ICPCSN

4th International Conference on Pervasive Computing and Social Networking (ICPCSN-2024)

3-4, May 2024 | Salem, India

PUBLISHED BY
IEEE COMPUTER SOCIETY
CONFERENCE PUBLISHING SERVICES

R P Sarathy Institute of Technology

Organized by

R P Sarathy Institute of Technology
Salem, India

ICPCSN

4th International Conference on Pervasive Computing and Social Networking (ICPCSN-2024)

3-4, May 2024 | Salem, India

PUBLISHED BY
IEEE COMPUTER SOCIETY
CONFERENCE PUBLISHING SERVICES

R P Sarathy
Institute of Technology

Organized by

R P Sarathy Institute of Technology
Salem, India

Message from Chairman

It is my great pleasure to announce that the “4th International Conference on Pervasive Computing and Social Networking”, ICPCSN 2024, is to be held at our institution on 3-4, May 2024.

I have firm belief that ICPCSN 2024 would be the starting point in exploring our institutional capabilities in the area of Engineering and Technology. It is my pleasure and privilege to extend a whole-hearted welcome to all the participants for sharing your research aptitudes through this conference. I appreciate the conference organizing committee for showing a keen interest in organizing a successful international conference and contributing new ideas and research findings. I wish them for their endeavours to spread knowledge.

I am sure the conference will provide you updated information on the recent researches in the field of Engineering. We all look forward to your presence and involvement in ICPCSN conference. I am delighted to be able to welcome you all.

Mr. B. Nitish Harihar,

Chairman,

R P Sarathy Institute of Technology,

Salem, Tamil Nadu, India.

Message from Director

On behalf of RPIT, I am delighted to invite you to the 4th International conference ICPCSN 2024 in May 2024. Research and development is always a sign of learning and development. In today's society, thinking and acting in a new and innovative way is highly essential to cope with the technological changes. I hope, this 4th ICPCSN will be a forum for scholarly discussions on computing and communication technologies.

The dedication shown by researchers towards this conference is highly encouraging. This motivates us to hold such conferences more regularly in the future. I would like to express my heartfelt appreciation to everyone who participated in this conference.

I wish the conference the best of luck.

Dr. K. Arutselvan,

Director- IQAC,

R P Sarathy Institute of Technology,

Salem, Tamilnadu, India.

Message from Conference Chair

On behalf of the conference organizing committee, I am pleased to welcome you to the 4th International Conference on Pervasive Computing and Social Networking. It is highly satisfying to witness that the conference is getting larger and more noteworthy each year. I am highly excited about the record number of participants and number of sessions planned to be organized in 4th ICPCSN 2024 event.

As the world is becoming more digital, there is an increasing need to enable rich networking and computing technologies. The theme of this year's conference is "Pervasive Social Computing". The 2024 conference aims to initiate research discussions on enhancing human social relationships with leading scientific communities in different disciplines from computer science to communication engineering. The technical program of ICPCSN 2024 is rich and varied with keynote speech and invited talks, technical paper presentation and oral sessions each day. We also expect to provide technical discussions, and numerous opportunities for informal networking..

The success of any conference depends on the people, who worked with us in planning and organizing the technical and non-technical arrangements. In that way, I extend my sincere thanks to the program chairs for their continuous guidance; Review Committee for the timely reviewing of the papers and local organizing committee for all the non-technical arrangements. Lastly, I would like to thank all of the conference participants for their contributions which are the foundation of this conference.

I hope that you will have a productive time at this 2023 conference.

Dr. Munusami Viswanathan,

Principal,

R P Sarathy Institute of Technology,

Salem, Tamil Nadu, India.

Performance Improvement of DC Motor by using ANFIS Controller

Mulla Gouse Basha
Assistant Professor
Department of EEE
Geethanjali College of Engineering and Technology, Hyderabad
m.gousebasha@gmail.com

Kalagotla Chenchireddy
Assistant Professor
Department of EEE
Geethanjali College of Engineering and Technology, Hyderabad
chenchireddy.kalagotla@gmail.com

Dr. Shabbier Ahmed Sydu
Lecturer, E&E Section
University of Technology and Applied Sciences
AI-musannah, Oman
shabbier@act.edu.om

waseem sultana
Research Scholar,
Dept. of EEE
Sathyabama University
Chennai, Tamilnadu
waseemsultana4@gmail.com

V Kumar
Assistant Professor
Dept. of EEE
Teegala Krishna Reddy Engineering College
Hyderabad, Telangana, India
kumarpoma@gmail.com

Lukka Bhanu Ganesh
Assistant Professor,
Dept. of Electrical and Electronics Engineering,
Vignan Institute of technology & science
bhanuganesh.lukka@gmail.com

Abstract: **This research paper is to improve a DC motor's performance by using two different control strategies: Sliding Mode Control (SMC) and ANFIS. Fuzzy Logic (FL) and Neural Networks (NN) are used by the ANFIS controller to show flexibility in the face of nonlinearities, and SMC shows resilience in the face of uncertainty. The dynamic response of the armature current and DC motor speed under each control technique is shown by the simulation results. By highlighting the benefits and drawbacks, comparative analysis helps to choose the best control strategy for a given set of application needs**.

Keywords: **DC motor, Sliding Mode Control, Adaptive Control, Robust Control, Performance Improvement, Simulation**

I INTRODUCTION

DC motors play a pivotal role in diverse industrial and technological applications. Efficient control mechanisms are essential to achieve optimal performance in terms of speed, torque, and energy consumption. The integration of intelligent control systems, such as ANFIS, presents a novel approach to enhance the overall performance of DC motors. Traditional PID controllers have been widely employed for DC motor control, but they may struggle to handle nonlinearities and uncertainties inherent in real-world systems. ANFIS, being a hybrid system combining NN and FL, exhibits superior adaptability and learning capabilities. The performance enhancement of DC motors is crucial for various applications, including robotics, automation, and renewable energy systems. This study investigates the utilization of an ANFIS controller to improve the performance of a DC motor. ANFIS, with its adaptive and self-learning capabilities, offers a promising approach to optimize the motor's speed control and efficiency.

The author focuses on the implementation of third-order SMC for a PMDC motor to regulate its speed. The designed controller is applied to the motor, which is powered by a Buck converter, and simulations are conducted using MATLAB software. The obtained results demonstrate the effectiveness of the proposed third-order sliding mode controller, and a comparison is made with a conventional proportional integral derivative controller, showing satisfactory performance [1]. The authors introduce an Adaptive SMC for a Buck converter-driven DC motor, incorporating a robust proportional derivative control law. The design, based on Lyapunov theory, aims to eliminate chattering in the control signal. The proposed scheme is simple for implementation, ensuring global asymptotic stability if signals remain uniformly bounded. Simulation studies demonstrate superior tracking performance compared to conventional controllers [2]. For a DC motor, the authors present a novel variable-structure position control law. The algorithm makes sure that a sliding mode is established as soon as the shaft motion begins by using a time-varying switching line. The design consists of an initial surface that moves gradually in the direction of a predefined desired surface after passing through the initial representative point. By doing this, the reaching phase is removed, making the motor behavior insensitive to modifications in the unknown mechanical configuration. Using a cheap 16-bit microcontroller, simulation and experimentation are used to test the suggested approach. The study further employs an advanced test bed to assess tracking properties and robustness under variations in mechanical configuration [3]. The authors addressed the position control of a brushed DC motor, a

common actuator in industrial and robotics applications. Two control techniques are compared: the widely used Proportional-Integral-Derivative (PID) controller known for its simplicity and reliability, and the Linear Quadratic Regulator (LQR) employing a state-space approach. The LQR design offers a systematic method for computing the state feedback control gain matrix, minimizing both weighted squared state error and control effort. The study evaluates and compares the performance of both controllers, focusing on disturbance resistance [4]. The authors present and put into practice an SMC algorithm for DC-DC buck converter closed-loop control. The variable switching frequency problem that traditional SMC buck converters have during load and supply changes is intended to be solved by the Improved Sliding Mode control (ISMC). The switching frequency's unpredictability makes filter design challenging. The ISMC seeks to address this problem by appropriately transforming the sliding mode control rule using a pulse width modulation technique. Comprehensive simulation studies in the MATLAB/Simulink environment validate the control algorithm. An extensive comparison analysis is carried out between the converter control systems that utilize ISMC and traditional SMC, emphasizing the fixed-frequency functionality of ISMC under various operating circumstances [5]. This research presents a fuzzy logic controller integrated phase-locked loop induction motor speed drive. The system seeks to combine the benefits of fuzzy logic—such as its simplicity, intuitiveness, ease of implementation, and low requirement for system dynamics knowledge—with the superior speed regulation capabilities of phase-locked loop approaches. The goal of the suggested system is to control the speed of an induction motor precisely, quickly, and robustly. The system's performance is assessed using an experimental setting, and the findings show that the induction motor drive's phase-locked loop and fuzzy logic controller work together to provide accurate speed control with a quick reaction [6].
The authors address the enhancement of output torque and rotational speed accuracy in reaction flywheels, essential for satellite attitude control. The study reviews the modeling and control approaches of DC–DC converters, employing variable structure system theory and SMC. The paper constructs the topology of the reaction flywheel and outlines the small signal linearization process for a buck converter. Using state averaging models and the Lee derivative, the general results of SMC are analyzed. The paper then deduces detailed analytical equivalent control laws for the reaction flywheel, considering various sliding surfaces during different operational stages. Numerical and experimental examples are presented, demonstrating favorable agreement between simulations and experiments. Strong rejection of mild disturbances, high-precision output torque, and rotational speed tracking are all accomplished by the suggested control technique [7]. It also produces preferred rotational speed regulation.

II BUCK CONVERTER WITH SMC

Investigating the process of creating a buck converter's sliding mode control system while taking control gains and switching function design into account [8]-[10]. Insights into the difficulties encountered in real-world applications of sliding mode control in buck converters. It offers strong load disturbance rejection and a quick transient response. The figure 1 shows the control mechanism is naturally resistant to changes in parameters and outside interference and motor performance.

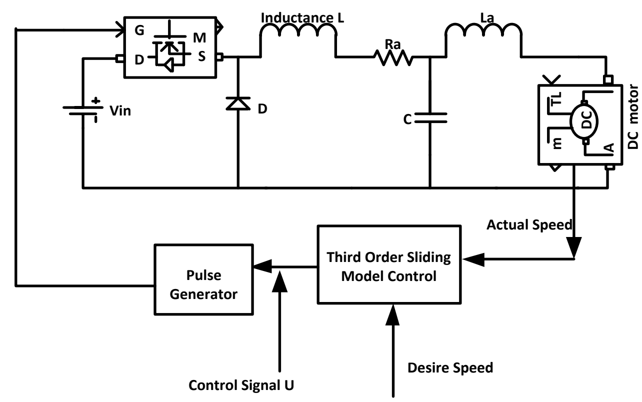


Figure 1: Buck Converter with Sliding Mode Control

The purpose of the sliding mode controller is to control speed within the specified time. The DC motor's speed error is,

$$e = \omega^* - \omega \tag{1}$$

Where,

ω^*=Reference speed

ω=Actual speed

The combination of the proportional, integral, and derivative of the speed error is known as the sliding function.

$$S = \int edt + e + \frac{de}{dt} \tag{2}$$

On the sliding surface S=0, the error variation is zero,

$$0 = \int edt + e + \frac{de}{dt} \tag{3}$$

The sliding function's derivatives are defined as

$$S=\int\omega^{*}dt+\omega-\int\omega dt+\left(\frac{F}{H}-1\right)\omega-\frac{k_t}{H}i_a \quad (4)$$

Where,

F =Friction constant

H =Inertia constant

i_a =Armature current

k_t =Torque constant

k_e =Back emf constant

$$\frac{dS}{dt}=S=\omega^{*}+y_1\omega+y_2 i_a-\frac{k_t}{HL_a}V_a$$

$$\frac{d^2S}{dt^2}=S=y_3\omega+y_4 i_a+\frac{a_2}{L_a}V_a-\frac{k_t}{HL_aC}i_L \quad (5)$$

$$\frac{d^3S}{dt^3}=y_5\omega+y_6 i_a+y_7V_a+\frac{y_2}{L_aC}i_L-\frac{k_tV_{in}}{HL_aLC}u$$

Where,

$$y_1=-1+\frac{F}{H}-\frac{F^2}{H^2}+\frac{k_ek_t}{HL_a} \quad (6)$$

$$y_2=\frac{k_t}{H}+\frac{k_tF}{H^2}+\frac{k_tR_a}{HL_a} \quad (7)$$

$$y_3=-\frac{F}{H}y_1-\frac{k_ey_2}{L_a} \quad (8)$$

$$y_4=\frac{k_t}{H}y_1-\frac{R_a}{L_a}y_2+\frac{k_t}{HL_aC} \quad (9)$$

$$y_5=-\frac{F}{H}y_3-\frac{k_e}{L_a}y_4 \quad (10)$$

$$y_6=\frac{k_t}{H}y_3-\frac{R_a}{L_a}y_4-\frac{y_2}{L_aC} \quad (11)$$

$$y_7=\frac{y_4}{L_a}+\frac{k_t}{HL_aLC} \quad (12)$$

The notation for the third-order derivative is,

$$\frac{d^3S}{dt^3}=M+Nu \quad (13)$$

Where,

$$M=y_5\omega+y_6 i_a+y_7V_a+\frac{y_2}{L_aC}i_L \quad (14)$$

$$N=\frac{k_tV_{in}}{HL_aLC}u \quad (15)$$

To achieve the necessary time of zero sliding function, the control law is selected as follows,

$$\frac{d^3S}{dt^3}=U \quad (16)$$

Where,

$$U=-\alpha\phi_{23} \quad (17)$$

$$\phi_{23}=sign\left(s+\beta\left(|s|^2+|s|^3\right)^{\frac{1}{6}}\left(\phi_{13}\right)\right) \quad (18)$$

$$\phi_{13}=sign\left(s+\beta_1|s|^{\frac{2}{3}}\,sign(s)\right) \quad (19)$$

$$U=M+Nu \quad (20)$$

The sliding mode control law itself is expressed as follows,

$$u=\frac{U-M}{N} \quad (21)$$

Where,

$$M=M^{*}+M_1 \quad (22)$$

$$M^{*}=y_5\omega^{*}+y_6 i_a^{*}+y_7v_a^{*}+\frac{y_2}{L_aC}i_L^{*} \quad (23)$$

$$M_1=y_5\left(\omega-\omega^{*}\right)+y_6\left(i_a-i_a^{*}\right)+y_7\left(v_a-v_a^{*}\right)+\frac{y_2}{L_aC}\left(i_L-i_L^{*}\right) \quad (24)$$

Sliding mode control law divided parts as,

$$u=\frac{U}{N}-\frac{M^{*}}{N}-\frac{M_1}{N} \quad (25)$$

Where,

$\frac{U}{N}=u_{sliding}=$ Sliding control

$-\frac{M_1}{N}=u_e=$ Error control

$-\frac{M^{*}}{N}=u_s=$ Steady state control

III ANFIS CONTROLLER

ANFIS utilizes fuzzy logic to model the system's nonlinearities and a neural network to adaptively tune the fuzzy rules. The controller's ability to self-adjust based on real-time feedback makes it well-suited for dynamic environments. The Methodology study involves the following steps: Modeling the DC motor system, Designing the ANFIS controller, Training the ANFIS model using experimental data, Implementing the ANFIS controller in the DC motor system. The figure 2 shows the data loaded and taring the ANFIS control. Using a learning technique, ANFIS uses this dataset to modify internal parameters during training, including rule weights and membership function parameters. The goal is to reduce the discrepancy between the actual outputs from the training data and the predictions made by the ANFIS model. The figure 3 shows the inputs Fuzzification is an important step before feeding these inputs into the ANFIS model [11]-[15]. Fuzzification is the process of assigning suitable fuzzy sets to the inputs to transform their clear, numerical values into linguistic variables. Membership functions, which specify the extent of belonging to certain linguistic categories, are assigned to each input. The contribution of each input value to the overall fuzzy logic systemis determined by these membership functions.

Figure 2: ANFIS Training Data

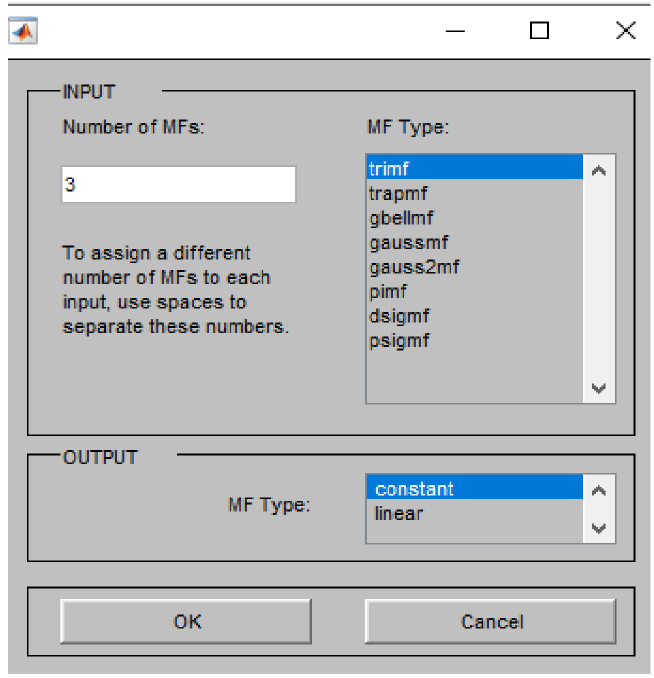


Figure 3: ANFIS Inputs and types

The figure 4 shows training error the degree to which the ANFIS model matches the observed behaviors of the DC motor is quantified by the training error. To lower this mistake, the learning method used during training tries to iteratively modify the ANFIS's parameters, such as membership function parameters and rule weights. The goal of popular optimization approaches like gradient descent and the least squares method is to optimize the model's parameters in order to increase the predicted accuracy of the model. It also shows the error tolerance and eprochs.

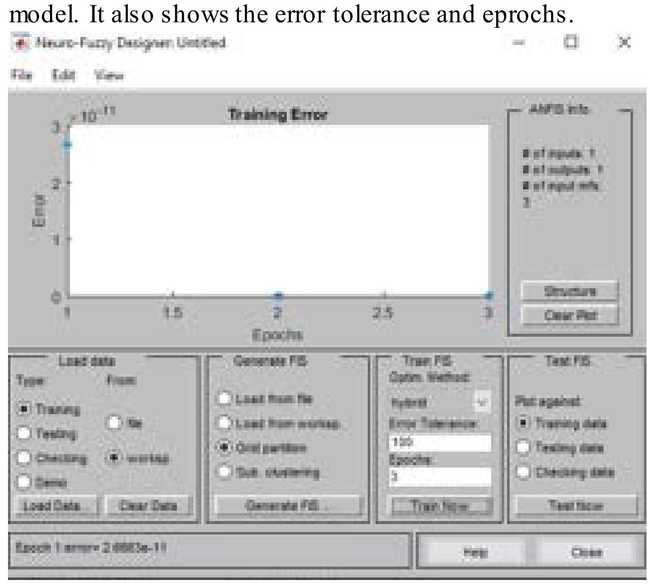


Figure 4: ANFIS Training Error

The figure 5 shows the input and output fuzzy rules are evaluated by the rule layer to handle the fuzzified inputs. These rules, which are derived from the antecedents of the rules, encapsulate the logical connections among the input variables. Rules may specify, for instance, how variations in voltage and load torque impact motor speed in a DC motor system. The degree to which the input variables belong to each fuzzy set determines the firing intensities of these rules.

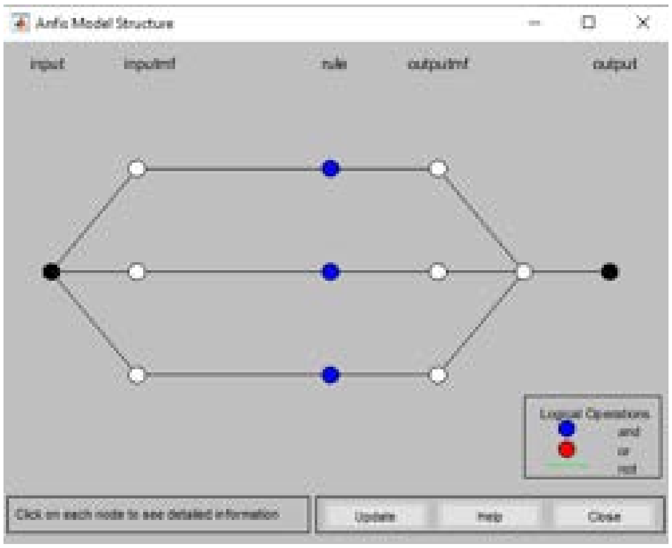


Figure 5: ANFIS input and output structure

The normalization layer, which comes after the rule layer, makes sure that the total firing strengths of all rules add up to 1. In the next steps, a balanced contribution from every rule depends on this normalization. The output layer creates the ANFIS model's final output by combining the normalized firing strengths. This output may indicate the anticipated motor speed in the case of a DC motor.

IV SIMULATION RESULTS

The figure 6 shows An Adaptive Neuro-Fuzzy Inference System (ANFIS) controller is used to simulate the performance of a DC motor. The simulation produces informative findings that demonstrate the efficacy of the control method that was designed. The ANFIS model is used in the simulation to control the behavior of the DC motor. The output, which might be the motor speed or position, is influenced by input variables like voltage, current, and load torque.

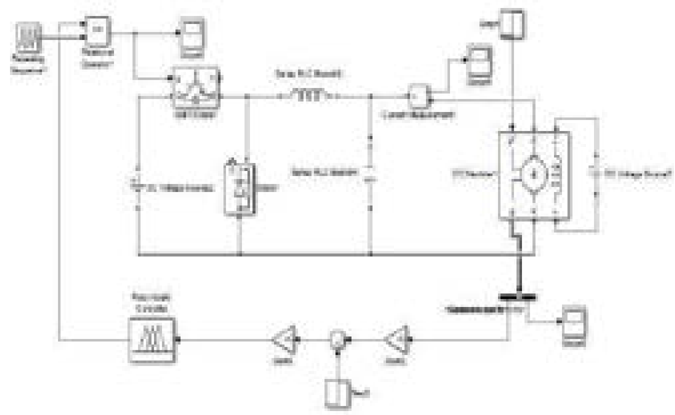


Figure 6: Simulation of ANFIS Control DC motor

The experimental results demonstrate significant improvements in the DC motor's performance when employing the ANFIS controller. Enhanced speed control accuracy, reduced overshooting, and improved energy efficiency contribute to the overall effectiveness of the ANFIS-based control system. A comparative analysis with traditional PID control showcases the advantages of ANFIS in handling nonlinearities and uncertainties. ANFIS proves to be more adaptive and responsive, especially in scenarios with varying load conditions. The findings of this study have practical implications for industries relying on DC motors. The integration of ANFIS controllers can lead to more robust and efficient motor control systems, thereby improving the overall performance and longevity of equipment.

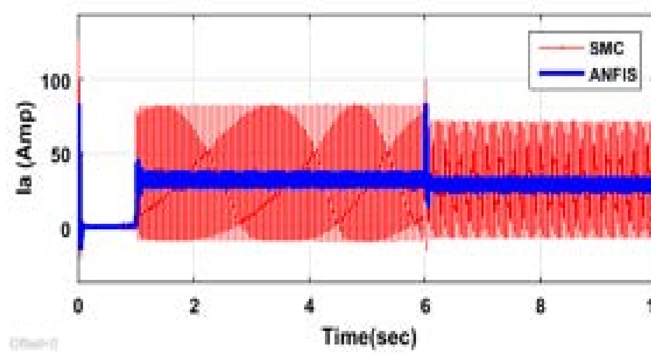


Figure 7: Simulation Output of armature current DC motor

The figure 7 shows ANFIS model is first trained on a dataset of input-output pairs, where the output is the armature current and the inputs might be parameters such as voltage, torque applied to the load, and speed reference. In order to reduce the discrepancy between the anticipated armature current and the actual armature current from the training data, the model's parameters, such as membership functions and rule weights, are adjusted throughout the training phase. On the other hand, the simulation output of the SMC-controlled DC motor shows how reliable the Sliding Mode Control approach is in keeping the armature current within the intended range in the face of outside uncertainties or disruptions. A sliding surface is used by the sliding mode controller to guide the system to the intended state while ensuring a stable and robust response. The output plot for SMC may demonstrate how quickly and precisely disturbances may be adjusted, highlighting the control strategy's capacity to guarantee stability and reliable operation.

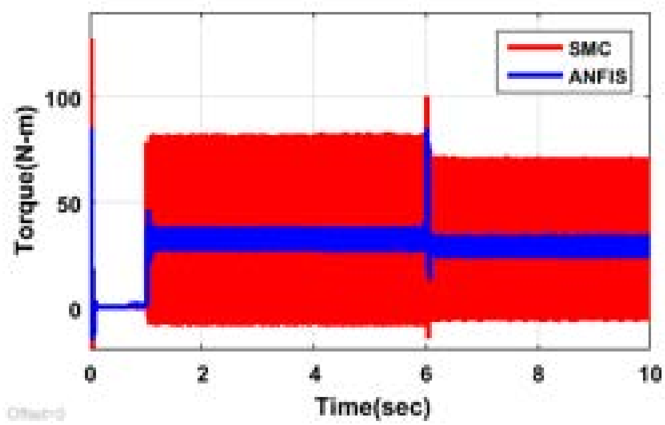


Figure 8: Simulation Output of DC motor Torque

The Figure 8 shows torque simulation result for DC motors shows how effectively the ANFIS controller captures the intricate and nonlinear connections present in the torque dynamics, as well as how well it generalizes to other operating circumstances. The simulation output's correctness is important because it shows how well the ANFIS model maintains the intended torque and adapts to changes in the input variables Conversely, the results of the simulation for DC motor torque under Sliding Mode Control demonstrate how resilient the SMC approach is to ambiguities and disruptions. SMC uses a sliding surface to keep the system on the intended course and ensure stability even when there are outside influences on torque.

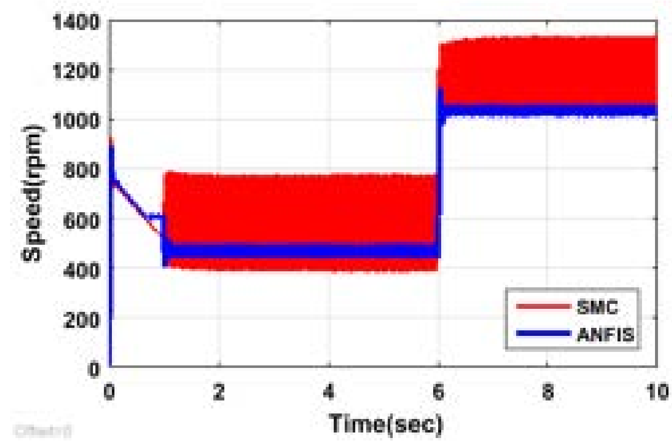


Figure 9: Simulation output of DC motor Speed

The Figure 9 shows simulation output for the ANFIS-controlled DC motor how the system responds to changes in speed while the ANFIS controller is in use. After training on input-output pairs, the ANFIS model adjusts its parameters to precisely forecast and regulate the motor speed. The output plot most likely demonstrates a responsive and smooth speed regulation, demonstrating the flexibility of the ANFIS controller in responding to different operating scenarios and the nonlinearities present in the DC motor system. The simulation result for the SMC-controlled DC motor shows how the robust sliding mode control method provides speed control. Known for its resilience to ambiguities and disruptions, SMC makes sure that the motor speed stays on the intended course even when outside influences are present. The output plot may highlight the sliding mode controller's stability and resistance to disturbances by displaying its fast and accurate speed adjustments.

V CONCLUSION

In conclusion, the utilization of an ANFIS controller for DC motor control demonstrates notable improvements in performance metrics. This study contributes to the growing body of research on intelligent control systems, offering a viable solution for optimizing DC motor operations in various applications. Future research directions may include exploring the adaptability of ANFIS controllers in different motor types, investigating real-time implementation challenges, and optimizing the ANFIS structure for specific industrial requirements. Continuous advancements in intelligent control systems hold the potential to revolutionize the field of motor control and automation.

REFERENCES

[1]. Dhanasekar, R., Kumar, S. G., & Rivera, M. (2018, October). Third order sliding mode control of buck converter fed permanent magnet DC motor. In 2018 IEEE International Conference on Automation/XXIII

Congress of the Chilean Association of Automatic Control (ICA-ACCA) (pp. 1-4). IEEE.

[2]. Chenchireddy, Kalagotla, and Varghese Jegathesan (2023) "CHB multilevel inverter with sliding mode controller for DSTATCOM applications." International Journal of Power Electronics and Drive Systems (IJPEDS) 14(4):195-2203.

[3]. Chouya, A. H. M. E. D. (2023). Adaptive sliding mode control with chattering elimination for Buck converter driven DC motor. WSEAS Trans. Syst. Control, 22, 19-28.

[4]. Babu, Nenavath Ramesh, et al. "Design of Modular Multi-Level Inverters by Using MATLAB." 2023 7th International Conference on Electronics, Communication and Aerospace Technology (ICECA). IEEE, 2023.

[5]. Betin, F., Pinchon, D., & Capolino, G. A. (2002). A time-varying sliding surface for robust position control of a DC motor drive. IEEE Transactions on Industrial Electronics, 49(2), 462-473.

[6]. Chenchireddy, Kalagotla, and Varghese Jegathesan (2023). "Cascaded H-bridge multilevel inverter-based DSTATCOM with an artificial neural fuzzy inference systembased controller." Indonesian Journal of Electrical Engineering and Computer Science 32(1): 43-51.

[7]. Chotai, J., &Narwekar, K. (2017, December). Modelling and position control of brushed DC motor. In 2017 international conference on advances in computing, communication and control (ICAC3) (pp. 1-5). IEEE.

[8]. Das, S., Qureshi, M. S., & Swarnkar, P. (2018). Design of integral sliding mode control for DC-DC converters. Materials Today: Proceedings, 5(2), 4290-4298.

[9]. Lai, M. F., Nakano, M., & Hsieh, G. C. (1996). Application of fuzzy logic in the phase-locked loop speed control of induction motor drive. IEEE Transactions on Industrial Electronics, 43(6), 630-639.

[10]. Liu, G., & Zhang, C. (2013). Sliding mode control of reaction flywheel-based brushless DC motor with buck converter. Chinese Journal of Aeronautics, 26(4), 967-975.

[11]. Chenchireddy, Kalagotla, and V. Jegathesan (2023). "ANFIS Based Reduce Device Count DSTATCOM." Journal of Applied Science and Engineering 26(11): 1657-1666.

[12]. Orlowska-Kowalska, T., & Szabat, K. (2004). Optimization of fuzzy-logic speed controller for DC drive system with elastic joints. IEEE Transactions on Industry Applications, 40(4): 1138-1144.

[13]. Chaudhary, H., Khatoon, S., & Singh, R. (2016, November). ANFIS based speed control of DC motor. In 2016 Second International Innovative Applications of Computational Intelligence on Power, Energy and Controls with their Impact on Humanity (CIPECH) (pp. 63-67). IEEE.

[14]. Chenchireddy, Kalagotla, et al. "Grid-Connected 3L-NPC Inverter with PI Controller Based on Space Vector Modulation." 2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS). Vol. 1. IEEE, 2023.

[15]. Jilani, A., Murawwat, S., & Jilani, S. O. (2015). Controlling speed of DC motor with fuzzy controller in comparison with ANFIS controller. Intelligent Control and Automation, 6(01), 64.

Exploring Cybersecurity Risks in Higher Education Environments with Machine Learning

Kenneth David Strang
W3-Research, Data Analytics
New York, 11011, USA
ORCID: 0000-0002-4333-4399
kenneth.strang@gmail.com

Narasimha Rao Vajjhala
University of New York Tirana
Tirana, Albania
ORCID: 0000-0002-8260-2392
narasimharaonarasimha@gmail.com

Abstract— **This paper uses unsupervised machine learning techniques (ML) to explore the risk of cybersecurity breaches or attacks in the higher education sector, including schools, universities, and support organizations. A large sample of higher education institutions (N=848) was analyzed by extracting hypertext data from the websites of educational institutions to identify links that may signal a future cybersecurity breach or attack. ML T-distributed Stochastic Neighbor Embedding (t-SNE) was used to train a model for identifying likely indicators of cybersecurity breach risks. The authors used additional techniques, including radial analysis and correspondence analysis, to visualize the cybersecurity breach signals in the data. In this way, decision-makers should be able to spot anomalies that indicate cybercrime activity or potential cyber-attacks. This paper also exposes the problems of using ML and how to use methods and data triangulation to check reliability and validity.**

Keywords— unsupervised machine learning, cybersecurity breaches, higher education sector, hypertext data analysis, t-distributed stochastic neighbor embedding, pervasive computing, cyber-attack indicators, radial analysis, correspondence analysis, anomaly detection, data triangulation.

I. Introduction

Cybercrime represents a significant financial burden to organizations and poses considerable detection challenges [1, 2]. Decision-making processes are imperative in complex scenarios, necessitating the employment of various analytical techniques. However, one of the significant challenges decision-makers face emerges when there is a suspicion of cybercrime within complex datasets [3]. This scenario demands an analytical approach that transcends simple binary evaluations of alternatives. Regrettably, the availability of methodologies capable of ranking choices and delineating improvements across a broad spectrum of variables and scenarios is limited. Existing techniques often restrict the data types that can be processed, accommodate only a finite set of independent or dependent variables, and typically categorize outcomes as positive or negative, rank performance metrics, or forecast optimal values post-model development. It is rare to find a single decision-making methodology that encompasses all these capabilities effectively.

Consider the complexity involved in a leader of a significant global company making decisions about closing or establishing various sections or sites, especially with the abundance of performance data at their disposal. How does such a leader pinpoint improvement in underperforming areas to avoid layoffs or new hires? Similarly, with the extensive information on marketing and benchmarks, how do students decide on the best universities for them? Or, how do regulatory bodies assess which entities under their watch are secure for public interaction, or conversely, identify which ones require enhancement and in what manner? These scenarios present intricate management research challenges that go beyond the realm of simple statistical correlations or predictive analytics for resolution.

This study seeks to demonstrate the application of cybersecurity management decision-making strategies within the higher education domain. It also aims to evaluate the pros and cons of employing a top-tier singular method for addressing management issues. This work will highlight critical elements for choosing and deploying managerial decision-making methodologies. It will detail how various relevant approaches operate and include practical examples from scholarly sources. The latter part of the document will focus on employing these methods to analyze American university websites.

This study examines explicitly the ML methodology, its foundational theories, benefits, and limitations within an empirical study comparing universities in higher education. It covers the diverse approaches to designing and implementing ML, offering guidance on selecting the most appropriate design based on the managerial research question and the data types. The central research question guiding this investigation explores how sampled universities can assess the risk of a cybersecurity breach by analyzing their websites. Furthermore, the study intends to detail the application of the ML method in such a context that is understandable to managers or individuals without a background in mathematics.

This study's findings and methodologies can be universally applied across various sectors and managerial challenges with similar decision-making variables. The primary audience for this research encompasses managerial decision-makers across all fields, extending from the realm of higher education to other industries. Such decision-makers range from executives, managers, and administrators to government officials and policy strategists. Additionally, a secondary audience includes consumers like students or clients of various services who face decisions amidst a plethora of complex numerical data for benchmarking or evaluation. Hence, this document is invaluable for anyone with intricate choices where numerous metrics are available yet challenging to sift through. It also holds significance for those in academic or professional pursuits within operations research or management science, offering insights into decision-making processes. Moreover, management science consultants and entities stand to gain substantially, as the strategies discussed are instrumental for benchmarking organizational performance against peers or competitors and uncovering avenues for enhancing productivity.

II. LITERATURE REVIEW

It is not unexpected that the roots of decision-making methods can be traced back to the foundational texts of management science. Rosacker and Olson [4] point out that initial decision-making strategies relied on qualitative data for assessing implicit or non-quantifiable factors through methods that appealed to visual cues and convincing narratives, including reliance on 'acts of faith' and 'instinct.' Early qualitative decision-making models incorporated approaches based on attributes like the 'sacred cow' (recommendations from top executives), operational necessity (emergencies), competitive necessity (sector demands), expansion of product lines (adaptation to market changes), and comparative assessments of advantages, such as the 'Q-Sort' technique for ranking staff priorities [5].

Various decision-making strategies have incorporated research methodologies that favor expert team-based brainstorming over individual decision-making to identify optimal solutions alongside consensus-building techniques such as nominal group or Delphi methods [6]. Many multi-criteria decision-making models and approaches have surfaced in financial and economic research, integrating financial metrics such as the payback period, average or internal rate of return, discounted cash flow, and profitability indices. Sophisticated methods have been introduced in finance, including Monte Carlo simulations, portfolio selection models, scenario planning, and sensitivity analysis.

Various weighting models have been introduced, each adhering to converting original numerical factors into more indicative indices. The analytical hierarchy process, a notable decision-making method developed by Saaty [7], facilitates the conversion of complex qualitative factors into numerical indices for prioritization through pairwise comparisons and normalization techniques. Statistical methodologies are frequently incorporated into these processes, particularly when dealing with quantitative data. A notable drawback of parametric statistical methods is their dependency on the assumption that sample data must follow a certain distribution, along with numerous other presuppositions. Nonparametric statistical methods also need help with the number of variables they can handle and the accuracy of their estimates. A common shortcoming across statistical approaches is their focus on providing answers to specific ranking or predictive research questions via hypothesis testing without offering insights on enhancing the influencing or outcome variables.

Universities have cybersecurity requirements due to the nature of their networks, the diverse range of users they cater to, and the valuable data they possess [8]. This data encompasses information as well as cutting-edge research. Safeguarding cybersecurity in environments is particularly challenging, given the increasing cyber threats for property, sensitive data, and critical infrastructure [9]. The global digitalization of universities has undoubtedly brought advantages like access to information learning materials and worldwide connectivity [10]. However, this digital transformation has also exposed institutions to various cybersecurity risks [11]. Some common threats universities face includes phishing attacks, ransomware incidents, distributed denial of service (DDoS) attacks, and insider threats.

Phishing attacks involve using fraudulent emails or messages that mimic legitimate institutions or individuals to trick victims into revealing sensitive information, such as login credentials or personal information [12, 13]. Universities are prime targets for phishing due to the sheer volume of users and their sensitive data [14]. Ransomware is malicious software designed to block access to a computer system or files until a sum of money is paid [15]. These attacks can encrypt crucial data and systems, crippling institutional operations. An instance of ransomware impacting universities is the attack on the University of California, San Francisco (UCSF) in June 2020. The university was forced to pay a ransom of over $1.14 million to regain access to encrypted data related to important academic research [16]. This event highlights ransomware's critical threat to research integrity and operational continuity in academic settings.

