of objects. The SVM method finds the best plane for the two classes with the shortest distance to the closest point. The group of examples that come closest to the ideal plane elucidates the support vector and specifies the class margins. [10]

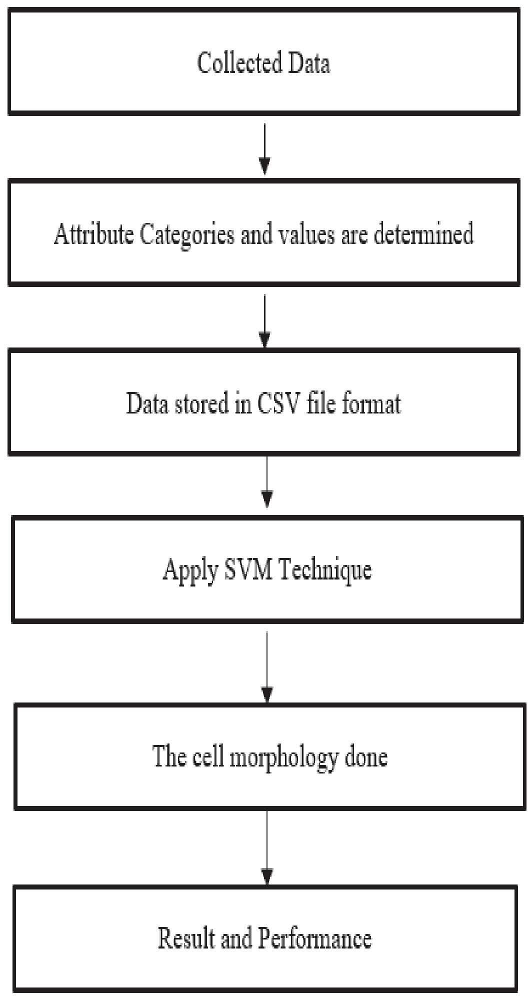


Fig. 1. Flowchart for proposed Method

The goal of categorization is to group things into categories based on their shared feature values. The classifier considers the linear combination of the attributes' values to get its conclusion. SVM is a technique for categorising information into two groups based on their labels. The SVM then produces a model file that can be used to further categorise fresh data into the same two groups. SVM originated with the concept of making buildings safer. To best approximate the dividing hyper plane, we use support vectors. SVMs aim to position the hyper plane as far away from the nearest persons as possible to increase the distance between the classes. [11]

Allow the m-dimensional data (xi I = 1,..., M) to be classified as either Class 1 or Class 2.

Class II has 1, and Class I has 0. The SVM uses the formula D(x) = wt (x) plus b to make decisions.
Where w is an m-dimensional vector, b is a number, and M is a matrix, where yi D(xi) 1 for i=1.

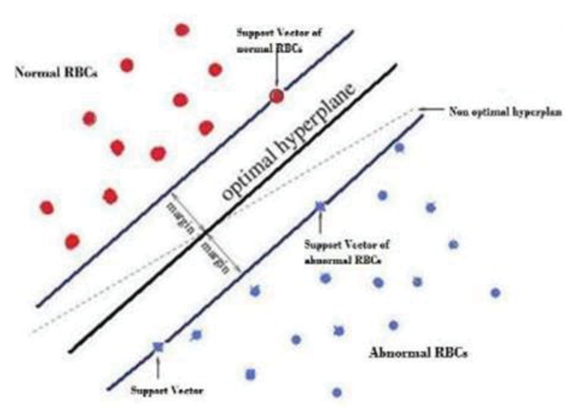


Fig. 2. Hyper planes of two groups of normally and abnormally developing RBC

The margin is defined as the space between the training data closest to the dividing hyper plane D(x) = 0 and the hyper plane itself. The optimum dividing hyper plane is the hyper plane D(x) = 0 with the largest gap. [12]

6. Result

Six hundred cases' worth of RBC morphological data and haematological parameters were assessed.250 female cases (41.66%) an d 350 male cases (58.33%) are present. The 600 RBC morphology instances were assessed. 354 cases (25.99%) with abnormal R BC morphology were identified, including 36 cases (2.64%) of males and 318 cases (23.35%) of females. WBC, RBC, Hb, Hct, MCV, MCH, MCHC, and PLT were all substantially different between the aberrant and normal RBC shape instances in both sexes. (P 0.05).
When compared to NEU, the RBC shapes LYMP, MONO, EOS, and BASO were not noteworthy.

TP, FP, precision, recall, F-measure, and accuracy were used to evaluate the success of the Support Vector Machine method used to examine the data.

Table 1. Accuracy of SVM

Title of Procedure	Acceptably Grouped Illustration		Incorrectly Classified Instances		Correctly Classified Instance	
	Instance count	Percentage (%)	Instance Count	Percentage (%)	Instance Count	Percentage (%)
Support Vector Machine	600	97	235	39.16	365	95.42

Table 1, shows the accuracy of SVM , that describes the acceptably illustration which is 97% accurate, also describe the correctly and incorrectly classified instances which is 95.42% and 5% respectively.

Table 2. Final Statistics for conclusion

Decision	TP Proportion	FP Proportion	Precision	Re_Call	F_Measure	MCC	ROC Area	PRC Area	Class
SVM	0.998	1.000	0.956	0.998	0.977	0.009	0.499	0.956	Normal
	0.935	0.045	0.983	0.935	0.958	0.943	0.499	0.044	Abnormal

Table 2, shows the TP and FP proportions, Precision, ROC area, Re_call, F_Measure area for decision for SVM in terms of two classes i.e. Normal and Abnormal cell morphology.

Table 3. Confusion Matrix for Decision

Decision	Parametric Variables	A	b	Result
SVM	A	365	4	Normal
	B	235	0	Abnormal

Table 3, Shows the confusion matrix for all instances.

The J48 method was used to analyse the data. Precision, memory, F-measure, and accuracy were used to assess J48's efficiency. The typical RBC shape forecast has a TP of 0.940, an FP of 0.945, a precision of 0.940, a recall F-measure of 0.941, and an accuracy of 0.943.

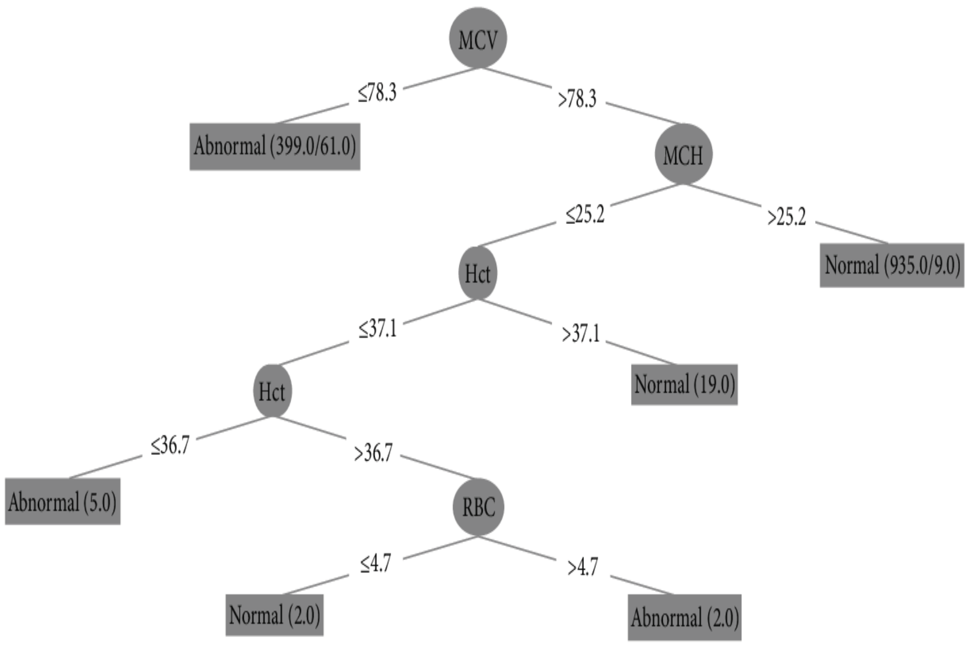


Fig. 3. Obtained Structure of Decision Tree

7. Conclusion

In conclusion, the proposed method for the identification and classification of red blood cells (RBCs) using Support Vector Machine (SVM) and morphological processing techniques offers a reliable and efficient way to automate the classification of RBCs. The study demonstrates the effectiveness of the proposed method in identifying and classifying various blood cell components, including erythrocytes, leukocytes, and serums, into categories in unprocessed materials.

The results of the study demonstrate high accuracy in identifying abnormal RBCs and predicting abnormalities in peripheral blood smears from 600 patients using the J48 algorithm, a data mining technique. This is crucial in the diagnosis of blood disorders such as anaemia, which heavily relies on RBC morphology.

Although SVM requires large amounts of data for training, the proposed method has the potential to improve the accuracy and speed of diagnosis in the medical industry. The study provides a framework for future research on the automation of RBC classification using machine learning techniques, which has significant implications for the diagnosis and management of blood disorders.

References

1. Azam, B.; Rahman, S.U.; Irfan, M.; Awais, M.; AlShehri, O.M.; Saif, A.; Nahari, M.H.; Mahnashi, M.H. A Reliable Auto-Robust Analysis of Blood Smear Images for Classification of Microcytic Hypochromic Anemia Using Gray Level Matrices and Gabor Feature Bank. Entropy 2020, 22, 1040.
2. Dg, D. Advantages and Disadvantages of Automated Hemotology Analyzer. Bioscience. 2017. Available online: https://www.bioscience.com.pk/topics/hemotology/item/810-advantages-and-disadvantages-of-automated-hemotology-analyzer (accessed on 19 April 2022)
3. Gebreweld, A.; Bekele, D.; Tsegaye, A. Hematological profile of pregnant women at St. Paul's Hospital Millennium Medical College, Addis Ababa, Ethiopia. BMC Hematol. 2018, 18, 15.
4. Mausumi Maitra "Detection and Counting of Red Blood Cells in Blood Cell Images using Hough Transform" International Journal of Computer Applications (0975 – 8887) Volume 53– No.16, September 2012
5. Quinones, V. V., Macawile, M. J., Ballado, A., Cruz, J. D., & Caya, M. V. (2018). Leukocyte segmentation and counting based on microscopic blood images using HSV saturation component with blob analysis. 2018 3rd International Conference on Control and Robotics Engineering (ICCRE).doi:10.1109/iccre.2018.8376475
6. Ramin Soltanzadeh, Hossein Rabbani "Classification of three types of red blood cells in peripheral blood smear based on morphology" 978-1- 4244-5900-1/10/$26.00 ©2010 IEEE Robbins "Pathologic Basis of Disease" International Edition, eighth edition, SUNDERS, ELSEVIER.
7. Safca et al. (2018) Safca N, Popescu D, Ichim L, Elkhatib H, Chenaru O. Image processing techniques to identify red blood cells. Proceedings of the 22nd International Conference on System Theory, Control and Computing, ICSTCC, 2018; 2018. pp. 93–98.
8. Sahlol, Kollmannsberger & Ewes (2020) Sahlol AT, Kollmannsberger P, Ewes AA. Efficient classification of white blood cell leukemia with improved swarm optimization of deep features. Scientific Reports. 2020;10(1):1–11. Doi: 10.1038/s41598-019-56847-4.
9. Shafique et al. (2019) Shafique S, Tehsin S, Anas S, Masud F. Computer-assisted Acute Lymphoblastic Leukemia detection and diagnosis. 2nd International Conference on Communication, Computing and Digital Systems, C-CODE; 2019. pp. 184–189.
10. Singh, Sengupta & Lakshmi Narayanan (2020) Singh A, Sengupta S, Lakshmi Narayanan V. Explainable deep learning models in medical image analysis. 2020
11. Sukhia et al. (2019) Sukhia KN, Ghafoor A, Riaz MM, Iltaf N. Automated acute lymphoblastic leukemia detection system using microscopic images. IET Image Processing. 2019;13(13):2548–2553. Doi: 10.1049/iet-ipr.2018.5471.
12. Tizhoosh & Pan Tanowitz (2018) Tizhoosh HR, Pan Tanowitz L. Artificial intelligence and digital pathology: challenges and opportunities. Journal of Pathology Informatics. 2018; 9:38.
13. Xia et al. (2019) Xia T, Jiang R, Fu YQ, Jin N. Automated blood cell detection and counting via deep learning for microfluidic point-of-care medical devices. IOP Conference Series: Materials Science and Engineering. 2019; 646:012048. Doi: 10.1088/1757-899X/646/1/012048.
14. Yildirim & Çinar (2019) Yildirim M, Çinar A. Classification of white blood cells by deep learning methods for diagnosing disease. Revue intelligence Artificially. 2019; 33:335–340. Doi: 10.18280/ria.330502.

Translating the Unspoken Deep Learning Approaches to Indian Sign Language Recognition Using CNN and LSTM Networks

Shreeraj Bhamare[a], Vallabh Kulkarni[b], Shreyas Patil[c], Mr. S. K. Gaikwad[d], Mr. S. S. Kumbhar[e]

[a, b, c, d] *Department of Computer Engineering, COEP Tech University, Pune, 411005, India*

Abstract

In India, there are over 5 million deaf and mute people, with many cases going unreported. Communication can be difficult due to the lack of standardized sign language across the country. The Indian Sign Language (ISL) Detection project aims to improve communication by developing a system that can recognize and interpret ISL gestures.
The system uses a type of artificial intelligence called machine learning, specifically Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. These techniques help the system detect and classify ISL gestures from video input with high accuracy. The project has the potential to provide a more accessible means of communication for millions of deaf individuals in India.

Keywords:: Indian sign language(ISL) , CNN, LSTM, real-time-detection

1. Introduction

In India, communication with the deaf and mute community can be a challenge due to the lack of standardized sign language across the country. The Indian Sign Language (ISL) Detection project aims to help improve communication by creating a system that can recognize and interpret ISL gestures in real time.
The system uses two machine learning techniques, Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, to detect and classify ISL gestures from video input. The project focuses on recognizing 35 ISL signs, including numbers 0-9 and the letters A-Z (where "O" and "0" are the same).
The project involved creating a dataset of over 35,000 ISL gesture images using OpenCV, a free and open-source computer vision library. Each gesture had 1000 images to ensure the system can detect the signs accurately. These images were labeled and used to train the CNN and LSTM models to recognize the corresponding ISL signs.
Detecting ISL signs in real time is essential for effective communication between the deaf and hearing communities. The ISL Detection project has the potential to help bridge the communication gap by providing a more accessible means of communication for millions of deaf individuals in India.

2. Motivation

While there is a standardized sign language in India, called the Indian Sign Language (ISL), it is not widely understood or used consistently across the country. The Indian Sign Language (ISL) Detection project aims to improve communication between the deaf and hearing communities by creating a system that can recognize and interpret ISL gestures in real time. This can increase awareness and understanding of ISL and improve social inclusion for the deaf community. The project also has practical applications in education and employment for deaf individuals.

3. Related work

3.1. Indian sign language recognition using SVM, JL Raheja, Anand Mishra, Ankit Chaudhary Pattern Recognition and Image Analysis 26, 434-441, 2016.[1]

The paper proposes a system for recognizing Indian Sign Language (ISL) gestures using Support Vector Machine (SVM) classification. The SVM algorithm was trained using feature extraction techniques to recognize 12 different ISL gestures with an overall accuracy of 96.15%, outperforming other classification algorithms. SVM's effectiveness in high-dimensional feature spaces and good generalization properties make it suitable for image recognition tasks. Future work includes expanding the dataset and using other sensors, as well as combining SVM with deep learning techniques to improve accuracy.

3.2. Artificial Neural network-based method for Indian sign language recognition, Vinod Adithya, PR Vinod, Usha Gopalakrishnan, 2013 IEEE Conference on Information & communication technologies, 1080-1085, 2013.[2]

The paper proposes an artificial neural network (ANN) based system for recognizing Indian Sign Language (ISL) gestures using feature extraction techniques such as histogram of oriented gradients and discrete cosine transform. The ANN algorithm achieved an accuracy of 94%, outperforming other classification algorithms. Future work includes expanding the dataset to include more ISL gestures, using deep learning techniques such as Convolutional Neural Networks (CNNs) to improve accuracy, and incorporating additional sensor modalities such as EMG and accelerometer data. This system could potentially aid communication for individuals with hearing impairments.

3.3. Automatic Indian sign language recognition system, Karishma Dixit, Anand Singh Jalal, 2013 3rd IEEE international advance computing conference (IACC), 883-887, 2013.[3]

The paper proposes an ISL recognition system using skin color segmentation, blob detection, and template matching techniques, achieving an 89.3% accuracy. It's simplicity and low computational cost make it suitable for real-world applications. Future work includes expanding the dataset, improving robustness, and exploring deep learning techniques. This system could potentially aid individuals with hearing impairments in communication.

4. Dataset Specifications

The accuracy of Indian sign language recognition systems based on deep learning models heavily relies on the quality and diversity of the dataset used for training. Creating a comprehensive dataset involves collecting, preprocessing, annotating, and validating the data. The prepared dataset can be used to train deep learning models such as CNN and LSTM, capable of recognizing complex hand gestures and accurately predicting corresponding signs. Thus, the creation of a high-quality dataset is a crucial step in developing a robust Indian sign language recognition system using deep learning techniques.