DDoS attacks aim to overwhelm a system's resources by flooding it with several requests, making the service unavailable to legitimate users [17]. DDoS attacks often target universities because they heavily rely on teaching, learning, and administrative services [18]. For example, in 2016, 63 UK universities had suffered from DDoS Attacks including, the University of Edinburgh faced a DDoS attack that disrupted its network, leading to access to its website and online resources [19, 20]. These kinds of attacks can impact operations, causing delays and difficulties for both students and faculty members. Insider threats come from individuals within the organization who may misuse their access to systems and data for malicious purposes [21]. This category includes intentional acts by disgruntled employees or students and unintentional incidents caused by negligence [22]. A notable example is the case at the University of Iowa, where a student manipulated his grades more than 90 times over 21 months by stealing professors' credentials. This incident shows the potential for academic fraud and highlights the critical need for robust access controls and monitoring within educational institutions.

The cybersecurity landscape of universities is uniquely challenging, shaped significantly by the complexity and openness of their IT infrastructures, the human factor, and resource constraints [23]. Each of these elements introduces vulnerabilities that cyber threats can exploit. Universities typically maintain an open, complex IT infrastructure to support various academic activities [24]. This openness is essential for academic freedom and collaboration, allowing students, faculty, and researchers easy access to information and systems from anywhere in the world. The complexity arises from the vast array of devices connected to the network, including the personal devices of students and staff and specialized equipment for research purposes. Additionally, universities often have decentralized IT policies, with individual departments managing their systems to some extent.

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a machine learning technique that helps simplify datasets by reducing their dimensionality [25, 26]. It is beneficial when trying to visualize datasets with variables. The primary objective of t-SNE is to transform data points from a space into a lower dimensional space [27]. By doing so, it arranges items in closer proximity and places dissimilar items further apart in the lower dimensional representation. This makes it a valuable tool for gaining insights into the structure of dimensional data. t-SNE begins by determining the likelihood that pairs of data points in the space are similar [28]. t-SNE considers each data point as the center of a distribution, calculating the probabilities of other points belonging to that distribution. This step generates a probability distribution for pairs of points in space. t-SNE then maps this data to a lower dimensional space and computes a similar probability distribution within this new space. The positions of points in the space are adjusted to minimize the Kullback-Leibler divergence between these two distributions. This ensures that similar points from the original high-dimensional space remain close together in their transformed lower-dimensional representation while dissimilar points are kept far apart.

III. METHODS

Scientific inquiry starts with formulating a research question and determining the explicit or implicit unit of analysis, which delineates the variables of interest and their data types. This study aims to evaluate quantitative data from multiple universities to address the managerial query: How do universities within the dataset rank in performance against established industry benchmarks, and what strategies could less efficient universities employ to enhance their performance? Here, the focus is on assessing university efficiency using predetermined variables for measurement. Except for the qualitative aspects of the university's identity and its geographical location, the variables in question are quantitative. Employing a comparative research design, this study analyzes the performance metrics of various universities. Should the objective have been merely to compare average values of specific metrics, such as student-faculty ratios, among the institutions, it would constitute a basic group comparison. For straightforward comparisons, statistical methods such as the T-Test (for two universities), ANOVA or MANOVA (for more than two universities), or CANOVA (if a numeric independent variable is known to influence another variable based on prior research) would be suitable. However, ANOVA and similar statistical methods are not equipped to address the secondary query regarding the improvement strategies for lower-performing institutions. Addressing this requires an operations research approach, specifically within linear programming, like machine learning.

The selection process for the sample utilized publicly available data from U.S. higher education institutions, resulting in a dataset comprising roughly 848 records from educational websites. The analysis focused on the individual college site level, leading to multiple entries for larger entities due to each campus providing separate data. The dataset featured over 4,000 elements identified as potential cybersecurity indicators, including URLs, XML embeddings, and various HTTP scripts. The sampling approach was initially purposive, aiming for a balanced representation of medium-sized public and private universities across different states and subsequently random. This decision was based on the presumption that universities of a smaller or larger size might display atypical metrics attributable to their distinct characteristics.

In contrast, medium-sized institutions were expected to offer more comparable data regarding efficiency and quality. The selection was narrowed down to 26 by excluding records lacking the required metrics, followed by the random selection of two private and two public universities from distinct states, resulting in a final sample comprising universities from Pennsylvania, Connecticut, New Jersey, and California. To maintain anonymity while illustrating the outcomes and interpretations of machine learning analyses, these institutions were anonymized with generic labels starting with 'Uni.' While ensuring confidentiality is essential, it was balanced with the scientific imperative of providing precise references to ensure transparency and reliability for the academic community and researchers worldwide, making the findings clear and verifiable.

IV. RESULTS AND DISCUSSION

The results of this study provide a compelling illustration of how ML can be leveraged to enhance the understanding of cybersecurity vulnerabilities within the educational sector. Through a detailed analysis of data from several university websites, the study identified significant cybersecurity indicators, including URLs, embedded XML, and HTTP scripts. Python-based ML algorithms enabled an examination of these indicators, culminating in a radial analysis highlighting the most critical vulnerabilities. This analysis was further supported by a comprehensive risk assessment, which employed a scoring system to quantify the likelihood of cybersecurity breaches. The findings revealed a diverse range of security postures across the sampled institutions, with the 'control' feature offering an insight into the aggregation of risk factors, thereby facilitating a targeted approach to cybersecurity enhancement.

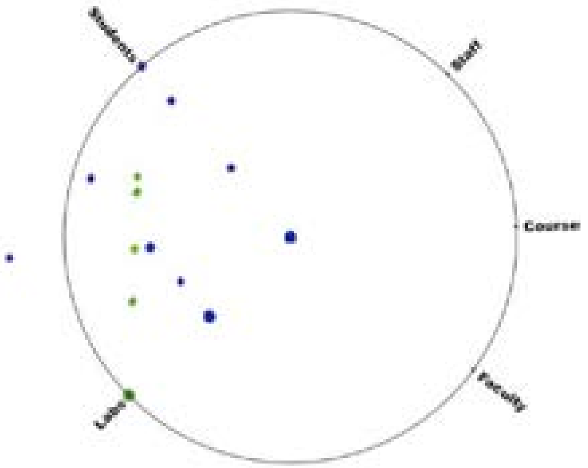


Fig. 1. Radial analysis of cybersecurity features related to potential breaches

Fig. 1 illustrates the top five features of potential cybersecurity breaches or shortcomings. These fields were based on using ML Python code to scan the website and evaluate the types of information found, including URLs, hypertext, and supporting scripts. Each field represents a different type of the above information. The Control field summarized the other indicators to calculate the combined cybersecurity breach risk due to any kind of risk found. This is the most useful feature in the ML analysis because it has identified the sample of potentially very high-risk components in the sample. The colors of the dots show the cybersecurity breach likelihood score on a scale from 0-3 (actually 1-3), where higher levels indicate higher risk, and the colors show this ranging from dark blue (low risk) to high risk (light yellow).

This multifaceted approach to risk assessment is further enhanced by the visual representation of risk levels through a color-coded system. Such a scheme not only simplifies the interpretation of results but also facilitates a quicker identification of areas requiring immediate attention. The gradation from dark blue to light yellow serves not just as a visual cue but also as a quantifiable measure of urgency, with the transition in colors reflecting an increasing scale of potential breach likelihood. This color mapping, grounded in the 1-3 risk score scale, allows for a more intuitive understanding of risk distribution across the analyzed websites, paving the way for targeted cybersecurity interventions.

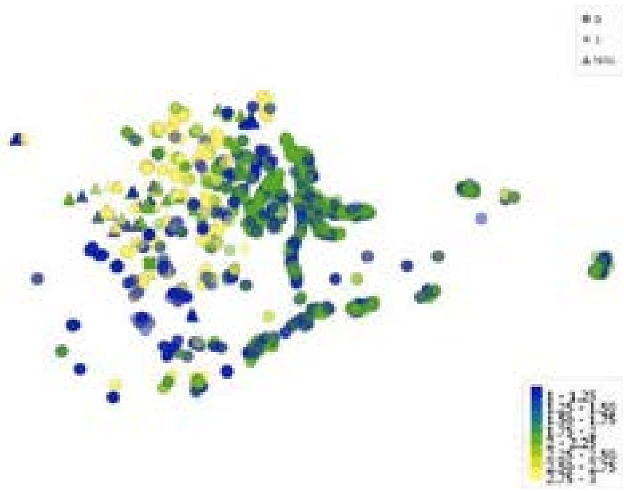


Fig. 2. ML t-SNE Analysis

Fig. 2 is the analysis done to explore the cybersecurity 'control' signal feature developed by examining various websites. Fig. 2 shows the cybersecurity breach likelihood score on a scale from 0-3 (actually 1-3), where higher levels indicate higher risk, and the colors show this ranging from dark blue (low risk) to high risk (light yellow). The '0' means these sites have no attacks, but the cybersecurity risk varies according to the color. The 'x' means there were one or more cybersecurity breaches due to malware discovered on the site (none were in this sample, so there is no 'x' here). The triangle means the site could not be verified because a URL was found that did not resolve to a working site; in other words, the NULL means a dead-end site or orphaned link. There is a general pattern of higher risks towards the middle and top.

The analysis presented in Fig. 2 further refines the understanding of the cybersecurity landscape across university websites by integrating a scoring mechanism that quantifies the vulnerability spectrum. This methodology, which distinguishes between varying degrees of risk from minimal to significant, employs a sophisticated visual taxonomy to differentiate between the operational status and security levels of the sites evaluated. Introducing specific symbols, such as the triangle, to indicate unverifiable areas due to non-resolving URLs adds layer of granularity to the analysis, allowing for a more detailed interpretation of the cybersecurity challenges these institutions face.

Importantly, the absence of the 'x' marker in the dataset underlines a positive note in this analysis—no instances of malware breaches were identified within the sample. This observation, however, does not detract from the critical need for vigilance, as the color-coded risk assessment reveals a discernible concentration of higher vulnerability scores in certain areas. Such a pattern suggests that while direct malware attacks may not have been detected, the potential for future breaches remains a significant concern, particularly in regions where the risk scores approach the upper end of the scale. The strategic placement of risks, notably more pronounced towards the central and upper sections of the visualization, hints at a correlation between the complexity of website features and the likelihood of security vulnerabilities. This insight underscores the necessity for a dynamic approach to cybersecurity, capable of adapting to the evolving digital ecosystem of higher education institutions. By mapping out the landscape of potential threats with such precision, the study highlights current security shortfalls. It lays the groundwork for developing more robust protective measures tailored to the specific vulnerabilities identified.

V. CONCLUSION

Future research should consider including a range of institutions, both large and small, to assess whether the identified cybersecurity risks and indicators hold true across different types and sizes of institutions. This would provide insights into the scalability and applicability of the machine learning models used and help understand cybersecurity risks. This could include analyzing network traffic, studying email patterns, and examining media presence. Refining the feature selection process is crucial to identify cybersecurity risk indicators. This involves exploring data processing techniques that effectively capture web content and infrastructure nuances. It is essential to investigate machine learning models and algorithms to determine their effectiveness in predicting cybersecurity breaches. Comparing machine learning models with deep learning approaches and ensemble methods can improve accuracy and reliability. Future studies could also include longitudinal studies to track the evolution of cybersecurity risks over time. Such studies enable us to identify trends and patterns, leading to measures in managing cybersecurity.

This paper has provided a novel examination of cybersecurity risks within the higher education sector by applying unsupervised machine learning techniques. Advanced analytical methods, such as t-SNE for model training and radial and correspondence analysis for data visualization, have enabled a deeper understanding of the cybersecurity landscape in this critical sector. The findings offer significant insights for decision-makers in educational institutions, highlighting the importance of proactive measures in identifying and mitigating potential cybersecurity threats. The ability to spot anomalies that could indicate cybercrime activity or impending cyber-attacks is crucial in an era where digital security is paramount.

Moreover, this study underscores the challenges inherent in leveraging ML for cybersecurity, particularly data reliability and model validity. The discussion on method and data triangulation presents a compelling case for the rigorous verification of ML outputs to ensure their accuracy and applicability. The exploration of cybersecurity risks in higher education environments with ML presents a promising avenue for enhancing digital security measures. However, it also calls for ongoing research and development to refine these techniques, improve data collection and processing methods, and fully integrate ethical considerations into ML applications. As cyber threats evolve, so must the approaches will evolve to detect and defend against them. This paper contributes to the growing body of knowledge in this field. It paves the way for further studies to build on these findings to develop more sophisticated and effective cybersecurity strategies for higher education.

REFERENCES

[1] V. Stanciu and A. Tinca, "Exploring cybercrime–realities and challenges," *Accounting and Management Information Systems,* vol. 16, no. 4, pp. 610-632, 2017.

[2] C. Horan and H. Saiedian, "Cyber crime investigation: Landscape, challenges, and future research directions," *Journal of Cybersecurity and Privacy,* vol. 1, no. 4, pp. 580-596, 2021.

[3] B. Dean and R. McDermott, "A research agenda to improve decision making in cyber security policy," *Penn St. JL & Int'l Aff.,* vol. 5, p. 29, 2017.

[4] K. M. Rosacker and D. L. Olson, "An empirical assessment of IT project selection and evaluation methods in state government," *Project Management Journal,* vol. 39, no. 1, pp. 49-58, 2008.

[5] S. Mantel, J. Meredith, and S. Shafer, *Project Management Core Textbook (With CD)*. John Wiley & Sons, 2006.

[6] K. D. Strang, "Applied financial non-linear programming models for decision making," *International Journal of Applied Decision Sciences,* vol. 5, no. 4, pp. 370-395, 2012.

[7] T. L. Saaty, "How to make a decision: the analytic hierarchy process," *Interfaces,* vol. 24, no. 6, pp. 19-43, 1994.

[8] M. Lehto, "Cyber security education and research in the Finland's universities and universities of applied sciences," in *Cyber Security and Threats: Concepts, Methodologies, Tools, and Applications*: IGI Global, 2018, pp. 248-267.

[9] U. P. D. Ani, H. He, and A. Tiwari, "Review of cybersecurity issues in industrial critical infrastructure: manufacturing in perspective," *Journal of Cyber Security Technology,* vol. 1, no. 1, pp. 32-74, 2017.

[10] A. Bozkurt *et al.*, "A global outlook to the interruption of education due to COVID-19 pandemic: Navigating in a time of uncertainty and crisis," *Asian Journal of Distance Education,* vol. 15, no. 1, pp. 1-126, 2020.

[11] S. Saeed, S. A. Altamimi, N. A. Alkayyal, E. Alshehri, and D. A. Alabbad, "Digital transformation and cybersecurity challenges for businesses resilience: Issues and recommendations," *Sensors,* vol. 23, no. 15, p. 6666, 2023.

[12] Z. Alkhalil, C. Hewage, L. Nawaf, and I. Khan, "Phishing attacks: A recent comprehensive study and a new anatomy," *Frontiers in Computer Science,* vol. 3, p. 563060, 2021.

[13] A. K. Jain and B. Gupta, "A survey of phishing attack techniques, defence mechanisms and open research challenges," *Enterprise Information Systems,* vol. 16, no. 4, pp. 527-565, 2022.

[14] K. Arlitsch and A. Edelman, "Staying safe: Cyber security for people and organizations," *Journal of Library Administration,* vol. 54, no. 1, pp. 46-56, 2014.

[15] A. Ali, "Ransomware: A research and a personal case study of dealing with this nasty malware," *Issues in Informing Science and Information Technology,* vol. 14, pp. 087-099, 2017.

[16] G. B. Liluashvili, "Cyber Risk Mitigation in Higher Education," *Law & World,* vol. 17, p. 15, 2021.

[17] S. T. Zargar, J. Joshi, and D. Tipper, "A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks," *IEEE communications surveys & tutorials,* vol. 15, no. 4, pp. 2046-2069, 2013.

[18] J. J. Santanna *et al.*, "Booters—An analysis of DDoS-as-a-service attacks," in *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)*, 2015: IEEE, pp. 243-251.

[19] Y. Bahtiri, E. BYTYÇI, F. Idrizi, S. Ismaili, and N. Sejfuli-Ramadani, "Cyber Security In Educational Institutions," *Journal of Natural Sciences and Mathematics of UT,* vol. 8, no. 15-16, pp. 307-314, 2023.

[20] B. Sieklik, R. Macfarlane, and W. J. Buchanan, "Evaluation of TFTP DDoS amplification attack," *computers & security,* vol. 57, pp. 67-92, 2016.

[21] I. H. Elifoglu, I. Abel, and Ö. Taşseven, "Minimizing insider threat risk with behavioral monitoring," *Review of business,* vol. 38, no. 2, pp. 61-73, 2018.

[22] A. Mahfuth, "Human factor as insider threat in organizations," *International Journal of Computer Science and Information Security (IJCSIS),* vol. 17, no. 12, p. December 2019 issue, 2019.

[23] M. Dunn Cavelty and A. Wenger, "Cyber security meets security politics: Complex technology, fragmented politics, and networked

science," *Contemporary Security Policy,* vol. 41, no. 1, pp. 5-32, 2020.

[24] D. J. Peurach and C. M. Neumerski, "Mixing metaphors: Building infrastructure for large scale school turnaround," *Journal of Educational Change,* vol. 16, pp. 379-420, 2015.

[25] D. M. Chan, R. Rao, F. Huang, and J. F. Canny, "GPU accelerated t-distributed stochastic neighbor embedding," *Journal of Parallel and Distributed Computing,* vol. 131, pp. 1-13, 2019.

[26] W. Zhu, Z. T. Webb, K. Mao, and J. Romagnoli, "A deep learning approach for process data visualization using t-distributed stochastic neighbor embedding," *Industrial & Engineering Chemistry Research,* vol. 58, no. 22, pp. 9564-9575, 2019.

[27] W. Song, L. Wang, P. Liu, and K.-K. R. Choo, "Improved t-SNE based manifold dimensional reduction for remote sensing data processing," *Multimedia Tools and Applications,* vol. 78, pp. 4311-4326, 2019.

[28] T. T. Cai and R. Ma, "Theoretical foundations of t-sne for visualizing high-dimensional clustered data," *The Journal of Machine Learning Research,* vol. 23, no. 1, pp. 13581-13634, 2022.

Improvements in Diagnosing Kawasaki Disease using Machine Learning Algorithms

Dr. Sheshang Degadwala
Professor & Head of Department
Department of Computer Engineering
Sigma University, Gujarat, India
sheshang13@gmail.com

Vinay Nagarad dasavandi krishnamurthy
Senior software engineer,
Bosch USA
Michigan-United state of America
Karnataka- India
vinayndk9@gmail.com

Dhairya Vyas
Research Scholar
The Maharaja Sayajirao University of Baroda,
Vadodara, Gujarat, India
dhairya.vyas-cse@msubaroda.ac.in

***Abstract*— This research explores the changing landscape of Kawasaki Disease (KD) diagnosis, specifically examining the Machine Learning (ML) algorithms to enhance diagnostic capabilities. KD, a complex pediatric vasculitis, presents diagnostic challenges due to its diverse clinical manifestations, often resulting in delayed treatment. The study investigates the incorporation of machine learning techniques to improve the accuracy and efficiency of KD diagnosis. Utilizing a comprehensive dataset encompassing various clinical parameters related to KD, this research aims to create a robust predictive model. The analysis involves exploring patterns and correlations within the dataset to identify key diagnostic markers for KD. The findings demonstrate the potential of machine learning algorithms to effectively distinguish KD cases from other febrile illnesses, assisting clinicians in making more informed and prompt decisions. The integration of these algorithms in the diagnostic process not only enhances accuracy but also shows promise in reducing diagnostic delays, ultimately contributing to improved patient outcomes.**

Keywords— Kawasaki Disease, Diagnosis, Machine Learning, Pediatric Vasculitis, Clinical Parameters, Predictive Model, Healthcare.

I. Introduction

Diagnosing Kawasaki Disease (KD), a complex pediatric vasculitic condition, has long been a concern in pediatric healthcare. The varied clinical presentation of KD poses challenges in early and accurate diagnosis, increasing the risks of delayed treatment and potential complications. Despite medical advancements, innovative solutions are needed to improve diagnostic precision and timeliness for KD. Recently, there has been a growing interest in exploring machine learning algorithms as a transformative tool in the diagnostic landscape. This research focuses on the intersection of Kawasaki Disease diagnosis and machine learning, aiming to contribute to the knowledge on enhancing pediatric healthcare.

Kawasaki Disease is characterized by acute inflammation of medium-sized arteries, primarily affecting children under five. The disease's manifestations include fever, mucocutaneous changes, lymphadenopathy, and potential cardiovascular complications, making its diagnosis intricate. Timely recognition is crucial, as delayed treatment may lead to severe cardiac issues. The current diagnostic paradigm relies on clinical expertise, but the complexity of KD often results in delayed or missed diagnoses.

Machine learning, a subset of artificial intelligence, has shown significant potential in healthcare, analysing vast datasets and identifying patterns for improved diagnostic accuracy. Its integration into the diagnostic process for KD holds promise for enhancing efficiency and reliability, potentially reducing diagnostic delays and improving patient outcomes.

This research addresses the need for developing more effective diagnostic approaches in Kawasaki Disease by exploring machine learning algorithms' application. Using a comprehensive dataset, the study aims to develop a predictive model discerning subtle patterns indicative of the disease. The investigation also examines the broader implications of integrating machine learning into pediatric healthcare, exploring its potential transformative impact on diagnostic practices and patient care. The research aims to contribute valuable insights to the ongoing discourse on the intersection of artificial intelligence and pediatric healthcare, specifically enhancing the diagnosis and management of Kawasaki Disease.

II. Related Works

The literature on Kawasaki Disease (KD) and its diagnosis through various methodologies reflects the evolving landscape of pediatric healthcare. H.-C. Kuo's comprehensive review provides insights into the diagnosis, progress, and treatment updates of KD [1]. The paper emphasizes the need for constant refinement in the understanding and management of this challenging condition. J. Y. Lam et al. propose a deep learning framework for image-based screening of KD, showcasing the potential of improving diagnostic processes [2]. H. Lee et al. contribute to the discourse with an explainable algorithm focused on distinguishing incomplete Kawasaki disease based on coronary artery lesions in echocardiographic imaging [3].

In the realm of applications, E. Xu, S. Nemati, and A. H. Tremoulet present a deep convolutional neural network for KD diagnosis, highlighting the algorithm's potential to enhance accuracy in identifying this complex condition [4]. C. M. Tsai et al. introduce a novel score system of blood tests to differentiate KD from febrile illnesses, offering an alternative approach to diagnosis [5]. V. Baronti and E. Araujo delve into the development of a diagnostic support expert system for symptomatic fever onset of KD, showcasing interdisciplinary efforts in advancing diagnostic tools [6]. J. Zhong et al. propose a unique approach by distinguishing KD from febrile infectious diseases using gene pair signatures, contributing to the molecular understanding of KD [7].

Furthermore, foundational works by A. Marchesi et al. [8], E. Marrani, J. C. Burns, and R. Cimaz [9], Y. Krishnamurthy [10], and J. W. Newburger, M. Takahashi, and J. C. Burns [11] provide essential guidelines, classifications, and clinical perspectives on KD. J. Sánchez-Manubens, R. Bou, and J. Anton address the diagnostic challenges and classification issues associated with KD [12][13]. J. Forsey and L. Mertens shed light on atypical KD, emphasizing the clinical challenges in its diagnosis [14]. Lastly, C. Manlhiot, K. Millar, F. Golding, and B. W. McCrindle contribute insights into coronary artery abnormalities based on z-scores after KD, offering valuable information for improved classification [15]. Collectively, this literature study underscores the interdisciplinary efforts and diverse approaches employed to advance the diagnosis and understanding of Kawasaki Disease.

III. Proposed Methodology

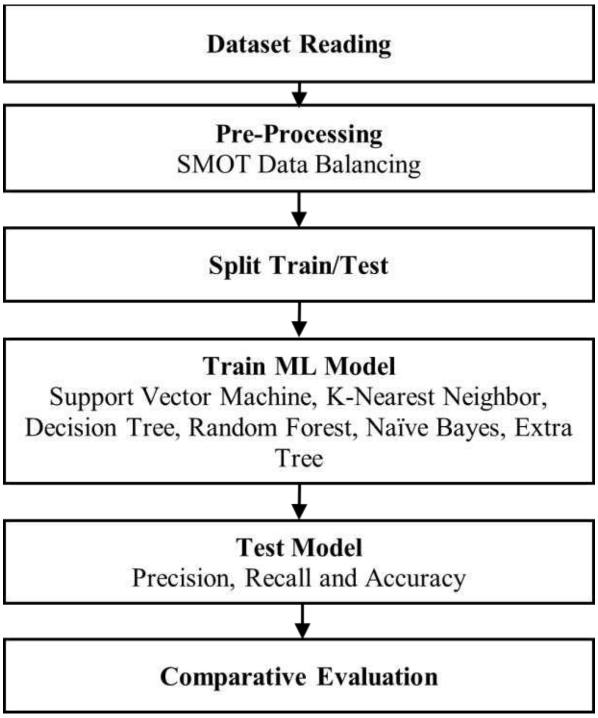


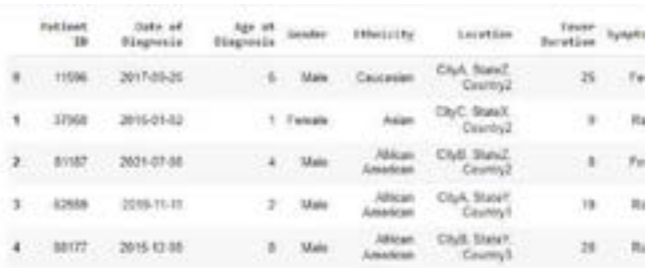
Fig. 1. Kawasaki Disease Classification System Flow

The classification process for Kawasaki Disease Classification, as depicted in Figure 1, progresses through a sequence of meticulous steps. Achieving successful classification involves a systematic approach, starting with a clear definition of the problem. The process includes collecting and preparing relevant data, with thorough cleaning and preprocessing being essential. To handle null values, functions like dropna() or fillna() in pandas for Python can be utilized. Balancing features through techniques like Synthetic Minority Oversampling Technique (SMOTE) contributes to the model's subsequent success. The dataset is then split into training and testing sets to accurately evaluate the model's performance, potentially incorporating a validation set for hyperparameter tuning. Selecting an appropriate classification algorithm, such as SVM, KNN, Decision Trees (DT), Random Forest (RF), or ensemble Extra Tree (ET), is a crucial step, with ensemble methods considered for improved performance. Training the model involves adjusting hyperparameters to optimize its capabilities. The model's performance is rigorously evaluated using metrics like accuracy, precision, recall, F1 score, and confusion matrices. Fine-tuning through hyperparameter adjustments is followed by validation on an independent dataset to ensure generalizability. Consideration is given to the interpretability of the model's decision-making process, especially when it is crucial for application understanding. Upon satisfactory results, the model is deployed for real-world use and integrated into the intended application or system. Continuous monitoring and updates are emphasized to adapt to changes in data patterns or shifts in the problem domain, ensuring the ongoing relevance and accuracy of the deployed model.

IV. Results And Findings

The dataset contains 2752 rows and 15 columns, encompassing crucial information on Kawasaki Disease (KD) cases. It includes Patient ID, Date of Diagnosis, Age at Diagnosis, Gender, Ethnicity, Location, Fever Duration, Symptoms, Laboratory Tests, Echocardiography, Treatment Approach, Clinical Outcomes, Complications, Follow-up Visits, and Long-Term Effects. This comprehensive dataset serves various purposes, offering valuable insights into the epidemiology, treatment patterns, and outcomes related to KD. With a focus on epidemiological insights, the dataset, comprising 2752 patient records, enables a thorough analysis of KD cases, revealing demographic patterns and potential risk factors.

Link: https://www.kaggle.com/datasets/cm037divya/kawasaki-dataset



	Patient ID	Date of Diagnosis	Age at Diagnosis	Gender	Ethnicity	Location	Fever Duration
0	[illegible]	[illegible]	5	Male	Caucasian	[illegible]	25
1	[illegible]	[illegible]	1	Female	Asian	[illegible]	[illegible]
2	[illegible]	[illegible]	4	Male	African American	[illegible]	[illegible]
3	[illegible]	[illegible]	2	Male	African American	[illegible]	[illegible]
4	[illegible]	[illegible]	[illegible]	Male	African American	[illegible]	[illegible]

Fig. 2. Dataset Read

```
[108] df.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 2752 entries, 0 to 2751
Data columns (total 13 columns):
 #   Column              Non-Null Count  Dtype
---  ------              --------------  -----
 0   Age at Diagnosis    2752 non-null   int64
 1   Gender              2752 non-null   object
 2   Ethnicity           2752 non-null   object
 3   Location            2752 non-null   object
 4   Fever Duration      2752 non-null   int64
 5   Symptoms            2752 non-null   object
 6   Laboratory Tests    2752 non-null   object
 7   Echocardiography    2752 non-null   object
 8   Treatment Approach  2752 non-null   object
 9   Clinical Outcomes   2752 non-null   object
 10  Complications       2752 non-null   object
 11  Follow-up Visits    2752 non-null   object
 12  Long-Term Effects   2752 non-null   object
dtypes: int64(2), object(11)
memory usage: 279.6+ KB
```

Fig. 3. After Pre-Processing

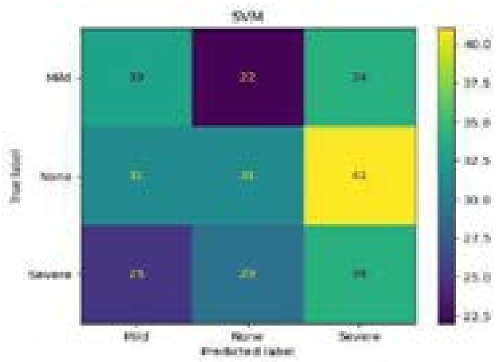
As Shown in Fig. 3. Pre-processing task null removal and data/label separation is done using python programming. The dataset remain after pre-process is 2752 rows.

```
[111] #Data Balancing over Sample
      from imblearn.over_sampling import SMOTE
      oversample = SMOTE()
      x1, y1 = oversample.fit_resample(x, y)
      y1.value_counts()
```

```
Mild      932
None      932
Severe    932
Name: Long-Term Effects, dtype: int64
```

Fig. 4. Data Balancing SMOT

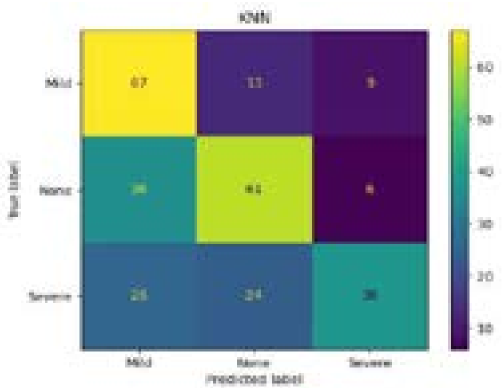
As Shown in Fig. 4. Data Balancing using SMOTE algorithm which is built in scikit-learn library in python programming.




```
              precision    recall  f1-score   support

        Mild       0.34      0.39      0.36        89
        None       0.43      0.32      0.37       103
      Severe       0.36      0.41      0.38        88

    accuracy                           0.37       280
   macro avg       0.38      0.37      0.37       280
weighted avg       0.38      0.37      0.37       280
```

Fig. 5. SVM Confusion Matrix

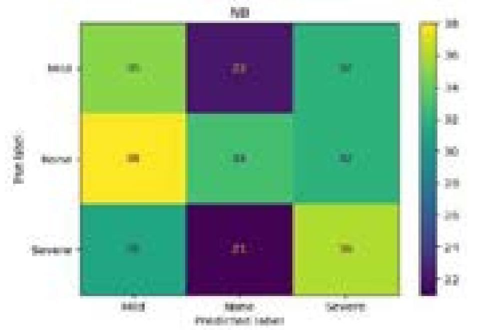
As Shown in Fig. 5 of SVM Confusion Matrix & Classification Report the class mostly is wrongly classified.




```
              precision    recall  f1-score   support

        Mild       0.34      0.39      0.36        89
        None       0.43      0.32      0.37       103
      Severe       0.36      0.41      0.38        88

    accuracy                           0.37       280
   macro avg       0.38      0.37      0.37       280
weighted avg       0.38      0.37      0.37       280
```

Fig. 6. KNN Confusion Matrix

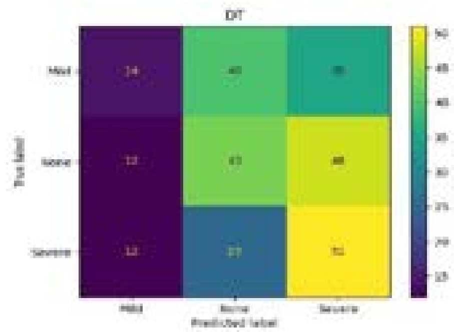
As Shown in Fig. 6 of KNN Confusion Matrix & Classification Report the class mostly is wrongly classified.




```
              precision    recall  f1-score   support

        Mild       0.34      0.39      0.36        89
        None       0.43      0.32      0.37       103
      Severe       0.36      0.41      0.38        88

    accuracy                           0.37       280
   macro avg       0.38      0.37      0.37       280
weighted avg       0.38      0.37      0.37       280
```

Fig. 7. NB Confusion Matrix

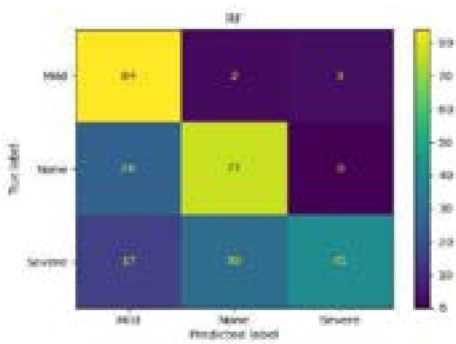
As Shown in Fig. 7 of NB Confusion Matrix & Classification Report the class mostly is wrongly classified.




```
              precision    recall  f1-score   support

        Mild       0.37      0.16      0.22        89
        None       0.40      0.42      0.41       103
      Severe       0.38      0.58      0.46        88

    accuracy                           0.39       280
   macro avg       0.38      0.38      0.36       280
weighted avg       0.38      0.39      0.36       280
```

Fig. 8. DT Confusion Matrix

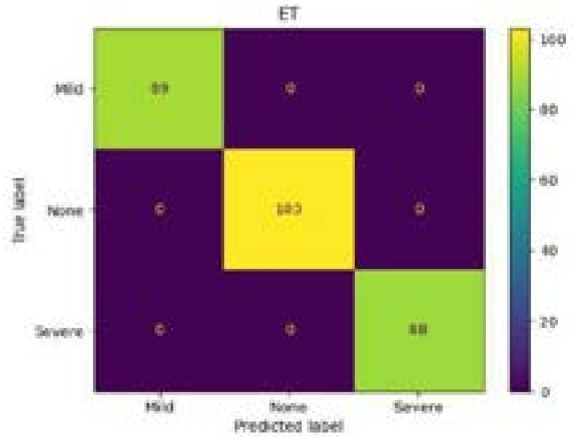
As Shown in Fig. 8 of DT Confusion Matrix & Classification Report the class mostly is wrongly classified.




```
              precision    recall  f1-score   support

        Mild       0.66      0.94      0.78        89
        None       0.71      0.75      0.73       103
      Severe       0.93      0.47      0.62        88

    accuracy                           0.72       280
   macro avg       0.77      0.72      0.71       280
weighted avg       0.76      0.72      0.71       280
```

Fig. 9. RF Confusion Matrix

As Shown in Fig. 9 of RF Confusion Matrix & Classification Report the class mostly is wrongly classified.



	precision	recall	f1-score	support
Mild	1.00	1.00	1.00	89
None	1.00	1.00	1.00	103
Severe	1.00	1.00	1.00	88
accuracy			1.00	280
macro avg	1.00	1.00	1.00	280
weighted avg	1.00	1.00	1.00	280

Fig. 10. ET Confusion Matrix

As Shown in Fig. 10 of Ensemble Extra Tree Confusion Matrix & Classification Report the no class is wrongly classified. Table 1 shows comparative analysis of all Ensemble Models, among them Ensemble Extra Tree perform best.

TABLE I. MODELS COMPARISION

Models	Accuracy (%)	Precision (%)	Recall (%)	F1 (%)
SVM	0.37	0.38	0.37	0.37
KNN	0.37	0.38	0.37	0.37
Navier Bayers	0.37	0.38	0.37	0.37
Dicision Tree	0.36	0.38	0.38	0.39
Random Forest	0.71	0.77	0.72	0.72
Extra Tree	0.99	0.99	0.99	0.99

CONCLUSION

In conclusion, the assessment of various models for Kawasaki Disease (KD) diagnosis reveals distinct performance metrics across different algorithms. Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Naive Bayes consistently demonstrate comparable accuracy, precision, recall, and F1 scores, all hovering around 0.37. Decision Tree exhibits a slightly lower accuracy at 0.36, with precision and recall at 0.38 and 0.38, respectively, while achieving a slightly higher F1 score of 0.39. Notably, Random Forest stands out among the assessed models, boasting a substantial accuracy of 0.71, with commendable precision, recall, and F1 scores at 0.77, 0.72, and 0.72, respectively. The Extra Tree model demonstrates remarkable performance, achieving near-perfect scores across all metrics with an accuracy of 0.99 and precision, recall, and F1 scores each at 0.99. These findings emphasize the importance of choosing an appropriate model for KD diagnosis, with Random Forest and Extra Tree emerging as promising candidates for further exploration and potential deployment in clinical settings. However, careful consideration of the specific requirements and nuances of the diagnostic task is essential in optimizing model selection for reliable and accurate KD diagnosis.

REFERENCES

[1] H.-C. Kuo, "Diagnosis, Progress, and Treatment Update of Kawasaki Disease," International Journal of Molecular Sciences, vol. 24, no. 18. 2023. doi: 10.3390/ijms241813948.

[2] J. Y. Lam et al., "A Deep Learning Framework for Image-Based Screening of Kawasaki Disease," in 2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2023, pp. 1–4. doi: 10.1109/EMBC40787.2023.10340801.

[3] H. Lee, Y. Eun, J. Y. Hwang, and L. Y. Eun, "Explainable deep learning algorithm for distinguishing incomplete Kawasaki disease by coronary artery lesions on echocardiographic imaging," Computer Methods and Programs in Biomedicine, vol. 223, p. 106970, 2022.

[4] E. Xu, S. Nemati, and A. H. Tremoulet, "A deep convolutional neural network for Kawasaki disease diagnosis," Scientific Reports, vol. 12, no. 1, p. 11438, 2022, doi: 10.1038/s41598-022-15495-x.

[5] C. M. Tsai et al., "A novel score system of blood tests for differentiating Kawasaki disease from febrile children," PLoS ONE, vol. 16, no. 1 January, pp. 1–13, 2021, doi: 10.1371/journal.pone.0244721.

[6] V. Baronti and E. Araujo, "Diagnosing Support Expert System for Symptomatic Fever Onset of Kawasaki Disease," in 2021 International Conference on Decision Aid Sciences and Application (DASA), 2021, pp. 706–711. doi: 10.1109/DASA53625.2021.9682283.

[7] J. Zhong et al., "Distinguishing Kawasaki Disease from Febrile Infectious Disease Using Gene Pair Signatures," BioMed Research International, vol. 2020, p. 6539398, 2020, doi: 10.1155/2020/6539398.

[8] A. Marchesi et al., "Kawasaki disease: guidelines of the Italian Society of Pediatrics, part I - definition, epidemiology, etiopathogenesis, clinical expression and management of the acute phase," Italian Journal of Pediatrics, vol. 44, no. 1, p. 102, 2018, doi: 10.1186/s13052-018-0536-3.

[9] E. Marrani, J. C. Burns, and R. Cimaz, "How Should We Classify Kawasaki Disease?," Frontiers in immunology, vol. 9, p. 2974, 2018, doi: 10.3389/fimmu.2018.02974.

[10] Y. Krishnamurthy, "Kawasaki Disease BT - Adult Congenital Heart Disease in Clinical Practice," in Circulation American Heart Association, D. DeFaria Yeh and A. Bhatt, Eds. Cham: Springer International Publishing, 2018, pp. 461–474. doi: 10.1007/978-3-319-67420-9_32.

[11] J. W. Newburger, M. Takahashi, and J. C. Burns, "Kawasaki Disease," Journal of the American College of Cardiology, vol. 67, no. 14, pp. 1738–1749, 2016, doi: https://doi.org/10.1016/j.jacc.2015.12.073.

[12] J. Sánchez-Manubens, R. Bou, and J. Anton, "Diagnosis and classification of Kawasaki disease.," Journal of autoimmunity, vol. 48–49, pp. 113–117, 2014, doi: 10.1016/j.jaut.2014.01.010.

[13] J. Sánchez-Manubens, R. Bou, and J. Anton, "Diagnosis and classification of Kawasaki disease," Journal of Autoimmunity, vol. 48–49, pp. 113–117, 2014, doi: https://doi.org/10.1016/j.jaut.2014.01.010.

[14] J. Forsey and L. Mertens, "Atypical Kawasaki disease—a clinical challenge," European Journal of Pediatrics, vol. 171, no. 4, pp. 609–611, 2012, doi: 10.1007/s00431-011-1629-9.

[15] C. Manlhiot, K. Millar, F. Golding, and B. W. McCrindle, "Improved Classification of Coronary Artery Abnormalities Based Only on Coronary Artery z-Scores After Kawasaki Disease," Pediatric Cardiology, vol. 31, no. 2, pp. 242–249, 2010, doi: 10.1007/s00246-009-9599-7.

Improved Generative Adversarial Network for Phishing Attack Detection

Shammi L
Research Scholar, Department of CSE,
Presidency University, Bengaluru, Karnataka
shammiazhar@gmail.com

Dr. C Emilin Shyni
School of CSE and IS
Presidency University, Bengaluru, Karnataka
emilinshyni2020@gmail.com

***Abstract*— A general form of attack that is happening over the internet is called phishing which could lead to identity theft and financial damages. Due to the increase of online electronic services and payment systems, the demand for accurate phishing detection tools has risen in recent times. However, the models often lead to high false detection rates. This research work introduces an Improved Generative Adversarial Network-based phishing attack detection, which has mainly two stages such as pre-processing and attack detection. Initially, for data pre-processing, a min-max normalization process is used. Following that, the attack detection process is done via Improved GAN, where a new discriminator loss function is adopted to enhance the detection performance. Finally, the performance of the proposed work is validated in terms of different performance measures.**

Keywords—*Phishing attack, Improved Generative Adversarial Network (IGAN), Attack detection, Min-max normalization.*

I. INTRODUCTION

Generally, phishing is a kind of social engineering attack. The users are tricked into disclosing their credentials like credit cards and passwords, along with other confidential information. It shows a great impact on individuals and industries in terms of personal and financial data. The increasing count of phishing attacks leads to the requirement for efficient attack detection approaches. Consequently, the victims are warned, while become a phishing campaign's target, to avoid any sensitive data loss [1] [2] [3].

Phishing is a dangerous cyber-attack that can result in financial losses and identity theft. The proliferation of online electronic services and payment systems has increased the need for high-accuracy phishing detection software. Because most phishing detection methods rely on elements found in webpage content, it requires crawling the website and using third-party services. Relying solely on webpage content-related criteria results in high false detection rates and poor detection accuracy. Deep learning has gained popularity recently as a method for identifying fraudulent websites. This phishing attack detection approach includes two kinds of approaches, which are software-based and human-based approaches. To enhance the end-user's knowledge and assist them in making good decisions while facing a phishing website [4] [5], human-based models are utilized. On the other hand, diverse techniques are adopted by the software-based models to decide whether a website is legitimate or phishing, without the interference of the end-user. The latter model has 5 types, which were, heuristic, visual similarity, blacklist/whitelist, ML, and DL [6] [7] [8].

The blacklist/whitelist-dependent approaches are mainly based on a list of recognized phishing websites, which includes data such as IP addresses, phishing URLs, and so on. Constantly updating this list is also crucial in the blacklist/whitelist-dependent approaches [9] [10]. Using diverse features, a phishing webpage's visual similarity is compared with its respective legitimate webpage, in this visual similarity-based method [11][12]. A webpage has been considered phishing if the similarity is greater than the preset threshold. The heuristic-dependent methods mainly rely on the characteristics of a phishing web page, experts' prior knowledge, or the similarity among the phishing web pages. Phishing detection has been considered a binary classification problem in the case of ML approaches. In the case of DL approaches, the detection is conducted with three steps, designing of DL model, model input selection, and analyzing the features used to categorize the websites [13][14][15]. However, only limited work is been done by GAN [16] architecture. On the other hand, the generative adversarial network (GAN) has received less attention. To address this, this work concentrates on proposing an improved GAN for automated phishing attack detection.

- To propose an effective phishing attack detection system using Improved GAN architecture adopting with trimmed factor-based loss function for accurate detection outcomes.

This work has been arranged as follows: a few recent publications related to Phishing attack detection have been reviewed in section II, section III details the proposed IGAN-based phishing attack detection system, section IV provides this work's implementation outcomes, and this work concludes in section V.

II. LITERATURE SURVEY

Some recent publications related to Phishing attack detection have been reviewed below

A phishing detection model called PDGAN was developed by Saad Al-Ahmadi *et al* [17] in 2022. For attaining reliable performance, this PDGAN depends on a website's URL. For synthetic phishing URL generation, LSTM was utilized. To categorize whether that URL was legitimate or phishing, CNN was utilized. For experiments, DomCop and Phish Tank datasets are utilized and the detection accuracy and precision of 97.58% and 98.02% were attained by PDGAN.

An ML-based phishing detection model was proposed by Ala Mughaid *et al* [18] in 2022. Initially, the data set was partitioned into training and testing data. With the training the data, the model was trained, and using the test data, the results were validated. With 3 diverse datasets, this approach was conducted, and based on the email text characteristics and other features, the attacks are categorized into non-phishing or phishing. With a boosted decision tree, accuracy of 0.88, 1.00, and 0.97 was attained.

In 2023, Muhammad Waqas Shaukat *et al* [19] proposed a multilayered ML-based approach. For phishing detection, this approach categorizes the websites and advertisements into phishing and legitimate websites. Based on the prominent features, diverse ML and DL models are trained. As per the URL features, the websites are categorized as phishing or legitimate by this predictive model. Outstanding results were attained by these models.

For fraudulent URL identification, SK Hasane Ahammad *et al* [20] proposed an ML-based model in 2022. In this work, RF, LGBM, DT, SVM, and LR are utilized. The URL's domain-based and linguistic properties are investigated. From these algorithms, LGBM attained the best outcomes. However, for accuracy enhancement, more features needed to be added since some URLs in the dataset were not present in the database.

An ML-based phishing detection system was proposed in 2023, by Abdul Karim *et al* [21]. This model was named the LSD Ensemble model. This model was compared with DT, RL, SVM, RF, GBM, SVM, K-Neighbour, NB, and hybrid (LR+SVM+DT) models. Higher performance was attained by this LSD Ensemble model. Furthermore, this model utilized the cross-fold validation along with canopy feature selection and Grid search hyperparameter optimization approach.

TABLE I: ADVANTAGES AND CHALLENGES OF SOME RECENT PHISHING ATTACK DETECTION MODELS

Citation	Techniques	Advantages	Challenges
Saad Al-Ahmadi et al [17]	GAN, LSTM,and CNN	Higher detection accuracy (97.58%) and precision (98.02%) were obtained	Depending on the batch size, the accuracy and loss ratings also change.
Ala Mughaid*et al* [18]	ML	For 50 features, it provides the highest accuracy	Limitation in finding the predefined dataset and while using 20 features, it cannot find the phishing attacks
Muhammad Waqas Shaukat *et al* [19]	ML	The promising results were attained by XGBoost	To provide a more robust detection process, this approach should be further optimized
SK Hasane Ahammad *et al* [20]	ML models (LGBM, RF, DT, LR and SVM)	Best outcomes were provided by LGBM	For accuracy enhancement, more features needed to be added since some URLs in the dataset were not present in the whois database
Abdul Karim *et al* [21]	Hybrid LSD model	Effectively detecting the Phishing attacks	To prevent and detect the attacks more effectively, this should be merged with list-based ML-dependent systems.

A.Problem Statement

Recently, governments, internet users, and service-providing organizations have faced an important attack named phishing attack. The client's sensitive data is collected by the attackers with the utilization of fake websites or spoofed emails. Diverse researches are conducted related to the detection of phishing attacks; however, also face challenges. Recently, the GAN [17] model was utilized for the phishing attack detection. Although higher detection accuracy (97.58%) and precision (98.02%) were obtained by this approach, the accuracy rating was greatly dependent on the batch size. Furthermore, DL models were utilized in [18], which provides higher accuracy for more features than was while using 50 features. Although it can provide sufficient accuracy for 20 features, these outcomes are not effective enough for phishing email detection. It's another drawback was finding the attack in the predefined dataset. For effective phishing attack detection, ML models were utilized in [19]. This work proved that promising results were attained by XGBoost. Although this model was optimized in terms of computational efficiency, to provide a more robust detection process, this approach should be further optimized. Similarly, ML models such as LGBM, RF, DT, LR, and SVM are utilized in [20]. Although the best outcomes were provided by LGBM, for accuracy enhancement, more features needed to be added since some URLs in the dataset were not present in the Who is database. Moreover, the Hybrid LSD model [21] was developed to effectively detect Phishing attacks. However, to prevent and detect the attacks more effectively, this should be merged with list-based ML-dependent systems (Table I). To tackle these shortcomings, new methods for phishing attacks should be developed. Consequently, an Improved GAN-based phishing attack detection model is developed in this work, which is detailed below.

III. PROPOSED IGAN-BASED PHISHING ATTACK DETECTION

Nowadays, for social interactions, the internet has been considered a powerful channel. Since, people are highly dependent on digital platforms, which pave the way for data theft or attacks. From online platforms like online businesses, online classrooms, online banking, e-commerce, digital marketplaces, etc, the users' credentials are stolen by the attacker, and this is known as a phishing attack. For phishing webpage detection, many tools are introduced by researchers, which are white-list, antivirus software, and blacklist. To penetrate the cyber defense, attackers utilize creative ways. Some detection models require high computational power to compute the features obtained from different sources. Hence, there need for a robust knowledge-based model to categorize the attack and nonattack data. In this work, an IGAN-based phishing attack detection model is introduced, which includes two stages: pre-processing and attack detection. Initially, for pre-processing, a min-max normalization process is utilized.

Afterwards, an IGAN-based phishing attack detection process is done, which is detailed below and its architecture is displayed in Fig 1.

A. Pre-processing

Initially, the input data IP_D is subjected to the pre-processing phase, where min-max normalization is utilized. A general technique utilized in data pre-processing is named min-max normalization. This transforms a dataset's value into a common scale. To preserve the relationships among the original values, have selected the min-max normalization process In the work, min-max normalization [22] performed a linear transformation on the IP_D and obtained the scaled data in the range (0, 1), that is numerically displayed in Eq. (1).

$$IP_{DS} = \frac{A - A_{\min}}{A_{\max} - A_{\min}} \qquad (1)$$

Here, the scaled data is denoted as IP_{DS} and A is the data. Furthermore, $A_{\max}$ and $A_{\min}$ denotes the maximum and minimum values of data. Subsequently, the obtained IP_{DS} is subjected to Improved GAN for detecting the attacks.

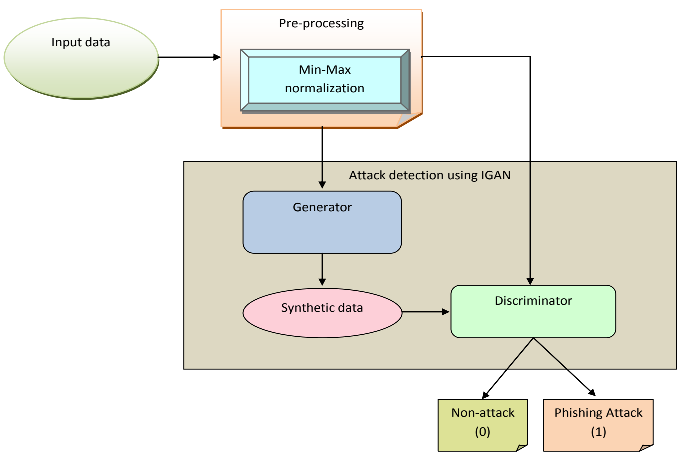


Fig.1. Architecture of IGAN-based phishing attack detection system

B. Attack detection using Improved GAN (IGAN)

The obtained IP_{DS} subjected to IGAN.IGAN is the unsupervised learning model that includes two NNs, a generator (GE) and a discriminator (DI). Both GE and DI models are probabilistic.

Generator (GE):

The IGAN's generator network includes the layer arrangement such as input layer, dense layer, leaky ReLU, reshape function, 2D convolution transpose, leaky ReLU, 2D convolution transpose, leaky ReLU, and afterwards a 2D convolution layer, that is displayed in Fig 2. Also, the reshape function has the size of (7, 7,128).

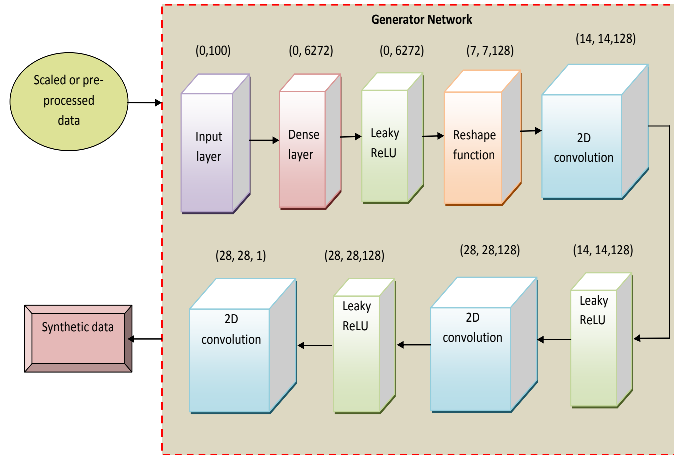


Fig.2. Architecture of IGAN's generator network

For capturing the data distribution, a generative model named generator is trained. From the learned distribution, the generator can output the generative and synthetic samples. To ensure the generated output sample's diversity, random noise(z) is present in this network. For learning the input data distribution, the joint probability distribution is applied by generative models using Eq. (2). Learning the data distribution facilitates the model to retrieve more prominent features and identify the data generation process. Therefore, artificial data having the identical distribution as the real data is provided by the generative models. These obtained synthetic data are plausible and dissimilar from the domain's real data.

$$p(IP_{DS}, Y) = p(IP_{DS}) \times p(Y) \qquad (2)$$

Where, p is probability and IP_{DS} is the input and Y denotes the class labels.

Discriminator (DI): IGAN's discriminator network includes the following layer arrangement. Firstly, the input layer, 2D convolutional layer, LeakyReLU, 2D convolutional, LeakyReLU, 2D convolutional, LeakyReLU, flatten, dropout and dense layer. The activation utilized in the network is lambda. The discriminator network's architecture is seen in Fig 3.

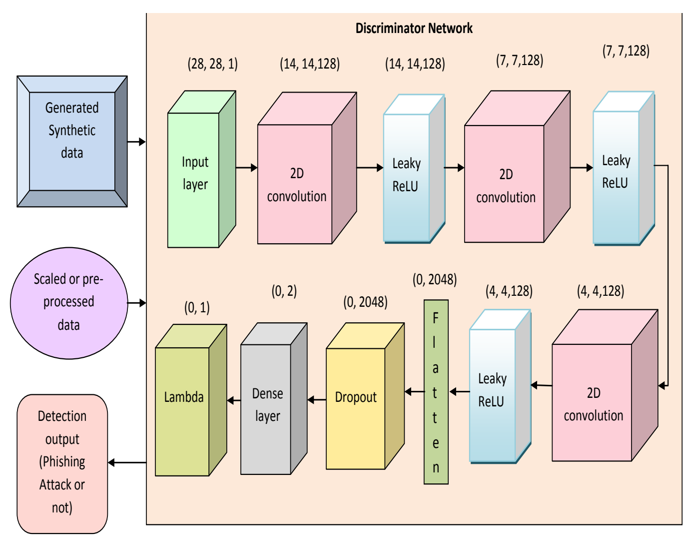


Fig.3. Architecture of IGAN's Discriminator network

A discriminative model has been trained to identify in which distribution the input data belongs. It helps to discover whether belong to real or artificial data distribution. Also, GAN's total performance can be enhanced by DI. GAN's training process has been considered as a min-max 2-player game.

To learn how to map input (IP_{DS})to their class labels (y), the conditional probability distribution is applied by the discriminative models using Eq. (3).

$$p\left(IP_{DS} \mid Y\right)=\frac{p\left(IP_{DS}, Y\right)}{p(Y)} \tag{3}$$

For maximizing the DI loss, the GE has been trained. The generator and discriminator losses of conventional GAN is in Eq. (4) and Eq. (5)

$$L^{(GE)}=\min\left\lfloor \log DI\left(IP_{DS}\right)+\log\left(1-DI\left(GE(z)\right)\right)\right\rfloor \tag{4}$$

$$L^{(DI)}=\max\left\lfloor \log DI\left(IP_{DS}\right)+\log\left(1-DI\left(GE(z)\right)\right)\right\rfloor \tag{5}$$

Here, $L^{(GE)}$ and $L^{(DI)}$ are generator and discriminator losses. Conventional discriminator loss function (in Eq. (5)), generally suffers from vanishing gradient issues and training instability. This will affect the detection performance of IGAN. For that reason, IGAN uses the trimmed factor-based loss function, which is mathematically expressed in Eq. (6). This trimming factor-based loss function (E_{TFL}) can provide accurate loss value.

$$E_{TFL}=-\frac{1}{H}\sum_{i=1}^{H} Q_{i:N} \tag{6}$$

where, H is the trimming factor that is considered during the training loss computation and it evaluates the smallest residual's quantity, N is the training set count.

This trimming factor-based loss function has additional functions based on constraints including data-specific characteristics and domain-specific knowledge. The model's ability to capture the relevant pattern in the data can be improved by this incorporation of a trimming factor-based loss function. Also, this loss function can help in ensuring the fair representation of all classes during training. For each observation, ordered losses are $Q_{1:N} \leq \ldots Q_{N:N}$, which can be written as in Eq. (7)

$$Q_i=\sum_{c=1}^{C} \lambda . PR_{ic} \log\left(\lambda T_{ic}\right) \tag{7}$$

Here, PR_{ic} symbolizes the predicted probabilities and C symbolizes the number of classes. Also, T_{ic} symbolizes the true values and λ denotes the robust factor, which can be evaluated using Eq. (8). Where δ =1.

$$\lambda=\delta\left(\left|PR_{ic}-T_{ic}\right|-\frac{1}{2}\delta\right) \tag{8}$$

This λ attains a smooth loss. Two versions of trimmed loss functions were tested in proposed work, which were when h=0.7 and h=0.9.

Finally, the discriminator provides the detection output based on the class labels, which might be 0 or 1 which means the class 0 indicates a non-attack and 1 indicates a phishing attack,

IV. RESULTS AND DISCUSSION

A. Simulation Procedure

The proposed phishing attack detection was simulated using PYTHON, specifically with "PYTHON 3.7." Further, the processor utilized was "11th Gen Intel (IR) Core (TM) i5-1135G7 @ 2.40GHz 2.42 GHz. Additionally, the system had "16.0 GB" of installed RAM. Moreover, the evaluation of phishing attack detection was carried out using the Phishing Dataset for Machine Learning [23].

B. Dataset Description

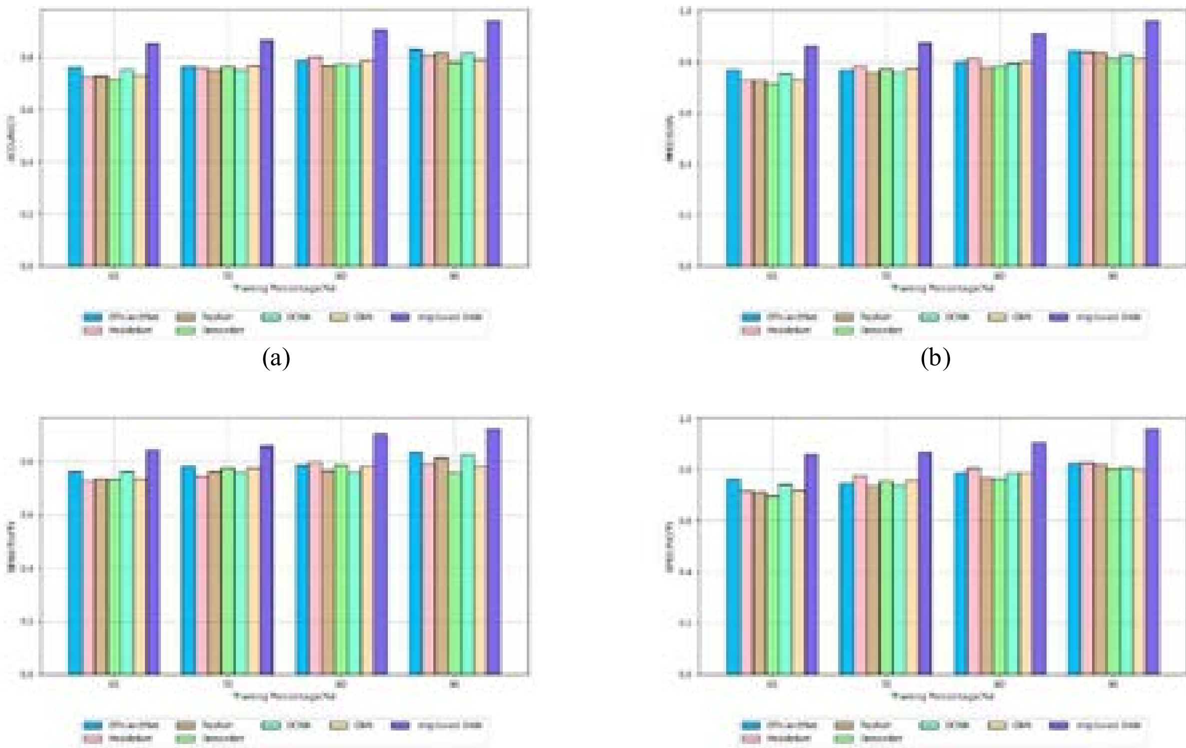
The dataset comprises 48 features extracted from a collection of 5000 phishing web pages and 5000 legitimate web pages. These web pages were acquired during two periods: from January to May 2015 and from May to June 2017. To enhance the extraction process, an advanced technique utilizing the browser automation framework, specifically Selenium WebDriver, was employed. This method proves to be more accurate and resilient when compared to the parsing approach reliant on regular expressions.

C. Performance Analysis

A thorough comparative analysis was conducted to evaluate the efficacy of the Improved GAN method in contrast to conventional approaches for detecting phishing attacks. This comprehensive assessment considered key metrics including Sensitivity, False Negative Rate (FNR), Negative Predictive Value (NPV), Specificity, F-measure, Precision, False Positive Rate (FPR), Matthews Correlation Coefficient (MCC), and Accuracy. Moreover, the performance of the Improved GAN scheme was compared with that of traditional approaches, including EfficientNet, MobileNet, ResNet, DenseNet, DCNN, and CNN [17].

D. Comparative Analysis on Positive Metric

In the context of phishing attack detection, a comparative analysis has been conducted to evaluate the Improved GAN strategy against conventional methods. Figure 4 illustrates the positive metric, a crucial factor in gauging the efficacy of the detection technique. The objective is to maximize the positive metric value for the efficient and effective identification of phishing attacks. To benchmark the Improved GAN strategy, a thorough comparison is made with established models such as EfficientNet, MobileNet, ResNet, DenseNet, DCNN, and CNN. Specifically, when considering a training percentage of 90, the performance of the Improved GAN model stands out with a notable detection accuracy of 0.938. In contrast, conventional methods demonstrate comparatively lower accuracy ratings: EfficientNet at 0.823, MobileNet at 0.809, ResNet at 0.816, DenseNet at 0.783, DCNN at 0.837, and CNN at 0.796, respectively. Furthermore, an additional noteworthy observation is the exceptional sensitivity achieved by the Improved GAN scheme, reaching a peak value of 0.895 at a training percentage of 80. This superior sensitivity performance surpasses that of established models such as Efficient Net, Mobile Net, ResNet, DenseNet, DCNN, and CNN, respectively. The results underscore the potential of the Improved GAN approach as a more effective solution in the realm of phishing attack detection, showcasing its ability to outperform established models under the specified training conditions.

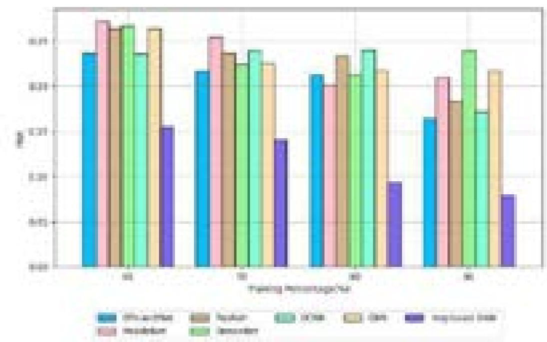


(a) (b) (c) (d)

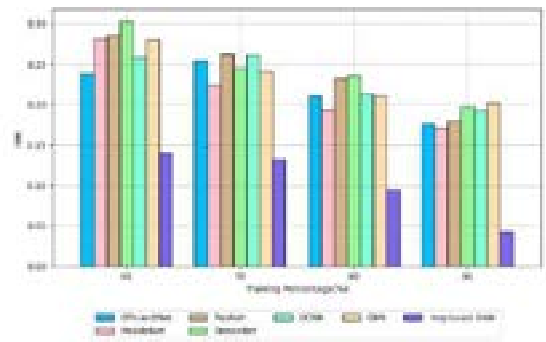
Fig.4. Assessment of Improved GAN and conventional strategies using Positive Metric

E. Comparative Analysis on Negative and Other Metrics

In the realm of phishing attack detection, the evaluation extends beyond positive metrics, encompassing negative metrics and other performance metrics to provide a comprehensive understanding of the Improved GAN model's efficacy. In this analysis, the Improved GAN model's performance is systematically contrasted with that of conventional methods, with a focus on minimizing negative values while concurrently enhancing other metrics. Figure 5 and Figure 6 serve as visual representations, highlighting the comparison of the Improved GAN model with the established methodologies. The objective is to showcase the Improved GAN model's proficiency in mitigating false negatives and optimizing other metrics critical to the accurate identification of phishing threats. Furthermore, the FPR of the Improved GAN scheme stands at 0.132 for a training percentage of 70. In contrast, established models exhibit higher FPR values, with Efficient Net at 0.254, Mobile Net at 0.223, ResNet at 0.262, Dense Net at 0.244, DCNN at 0.261, and CNN at 0.240, respectively. This discrepancy underscores the superior performance of the Improved GAN scheme in minimizing false positives compared to conventional methods. Similarly, the NPV of the Improved GAN approach is 0.913, whereas the Efficient Net, Mobile Net, ResNet, DenseNet, DCNN and CNN yielded the greatest NPV ratings of 0.813, 0.775, 0.795, 0.745, 0.803 and 0.762, respectively. Furthermore, the Improved GAN approach consistently achieved superior ratings compared to conventional methodologies across all metrics, ensuring precise detection of phishing attacks.

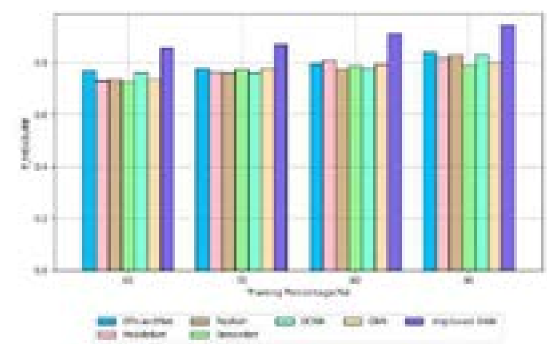


(a)

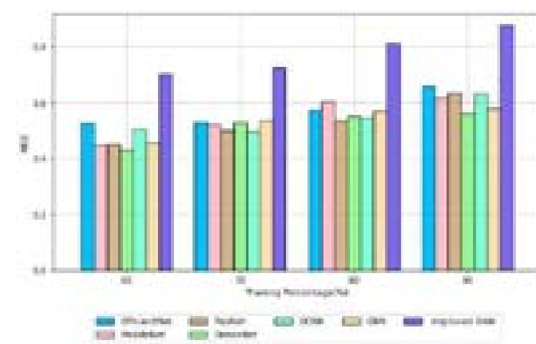


(b)

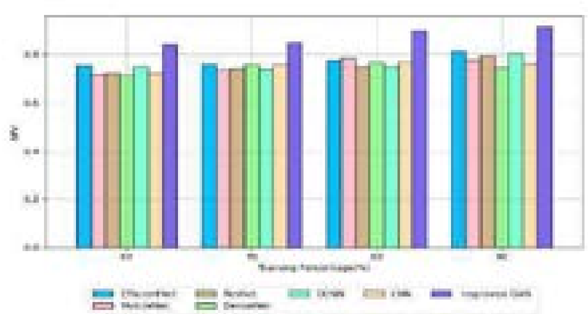
Fig.5. Assessment on Improved GAN and conventional strategies using Negative Metric



(a)



(b)



(c)

Fig.6. Assessment of Improved GAN and Conventional Strategies using Other Metric

F. Statistical Analysis on Accuracy

To ensure accurate results, each method undergoes a thorough statistical evaluation, including a comprehensive examination of key statistical parameters such as "Maximum, Minimum, Mean, Standard Deviation, and Median." Table II provides a comprehensive statistical assessment comparing the Improved GAN model with Efficient Net, Mobile Net, Res Net, Dense Net, DCNN, and CNN for phishing attack detection. Notably, in terms of the maximum statistical metric, the Improved GAN scheme demonstrates a remarkable accuracy of 0.938, meanwhile, the Efficient Net is 0.830, Mobile Net is 0.808, Res Net is 0.818, Dense Net is 0.780, DCNN is 0.818 and CNN is 0.790, respectively. Likewise, across most statistical metrics, the Improved GAN model consistently achieved higher accuracy ratings.

TABLE II: STATISTICAL EVALUATION OF ACCURACY

Statistical Metrics	Efficient Net	Mobile Net	ResNet	Dense Net	DCNN	CNN	Improved GAN
Standard Deviation	0.027	0.034	0.034	0.026	0.027	0.025	0.034
Mean	0.786	0.773	0.765	0.759	0.774	0.768	0.890
Minimum	0.763	0.723	0.726	0.716	0.750	0.728	0.852
Median	0.777	0.781	0.759	0.771	0.763	0.777	0.885
Maximum	0.830	0.808	0.818	0.780	0.818	0.790	0.938

V. CONCLUSION

A social engineering attack, which targets users' emails to steal sensitive and confidential information, is known as a phishing attack. It can be utilized as an enormous attack's part that is launched to invade government or corporate networks. Numerous anti-phishing attacks have been developed by researchers, but are inaccurate and inefficient. Consequently, an IGAN-based phishing attack detection system was introduced in this work. Initially, for data pre-processing, min-max normalization process was utilized. Afterwards, a phishing attack detection process takes place, which makes use of IGAN to provide accurate and efficient attack detection through the proposed trimming factor-based loss function. In the future, the suggested method's next steps will require classifying into two classes in order to improve the outcomes.

REFERENCES

[1] D. Rathee, and S. Mann 2022, Detection of E-mail phishing attacks–using machine learning and deep learning. *International Journal of Computer Applications*, vol. 183, no 1, pp.1-7.

[2] E. Benavides-Astudillo, W. Fuertes, S. Sanchez-Gordon, D. Nuñez-Agurto, and G. Rodríguez-Galán 2023. A Phishing-Attack-Detection Model Using Natural Language *Processing and Deep Learning. Applied Sciences*, vol. 13, no. 9, pp.1-23.

[3] U.A. Butt, R. Amin, H. Aldabbas, S. Mohan, Alouffi, and A. Ahmadian, 2023. *Cloud-based email phishing attack using machine and deep learning algorithms. Complex & Intelligent Systems*, vol. 9, no. 3, pp.3043-3070.

[4] M.J. Nabet, and L. E. George, 2023. Phishing Attacks Detection by Using Support Vector Machine. Journal of Al-Qadisiyah for Computer Science and Mathematics, vol. 15, no. 2, pp.180.

[5] G. Mohamed, J. Visumathi, M. Mahdal, J. Anand, and M. Elangovan 2022, An effective and secure mechanism for phishing attacks using a machine learning approach. *Processes*, vol. 10, no. 7, pp.1-14.

[6] S. Hossain, D. Sarma, and R. J. Chakma 2020, Machine learning-based phishing attack detection. International Journal of Advanced Computer Science and Applications, vol. 11, no. 9.pp.378-388.

[7] Y. Wang, W. Ma, H. Xu, Y. Liu, and P. Yin, 2023. A Lightweight Multi-View Learning Approach for Phishing Attack Detection Using Transformer with Mixture of Experts. *Applied Sciences*, vol. 13, no. 13, pp.1-17.

[8] T. Choudhary, S. Mhapankar, R. Bhddha, A. Kharuk, and R. Patil, 2023. A Machine Learning Approach for Phishing Attack Detection. *Journal of Artificial Intelligence and Technology*, vol.3, no. 3, pp. 108–113.

[9] Y.A. Alsariera, A.O. Balogun, V. E. O. H. Adeyemo, Tarawneh, and H.A. Mojeed, 2022. Intelligent tree-based ensemble approaches for phishing website detection. *J. Eng. Sci. Technol*, 17, pp.563-582.

[10] M. Sánchez-Paniagua, E.F. Fernández, E. Alegre, W. Al-Nabki, and V. Gonzalez-Castro 2022, Phishing URL detection: A real-case scenario through login URLs. *IEEE Access*, vol.10, pp.42949-42960.

[11] A.T.G. Tapeh, and M. Z. Naser 2023, Artificial intelligence, machine learning, and deep learning in structural engineering: a scientometrics review of trends and best practices. Archives of Computational Methods in Engineering, vol. 30, no. 1, pp.115-159.

[12] F. Salahdine, Z. El Mrabet, and N. Kaabouch, 2021, December. Phishing Attacks Detection A Machine Learning-Based Approach. *In 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON) (pp. 0250-0255). IEEE.*

[13] A. A. Orunsolu, A. S. Sodiya, and A.T. Akinwale, A predictive model for phishing detection 2022. *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no.2, pp.232-247.

[14] M..Bhurtel, Y.R. Siwakoti, and D. B. Rawat, Phishing Attack Detection with ML-Based Siamese Empowered ORB Logo Recognition and IP Mapper. *In IEEE INFOCOM 2022-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)* (pp. 1-6). IEEE. 2022.

[15] G. Sonowal, and K.S. Kuppusamy 2020, PhiDMA–A phishing detection model with multi-filter approach. Journal of King Saud University-Computer and Information Sciences, vol. 32, no. 1, pp.99-112.

[16] N. Elaraby, S. Barakat, and A. Rezk 2022, "A conditional GAN-based approach for enhancing transfer learning performance in few-shot HCR tasks". *Scientific Reports*, vol. 12, no.1, p.16271.

[17] S. Al-Ahmadi, A. Alotaibi, and O. Alsaleh 2022, PDGAN: Phishing detection with generative adversarial networks. *IEEE Access*, vol. 10, pp.42459-42468.

[18] A. Mughaid, S. AlZu'bi, A. Hnaif, S. Taamneh, A. Alnajjar, and E.A. Elsoud 2022, An intelligent cyber security phishing detection system using deep learning techniques. *Cluster Computing*, vol. 25, no.6, pp.3819-3828.

[19] M.W. Shaukat, R. Amin, M.M.A. Muslam, A.H. Alshehri, and J. Xie, 2023. A hybrid approach for alluring ads phishing attack detection using machine learning. *Sensors*, vol. 23, no. 19, pp.1-27.

[20] S.H. Ahammad, S.D. Kale, G.D. Upadhye, S.D. Pande, E.V. Babu, A.V. Dhumane, and M.D.K.J. Bahadur2022, Phishing URL detection using machine learning methods. *Advances in Engineering Software*, vol. 173, no.103288.

[21] A. Karim, M. Shahroz, K. Mustofa, S.B. Belhaouari, and S.R.K. Joga, 2023. Phishing Detection System Through Hybrid Machine Learning Based on URL. *IEEE Access*, 11, pp.36805-36822.

[22] https://www.geeksforgeeks.org/data-normalization-in-data-mining/

[23] https://www.kaggle.com/datasets/shashwatwork/phishing-dataset-for-machine-learning?resource=download

A Comprehensive Study on Alcohol Detection and Engine Locking System using Arduino

Prabu.U
Department of CSE,
Velagapudi Ramakrishna Siddhartha Engineering College,
Vijayawada, India
uprabu28@gmail.com

Meenakshi Karre
Department of CSE,
Velagapudi Ramakrishna Siddhartha Engineering College,
Vijayawada, India
karremeenakshi06@gmail.com

Maheswari Kolli
Department of CSE,
Velagapudi Ramakrishna Siddhartha Engineering College,
Vijayawada, India
maheswarikolli67@gmail.com

Malarvizhi.N
Department of CSE,
Velagapudi Ramakrishna Siddhartha Engineering College,
Vijayawada, India
nmalarvizhi16@gmail.com

***Abstract* — Currently, there is a significant increase in incidents of drunk driving and road accidents, primarily caused by drivers who lack control over their vehicles. Approximately 31% of traffic crash fatalities in the United States are attributed to drivers who were under the influence of alcohol. The number of fatalities resulting from these avoidable collisions in 2021 amounted to 13,384 individuals. Over a span of 10 years, the average number of fatalities amounted to 10,850 individuals. Implementing an engine locking system that incorporates alcohol detection can effectively mitigate the occurrence of accidents caused by intoxicated driving. An engine locking system that incorporates alcohol detection technology can prove highly beneficial for law enforcement agencies. It offers an automated safety solution for automobiles and other vehicles. While it is not possible for a person to expedite their sobriety, there are several methods that can enhance their alertness and give the appearance of being more sober. One such solution is the implementation of alcohol detection and engine locking system. When a driver attempts to operate the vehicle while heavily intoxicated, the MQ3/BAC sensor detects the presence of alcohol. If the alcohol concentration exceeds a certain threshold, a signal is transmitted to the Arduino UNO. Based on this signal, the engine will be deactivated. This measure has the potential to decrease the occurrence of accidents caused by driving under the influence of alcohol, hence minimizing the risk of fatalities and property damage resulting from drunk driving. In this article a comprehensive study is carried out on various alcohol detection and engine locking systems using Arduino. The various components used in each of the systems are analyzed.**

Keywords* — *Arduino UNO, MQ3 sensor, Alcohol detection, Engine locking

I. INTRODUCTION

Over the years, Arduino has served as the central processing unit for numerous applications. Arduino is a programmable board that is openly accessible. It is utilised in embedded systems, home automation, and industrial automation for numerous real-time applications. It is extensively utilised in several real-world contexts, encompassing embedded systems and medical applications. Its affordability and ease of upkeep make it a popular choice among professionals. Arduino's qualities have rendered it indispensable in the realm of IoT (Internet of Things) and automation, where it excels in its ability to interface with the physical world in real-time [1].