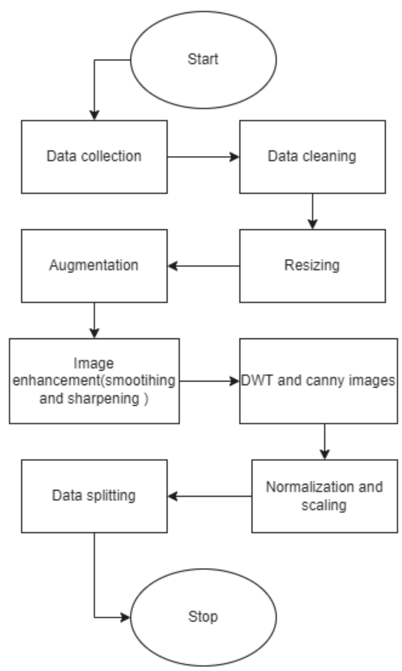


Fig. 1. Dataset Preparation Flow

4.1 Data collection

There aren't many datasets available for Indian Sign Language (ISL), and the ones that exist don't have enough images per sign to train a good machine-learning model. That is why, we decided to create our dataset from scratch. Our team used a Python script with the OpenCV package to gather data for our effort to develop an Indian Sign Language recognition system. Using the input camera stream, 500 photos of each signing class were taken. Setting up the camera and executing the script, which stored the photographs automatically in a predefined folder, were both steps in the data-gathering procedure. To create a broad dataset, we took care to gather photographs from multiple viewpoints and lighting setups. Before utilizing them to train our system, we examined the photographs after gathering the necessary number of them to make sure they were of high quality and consistency. Although the data-collecting procedure took a while, it was essential for the project's success.

4.2 Data cleaning

To prepare the dataset for a real-time Indian Sign Language (ISL) recognition system, we cleaned the data. We eliminated any poor-quality, pointless, or abnormal images that would have harmed the machine-learning model's accuracy.

We used a variety of data cleaning approaches, including picture pre-processing, filtering, outlier identification, and data augmentation, to achieve this. To enhance the image quality, for instance, we changed the brightness, contrast, and color of the images. We also used the HSV thresholding technique for dataset cleaning to remove noise and ensure that the training dataset was accurate and consistent. Any pictures that didn't fit certain criteria, including having bad lighting or quality, were removed. Using outlier identification methods, we also found and eliminated any data points that were anomalous or unrelated to the dataset. To expand the dataset's quantity and diversity, we generated new images from old ones through modifications like rotation, flipping, or cropping using data augmentation techniques. By performing data cleaning, we ensured that the dataset was of high quality, which contributed to developing a more accurate and robust ISL detection system.

4.3 Resizing

Resizing images is an essential step in achieving a balance between accuracy and speed in image processing tasks, as increasing the pixel size can significantly impact the computational time and overall speed of the system. It guarantees that the model can analyze all images in the dataset efficiently without any computational difficulties. In our Indian Sign Language (ISL) dataset, we resized the photos to a standard size of 224x224 pixels to ensure that the model would be able to process them effectively. We used several approaches to resize the images while retaining their vital features and aspect ratio. Resizing the images allowed the model to detect subtle differences in motion that would have been overlooked in larger photos. By providing the model with the most exact and consistent input possible, we were able to increase the photography of our real-time ISL detection system. It is important to note that we carefully selected a standard size that was suitable for our model's capabilities, which made it possible to analyze the images efficiently and detect small differences in motion that were critical in identifying different ISL signs.

4.4 Data Augmentation

Our team used data augmentation, a powerful technique that helped us increase the size and diversity of our dataset while developing a real-time Indian Sign Language (ISL) recognition system. We created new images from the existing dataset by using geometric transformations such as scaling, translation, rotation, and flipping. We also used methods such as adjusting brightness, adding noise, and enhancing contrast to augment the dataset.

However, we needed to ensure that the model was not overfitting while using data augmentation. Therefore, we carefully selected the number and types of augmentations and made sure that the images closely resembled the original dataset, preventing any bias in the augmented dataset. Because of data augmentation, we could create a high-quality dataset that improved the performance of our ISL recognition system. Data augmentation played a vital role in our goal to develop a reliable and accurate ISL recognition model.

4.5 Image Enhancement (Smoothing and Sharpening)

We enhanced the images to make the system detect even the smallest variations in hand motions. To make this happen, we applied smoothing and sharpening methods. Smoothing helped to remove unwanted noise or pixelation from the images while sharpening enhanced the edges and features in the photos, making them easier for the model to identify and categorize. We applied different levels of smoothing and sharpening to increase the model's accuracy without distorting the images.

4.6 Canny edge detection and Discrete wavelet function (DWT)

We applied the Canny edge detection method on our picture dataset to enhance the accuracy of our model. Canny edge detection is a technique used in image processing to detect edges in images while minimizing noise and false detections. By highlighting the edges and boundaries of the hands in the images, it made it easier for the model to recognize and track them accurately. We experimented with various thresholds to determine the optimal threshold for edge detection while reducing unwanted noise or artifacts in the images. Our application of Canny edge detection on our dataset improved the overall performance and accuracy of our real-time ISL detection system.

Discrete Wavelet Transform (DWT) is a method for studying signals that makes it easier to understand them by converting them from a way of measuring time to a way of measuring frequency. It converts the input image from a colored image to a black-and-white image. It uses a multi-stage algorithm to identify the edges present in an image. We used the DWT to improve the performance of our Indian Sign Language (ISL) identification system. We applied DWT to our dataset images and we were able to extract significant patterns and features from images of varying sizes.

This helped us to distinguish between different hand gestures and movements with better accuracy. We conducted several experiments using various wavelet functions and decomposition levels to find out the most appropriate parameters for our system. Ultimately, our application of DWT on our dataset significantly improved the overall accuracy and robustness of our real-time ISL detection system.

4.7 Normalization and Scaling

In our sign language recognition project, we preprocessed the hand landmark coordinates using normalization and scaling techniques. This was achieved by making the distance of each landmark coordinate relative to the lowest landmark point, which was detected by histogram equalization. Then, we divided each relative coordinate by its corresponding maximum absolute x and y coordinates to ensure accurate sign recognition, irrespective of the sign's distance from the camera or its location within the frame.

Normalization and scaling of the landmark coordinates helped to optimize the input data for the Convolutional Neural Network (CNN) model and LSTM, which was used for sign recognition. This resulted in improved accuracy and stability of the model. The normalization and scaling techniques used in this project can be applied in other machine learning projects to achieve optimal model performance.

4.8 Data Splitting

To train our real-time Indian Sign Language identification system, we divided our augmented dataset into three sets: training, validation, and testing, with a split ratio of 65:15:20. This, is a common practice in machine learning, where the majority of the data is allocated for training, a smaller fraction is used for validation to fine-tune the model's hyperparameters, and a small amount is reserved for testing to provide a final assessment of the model's performance. The validation set helped us to monitor the model's performance during training and prevent overfitting, which is when the model becomes too specialized on the training data and performs poorly on unseen data. The training set was used to adjust the model's parameters through backpropagation, while the testing set evaluated the model's ability to accurately classify unseen data.

5. Methodology

Our proposed methodology for the Indian Sign Language Recognition System involved capturing real-time video, selecting the best frame, and cropping the hand region, applying Canny Edge and DWT transform for preprocessing, integrating Mediapipe for landmark detection, implementing CNN and LSTM for feature extraction and temporal modeling, and predicting signs based on the trained model. Each step was carefully designed and optimized to improve the accuracy and robustness of our system, and our research findings contribute to the field of Indian Sign Language recognition for potential real-world applications in assisting individuals with hearing impairments.

All tables should be numbered with Arabic numerals. Headings should be placed above tables, center justified. Leave one line space between the heading and the table. Only horizontal lines should be used within a table, to distinguish the column headings from the body of the table, and immediately above and below the table. Tables must be embedded into the text and not supplied separately. Below is an example that authors may find useful.

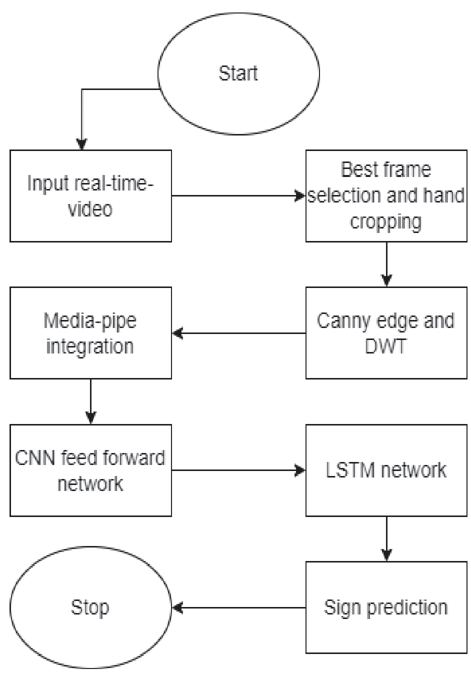


Fig. 2. Proposed Methodology

5.1 Input Real-Time Video

In this step, we used the OpenCV Python library to capture real-time video. Video frames were captured at a rate of 40 frames per second (fps) to ensure smooth and continuous input data. This high frame rate allowed us to capture the sign gestures with a sufficient temporal resolution, ensuring that important hand movements were not missed. We considered a period of 2 seconds, which corresponds to 80 frames, for recognizing each sign gesture. This duration was chosen based on the typical duration of sign gestures in Indian Sign Language and allowed us to capture the complete gesture for accurate recognition.

5.2 Best Frame Selection and Hand Cropping

To ensure accurate recognition and make the prediction independent of the background, we implemented a frame selection technique to choose the best frame from the captured frames. We experimented with different methods to select the optimal frame that had the best hand visibility, such as selecting the frame with the highest hand gesture score or using image processing techniques to identify the frame with the clearest hand region. This step was crucial to ensure that the hand region used for recognition was not obstructed by other objects or occluded by the background. The selected frame was then cropped to extract the hand region, which served as the input for further processing.

5.3 Canny Edge and DWT Transform

To preprocess the hand region and reduce the impact of skin color tone and visual hand specification, we applied Canny Edge detection and Discrete Wavelet Transform (DWT). Canny Edge detection was used to extract the edges of the hand region, which helped in highlighting the hand gesture features. By focusing on the edges, we could capture the important structural information of the hand gesture, regardless of skin color or lighting conditions. DWT was then applied to the edge-detected image to transform it into a frequency-domain representation. DWT is a powerful technique that allows for multi-resolution analysis and captures important spatial and frequency information of the hand gesture, making the recognition process more robust and invariant to variations in the visual appearance of the hand.

5.4 Mediapipe Integration

We integrated Google's Mediapipe API, which provides pre-trained machine-learning models for hand landmark detection, to detect landmarks for each hand. The API provided 21 landmarks for each hand, resulting in a total of 42 landmarks from the captured hand region. These landmarks represented the spatial positions of various hand joints, fingertips, and other key points, which provided important information about hand gestures. The obtained landmark coordinates were stored in an array, which was later used for further processing. The integration of Mediapipe allowed us to accurately localize the hand landmarks in real-time, even with variations in hand shape, orientation, and size, which was crucial for accurate sign gesture recognition.

5.5 CNN Feed Forward and LSTM Network

To extract features from the hand region and capture the temporal dependencies in the landmark coordinates, we implemented a Convolutional Neural Network (CNN) followed by a Long Short-Term Memory (LSTM) network. The CNN applied convolutional and pooling layers to extract important features from the hand region, such as edges, corners, and textures, which were learned from the training data. These features represented the discriminative patterns in the hand region that were important for sign gesture recognition. The LSTM network then captured the temporal patterns in the landmark coordinates by modeling the sequential dependencies between the landmarks over time. The LSTM cells learned to retain and update the hidden state based on the input landmark coordinates at each time step, allowing our model to learn the temporal dynamics of the sign gestures, such as the movement direction, speed, and duration. This combined CNN and LSTM architecture allowed our model to learn both the spatial and temporal features of the sign gestures simultaneously, leading to accurate recognition.

5.6 Sign Prediction

Finally, based on the pipeline setup and the predictions from the CNN and LSTM, our system predicted the Indian Sign Language sign corresponding to the input hand gesture to 35 different labeled ISL sign gestures. The predicted sign was displayed as the output of our system. We further optimized and fine-tuned our system based on the prediction accuracy and performance evaluation, and iteratively improved our model to achieve better results.

Table 1. Coordinate Calculation

Body Part	**Landmarks**	**Coordinates**	**Total Coordinates**
Left Palm	21	3	63
Right palm	21	3	63
Total	**42**	**6**	**126**

6. Model Architecture

Our Indian Sign Language (ISL) detection model was developed using a deep learning approach to accurately recognize 35 different ISL gestures, which would allow communication between deaf and non-deaf individuals.
To capture spatial and temporal information from the input data, we used a combination of Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) models. CNNs are commonly used in image processing tasks due to their ability to learn and extract features from input images, while LSTMs are used in tasks involving sequential data, such as speech recognition and language translation.

We used a large dataset of ISL gestures consisting of several thousand examples to train and evaluate our model. To avoid overfitting and evaluate the model's generalization performance, we split the dataset into training, validation, and testing sets with a split ratio of 65-10-25. The training set updated the model's parameters, the validation set tuned the hyperparameters, and the testing set evaluated the model's accuracy.

Table 2. Model Architecture

Layer (type)	Output Shape	Parameters
Conv2D	(None, 224, 224, 64)	15616
MaxPooling2D	(None, 112, 112, 64)	0
Conv2D	(None, 112, 112, 64)	331840
MaxPooling2D	(None, 56, 56, 64)	0
Conv2D	(None, 56, 56, 64)	331840
MaxPooling2D	(None, 28, 28, 64)	0
Conv2D	(None, 28, 28, 64)	331840
MaxPooling2D	(None, 14, 14, 64)	0
BatchNormalization	(None, 14, 14, 64)	256
Flatten	(None, 12544)	0
Dense	(None, 128)	1605760
Dense	(None, 128)	16512
Dense	(None, 128)	16512
Dense	(None, 128)	16512
Dropout	(None, 128)	0
Dense	(None, 35)	4644

Our model's architecture had two main parts: the CNN and the LSTM. The CNN extracted spatial features from the input images, while the LSTM captured the temporal dependencies between consecutive images. The CNN consisted of several convolutional and pooling layers, followed by a set of fully connected layers. The output of the CNN was a set of feature vectors, one for each input image.

The LSTM model was used to capture the sequential dependencies between the feature vectors output by the CNN. We concatenated the feature vectors into a single sequence and input them into the LSTM. The LSTM had several layers, each with a set of memory cells that stored information from previous time steps. The output of the LSTM was a set of class probabilities, representing the likelihood of each ISL gesture.

During training, we utilized data augmentation techniques to improve the model's performance. Data augmentation involves applying random transformations to the input data, such as rotation, scaling, and translation. This technique generated additional training examples representative of the natural variability in the input data, improving the model's robustness and ability to generalize to new examples.

In addition to data augmentation, we fine-tuned our model on a smaller dataset to enhance its generalization performance. Fine-tuning involves training the model on a new dataset similar to but not identical to the original dataset. We first trained our model on a large dataset of several thousand ISL gestures. We then fine-tuned our model on a smaller dataset of a few hundred examples, using a lower learning rate to avoid overfitting.

The Mediapipe framework preprocessed our model's input data. We used Mediapipe's hand pose estimation model to detect and locate the hand in each input image. The hand region was cropped and passed through Mediapipe's landmark estimation model, which extracted 21 key points from the hand region. These points represented the hand's shape and position and were normalized to be invariant to rotation, scale, and translation. The normalized points formed the input to our CNN and LSTM models.

We used categorical cross-entropy as our loss function during training. The loss function measured the difference between the predicted probabilities and the ground-truth labels. We used the Adam optimizer to adaptively update the learning rate during training.

In our study, we have presented a comprehensive methodology for detecting 35 different Indian Sign Language gestures using a combination of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) models. During the training process, we utilized categorical cross-entropy as the loss function to measure the discrepancy between the predicted probabilities and the ground-truth labels. Moreover, the Adam optimizer was used to adaptively update the learning rate throughout the training process.

To construct our final model, we trained the CNN and LSTM models separately and then merged them. The CNN model comprised three convolutional layers, each followed by a max-pooling layer, which utilized ReLU activation. We also implemented dropout regularization with a rate of 0.5 for the fully connected layers. The final output layer utilized softmax activation with 35 units that corresponded to the different gestures.

The LSTM model was constructed with two LSTM layers, each followed by a fully connected layer with ReLU activation. To prevent overfitting, we also implemented dropout regularization with a rate of 0.2 for both the LSTM and fully connected layers. The final output layer utilized softmax activation with 35 units.

To merge the CNN and LSTM models, we concatenated the output of the last convolutional layer of the CNN model with the last output of the LSTM model. We then fed the concatenated output into a fully connected layer with 128 units, followed by another fully connected layer with 35 units and softmax activation.