Currently, the intake of alcohol has significantly escalated. This has a profound influence on society in numerous ways. Alcohol-related accidents lead to fatalities and property damage. The primary issue associated with alcohol drinking is the occurrence of accidents, which have a detrimental impact on individuals in close proximity. Several studies have also concluded that traffic accidents are caused by individuals who consume alcohol excessively. This significantly heightens the primary peril of fatality for the intoxicated individual, as well as impacting those in close proximity. The combination of alcohol intake and driving has a profound impact on society. Engaging in the act of drinking and driving poses a significant threat to both the individual and everyone involved [2]. This irresponsible conduct results in a significant number of accidents, injuries, and financial detriment. Furthermore, alcohol-related accidents, such as drunk driving, result in several fatalities annually and do immense harm to society.

The utilisation of Arduino in an autonomous engine locking system is a sophisticated and refined approach. Implementing an automated locking system can significantly improve security. The Arduino has the capability to monitor periods of inactivity and recognise when the user shuts down the engine. Upon the expiration of a designated duration of inactivity or when the user shuts down the car, the Arduino has the capability to autonomously transmit a signal to re-engage the engine lock. Upon successful authentication, the Arduino initiates the unlocking procedure to enable the engine to start. Implementing optional automatic locking provides an additional level of protection, however ensuring consistent performance requires careful management of power supply. Consistent testing and debugging are essential for preserving the system's dependability and efficacy in improving vehicle security.

The engine locking system, built on an Arduino-based platform, is a highly effective security feature for automobiles. The system commences by initiating user authentication, which necessitates the input of a PIN code or analogue data. The Arduino validates the authentication data and, following successful verification, triggers a relay module. The relay module is responsible for regulating a locking mechanism, typically a DC motor, that physically secures or releases the engine. Once the engine is disengaged, the Arduino activates the ignition system to initiate the vehicle's operation [3]. A discretionary automated locking function improves security by automatically re-locking the engine after a period of inactivity or when the user turns off the vehicle. Prudent handling of power supply guarantees steady operation.

The current situation clearly illustrates that intoxicated driving is the main factor behind the majority of traffic collisions. Intoxicated drivers often exhibit irregular driving behaviour on roadways as a result of their lack of stability, posing a threat to all individuals on the road, including themselves. The magnitude of reckless driving surpasses all limits. India has implemented legislation that prohibits drivers from consuming alcohol and operating a vehicle. Police officers and road safety advocates may struggle to identify intoxicated drivers due to the inherent limitation of human beings being unable to occupy multiple locations simultaneously. Despite the efforts of law enforcement personnel, every attempt to reduce drinking and driving through individual actions is unsuccessful. Therefore, it is necessary to develop a system that can accurately identify the presence of alcohol in any location and at any given moment.

The World Health Organisation (WHO) has provided statistics on traffic deaths. Based on the obtained statistics, in 2013, low-income and middle-income countries experienced higher death rates per 100,000 inhabitants (24.1% and 18.4% respectively). Additionally, a significant number of economic vehicle drivers in India acknowledged consuming alcohol while on duty. In 2013, a total of 1.25 million vehicle homicides were reported worldwide. This illustrates the prevalence of drunk driving among drivers, particularly those operating large trucks and commercial vehicles.

Later in 2023 WHO provided key facts such as Road traffic crashes claim the lives of almost 1.19 million individuals annually. Motor vehicle accidents are the primary cause of mortality among individuals aged 5-29 years. Ninety-two percent of global road fatalities are concentrated in low- and middle-income nations, despite these countries accounting for only 60% of the world's automobiles. Vulnerable road users, such as walkers, cyclists, and motorcyclists, account for almost 50% of all road traffic fatalities. Most countries incur a loss of 3% of their gross domestic product due to road traffic crashes.

Alcohol impairs the driver's cognition, vigilance, attention, judgement, and reaction, all of which are crucial for their driving abilities. Recent findings indicate that even minimal alcohol use has a substantial impact on driving abilities, specifically impairing crucial skills such as visual perception, braking proficiency, and focus. The car's engine locking mechanism prohibits intoxicated drivers from starting it.

This article presents a comprehensive investigation of different alcohol detection and engine locking systems implemented with Arduino. Section 1 provides introduction, Section 2 provides an overview of various alcohol detection and engine locking systems, Section 3 provides summary on block diagrams of various alcohol detection and engine locking systems, Section 4 provides a comprehensive analysis of boards, sensors, modules, and displays utilized in various alcohol detection and engine locking systems, and Section 5 concludes the article.

II. An Overview of Various Alcohol Detection and Engine Locking Systems

In this section, an overview of various alcohol detection and engine locking systems is discussed.

Bhuta *et al.* [4] proposed an alcohol detection and engine control system. The objective of this system is to enhance the safety of vehicle driving. The implementation is carried by utilising Arduino. In a real-time environment, a system is developed to determine the condition of the driver. The proposal involves using an alcohol detector connected to an Arduino. When the alcohol level exceeds a certain limit, the vehicle's ignition system will be deactivated, and the GPS module will record the vehicle's current location. Additionally, the GSM module will autonomously transmit a danger message to the police or designated family members.

Pranjali *et al.* [5] proposed an alcohol detection system. The objective of system is to enhance the safety of human driving and mitigates accidents. The system is created by combining an alcohol sensor with an Arduino board. The Arduino processor, namely the ATmega328, possesses a greater capacity to execute a wider range of functions compared to traditional microcontrollers. The alcohol sensor utilised in the system is the MQ3 sensor, which is designed to detect the alcohol concentration present in human breath. Due to its precise sensitivity range of around 2 metres, the sensor is compatible with any type of vehicle and may be conveniently concealed from potential suspects.

Gbenga *et al.* [6] proposed an automatic car engine locking system. A prototype of alcohol detection and engine locking system is created by utilising an Arduino UNO connected to an alcohol sensor, an LCD screen, and a DC motor to showcase the idea. The system employs a MQ-3 alcohol sensor to constantly monitor the blood alcohol concentration (BAC) in order to detect the presence of alcohol in a driver's exhaled breath. The device is capable of continuously monitoring the driver's breath alcohol level by positioning the sensor on the steering wheel. If the sensors detect alcohol in the driver's breath, the ignition will not start. If the driver becomes intoxicated while operating the vehicle, the sensor will continue to detect the presence of alcohol in their breath. As a result, the engine will be automatically disabled, preventing any further acceleration of the automobile. This allows the driver to safely park the vehicle on the side of the road.

Shinde *et al.* [7] proposed an engine locking system to address the issue of fatalities and property damage caused by intoxicated driving. The solution utilises an Arduino UNO3 microcontroller connected to an alcohol sensor that detects the presence of alcohol by analysing the breath of a person operating the vehicle. The vehicle's engine is deactivated, and the emergency siren is activated immediately upon detecting alcohol, therefore reducing the likelihood of any potential accidents. Consequently, the prevention of casualties and damage to assets is ensured.

Banerjee *et al.* [8] proposed an automatic engine locking system. The device receives input in the form of breath that contains alcohol. The controller remains in a state of anticipation for the output signal from the alcohol sensor. A stimulating procedure is initiated by utilising a direct current (DC) motor with the assistance of a freewheeling diode. The entire process is closely monitored by an intelligent Atmega 328 microcontroller. Although the traditional approaches have implemented efficient setup requirements, it may be more advantageous to incorporate the whole state-of-the-art design into the system. The majority of conventional systems tend to rely heavily on the operator and are susceptible to failure due to issues such as battery life, power consumption, and external disruptions that cannot be avoided. Drunk driving is a significant cause of accidents in nearly all countries worldwide. If the system detects many instances of alcohol exceeding the predetermined threshold, it will immobilise the engine and activate the buzzer simultaneously, so preventing potential accidents.

Swetha *et al.* [9] proposed an engine locking system. The objective of this system is to detect alcohol using the alcohol sensor MQ3, which is installed in the car. In the event that alcohol is consumed, the sensor will send a signal to the microcontroller, which will then transmit a specific signal to the GSM. The GSM system transmits a message to a specific pre-registered mobile device indicating that the driver is intoxicated. This is accompanied by an audible alert and a red LED indicator, after which the engine will be immobilised.

Shukla *et al.* [10] proposed an automatic engine locking system. It showcases the development and execution of a system that detects alcohol levels in cars and automatically locks the engine. The system uses an Ultrasonic Sensor and an Arduino UNO as the Master Control Unit (MCU). The system consistently monitors the alcohol concentration level in the alcohol detection sensor and subsequently deactivate the vehicle's engine if the alcohol concentration exceeds the predetermined threshold level. The model will transmit the vehicle's location information via SIM900A. The idea offers a highly effective approach for managing and preventing accidents caused by intoxicated driving.

Anthony *et al.* [11] proposed an alcohol detection system. As the number of vehicles on roadways is rapidly increasing, the likelihood of road accidents is also sharply growing. Drunk driving is well recognised as a significant contributor to road accidents worldwide. The primary objective is to create a system capable of detecting the quantity of alcohol drunk by the driver of a vehicle. The proposed system seeks to deter users from operating a vehicle while under the influence of alcohol, with the goal of diminishing the frequency of accidents caused by drunk driving. The proposed system is constructed utilising an Arduino UNO microcontroller and a MQ-3 alcohol detection sensor as its primary components. As a precautionary measure, when the alcohol level exceeds the acceptable threshold, the ignition system of the car (DC Motor) will be deactivated, and the relevant authority will be notified using the GSM module.

Shreyas *et al.* [12] proposed an alcohol detection system. It outlines the design and execution of an Alcohol Detection system with motor locking for motorcycles and cars. The Arduino UNO is chosen as the Master Control Unit (MCU) for this purpose. The device will continuously display the alcohol concentration detected by the alcohol detection sensor and subsequently disable the vehicle's engine if the alcohol concentration exceeds the threshold level. The task provides a sustainable solution to regulate alcohol-induced damage.

Singh *et al.* [13] proposed an alcohol detection and engine locking system. This endeavours to tackle the problem of fatalities and property destruction caused by intoxicated driving. The objective of this system is to showcase the efforts in enhancing the safety of human driving and mitigating the occurrence of accidents. In this system, an Arduino Nano microcontroller equipped with an alcohol sensor is employed to identify alcohol by analysing the breath of the vehicle operator. Upon the detection of alcohol, the vehicle's engine is promptly turned off and a warning siren is activated, thereby minimising the chances of any potential calamities. Consequently, there is no loss of human lives or damage to property.

Kumarasamy *et al.* [14] proposed an alcohol detection and engine locking system. This system utilises the Alcohol Sensor and Arduino-UNO as the controlling unit to create an Alcohol Detection system with Engine Locking for automobiles. An alcohol detection sensor will be employed by the system to continuously monitor the level of alcohol concentration. If the concentration surpasses the predetermined threshold, the engine of the vehicle will be deactivated. This variant utilises the SIM800A and NEO6M GPS modules to transmit the vehicle's position to the registered mobile number via an SMS.

Sandeep *et al.* [15] proposed an alcohol detection and engine locking system as there is a requirement for an alcohol detection system that can operate without limitations in terms of location and time. It showcases the development and execution of a system that detects alcohol levels in cars and automatically locks the engine. The system utilises an Ultrasonic Sensor and Arduino UNO. The technology will consistently monitor the alcohol concentration level in the alcohol detection sensor and subsequently deactivate the vehicle's engine if the concentration exceeds the predetermined threshold level.

Vaishali *et al.* [16] proposed an alcohol detection and engine locking system. The components used in the system include an Arduino UNO microcontroller, a MQ3 sensor for alcohol detection, a resistor for current control, a DC motor, a battery, a transistor, an LED, and a buzzer for output detection. A mini breadboard is also utilized for making the necessary connections and jumper wires to interconnect all the components. If the driver is intoxicated, the MQ3 sensor will identify this and activate the LED, causing it to emit light, while simultaneously triggering the buzzer to produce a sound.

Garg *et al.* [17] proposed an alcohol detection system. The system involves the integration of alcohol concentration detection sensors with a Microcontroller board such as Arduino. The MQ3 module is utilised for the detection of alcohol particles, exhibiting a reasonable sensitivity range of around two metres.

III. A Summary on Block Diagrams of Various Alcohol Detection and Engine Locking Systems

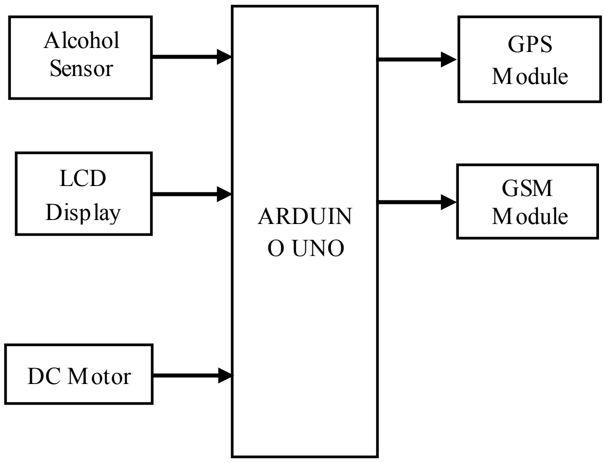
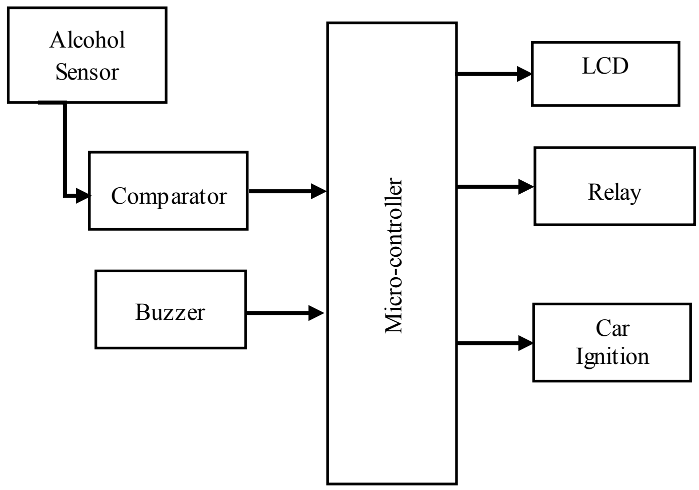
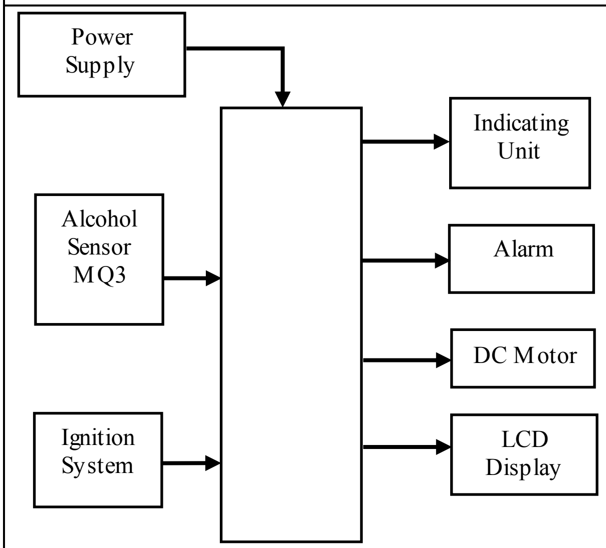
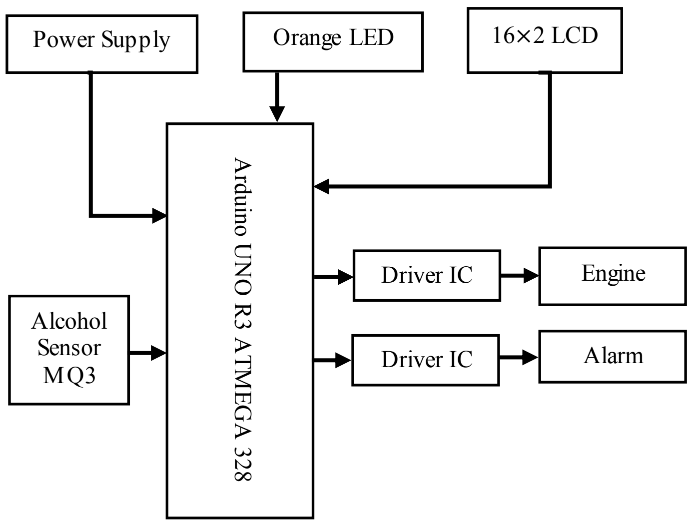
In this section, a summary on block diagrams of various alcohol detection and engine locking systems is studied.

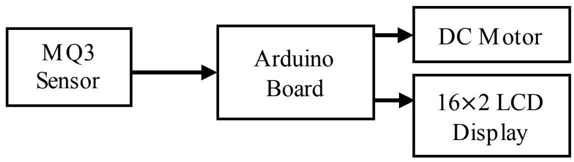
Block diagrams provide a clear and concise overview of a system's architecture, helping to convey information in a more accessible and understandable manner. The summary on various block diagrams facilitates the comparison of different configurations or versions of Arduino UNO setups, enabling a clearer understanding of how changes impact the overall system.

To have an in-depth understanding on the state-of-the-art alcohol detection and engine locking systems, the block diagrams from the different proposed works are tabulated in Table 1. Each of these block diagrams provide a diverse functionality based on the scenarios and case studies considered.

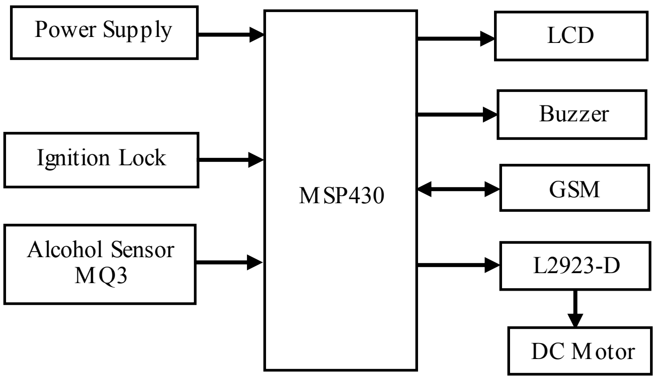
Thus, a comprehensive interpretation of different block diagrams of alcohol detection and engine locking system provides a clear understanding about the usage of various boards, sensors, and other components utilized for making the system effective.

TABLE I. BLOCK DIAGRAMS OF VARIOUS ALCOHOL DETECTION AND ENGINE LOCKING SYSTEM

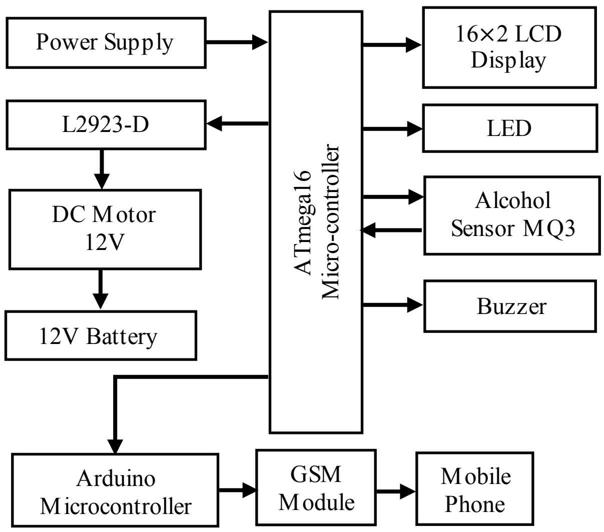
 [4]	 [5]
 [6]	 [7]


MQ3
Sensor
Arduino
Board
DC Motor
16×2 LCD
Display

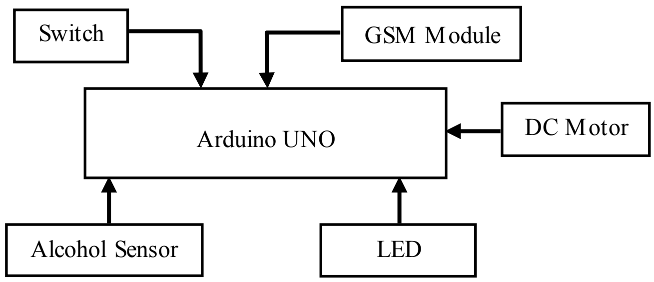
[8]


Power Supply
Ignition Lock
Alcohol Sensor
MQ3
MSP430
LCD
Buzzer
GSM
L2923-D
DC Motor

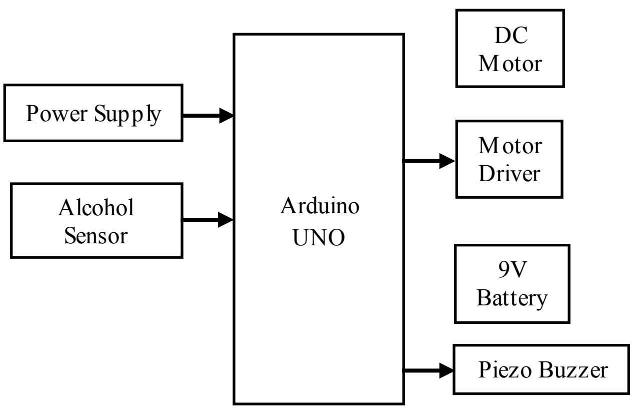
[9]


Power Supply
L2923-D
DC Motor
12V
12V Battery
ATmega16
Micro-controller
16×2 LCD
Display
LED
Alcohol
Sensor MQ3
Buzzer
Arduino
Microcontroller
GSM
Module
Mobile
Phone

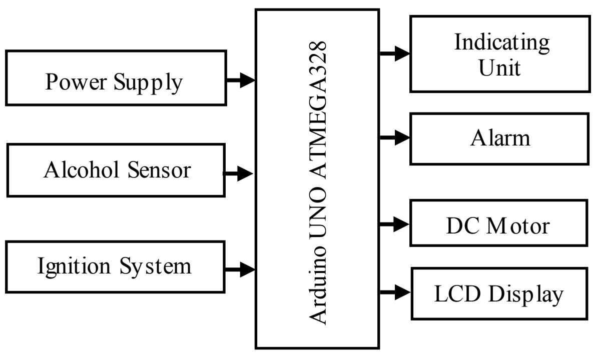
[10]


Switch
GSM Module
Arduino UNO
DC Motor
Alcohol Sensor
LED

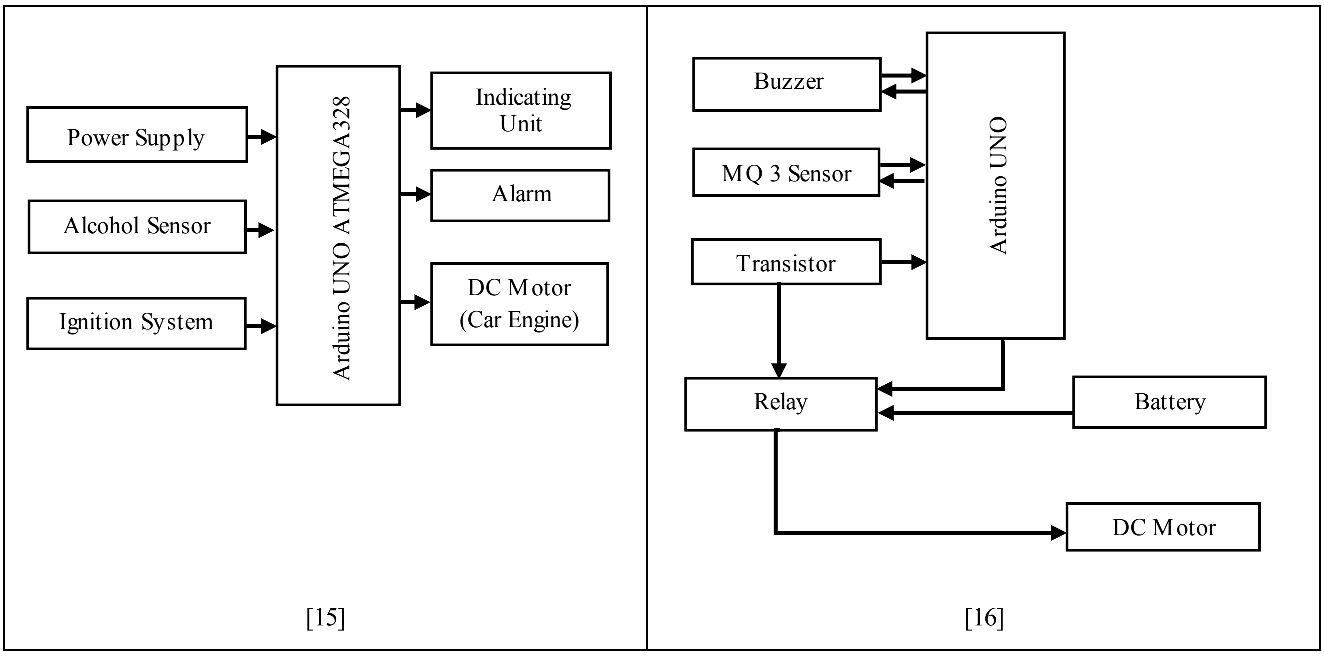
[11]


Power Supply
Alcohol
Sensor
Arduino
UNO
DC
Motor
Motor
Driver
9V
Battery
Piezo Buzzer

[12]


Power Supply
Alcohol Sensor
Ignition System
Arduino UNO ATMEGA328
Indicating
Unit
Alarm
DC Motor
LCD Display

[13]



IV. A Comprehensive Analysis of Boards, Sensors, Modules and Displays Utilized in Various Alcohol Detection and Engine Locking Systems

The alcohol detection and engine locking systems are composed of different components. A brief overview of each component is provided in the forthcoming sub-sections.

A. Boards and Motors

In the state-of-the-art research various boards and motors such as Arduino UNO, L2923D Motor Driver Integrated Circuit (IC), DC Motor, LM324 Single Supply Quad Operational Amplifiers, Mini Pump Motor, and Comparator.

1) Arduino UNO: The Arduino Uno is a widely-used microcontroller board belonging to the Arduino family of open-source electronics. It is extensively utilised by enthusiasts, learners, and experts for the purpose of prototyping and fabricating diverse electronic products. The Arduino Uno utilises the ATmega328P microcontroller, which is furnished with digital and analogue input/output pins, diverse communication ports, and other important components.

2) L2923D Motor Driver Integrated Circuit (IC): The L293D motor driver IC is a compact electronic component capable of concurrently driving two motors, commonly employed for motor control in autonomous systems. This motor driver integrated circuit allows for bidirectional control of a DC motor as well as the ability to regulate its speed. It functions as a dual H-bridge motor driver. The H-bridge is a basic circuit used to regulate a motor with a low current rating. A single H-bridge has the ability to control the movement of a DC motor in both forward and reverse directions. L293D is an integrated circuit that amplifies electric current. Additionally, it can function as a switching device.

3) DC Motor: In order to employ a DC motor with an Arduino Uno, it is necessary to build a physical connection between the motor and the board. The operation of the motor can be controlled by utilising the digital output pins of the Arduino, in conjunction with a motor driver or an H-bridge circuit as required.

4) LM324 Single Supply Quad Operational Amplifiers: The LM324 is a quadruple operational amplifier (op-amp) that is extensively utilised in electronic circuits. It is a low-cost, general-purpose operational amplifier that belongs to the LMx24 series. It is manufactured by numerous companies. The presence of the "324" in the part number indicates that the IC package comprises four distinct op-amps.

5) Mini Pump Motor: In order to employ a mini pump motor with Arduino, it is necessary to establish a physical connection between the motor and the Arduino. Then, the motor's operation can be controlled by utilising the digital output pins of the Arduino.

6) Comparator: A comparator can be employed to compare the output voltage from an alcohol sensor with a pre-established threshold. The purpose of this system is to inhibit the ignition or operation of a vehicle when the concentration of alcohol in the driver's breath beyond a specific threshold. This system is designed to enhance safety and mitigate the likelihood of events related to driving under the influence of alcohol.

B. Sensors

Sensors such as IR Sensor, MQ-3 Sensor, and BAC Sensor are used in various state-of-the art works.

1) IR Sensor: The process of incorporating an Infrared (IR) Sensor into an alcohol detection and engine locking system entails integrating the sensor to identify the existence of alcohol vapour and integrating it into a control system that regulates the engine locking mechanism according to the concentration of alcohol.

2) MQ-3 Sensor: The MQ-3 sensor is a gas sensor frequently employed for the detection of alcohol vapour, rendering it well-suited for applications such as alcohol detection systems. The process of incorporating it into an engine locking system entail establishing a connection between the sensor and a microcontroller, analysing the data obtained from the sensor, and regulating the engine immobilisation mechanism in accordance with the concentration of alcohol detected.

3) BAC Sensor: A sensor designed specifically for measuring Blood Alcohol Concentration.

C. Modules

Modules such as GSM, GPS, Relay, and Transistor are used for various purposes in the system.

1) GSM: By incorporating a GSM (Global System for Mobile Communications) module into an alcohol detection and engine locking system, it becomes possible to enable remote communication and monitoring functionalities. This functionality can be highly advantageous for transmitting notifications or warnings to specified individuals when the level of alcohol concentration above a predetermined limit or when the engine immobilisation systemis engaged.

2) GPS: Incorporating a GPS module into an alcohol detection and engine locking system offers location-based data and can improve the system's capabilities. It can be advantageous for monitoring the precise location of the vehicle, recording occurrences, and incorporating other functionalities.

3) Relay: The relay module functions as a controllable switch, operated by a low-power microcontroller, to regulate the high-power components responsible for the activation or deactivation of the vehicle's engine.

4) Transistor: A transistor is a semiconductor device with the ability to function as an amplifier, a switch, or in diverse electronic circuits. It is also used based on the requirement.

D. Displays

Displays such as LED, LCD, and Buzzer are used for indicating the status for the end user.

Table 2 shows the boards and motors used in the various alcohol detection and engine locking systems.

TABLE II. BOARDS AND MOTORS USED IN THE VARIOUS ALCOHOL DETECTION AND ENGINE LOCKING SYSTEMS

Research Work	Arduino UNO	L2923D	DC Motor	LM324	Mini Pump	Comparator
[4]	✓		✓			
[5]	✓					
[6]	✓					
[7]	✓					✓
[8]	✓		✓			
[9]		✓				
[10]	✓		✓			
[11]	✓		✓			
[12]	✓		✓			
[13]						✓
[14]	✓		✓	✓		
[15]	✓		✓			
[16]	✓				✓	
[17]	✓					

TABLE III. SENSORS USED IN THE VARIOUS ALCOHOL DETECTION AND ENGINE LOCKING SYSTEMS

Research Work	IR Sensor	MQ-3 Sensor	BAC
[4]		✓	
[5]		✓	
[6]		✓	✓
[7]		✓	
[8]		✓	
[9]		✓	
[10]		✓	
[11]		✓	
[12]		✓	
[13]		✓	
[14]	✓		
[15]		✓	
[16]		✓	
[17]		✓	

TABLE IV. MODULES USED IN THE VARIOUS ALCOHOL DETECTION AND ENGINE LOCKING SYSTEMS

Research Work	SIM900A	GPS Module	GSM Module	Relay Module	Transistor
[4]		✓	✓		
[5]				✓	
[6]					
[7]				✓	
[8]				✓	✓
[9]			✓		
[10]	✓				
[11]			✓		
[12]	✓				
[13]			✓	✓	
[14]			✓		
[15]			✓		
[16]			✓	✓	
[17]			✓		

TABLE V. DISPLAYS USED IN THE VARIOUS ALCOHOL DETECTION AND ENGINE LOCKING SYSTEMS

Research Work	LED	LCD	Buzzer
[4]		✓	
[5]		✓	✓
[6]		✓	
[7]		✓	✓
[8]		✓	✓
[9]		✓	✓
[10]	✓		✓
[11]	✓		
[12]			✓
[13]		✓	✓
[14]		✓	✓
[15]		✓	✓
[16]		✓	✓
[17]		✓	✓

Table 3 shows the sensors used in various alcohol detection and engine locking systems. Table 4 shows the modules used in various alcohol detection and engine locking systems. Table 5 shows the displays used in various alcohol detection and engine locking systems. From all those tables it can be observed that different types of components are used in varied alcohol detection and engine locking systems.

V. CONCLUSION

Drunken driving significantly aids in accidents; hence it is imperative that every vehicle be equipped with an automatic engine locking system to prevent intoxicated drivers from operating the vehicle. To have a significant impact on the current situation, it is imperative that the implementation of this technology is prioritized in both traditional internal combustion engine vehicles and electric vehicles. There have been several research carried out in this area for over a decade. Many prototypes are developed alcohol detection and engine locking systems. In this article a comprehensive study is carried out on those systems. The block diagrams, components used, and other functionalities are analyzed and summarized. Through study the insights of alcohol detection and engine locking systems are gained.

REFERENCES

[1] Francis, J., Joseph, J., Joseph, J., Reji, J., & Divya, R. S. (2023). Intelligent Accident Prevention System using IoT. AIJR Proceedings, 246-255.

Negash, B. A., & Che, Y. (2023). Design and Simulation of Vehicle Engine Lock System by Alcohol Detection Sensor for Ethiopia Driver. Journal of Control, Measurement & Instrumentation, 1(1), 1-6.

[2] Sireesha, R., Venkataiah, C., Raju, K., Rao, Y. M., & Jayamma, M. (2023). Internet of Things (IoT) based Engine Locking System. Research and Applications: Emerging Technologies, 5(3), 14-19.

[3] Bhuta, P., Desai, K., & Keni, A. (2015). Alcohol detection and vehicle controlling. International Journal of Engineering Trends and Applications (IJETA), 2(2), 92-97.

[4] Pranjali, I., Priyanka, B., Bhagyashri, N., & Chaudhari, V. D. (2017). Alcohol Detection System in Vehicle Using Arduino. International Research Journal of Engineering and Technology (IRJET), 4(6), 287-291.

[5] Gbenga, D. E., Hamed, H. I., Lateef, A. A., & Opeyemi, A. E. (2017). Alcohol detection of drunk drivers with automatic car engine locking system. Nova Journal of Engineering and Applied Sciences, 6(1):1-15.

[6] Shinde, N., Giri, A., Rima, S., & Singh, P. (2018). Alcohol detection and vehicle engine locking system. International Journal of Industrial Electronics and Electrical Engineering, 6(3):1-4

[7] Banerjee, A., Basu, U., Paul, S., Biswas, S., & Kumari, R. (2018). Alcohol Sensing and Automatic Engine Locking System, 6(2):464-469.

[8] Swetha, A., Pallavi, B., Pravallika, B., Madhuri, C., Jayasheela, T., & Venkatesh, B. (2020). GPS tracker and alcohol detector with engine locking system using GSM. International Journal of Engineering Applied Sciences and Technology, 4(1), 615-618.

[9] Shukla, D. P., Srivastava, U., Singh, S., Tripathi, R., & Sharma, R. R. (2020). Automatic Engine Locking System Through Alcohol Detection. International Journal of Engineering Research & Technology (IJERT), 9(5), 634-637.

[10] Anthony, M., Varia, R., Kapadia, A., & Mukherjee, M. (2021). Alcohol detection system to reduce drunk driving. International Journal of Engineering Research & Technology, 9(3), 360-365.

[11] Shreyas C, Vakkalagadda Anil, Mrs Aruna Rao B. (2021). Alcohol Detection using Arduino with Motor Locking. International Journal of Research Publication and Reviews. 2(7), 1521-1525.

[12] Singh, A., Rathor, A., Nagar, A., Shinde, A. K., Barfa, L., & Jadaun, A. P. S. (2022). "Alcohol detection alert and vehicle engine blocking system." Journal of Emerging Technologies and Innovative Research (JETIR), 9(6),49-62.

[13] Ramaswamy, V., Srikanth, M., Veerabhadra, Y., & Navaneeth, S. (2022). Alcohol detection and engine locking system using Arduino Uno. International Research Journal of Engineering and Technology (IRJET), 9(05), 1426-1430.

[14] Sandeep G, Tejaswini K, Shweta H K, Nagaraj, Yathish Babu AM (2022). Alcohol detector with engine locking system. International Journal of creative research Thoughts (IJCRT),10(07),2320-2882.

[15] Vaishali S. Shrinivas G. Jadhav, Vedanti M. Jadhav, Shivam S. Jadhav, Yogesh K. Jadhav, Deepak N. Jadhavar (2022). Alcohol Detection and Engine Locking System. YMER, 21(11),0044-0477.

[16] Deepak Garg, Ayush Kumar Srivastava, Deep Paliwal, Shashank Shekhar, Ashutosh Singh Chauhan (2023). Alcohol Detection System in Vehicle Using Arduino. Journal of Xi' an Shiyou University, Natural Science Edition.16(9);105-108.

Implementing an Interactive Microcontroller Toy for Children with Autism

Bala Subramanian C
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
c.balasubramanian@klu.ac.in

Nagaraj P
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
nagaraj.p@klu.ac.in

Stephi Jacob
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
9920004130@klu.ac.in

Pommayan Thileja
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
9920004107@gmail.com

Nithiya Lakshmi A
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
9920004094@klu.ac.in

Rathieeswari. R
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
9920004198@klu.ac.in

***Abstract*—**

Autism Spectrum Disorder (ASD) is a complex condition characterized by a wide range of symptoms and challenges. Children with autism often experience difficulties in social interaction, and communication, and may exhibit repetitive behaviors. Many also struggle with understanding emotions and recognizing shapes, which can impact their daily lives and learning experiences. To address these challenges and enhance their skills, a novel toy has been developed based on a microcontroller system. This innovative toy utilizes the ATmega328P microcontroller, a popular open-source platform known for its versatility and accessibility. The toy is programmed using C language within the Arduino Integrated Development Environment (IDE), providing a user-friendly and customizable platform for development. The toy offers a variety of interactive tasks and activities designed to engage and stimulate children with autism. For example, it can generate arithmetic problems for the child to solve, helping to improve their mathematical skills and cognitive abilities. Additionally, the toy includes features that allow children to identify geometric shapes, supporting their learning. By providing a structured and predictable environment, the toy helps alleviate anxiety and sensory overload often experienced by children with autism. The interactive nature of the toy encourages social interaction and communication, promoting meaningful connections with peers and caregivers.

***Keywords*— Microcontroller, Atmel Corporation Mega AVR (ATMEGA328P), Autistic, Spectrum, Arduino Integrated Development Environment (IDE), Repetitive, Geometry**

I. INTRODUCTION

Autism spectrum disorder (ASD) is a complex neurodevelopmental condition characterized by challenges in social interaction, communication, and repetitive behaviors. It manifests differently in each individual, ranging from mild to severe symptoms [1]. Typically diagnosed in early childhood, signs may include delayed speech or language skills, difficulty with social cues, repetitive movements or behaviors, and intense focus on specific interests. The diagnostic process often involves a multidisciplinary approach, with healthcare professionals such as pediatricians, child psychologists, and developmental specialists working together [2]. Parents and caregivers play a crucial role, providing valuable insights into the child's behavior and development. While there's currently no cure for autism, various interventions and therapies aim to improve quality of life and help individuals reach their full potential [3]. Behavioral therapy, including Applied Behavior Analysis (ABA), helps individuals learn and apply positive behaviors and social skills. Speech therapy assists in developing communication skills, while occupational therapy addresses sensory sensitivities and enhances daily living skills [4]. Education and support services are also vital components of autism management. Specialized educational programs tailored to individual needs provide structure and support, fostering academic and social development [5]. Additionally, support groups offer guidance and connections for families navigating the challenges of autism. Innovative approaches such as sensory integration therapy and the use of sensory toys and tools can further enhance individuals' sensory experiences, promoting relaxation, focus, and self-regulation [6]. These tools, ranging from textured objects to weighted blankets, provide comfort and support in managing sensory sensitivities. Overall, the goal of interventions and support strategies is to empower individuals with autism to thrive and lead fulfilling lives [7]. Understanding child development is crucial for effectively monitoring progress and identifying potential delays or abnormalities, especially in the case of autism, a leading cause of disability among mental disorders in children. Autism Spectrum Disorders (ASD) encompass a wide range of symptoms and severity levels, characterized by qualitative impairments in social functioning and communication, often accompanied by repetitive behavioral patterns [8]. Social interaction difficulties are a hallmark of autism, the affected children struggling to express emotions verbally and navigate social situations. Consequently, their communication skills may be significantly impaired. It is imperative to provide a support to autistic children so that the children can understand socially acceptable behaviors across various contexts [9]. To address the unique needs of children with autism, specially designed toys play a vital role. These toys are crafted to facilitate sensory stimulation, promote social interaction, and encourage communication development. Sensory toys, for instance, provide tactile, auditory, or visual feedback, aiding in sensory integration and regulation. Interactive toys encourage engagement and communication through structured play activities, fostering social skills and emotional expression [10]. Moreover, these toys can be used in therapeutic settings, complementing interventions such as behavioral therapy and speech therapy. By incorporating play-based learning into treatment plans, children with autism can enhance their skills in a supportive and enjoyable environment.

II. LITERATURE REVIEW

1.	The study investigates children's perceptions of manipulatives as toys or math tools and their influence on learning. A mixed-methods approach involving interviews and behavioral tasks was employed.	Children's perceptions of manipulatives impact their learning outcomes; when seen as math tools, manipulatives enhance learning. Findings highlight the importance of considering children's perspectives in designing effective educational interventions.	While the study provides insights into children's perceptions of manipulatives, a research gap might involve exploring the implications of such perceptions across diverse age groups, educational settings, and subjects beyond mathematics, as well as potential strategies to enhance children's positive attitudes towards manipulatives as learning tools.
2.	The study explores make-believe play with the Internet of Toys through a qualitative case study approach. Multimodal playscapes are investigated to understand their role in enhancing children's play experiences with IoT toys.	The research emphasizes the potential of multimodal playscapes in fostering imaginative play using Internet of Things (IoT) toys. Findings highlight the importance of providing diverse and engaging environments for children's play activities with IoT-enhanced toys.	While the study contributes insights into the role of multimodal playscapes and IoT toys in children's imaginative play, a potential research gap could involve further investigation into the long-term effects on children's creativity, cognitive development, and the implications for designing interactive and educational play environments.
3.	The study utilized a design-based research method involving several iterative phases. It started with a needs analysis to determine attributes for effective concept development in children. A smart toy prototype (Maya) was designed, developed, and refined based on preschool teachers' interviews and feedback, leading to the creation of an interactive smart toy using artificial intelligence technology.	The paper introduces the development of a smart toy utilizing artificial intelligence technology, catering to preschool children's interaction and engagement needs.	While the details are limited, a potential research gap might involve further exploration of the toy's effectiveness, usability, and potential impacts on early childhood education and development, as well as considerations of privacy and ethical concerns associated with AI-based toys for young children.
4.	The study employs a mixed-methods approach, combining qualitative interviews with stakeholders	The findings suggest significant concerns about the implications of emotional AI in children's	While the study effectively analyzes stakeholder and parental perspectives on emotional

	and a demographically representative UK national survey. Interviews explore professional perspectives on emotional AI in children's toys, while the survey gathers parental views on networked toys utilizing emotional data. The research investigates the social acceptability and governance aspects of emotional AI in child-oriented devices.	toys and devices, including issues related to data privacy, manipulation, media literacy, and fairness. The study underscores the need for improved governance and regulation to address potential ethical and societal challenges arising from the use of emotional AI technology in the context of children's play.	AI in children's toys, a potential research gap could involve exploring potential technological solutions or strategies to mitigate the identified concerns, as well as investigating the long-term societal impacts and potential benefits of integrating emotional AI in child-oriented technologies.
5.	The study utilized a cross-lagged panel analysis to investigate developmental associations between joint engagement variables (higher order supported joint engagement and higher order supported joint engagement + follow-in) and expressive and receptive vocabulary in young autistic children (n = 91) with language delay, measured twice over 8 months. Structural equation models were employed to derive coefficients for cross-lagged variable pairs.	The findings indicate that early higher-order supported joint engagement and higher-order supported joint engagement + follow-in are significantly associated with later expressive and receptive vocabulary. However, early vocabulary did not significantly predict later joint engagement. This suggests that these specific forms of joint engagement could be valuable targets for early intervention strategies aimed at enhancing language development in initially language-delayed autistic children.	While the study provides insights into the associations between joint engagement and vocabulary development in autistic children, a potential research gap could involve further exploration of the underlying mechanisms that drive these associations. Additionally, future research could investigate the long-term effects of targeted interventions focused on enhancing joint engagement in children with language delays, potentially bridging the gap between research findings and effective clinical practices.
6.	The study employs a Design Science framework and a Human-Centered Design (HCD) perspective to develop and assess a collection of design tools for creating Toy User Interfaces (ToyUIs). The tools encompass various stages, from user research to prototyping, and are evaluated through qualitative assessment involving 255 creators in an educational setting over six years.	The paper presents a comprehensive approach to ToyUI design by offering interdisciplinary creators a range of tools that facilitate the integration of physical toy components with technology, such as social robots and computing devices. The emphasis on Human-Centered Design and privacy considerations underscores the intent to produce innovative and user-centric solutions.	While the paper contributes valuable design tools for ToyUI development and highlights their applicability in educational contexts, a potential research gap could lie in the exploration of broader real-world implementation and scalability of these tools, considering factors beyond educational settings, and addressing potential challenges arising from evolving technology trends and user needs.
7.	The article employs a qualitative approach, drawing on a literature review and expert analysis from the gifted education field. It identifies and categorizes diverse enrichment pedagogical strategies, such as interest-based learning,	The article underscores that leveraging enrichment pedagogy, rooted in the gifted education domain, offers a compelling strategy to motivate student learning. It highlights a range of engaging instructional	While the article celebrates the contributions of enrichment pedagogy to education, a potential research gap could lie in the need for more comprehensive studies exploring the widespread implementation and impact of these

	differentiation, and project-based learning. Practical examples are provided for implementation in classrooms, and a review of longitudinal research supports positive student outcomes.	methods applicable to diverse students, promoting deeper engagement and learning. Longitudinal research findings suggest that exposure to these methods correlates with favorable student outcomes over time.	strategies across various educational contexts, especially beyond academically talented students, to address equitable and effective application.
8.	The study utilized Vygotsky's mediation theory and a design-based research approach to investigate the use of AI-interfaced robotic toys in early childhood settings. Teachers engaged 4-5-year-old children in play experiences, analyzing data from interviews, observations, and artifact analysis to uncover how children creatively collaborated, fostering inquiry literacies such as creative, emotional, and collaborative inquiry.	The findings revealed that incorporating AI-interfaced robotic toys in early childhood education promoted inquiry literacies. Children collaboratively constructed a sustainable city for a robot and its family, demonstrating creative, emotional, and collaborative inquiry skills during play, thus substantiating the potential of AI technology to enhance children's learning experiences.	This study bridges a research gap by exploring the integration of AI toys and physical environments in early childhood education. It highlights the development of children's inquiry literacies through AI-enhanced play, emphasizing the need to train educators to leverage AI-interfaced robotic toys in fostering children's inquiry literacy development.
9.	The study utilized data from the Autism and Developmental Disabilities Monitoring (ADDM) Network, focusing on 11 U.S. communities. It analyzed 2016 data to estimate the prevalence of autism among 8-year-old children, comparing it to the previous estimate from 2014. The study also explored age-related trends in autism identification and evaluated potential factors contributing to the observed increase in prevalence.	The study's findings suggest an elevated prevalence of autism among 8-year-olds, with a rate of 1 in 54, higher than the previously estimated 1 in 59. Additionally, the observed increase in prevalence might be attributed to improved identification, diagnosis, and service provision within communities, potentially indicating enhanced awareness and accessibility of services for children with autism. The study also highlights a noteworthy reduction in racial disparities in autism identification, evidenced by similar rates among black and white 8-year-olds.	The study's potential research gap lies in the need for further investigation into the underlying factors contributing to the increase in autism prevalence, including a more comprehensive understanding of changes in diagnostic practices and healthcare access. Additionally, future research could delve into the long-term impact of early evaluation and intervention on the developmental trajectories and outcomes of children identified with autism at younger ages. Lastly, exploring variations in prevalence and identification across diverse geographic and demographic contexts could provide a more comprehensive understanding of autism's complex epidemiology.
10	The paper likely employs an interdisciplinary approach, integrating insights from design engineering,	The paper suggests that design engineering can significantly enhance the well-being of individuals	Potential limitations include challenges in universal applicability due to varying autism profiles,

	neuroscience, medicine, and psychology. It involves literature review, needs assessment, concept development, prototyping, testing, and iterative refinement to create a multifunctional smart toy for individuals with autism spectrum disorder.	with autism by developing innovative products that facilitate social interactions and support cognitive, motor, and sensory skills, addressing a gap in specialized smart products for this population.	ethical considerations in designing for vulnerable populations, and potential resource constraints in producing and implementing specialized smart toys.

III. PROPOSED WORK

The proposed toy, "Why This Face," presents an innovative approach to aiding the understanding of emotions among autistic children, integrating shape recognition and arithmetic tasks within an engaging framework. With an emphasis on visual and tactile stimuli, this toy aims to facilitate comprehension of various emotional states, such as happiness, sadness, anger, and surprise, through the exploration of corresponding facial expressions. By pressing an emoji pushbutton, children initiate the journey of deciphering the reasons behind different emotional displays [21].

Structured into two levels, the toy challenges children to recognize shapes and solve single-digit arithmetic operations, providing a multi-dimensional learning experience. The incorporation of a start button activates the toy's functionalities, initiating a dynamic interaction that captivates and educates. Featuring six buttons, the toy offers a blend of shape and emotion representations, fostering cognitive development and sensory exploration. Through an LCD, children receive guidance and feedback, enhancing their understanding of the game's objectives and commands [22]. While play is crucial for children's development, the response of autistic children varies, highlighting the importance of adaptable design and individualized engagement strategies. The "Why This Face" toy is designed to appeal to autistic children through its vibrant colors, diverse shapes, tactile push-button interface, and emotive emoji representations. However, recognizing the diverse needs and preferences within the autism spectrum, the toy's effectiveness may vary, necessitating parental or teacher support for optimal engagement [23]. The integration of educational elements within an interactive and visually stimulating toy offers a promising avenue for promoting emotional understanding and cognitive skills among autistic children. By leveraging technology and sensory-rich experiences, the "Why This Face" toy seeks to empower children with autism to navigate and comprehend the complexities of human emotions, fostering growth, and social integration.

The Emotion and Shape Recognition Toy Algorithm is designed to engage users in a series of interactive tasks aimed at enhancing their emotional understanding, shape recognition skills, and arithmetic abilities [24].

The algorithm begins by initializing lists for emotions, shapes, and arithmetic answers. These lists serve as the basis for the tasks presented to the user. Upon receiving input from the user, which triggers a start signal, the algorithm presents shapes along with corresponding emotions. The user is prompted to match the shapes with their associated emotions, fostering cognitive development and emotional recognition. After completing the emotion and shape-matching task, the algorithm asks the user to provide arithmetic input. The user is required to solve arithmetic problems, and the algorithm verifies the correctness of their answers. This task aims to improve the user's mathematical skills and problem-solving abilities. Following the arithmetic task, the algorithm prompts the user to identify shapes independently. The user's input is evaluated, and feedback is provided based on the correctness of their responses. This task reinforces shape recognition skills and encourages independent thinking. Upon completion of the game, the algorithm displays an output message congratulating the user. This message serves as positive reinforcement and indicates the successful completion of the interactive tasks [25].

Algorithm.1 Emotion and Shape Recognition Toy Algorithm

1. Initialize emotion_list = ["happy", "sad", "angry", "surprised"]
2. Initialize shape_list = ["circle", "square", "triangle"]
3. Initialize answer_list = [4, 9, 7]
4. Display "Welcome to Why This Face!"
5. Wait for start button press
6. For each shape and emotion in shape_list and emotion_list:
7. Display shape and corresponding emotion
8. Wait for button press
9. If the button press matches the shape, display "Correct"
10. Else, display "Incorrect"
11. Display "Let's solve some arithmetic!"
12. For each arithmetic operation in answer_list:
13. Display arithmetic operation
14. Wait for user input
15. If the user input matches the correct answer, display "Correct" Else, display "Incorrect"
16. Display "Now, identify the shapes!"
17. For each shape in shape_list:
18. Display shape
19. Wait for a button to press
20. If the button press matches the shape, display "Correct"
21. Else, display "Incorrect"
22. Display "Congratulations! Game over."

IV. MODEL DEVELOPMENT AND DISCUSSION

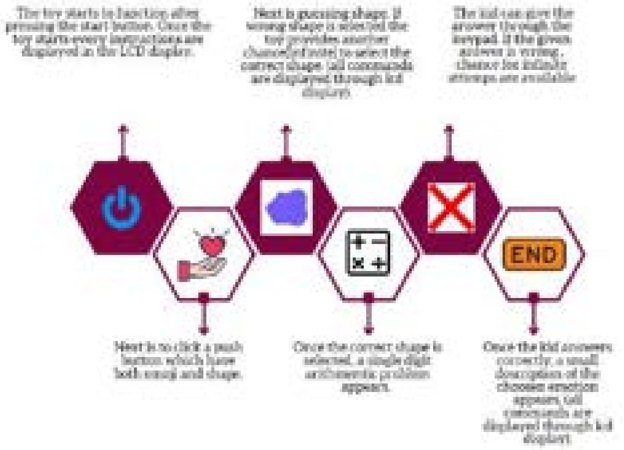


Fig 1: Process Flow chart.

The Process of playing with the micro-controller toy for an autistic kid. This flow chart helps to understand the rules of the game quickly.

In the implementation phase of the project, the focus shifted towards creating a tangible and functional prototype of the "Why This Face" toy, catering to the specific needs of autistic children. The hardware setup involved the assembly of components without the need for a constant connection to a computer, emphasizing portability and ease of use. Utilizing a battery power source, the Arduino ATmega328P microcontroller, LCD, and push buttons were integrated into a compact and durable enclosure. The Arduino code was meticulously crafted to orchestrate the interactive experience envisioned for the toy. The programming logic revolved around initiating the game upon pressing the start button, and guiding users through various tasks to enhance emotional understanding and cognitive skills. Functions were implemented to detect button presses, display messages and prompts on the LCD screen, and evaluate user responses in real time. The game flow was designed to engage users in a series of challenges, starting with the identification of shapes and corresponding emotions. Users were prompted to match shapes with their associated emotions using the push buttons, fostering cognitive development and emotional recognition. Subsequently, arithmetic operations were presented, requiring users to solve them by inputting their answers through the push buttons. Finally, users were tasked with identifying shapes independently, further reinforcing their cognitive abilities. Thorough testing and debugging were conducted to ensure the seamless operation of the toy. Each component was rigorously evaluated to verify proper functionality and responsiveness. Any issues or errors encountered during testing were promptly addressed through iterative refinement of the Arduino code and hardware configuration. Upon successful integration and finalization of the prototype, the "Why This Face" toy was ready for deployment and usage in environments catering to autistic children. Its portable and self-contained design made it easily accessible for use in homes, schools, and therapy sessions. The "Why This Face" toy represents a significant advancement in interactive toys for autistic children, offering a unique blend of educational and therapeutic benefits. One of the key advantages of the toy lies in its ability to seamlessly integrate emotional understanding with cognitive skill development. By using shapes as a medium to convey emotions, the toy engages children in a meaningful and relatable way, fostering a deeper understanding of their own and others' emotions. This innovative approach not only enhances emotional intelligence but also provides a platform for practicing social skills in a safe and controlled environment. Another aspect of the toy is its emphasis on independent learning and problem-solving. By presenting arithmetic operations and shape recognition tasks, the toy encourages children to think critically and solve problems on their own. This promotes a sense of autonomy and self-confidence, which are essential for academic and personal growth. Additionally, the toy's interactive nature keeps children engaged and motivated, creating a positive learning experience that is both enjoyable and effective. The "Why This Face" toy stands out for its versatility and adaptability. The modular design allows for easy customization, enabling caregivers and educators to tailor the toy to suit the specific needs and abilities of each child. This adaptability ensures that the toy remains engaging and challenging, providing long-term benefits for children with autism. Overall, the "Why This Face" toy represents a groundbreaking approach to interactive play for autistic children, offering a range of benefits that promote emotional, cognitive, and social development. Its innovative design and engaging gameplay make it a valuable tool for therapy, education, and recreation, setting a new standard for interactive toys in the field of autism care.

Figure 2 Implementation of toy

Fig.2 Describes the overall circuit implementation. It includes hardware components like Arduino UNO, LCD, Pushbuttons, Keypad, and connecting wires.

Figure 3 Starting command

The command "Let's start selecting an emoji" is designed to attract autistic children to play by engaging them with simple and clear instructions. The use of the word "emoji" taps into the visual appeal and familiarity of these symbols, which are often colorful and expressive. The phrase "Let's start" is encouraging and sets a positive tone, inviting the child to begin a fun activity. By keeping the command short and direct, it is easier for autistic children to understand and follow, helping them feel more confident and motivated to participate.

Figure 4 Guessing the emotion

The command "Guess the shape of the emoji" is designed to engage and encourage interaction from autistic kids. By prompting them to select the correct shape in the emoji section, the game taps into their visual and cognitive skills, providing a fun and interactive way to learn. The colorful emojis and the challenge of identifying shapes can captivate their attention, fostering engagement and learning. Additionally, the game's positive reinforcement and rewards for correct answers can boost their confidence and motivation to continue playing, creating a rewarding experience that encourages repeated engagement.

Figure 4 Implementation of arithmetic operation

After completing each arithmetic operation, the game provides a detailed result screen. This screen congratulates the player on their attempt and displays the correct answer for the operation. It also offers encouragement to keep playing and improve their skills.

V. CONCLUSION

In today's technological era, children are increasingly drawn to intelligent frameworks, signaling a growing interest in utilizing such tools. In line with this trend, a microcontroller-based toy has been developed. Studies have demonstrated the efficacy of this toy in enhancing the understanding of facial expressions, object shapes, and mathematical skills among children with varying degrees of autism spectrum disorder (ASD). Understanding facial expressions enables ASD children to perceive human emotions intuitively, contributing to improved social interactions. Moreover, focusing on mathematical tasks capitalizes on the strengths of some autistic children in this area, potentially leading to the mastery of complex problem-solving. The "Why This Face" toy is designed to cater to the specific needs of autistic children. It offers a unique blend of educational and therapeutic benefits, making it a valuable tool for therapy, education, and recreation. One of the key novelties of the toy lies in its ability to seamlessly integrate emotional understanding with cognitive skill development. By using shapes as a medium to convey emotions, the toy engages children in a meaningful and relatable way, fostering a deeper understanding of their own and others' emotions. Moreover, the toy's emphasis on independent learning and problem-solving sets it apart. By presenting arithmetic operations and shape recognition tasks, the toy encourages children to think critically and solve problems on their own. This promotes a sense of autonomy and self-confidence, which are essential for academic and personal growth.

VI. FUTURE SCOPE

To enhance user experience, two technical modifications are suggested. Firstly, increasing the number of push buttons would add diversity and excitement to the gameplay. Secondly, incorporating a variety of arithmetic questions instead of repetitive ones would prevent monotony and sustain engagement. Despite their challenges, autistic children often display remarkable talents, and the "WHY THIS FACE" toy aims to nurture these abilities for a brighter future. It caters to individuals capable of comprehending commands displayed on the LCD screen, making it accessible to a wide range of users. In today's modern era, voice assistants have seamlessly integrated into the lives of individuals across the globe. However, for autistic children who struggle with reading, conventional interactive toys may pose challenges. To address this issue, integrating voice assistant functionality into the designed toy could offer substantial benefits to a wide range of autistic children, transcending the limitations posed by traditional input methods. Presently, the toy employs a keypad as its primary input unit, limiting its interactive potential. By incorporating a microphone as an alternative input mechanism, the toy could transform into an interactive tool, facilitating improved communication skills for autistic children. The microcontroller-based toy, featuring an LCD, serves as a valuable educational tool for autistic children, facilitating learning experiences centered around emotions, shapes, and mathematical concepts. Looking ahead, the project's future endeavors involve replacing the LCD with a microphone, thereby enhancing the toy's functionality and addressing specific challenges faced by autistic individuals. Many autistic children exhibit monotone speech patterns, making it challenging for them to express emotions vocally or modulate their voices appropriately. By leveraging a microphone within the toy, post-game feedback can be delivered in various tones corresponding to the emotions chosen during gameplay. This innovative approach aims to foster greater vocal variation and communication skills among autistic children, ultimately promoting their social and emotional development.

REFERENCES

1. Gretschel, P., Munambah, N., Campodonico, K., Jacobs, M., Mabasa, N., Masinyana, A., ... & Nghulele, T. (2022). Promoting the play of children with autism spectrum disorders: Contributions of teachers and caregivers. *South African Journal of Occupational Therapy*, *52*(3), 44-51.

2. Leoste, J., Tammemäe, T., Eskla, G., San Martín López, J., Pastor, L., & Blasco, E. P. (2022). Bee-bot educational robot as a means of developing social skills among children with autism spectrum disorders. In Robotics in Education: RiE 2021 12 (pp. 14-25). Springer International Publishing

3. Canoy, A., LPTa, J. C. B., & LPTb, W. E. C. (2022). The Lived Experiences of Teachers and Parents Having Children with Autism Spectrum Disorder During The Pandemic. *The Lived Experiences of Teachers and Parents Having Children with Autism Spectrum Disorder During The Pandemic*, *115*(1), 17-17.

4. Wang, Z., Gui, Y., & Nie, W. (2022). Sensory integration training and social sports games integrated intervention for the occupational therapy of children with autism. *Occupational Therapy International*, *2022*.

5. Mehralizadeh, B., Baradaran, B., Nikkhoo, S., Soleiman, P., & Moradi, H. (2023). A Sensorized Toy Car for Autism Screening Using Multi-Modal Features. Sustainability, 15(10), 7790

6. Sundas, A., Badotra, S., Rani, S., & Gyaang, R. (2023). Evaluation of autism spectrum disorder based on the healthcare by using artificial intelligence strategies. Journal of Sensors, 2023, 1-12.

7. Hosseinioun, A., & Dolah, S. M. (2022). Designing Factors to Improve Verbal Ability among Children with Autism in Subang Jaya, Malaysia. Specialusis Ugdymas, 2(43), 752-761.

8. Sulaiman, A., Rahman, H., Ali, N., Shaikh, A., Akram, M., & Lim, W. H. (2022). An augmented reality PQRST based method to improve self-learning skills for preschool autistic children. *Evolving Systems*, 1-14.

9. Cañete, R., & Peralta, E. (2022). Assistive Technology to Improve Collaboration in Children with ASD: State-of-the-Art and Future Challenges in the Smart Products Sector. Sensors, 22(21), 8321.

10. Bondioli, M., Chessa, S., Kocian, A., & Pelagatti, S. (2023). A survey on technological tools and systems for diagnosis and therapy of autism spectrum disorder. Human–Computer Interaction, 1-29

11. Junnarkar, V. S., Tong, H. J., Hanna, K. M. B., Aishworiya, R., & Duggal, M. (2022). Occupational and speech therapists' perceptions of their role in dental care for children with autism spectrum disorder: A qualitative exploration. *International Journal of Paediatric Dentistry*, *32*(6), 865-876.

12. Binns, A. V., Casenhiser, D. M., Shanker, S. G., & Cardy, J. O. (2022). Autistic preschoolers' engagement and language use in gross motor versus symbolic play settings. *Autism & Developmental Language Impairments*, *7*, 23969415221115045.

13. Karrim, S. B., Flack, P. S., Naidoo, U., Beagle, S., & Pontin, A. (2022). The experiences of speech-language therapists providing telerehabilitation services to children with autism spectrum disorder. *South African Journal of Communication Disorders*, *69*(2), 917.

14. Burghardt, G. M., & Palagi, E. (2023). The evolution and ontogeny of play: comparative perspectives. *International Journal of Play*, *12*(1), 1-3.

15. Singh, A., Raj, K., Kumar, T., Verma, S., & Roy, A. M. (2023). Deep learning-based cost-effective and responsive robot for autism treatment. *Drones*, *7*(2), 81.

16. Shahane, V., Kumavor, P., Morgan, K., Friel, K. M., & Srinivasan, S. M. (2023). Protocol: A protocol for a single-arm interventional study assessing the effects of a home-based joystick-operated ride-on-toy navigation training programme to improve affected upper extremity function and spontaneous use in children with unilateral cerebral palsy (UCP). *BMJ Open*, *13*(5).

17. Bouhali, R., Al-Tabaa, H., Abdelfattah, S., Atiyeh, M., Esmaeili, S., & Al-Tabbakh, H. (2023). Otisma: an integrated application and humanoid robot as an educational tool for children with autism. *Journal of Medical Engineering & Technology*, *47*(1), 44-53.

18. EL ARBAOUI, F. E. Z., ELHARI, K., & SAIDI, R. (2023). A systematic literature review on the internet of things applications used by people with autism spectrum disorder.

19. Khan, N., Plunk, A., Zheng, Z., Adiani, D., Staubitz, J., Weitlauf, A., & Sarkar, N. Real-Time Analysis of Precursors to Challenging Behaviors for Intervention and Detection (Rapid) in Children with Autism Spectrum Disorder. *Available at SSRN 4431502*.

20. Asahi, R., Yoshimoto, S., & Sato, H. (2023, June). Tomographic tactile sensor-based finger motion analysis system to identify number of grasping fingers for evaluating fine motor skills. In *2023 IEEE International Symposium on Medical Measurements and Applications (MeMeA)* (pp. 1-6). IEEE.

21. Perumal, B., Nagaraj, P., Rajesh, V., Reddy, K. K., Shoaib, M., & Krishna, B. V. (2023, August). Vehicle Controlling System for Traffic Law Enforcement using Internet of Things. In *2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS)* (pp. 1564-1568). IEEE.

22. Nagaraj, P., Subramanian, C. B., Reddy, B. N. K., Reddy, L. M., Reddy, G. G., & Reddy, C. Y. (2023, July). Smart Water Controller in Metro Water Supply Lines Using Ultrasonic Sensor and Flow Sensor. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)* (pp. 127-132). IEEE.

23. Muneeswaran, V., Nagaraj, P., Anuradha, L., Lekhana, V., Vandana, G., Devi, K. G., & Tejaswee, A. (2023, January). Automatic Pressure Cooker Risk Detection And Control Using Arduino. In *2023 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-5). IEEE.

24. Nagaraj, P., Muneeswaran, V., Rupesh, C., Kartheek, B., Lekhya, Y., & Swetha, K. (2023, July). A Robust Recession Detective Analysis System using IoT Smart Sensor Devices. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)* (pp. 1260-1264). IEEE.

25. Vamsi, A. M., Deepalakshmi, P., Nagaraj, P., Awasthi, A., & Raj, A. (2020). IOT based autonomous inventory management for warehouses.

In *EAI International Conference on Big Data Innovation for Sustainable Cognitive Computing: BDCC 2018* (pp. 371-376). Springer International Publishing.

IoT-based Implementation of Eye-Gaze Technology for Enhanced Computer Accessibility and Control in Differently-Abled Communities

Nagaraj P
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamil Nadu, India
nagaraj.p@klu.ac.in

Dhiliphan Rajkumar T
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamil Nadu, India
t.dhiliphan@klu.ac.in

Reshma Joselin S
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamil Nadu, India
9920004115@klu.ac.in

Ram Balaji S
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamil Nadu, India
9920004272@klu.ac.in

Suresh Kannan S
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamil Nadu, India
9920004275@klu.ac.in

Praveena R
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamil Nadu, India
9920004196@klu.ac.in

Abstract*— In the current computer era, many computer technologies have been developed such as Artificial Intelligence, the Internet of Things, Machine Learning, Data Science, and so on. Mainly IoT contributes to improving accessibility and connectedness in this system, which makes the system more adaptable and user-friendly. These technologies are also developed to be used by physically disabled people throughout the world. This system has leveraged a way for disabled people to use the computer easily, by a method called Eye-Gaze. Eye-Gaze is an alternative approach that lets the user interact with the computer with only their eye movements, who do not have access to the mouse. This system uses a webcam to improve the accuracy, mobility, and utility of the eye-tracking technology. In short, this system offers an up-front and convenient interaction mode that uses only eyeballs, which is crucial for those with physical disabilities. It is implemented to overcome the accuracy, reliability, and calibration issues that already exist in the system. The proposed system outlines the implementation of the cursor control and the system is set up in a way that perfectly mimics the everyday routine of the people.

***Keywords*— Computer-Era, Artificial Intelligence, Internet of Things, Machine Learning, Eye-gaze Cursor Control.**

I. INTRODUCTION

The importance of human-computer interactions cannot be overstated in the fast-paced world of computer technology. Human-computer interaction plays an important role in the development of the information society. Human-computer interfaces (HCI) focus on the utilization of computer technology to develop effective interfaces. Finding the right technology is essential for smooth human-computer interactions, which in turn are essential for the continuous growth of the information society. It is crucial to guarantee inclusivity in the information society, especially for those with physical impairments. It is imperative to advance alternate techniques of human-computer communication, to give them equal possibilities. Researchers' interest in human-computer interaction (HCI) has grown recently on a global level, showing the significance of coming up with novel solutions, where eyeball movements, face recognition, and hand gestures are common means of communication. [1] In this interaction, hand gesture recognition is essential for device control. Certain research aims to enhance the user experience by providing natural and immersive computing environments through the utilization of hand gestures for intuitive and effective interactions. [2] It imitates mouse control by tracking finger movements, using Mediapipe and openCV2, the goal is to improve computer interaction at a reasonable cost. [3]

The use of vision-based eye movement detection systems in Human-Computer Interfaces is an important example. Eye gazing is a crucial hands-free display control technique that is integrated into HCI. [4] It uses Python's image processing and machine learning capabilities to track the movements of the user's eyes. [5] People with disabilities can operate with support, overcoming obstacles they might encounter when using a standard computer, With the help of this eye control technology. The need for efficient human-computer interaction is growing as computer technology develops. Yet, people with disabilities still encounter difficulties when using computers. Users can collaborate by using this eye control technology, which is integrated into computers. The main objective of a human-computer interface is, to create user-friendly interfaces with computer technology. The significance of human-computer contact is imperative to support. This advances their status in the Information Society and advances the larger objective of developing a range of technological innovations that are both approachable and diverse. Recently, there has been a lot of application of vision-based eye movement detection systems for people with disabilities. This system is done by, proving that eye movements may be used to control the cursor. An integrated hands-free human-computer interaction system has taken the place of traditional input devices. Those with specific diseases or disabilities that make it difficult for them to operate a traditional computer would especially benefit from this situation. The use of a physical computer component to operate the cursor satisfies the requirements of those with physical impairments. Modern technology uses Open CV to interpret eye movements in place of conventional mouse actions like clicks and movements. Facial expressions like eye blinking, mouth opening, and head movements are used to activate mouse clicks. This model presents a webcam mouse that uses 3D model-based biased face-tracking technology. The computer technique known as facial recognition is used widely in different fields to identify faces in digital images. The suggested approach makes use of a basic face-tracking system to identify facial features. This makes it possible to analyze facial gestures without the requirement for the user to interact directly. The human face provides crucial information for applications like emotion recognition and human-computer interface. It is possible to accomplish feature extraction with the Local Binary Pattern algorithm. Most of the eye movements are classified using a support vector machine (SVM) classifier. The movements of the eyes to the left, right, and open are all captured by a web camera. SVM performs data analysis and works well for regression analysis as well as classification.

II. LITERATURE REVIEW

An EOG-based HCI system that uses real-time processing for mouse cursor control after EOG signals are acquired, amplified, and digitized. This allows users to control a computer interface with their eye movements. [6] The system combines computer vision algorithms, Raspberry Pi, and IP camera technologies to accurately detect eye movements and provide precise control over the computer cursor based on the direction of the user's eyeball movement. [7] The system uses OpenCV and a Raspberry Pi to detect pupils, which allows for accurate control over cursor movement by creating a direct link between the cursor's movement and the pupil's center. Firstly, the application program for pupil detection is executed, and then the SD card is loaded into the Raspberry Pi. [8]A specific human-computer interaction system that depends on various input devices to receive user input. The movement of the computer cursor is technically connected to the center of the pupil. An electronic device governs the cursor's movement. [9] It is a method of dividing the screen into multiple tiny rectangular regions to increase the precision and stability of eye-controlled cursor movement. Upon entering any of these places, the cursor's position is locked in the middle of the area. This method pursues to reduce the cursor's slight shaking that results from the pupil's natural swaying within a limited range. [10] The most recent developments in 3D visualization techniques and instruments for studying eye movement data in virtual spaces. Especially in light of the accessibility of non-intrusive eye-tracking devices that are often combined with head-mounted displays. [11] The usability of eye-tracking in human-computer communication, mainly for those with physical impairments, the study presents an eye control system. This device combines keyboard and mouse functionality into one whole interface that responds only to the user's eye motions. Accuracy is guaranteed by a magnification module, and the design adapts to human behavior. [12] This work overcomes restrictions on gaze detection accuracy and presents a unique gaze-based virtual keyboard with portable eye trackers. The virtual keyboard provides instant access to every letter with a single command. Moreover, it is suggested that severely disabled people can get around the Midas touch problem by using a USB mouth switch attached to a computer mouse. [13] A HCI system is introduced, which is useful for people with limited mobility. It replaces conventional mouse controls with eyeball movements. It does this without the need for additional hardware by using OpenCV and NumPy, and simple webcam technology. The main component of the system is a face detector that combines a linear classifier and HOG to provide accessibility without requiring direct human engagement. [14]

A decision tree algorithm guarantees accurate eye location detection for accurate mouse movement. The system provides a user-friendly interface that is accessible to individuals with motor impairments. They enable smooth computer interaction through natural eye and facial movements and gestures. [15] The study found that, for old people, the eye-gaze input system which uses eye-tracking technology produced faster pointing times than traditional mouse input. [16] This study aims to improve VR usability by combining bio-sensors such as EEG and eye tracking. It is discovered that hand-tracking works well for teleport triggering and eye-tracking works well for location targeting. [17] In UX research, eye tracking and mouse tracking provide useful data. Both eye trackers performed similarly in a study that used a menu navigation task. By combining these techniques, insights can be obtained at a low cost. [18] The users can integrate eye tracking with virtual reality (VR). It describes immersive vs non-immersive VR and emphasizes the advantages of IVR, particularly when combined with eye tracking. It analyses recent research and makes references for future research to enhance the use of these technologies in education. [19]

The eye-tracking interactions are assessed with a Meta Quest 2 VR headset in this study. In terms of, unmodified eye tracking performed similarly to controller input and outperformed head tracking. Eye tracking was rated positively by each user in terms of adoption and usability, representing that it may improve AR/VR interactions. [20] A method for examining mental representation patterns is eye-tracking technology, especially when it is used for tasks like mental rotation. It requires the user to mentally rotate three-dimensional objects to gauge visuospatial ability. Though visual neuroscience has traditionally used 2D images on computer screens, recent trends have responses to naturalistic stimuli. On the other hand, there is a lack of research on mental rotation tasks in immersive VR with 3D objects. [21] The research presents an eye-tracking controlled mouse cursor that allows to use of computers independently without the need for a conventional mouse, for people with physical impairments. It is ergonomic, lightweight, and works with most operating systems. It improves life quality and productivity. It is a big step toward increased independence for people with disabilities. [22]

Three experiments using mouse cursor-tracking investigated predictive sentence processing speed. Participants heard predictive or non-predictive sentences while viewing visual arrays and moved the cursor accordingly before hearing the nouns. This effect continued across different speech rates. [23] A transparent and adaptable electrostatic sensing interface for active eye tracking (AET) systems is suggested in this system. This interface improves charge storage capacity and permits real-time interpreting of eye movements for a variety of applications such as virtual reality, medical monitoring, and eye-controlled HCI, with an angular resolution of 5°. [24] People with physical impairments encounter difficulties when using computers, and it suggests using facial gesture recognition in place of human-computer interaction (HCI). This strategy could take the place of conventional HCI tools. The study describes the procedures used to carry out, the experiments that validate this idea. [25]

III. PROPOSED METHODOLOGY

The system manages contactless human-computer interfaces by employing face tracking, eye detection, and interpreting blinking in real time [26]. The conventional method of computer interaction via a mouse is replaced by leveraging human eyeball movements. There are some software and hardware used in this system. The software used is python3, and there are some libraries included in it, they are CV2, Mediapipe, and PyautoGUI. The hardware used in this system is Raspberry Pi, LCD, and webcam. Using these software and hardware a convenient system is created mainly for the physically disabled people to use [27].

A. Software

The software that is used in this system is Python3. Python is a programming language frequently employed for constructing websites and software, automating tasks, and analyzing data. Python has resources, guides, and community support, so working in Python is easier compared to other programming languages. The libraries that have been included are:

1) CV2
2) MediaPipe
3) PyAutoGUI

a. CV2

CV2 stands among the frequently employed libraries for handling image and video processing tasks. Its central module, referred to as cv2, provides developers with an accessible interface to seamlessly engage with a variety of image and video processing functionalities. Since Integration with Other Libraries is easier CV2 is used as one of the libraries.

b. MediaPipe

A flexible pipeline framework called Mediapipe was created to create customized machine-learning solutions for streaming and real-time media. It is an open-source tool that is used to build pipelines that perform computer vision inference on different types of sensory data, such as audio and video. Using Mediapipe, users can create a modular graph representation of a perception pipeline, adding or removing components as needed. Since High-Level API can be created and ready-to-use Models are present media pipe is used.

c. PyAutoGUI

PyAutoGUI is a toolkit that enables users to develop scripts capable of emulating mouse movements, object clicks, text transmission, and the activation of hotkeys. It proves useful for navigating through systems that restrict automated browser usage. Since there is no need for External Dependencies and cross-platform Compatibility is supported pyautoGUI is used.