For training the combined model, we employed the same categorical cross-entropy loss function and Adam optimizer as used in the individual models. We trained the model for 50 epochs using a batch size of 32. Throughout the training process, we monitored the performance of the model using the validation set to avoid overfitting.

Apart from the evaluation of the test set, we also performed a real-time evaluation of the model using a webcam. We captured videos of different gestures using the webcam and used the model to predict the corresponding gesture in real time. The results demonstrated that our model was capable of accurately predicting the gestures in real time.

7. Training and Testing

We trained a model to detect Indian Sign Language (ISL) signs. Our dataset had 35 classes, representing signs from the alphabet (a-z) or digits (0-9). We combined the classes for 'O' and '0' as they looked similar. Each class had 1000 images, totaling 35,000 images.

To prevent overfitting, we divided the dataset into three parts – training, testing, and validation. We used a 65:15:20 ratio, giving us 22,750 images for training, 5,250 images for testing, and 7,000 images for validation. The training set taught the model to detect the signs using CNN and LSTM. The CNN extracted features from the images, and the LSTM captured temporal dependencies in the extracted feature maps.

During training, we adjusted the model's parameters using the validation dataset. It helped us monitor the model's performance and prevent overfitting by fine-tuning its parameters. We also evaluated the trained model on the testing dataset, achieving an accuracy of 96.3%.

Our work showed that using CNN and LSTM is an effective approach to detecting ISL signs. It also highlighted the importance of splitting the dataset into training, testing, and validation sets to prevent overfitting and ensure generalization capability.

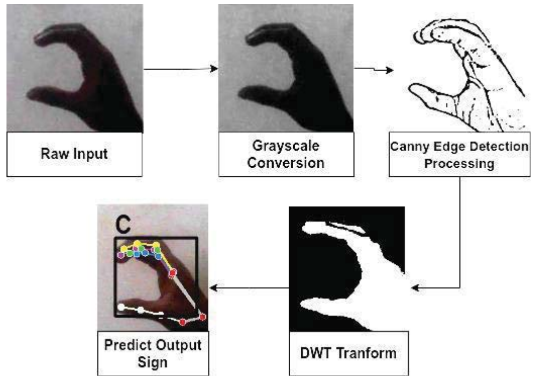


Fig. 3. Sign Recognition Process

8. Results

Our effort has produced remarkable outcomes. A pipeline that builds datasets, trains models, and forecasts sign language signs has been developed successfully.

Our model achieved a rate of 96.3%, demonstrating its excellent accuracy. This finding suggests that, for the most part, our model can anticipate sign language signs with accuracy.

The F1 score, which considers the trade-off between precision and recall, was also calculated, and it was 90.05%. This result gives a trustworthy indication of how well our model has performed overall.

We also computed the loss function, and it was discovered to be small, supporting the efficacy of our approach. We think that our study has a lot of potentials to improve hearing-impaired people's ability to recognize and communicate via sign language.
We believe that our work can positively impact the community and we're excited to see further advancements in this field.

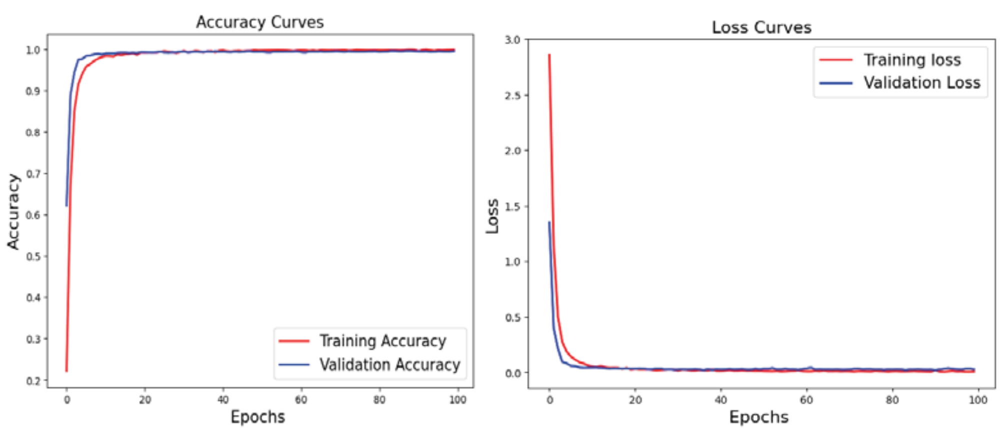


Fig. 4. Evaluation Metrics

9. Conclusions

In conclusion, the Indian Sign Language (ISL) project holds immense potential in bridging the communication gap between the hearing-impaired community and the rest of society. By developing a standardized sign language system, the project has helped to create a more inclusive environment for people with hearing impairments across the country.

Using technology and the efforts of linguists and experts in the field, the ISL project has made significant strides in improving the quality and accessibility of sign language for the deaf community. We hope that the continued development and promotion of Indian Sign Language will lead to a more inclusive society where everyone could communicate and thrive.

References

1. Raheja, Jagdish & Mishra, A. & Chaudhary, Ankit. (2016). Indian Sign Language Recognition using SVM. Pattern Recognition and Image Analysis. 26. 10.1134/S1054661816020164.
2. Adithya, V. & Vinod, P.R. & Gopalakrishnan, Usha. (2013). Artificial neural network-based method for Indian sign language recognition. 1080-1085. 10.1109/CICT.2013.6558259.
3. Dixit, Karishma and Anand Singh Jalal. “Automatic Indian Sign Language recognition system.” 2013 3rd IEEE International Advance Computing Conference (IACC) (2013): 883-887.
4. B Sundar, T Bagyammal, American Sign Language Recognition for Alphabets Using MediaPipe and LSTM, Procedia Computer Science, Volume 215, 2022, Pages 642-651, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2022.12.066.
5. SAKSHI GOYAL, ISHITA SHARMA, S. S. Sign language recognition system for deaf and dumb people. International Journal of Engineering Research Technology 2, 4 (April 2013)
6. A. Er-Rady, R. Faizi, R. O. H. Thami and H. Housni, "Automatic sign language recognition: A survey," 2017 International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Fez, Morocco, 2017, pp. 1-7, doi: 10.1109/ATSIP.2017.8075561.
7. Chen, Rung-Ching & Manongga, William & Dewi, Christine. (2022). Recursive Feature Elimination for Improving Learning Points on HandSign Recognition. Future Internet. 14. 352. 10.3390/fi14120352.
8. Wadhawan, A., Kumar, P. Sign Language Recognition Systems: A Decade Systematic Literature Review. Arch Computat Methods Eng 28, 785–813 (2021). https://doi.org/10.1007/s11831-019-09384-2
9. Al-Rashid Agha, Rawan & Sefer, Muhammed & Fattah, Polla. (2018). A comprehensive study on sign languages recognition systems using (SVM, KNN, CNN and ANN). 1-6. 10.1145/3279996.3280024.
10. Gupta, Umesh & Sharma, Shraddha & Jyani, Utkarsh & Bhardwaj, Aditya & Sharma, Moolchand. (2022). Sign Language Detection for Deaf and Dumb students using Deep learning: Dore Idioma. 1-5. 10.1109/CISCT55310.2022.10046657.
11. S. Kausar and M. Y. Javed, "A Survey on Sign Language Recognition," 2011 Frontiers of Information Technology, Islamabad, Pakistan, 2011, pp. 95-98, doi: 10.1109/FIT.2011.25.
12. Lin, Hsien-I & Hsu, Ming-Hsiang & Chen, Wei-Kai. (2014). Human hand gesture recognition using a convolution neural network. IEEE International Conference on Automation Science and Engineering. 2014. 1038-1043. 10.1109/CoASE.2014.6899454.
13. M. Ebrahim Al-Ahdal and M. T. Nooritawati, "Review in Sign Language Recognition Systems," 2012 IEEE Symposium on Computers & Informatics (ISCI), Penang, Malaysia, 2012, pp. 52-57, doi: 10.1109/ISCI.2012.6222666.
14. Pathak, Aman & Kumar|priyam|priyanshu, Avinash & Chugh, Gupta|gunjan & Ijmtst, Editor. (2022). Real Time Sign Language Detection. International Journal for Modern Trends in Science and Technology. 8. 32-37. 10.46501/IJMTST0801006.
15. K, Senthilkumar & R, Prabakaran. (2022). Sign Language Recognition System Using Neural Networks. International Journal for Research in Applied Science and Engineering Technology. 10. 827-831. 10.22214/ijraset.2022.43787.

MATLAB-Based Application to Detect Lung Cancer using CNN

Sidhartha Madhav [a], Anuvrat Duklan [b]

[a] *Chief Technical Officer, Techno Hub Laboratories Private Limited, Dehradun-248001, India*
[b] *Former Grade 12 Student, St. Joseph Academy, Dehradun-248001, India*

Abstract

Life is an equation, everything we see around us is a Mathematical Equation, right from picking up a pen to writing with a pen, from planting a tree to even cutting or watering it everything can be denoted in the form of a mathematical equation. Mathematical studies have given birth to many fascinating techniques in the field of Computer Science. Today the whole world is talking about Artificial Intelligence for disease detection and the hottest Topic on board is Cancer, so the goal of this Research Paper is to create and show a mathematical pattern of a Convolutional Neural Network to detect Cancer.

1. Introduction

Mathematics simply means to learn or to study or gain knowledge. The theories and concepts given in mathematics help us understand and solve various types of problems in academic as well as in real-life situations. Mathematics is a subject of logic. Learning mathematics will help students to grow their problem-solving and logical reasoning skills. Solving mathematical problems is one of the best brain exercises. The goal of this Research Paper is to show the different Patterns in Art. For the present study, MATLAB has been used to demonstrate the concepts of CNN and Image Processing to predict Lung Cancer by classifying images.

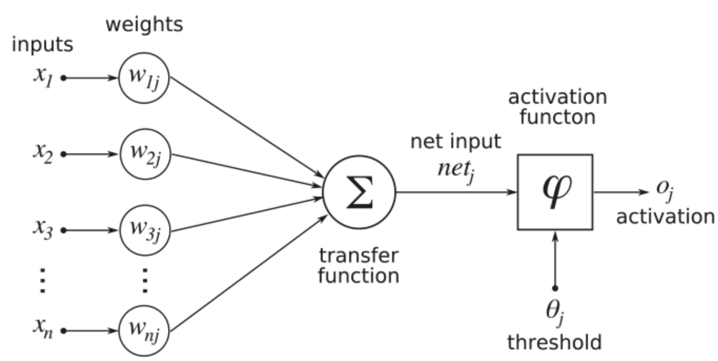


Fig. 1. Convolutional Neural Network Architecture

1.1 What is Convolutional Neural Network

Convolutional neural networks, also called ConvNets, were first introduced in the 1980s by Yann LeCun, a postdoctoral computer science researcher. LeCun had built on the work done by Kunihiko Fukushima, a Japanese scientist who, a few years earlier, had invented the recognition, a very basic image recognition neural network. The early version of CNNs, called LeNet (after LeCun), could recognize handwritten digits. CNN found a niche market in banking and postal services and banking, where they read zip codes on envelopes and digits on checks. But despite their ingenuity, ConvNets remained on the sidelines of computer vision and artificial intelligence because they faced a serious problem: They could not scale. CNNs needed a lot of data and computing resources to work efficiently for large images. At the time, the technique was only applicable to images with low resolutions. In 2012, Alex Net showed that perhaps the time had come to revisit deep learning, the branch of AI that uses multi-layered neural networks. The availability of large sets of data, namely the ImageNet dataset with millions of labeled pictures, and vast compute resources enabled researchers to create complex CNNs that could perform computer vision tasks that werepreviously impossible.

1.2 How Does Convolutional Neural Network Work

Convolutional neural networks are composed of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and output an activation value.

The behavior of each neuron is defined by its weight. When fed with the pixel values, the artificial neurons of a CNN pick out various visual features. When you input an image into a Conv Net, each of its layers generates several activation maps. Activation maps highlight the relevant features of the image. Each of the neurons takes a patch of pixels as input, multiplies their color values by its weights, sums them up, and runs them through the activation function. The first (or bottom) layer of the CNN usually detects basic features such as horizontal, vertical, and diagonal edges. The output of the first layer is fed as input of the next layer, which extracts more complex features, such as corners and combinations of edges. As you move deeper into the convolutional neural network, the layers start detecting higher-level features such as objects, faces, and more. As you move deeper into the convolutional neural network, the layers start detecting higher-level features such as objects, faces, and more.

The operation of multiplying pixel values by weights and summing them is called "convolution" (hence the name convolutional neural network). A CNN is usually composed of several convolution layers, but it also contains other components. The final layer of a CNN is a classification layer, which takes the output of the final convolution layer as input (remember, the higher convolution layers detect complex objects). Based on the activation map of the final convolution layer, the classification layer outputs a set of confidence scores (values between 0 and 1) that specify how likely the image is to belong to a "class." For instance, if you have a Conv Net that detects cats, dogs, and horses, the output of the final layeris the possibility that the input image contains any of those animals.

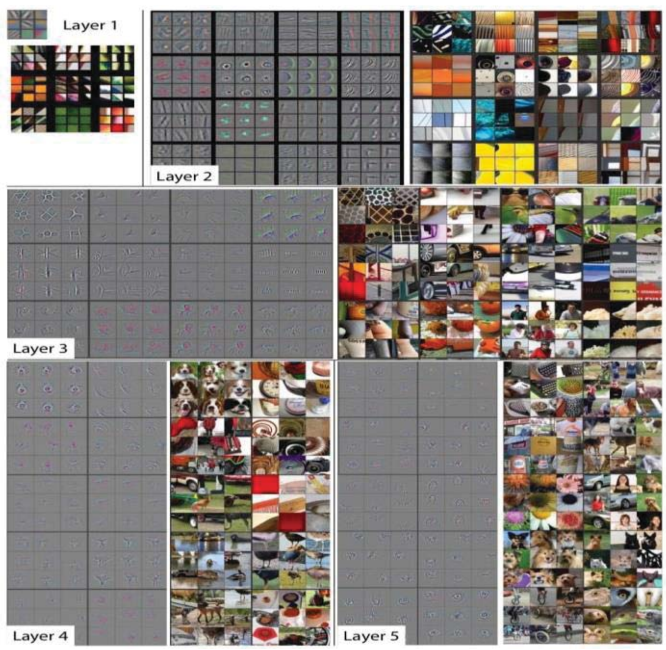


Fig. 2.

1.3 Training of a Convolutional Neural Network

In the beginning, CNN starts with random weights. During training, the developers provide the neural network with a large dataset of images annotated with their correspond- ing classes (cat, dog, horse, etc.). The ConvNet processes each image with its random values and then compares its output with the image's correct label. If the network's output does not match the label—which is likely the case at the beginning of the training process—it makes a small adjustment to the weights of its neurons so that the next time it sees the same image, its output will be a bit closer to the correct answer. The corrections are made through a technique called back-propagation (or backdrop).

Essentially, backpropagation optimizes the tuning process and makes it easier for the network to decide which units to adjust instead of making random corrections [1].

Every run of the entire training dataset is called an "epoch." The ConvNet goes through several epochs during training, adjusting its weights in small amounts. After each epoch, the neural network becomes a bit better at classifying the training images. As the CNN improves, the adjustments it makes to the weights become smaller and smaller. At some point, the network "converges," which means it essentially becomes as good as it can. After training the CNN, the developers use a test dataset to verify its accuracy. The test dataset is a set of labeled images that were not part of the training process. Each image is run through ConvNet, and the output is compared to the actual label of the image.

Essentially, the test dataset evaluates how good the neural network has become at classifying images it has not seen before. If a CNN score well on its training data but scores badly on the test data, it is said to have been "overfitted." This usually happens when there's not enough variety in the training data or when the ConvNet goes through too many epochs on the training dataset.
The success of convolutional neural networks is largely due to the availability of huge image datasets developed in the past decade. ImageNet, the contest mentioned at the beginning of this article, got its title from a namesake dataset with more than 14 million labeled images. There are other more specialized datasets, such as the MNIST, a database of 70,000 images of handwritten digits [2].

You don't, however, need to train every convolutional neural network on millions of images. In many cases, you can use a pre-trained model, such as the AlexNet or Microsoft's ResNet, and finetune it for another more specialized application. This process is called transfer learning, in which a trained neural network is retrained with a smaller set of new examples [3].

1.4 What is Bioinformatics

Bioinformatics is an interdisciplinary field that develops methods and software tools for understanding biological data, in particular when the data sets are large and complex. As an interdisciplinary field of science, bioinformatics combines biology, computer science, information engineering, mathematics, and statistics to analyze and interpret biological data. Bioinformatics has been used for in silico analyses of biological queries using mathematical and statistical techniques. [clari f icationneeded]

Bioinformatics includes biological studies that use computer programming as part of their methodology, as well as specific analysis "pipelines" that are repeatedly used, particularly in the field of genomics. Common uses of bioinformatics include the identification of candidate genes and single nucleotide polymorphisms (SNPs). Often, such identification is made with the aim of better understanding the genetic basis of disease, unique adaptations, desirable properties (esp. in agricultural species), or differences between populations. Less formally, bioinformatics also tries to understand the organizational principles within nucleic acid and protein sequences, called proteomics [4, 5].