B. Hardware

a. Webcam

A webcam is a regular video camera that can record videos and take still images when it is connected to a computer. Frequently, it is utilized for video conferences via Zoom, Google Meet, and other similar apps. In this system, the webcam is utilized to identify the eyeball so that it can make the mouse cursor move as per the movement of the eyeball. In Figure 3, a USB Webcam is given.

Figure 3. Webcam

b. Raspberry Pi

The Raspberry Pi is a compact, cost-effective, designed to connect to a computer desktop or TV using standard peripherals like a mouse and keyboard. It possesses dedicated components, including a processor, memory, and graphics driver, akin to a traditional PC. Additionally, it is equipped with its operating system, Raspberry PiOS, which is a modified version of Linux. The primary purpose of the Raspberry Pi is to serve as a low-cost, low-power device with various applications. The Raspberry Pi is to design an Operating System. In Figure 4, Raspberry Pi Model 4B is given.

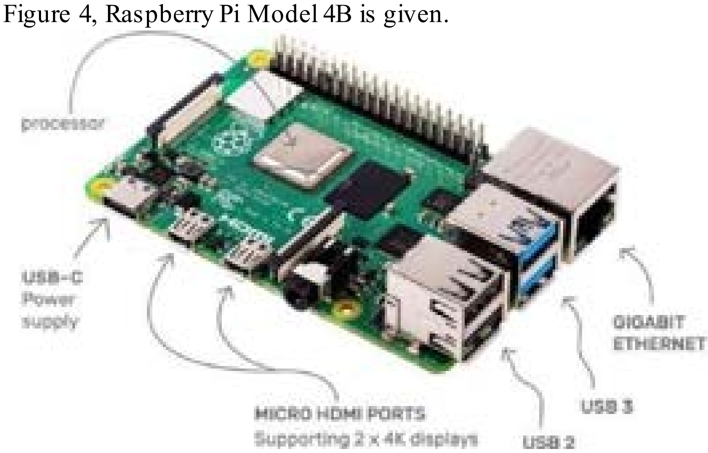


Figure 4. Raspberry Pi

c. LCD Display

Using the DSI connector, the Raspberry Pi Touch Display connects to the Raspberry Pi as an LCD screen. You can use the Touch Display and the regular HDMI display output at the same time when the panel is connected. All Raspberry Pi models are compatible with the Touch Display. Therefore, using the Raspberry Pi and the display a personal computer can be created. In Figure 5, an LCD that is compatible with Raspberry Pi is given.

9) Transform the frame to RGB format
10) Using face mesh, process facial landmarks and extract landmark_points from the result
11) If landmark_points is not null:
12) Remove landmarks so you can process them later
13) Landmark in enumerate (landmarks [474:478]) for eachid:
14) Utilizing normalized coordinates, compute (x, y)
15) On the frame, draw a circle at (x, y)
16) When id = 1:
17) Determine screen_xand screen_y using the normalized coordinates
18) Point and click on (screen_x, screen_y)
19) A list of specific landmarks, such as [landmarks [145], landmarks[159]], is set to the left
20) For every landmark on the left:

Figure 5. LCD Display

C. Pseudocode

1) First, import cv2 and then import media pipe as mp and Pyautogui
2) Select cv2.VideoCapture(0) as the camera
3) Refine_landmarks =True
4) Set face_mesh to mp.solutions.face_mesh.FaceMesh
5) pyautogui.size() to set screen_w and screen_h
6) If True:
7) Read the camera's frame
8) Turn the frame to a horizontal position
21) Utilize normalized coordinates, and compute (x,y)
22) On the frame, draw a circle at (x, y).
23) If left [0] and left [1] are separated vertically by less than 0.004:
24) Use pyautogui.click() to execute a mouse click.
25) Sleep 1 second
26) Show the frame with the landmarks drawn
27) Await the pressing of a key (cv2.waitKey(1))

IV. RESULTS AND DISCUSSION

In Figure 6, the architecture of the system is given. The computer is created by the Raspberry Pi together with the LCD. A webcam is fixed with the computer. The webcam is analyzing the eyeball movements and face recognition of the user. Therefore, the user can see the cursor movement on the computer.

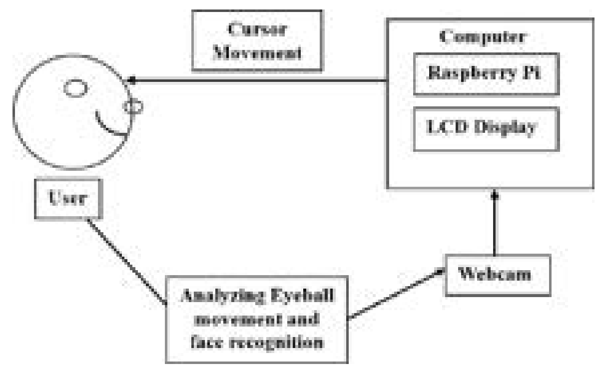


Figure 6. Schematic Diagram

In Figure 7, the primary steps of the eye-tracking project are defined as an algorithm. The system is initialized at first. It then records video input from the webcam. The facial features, especially the eyes are detected in video recordings. The system determines the direction of gaze based on the detected eye positions. Based on this data, the screen's cursor is shifted appropriately. The user sees the processed video with gaze overlays at the same time. The system keeps an eye out for user input during this process, and also for commands to end the program. The video processing is carried out by the system if no quit commands are found. Finally, the system shuts down when it detects a quit command.

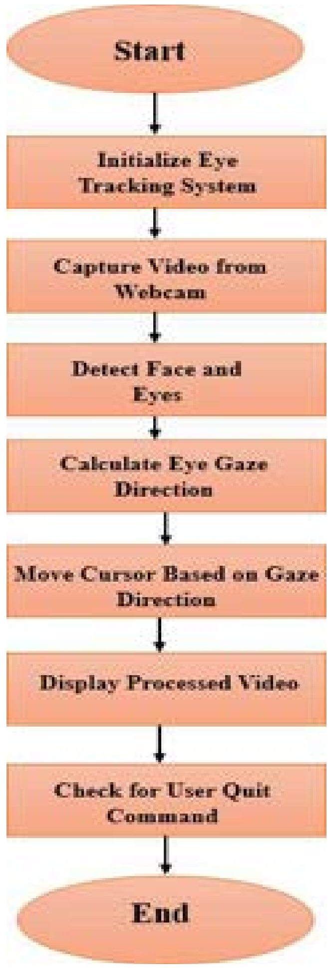


Figure 7. Flow chart

The development of an affordable and portable mouse control system using eye-tracking technology marks notable progress in the realm of human-computer interaction. The implementation of this system includes a few steps. Eye tracking is identifying where the eyes are for various purposes, such as analyzing behavior or operating a computer. Firstly, the libraries in the Python program are included they are, CV2, MediaPipe, and PyAutoGUI so that mouse movements can be emulated, build pipelines that perform computer vision inference on different types of sensory data. To identify faces in the video frame, the cv2 is used. It detects the face by comparing every content in the frame that matches the real facial traits. After the face has been detected the eye is detected similarly by the cv2. Then the processed video frame is displayed on screen with a drawn rectangle around the facial area. This enables real-time observation of the face and eye detection.

Once the face and eye-gaze are detected using face landmark detection, and eye-gaze estimation, they are statistically analyzed based on any meaningful trend or data collection. This analysis makes the cursor move according to the eyeball movement in the video frame. The analyzed data are not stored anywhere in the memory of the operating system. Only during the interpretation, the eyeball movement data is stored locally and then it is deleted quickly, and stores the next data locally for a few seconds. Lastly, testing and calibration were done. It represents vital phases in enhancing the accuracy, precision, dependability, and overall effectiveness of eye-tracking systems. The testing is done in various environmental conditions to ensure accuracy and precision. Depending on the test the calibration is done in this system for increasing the accuracy and precision. In Figure 8, a screenshot of the system when it is analyzing the eyeball and face is given. In Figure 9, the screenshot of the cursor moving by the eyeball is given. In Figure 10, the prototype is given.

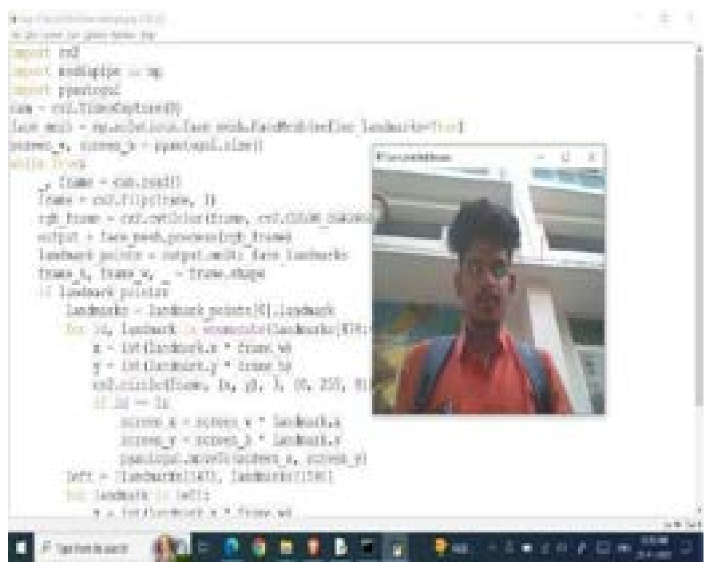


Figure 8. Analyzing Face and Eyeball

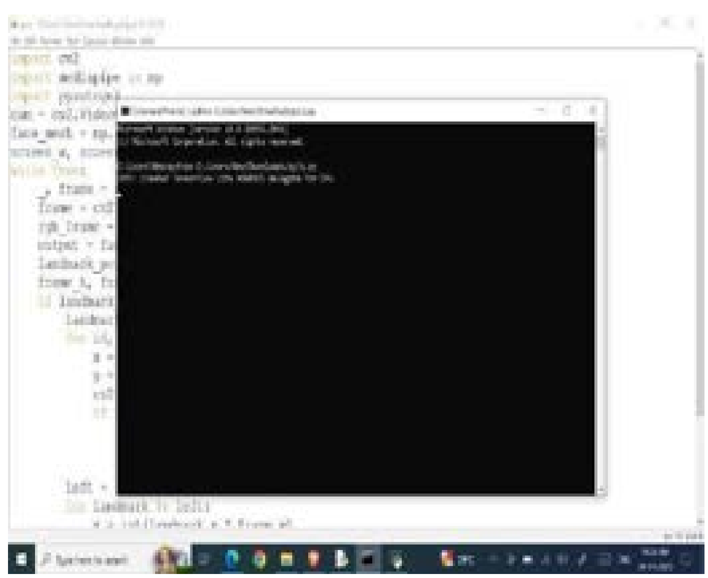


Figure 9. Cursor movement

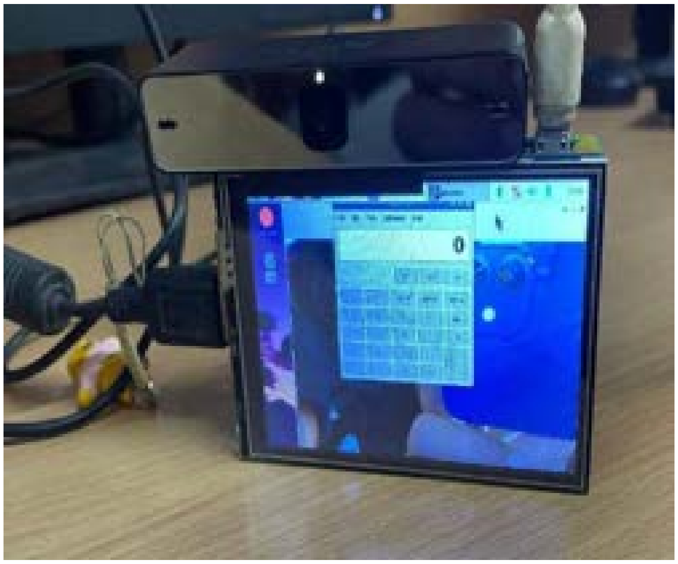


Figure 10. Prototype

Table 1. Accuracy Table

User Position	**Light Condition**	**Accuracy (%)**
Front Facing	Normal Lighting	95%
Front Facing	Low Lighting	85%
Side Facing	Normal Lighting	65%
Side facing	Low lighting	50%

In Table 1. An accuracy table is given. The table consists of the survey taken from 20 different people to check the accuracy of the system in 2 different light conditions either during normal lighting or in low lighting and 2 different user positions either front facing or side facing. The accuracy percentage is higher when the user is front-facing, under normal lighting, compared to other scenarios. The accuracy percentage is less when the user is side-facing under lowlighting compared to other scenarios. The accuracy percentage is found by calculating the number of correct matches divided by the totalnumber of test cases and then multiplying it by 100.

In the first scenario, when the users are facing the front under normal lighting 19 out of 20 people experienced the system with good accuracy, except 1 person, because of the crowded environment. In the second scenario, when the users are facing the front under low lighting 17 out of 20 people experienced good accuracy, except 3 people because the lighting was not enough to detect the eyeball. In the third scenario, when the users are side-facing under normal lighting 13 out of 20 people only experienced good accuracy, except 7 people because, the eyeballs are detected only up to a certain angle of turn and also facial feature varies for everyone, that is for some people both eyes are visible upto 45 degrees and for others only one eye was visible. In the last scenario, when users are side-facing under low lighting only 10 out of 20 people experienced good accuracy, except 10 people because the light was not enough for some people and the angle of turn was not perfect for the eyes to be detected. Therefore, these drawbacks can be overcome by using different tools, like using an IR webcam for low-lighting places, and using multiple cameras for detecting eyeballs from many angles, help to achieve higher accuracy.

V. CONCLUSION

A significant advancement in the realm of human-computer interaction involves the establishment of a cost-effective, portable system for controlling a mouse through eye-tracking. To sum up, the goal is to create an eye-gaze tracking system using computer vision and Python. Using the libraries the face and eyeball movement are detected and it is analyzed and interpreted as the mouse cursor movement. Then testing is done to ensure accuracy and precision. This system presents promising avenues for applications in ATMs, Driving Assistance health monitoring, gaming, education, and telemedicine. The system's effectiveness and influence across many fields will increase in the future as its accuracy, precision, and usability are further improved. This breakthrough signifies a revolutionary step towards improving the overall quality of life for individuals with disabilities and advancing the field of assistive technology.

REFERENCS

1. Ranjith, G. C., and Saritha Shetty. "Implementing A Real Time Virtual Mouse System Using Computer Vision." *International Research Journal of Modernization in Engineering Technology and Science* (2023).
2. Annachhatre, R., Tamakuwala, M., Shinde, P., Jain, A., & Kulkarni, P. V. (2022, September). Virtual Mouse Using Hand Gesture Recognition-A Systematic Literature Review. In *2022 IEEE International Conference on Blockchain and Distributed Systems Security (ICBDS)* (pp. 1-6). IEEE.
3. Sonawane, T., Waghmare, S., Dongare, A., Joshi, A., & Birajdar, A. (2023, April). Computer Vision-Based Virtual Mouse Cursor Using Hand Gesture. In *International Conference on Paradigms of Communication, Computing and Data Analytics* (pp. 713-722). Singapore: Springer Nature Singapore.
4. Miah, P., Gulshan, M. R., & Jahan, N. (2022, December). Mouse Cursor Movement and Control using Eye Gaze-A Human Computer Interaction. In *2022 International Conference on Artificial Intelligence of Things (ICAIoT)* (pp. 1-6). IEEE.
5. Dhanasekar, J., KB, G. A., Kiren, A. S., & Ahamath, F. (2023, May). System Cursor Control Using Human Eyeball Movement. In *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 2495-2500). IEEE.
6. Hossain, Z., Shuvo, M. M. H., & Sarker, P. (2017, September). Hardware and software implementation of real time electrooculogram (EOG) acquisition system to control computer cursor with eyeball movement. In *2017 4th international conference on advances in electrical engineering (ICAEE)* (pp. 132-137). IEEE.
7. Khare, V., Krishna, S. G., & Sanisetty, S. K. (2019, March). Cursor control using eye ball movement. In *2019 Fifth International Conference on Science Technology Engineering and Mathematics (ICONSTEM)* (Vol. 1, pp. 232-235). IEEE.
8. Sivasangari, A., Deepa, D., Anandhi, T., Ponraj, A., & Roobini, M. S. (2020, July). Eyeball based cursor movement control. In *2020 International Conference on Communication and Signal Processing (ICCSP)* (pp. 1116-1119). IEEE.
9. Narayanan, P., & Redy, S. K. (2022, December). A Generic Algorithm for Controlling an Eyeball-based Cursor System. In *2022 International Conference on Automation, Computing and Renewable Systems (ICACRS)* (pp. 117-120). IEEE.
10. Chen, C., Yin, J., & Fu, T. (2023, April). Investigation to Screen Region Division for Stable Input of Eye-controlled Cursor. In *2023 IEEE International Conference on Control, Electronics and Computer Technology (ICCECT)* (pp. 535-539). IEEE.
11. Sundstedt, V., & Garro, V. (2022). A systematic review of visualization techniques and analysis tools for eye-tracking in 3d environments. *Frontiers in neuroergonomics*, *3*, 910019.
12. Zhang, X., Liu, X., Yuan, S. M., & Lin, S. F. (2017). Eye tracking based control system for natural human-computer interaction. *Computational intelligence and neuroscience*, *2017*.

13. Soundarajan, S., & Cecotti, H. (2018, July). A gaze-based virtual keyboard using a mouth switch for command selection. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (pp. 3334-3337). IEEE.
14. Vasisht, V. S., Joshi, S., & Gururaj, C. (2019, June). Human computer interaction-based eye-controlled mouse. In *2019 3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)* (pp. 362-367). IEEE.
15. Mahalakshmi.S, S.Nirmal, Dr.L.Suriya Kala .Eye-Controlled Mouse Cursor . Ijetms;7(3):407-414.
16. Murata, A. (2006). Eye- gaze input versus mouse: Cursor control as a function of age. *International Journal of Human- Computer Interaction*, *21*(1), 1-14.
17. Kim, J., Jang, H., Kim, D., & Lee, J. (2023). Exploration of the virtual reality teleportation methods using hand-tracking, eye-tracking, and EEG. *International Journal of Human–Computer Interaction*, *39*(20), 4112-4125.
18. Stone, S. A., & Chapman, C. S. (2023). Unconscious Frustration: Dynamically Assessing User Experience using Eye and Mouse Tracking. *Proceedings of the ACM on Human-Computer Interaction*, *7*(ETRA), 1-17.
19. Shadiev, R., & Li, D. (2023). A review study on eye-tracking technology usage in immersive virtual reality learning environments. *Computers & education*, *196*, 104681.
20. Fernandes, A. S., Murdison, T. S., & Proulx, M. J. (2023). Leveling the playing field: A comparative reevaluation of unmodified eye tracking as an input and interaction modality for VR. *IEEE Transactions on Visualization and Computer Graphics*, *29*(5), 2269-2279.
21. Tang, Z., Liu, X., Huo, H., Tian, S., & Fan, Y. (2023). Eye movement characteristics in a mental rotation task presented in virtual reality. *Frontiers in Neuroscience*, *17*, 1143006.
22. Kanakaprabha, S., Arulprakash, P., Praveena, S., Preethi, G., Hussain, T. S., & Vignesh, V. (2023, May). Mouse Cursor Controlled by Eye Movement for Individuals with Disabilities. In *2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS)* (pp. 1750-1758). IEEE.
23. Kukona, A. (2023). Predictive sentence processing at speed: evidence from online mouse cursor tracking. *Cognitive Science*, *47*(4), e13285.
24. Shi, Y., Yang, P., Lei, R., Liu, Z., Dong, X., Tao, X., ... & Chen, X. (2023). Eye tracking and eye expression decoding based on transparent, flexible and ultra-persistent electrostatic interface. *Nature Communications*, *14*(1), 3315.
25. Raja, A. K., Sugandhi, C., Nymish, G., Havish, N. S., & Rashmi, M. (2023, April). Face Gesture Based Virtual Mouse Using Mediapipe. In *2023 IEEE 8th International Conference for Convergence in Technology (I2CT)* (pp. 1-6). IEEE.
26. Raja, M., Nagaraj, P., Sathwik, P., Khan, K. M. A., Kumar, N. M., & Prasad, U. S. (2023, March). Voice Assistant and Gesture Controlled Virtual Mouse using Deep Learning Technique. In *2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS)* (pp. 156-161). IEEE.
27. Muneeswaran, V., Nagaraj, P., Anuradha, L., Lekhana, V., Vandana, G., & Sushmitha, K. (2023, March). An Elucidative Review on the Current Status and Prospects of Eye Tracking in Spectroscopy. In *Congress on Control, Robotics, and Mechatronics* (pp. 225-236). Singapore: Springer Nature Singapore.

Design of Internet of Vehicles (IoV) based Vehicle to Vehicle Communication System

Nagaraj P
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Srivilliputhur, India
nagaraj.p@klu.ac.in

Muneeswaran V
Department of Electronics and Communication Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Srivilliputhur, India
munees.klu@gmail.com

G Haneesh
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Srivilliputhur, India
guntupallihaneesh@gmail.com

M Sripadh
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Srivilliputhur, India
Msripadhlion123@gmail.com

S Chaitanya
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Srivilliputhur, India
Skrishnachaitanya2003@gmail.com

C Ganesh
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education
Krishnankoil, Srivilliputhur, India
ganeshchandra3456@gmail.com

Abstract— Vehicle-to-vehicle (V2V) communication involves remote communication in which vehicles communicate with each other using interactive messages. Information will include speed, fire warning, problem management, loss of area, direction, response, and safety. The most important objective of progressing vehicle-to-vehicle communication through LI-FI is to dispose of cheap and life-threatening issues. Li-Fi requires Drove lights, such as those currently utilized in numerous energy-efficient homes and workplaces. These bulbs are prepared with a chip that subtly controls the light for optical information transmission. Li-Fi information is at that point transmitted by the Drove seed and gotten by the photoreceptor. Standard microcontrollers enable V2V communication via LI-FI. Node MCU is a microcontroller that controls the entire module. Data such as distance, fire, and emergency preparedness are processed using ultrasonic sensors, fuel sensors, and marking and disabling emergency switches that alert multiple vehicles when activated. The framework has tall information rates and complies with the 802.11bb determination. There's moreover information security. Car-to-car communication through Li-Fi is done on the Arduino Uno microcontroller. This comes about given the network of Bluetooth frameworks, the removal of the vehicle that can be sent, fire readiness, and the capacity to reply to crises. The farthest distance is about 10cm, and the compartment is controlled or controlled by ultrasonic sensors and an LCD screen, for example, if a fire situation occurs, the fire signal is recommended and if there is an emergency, the LCD screen is ready for fire, the LCD screen is emergency warning and ready to provide information to other vehicles.

Keywords— Sensors, V to V interaction, Li-Fi technology, Node MCU, IoT.

I. INTRODUCTION

As the number of vehicles on the highway continues to increase, traffic management becomes a critical daily test. Vehicle-to-Vehicle (V2V) communication may be a progressive innovation that makes a difference makes transportation frameworks more obvious. It permits vehicles to trade data at almost their speed, area, and activity over an amplified period. Later development of V2V communication allows the vehicle to send and receive various information. The plan will then address the use of visuals, devices, and visual alerts to alert drivers to specific actions to prevent accidents. Complete V2V is a reliable space for the coordination of traffic accidents by sending traffic information and offering drivers the fastest route. Various methods for vehicle switching, including Wi-Fi, Bluetooth, and radio adaptive networks, can guarantee long-distance and detailed information because radios do not have such happy ways. Here are some simple precautions to prevent accidents for different vehicles. Vehicles in this traffic zone are equipped with warning lights. These explorers took the connection of light and extracted information from the relevant environment in which light could be imagined. Li-Fi is common and stable in optical communication because it has a good transmission (10 Gbps), high speed, higher communication speed, and high limit of different wave types. Li-Fi was noted as an advance. Development is rapid and ecologically responsible. Li-Fi not only improves quality, it also provides a non-electromagnetic, safe environment. Use it quickly, without complex information or risk of attack by hackers. V2V communication is the best solution to reduce traffic accidents.

II. LITERATURE SURVEY

In [1], the V2V model makes sense. Using V2V communications, the vehicles can distinguish between locations and developments of different vehicles up to a quarter mile away. In practice, cars are equipped with individual radios, PC chips, and GPS innovations. Cars will know where most other cars are, and many cars will also know where you are, even if you are on the weak, invisible side of the highway, in a corner with external barriers, or blocked by various pressures on the car. If the driver does not react to the warning, the vehicle can be protected by staying a safe distance from the incident.

V2V communication in Wi-Fi was proposed in [2]. IEEE 802.11p is additionally known as Vehicle Conditional Inaccessible Get to (WAVE). This model will be utilized within the arrangement of the Short-Term Communications Contract (DSRC). The car sent direct messages from buyers to experts, facilitating close collaboration through the convenient DSRC. This will be a future level because it allows communication between vehicles. Considering the development of IEEE802.11, remote organization can be widely used in home organizations and workplaces with free Wi-Fi or low cost. The original concept of DSRC is to provide off-the-shelf mobility for the construction of tools such as IEEE802.11a. This development is efficient and can be improved and new models are created. The newest standard for wireless LAN (WLAN) is IEEE 802.11. The IEEE 802.11n standard endeavors to make strides and expand the best-known WLAN standard through expansion, optimization, and scope. V2V communication utilizing the ZigBee convention is portrayed in [3]. ZigBee could be a keynote for sensor organizations due to its long battery life, low operating costs, easy support, and high performance. Zigbee supports network organization by allowing remote communication between multiple organizations, switches, and receivers when testing multiple applications. The network connection is a perfect world, self-created and therefore exclusive, used mostly by organizations in the telephone network. Multiple light communication (VLC) is suggested in the article [4]. VLC offers many features as it can deliver unlimited, unencumbered, license-free content. Driving has ended up well known in car lighting due to its long life, moo taking a toll, capacity to resist vibration, and being more advantageous. VLC is another capable choice for electronic playback. It is a remote optical communication system that uses non-repeating light (400nm to 700nm). Metin [5] suggested knowing the distance information of two frames from visible light. This approach is simple and allows information to be sent faster than current innovations in long-distance communication. Using light for long-distance communication can solve the problem of the lack of radio waves in many places so that the message can be transmitted effectively. The proposed framework illustrates information transmission and collection by turning the drive on and off with a variable control as well quick to be obvious to the bare eye. Distinctive information can be changed by squinting the Driven and changing the communication. Viable and commercial innovation for vehicle-to-vehicle (V2V) communication utilizing light is portrayed in [6]. It was where two circumstances were inspected: when the car moderated down out of the blue, it made the cars below aware of this change of speed; When a fast car approaches an intersection, it warns other cars who can't tell the difference. The Li-Fi-based vehicle communication system aims to further improve the vehicle by eliminating traffic violations (such as speeding violations) and sending this information to the checkpoint, thus dispatching the vehicle using lights and lights. Health and transportation. Optometry correspondence with the rules to be followed from now on has been prepared in [7]. Light-emitting diodes will enter numerous ranges in daily lives. A curiously included of these instruments is their broad capabilities that can be utilized for information exchange [8]. LEDs require moo control to function and have exceptionally quick exchange times [9]. The venture points to forming a system for communication by utilizing lighting in indoor lighting. Data increase can be achieved by using the appropriate light source and OFDM can be used instead of cold interrogation [10].

In existing systems Vehicles are becoming more prevalent because of industrialization and urbanization, while road accidents are also becoming more prevalent. When accidents occur, someone should send a message to emergency services to save lives. When accidents occur, traffic is often a huge issue [11].

III. PROPOSED METHODOLOGY

V2V communication utilizing Li-Fi innovation comprises a transmitter and a collector, which shows the V2V message sent by the driver's front light. Li-Fi development employs light to transmit messages, hence diminishing the requirement for capacity space. The outline is controlled by the communication piece profile of the two vehicles. Both cars have Arduino, LCD Show, Ultrasonic Separate Estimation (UDM) Sensor, Fire Sensor, and Caution.

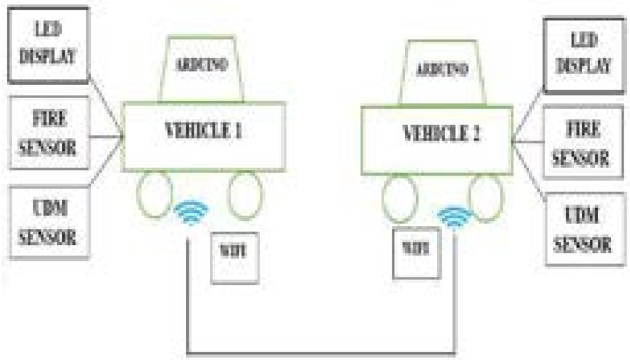


Fig 1: Communication between vehicles

Li-Fi modules contain LEDs. The advantage of consolidating LEDs into the Li-Fi system is that it increments the information rate compared to RF flag transmission.

The collector contains a photodiode that acts as a demodulator for comparing flags in a coordinated plan. The demodulated flag is at that point sent out of the channel, disposing of undesirable signals. Flag partition is presently moved forward by the flag improvement framework.

Filtered and reliable signals are provided for display devices such as LCD monitors.

Arduino can be used to create smart things through the collaboration of sensors and control of various lights, engines, and other benefits.

Ultrasonic sensors are utilized to degree the separation between two vehicles and recognize objects.

Electrical sensors are used to detect and identify objects between two vehicles. Respond in the presence of fire, caution should be exercised.

Fig. 2 and 3 describe the architectural design of sender and receiver information.

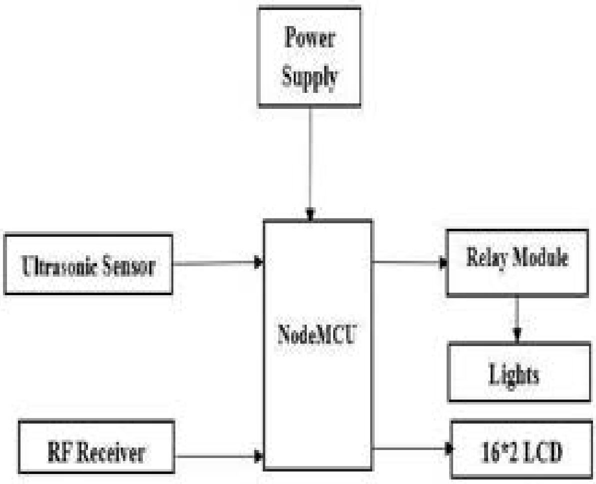


Fig 2: Block diagram of sender

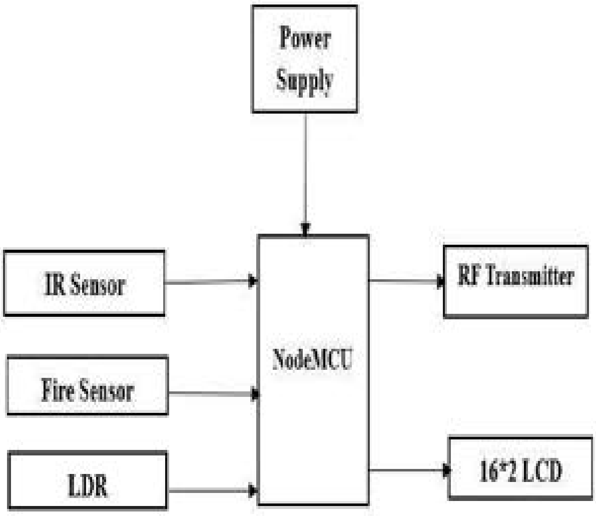


Fig 3: Block diagram of receiver

V2V data is transmitted via LED lighting. Since Li-Fi technology uses light to transfer data, it requires less storage space. The control system uses a device called a microcontroller, which can be implemented using Arduino, thus reducing the time. Show a communication diagram between two cars, each with Arduino, LCD, ultrasonic range cut sensor, electronic detector, and anti-theft alarm. Li-Fi modules are made of LEDs. Using LEDs in Li-Fi systems has the advantage of increasing throughput compared to Bluetooth signal transmission. The recipient encompasses a photodiode and a demodulator to get parallel signals. The demodulated flag is at that point sent through a channel that evacuates undesirable signals. Presently utilize the flag enhancement strategy to grow the channel flag. The sifted and increased signal is sent to the yield gadget LCD. Arduino can be utilized to make intelligent gadgets that acknowledge sensor input and control different lights, engines, and other outside components. Ultrasonic sensors are utilized to degree the separation of the vehicle and identify its current. Fire sensors react to the presence of fire and provide warning signals.

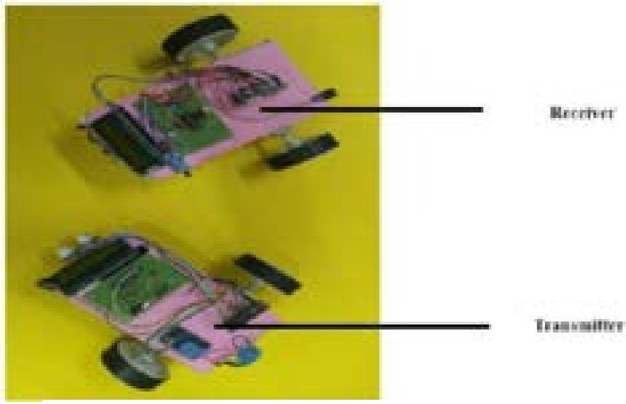


Fig 4: Model transceivers used in the prototype implementation.

IV IMPLEMENTATION AND RESULT

Table 1. Vehicle and the Respective Sensor Output

Vehicle	Fire Sensor	LED Display	Ultrasonic Sensor	Buzzer	Relay Module
Vehicle A	Yes	Warning: Fire Detected!	3 meters	Yes	Yes
Vehicle B	No	N/A	1.5 meters	No	No
Vehicle C	Yes	Warning: Fire Detected!	2 meters	Yes	Yes

1. *Move information starting with one vehicle and then onto the next vehicle.*

The foremost common strategy of exchanging information employing light is called Li-Fi. It appears how to exchange information initially from one car and after that to another car. Border events such as traffic accidents, fires, and emergencies are constantly monitored, and communication is provided between vehicles. In case of distance limit, ultrasonic sensors can distinguish and measure the distance before using the car light and can be used to find the option that controls the distance measurement. Create two-car ringtones. If an emergency occurs, sensors will be activated, and both vehicles will be notified of the emergency and can take the necessary steps to reduce damage. All frames are very low power and only need five volts of power.

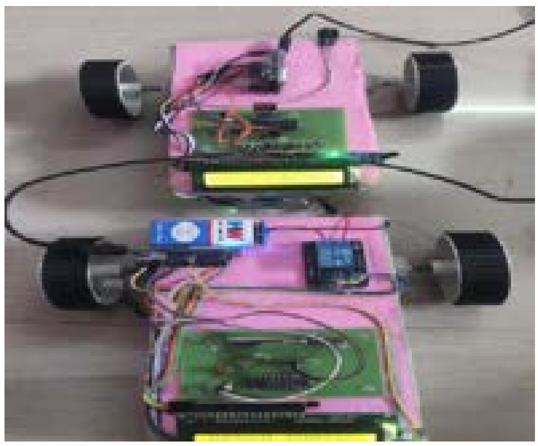


Fig. 5 Transmission of data beginning with one vehicle and then onto the following vehicle

2. Actuation of the bell in the event of a fire

The Li-Fi system is planned to empower car-to-car communication. On the off chance that a fire breaks out within the car, the fire finder will recognize the flag delivered by the sensor. The Arduino microcontroller controls this and sends the message to the main car alarm and the second car using the Li-Fi framework. The following car receives this flag and checks this information in its microcontroller and takes necessary actions. Figure 6 shows the alarm sound in case of fire.

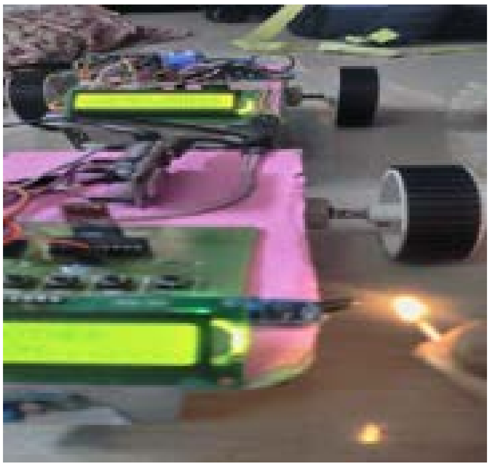


Fig 6. Buzzer sounding in case of a fire emergency.

3. Ultrasonic sensor for Traffic alert

If the Li-Fi framework could not able to pass the information to other vehicles then the system do not have sufficient time to remove other vehicle which may lead to collision. To overcome this ultrasonic sensor are used in the proposed model. Figure 7 shows how the ultrasonic sensor and Li-Fi system work. The ultrasonic sensor always sends a flag, and in case of an obstruction, the flag is reflected to the sensor. Based on the flag, the microcontroller takes decision and alerts the user to take another direction.

Fig. 7 Display Guidance for driver

4. Fire Accident Alert using RF.

The RF framework additionally helps in the quicker transmission of information. The IR Sensor exploits this property. The IR Sensor upon initiation conveys a message through the RF transmitter on which a emergency alert is shown on one more vehicle as displayed in Figure 8. Similar information is sent to the next vehicle. This is exceptionally valuable if the driver of the vehicle is encountering some sort of issue like a stir-up or tipsiness halfway through travel. The travelers in the two vehicles get cautioned, help each other immediately, and avoids the damage or predict a significant issue.

Fig. 8 Display of caution about fire accident

V. CONCLUSION

Approximately 150,000 people die every year due to driving impairment. The impact of these emergencies considers the lack of human driving that causes many accidents of central importance. As the vehicle moves more and more, driver and passenger comfort and convenience become increasingly important. The project aims to connect multiple parking areas to create a structure that will assist drivers on both highways and city streets. The concept incorporates an ultrasonic sensor to distinguish unsettling influences, a fire sensor to distinguish fire, and a caution button that can be squeezed in straightforward circumstances. The task was transferred to the center MCU controller, which is known for its great control and response time. In this way, using the proposed structure will reduce the damage and strengthen the safety and comfort of the driver.

REFERENCES

[1] I. Takai, T. Harada, M. Andoh, K. Yasutomi, K. Kagawa, and S. Kawahito, "V2V communication in intelligent transport system" in IEEE Photonics Journal, vol. 6, no. 5, pp. 1-14, Oct. 2014, Art no. 7902513.
[2] S. Haruyama, "Vehicle-to-vehicle communication using Wi-Fi" in Proc. 38th Eur. Conf. Exhib. Opt. Commun., Sep. 2012, pp. 1–3.
[3] T. Yamazato et al., "Survey on different wireless technologies used in V2V communication" IEEE Commun. Mag., vol. 52, no. 7, pp. 88–97, Jul. 2014.