Even many Technology and Computer Science Topics such as Artificial Neural Networks and Convolutional Neural Networks are used in Bioinformatics.

1.5 CNN-BASED CAD

CNN models are computational models that are composed of multiple processing layers to retrieve features from raw data with multilevel representations and hierarchical abstraction. As shown in the right of Figure 1, the general architecture of CNN models is made up of convolutional layers, full-connection layers, and pooling layers in addition to the input and output layers. Specifically, Figure 2 shows the architecture of VGG16 which consists of 13 convolutional layers, 3 full-connection layers, 5 pooling layers, and 1 softmax layer. For further improvement in object classification, many techniques can be embedded, including nonlinear filtering, data augmentation, local response normalization, hyperparameter optimization, and multiscale representation. At present, widely used deep learning models include but are not limited to, VGG, LeNet, AlexNet, GoogLeNet, ResNet, YOLO, faster R-CNN, and LSTM [6].

Mathematically, the procedure of using a pre-trained CNN- based CAD model for the prediction of lesion malignancy can be described as follows. Given a suspicious region (I_x), the output of a CNN-based model can be formalized as where n stands for the number of hidden layers and f_i denotes the activation function in the corresponding layer i. Furthermore, how to design the architecture of deep learning models in addition to the comprehensive analysis and systematic methodologies of learning representation can be referred from.

It should be noted that CNN models are data-driven and can be trained end-to-end. The models enable the integration of feature extraction, feature selection, and malignancy prediction into an optimization procedure. Therefore, these retrieved features are not designed by human engineers but learned from the input data. In general, the remarkable performance of CNN- based CAD models comes from advanced computing hardware resources (i.e., graphic processing units and distributed computing systems), open-source software, such as TensorFlow (https://www.tensorflow.org/), and open challenges based on millions of high-quality labeled images, such as ImageNet (http://www. image-net.org/). Its success also benefits from the novel designof architectures for deep learning, such as inception and identity mapping [8].

1.6 Data Set Training

The purpose of this layer is to receive a feature map. Usually, we start with a low number of filters for low-level feature detection. The deeper we go into the CNN, the more filters we use to detect high-level features. Feature detection is based on ‘scanning’ the input with the filter of a given size and applying matrix computations to derive a feature map.

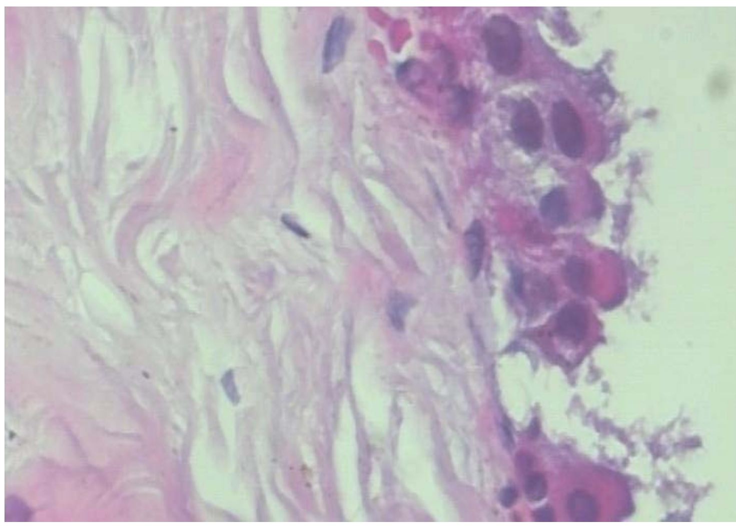


Fig. 3. Benign Tumor

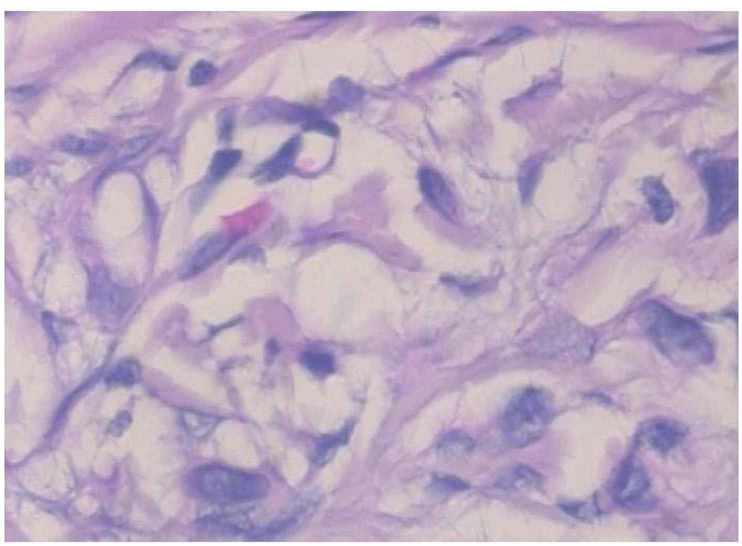


Fig. 4. Malignant Tumor

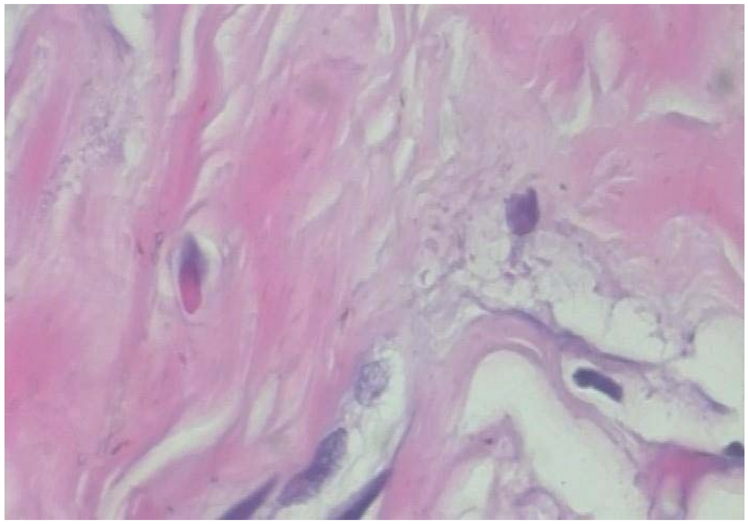


Fig. 5. Transformation of Benign to Malignant Tumor

1.7 Training of The Dataset

Split the data set into two sets — train and test sets with 80% and 20% images respectively. Let's see some sample benign and malignant images.

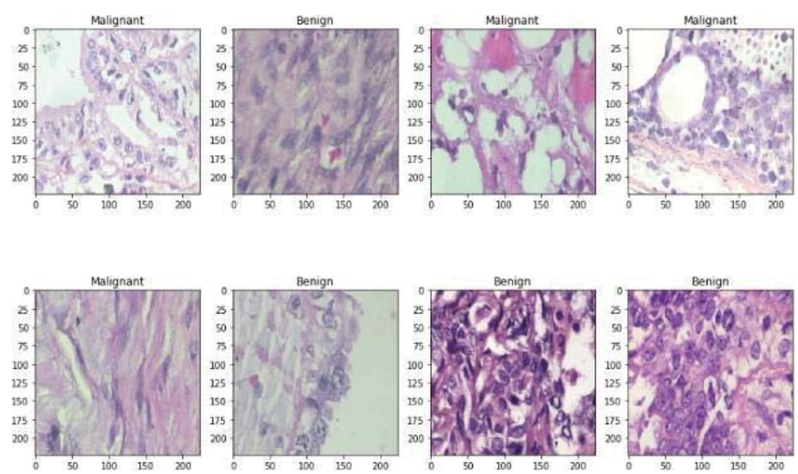


Fig. 6.

I used a batch size value of 16. Batch size is one of the most important hyperparameters to tune in deep learning. I prefer to use a larger batch size to train my models as it allows computational speedups from the parallelism of GPUs. However, it is well known that too large of a batch size will lead to poor generalization. On the one extreme, using a batch equal to the entire dataset guarantees convergence to the global optima of the objective function. However, this is at the cost of slower convergence to that optima. On the other hand, using smaller batch sizes has been shown to have faster convergence to good results. This is intuitively explained by the fact that smaller batch sizes allow the model to start learning before having to see all the data. The downside of using a smaller batch size is that the model is not guaranteed to converge to the global optima. Therefore, it is often advised that one starts at a small batch size reaping the benefits of faster training dynamics and steadily growing the batch size through training.

I also did some data augmentation. The practice of data augmentation is an effective way to increase the size of the training set. Augmenting the training examples allow the network to see more diversified, but still representative data points during training.

The next step was to build the model. This can be describedin the following 3 steps:

I used DenseNet201 as the pre-trained weights which are already trained in the Imagenet competition. The learning rate was chosen to be 0.0001. On top of it, I used a global aver- pooling layer followed by 50% dropouts to reduce over-fitting. I used batch normalization and a dense layer with 2 neurons for 2 output classes ie benign and malignant with softmax as the activation function. I have used Adam as the optimizer and binary-cross-entropy as the loss function.

Before training the model, it is useful to define one or more callbacks. Pretty handy one, are ModelCheckpoint and ReduceLROnPlateau.

ModelCheckpoint: When training requires a lot of timeto achieve a good result, often many iterations are required.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
densenet201 (Model)          (None, 7, 7, 1920)        18321984
_________________________________________________________________
global_average_pooling2d_1 ( (None, 1920)              0
_________________________________________________________________
dropout_1 (Dropout)          (None, 1920)              0
_________________________________________________________________
batch_normalization_1 (Batch (None, 1920)              7680
_________________________________________________________________
dense_1 (Dense)              (None, 2)                 3842
=================================================================
Total params: 18,333,506
Trainable params: 18,100,610
Non-trainable params: 232,896
_________________________________________________________________
```

Fig. 7.

In this case, it is better to save a copy of the best perform- ing model only when an epoch that improves the metrics ends.**ReduceLROnPlateau**: Reduce the learning rate when a metric has stopped improving. Models often benefit from reducing the learning rate by a factor of 2–10 once learning stagnates. This callback monitors a quantity and if no improvement is seen for a 'patience' number of epochs, the learning rate is reduced.

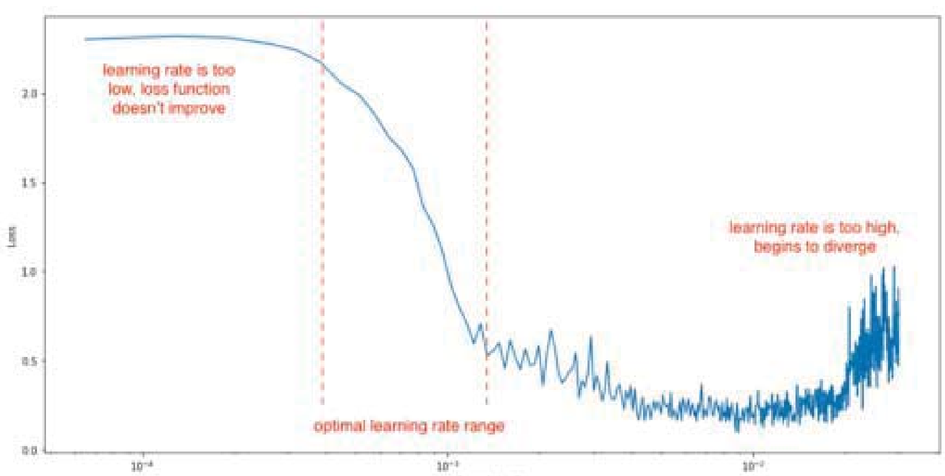


Fig. 8

1.8 Performance Metrics

The most common metric for evaluating model performance is accuracy. However, when only 2% of your dataset is of one class (malignant) and 98% of some other class (benign), misclassification scores don't make sense. You can be 98% accurate and still catch none of the malignant cases which could make a terrible classifier. However, when only 2% of your dataset is of one class (malignant) and 98% of some other class (benign), misclassification scores don't make sense. You can be 98% accurate and still catch none of the malignant cases which could make a terrible classifier.

1.9 Precision, Recall, and f1-score

For a better look at misclassification, we often use the following metric to get a better idea of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN).

Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

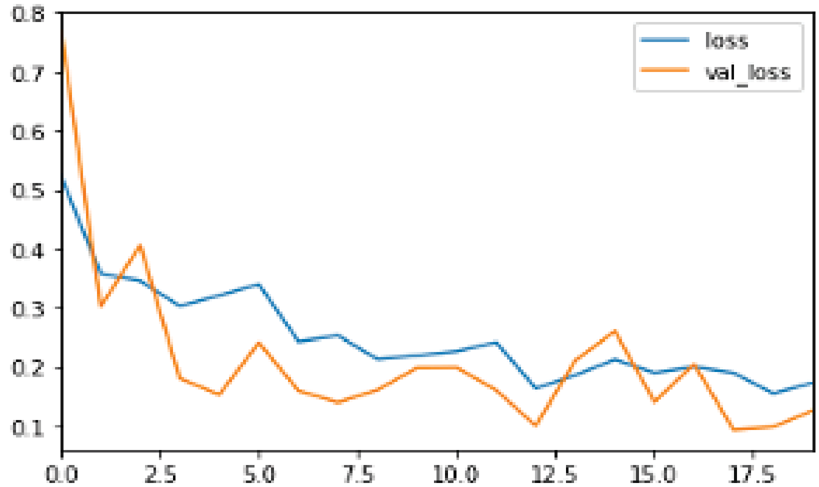


Fig. 9.

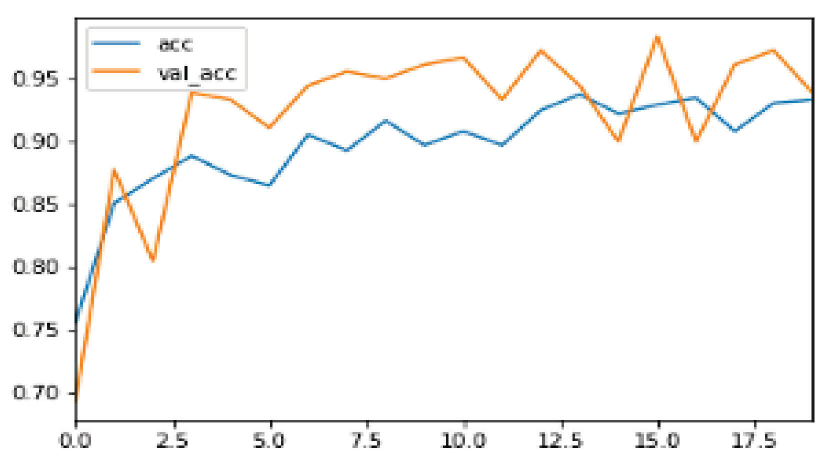


Fig. 10.

The recall is the ratio of correctly predicted positive observations to all the observations in the actual class. **F1-Score** is the weighted average of Precision and Recall.

$$F1 = \frac{2 * (Recall * Precision)}{(Recall + Precision)}$$

Fig. 11.

The higher the F1-Score, the better the model. For all threemetrics, 0 is the worst while 1 is the best.

1.5 *Confusion Matrix*

The Confusion Matrix is a very important metric when analyzing misclassification. Each row of the matrix represents the instances in a predicted class while each column represents the instances in an actual class. The diagonals represent the classes that have been correctly classified. This helps as we not only know which classes are being misclassified but also what they are being misclassified as.

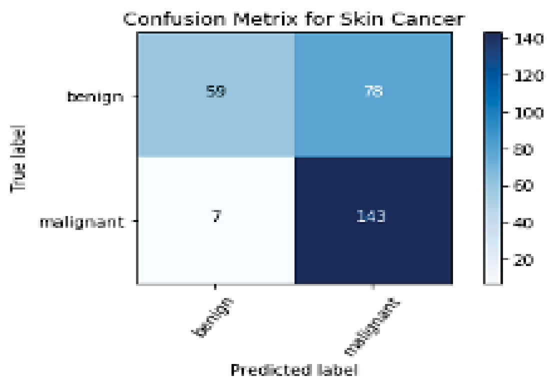


Fig. 12.

2. Rock Curves

The 45-degree line is the random line, where the Area Under the Curve or AUC is 0.5. The further the curve from this line, the higher the AUC and the better the model. The highest a model can get is an AUC of 1, where the curve forms a right-angled triangle. The ROC curve can also help debug a model. For example, if the bottom left corner of the curve is closer to the random line, it implies that the model is misclassifying at Y=0. Whereas, if it is random on the top right, it implies the errors are occurring at Y=1.

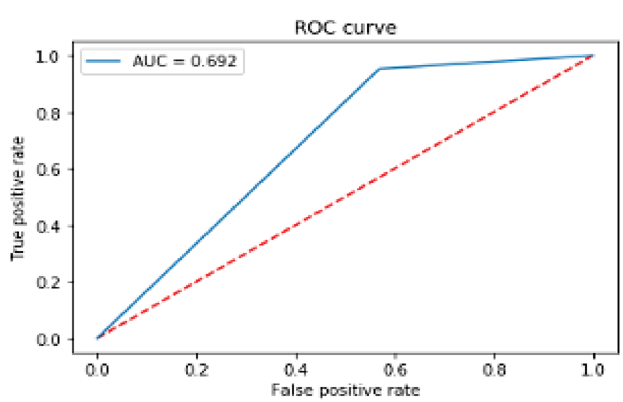


Fig. 13.

3. Results

Accuracy	Precision	Recall	F1 score	ROC-AUC
98.3%	0.65	0.95	0.77	0.692

Fig. 14.

4. Conclusions

By this Paper, we can once again conclude that Breast Cancer is a very vital life-threatening situation and with the help of Artificial Intelligence and tools like MATLAB and CNN, we can make programs to detect Lung Cancer at an early stage and also save many lives.

References

1. Amos, C.I., Xu, W., and Spitz, M.R. (1999). Is there a genetic basis for lung cancer susceptibility? Cancer Res. 151, 3–12.
2. Belinsky, S.A., Nikula, K.J., Palmisano, W.A., Michels, R., Saccomanno, G., Gabrielson, E., Baylin, S.B., and Herman, J.G. (1998). Proc. Natl. Acad. Sci. USA, 95, 11891–11896.
3. Bunn, P.A., Jr. (2001). Triplet combination chemotherapy and targeted therapy regimens. Oncology (Huntingt.) 15, 26–32.
4. Burbee, D., Forgacs, E., Zöchbauer-Müller, S., Shivakuma, L., Fong, K., Gao, B., Randle, D., Virmani, A., Bader, S., Sekido, Y., et al. (2001). RASSF1A in the 3p21.3 homozygous deletion region: epigenetic inactivation in lung and breast cancer and suppression of the malignant phenotype. J. Natl. Cancer Inst. 93, 691–699.
5. Dammann, R., Li, C., Yoon, J.H., Chin, P.L., Bates, S., and Pfeifer, G.P. (2000). Nat. Genet. 25, 315–319.
6. Fiore, M.C., Jorenby, D.E., and Baker, T.B. (1997). Smoking cessation: principles and practice based upon the AHCPR Guideline, 1996. Agency for Health Care Policy and Research. Ann. Behav. Med. 19, 213–219.
7. Fong, K.M., Sekido, Y., and Minna, J.D. (2001). The molecular basis of lung carcinogenesis. In The Molecular Basis of Human Cancer, W. B. Coleman and G. Tsongalis, eds. (Totowa, NJ: Humana Press), pp. 379–405.
8. G. Wu, M. Kim, Q. Wang, Y. Gao, S. Liao, and D. Shen, "Unsupervised deep feature learning for deformable registration of mr brain images.," Medical Image Computing and Computer-Assisted Intervention, vol. 16, no. Pt 2, pp. 649–656, 2013.
9. Y. Xu, T. Mo, Q. Feng, P. Zhong, M. Lai, and E. I. Chang, "Deep learning of feature representation with multiple instance learning for medical image analysis," in IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP, pp. 1626–1630, 2014.
10. D. Kumar, A. Wong, and D. A. Clausi, "Lung nodule classification using deep features in ct images," in 2015 12th Conference on Computer and Robot Vision, pp. 133–138, June 2015.
11. Y. Bar, I. Diamant, L. Wolf, S. Lieberman, E. Konen, and H. Greenspan, "Chest pathology detection using deep learning with non-medical training," Proceedings - International Symposium on Biomedical Imaging, vol. 2015-July, pp. 294–297, 2015.
12. W. Sun, B. Zheng, and W. Qian, "Computer-aided lung cancer diagnosis with deep learning algorithms," in SPIE Medical Imaging, vol. 9785, pp. 97850Z–97850Z, International Society for Optics and Photonics, 2016.
13. J. Tan, Y. Huo, Z. Liang, and L. Li, "A comparison study on the effect of false positive reduction in deep learning-based detection for juxta pleural lung nodules: Cnn vs den," in Proceedings of the Symposium on Modeling and Simulation in Medicine, MSM '17, (San Diego, CA, USA), pp. 8:1–8:8, Society for Computer Simulation International, 2017.
14. R. Golan, C. Jacob, and J. Denzinger, "Lung nodule detection in ct images using deep convolutional neural networks," in 2016 International Joint Conference on Neural Networks (IJCNN), pp. 243–250, July 2016.
15. Kaggle, "Data science bowl 2017." https://www.kaggle.com/c/data science-bowl-2017/data, 2017.
16. LUNA16, "Lung nodule analysis 2016." https://luna16.grand challenge.org/, 2017.

17. M. Firmino, A. Morais, R. Mendoza, M. Dantas, H. Hekis, and R. Valentim, “Computer-aided detection system for lung cancer in computed tomography scans: Review and prospects,” BioMedical Engineering OnLine, vol. 13, p. 41, 2014.
18. S. Hawkins, H. Wang, Y. Liu, A. Garcia, O. Stringfield, H. Krewer, Q. Li, D. Cherezov, R. A. Gatenby, Y. Balagurunathan, D. Goldgof, M. B. Schabath, L. Hall, and R. J. Gillies, “Predicting malignant nodules from screening ct scans,” Journal of Thoracic Oncology, vol. 11, no. 12, pp. 2120–2128, 2016.
19. M. S. AL-TARAWNEH, “Lung cancer detection using image processing techniques,” Leonardo Electronic Journal of Practices and Technologies, pp. 147–158, June 2012.
20. O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” CoRR, vol. abs/1505.04597, 2015.
21. M. D. Zeiler, M. Ranzato, R. Monga, M. Mao, K. Yang, Q. V. Le, P. Nguyen, A. Senior, V. Vanhoucke, J. Dean, and G. E. Hinton, “On rectified linear units for speech processing,” in IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3517– 3521, May 2013.

.

Lightweight Cryptography in IoT: Impact Survey, Challenges and Solution

Amita V Shah[a], Dr. Sanjay M Shah[b], Dr. Dhaval A Parikh[c], Namit V Shah[d]

[a]*Ph.D Scholar, Computer/IT Engineering, Gujarat Technological University, Gujarat, India*
[b]*Professor & Head, Computer Engineering Dept., Government Engineering College, Rajkot, Gujarat, India*
[c]*Professor & Head, Computer Engineering Dept., Government Engineering College, Gandhinagar, Gujarat, India*
[d]*Student, Computer Engineering Dept., L D College of Engineering, Ahmedabad, Gujarat, India*

Abstract

The purpose of this survey is to investigate the cryptographic system for low-power Internet of Things (IoT) devices. The Internet of Things is a network in which people cooperate with nature using sensors and actuators. IoT enables remote detection and control of objects over the internet. The growing importance of IoT in our daily lives, such as in healthcare and home automation, presents new challenges surrounding the management of personal and sensitive data and the protection of the data. It is considered that the majority of issues with IoT devices include data security and data integrity. This survey will focus on a variety of security techniques, including ECC, RSA, AES, BCC, and CP-ABE. Parameters such as time, data size, memory space, battery power, and processing power should be considered in the improvement of the IoT systems' security. A high number of resource-constrained devices necessitates that IoT security solutions be extremely lightweight. Based on their key size, block size, number of rounds, and structures, we examined a wide range of lightweight cryptographic methods. We also go through IoT security architecture for confined device environments, with a particular emphasis on research problems, challenges, and solutions. Finally, a hybrid security approach has been suggested for improving security in resource-constrained IoT environments and several remaining issues are highlighted.

Keywords: IoT Devices, Security, Lightweight Cryptography.

1. Introduction

Everything we use in our daily lives can be connected thanks to the Internet of Things. It is similar to every non-exclusive thing, communicating and demonstrating in the same ways. For instance, IoT will wake you up in the morning via an alarm, by making reference to a schedule, and by turning off lights, fans, air conditioning, and other appliances when you leave the house and in accordance with the outside environment. It shows you the most likely and efficient path to your office as soon as you get into your car. This is how a fully automated life is possible in an Internet of Things environment [1].

The Internet of Things encompasses every connection between different kinds of devices; therefore we have machine-to-machine (M2M) systems in which each machine communicates data and information in real-time with other machines. The IoT connects many physical items that interact without human involvement and security is a major concern whenever people or devices are communicating or interacting via a network. System security prevents unwanted access, data misuse, monitoring, and more. Applications of IoT like healthcare, smart grids, smart homes, and smart parking systems are already part of our daily lives and benefit everyone but IoT systems cannot protect users data from various threats [2]. Security is the major concern while communicating in such a heterogeneous environment. A form of a traditional cryptographic method known as lightweight encryption is useful in Internet of Things (IoT) devices with limited resources.

Cryptography is the process of changing plaintext into unintelligible text and vice versa. In addition to safeguarding data from attack or alteration, cryptography may be used to authenticate users and for the privacy of data. Lightweight Cryptography is a subfield of cryptography that focuses on the creation of algorithms for use in devices with insufficient resources (memory, power, or size) to perform the operation. For smart devices with limited resources, the traditional cryptographic primitives might not be appropriate. For instance, RFID tags cannot use the 1204-bit RSA algorithm [3]. This paper describes related work for lightweight methods used for secure data transfer.

Two primary justifications for implementing new technologies for IoT are: *1) End-to-end communication effectiveness* the lightweight symmetry key technique will provide end-to-end security while consuming less power in low-resource devices and *2) Low-resource smart device Adoptability* The footprint of lightweight cryptography is substantially smaller than that of traditional cryptography. It offers the potential for additional network connections with smart devices with fewer resource requirements.

NIST defines lightweight cryptography as a subcategory of cryptography with the goal of offering solutions for quickly expanding

**Amita V Shah*
Email address: amitashah@ldce.ac.in

applications that frequently make use of smart, low-power devices [4]. It is intended for a wide range of hardware and software based devices. A standard encryption algorithm may run smoothly on PCs, servers, and some mobile devices. Yet, the lower ends of the spectrum are made up of things like embedded systems, sensing devices, and RFID tags so these devices and networks demand lightweight cryptography systems.

In Cryptography, both symmetric and asymmetric cryptography algorithms exist. Symmetric algorithms communicate using a single private key. The communication key between the sender and the recipient is shared. Symmetric keys guarantee data confidentiality and integrity but do not guarantee authentication. Symmetric encryption has the advantage of requiring fewer keys with smaller key sizes. In this paper, we have analyzed different symmetric and asymmetric algorithms of cryptography.

This paper's main goal is to present an extensive overview of IoT security issues and potential solutions. The relevant work on lightweight IoT security techniques has been discussed. All of the IoT-related challenges are covered. The review then emphasizes the most significant IoT problem and the suggested solution.

2. Challenges and countermeasures for IoT

There are several issues with IoT. We categorize them according to numerous IoT-related technological factors.

2.1 Technology for Sensing and Communication

There are numerous IoT standards, but there isn't a single framework attempt that allows for their incorporation. RFID operates through EPCglobal. Even though Zigbee cannot be regarded as a standard technology, 6LoWPAN coordinates low power IEEE 802.15.4 devices in our network. Bringing together all intriguing developments on one stage is not possible. We have seen that IoT devices can be identified using NFC, WSN, RSN, or RFID. Therefore, there are different working and hailing structures for each situation. Interoperability is a problem for both specificity and semantics [5].

Since RFID relies on radio communication, which is typically monotonous, the data discovered with this technology are typically raucous. Therefore, a data cleaning technique should be powerful and sufficiently advantageous to remove important and lossless data. For RFID data, there are a number of data cleaning and isolating techniques available, including various sliding window methods and sequential deduction calculations.

There are various factors that affect how well a framework is being used, such as the arrange factor, portability rate, heterogeneous device count, and enormous number of occasions. As the number of hubs in the IoT network increases, more sophisticated tools are needed to manage the network's coherence.

2.2 Challenges in Communication Protocol

Although Zigbee uses a mesh topology to solve the problem of short-distance communication, nodes are required to manage the routing table, which raises the expense of routing. In addition, the nodes are unable to create super frame, making sleep mode impossible. In order to find the path for additional nodes, they must be in the open state.

Owing to the characteristics of Zigbee, the benefit of long battery life in sensors is rendered null and void. Particular gate necessity: When connecting a Zigbee network to a standard IP network, a particular gateway is required for conversion. There are ongoing efforts to integrate IP with Zigbee. Since they are too complex and hence inappropriate for small, low-power devices, conventional security encryption approaches can't be applied directly to 6LowPAN devices [6]. Due to the high rate of device movement, routing information quickly becomes outdated, resulting in connection failure. Thus, processing the routing information of a highly dynamic network imposes expenses.

2.3 Challenges in Middleware

Standardization: IoT applications vary; hence there is no common middleware. RFID/WSN middleware does not provide security, privacy, or hardware abstraction, according to the research. Alternative methods may lack SOA's interoperability, scalability, hardware abstraction, dynamic infrastructure, security, and privacy. Each Internet of Things device will contain hardware, data retrieval and storage apps, and communication technologies like wifi, zigbee, etc. Middleware abstraction of heterogeneous devices is tough.

Scalability: When an IoT network expands, interaction events increase, causing traffic, data processing, and flooding.

Middleware should handle changing infrastructure and large data. Zero infrastructures: Middleware should help the Internet of Things' zero infrastructure by letting nodes broadcast their capabilities [6].

The Internet of Things connects diverse devices. Each device handles and formats data differently. Hence, middleware must handle varied data types.

2.4 Challenges in QoS

Included in resource restrictions are energy, processing, and storage capacities. The volume of data collected necessitates an increase in processing capacity to manage them. Reliable transmission needs a high-performance algorithm and signaling protocol, which uses additional energy and requires computing power. When the sensors are dispersed throughout the IoT network, it may not always be possible to recharge or replace their batteries. Because of their dynamic nature, IoT networks require high-performance interfaces and routing algorithms. That makes it more difficult to provide QoS support. Expanding an IoT network causes a rise in network traffic, which affects the network's reliability, consistency, and transmission time quality of service characteristics [6].

Diverse traffic: Communication between diverse sensing devices and the integration of data from these sensors might compromise the Internet of Things' Quality of Service criteria for precision.

2.5 Challenges in Security and Privacy

Data protection is impractical for low-powered IoT sensing devices. The sensor devices are also unobserved, making outside meddling and eavesdropping possible.
Authentication, verification, data integrity, and network attacks including spoofing, DDoS, jamming, and shielding are security issues. IoT apps help users, but hackers, attacks, and flaws can compromise personal data [7].