[4] Shivaji Kulkarni, Amogh Darekar, and Pavan Joshi, "A Survey on Li-Fi Technology," 2016 International on Wireless Communications, Signal Processing and Networking (WiSPNET), March 2016.
[5] T. Komine et al., "Basic Study on Visible-Light Communication Using Light Emitting Diode Illumination," Proc. 8th International. Symp. Microwave and Optical Technology, 2001, pp. 45–48.
[6] A. Adwani and S.Nagtode, "Li-Fi information transferring through LED," in Proceedings of the 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT), pp. 2125–2127, IEEE, Chennai, India, March 2016.
[7] J. Karedal, N. Czink, A. Paier, F. Tufvesson, and A. F. Molisch, "High-speed data communication using LightFidility", IEEE Transactions on, vol. 60, no. 1, pp. 323 –328, 2011.
[8] A. Bazzi, B. Masini, G. Pasolini, and P. Torreggiani, "V2V communication for traffic safety" in Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on, 2010, pp. 1057 – 1064.
[9] Yunlu Wang, Stefan Videv, Harald Haas, "Dynamic load balancing with handover in hybrid Li-Fi and Wi-Fi networks", 2014 IEEE 25th Annual International Symposium on Personal, Indoor, and Mobile Radio Communication (PIMRC), pp. 575- 579, 2014.
[10] Yu. A. Vershinin, Yao Zhan, "Vehicle to Vehicle Communication: Dedicated Short Range Communication and Safety Awareness", 19-20 March 2020, DOI:10.1109/IEEECONF48371.2020.9078660.
[11] Nagaraj, P., Muneeswaran, V., Rupesh, C., Kartheek, B., Lekhya, Y., & Swetha, K. (2023, July). A Robust Recession Detective Analysis System using IoT Smart Sensor Devices. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)* (pp. 1260-1264). IEEE.

The Role of Digital Twins and Estimating their Impact on the Field of Agriculture in Promoting Sustainability

Shirley C P
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India.
shirleydavidlivingston@gmail.com

Brijit Benny
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
brijitbennyprofl@karunya.edu

Vidhya K*
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India.
vidhyak@karunya.edu

Berin Jeba Jingle I
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
berinjeba@karunya.edu

***Abstract*—Agriculture is the activity of cultivating natural resources that humans utilize for survival and economic benefit. It has evolved into one of the most significant industries in the day-to-day existence of all living organisms, providing access to food, a means of subsistence, etc. In reality, a country's agricultural output may have a direct or indirect impact on its food security and overall health. Despite the widespread adoption of these agricultural methods, various obstacles, including soil fertility and climate, can negatively impact the development and output of agricultural goods. As every living thing on Earth depends on agriculture for its existence, there is a constant need to find ways to overcome the problems faced. It is a well-acknowledged reality that contemporary issues demand modern solutions. Thus, it is critical to use developing technology to tackle the challenges that exist. The quality of life has been greatly enhanced by modern technologies. This concept has led to the introduction of several contemporary inventions and technologies that enhance the standard of life and advance the welfare of every individual. The demand for new technologies has been growing over time, and one promising technology with the potential to improve living standards and address long-standing issues in the agricultural sector is digital twin technology. It has made it possible to reduce labor and energy waste, among other things. Examining how digital twins might improve agricultural productivity and promote sustainability is part of the process of understanding digital twins and their place in the agricultural industry. The prerequisites and implementation strategies for encouraging the use of digital twins are covered. Since this is a new technology, a careful analysis of its benefits and limitations is necessary to properly address identified issues.**

***Keywords*—sustainability, Digital Twins, agriculture, sensors, IoT, Artificial Intelligence**

I. Introduction

The agricultural sector is the very backbone keeping a country alive and healthy. It is essential for food production, economic stability, and environmental sustainability, impacting land use, water resources, and biodiversity [1, 2]. When agriculture thrives, the economy booms along. But real prosperity shows when green farms nourish entire communities year after year, not just through economic measures. Still, challenges hamper further betterment of rural belts, from climate threats to crops to promoting nature-friendly techniques that last over time.

With such tricky real-world problems, new-age technologies provide rays of hope instead of doom. Artificial Intelligence or AI is one such forward-thinking innovation now being deployed everywhere - factories, colleges, government - to drive progress. At the core, AI refers to creating smart systems displaying almost human-style analytical smarts, just minus the emotions. These methodical AI tools can continuously swallow huge amounts of raw data, identify hidden patterns and even predict complex behaviours beyond human specialists. Having such 'digital assistants' means making sense of messy data now needs much lower investment of time, money and manpower.

One groundbreaking application involves creating digital replicas of physical assets, such as jets, rails, or entire smart city power networks, known as digital twins. These digital twins are designed to mimic their real counterparts through the integration of cutting-edge technologies, including IoT, cloud computing, AI, and XR, facilitating their widespread application [3]. The incorporation of IoT is particularly notable in these virtual replicas, enabling seamless real-time connectivity and data exchange between physical assets and their digital counterparts. Equipped with IoT capabilities, these 'digital twins' act as round-the-clock digital sensors, capturing intricate performance statistics without the need for expensive manual tracking or risky physical experiments on valuable infrastructure.

Think - testing changes on a live train disrupts rail movement crucially. But its digital clone allows endless software simulations to assess upgrades in isolation. Such cyber counterparts enable closely inspecting minute material behaviours, wear-and-tear issues impossible to regularly test in the field. Infrastructure authorities from defence to urban planning are now using such helpful digital shadows to light up smarter direction. And as computing strengths surge rapidly, the promise of such realistic digital environments to optimize physical realities will only swell.

Let's consider urban planning authorities aiming to expand public transport infrastructure. Physical trials involve massive investments for deploying trains, fuel systems, stations along with severe operational risks. A digital twin prototype allows modelling countless options spanning fares, routes, wagon capacities all within flexible software environments. Experts can identify optimal configurations maximizing affordability, accessibility, and sustainability goals before even allotting physical land. Such virtual mock-ups powered by real-time data also continue advising maintenance needs, rider estimates over decades rather than a one-time fix.

Now think of healthcare - a digitized patient twin incorporating genetics, diagnosis history, and treatment response can enable customized therapy. Pharma researchers can investigate countless drug combinations tailored to individual DNA strains using such profiles, considerably accelerating successes. The future possibilities with such made-to-order digital blueprints remain manifold. No wonder pioneering applications have already secured strong funding. As computing power and design thinking both witness exponential gains, the problem-solving potential of digital shadows mimicking and optimizing physical systems seems all set to illuminate solutions faster than otherwise possible.

The main objectives of the research are:

- Understanding Digital Twins
- Understanding how digital twins can impact on the field of agriculture
- Understanding how sustainability can be achieved using digital twins
- Analysing various existing limitations

II. BACKGROUND

A. *Digital Twins*

Digital twins are basically intelligent virtual representations of physical entities that offer real-time monitoring, analysis, evaluation, and prediction of physical systems [4]. In other words, they are very detailed virtual copies of real physical things which allows the testing of improvements on them without touching the actual item and allow physical and remote monitoring of the livestock [5]. For example, think of digitally reproducing an actual running jet engine into a software environment - matching nuts, bolts, pipes etc. Now engineers can tweak this cyber clone endlessly to push limits without taking the real engine offline.

These software duplicates assimilate both static and dynamic inputs from multiple measuring tools around the assets like high-resolution cameras, heat and vibration sensors etc. So, they transform from inert 3D diagrams to almost live, breathable models of the physical equipment.

Consider the jet engine twin receiving live temperature, pressure and speed data every fraction of a second from the operating machine. By processing this running stream of numbers, the digital copy starts mirroring the real-time performance parameters of the engine within the simulation world. So, without expensive and risky real-world trials, engineers can now stress-test virtual components to extreme boundaries for finding design upgrades.

Now consider digital twin powered control centres for entire cities! Facets like traffic signals, power distribution grids, water pipelines get monitored 24x7 through such tailored virtual replicas assimilating real-time feeds. Municipal authorities can now catch emerging urban chokepoints, infrastructure risks much earlier by leveraging such city-scale digital estates.

Creating digital shadows begins by digitally modelling the physical asset down to fine details like shape, material types, structural assembly etc. However, mirroring operating dynamics like asset deterioration rates, breakdown risks require continuous data integration from surrounding sensors. So, experts carefully install measurement hardware across critical machines and infrastructure for this.

Advanced analytical software thereafter converts accumulated sensor data into meaningful insights - be it detecting anomalies, predicting maintenance needs or simulating overhaul responses for industrial equipment or public amenities before actual breakdowns. Digital twins have thus become integral for diverse sectors to simulate expansions, quantify risks and boost efficiency.

In healthcare, they enable patient-specific diagnostics and treatment evaluations before deploying actual interventions. For smart city projects, virtual models validate different design proposals to find optimal configurations even before laying the first bricks.

By enabling virtual prototyping of intricate systems, digital twins can drive productivity across areas while minimizing experimentation risks. As computing and analytical abilities accelerate, such cyber twins hold enormous potential in advancing how anything is setup and managed - from small appliances to complex machinery to even entire smart cities!

B. *Digital Twins In Agriculture*

For generations, the process of improving agriculture has been gradual and resource intensive, relying on farmers' practical learning through trial-and-error experiments. However, given the increasing urgency of global food security and sustainable environmental stewardship, persisting on the current trajectory is not viable if the aim is to meet expanding needs. The solution may lie in emerging digital twin technology - highly sophisticated computer simulations that create "living" virtual crop environments to rapidly test innovations at a fraction of the cost, risk and time. They integrate digital models of physical objects to optimize performance and prevent failures using real-time data [6]. As mentioned in [7], they have the potential to revolutionize agriculture by creating a virtual representation of a physical system, allowing for the simulation of the impact of various events and actions. Fig. 1 illustrates the prevalence and impact of digital twins in agricultural settings.

Just as architects develop 3D building information models with painstaking detail before physical construction, it is now possible to leverage massive agricultural datasets and computing power to digitally prototype the next era of farming. Teams of researchers worldwide could collaborate to examine endless variations of seeds, soils, weather patterns, watering schedules, and other growth factors all within these breathtakingly intricate simulated environments. Minute genetic differences between seed strains could be cross-analysed with geographic climate models, seasonal solar cycles, soil nutrient compositions and water conservation strategies to reveal optimum combinations for maximizing crop quality and yield over a full life cycle. It also allows farmers to remotely monitor and control greenhouse operations, making agriculture more efficient and sustainable while addressing labour shortages [8].

Where physical fields may allow dozens of test trials across a few acres per growing season, continuously run simulations permit limitless controlled experiments across millions of virtual acres globally. The most promising solutions are flagged for targeted real-world prototyping while the bulk of inferior ideas are discarded outright without wasted land or resources. And as digital platforms ingest more observational data from these selective physical validations, the virtual crop representations become increasingly sophisticated and reliable for future testing. It is a self-improving cycle between physical and digital research where each new iteration is more productive than the previous as virtual and reality converge.

This integrated virtual-to-real approach may be how rapidly agriculture is advanced to address impending global food security challenges. Optimizing yields through digitally-enhanced innovation pipelines is possible without expanding farmland footprints. Virtual cropping technology buys enough time to balance mounting population needs with environmental sustainability in the critical decades ahead.

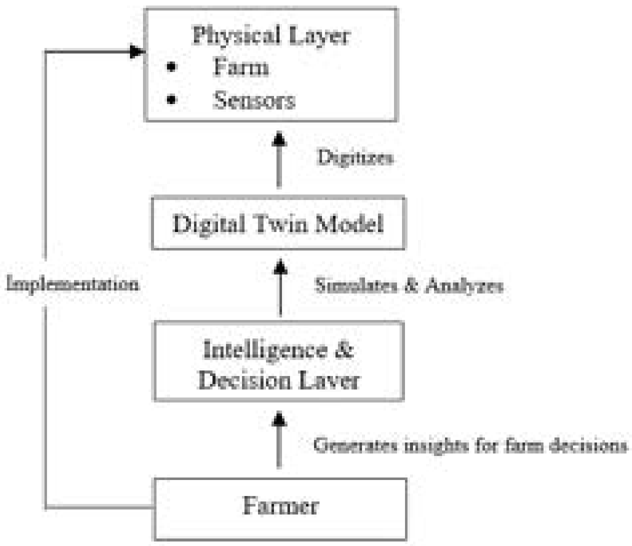


Fig. 1. A simplified representation of Digital Twins in Agriculture

C. Understanding the importance

When implementing digital twins in agriculture, it is important to understand the possible outcomes and its importance.

- *Significance for farmers*

Digital twins provide farmers various layers of actionable advice to help them connect with sustainability goals without sacrificing output or profitability. It is now possible to conduct real-time monitoring and analysis of crop conditions, leading to improved decision-making and resource management. Dynamic feedback is supplied into the digital platform via ongoing soil, crop, and meteorological condition gathering. This makes it possible to precisely adjust important inputs, like as water, fertilizer, herbicides, etc., to the ideal quantities in order to maximize production. In order to implement preventative care, the system furthermore promptly notifies users in the event of an approaching insect assault, equipment malfunction, or extreme weather occurrence. As a result, farmers have a wise partner that uses technology to offer guidance on infrastructure upkeep, operations, and productivity.

- *Environmental Benefits*

The digital twin solution's suggestions include provisions for responsible resource management. As a result, even when output objectives are appropriately reached, there is no ecological impact. When regenerative agricultural methods are gradually used, sustainability indicators such as biological richness, groundwater levels, and soil quality really improve over several harvest cycles. Climate-smart agriculture thereby gains traction without requiring farmers to take on significant risks or expenditures. It is thereby understood that digital twins enable efficient management and optimization, leading to reduced energy consumption and environmental impact in agricultural operations [9].

- *Societal Impact*

The agricultural community may more easily and affordably embrace high-tech, precise procedures by mainstreaming innovative technologies like digital twins. Progressive farmers may justify their expenses with higher yields and revenue, and increased capacity can help meet the growing demand for food. Sustainable food production enhances the well-being, independence, and prosperity of rural communities, who comprise a significant portion of India's population. Digital twins have now become a reason for the improved sustainability, decision support and supply chain transparency [10].

- *Promoting regenerative agriculture*

Indicators of soil health, biodiversity, and carbon sequestration are tracked via the digital twin platform. This makes it possible to use corrective measures like polyculture cropping, organic fertilizers, low-cost natural farming methods, etc. Such regenerative techniques can, over time, improve long-term resilience by aiding in the farm ecosystem's own regeneration.

Using digital twins in farming doesn't just have one set of positive results; it can bring various benefits as techniques improve. So, it's fair to say that adopting digital twins in agriculture can make life better not only for individuals but also for the well-being and lifestyle of future generations.

D. Achieving Sustainability Using Digital Twins

Digital twin technology brings an innovative revolution that can transform sustainable agriculture in India. Referencing [11], it’s understood that digital twins in agriculture help achieve sustainability by providing real-time monitoring, precise control, and accurate prediction for accurate and efficient smart agriculture. It can assist in finding possibilities for waste reduction, energy savings, and general sustainability improvements by modelling and forecasting the behaviour of complex systems [12,13]. These advanced computer simulations mirror real farms, ingesting massive data from soil moisture levels to insect populations to individual cow movements. Utilizing decision layers, digital twins function as hyper-attentive virtual assistants, carefully analysing this flood of farm information to generate targeted interventions that simultaneously maximize yields and minimize environmental impact.

Envision a digital twin reviewing real-time data on crop growth rates, soil nutrient composition, rainfall distribution, grazing patterns, and more. It swiftly prescribes precise application of fertilizers, water, and pesticides down to square meter granularity across fields for optimal results. When sensor data indicates livestock overgrazing in a certain area, the twin rapidly recalibrates grazing zones and schedules to prevent long-term soil degradation. It examines crop cycles, biodiversity trends, and weather models to predict how farming choices today will cascade over decades, allowing farsighted stewardship of the land for future generations.

But more than just superior production, digital twins enable protection of vital ecosystems thriving alongside farmland by tracing interconnected relationships. The simulated advisor studies pollinator populations, wildlife migrations between habitats, soil food web health, and forest conservation to implement mutually beneficial strategies like planned biodiversity zones and biological pest control. It perpetually fine-tunes this balance between agricultural output and environmental flourishing through exacting analysis of factors across the ecological spectrum. As costs inevitably decline, these artificial intelligence allies will become accessible for smallholder farms across India. Digital twins in an agricultural era provides sustainability and productivity, technology and tradition are integrated in harmony. They form the foundation for communities to prosper in sync with the living landscapes around them. By deducing precise insights from a deluge of data, digital twins show a path where nourishing people and nature are not opposing forces to be traded off but rather complementary forces to be optimized in unison.

E. Existing Limitations Of Digital Twins

As digital twins become more widely used in sectors such as manufacturing, healthcare, and smart cities, legitimate privacy and security issues are brought up. While digital twins have many advantages, it is important to be aware of their drawbacks.

Even though the limitations of digital twins in agriculture also include data management challenges, model accuracy, scalability issues, and the need for high-quality data [14], permission and privacy can be considered as the most important ones. Large amounts of data must be collected from several sources, including sensors on physical assets, as data is the basis of digital twins. Selecting the right sensors is crucial for the effectiveness of digital twins. Considerations such as accuracy, reliability, and compatibility with the agricultural environment must guide the sensor selection process. Additionally, ensuring scalability and the ability to withstand varying environmental conditions are paramount for sustained and effective data collection. That being said, a significant amount of this data collecting violates people's right to privacy because it is done without their consent. For example, a digital duplicate of a public park may track foot movement without users' permission by using motion sensors and cameras. Digital twins developed for the healthcare sector using patient medical records may also violate privacy if appropriate data protections are not implemented. Even if the data is anonymized, patients' agreement was not secured for secondary usage in digital twin creation and analytics.

Another major limitation stemming from increased connectivity and data aggregation in digital twins are cybersecurity issues. The more operational and sensory data that is moved from physical assets to their digital counterparts, the bigger the attack surface becomes for malicious actors. In the event that operational data is breached or personally identifiable information is stolen, inadequate security mechanisms for the transmission and storage of digital twin data provide risks to both privacy and safety. Inaccurate information obtained through a hack, for example, that alters data going from a power plant to a digital replica, might lead to catastrophic malfunctions. Parallel to this, hacked digital twin data from wearables or medical devices can reveal personal user information.

Furthermore, digital twins are intrinsically more susceptible to cyberattacks due to their intricate interfaces with many IT/OT systems and data sources. These hazards are increased by the multifaceted nature of their connections between several virtual and actual components. Sophisticated hackers may be able to take advantage of more attack avenues due to the lack of common designs across various digital twin implementations. An example of a smart city digital twin is a virtual layer that is networked and combines data from emergency response systems, electricity grids, water networks, and transportation systems. It is possible to go further into the network and obtain access to other vital municipal infrastructure by exploiting obscure weaknesses in even one component's digital incarnation.

Since digital twins serve as the foundation for AI/ML-based autonomous decision making, inaccurate analytics resulting from hacked data or algorithms provide many more dangers than simply privacy breaches. By manipulating the virtual model, one may fool the twin into recommending risky operations on the real asset by changing the simulation results. Attackers may, for instance, be able to induce catastrophic failures on a linked power system by infiltrating its digital form with malware and then suggesting dangerous sequences that, when followed, result in blackouts.

While conventional industrial control systems are still vulnerable to cybersecurity attacks, the attack surface is greatly expanded by the increased connectivity and reliance on potentially sensitive data that come with digital twins. To proactively identify threats to IT/OT systems and prevent breaches, a robust, multilayered security architecture must be implemented as digital twins grow more interconnected and appear in public places such as smart cities.

In addition, results from digital twin simulations may differ from actual results due to unclear data or incorrect modelling assumptions. These differences are concerning as businesses are depending more and more on digital twins for automation and forecasting. It is important to understand that the accuracy of predictions depends on model adequacy [15], which can be enhanced with the use of appropriate sensor, satellite, historical, simulation, and contextual data [16]. This involves ensuring the quality of data, including high-quality sensor data, satellite imagery, and contextual information. Continuous monitoring, regular updates, and real-world observations are essential to refining and enhancing the accuracy of the digital twin model over time. In summary, the interpretability, accuracy and real-world applicability of AI/ML underpinning digital twins needs to be enhanced to enable trust, manage risks from erroneous insights, and ensure successful deployments at scale.

So, in conclusion, it is important to tackle the technical, social, and ethical challenges, promote inclusive design, transparency, and open dialogue, and consider the availability and quality of real-time data in order to create accurate digital twins as noted in [17]. While digital twins offer transformative potential, challenges such as data management, model accuracy, and scalability must be acknowledged. Privacy concerns and cybersecurity issues, including the interpretability of AI/ML underlying digital twins, need to be addressed to ensure successful deployments at scale.

F. Overcoming The Privacy And Security Concerns : Digital Twins In Agriculture

The idea of digital twins has completely changed how farmers handle crop management in the agricultural sector. Farmers may simulate various scenarios and test novel strategies before putting them into practice in the field by building virtual representations of real-life plants. Nonetheless, privacy and safety are issues that come with every technology development. Digital twins in agriculture generate and rely on large volumes of data, making data security, cybersecurity, and supply chain security crucial to prevent potential vulnerabilities and ensure operational integrity [18]. In the coding process of building a digital twin for agriculture, a number of creative techniques may be used to overcome these issues.

One strategy is to put strong encryption techniques in place and give data security and privacy a priority. This guarantees the protection of critical agricultural data from unwanted access, including growth trends and production forecasts. Farmers may be safe in the knowledge that their data is private and secure by employing robust encryption mechanisms.

A different approach is to include sophisticated authentication procedures into the digital twin framework. The digital twin can only be accessed and modified by authorized personnel by employing strategies like biometric identification and multi-factor authentication. This guarantees the integrity of the data and prohibits unauthorized manipulation with the virtual plants.

The digital twin system must be equipped with real-time monitoring and anomaly detection algorithms in order to improve safety. All anomalous behaviour, including abrupt alterations in growth patterns or anomalous environmental circumstances, may be quickly identified by continually evaluating data from sensors positioned throughout the fields. Farmers can respond quickly to reduce hazards and safeguard their crops using this early detection.

Furthermore, the safety and privacy features of the digital twin system may be further improved by using machine learning techniques. Through ongoing data collection and pattern analysis, the digital twin can detect possible weaknesses and recommend enhancements to security protocols. By being proactive, farmers can keep their digital twin system safe and stay one step ahead of possible dangers.

In addition, cooperation and information exchange between farmers and agricultural specialists can be extremely important in resolving privacy and security issues. Farmers may collaborate to create standardized security rules and regulations unique to digital twin systems in agriculture by creating safe spaces for information sharing and best practices. Through this cooperative endeavour, the agricultural community as a whole will gain from the collective knowledge and experiences, resulting in the adoption of digital twins that are safer and more secure.

In conclusion, eensuring the security of digital twins in agriculture involves implementing robust encryption, authentication procedures, and real-time anomaly detection. Machine learning techniques and collaborative efforts among farmers and specialists are essential components for strengthening the overall security framework. As noted in [19], data access should be restricted, encryption protocols employed, and clear policies for data handling established. Farmers may get over these worries and fully profit from digital twin technology by putting into practice cutting-edge techniques like strong encryption, sophisticated authentication, real-time monitoring, anomaly detection, machine learning, and encouraging teamwork. The agriculture sector may use digital twins with confidence, resulting in more effective and sustainable farming operations, by guaranteeing data security, safeguarding privacy, and improving safety precautions.

III. Discussion

This study investigates the application of digital twin technology in agriculture, exploring its potential impact on agricultural practices and sustainability. The research focuses on the unique advantages of digital twins, aiming to understand their role in revolutionizing crop management and precision agriculture.

The findings of this research highlight the transformative implications of digital twins, offering an alternative to traditional, resource-intensive methods. Unlike conventional approaches relying on trial-and-error experiments, digital twins enable rapid testing of innovations at a reduced cost, minimizing risks and saving time. The ability to create digital replicas of farming processes allows for endless simulations, facilitating collaborative exploration of diverse variables such as seeds, soils, weather patterns, and growth factors.

A key contribution of this research lies in its exploration of how digital twins can enhance crop management and precision agriculture. Researchers can leverage massive datasets and computing power to digitally prototype farming processes, optimizing inputs with precision. This not only enhances productivity but also addresses labour shortages and promotes sustainable agricultural practices.

The study identifies the significance of digital twins for farmers, environmental benefits, societal impact, and their role in promoting regenerative agriculture. Farmers, equipped with real-time monitoring and analysis, receive actionable insights for resource management, infrastructure upkeep, and improved productivity. The environmental benefits extend to responsible resource management, contributing to climate-smart agriculture without compromising output. The societal impact is evident in the increased adoption of high-tech farming practices, meeting the growing demand for food while enhancing the well-being of rural communities.

To address the novelty aspect and improve upon existing models, the research emphasizes the need for a nuanced understanding of digital twins in agriculture. The study delves into the complexity of digital twin technology, detailing its operation, assimilation of static and dynamic inputs, and the role of IoT in creating comprehensive virtual replicas. By drawing parallels with applications in healthcare and urban planning, the research establishes the versatility and scalability of digital twins, showcasing their potential in diverse domains.

Despite these promising outcomes, the study acknowledges existing limitations, including data management challenges, model accuracy, scalability issues, and privacy concerns. To enhance the novelty and practicality of digital twin applications, future research could focus on addressing these challenges. Emphasizing the importance of accurate, real-time data, the research suggests ongoing monitoring, regular updates, and collaboration to refine digital twin models. Additionally, the study underscores the need for a robust, multilayered security architecture to mitigate cybersecurity risks, ensuring privacy and safety in the era of interconnected digital twins.

In conclusion, this research sheds light on the transformative potential of digital twins in agriculture, emphasizing their role in sustainability, precision farming, and societal well-being. The findings contribute to the existing body of knowledge by providing insights into the nuanced applications of digital twin technology and offering valuable recommendations for overcoming challenges. Future research endeavours can build upon these foundations, further advancing the field of digital twins in agriculture.

IV. Conclusion

Sustainable agriculture has advanced significantly with the advent of digital twins. A digital twin is essentially a sophisticated virtual model created through simulations and real-time data integration that replicates every facet of an actual farm. This convergence of the digital and physical worlds opens up new applications to improve farming's accuracy, predictability, and environmental responsibility.

Digital twins propel modernization in harmony with sustainability to concurrently improve rural welfare, ecological stability, and production. Next-generation agriculture is enabled by the clever combination of sensors, software, and analytics, which creates a platform for digitally replicating agricultural systems as living models. Thus, the secret to feeding India's expanding population sustainably while preserving the planet's resources is digital transformation. In conclusion, there is a great deal of promise for output increases without causing ecological damage via digital replication and holistic agriculture optimisation. Therefore, using digital twins to integrate advanced technology with moral obligations is the path forward for sustainable farming.

V. Future scope

The future scope of digital twins in agriculture holds promise for further advancements. Ongoing research can focus on refining decision layers, improving data integration techniques, and addressing evolving cybersecurity challenges. The integration of emerging technologies and the development of standardized frameworks for digital twins could lead to more widespread adoption, fostering sustainable practices in agriculture.

References

[1] Mark Ryan, "The social and ethical impacts of artifcial intelligence in agriculture: mapping the agricultural AI literature", AI & SOCIETY, 38:2473-2485, 20203. https://doi.org/10.1007/s00146-021-01377-9.

[2] Christos Pyliandis, Sjoukje Osinga, Ioannis N, Athanasiadis, "Introducing digital twins to agriculture", Computers and Electronics in Agriculture, 184:1-25, 2021. https://doi.org/10.1016/j.compag.2020.105942.

[3] Mohsen Attaran a, Bilge Gokhan Celik, "Digital Twin: Benefits, use cases, challenges, and opportunities", Decision Analytics Journal, 6:1-10, 2023. https://doi.org/10.1016/j.dajour.2023.100165.

[4] Stefan Mihai, Mahnoor Yaqoob, Dang V. Hung , William Davis, Praveer Towakel, Mohsin Raza, Mehmet Karamanoglu, Balbir Barn, Dattaprasad Shetve, Raja V. Prasad, Hrishikesh Venkataraman, Ramona Trestian , Huan X. Nguyen, "Digital Twins: A Survey on Enabling Technologies, Challenges, Trends and Future Prospects", IEEE Communications Surveys & Tutorials, 24(4): 2255 – 2291, 2022. doi: 10.1109/COMST.2022.3208773.

[5] https://geopard.tech/blog/what-are-digital-twins-in-agriculture/

[6] Stefano Cesco, Paolo Sambo, Maurizio Borin, Bruno Basso, Guido Orzes, Fabrizio Mazzetto, "Smart agriculture and digital twins: Applications and challenges in a vision of sustainability", European Journal of Agronomy, 146:1-9, 2023. https://doi.org/10.1016/j.eja.2023.126809.

[7] Natasja Ariesen-Verschuur, Cor Verdouw, Bedir Tekinerdogan, "Digital Twins in greenhouse horticulture: A review" Computers and Electronics in Agriculture, 199:1-17, 2022. https://doi.org/10.1016/j.compag.2022.107183.

[8] A F Rogachev, N N Skiter, N V Ketko, A B Simonov, I V Makarevich, "Digital twins as a tool for systemic integration of innovative digital technologies in agriculture", Agricultural Systems, 189:1-19, 2021. doi:10.1088/1755- 1315/1069/1/012042.

[9] Warren Purcell, Thomas Neubauer, Kevin Mallinger, "Digital Twins in agriculture: challenges and opportunities for environmental sustainability", Current Opinion in Environmental Sustainability, 61:1-7, 2023. https://doi.org/10.1016/j.cosust.2022.101252.

[10] Jing NIE, Yi WANG, Yang LI, Xuewei CHAO, "Artificial intelligence and digital twins in sustainable agriculture and forestry: a survey", Turkish Journal of Agriculture and Forestry, 46(5):642-661, 2022. https://doi.org/10.55730/1300-011X.3033.

[11] Robert Miehe, Lara Waltersmann, Alexander Sauer, Thomas Bauernhansl, "Sustainable production and the role of digitaltwins–Basic reflections and perspectives", Journal of Advanced Manufacturing and Processing, 3(2):1-21, 2021. https://doi.org/10.1002/amp2.10078.

[12] Rui Carvalho, Alberto Rodrigues da Silva, "Sustainability Requirements of Digital Twin-Based Systems: A Meta Systematic Literature Review", Applied Sciences, 11(12):1-14, 2021. https://doi.org/10.3390/app11125519.

[13] Nikolaos Peladarinos, Dimitrios Piromalis, Vasileios Cheimaras, Efthymios Tserepas, Radu Adrian Munteanu, Panagiotis Papageorgas, "Enhancing Smart Agriculture by Implementing Digital Twins: A Comprehensive Review", Sensors, 23(16):1-38, 2023. https://doi.org/10.3390/s23167128.

[14] P O Skobelev, I v Mayorov, E V Simonova, O I Goryanin, A A Zhilyaev, A S Tabachinskiy, VV Yalovenko, "Development of models and methods for creating a digital twin of plant within the cyber-physical system for precision farming management", Journal of Physics: Conference Series, 1703:1-18, 2020. https://doi.org/10.1088/1742- 6596/1703/1/012022.

[15] C.N Verdouw, J.W. Kruize, "Digital twins in farm management: illustrations from the FIWARE accelerators SmartAgriFood and Fractals", 7 th Asian-Australasian Conference on Precision Agriculture 2017, (pp.1-7) https://api.semanticscholar.org/CorpusID:64574388.

[16] Asaf Tzachor, Soheil Sabri, Catherine E. Richards, Abbas Rajabifard and Michele Acuto. "Potential and limitations of digital twins to achieve the Sustainable Development Goals", nature sustainability, 5:822–829, 2022. https://doi.org/10.1038/s41893-022-00923-7.

[17] M. Gabriele, A. Cazzani, C. M. Zerbi, R Brumana, "Digital Twin to monitor, understand and preserve the complexity of multi-scale natural, agricultural, designed landscapes and architecture : Biodiversity Conservation, Transformation and Decline at Villa ArconatiSite at Castellazzo of Bollate(MI)", The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 613—620, https://doi.org/10.5194/isprs-archives-XLVIII-M-2-2023-613-2023.

[18] Abozar Nasirahmadi, Oliver Hensel, "Toward the Next Generation of Digitalization in Agriculture Based on Digital Twin Paradigm", Sensors, 22(2):1-16, 2022. https://doi.org/10.3390/s22020498.

[19] Cor Verdouw, Bedir Tekinerdogan, Adrie Beulens, Sjaak Wolfert, "Digital Twins in smart farming", Agricultural Systems, 189:1-19, 2021. https://doi.org/10.1016/j.agsy.2020.103046.

Object Color Identification and Classification using CNN Algorithm and Machine Learning Technique

Mrs.R.Kiruba
Assistant Professor, EIE
Sri Ramakrishna Engineering College
Coimbatore,India
kiruba.r@srec.ac.in

Ms.V.Sneha
UG Student, Sri Ramakrishna Engineering College
Coimbatore,India
sneha.2206047@srec.ac.in

Ms.S.Anuragha
UG Student, Sri Ramakrishna Engineering College
Coimbatore,India
anuragha.2206002@srec.ac.in

Ms.S.K.Vardhini
UG Student, Sri Ramakrishna Engineering College
Coimbatore,India
vardhini.2206052@srec.ac.in

Ms.V.Vismaya
UG Student, Sri Ramakrishna Engineering College
Coimbatore,India
vismaya.2206053@srec.ac.in

Abstract—The need for automation in the textile industry is growing rapidly today. Color based object sorting is a highly challenging process to be considered and needs to be addressed. It involves an automated material handling system. It synchronizes the movement of robotic arm to pick up objects moving on a mobile robot. It aims in classifying the colored objects then picking and placing the objects in its respective pre-programmed place. Thereby eliminating the monotonous work done by human, achieving accuracy and speed in the work. The core objective of the project is to propose an intelligent color-based object sorting system using deep learning technique like Convolution Neural Network for extraction of feature embedded with the machine learning algorithm. The two classifiers Random Forest and K-NN algorithm were implemented and studied for better classification. Based on the performance metrics, the Radom Forest algorithm out performs in classification. The project module involves cameras that captures the object's color through the computer vision Library and sends the signal to the controller. The dataset of the captured images has been uploaded and compared with the trained data set. The ESP 32 Module transmit a signal to relay circuit, which then drives the robotic arm's multiple motors to grip the object and position it in the given area. Based on the color observed, the robotic arm goes to the given point, releases the object, and returns to its original position.

Keywords — Computer Vision Library, Robotic Arm, Convolution Neural Network (CNN), Random Forest (RF), k means Nearest Neighbor (KNN)

I. INTRODUCTION

The way that things are differentiated, sorted, and used in many industrial applications depends heavily on their color. However, if colour sorting is done manually, it would be a highly monotonous, time-consuming, and laborious task, thus it is crucial to develop machines and automate them to make every task more exact and optimal while also making the job of humans easier. Automation is not the newest innovation, but it has undoubtedly had the biggest impact because it is a good step towards handling different sorts of machinery and procedures to limit the need for human involvement, which saves enough time. On a related point, use of this automation to classify objects based on colour in an industrial setting where many products are produced every day, it might play a significant role. Already there exists a traditional way of sorting an object based on the color that are to be picked by the conveyor belt and then placed in the designated portion based on its color feature.

Here the main goal of the project is to flexibly sort the object based on the color and it should act intelligently like human for color discrimination through utilization of deep learning techniques. The Convolution Type Neural Networks has been proposed to implement the system to act like an intelligent system that differs from the traditional system that uses manual method of sorting. After detecting the color, a signal has been sent to the ESP 32 controller which directs the signal to operate the servo motors and grippers for picking and placing of the objects. By this a system of automated color-based object sorting is thus achieved. Color-based item sorting has several advantages over standard sorting algorithms. It eliminates the need for complicated and expensive mechanical setups, speeding the sorting process and lowering operational costs. Furthermore, its versatility enables smooth integration with a variety of industries, ranging from manufacturing and logistics to recycling and waste management. This cutting-edge technology enables perfect sorting, minimizing errors and enhancing productivity with unsurpassed accuracy and speed. Its capacity to handle objects of various shapes and sizes makes it an invaluable tool for a wide range of applications.
Currently, the color-based object sorting system based on deep learning provides the way for a more efficient, cost-effective, and sustainable approach to object sorting, moving industries forward. The color-based object sorting system makes use of CNNs, a type of deep neural network noted for its superior image processing feaures. CNN architecture is inspired by human visual processing methods, making them

extremely successful in extracting complicated patterns and characteristics from images. The system can detect small color differences in objects with unprecedented accuracy and efficiency by exploiting this neural network design. The CNN's feature detectors experts in recognizing colors and their variations as training advances, resulting in a robust internal representation of color information. This format allows the algorithm to differentiate items based on their distinct color profiles, laying the groundwork for correct object sorting. As deep learning and computer vision research advances, the color-based item sorting system will evolve further, opening new possibilities and redefining the automation landscape of industries. It holds the possibility of altering enterprises, supporting sustainability, and boosting human capacities through responsible development in an increasingly networked and automated world.

Finally, the color-based object sorting system that employs CNN deep learning algorithms shows the convergence of innovation and intelligence, providing exceptional accuracy, efficiency, and versatility in object sorting tasks. This extraordinary technology demonstrates the power of artificial intelligence in revolutionizing industries, propelling innovation, and contributing to a future in which automation and human inventiveness coexist happily for the benefit of society.

OBJECTIVES

The main objective of the proposed methodology is to design a prototype that sorts the objects based on the feature extracted from the objects using deep learning convolution neural networks and classification of the objects with respect to colour by utilizing the machine learning supervised algorithms such as random forest and KNN.

II. RELATED WORK

Ren et al. [1] offer a novel approach for locating and tracking border contours in a binary picture, which is frequently applied in digital image processing. Sun et al. provide a unique method for locating and recognizing objects belonging to a certain visual category in most crowded spaces. The visual mode of classification requires more time to process.

It is well known that a generic location strategy put out by Seo et al. [2] may be used to come across a visual item of interest without the need for training. The suggested method makes use of an item of interest to find related things without any prior knowledge of the objects being sought for (learning) or the need for preprocessing. In this classical method of learning only have dealt. Comparison with various algorithms to evaluate the performance of the model haven’t done.

Deep learning applications for the automobile industry were proposed by A. Luckow et al. [3]. Deep learning algorithm especially for automotive industry had been discussed. Machine learning was suggested by Wuest et al. [4] as a solution to the industrial challenges. This article focusses solution only for the industrial challenges.

Three deep convolutional neural network architectures have been evaluated and analysed regarding computer-aided identification issues, dataset characteristics (such as selecting either a big dataset or finer item detection models), and CNN transfer learning from non-medical to medical picture fields [5]. CifarNet, AlexNet, and GooLeNet are the three primary CNN architectures utilized. This paper limits the suggestion for appropriate algorithm to be implemented.

Garad [6] evaluated changes in the precise geometric measurements for item sorting based on colour, size, and form recognition using MATLAB. This paper provides with the information for features extraction. Any surface flaws identified were not, however, remedied.

On a faster R-CNN model that was trained on pictures of mango and pitaya, researchers were able to use the model to recognize and sort several fruits with 99% accuracy while still maintaining fruit quality [7-9].

III. EXISTING METHODOLOGY

A color-based item sorting system is conceptually simple, but it has several drawbacks that might reduce its usefulness, especially when compared to the usage of deep learning methods like Convolutional Neural Networks (CNNs). Color-based sorting cannot distinguish between objects that have similar colours but distinct forms or textures since it solely takes colour information into account. Additionally, traditional sorting techniques are quite sensitive to variations in lighting, which can lead to misclassification as a result of shifting perceptions of colour.