Cryptography secures data. Standard cryptography requires significant resources. IoT devices have limited processor, memory, power supply, and battery life [8]. Ciphers are evaluated based on security. Whether a cipher can withstand several attacks determines its security.

As the devices/person may be readily followed by the sensing devices, privacy and confidentiality cannot be maintained for the person/object. Before tracking devices, a valid policy must be established. Yet, it is difficult to govern the sensors in terms of what data should be gathered and by whom.

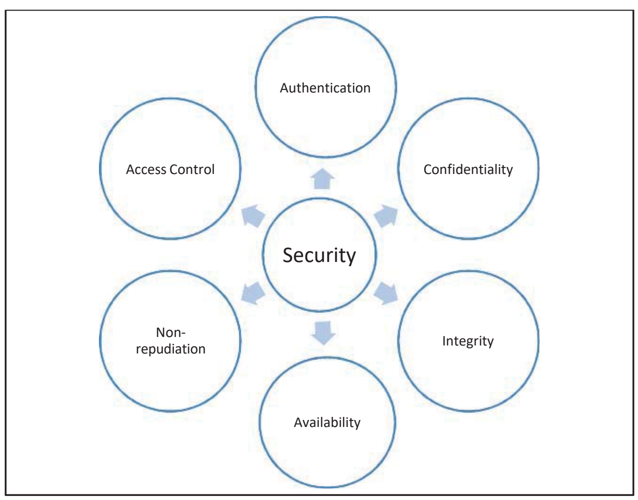


Fig. 1. Primary Goal of Security

2.5.1 Primary Security Goals:

Confidentiality: This security service guarantees that only authorized users have access to the information's content. Hence, it prohibits an unauthorized user from accessing the protected information's content. It is occasionally called secrecy.

Data Integrity: It can detect if an unauthorized party has changed data since an authorized user produced, stored, or transferred it.

Insert, delete, and replace data manipulation.

Authentication: Data origin and entity authentication are strongly related to identification. Data origin authentication checks message origination. Data integrity is implicitly provided through data origin authentication. Occasionally, it is called message authentication. Authentication of entities guarantees the authenticity of a second entity with which a first entity interacts. Often, entity authentication is synonymous with data origin authentication.

Non-Repudiation: This security service prohibits an organization from rejecting a prior action or promise. It is especially beneficial in circumstances that might potentially escalate to conflict. When a disagreement emerges, a reliable third party can give the necessary evidence to resolve it.

3. Research Motivation

IoT enables in establishing connections between different facets found in varied environments. Because of its openness and the scarcity of human participation, the Internet of Things (IoT) is vulnerable to a variety of attacks, including man-in-the-middle attacks and denial-of-service (DoS) attacks. In addition, any device can connect to the network, allowing for illegal access. These threats have the potential to physically harm devices as well as network connections. This will ultimately jeopardize the security and privacy of IoT. Considering IoT have limited resources (low power, bandwidth, and storage), an effective security solution is needed that won't hog up IoT resources.

4. Existing Research & Solution in IoT: Related Work

Duy-Hieu Bui Edith Beigne et al. in [9] Developed several AES equipment enhancement solutions for high-speed, ultralow-power, ultralow-energy Internet of Things (IoT) applications with variable security based on key sizes, power and energy usage for data flow and key expansion, and low power consumption.
In [10], Prachi Surbhi et al. compared methods that manage communication between the server and limited IoT devices. It is necessary to select protocols with characteristics such as low packet loss, high packet generation time, and low packet response time, among others.

In Darshan Pritam et al. in [11] proposed Circular Bend Web of Things (ECIOT) protocol, the IoT nodes would produce a secret key for each session with the server using ECDH and EX-OR, however it does not employ a reliable mechanism or algorithm for Message Encryption. The Diffie–Hellman key agreement and DSA algorithms, which provide strong authentication but are only used for short communications, have supplanted the Elliptic Curve Diffie–Hellman (ECDH) and ECDSA.
According to Lara Carlos et al. in [12], ECC-163 will give the same degree of security as RSA-1024, allowing it to be deployed inenvironments with limited resources.

Sharafi Masoumeh et al. in [13] showed that MBCC improves BCC's energy and security. Chaos theory underpins energy-constrained wireless sensor devices. MBCC encrypts 32-bit data in 13.44% less memory and 6.4 and 6.6 times less time and energy than BCC. Installed MBCC, BCC, RC5, and Skipjack in SunSPOT sensors and analyzed performance. MBCC performed best in security and resource utilization.

In [14], Saurabh Singh et al. described a scheme that may be applied in the context of a smart home and investigated outstanding problems with cipher construction, implementation, block size, key size, unique threats, and security metrics.
Wenfeng Yajun Ha Massimo Alioto et al. in [15] offered the structure analysis of lightweight AES accelerators with the objective of restricting the force utilization and demonstrated a minimal exertion ultra-energy-capable AES encryption location for cubic-millimeter stages for IoT.

Xuanxia Yaoa Zhi Chena et al. in [16] proposed ABE scheme for IoT. It is a lightweight no-pairing ECC-Based ABE scheme which improves execution efficiency with low communication costs but it has some disadvantages like poor flexibility in revoking attribute and poor scalability.

Salvador Perez et al. in [17] Presents an encryption method based on the lightness of the symmetric cryptography and expressiveness of attribute-based encryption. Just approved administrations will actually access specific pieces of data, so that client's protection isn't undermined. It protect lots of sensitive data of particular building based on IoT.

Because of IoT devices' resource limitations, many IoT devices cannot use basic CP-ABE Scheme. ABE uses number of exponentiations and pairing operations which makes the entire scheme very complex. We can also migrate complex arithmetic operations to hardware accelerators which can enhance energy efficiency and also reduce execution time[18,19].

Oualha Nouha Kim et al. in [20] applied Pre-count framework to the CP-ABE which diminishes the cost of computations for *encryption however require additional memory for storage.*

Moffat, Steve, Mohammad Hammoudeh, and Robert Hegarty et al. in [21] indicates that the recommendations should also examine architecture choices such proxy devices or helper nodes to handle heavy computing and operate as security gateways to limit or prevent IoT device-internet connectivity.

5. Three Layer Architecture of IoT and Related Protocols

Application Layer: Delivering application-specific services to the user is the responsibility of the application layer. At this layer, the real implementation of IoT intelligence is understood. It can be utilized for a variety of purposes, including retail, social interaction, health, and personal use. For the constrained IoT devices, this layer supports the Constrained Application Protocol (CoAP)[24]. CoAP previously used IPSec and DTLS security. The various threats can compromise the present security systems. Hence, cryptography techniques can be added into them [22].

Network Layer: Connecting to servers, network devices, and other smart objects is the responsibility of the network layer. Network layer receives data from physical layer. Using the IPv6 addressing technique, the network layer is utilized to split the message into packets and route the packets from source to destination. As the number of Internet-connected devices grows, IPv4 address space is being replaced with IPv6, which has larger address space [23]. IPSec at the network layer can be used to implement built-in cryptography protocols like AES and DES.

Perception Layer: The physical layer, which has sensors for detecting and receiving information about the surroundings, is the perception layer. It is the Internet of Things' bottom layer, which combines the physical and MAC layers. It is used to gather data utilizing RFID, sensors, or GPRS. In this layer of the Internet of Things, IEEE 802.15.4 is utilized as the standard specification. IEEE 802.15.4 is effective for inexpensive, battery-powered devices [23]. At this layer, a security solution based on IEEE 802.15.4 is available; however it is still open to intrusion. Table 1 provides a summary of each layer's current protocol, security protocol, and attacks.

Table 1. Protocol used at each Layer

Layer	Protocol Used	Security Protocols	Attacks
Application	COAP	Not fixed designed by user	Depend on Protocol
Network	IPv6, RPL	IPSec	DoS Attack, data privacy
Physical	IEEE 802.15.4, PHY, MAC	IEEE 802.15.4 Security	DoS, Attack on authentication, integrity

6. Symmetric Lightweight Algorithms for IoT Devices

PRESENT: The block size of PRESENT is 64 bits, while the key size might be 80 bits or 128 bits. A single 4-bit S-box that was created with hardware optimizations in mind serves as the foundation for the non-linear layer. When low power consumption and excellent chip efficiency are required, PRESENT should be employed. PRESENT was incorporated into the new international standard for light-weight cryptographic techniques by the International Organization for Standardization and the International Electro technical Commission [24]. In 2014, a reduced differential attack on 26 of the 31 rounds of PRESENT was proposed. PRESENT is not software-friendly and employs bit-oriented permutations.

AES (Advance Encryption Standard): At the application layer of COAP, AES can be employed as an integrated solution. It isa symmetric block cipher algorithm that has been approved by NIST. It employs a substitution permutation network and operates on a 4x4 matrix with a 128 bit block length. Subbytes, shiftrows, MixedColumns, and AddRoundKey, all the components have an impact on every byte of the block [25]. The available key sizes are 128, 192, and 256 bits.

BCC (Block Cipher based on Chaos): Chaotic ciphers can produce chaotic points to accommodate different plaintext lengths and are useful for encrypting enormous amounts of information. When it comes to textual material, traditional encryption techniques are typically used rather than chaos-based alternatives.

TWINE: The 64-bit TWINE block cipher, which may accommodate 80 and 128-bit keys. Similar to the other approaches, it permits fairly modest hardware implementations while yet allowing for effective embedded software implementations. Moreover, it enables a tiny implementation of combined encryption and decryption. The usage of generalized Feistel with several subblocks and a recent advancement in the diffusion layer are the main causes of this phenomenon [26].
TEA (Tiny Encryption Algorithm): TEA is useful in confined surroundings, such as sensor networks or smart devices. It only a

few lines of code. It makes use of a straightforward programme that only needs to do the operations XOR, adding, and shifting.It does not use any predefined calculations or existing tables; instead, it employs 128 bit keys and a block size of 64 bits. ManyTEA variants exist, including extended TEA19, Block TEA, and others. These extensions attempt to fix issues with the original TEA, such as similar keys. TEA and its variants are still vulnerable to a variety of assaults because of how simply theyoperate [26].

RC5: Rivest was the first to use it for rotations that are not dependent on data. It has a Feistel structure and can function effectively as a lightweight algorithm when employed in wireless sensor applications. RC5 is regarded as w/r/b, where w stands for word size, r for number of working rounds, and b for number of bytes in encryption key. RC5 typically operates on a 32-bit size, although it also has 16, 32, and 64-bit variations. With 0, 1,..., 255 key bytes, it can function for 0 to 255 rounds. The standard key size is 16 bytes with 20 rounds of operation. Differential attacks are possible to RC5 [27].

HIGHT: For the Feistel network, Hight utilizes extremely simple operations like addition mod 28 or XOR. It operates on 128- bit keys with a block size of 64 bits and 32 rounds. During the encryption and decryption phases, its keys are generated. In 2017, a parallel implementation of height was developed that uses less power, is specified in less lines of code, and increases speed for RFID systems. Saturation attacks can harm Hight encrypted data [28].

Based on a review of the literature, comparisons of the symmetric lightweight algorithms listed above are made in Table 2 which are based on the table's descriptions of the code's structure, number of rounds, key size, block size, and attacks.

Table 2. Comparison of symmetric lightweight cryptography algorithms in IoT

Symmetric Algorithm	Block Length(bits)	Key Length (bits)	No of Rounds	Structure	Advantages	Disadvantages
AES	128	128/192/256	10/12/14	SPN	Appropriate performance in energy and processing time, no proven attack found	Large code and data memory size
PRESENT	64	80/128	31	SPN	RAM/ROM performance and code size-related memory consumption	Vulnerable to discrepancy attack, Key Recovery attack
TWINE	64	80/128	23/24/36	Feistel	Better code size, processing speed, and implementation complexity.	Meet-in-the-middle attack, discrepancy fault analysis.
TEA	64	80	32	Feistel	Practice the features of disorder philosophy High Security(as it create confusion and diffusion, great avalanche effect)	High RAM, ROM, energy, time, and key length.
RC5	32	16	20	Feistel	Appropriate performance in code memory and data memory	Vulnerable to discrepancy attack
BCC	Unlimited	128 or greater	2	Simultaneous confusion-	Apply stochastic theory. High-	High RAM, ROM, energy, time, and key
				diffusion operatio n	security (as it create confusion and diffusion, great avalanche effect)	length. (tolerable if security is higher importance than performance)
HIGHT	64	128	32	Feistel	Appropriate performance w.r.t. code size	Vulnerable to related key attack

7. Asymmetric Lightweight Algorithms for IoT Devices

RSA: Ron Rivest, Adi Shamir, and Leonard Adleman created it in 1978. RSA generates a public and private key combination by choosing two significant prime numbers. Calculate the decryption key by determining their modulus and randomly selecting the encryption key. The private key is made secure while the public key is made public. To protect user privacy, a more secure version of RSA encryption is highly recommended used to encrypt and decrypt information [29].

ECC: Compared to RSA, it requires a smaller key size. As a result, it requires less storage and processes data quickly. It is based on an algebraic system and uses two points on an elliptic curve as inputs. A discrete logarithm problem is used to create the key that is utilized to compute the key. a safe hardware implementation of ECC is suggested for small messages, which will speed up real-time computations. By improving its complex multiplication operation, ECC is optimized for 6LoWPAN nodes. In order to maximize the utilization of low-power devices, bit shifting is also used instead of multiplication in the microprocessors [30].

ECDH: Elliptic Curve Diffie Hellman (EC-DH) Algorithm has qualities that are appropriate for Internet of Things (IoT) devices, such as low power consumption, small weight, and robustness so the EC-DH Algorithm is gaining significance. One of the cryptosystems for exchanging keys is the Diffie-Hellman scheme. Since messages are excluded from this scheme, we take advantage of it by using the key being exchanged as a secret key; we generate the key using elliptic curve cryptography and Diffie-Hellman. Our findings lead us to the conclusion that in terms of power and area, ECDH is better to other algorithms under consideration [31].

Table 3 lists the similar bit count for each algorithm needed to achieve the same level of security for each algorithm; this is the bit size for which the key can be reverse computed in the same amount of time. Here, AES has been taken as a base algorithm. Table 4 shows the difference and usage of Symmetric and Asymmetric Algorithms.

Table 3. Comparison of Level of Security for the Asymmetric lightweight cryptography algorithms

Key Size (No of Bits)	AES Key Size	RSA Key Size	ECC Key Size	ECDH Key Size
5	5	16	10	10
27	27	128	64	54
57	57	512	128	114
80	80	1024	192	160
110	110	2048	256	220

Table 4. Comparison of Symmetric and Asymmetric Methods

Type	Asymmetric Key Cryptography	Symmetric Key Cryptography
Keys	Uses Public Key for Encrypting and Private Key for Decrypting	Permits Encrypting and Decrypting of the message with same key.
Speed	Slow in Execution	Fast in Execution
Use	Used for securely exchanging security key	Used for Bulk Data Transmission
Number of Keys	Depends on the number of users connected (Linear Proportion)	Depends on the number of users connected (Exponential Proportion)
Implementation	Complex Hardware Implementation	Simple in Comparison
Examples	RSA,DSA,ECC	AES, DES, Blowfish, Twofish

8. CP-ABE in IoT

To protect the IoT, it is crucial to establish a data access control solution that provides the necessary flexibility to handle a large number of IoT devices. Those needs are met by ciphertext policy attribute-based encryption. Because ABE is not energy efficient, many IoT devices cannot be included in CP-ABE, however, several CP-ABE variants can show how applicable CP-ABE is to the IoT [32].

It provides flexible and fine-grained access control. It removes the problem of encrypting data for a subset of users. In such a system, the decryption of a ciphertext is possible only if the set of attributes of the user key matches the attributes of the ciphertext. An access tree is generated which describes the attributes of a user required to decrypt a cipher text. For example : A message may be encrypted for users who are “PhD” “student” of “GTU” enrolled in “2017”. User must possess attributes: {PhD, student, GTU , 2017} to decrypt the target message.The implementation of CP-ABE is based on four algorithms: (i) Set-up (ii) Key Generation, (iii) Encryption, (iv) Decryption (v) Global set-up [33].

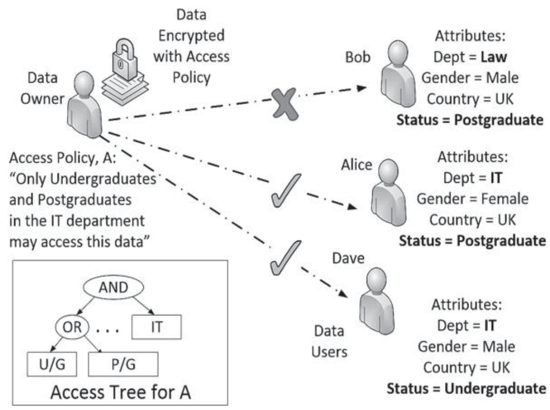


Fig. 2. A functional overview of CP-ABE

9. Attacks on Existing Algorithms

The following attacks can still compromise IoT security solutions: [34]

Denial of Service (DoS): This will stop network services for authorized users as a result of requests from unauthorized users to access the network.

Man-in-the-Middle: In this scenario, a third party user can access one side's key and initiate communication on that side's behalf. Eavesdropping: An intruder can hear what the transmitter and receiver are saying. Therefore this is an attack on confidentiality. Masquerading occurs when an unauthorized user assumes the identity of another authorized user. So that IoT resources can be destroyed.

Saturation: In this situation, the intruder will attempt to use the authorized party's physical and mental abilities through excessive use.

Differential: A change in input behavior will have an impact on the output. The key can thus be discovered through network transitions using this attack.

10. Research Challenges and Solution

IoT is now accepted in homes, workplaces, public spaces, and commercial organizations, which will present security and privacy challenges. As a result, concerns about security and privacy are growing in importance when using the Internet of Things. If any threat is introduced into the Internet of Things, the potential loss is obvious to imagine. There are many different types of IoT attacks, including signal spoofing, DoS, replay attacks, and false signal injection. These attacks will undermine IoT security features including confidentiality, integrity, and authentication; they will also have an effect on users' privacy. The Internet of Things (IoT) has built-in, rudimentary security measures at each layer, yet these are still open to intrusion.

Due of its limited resources, like as power and real-time execution, traditional cryptography and authentication algorithms do not work effectively in an IoT situation. As a result, IoT solutions for lightweight cryptography frequently function effectively.

Many lightweight Symmetric and Asymmetric cryptography methods, including AES, HIGHT, RC5, PRESENT, RSA, and ECC, are available in the literature. Due to longer execution times, longer codes, and higher memory requirements, these current solutions cannot guarantee the highest level of security in real-time communication. Execution time takes into account time for key management and distribution, encryption, and decryption, which determines how well the protocol works. Due to their high key sizes, asymmetric algorithms are extremely slow, whereas symmetric algorithms can only provide secrecy and integrity but no authentication, which makes them vulnerable to availability attacks.

This may affect real-time information gathering and processing and will waste IoT resources. This demand for a safe algorithmfor IoT that would guarantee services like confidentially, integrity and authentication in optimal time.

Many authors have presented lightweight symmetric and asymmetric security algorithms for IoT based on the findings of a literature review that was conducted. Symmetric algorithms give confidentiality, integrity, have tiny key sizes, and reduced complexity, but they do not provide authenticity, and it is difficult to distribute keys in them. Asymmetric algorithms, on the other hand, offer confidentiality, integrity, and authenticity, but their key size is too huge, making them more complex and unsuitable for restricted IoT scenarios. So, a secure algorithm is required that maps the best aspects of lightweight symmetric and asymmetric algorithms in a way that requires the least amount of energy and ensures all security services, such as confidentiality, integrity, and authenticity.

11. Conclusion

IoT has a variety of challenges, including those related to power, bandwidth, scalability, heterogeneity, security, and privacy. To keep people's trust in IoT, security and privacy are the most important problems to overcome. Predefined security solutions at each layer are still vulnerable to attacks. As a result, security can be ensured using cryptographic techniques. But because of certain restrictions, IoT is not a good fit for traditional, heavyweight algorithms. As a result, one hybrid and strong Encryption Scheme must be presented, incorporating the advantages of lightweight symmetric and asymmetric algorithms, ensuring confidentiality, integrity, and authenticity with minimal execution time and energy.

References

1. Naru, Effy Raja, Hemraj Saini, and Mukesh Sharma. "A recent review on lightweight cryptography in IoT." 2017 international conference on I-SMAC (IoT in social, mobile, analytics and cloud) (I-SMAC). IEEE, 2017.
2. Xin, Mingyuan. "A mixed encryption algorithm used in internet of things security transmission system." 2015 international conference on cyber-enabled distributed computing and knowledge discovery. IEEE, 2015.
3. Goyal, Tarun Kumar, and Vineet Sahula. "Lightweight security algorithm for low power IoT devices." 2016 international conference on advances in computing, communications and informatics (ICACCI). IEEE, 2016.
4. Mazumder, Rashed, Atsuko Miyaji, and Chunhua Su. "A simple construction of encryption for a tiny domain message." 2017 51st Annual Conference on Information Sciences and Systems (CISS). IEEE, 2017.
5. Pandya, Hetal B., and Tushar A. Champaneria. "Notice of Removal: Internet of things: Survey and case studies." 2015 international conference on electrical, electronics, signals, communication and optimization (EESCO). IEEE, 2015.
6. Bertino, Elisa. "Data Security and Privacy in the IoT." EDBT. Vol. 2016. 2016.
7. Zhao, Wenfeng, Yajun Ha, and Massimo Alioto. "AES architectures for minimum-energy operation and silicon demonstration in 65nm with lowest energy per encryption." 2015 IEEE International Symposium on Circuits and Systems (ISCAS). IEEE, 2015.
8. Garzia, Fabio, and Luca Papi. "An Internet of Everything based integrated security system for smart archaeological areas." 2016 IEEE International Carnahan Conference on Security Technology (ICCST). IEEE, 2016.
9. Bui, Duy-Hieu, et al. "AES datapath optimization strategies for low-power low-energy multisecurity-level internet-of-things applications." IEEE Transactions on Very Large Scale Integration (VLSI) Systems 25.12 (2017): 3281-3290.
10. Dewan, Surbhi. "Comparative Study of Security Protocols to Enhance Security." 2015 Fifth International Conference on Advanced Computing & Communication Technologies. IEEE, 2015.
11. Darshana Pritam Shah, Pritam Gajkumar Shah ," Revisting of Elliptical Curve Cryptography for Securing Internet of Things (IOT)" 2018 Advances in Science and Engineering Technology International Conferences (ASET) IEEE-2018
12. Lara-Nino, Carlos Andres, Arturo Diaz-Perez, and Miguel Morales-Sandoval. "Elliptic curve lightweight cryptography: A survey." IEEE Access 6 (2018): 72514-72550.
13. Sharafi, Masoumeh, et al. "A low power cryptography solution based on chaos theory in wireless sensor nodes." IEEE Access 7 (2019): 8737-8753.
14. Singh, Saurabh, et al. "Advanced lightweight encryption algorithms for IoT devices: survey, challenges and solutions." Journal of Ambient Intelligence and Humanized Computing (2017): 1-18.
15. Wenfeng Zhao, 1;2Yajun Ha, 1Massimo Alioto, "AES Architectures for Minimum-Energy Operation and Silicon Demonstration in 65nm with Lowest Energy per Encryption ",978-1-4799-8391-9/15/$31.00 ©2015 IEEE
16. Yao, Xuanxia, Zhi Chen, and Ye Tian. "A lightweight attribute-based encryption scheme for the Internet of Things." Future Generation Computer Systems 49 (2015): 104-112.
17. Pérez, Salvador, et al. "A lightweight and flexible encryption scheme to protect sensitive data in smart building scenarios." IEEE Access 6 (2018): 11738-11750.
18. Ambrosin, Moreno, et al. "On the feasibility of attribute-based encryption on internet of things devices." IEEE Micro 36.6 (2016): 25-35.
19. Ambrosin, Moreno, Mauro Conti, and Tooska Dargahi. "On the feasibility of attribute-based encryption on smartphone

devices." Proceedings of the 2015 Workshop on IoT challenges in Mobile and Industrial Systems. 2015.
20. Oualha, Nouha, and Kim Thuat Nguyen. "Lightweight attribute-based encryption for the internet of things." 2016 25th International Conference on Computer Communication and Networks (ICCCN). IEEE, 2016.
21. Moffat, Steve, Mohammad Hammoudeh, and Robert Hegarty. "A survey on ciphertext-policy attribute-based encryption (CP-ABE) approaches to data security on mobile devices and its application to IoT." Proceedings of the International Conference on Future Networks and Distributed Systems. 2017.
22. Çorak, Burak H., et al. "Comparative analysis of IoT communication protocols." 2018 International symposium on networks, computers and communications (ISNCC). IEEE, 2018.
23. Al-Fuqaha, Ala, et al. "Internet of things: A survey on enabling technologies, protocols, and applications." IEEE communications surveys & tutorials 17.4 (2015): 2347-2376.
24. Thapliyal, Himanshu, T. S. S. Varun, and S. Dinesh Kumar. "Adiabatic computing based low-power and DPA-resistant lightweight cryptography for IoT devices." 2017 IEEE Computer Society Annual Symposium on VLSI (ISVLSI). IEEE, 2017.
25. Eisenbarth, Thomas, et al. "A survey of lightweight-cryptography implementations." IEEE Design & Test of Computers 24.6 (2007): 522-533.
26. Sevin, Abdullah, and Abdu Ahmed Osman Mohammed. "A survey on software implementation of lightweight block ciphers for IoT devices." Journal of Ambient Intelligence and Humanized Computing (2021): 1-15.
27. Shahzadi, Romana, et al. "Chaos based enhanced RC5 algorithm for security and integrity of clinical images in remote health monitoring." IEEE Access 7 (2019): 52858-52870.
28. Aguilar, Jeremy, Stephanie Sierra, and Edwar Jacinto. "Implementation of 'HIGHT'encryption algorithm on microcontroller." 2015 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON). IEEE, 2015.
29. Suárez-Albela, Manuel, et al. "A practical performance comparison of ECC and RSA for resource-constrained IoT devices." 2018 Global Internet of Things Summit (GIoTS). IEEE, 2018.
30. Hammi, Badis, et al. "A lightweight ECC-based authentication scheme for Internet of Things (IoT)." IEEE Systems Journal 14.3 (2020): 3440-3450.
31. Shah, Pooja, Mukesh Arora, and Kinjal Adhvaryu. "Lightweight Cryptography Algorithms in IoT-A Study." 2020 Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). IEEE, 2020.
32. Odelu, Vanga, et al. "Expressive CP-ABE scheme for mobile devices in IoT satisfying constant-size keys and ciphertexts." IEEE Access 5 (2017): 3273-3283.
33. Das, Sangjukta, and Suyel Namasudra. "Multiauthority CP-ABE-based Access Control Model for IoT-enabled Healthcare Infrastructure." IEEE Transactions on Industrial Informatics 19.1 (2022): 821-829.
34. Deogirikar, Jyoti, and Amarsinh Vidhate. "Security attacks in IoT: A survey." 2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC). IEEE, 2017.

An Adaptive PID based Water temperature control of a Thawing Unit for Artificial Insemination of Livestock Animals

Mehul Shah[a], Amit Choksi[b], Rajendra Chhasatiya[c], Ritesh Prajapati[c], Dipen Parmar[d]

[a]Asso. Professor,GCET, EC Dept, [b]Assist. Professor,BVM, EC Dept.[c] System Level Solution, V.V.Nagar [d] Gavyam

Abstract

Thawing is the process to liquefy frozen or sex-sorted semen precisely at 37DegC precisely taken out from ~200DegC from Liquid Nitrogen Container during the artificial insemination process of livestock animals like cow and buffalo. Traditional method employed in the field does not guarantee temperature control of this process. Use of a Thawing unit is encouraged to improve the success rate of breeding livestock animals. This work presents a low cost and precise solution of design of a thawing unit. A prototype of the thawing unit is developed to control water temperature for liquefy frozen semen using Arduino Uno. Simple on-off control does not regulate water temperature at precise 37DegC, hence a PID control mechanism is used to provide automated temperature control. The prototype of thawing unit is tested with various PID parameters to see the temperature fluctuation with respect to set temperature of 37DegC. Experimental result suggests using an adaptive PID mechanism to control water temperature precisely at 37DegC instead of a basic PID control. With the use of adaptive PID control, water temperature of the thawing unit is controlled with the limit less than +-0.5 degree centigrade of set value temperature.

Keywords: Thawing Unit , PID Control, Temperature Control, Water Heater Design, Adaptive control

1. Introduction

Artificial Insemination (AI) in livestock animals is a technique in which sperm are collected from the male animal, processed, stored, and artificially introduced into the female animal's reproductive tract at proper time for purpose of conception [1]. Phenomenal growth of Artificial Insemination is due to improvement in fertility resulting from various methods [2],[3].

Thawing is the process to liquefy frozen or sex-sorted semen at 37DegC precisely taken out from ~200DegC from Liquid Nitrogen Container during the artificial insemination process of animals like cow and buffalo. Thawing semen in warm water (35-38 °C) for 40 seconds is the most used thawing procedure [4]. Use of a Thawing unit for controlling water temperature is the most effective way for the artificial insemination process. It equips users to set the correct temperature for the thawing process and can also display live temperature readings to minimize errors during semen thawing process. Various manufacturers have made such products, but none guarantees accurate temperature control of water while artificial insemination process. In this paper accurate temperature control of water with the help of proportional-integral-derivative (PID) method is implemented using Arduino Uno board.

Many researchers have applied PID control methods for patents, software packages and commercial hardware modules [5]. Authors in [6] has applied PID control of room temperature using microcontrollers. PID control methods are successfully applied to control temperature of water inside electric kettle for making tea and coffee[7]. Authors in [8] have applied PID temperature control using LABVIEW and Arduino and compared their method with on-off control. In [9] a PID based temperature of water dispenser is controlled using Arduino Uno and DS18B20 temperature sensor using Ziegler-Nichols tuning method. The experimental work in this paper uses an adaptive PID control method for precise temperature control of a thawing process used for AI.

To achieve precise temperature control, a PID control algorithm is implemented using Arduino Uno. A basic PID control algorithm is implemented to achieve precise water temperature control with different values of proportional, integral, and derivative parameters. An adaptive PID control algorithm is also implemented to achieve further accuracy of water temperature control within range of +-0.5 degrees of temperature.

Section II discusses heater design for a thawing unit and various issues related to its design. Section III presents basics of PID control, its implementation on Arduino Uno and two algorithm implementations on Arduino for precise temperature control. Section IV discusses various experimental results and analysis and section V conclude the paper.

2. Design of Heater for a Thawing Unit

In this section first various issues related to heater design are described and then heating time required for practical Heater design are described.

Once semen straws are injected for AI process from liquid nitrogen container at -200 degree Celsius, water temperature of a thawing unit momentarily reduced to very low temperature level. The thawing unit therefore uses a heater for keeping water temperature precisely at 37DegC for a pre-defined time of 30-40 second after insertion of semen straws.

An aluminum pipe is used as the core material of a heater in which the water is going to fill. Aluminum has high thermal conductivity and can be heated very fast. A nickel-chromium wire is used as a heating coil. Between the pipe and coil, an insulation made from Teflon is placed. Teflon paper is used between pipe and coil as insulation so that the current through the coil does not pass through the aluminum pipe and current leakage does not happen.

At the bottom of the pipe, the DS18B20 Temperature sensor [10] has been fixed to sense the water temperature for control purposes. The core functionality of the DS18B20 is its direct-to-digital temperature sensor. The resolution of the temperature sensor is user- configurable to 9, 10, 11, or 12 bits, corresponding to increments of 0.5°C, 0.25°C, 0.125°C, and 0.0625°C, respectively. The default resolution at power-up is 12-bit. The main objective of the heater is to supply heat to the water inside the aluminum tube and distribute it uniformly throughout its length.

Heat supply to the water is controlled through controlling current to the heater. An Arduino Uno is used to supply Pulse Width Modulated (PWM) current on one of its General-Purpose Input-Output (GPIO) pins as shown in fig-1. The heater has a maximum current carrying capacity of 2 Ampere, while Arduino can supply a maximum of 40 milliamp, therefore a TIP120 Darlington transistor is used as shown in fig-1 to increase the current capacity up to 2 amperes.

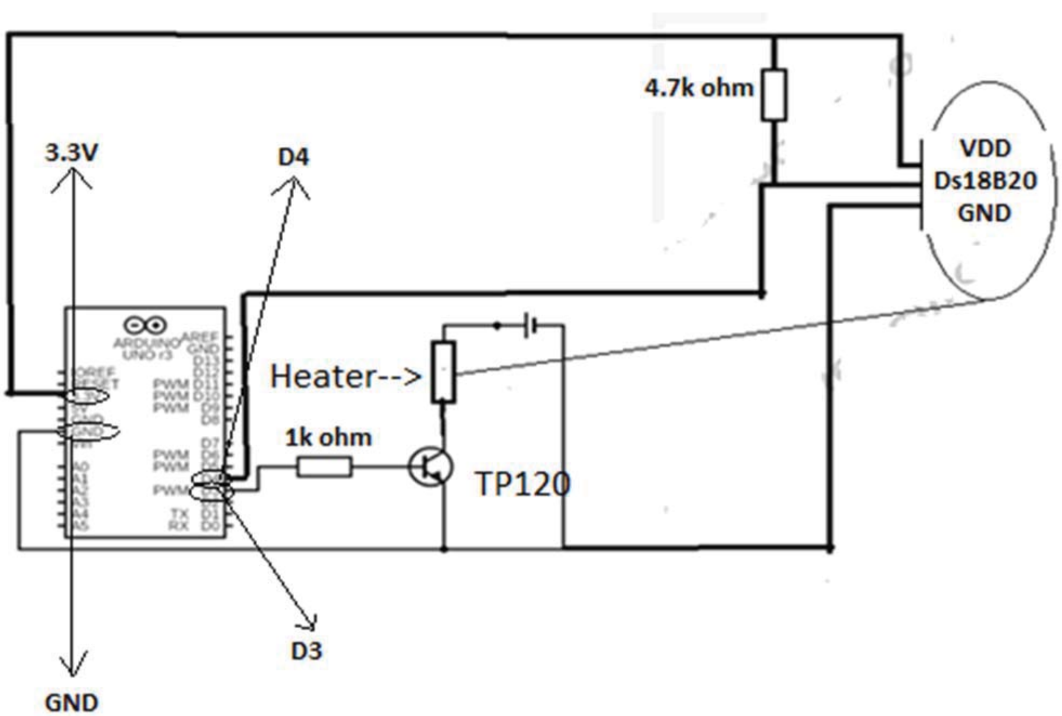


Fig. 1. Basic Circuit for Temperature control

2.1. Heater Design and Heating time calculation

Aluminum pipe is used as a core material of the Heater in which water will be filled. The pipe is 14 cm in length. Its outer Diameter is 20mm and inner diameter is of 15.5mm so thickness is of 2.25mm. Total volume of water inside the pipe is 25 ml. (volume = pi*r^2*length).
For this dimension of pipe, a 1.2-meter long 21-gauge nichrome wire is wounded as a heating coil, having a resistance of 3.1 ohms. 21-gauge nichrome wire has 2.6 ohm/meter resistance. So, 1.2-meter wire will have 3.1-ohm resistance.

To calculate the time required to heat 25ml of water from 19 degrees Celsius to 37 degrees Celsius using 21-gauge nichrome wire wounded on an aluminum pipe, following calculations are performed.

The heat Q required to raise the temperature of the water, Q = m•C•ΔT, Where m = mass of water, C = specific heat capacity of water and ΔT = Temperature difference. In this case, m = 25ml = 0.025kg (because the density of water is 1g/ml), C = 4.18 joule/g°C (Specific heat capacity of water at room temperature) and ΔT = 37-19 = 18°C
So, Q = (0.025kg)*(4.18 j/g°C)*(18°C) = 1.88kJoule

Now the power output of the heating coil P = VI, so P = (5.8) *(1.8) (in our case V=5.8, I=1.8) = 10.44W
Now practically the heating coil will lose some of this power to the aluminum pipe, So useful power P will be 6.264W, after assuming 40% of power dissipation.

Therefore, now the time t = Q/P=1.88kJ/6.264W=301 seconds to heat water pipe at level of 37 deg C.
Therefore, the heater will take approximately 301 seconds or 5 minutes to raise the temperature from 19 to 37 degree Celsius.

3. PID based Temperature control of the Heater

PID control is a feedback mechanism used in a control system. This type of control is also termed as three-term control and is implemented by a PID Controller. The proportional, derivative, and integral parameters can be expressed as Kp, Kd and Ki respectively. By these three parameters, how much a process variable (temperature for example) deviates from the desired set point value can be calculated and controlled. All these three parameters have effect on rise time, settling time, overshoot, and the steady state error of the process variable. Therefore, different control actions for specific work can be achieved with these parameter settings. For PID control the actuating signal at the input of a process block consists of proportional error signal added with derivative and integral of the error signal as shown in fig.2.

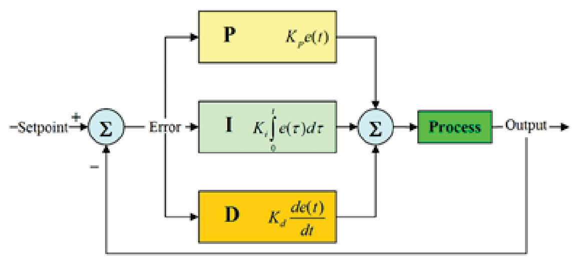


Fig. 2. PID Control block diagram

PID controller consists of three sub-units: an Arduino Uno R3; DS18B20 temperature sensor; and heater with transistor driver TIP 120 circuit as shown in fig. 3. The TIP120 acts like a switch, which is turned on and off by the PWM signal (Pin 3 of Arduino) at the base. A 1000-ohm resistor is added between the base of TIP120 and the PWM pin NO. 3. DS18B20 sensor read the temperature data from the heater. The power is provided to the circuit with the help of a +12V external power supply. The value set by push button is set by set value variable defined in the program in present implementation.

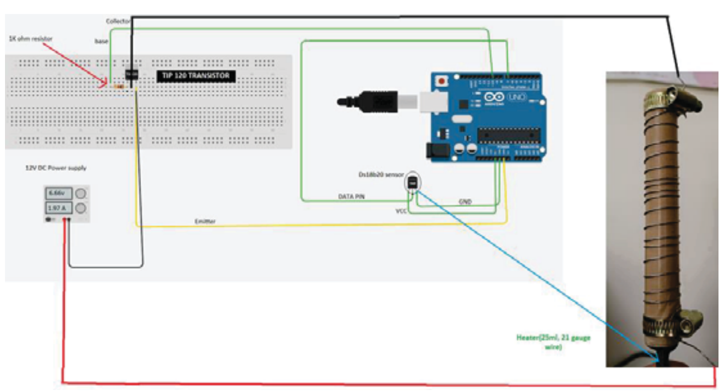


Fig. 3. Arduino based Temperature control of Heater

3.1 PID Control Algorithm

The PID temperature control works using a mathematical formula to calculate the difference between the current temperature and setpoint and then it tries to deliver the required power to ensure the target temperature remains constant. This also reduces overshoots that can be difficult to found in the traditional on-off control mechanism.

The operation of the PID based water temperature control system is as follows:

- The measured temperature with DS18B20 is compared with the setpoint temperature.
- The PID algorithm calculates the output signal based on the error between the setpoint and the measured temperature.
- The output control signal is sent to the PWM pin of Arduino, which controls the heater temperature.

A basic PID Control algorithm is used to implement a simple PID control strategy. Based on experimentation of this basic PID control with various Kp, Ki and Kd values an adaptive PID algorithm is also implemented to achieve precise water temperature control. Both of these algorithms are described next using pseudocode.

```
Algorithm 1: Basic PID Controller

    Require: Sampling_Time, Temperature_Set_point, kp, ki, and kd
    Start
    Set Error = 0
    Set Sampling_Time =0
    Set Sum_Error = 0
    Set Last_Error = 0
    Set Temperature_Set_point = 37
    Define kp, ki, and kd
    Pass the Temperature_Set_point, kp, ki, and Kd to the standard PID controller
    Find Sampling_Time = (Current_Time – Last_Time) * 1o^-3
    Compute Error = Temperature_Set_point – Past_Output
    Compute Sum_Error = Sum_Error + (Error * Sampling_Time)
    Compute Rate_Error = (Error - Last_Error)/Sampling_Time
    Set Last_Error = Error
    Set Last_Time = Current_Time
    Find Output = kp * Error + ki * Sum_Error + kd * Rate_Error
    Return Output
    End PID Control Process

Algorithm 2: Adaptive PID Controller

    Require: Sampling_Time, Temperature_Set_point
    Require: aggkp, aggki, and aggkd
    Require: conskp, conski and conskd
    Start
    Set gap = 0
    Set PID Output
    Set Input
    Set Temperature_Set_point = 37
    Compute Error = Temperature_Set_point – Input
    if (gap < 2)
    Compute PID Output using conservative conskp, conski and consKd
    else
    Compute PID Output using aggressive aggkp, aggki and aggkd
    end
    Return Output
    End Adaptive PID Control Process
```

4. Results and Analysis

An experimental setup as shown in figure 4 is implemented on Arduino Uno to control temperature of the heater using PID Control algorithm. Temperature of the water and current consumption data are collected using cool term software through UART interface of the board. Collected data are stored in csv files and then analyzed using MATLAB software.

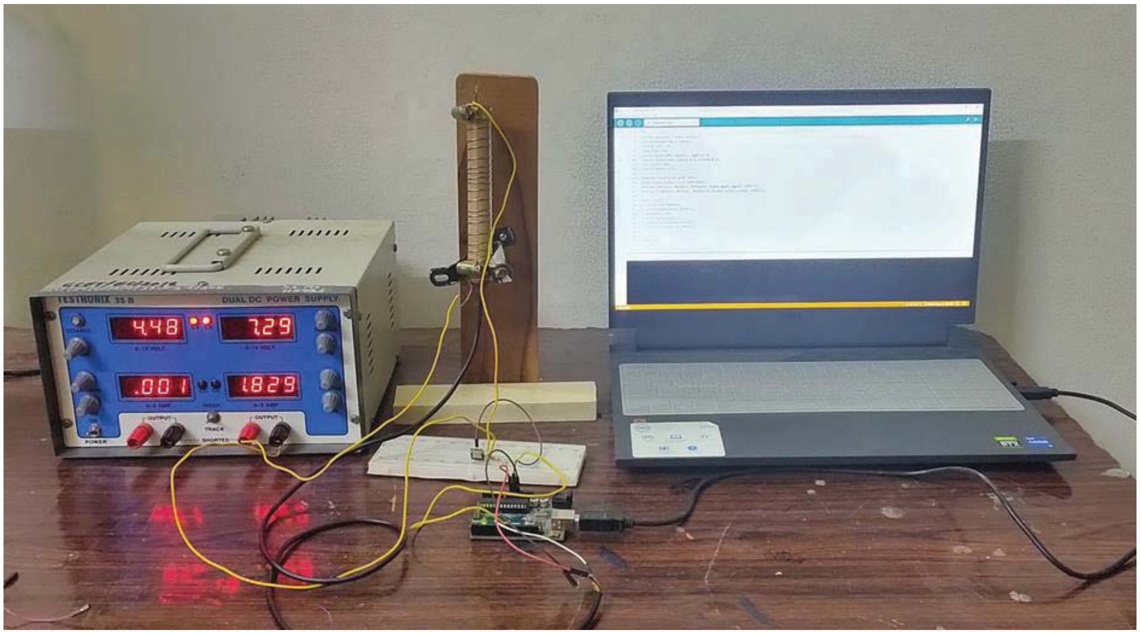


Fig. 4. Experimental setup for Temperature control

A basic PID Control algorithm is implemented first by varying different values of Kp, Ki and Kd, while keeping others values constant to see parameter's effect on rise time and over shoot values of the control temperature. Experiments are performed with below mentioned values of all the parameters.

- Values of Kp are changed from 400 t0 450 while keeping Ki at 5 and kd at 0.
- Value of Ki is changed from 2 to 50 while keeping Kp at 400 and kd at 0.
- Value of Kd is changed from 1.5 to 2 while kp and ki fixed at 600 and 8 respectively.

Results of above experiments are analysed with MATLAB on different CSV files of collected data and final summary of important observations are made in table-1.

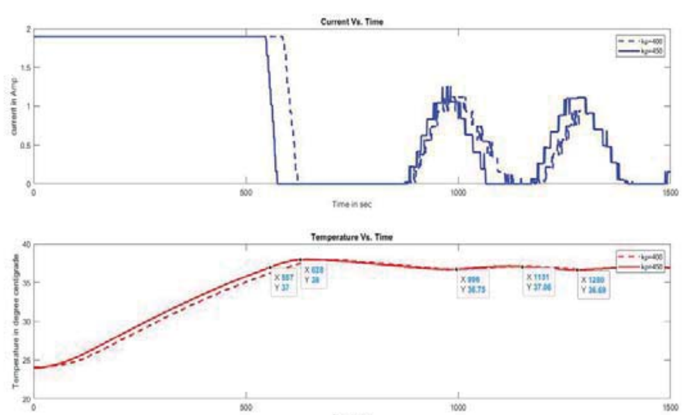


Fig. 5. Effect of increasing kp on rise time and overshoot (ki and kd constant at 5 and 0)

Experimental result as shown in fig. 5 suggest that as Kp values are increased, the rise time (Time to reach set temperature value) is reduced and first overshoot is marginally increased. This is in accordance to standard observation in the reference literature [5].

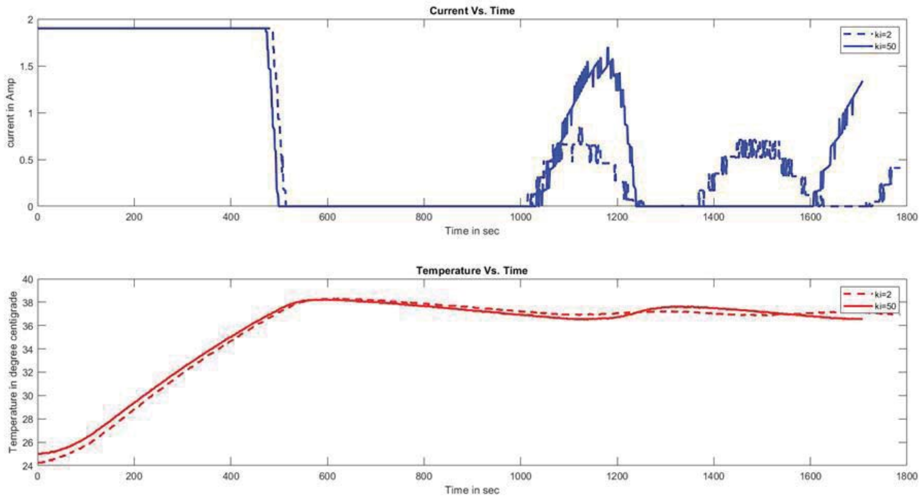


Fig. 6. Effect of increasing Ki on rise time and overshoot (Kp and Kd constant at 400 and 0)

Experimental result as shown in fig. 6 suggest that increase in ki value reduces rise time only marginally as compared to Kp. Overshoot has also increased marginally.

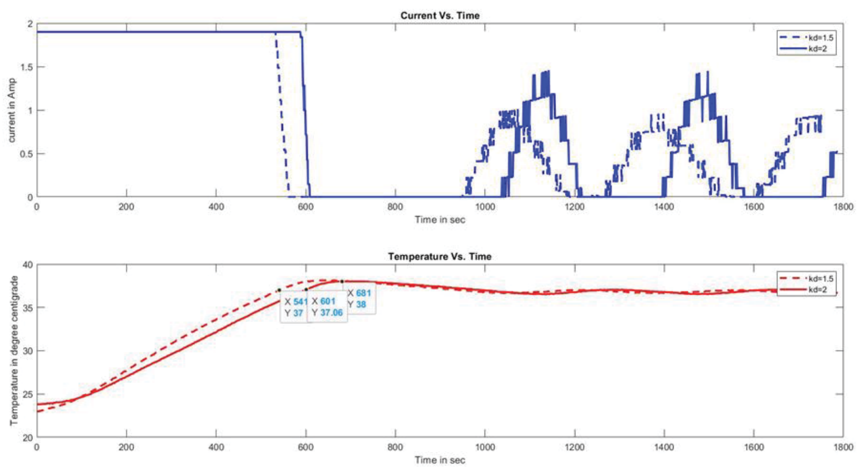


Fig. 7. Effect of increasing Kd on rise time and overshoot (Kp and Ki constant at 600 and 8)

As shown in fig.7 the overshoot has decreased.

Table 1. Summary of results plotted in fig. 5,6 and 7

Kp	Ki	Kd	Rise time	Maximum Overshoot	Maximum Undershoot
400	5	0	4:42 min	37.94 °C	36.75 °C
450	5	0	4:40 min	38 °C	36.75 °C
400	2	0	4.08 min	38.31°C	36.88°C
400	50	0	4.08 min	38.19 °C	36.56°C
600	8	1.5	4.36 min	38.13 °C	36.63-°C
600	8	2	5.02 min	38.00 °C	36.5-°C

Next an adaptive PID control is implemented for accurately controlling the temperature of the water temperature.

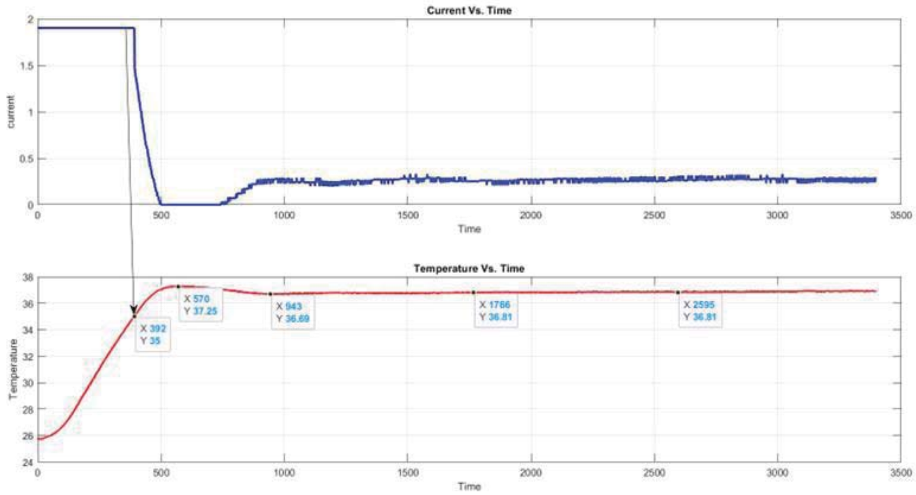


Fig.8 Adaptive PID control within Gap of 2degreeC

Experimental results suggest a precise control of temperature within the range of +- 0.5 degC.

5. Conclusion

Designing temperature control systems can be made easier with the assistance of Arduino Uno Hardware and software. Based on the experiments, it can be concluded that the system can maintain the temperature of water in the aluminum pipe at set level of 37 degrees and can maintain it stable with a range from 37.5 -36.5 degreec. Proper tuning of Kp, Ki and Kd values are crucial for accurate temperature of water at set value. Experimental result shows variation of Kp, Ki and Kd as per independent control of these parameters. However, for precise control of temperature with in a range of -+ 0.5 degC a use of Adaptive PID is required. The Kp, Ki and Kd parameters need to be fine-tuned for an optimum temperature control with both PID methods.

References

1. Foote, R. H. "The history of artificial insemination: Selected notes and notables." J. Anim. Sci 80 (2010): 1-10
2. Selk, Glenn. Artificial insemination for beef cattle. Oklahoma Cooperative Extension Service, 2004.
3. Yadav, Sarvajeet, et al. "Artificial Insemination in bovine and livestock management." 2022.
4. kumar Patel, Gaurang, et al. "Artificial insemination: A tool to improve livestock productivity." Journal of Pharmacognosy and Phytochemistry 6.6S (2017): 307-313.
5. Ang, Kiam Heong, Gregory Chong, and Yun Li. "PID control system analysis, design, and technology." IEEE transactions on control systems technology 13.4 (2005): 559-576.
6. Ranjan, Sudhir, Abhishek Sharma, and Puneet Chaudhary. "An effective temperature controller system using PID mechanism." 2014 Innovative Applications of Computational Intelligence on Power, Energy and Controls with their impact on Humanity (CIPECH). IEEE, 2014.
7. Shah, M. B. N., Zailany, N., Abidin, A. F. Z., Halim, M. F., Annuar, K. A., Azahar, A. H., & Yaakub, M. F. (2019). PID-based temperature control device for electric kettle. International Journal of Electrical and Computer Engineering, 9(3), 1683-1693.
8. Asraf, H. Muhammad, et al. "Development of experimental simulator via Arduino-based PID temperature control system using LabVIEW." Journal of Telecommunication, Electronic and Computer Engineering (JTEC) 9.1-5 (2017): 53-57.
9. Aisuwarya, Ratna, and Yulita Hidayati. "Implementation of ziegler-nichols PID tuning method on stabilizing temperature of hot-water dispenser." 2019 16th International Conference on Quality in Research (QIR): International Symposium on Electrical and Computer Engineering. IEEE, 2019.
10. Semiconductor, Dallas. "Programmable resolution 1-wire digital thermometer." Data Sheet DS18B20 (2002).

Acknowledgement

The authors would like to thank System Level Solutions and Gavyam Gentech Private Limited for helping this project work.