The difficulty of color-based systems to handle objects with complex patterns or different colours is another notable flaw of these systems. This results in incorrect sorting conclusions, unlike CNNs, which can recognize intricate visual clues for exact categorization.

Though the principle behind the TA color-based item sorting system is straightforward, it has several flaws that might limit its utility, especially when contrasted to the use of deep learning techniques like Convolutional Neural Networks (CNNs). Since color-based sorting only considers colour information, it is unable to discriminate between things that have similar hues but different shapes or textures. Traditional sorting methods are also quite susceptible to changes in lighting, which can result in misclassification due to altered perceptions of colour.[10-14]

Another obvious drawback of color-based systems is their inability to manage objects with intricate patterns or many hues. Contrary to CNNs, which can detect complex visual cues for precise categorization, this leads to inaccurate sorting judgements.

IV. PROPOSED METHODOLOGY

A robotic arm that can sort objects by color is able to pick out objects of a particular color and position them where they are needed. Initially the object will be captured by the deep learning network which does the feature extraction of the image for an object. The feature extraction is processed by utilizing the concept of Convolution Neural Network, a deep learning algorithm. First it undergoes preprocessing, convolution of an image, Rectified Linear Unit and finally does the classification based on the fully connected network. This layer uses Multilayer Feed forward network for classification based on its color. To detect the colour of the object, the OpenCv computer vision is employed. The objects are acquired by the camera and they are converted into the respective RGB index. The process of Deep learning algorithm along with the machine learning algorithm that used for classifying the object will be run in GPU processor, which provides us the information about the object based on its color. The controller, which controls the motion of DC and servo motors. The proposed system has a limitation that the object that can be handled by the robot must have specific dimensions which is provided in the table. The camera that captures the image and provides the images as input to the ConvNets. The ConvNet's preprocess and extract the features. The model has been trained with three colors namely Red, Blue, Green.

Table 1: Constraints of Objects handled by Robot

Parameter	**Ranges**
Weight	<100 grams
Size	40 x 40 x 40
Shape	Cubic

Fig 1 Prototype of the Proposed Methodology

The Fig 1 shows the working model of the pick and place robot setup with the all-electronic circuit arranged in a manner to compute the automation task.

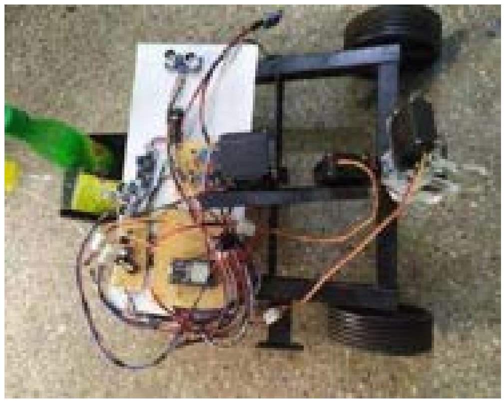


Fig 2 Prototype of the model for sorting object

The Fig 2 shows the model that automatically picks the objects that falls under the range of distance covered. Based on the object input features that has been processed under the deep learning algorithm like CNN, the classification has been done.

Fig 3. Block Diagram of the proposed System

The block diagram for the proposed methodology that developed is given in the Fig 3. The entire process uses the DC and Servo motors, the camera or colour sensor interfaces with the controller and process runs with the aid of the computer which has high processing speed. The computer runs the deep learning that built with the Convolution Neural Network for classification of an object.

The process involves the three major parts, which are listed below

A. CONTROLLER AND DATA ACQUISITION

ESP 32 controls the entire mechanism of process, from controlling the relays to the servo grippers. The object sorting involves the sensing of the object through ultrasonic sensor

B. SORTING AND CLASSIFICATION MECHANISM

The object that placed in the basket is picked by the Robot gripper and moves certain along based on the colour. Based on the colour picked by the robot gripper, it follows different angle patterns and place it in the appropriate bins. For employing this operation, the ultrasonic and proximity sensors are employed for moving and sorting of the objects.

C. SOFTWARE REQUIREMENT

i. *Image Capture:* Object is captured initially by the camera in the form of RGB format. OpenCv Lib (Computer Vision Library) is employed to convert the RGB into binary. The pseudocode performs this function by setting up the threshold value for the images.

ii. *Python IDE:* To make code to run on the processor to perform the classification and feature extraction process and other computer vision process.

D. DEEP LEARNING TECHNIQUE

The most recent technology CNN, a kind of deep neural networks known for their excellent image recognition capabilities, are used in the color-based item sorting method. CNNs are incredibly adept at extracting intricate patterns and attributes from pictures since their design is modelled after human visual processing techniques. Covnets, also known as convolution neural networks, are neural networks with shared parameters. They have,

(i) Convolutional Layer:

The central component of a CNN is the convolutional layer, which is also where the majority of processing takes place. It needs input data, a filter, and a feature map, among other things. Convolution layers are made up of a group of learnable filters (or kernels) with minor widths and heights and the same depth as the input volume (or three if the input layer is an image input). It has three layers namely, convolutional, pooling, and a fully connected layer. The convolution layer is the building block of CNN carrying the main responsibility for computation.

(ii) Padding:

The padding procedure is an important feature of CNNs because it improves the network's performance. Padding refers to the addition of extra pixels around the borders of an image before it is sent through the network in the context of CNNs. This strategic pixel increase allows for greater spatial information preservation during the various levels of convolution and pooling operations, resulting in improved item detection and sorting accuracy. The padding procedure in a Color-Based Object Sorting System ensures that the CNN can successfully identify and classify objects based on their color and other visual properties. Padding improves the network's ability to reliably recognize and sort objects by conserving spatial information and preventing information loss at the edges of objects, contributing to the overall efficiency and effectiveness of the sorting system.

(iii) Pooling layers:

Its major goal is to shrink the input feature map's spatial dimensions while preserving crucial data. It lowers the computational complexity of the network and increases its invariance to tiny changes in the input data as well as translation and rotation. The size of the pooling window, which is typically expressed as a 2x2 or 3x3 window, and the stride, which refers to how much the window is shifted throughout the pooling process, are two hyper parameters that affect how pooling layers behave.

Overall, pooling layers are essential in lowering the spatial dimensions of the feature maps, allowing CNN to concentrate on the most significant features while preserving the network's capacity to detect complicated patterns.

(iv) Fully Connected Layer:

To connect the neurons between two layers, the Fully Connected (FC) layer, which also includes weights and biases, is utilized. These layers make up the final few layers of a CNN architecture and are often positioned before the output layer.

Utilizing the features that the earlier layers and filters gathered, the FC layer performs classification jobs. The FC layer often employs a softmax function—which classifies inputs more accurately and generates a probability score between 0 and 1—instead of ReLu functions.

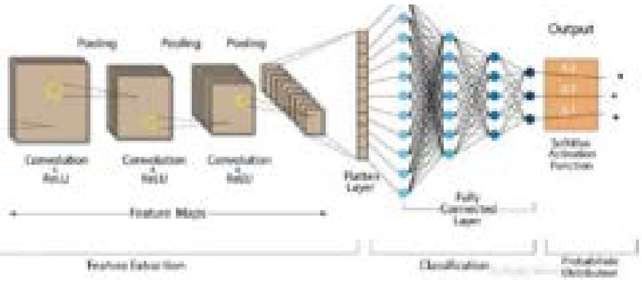


Fig 4 Layers involved in Convolution Neural Network for Object Classification [Source: https://developersbreach.com]

E. MACHINE LEARNING TECHNIQUE

Machine learning technique is implemented for training the model with appropriate training sets. When the model is fed with new data based on the training by machine learning technique and feature extraction by CNN deep learning technique, it sorts the objects effectively. The training data set with respect to input features labelled to output features were trained for 250 iterations. The model trains and classify with the greater accuracy. 100 labelled data set were taken as data set. In those 70 data had been used for training the model and remaining 30 data set had been used for testing the model for the validating the performance.

Since the labelled data set are used, the supervised type of learning was adopted. The input features such as colour with its RGB values are labelled with the color classification of objects such as Red, Green and Blue. The detection of color is achieved by acquiring images of the objects and based on RGB pixel acquisition it has been classified.

The overall block diagram of deep learning and machine learning technique is presented in the Fig 5.

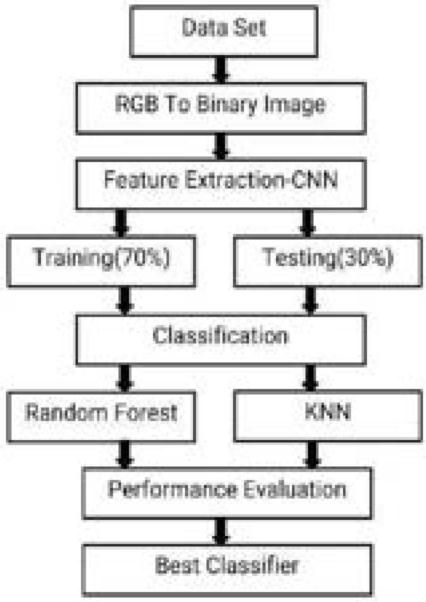


Fig 5 Flowchart procedure of Machine Learning and Deep Learning for classification

V. RESULTS AND DISCUSSIONS

Deep learning algorithm was applied for analyzing the object's features. The features taken in consideration for classification or sorting of objects are shape and colour of the object.

A. PERFORMANCE ANALYSIS OF SYSTEM

Many performance analysis matrixes are available in the Machine Learning algorithm for assessing the performance of the model. In this project, the performance has been analyzed using the Confusion Matrix Performance analysis, accuracy, recall and F1 score methods for evaluating the performance in detecting and sorting the object based on the colour.

The confusion Matrix will be assessed based on the True Positive Rates (TPR) and False Negative Rates (FNR). The confusion matrix parameters that derived from the Convolution Neural Network's output are presented in the Table I.

Table I: Performance Metrics of Confusion Matrix

Parameters	Values of RF	Values of KNN
Total Images	100	100
Training Images	70	70
Testing Images	30	30
True Positive	94	87
True Negative	6	13
False Positive	5	20
False Negative	95	80

From the above table, the confusion matrix gives the True Positive, True Negative, False Positive and False Negative. From the performance analysis of the confusion matrix, true positive in the confusion matrix parameter gives the correct classification and sorting of the object based on the colour. The false positive parameter gives the misclassification or wrong identification of the object based on the colour. This kind of performance in the network affects the accuracy of the network in appropriate detection and classification of objects. Recall is the percentage of correctly classified items in a particular class.

$$recall = \frac{TP}{(TP+FP)} \quad (1)$$

F1 score measured by the precision and recall to analyze the performance of the object classification. Table 3 highlights the performance evaluation of the algorithm. After the classification of objects based on the colour the precision of the Random Forest is 92% and recall of 90% and F1 measure of 92% compared to the KNN algorithm with 81%, 85% and 70% respectively.

Table II Comparison of Proposed Algorithms in Sorting System

Classifier Algorithm	Precision %	Recall %	F1 Measure %
Random Forest	92	90	92
KNN	81	85	70

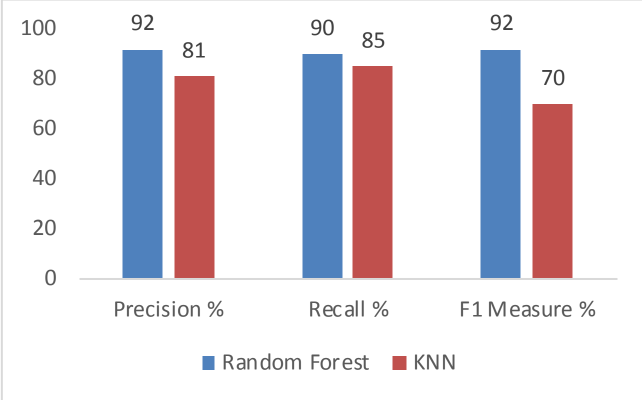


Fig 6 Comparison Results of Classification with Random Forest and KNN

VI. CONCLUSION AND FUTURE WORK

A system based on Convolution type Neural Networks for feature extraction was created to detect the object based on the colour and shape of the object that present in any industrial to domestic applications. The object that differs based on the shape, colour and size are considered as the input features and it is fed to the neural networks for training the Convnet to make the system to learn. The model learns for the various set of training datasets that differ from each set based on its feature vector. For various trails, the proposed system detects and sorts the objects accurately. By utilizing the Convolution Neural Network algorithm employed with Random Forest technique, the classification has been done very accurately and efficiency is very high with very negligible error. The proposed model is limited to handle only three colored objects, for handling more colors it requires concentration in feature extraction. By considering more input features or employing better classifiers the classification efficiency can be improved that Random Forest technique.

Since industries are running towards industrial revolution 4.0, the future industries and even domestic application will adapt deep learning technologies to enable the system to act intelligent. Industries future will be deep learning. This project not only limited for sorting objects in textile industries. It can be extended to the domestic applications to the large-scale industrial applications where sorting can be carried out.

REFERENCES

[1] Mingwu Ren, Jingyu Yang, Han Sun, “Tracing boundary contours in a binary image,” Image and Vision Computing”, vol.20, pp.125-131,2012.

[2] H. J. Seo, S. Member, and P. Milanfar, “Training-Free, Generic Object Detection Using Locally Adaptive Regression Kernels,” *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 32, no. 9, pp.1688 - 1704, 2010.

[3] A. Luckow, M. Cook, N. Ashcraft, E. Weill, E. Djerekarov and B. Vorster, "Deep learning in the automotive industry: Applications and tools,", IEEE International Conference on Big Data (Big Data),

Washington, DC, 2016, pp. 3759-3768, 2016.

[4] Thorsten Wuest, Daniel Weimer, Christopher Irgens & Klaus-Dieter Thoben, "Machine learning in manufacturing: advantages, challenges, and applications, Production & Manufacturing Research", 4:1, 23-45, DOI: 10.1080/21693277.2016.1192517, 2017.

[5] H. Shin et al., "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning", IEEE Transactions On Medical Imaging, vol. 35, no. 5, pp.1285-1298, doi: 10.1109/TMI.2016.2528162, May 2016.

[6] Priya Vinayak Garad, "Object Sorting Robot Based on the Shape", International Journal Of Advance Research, Ideas And Innovations In Technology, Vol.3, Issue 5, pp. 129-134.

[7] H. Basri, I. Syarif and S. Sukaridhoto, "Faster R-CNN Implementation Method for Multi-Fruit Detection Using Tensor Flow Platform", International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC), Bali, Indonesia, pp. 337 – 340, 2018, doi: 10.1109/KCIC.2018.8628566, 2018.

[8] Fan, J., Li, G., Zheng, X., & Yu, Z. "Real-time color sorting of fruits using deep learning," Journal of Food Engineering, 229, pp. 1-7, 2018.

[9] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137 - 1149, 1 June 2017, doi: 10.1109/TPAMI.2016.2577031.

[10] Zhang, H., Xu, Y., & Liu, X, "A color sorting system based on deep learning for plastic recycling," IEEE International Conference on Robotics and Automation, pp. 2728-2733, 2017.

[11] Pratik Roy et al., "Color Based Object Sorting System using Deep Learning", International Journal of Engineering and Advanced Technology, Vol.9, Issue 5, pp no 964 – 968, 2018.

[12] Fadhil A T et al., "Computer Vision – Based System for Classification and Sorting Color Objects", IOP Conference Series: Materials Science and Engineering, DOI 10.1088/1757-899X/745/1/012030, 2019.

[13] Huang J Rathod V,Sun C e.al., "Speed/accuracy trade-offs for modern convolutional object detectors., CVPR 2017.

[14] . L. BinYan, W. YanBo, C. ZhiHong, L. JiaYu and L. JunQin, "Object detection and robotic sorting system in complex industrial environment", Chinese Automation Congress (CAC), Jinan, 2017, pp. 7277-7281, doi: 10.1109/CAC.2017.8244092.

A Survey paper on Understanding the Rise of AI-driven CyberCrime and Strategies for Proactive Digital Defenders.

Gogu Venkata Sai Meghana
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, A.P., India-522502
gvsaimeghana@gmail.com

Shaik Saqlain Afroz
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, A.P., India-522502
shaiksaqlainafroz2003@gmail.com

Rajesh Gurindapalli
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, A.P., India-522502
rajeshgurindapalli15@gmail.com

Subhash Katari
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, A.P., India-522502
subhashkatari1234@gmail.com

Kolachana Swetha
Department of Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Vaddeswaram, A.P., India-522502
swetha.k@kluniversity.in

***Abstract*: As artificial intelligence (AI) continues to evolve, so too does its integration into cyber criminal activities, presenting a formidable challenge to digital security. These findings investigate the escalating nexus between AI and cybercrimes, highlighting the emergent dangers posed by AI-driven malicious activities. The study delves into the various ways in which AI technologies are leveraged by cyber criminals to orchestrate sophisticated attacks, incorporating, data breaches, malware propagation, phishing, and social engineering tactics. Furthermore, the methodologies propose proactive research strategies aimed at mitigating the threats posed by AI-facilitated cybercrimes through the lens of digital forensic techniques. By analyzing current trends in AI-driven cyber offenses and their repercussions on digital security frameworks, this research endeavors to elucidate the imperative for novel approaches in digital forensics. Such proactive strategies encompass the establishment of AI-powered forensic tools, the enhancement of detection and attribution methodologies, and the augmentation of cyber resilience through predictive analytics and preemptive measures. Through a comprehensive review of existing literature, case studies, and empirical data, paper analysis seeks to offer insights into the changing landscape of AI-facilitated cybercrimes and the critical importance of digital forensics in countering these threats. By Gaining a more comprehensive grasp of the combined effects or interactions between AI technologies and cyber criminality, this research endeavors to inform stakeholders in the realms of cybersecurity, law enforcement, and policy-making, thereby contributing to the progress or development in proactive measures aimed at safeguarding digital ecosystems against emerging threats.**

***Keywords**: Artificial Intelligence, Cyber Crime, Digital Security, Digital Forensic, Cyber Resilience, Law enforcement, AI-Driven vehicles, Deep-Fake technology, Human Safety.*

I. INTRODUCTION

The study on numerous published research papers, articles, and search engine results highlights a significant shift in cybercrime from traditional physical methods to automated approaches. The integration of AI into illicit activities poses a considerable challenge to digital security, potentially leading to unprecedented threats. As AI technologies evolve, they provide cybercriminals with new capabilities and tools, enabling them to execute more sophisticated and covert attacks. This dynamic realm of AI-enabled cybercrimes encompasses various malicious activities, such as data breaches, malware propagation, phishing schemes, and social engineering strategies. The rapid proliferation of AI-driven cybercriminal activities highlights the crucial need for proactive measures to mitigate these threats. Traditional approaches to digital forensics and cybersecurity must adapt to the complexities introduced by AI-driven offenses.

By analyzing current trends and emerging patterns in AI-facilitated cybercrimes, researchers can empower authorities to identify vulnerabilities and develop innovative solutions to safeguard digital ecosystems. Additionally, the convergence of AI and cybersecurity presents both challenges and opportunities for policymakers, law enforcement agencies, and technology developers. Effective responses to AI-enabled cybercrimes necessitate interdisciplinary collaboration, technological innovation, and robust regulatory frameworks. Encouraging a deeper understanding of the complexities of AI cybercriminality enables stakeholders to formulate proactive measures for risk reduction and enhance digital resilience. In cybersecurity, AI plays a pivotal role across various applications including threat detection, response, and prevention, enabling organizations to proactively predict and combat evolving cyber threats. However, the advancement of AI-powered criminals highlights the importance of a robust understanding of predictive AI research to effectively identify and mitigate emerging threats. By analyzing the evolving strategies, techniques, and procedures employed by cybercriminals, this study aims to analyze and present the complex integration of AI technologies and contemporary cyber threats. Despite extensive research in the cybersecurity domain, there remains a gap in investigating how AI perpetuates a feedback loop, generating new challenges and necessitating innovative mitigation strategies for solving AI-powered cyberattacks [11]. Through an extensive review of over a hundred articles collected via Google search, each focusing on AI-enabled attacks and real-world threats, this study has observed a pattern of large-scale damage incurred by organizations, indicative of organizational crime. Cybersecurity emerges as a critical aspect of security in large and well-organized industries, making them prime targets for cybercriminals aiming to exploit vulnerabilities. Additionally, AI-powered cyber attackers exploit deep-fake technology to create misleading videos, manipulating innocent individuals for malicious gains [12]. These attackers leverage AI-based bots on popular social media platforms like Telegram and Snapchat to facilitate data leaks and perpetrate face recognition scams, thereby compromising user privacy and security [21]. The advent of AI-driven vehicles brings forth new risks, where maliciously programmed automated vehicles could be weaponized to cause illegal accidents. Despite challenges in identifying culprits, enhanced digital forensic techniques offer opportunities for identifying perpetrators and holding them accountable. Upon conducting an enhanced survey of various research papers and articles, it becomes evident that while AI holds significant potential in cybersecurity, its misuse may result in dire consequences. The misuse of AI techniques in cybercrime, as highlighted in references [3] and [6], can damage any individuals' reputation.

To overcome these challenges, forensic analysts must adopt a forward-thinking strategy and understand the methodologies employed by cybercriminals. By thinking like criminals, analysts can adeptly act through the complexities of AI-driven cybercrimes and protect innocent individuals from unwarranted accusations and damaging tactics. Framing the entire problem into a single question, it can be stated: "What if cybercrime enabled by AI is encountered by an educated criminal in the fields of AI and digital forensics?" This question encapsulates the potential scenario where individuals with knowledge in both AI and digital forensics may exploit AI-driven cybercrime for malicious purposes, highlighting the importance of proactive measures and interdisciplinary collaboration in combating emerging threats.

II. RESEARCH METHODOLOGY

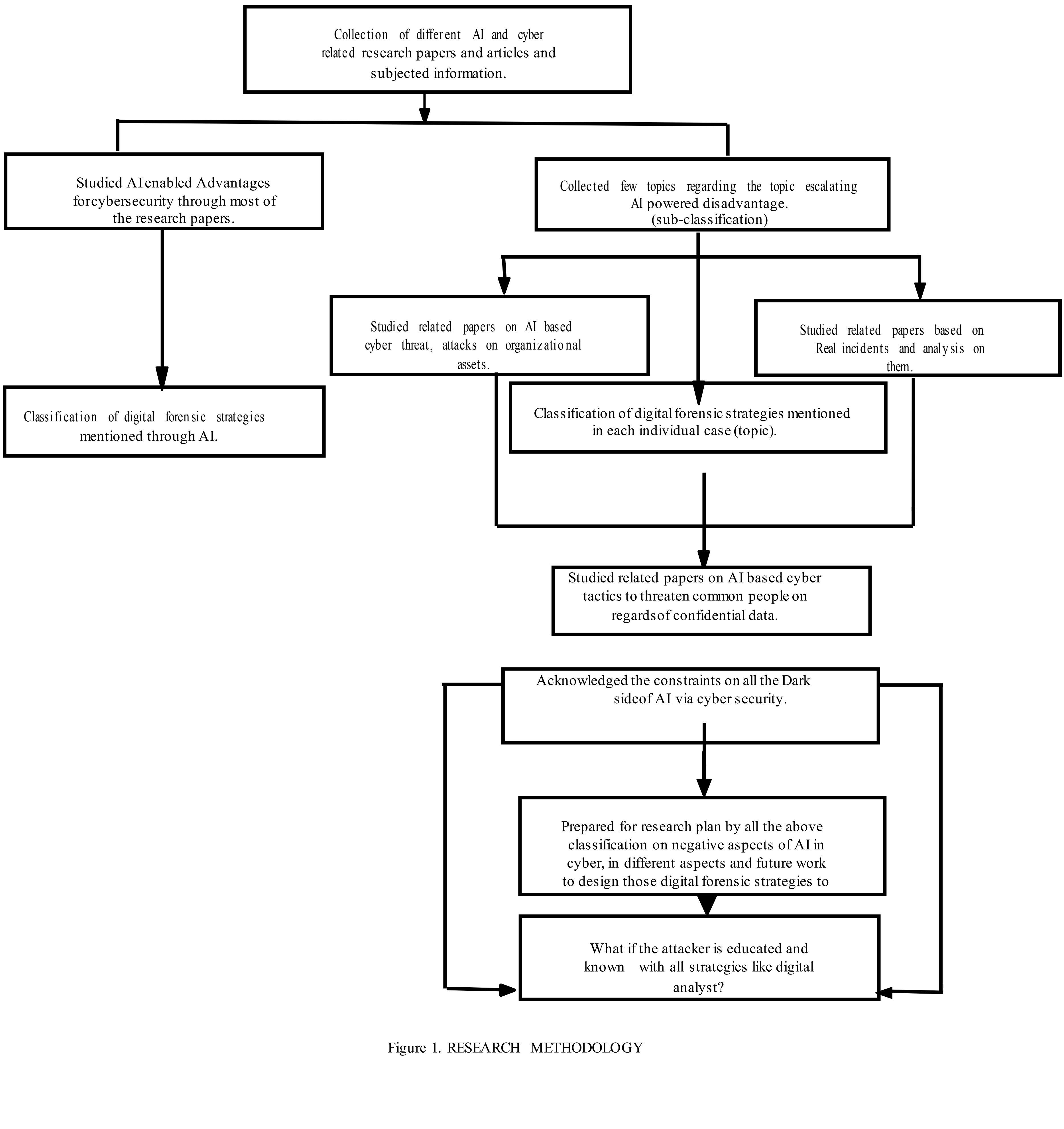


Figure 1. RESEARCH METHODOLOGY

Figure 1 shows the proposed research methodology to perform a comprehensive research study

Step – 1: Collection of Different AI and Cyber Related Research Papers and Articles and Subjected Information:
This section compiles various research papers and articles related to artificial intelligence (AI) and cybersecurity, providing a comprehensive overview of the existing literature and subject matter
Step – 2: Studied AI-Enabled Advantages for Cybersecurity Through Most of the Research Papers:
Examines the advantages offered by AI in enhancing cybersecurity measures as identified in the reviewed literature. Subsequently, it delineates specific topics highlighting the escalating disadvantages associated with AI-powered cybersecurity.
Step – 3: Studied Related Papers on AI-Based Cyber Threats, Attacks on Organizational Assets:
Analyzes research papers focusing on AI-based cyber threats and attacks, particularly those targeting organizational assets. Real incidents are scrutinized, and thorough analysis is conducted to understand the evolving landscape of threats
Step – 4: Classification of Digital Forensic Strategies Mentioned Through AI:
Categorizes digital forensic strategies outlined in the literature that leverage AI technologies to investigate cyber incidents, aiding in the identification and resolution of security breaches.
Step – 5: Studied Related Papers on AI-Based Cyber Tactics to Threaten Common People Regarding Confidential Data:
Explores research papers detailing AI-driven cyber tactics aimed at threatening common individuals and compromising with confidential data. It evaluates the implications of such tactics on privacy and security.
Step – 6: Classification of digital Forensic Strategies Mentioned in Each Individual Case:
Provides a systematic classification of digital forensic strategies discussed in individual case studies, examining their application and effectiveness while addressing cyber threats.
Step – 7: Acknowledged the Constraints on All the Dark Side of AI via Cyber Security:
Acknowledges the limitations and constraints associated with the dark side of AI in cybersecurity, highlighting ethical, legal, and technical challenges that need to be addressed.
Step – 8: What if the Attacker is Educated and Known with All Strategies Like Digital Analyst:
Explores the scenario wherein attackers possess comprehensive knowledge of digital forensic strategies, akin to a digital analyst, and discusses potential implications and countermeasures.

III. ANALYSIS OF EXISTING METHODOLOGIES

The different types of online threats faced by children, including cyberbullying, online harassment, grooming, and exploitation, has been extensively highlighted in the recently published research articles and reports. This study has examined the scholarly works within Pakistan and examined the prevalence of cybercrimes against children and highlighted the need for preventative measures to safeguard their digital well-being [1]. Studies have explored the potential utilization of Artificial Intelligence (AI) innovations, including Natural Language Processing (NLP), machine learning algorithms, and image recognition, in detecting and preventing online threats targeting children. Furthermore, initiatives by governmental bodies, non-profit organizations, and advocacy groups in Pakistan have emphasized the significance of raising awareness, promoting digital literacy, and fostering collaboration among stakeholders to address cybercrimes against children effectively.

Artificial intelligence (AI) is increasingly intertwined with cybersecurity, as highlighted in various scholarly articles [2]. Researchers emphasize AI's central role in automating tasks, accelerating risk detection, and enhancing response mechanisms within cybersecurity frameworks. The recent research reviews explore AI's potential to address cybersecurity challenges by securing communication channels and mitigating cyber threats. Additionally, the idea of explainable AI (XAI) is investigated to clarify cybersecurity using AI applications and enhance security analysis. However, there is growing concern regarding the emergence of AI-driven cyberattacks, prompting studies to analyze cybercriminals' techniques and the changing risk landscape. As cybercrime escalates, researchers highlight the necessity for comprehensive studies on AI-based security frameworks, particularly within emerging digital environments like the metaverse. Moreover, systematic literature reviews investigate AI strategies in crime prediction, underscoring the significance of understanding AI's implications across various domains.

The evolving trend of cybercriminals integrating AI into their malicious operations outlines the diverse ways in which AI technologies empower adversaries to conduct efficient and stealthy attacks, posing significant challenges to traditional cybersecurity paradigms [3]. Furthermore, surveys delve into the proactive measures adopted by cybersecurity professionals to utilize AI for threat detection, incident response, and vulnerability management, aiming to counteract the adversarial exploitation of AI to safeguard organizations.

The abstract of the paper "How AI Is Shaping the Future of Cybercrime" highlights the increasing utilization of AI by cybercriminals to perpetrate advanced attacks and underscores the urgent need for defenders to employ AI-driven defensive mechanisms to safeguard against evolving cyber threats and maintain the integrity of digital ecosystems. Cybercriminals are using AI tools to orchestrate sophisticated attacks, exploiting its capabilities to enhance malware development, automate assaults, and craft effective social engineering plans. However, cybersecurity professionals are responding with proactive measures to mitigate these threats by harnessing AI for defensive strategies.

"Artificial Intelligence-based Cybersecurity within the Framework of the New Industrial Revolution" presents a comprehensive analysis of existing research, systematically examining literature pertaining to AI-driven cyber-attacks and their implications for Industry 4.0 ecosystems [4]. The study explores the use of artificial intelligence in cybersecurity, encompassing threat detection, anomaly detection, and tailored incident response mechanisms for Industry 4.0. Additionally, it delves into emerging trends in AI-driven cyber threats, such as adversarial machine learning and autonomous attack vectors, highlighting the importance of adaptive and resilient cybersecurity frameworks. Furthermore, the survey underscores the integration of AI technologies with traditional cybersecurity measures to bolster threat intelligence, mitigate risks, and safeguard critical infrastructures within the new industrial revolution framework. Overall, the literature review furnishes valuable enlightenment into the evolving dynamics of artificial intelligence and cybersecurity within the context of Industry 4.0, offering a basis for informed decision-making and guiding future research endeavors. Recent advancements in applying Artificial Intelligence (AI) and Machine Learning (ML) to enhance cybersecurity measures have been a focal point of research. This study analyzes various facets, including the utilization of AI/ML algorithms for detecting potential risks, anomaly identification, and behavior analysis [5]. It explores emerging trends such as Explainable AI (XAI) for interpreting AI-driven decisions in cybersecurity contexts. Furthermore, the literature survey explains the convergence of AI and ML techniques with regular security systems to strengthen the defense mechanisms against evolving cyber threats. It also examines key research areas like adversarial machine learning and the application of AI/ML in network intrusion detection systems. These challenges and prospects associated with AI/ML adoption in cybersecurity domains offer perspectives on prospective research paths and strategic implementations for robust cyber defense strategies.

The implications of AI policy-making on the administration of criminal justice systems, particularly in combating cybercrimes, have been thoroughly discussed [6]. It highlights the collaborative efforts of international organizations in developing applicable instruments to address cybercrimes effectively. Moreover, it explores the intersection of artificial intelligence and cybercrime, showcasing the difficulties and opportunities arising from the integration of AI technologies in criminal investigations and law enforcement. The changing terrain of cyber threats and the significance of AI in enhancing forensic techniques and crime prevention strategies are also explored. COVID-19 Pandemic accelerated the dependence on AI-based technologies, but it also encouraged organized criminal groups to exploit vulnerabilities, increasing cyber threats against various sectors and individuals [6]. Cybercriminals deploy AI to orchestrate sophisticated attacks, exploiting IoT vulnerabilities and perpetuating misinformation campaigns through bots and deepfakes. This requires enhanced international cooperation, legal frameworks, and technological capabilities to combat evolving cyber threats facilitated by AI-enabled criminal activities effectively, safeguarding the society against emerging threats. Another study has conducted an extensive survey exploring the connection between cybersecurity and Artificial Intelligence (AI), outlining the rapidly changing landscape of cyber threats [7]. It highlights the adaptability of cyberattackers, who are increasingly utilizing AI-driven techniques in their attacks. The survey investigates into the multifaceted implications of AI across different security domains, encompassing physical, digital, and political security contexts. Furthermore, it illuminates the potential risks linked with the malicious application of AI technology. Through a methodical analysis, the paper pinpoints research gaps and challenges, with a particular focus on addressing accident dangers linked to AI and other critical cybersecurity issues. By furnishing insights into both the opportunities and perils presented by AI in cybersecurity, the survey stands as an informative resource for understanding how cyber threats are changing and how AI may reduce or increase these problems. Additionally, the study scrutinizes potential risks emanating from the malevolent exploitation of AI across various security domains, including physical, digital, and political security. It also emphasizes research domains related to potential accident risks linked with AI and other relevant cybersecurity matters. AI technologies within cybersecurity, spotlighting the necessity for strong governance frameworks to guarantee responsible AI implementation and address the risks of unintended outcomes. It motivates interdisciplinary collaboration between policymakers, technologists, and ethicists to develop holistic approaches to AI governance that balance innovation with accountability.

TABLE 1: SUMMARY OF EXISTING METHODOLOGIES

S. No	Paper Title	Year of Publication	Objective of AI based cyberattacks via Digital Forensics	Conclusion	Limitations
1	Artificial Intelligence Implementation to Counteract Cybercrimes Against Children in Pakistan	2022 (10th October)	Detect, takedown, and trace online violence by AI, against children	. Updating cybersecurity measures and leveraging AI alongside traditional methods areessential to safeguard children from evolving online risks. As technology progresses, the call for anticipatory actions becomes paramount to protect children from the growing threats of sexual abuse, bullying, and exploitation online.	Lack of primary data limits the scopeof the study .and the Lack of data sufficiently witnessing cyberattacks
2	AI for Cybersecurity and Cybercrime: How Artificial Intelligence Is Battling Itself	2023 (9th June)	Detect Deepfake-AI-powered forensic tools and establish legal frameworks for AI crime investigation.	The paper emphasizes the needfor standardized methodologies and ethical considerations in combating AI-driven cybercrime	Lack of standardized methodologies for AI crime detection, ethical considerations inAI-driven forensic analysis
3	AI Is Shaping the Future of Cybercrime	2023 (21st December)	Not mentioned via digital services criteria but objective Isto discussed about different realtime mitigations of AI power- cyberattacks.	AI tools are opening society's eyes to new possibilities in virtually every field of work. Ashackers take fuller advantage oflarge language model technologies, the industry will need to keep pace to keep the AIthreat under control.	The paper is only limited to real world examples and malicious cases of AI application and mitigations due to AI powered cyberattacks.
4	Artificial Intelligence-Based Cyber Security within the Situation of Industry 4.0—A Survey	2023 (19th April)	With the increasing necessity of networked information technology, the surface area for cyberattacks has increased. Because of this, studies aiming at understanding the actions of cybercriminals are essential for producing data for cybersecurity policies.	Mentioned variations in cyberattacks, defense counterstrategies, implementation of AI, ML and DL for cyber security in Industry 4.0, advantages, and disadvantages through AI to preserve security.	The study does not complete into the practical implementation and effectiveness of the derived cyber security challenges in real-world scenarios.
5	Current trends in AI and ML for cybersecurity: A state-of-the-art survey	2023 (25th October)	Not mentioned about digital forensicsOr Services but mentioned about future work and research based on digital forensic methodologies.	Out of 100 percent 20 to 30 percent organizations are still facing challenges to application of AI with cyber-security.	The discussion in involved only with different real scenarios but no implementation or methodology which were left for future work or research.
6	Cybercrime and Artificial Intelligence. A conclusion of the work of international organizations on criminaljustice and the international applicableinstruments	2022 (22nd February)	Discussed about public and national security objectives and challenges when drones auto vehicles are not properly managed.disruption of AI from the times of Covid 19 and ec3 Europol.	The paper finalizes with a conclusion that offers an alternative to create effective policy responses to counter cybercrime committed through AI systems.	The entire research is an overviewof organizational conventions butno practical plan and even no future work.
7	The AI-Based Cyber Threat Landscape: A Survey	2020 (February)	To raise awareness about the progressive landscape of cybersecurity threats driven by AI technologies and to facilitate the development of strategies and defenses to mitigate these risks effectively.	Cyberattacks merging AI with conventional methods, introducing an analytic framework for modelingattacks and strategizing defenses. It highlights exposure and bugs in infrastructures like smart grids.	Findings and frameworks presented may not universally apply to all cyberattack situations, as the effectiveness ofdefenses can vary based on specific contexts.
8	Artificial Intelligence Safety and Cybersecurity: aTimeline of AI Failures	2016 (25th October)	To compile and analyze a timelineof AI failures in the context of safety and cybersecurity, document instances where AI systems have malfunctioned or caused harm dueto safety and security vulnerabilities	Importance of addressing safety and cybersecurity in AI development and deployment. Analysis on timeline of AI failures, it underscores the need forrobust safety mechanisms, rigorous testing, and proactive risk mitigationstrategies to prevent future incidents.	Researches findings may include the relianceon publicly available information and documented cases of AI failures, which may not represent the entire spectrum of incidents or accurately capture the extent of the problem.
9	The Emerging Threat of Ai-driven Cyber Attacks: A Review	2022 (4th March)	To examine the emerging threatposed landscape of AI-driven cyberattacks, analyze the techniques employed by threat actors, and assess the potential risks and impacts on cybersecurity (digital service)	The significant challenges posed by AI-driven cyberattacks and underscores the importance of proactive measures to mitigate these threats, emphasize need for continuous monitoring, adaptationof cybersecurity strategies (digital forensic methodology)	Focus on existing AI-driven cyberattack techniques without extensive coverage of emergingor novel threats.
10	Fighting malevolent AI: artificial intelligence, meet cybersecurity	2016 (13th June)	To explore the intersection of artificial intelligence and cybersecurity, focusing on mitigating malevolent uses of AIin cyber threats	Emphasizes the critical need for proactive measures to address the growing challenges posed by malevolent AI in cybersecurity. It underscores the main aspect of ongoing research, collaboration, and innovation to stay ahead of emerging threats and safeguard digital systems and infrastructure	The paper may not provide comprehensive solutions to all potential threats posed by malevolent AI, as the plot of AI–enabled cybersecurity is rapidly evolving

VI. CONCLUSION

This research study has reviewed numerous AI-driven cybersecurity research articles across various research platforms, revealing significant advantages in the implementation of AI for cybersecurity. While many research works focus on the benefits, measures have been identified, termed the "Dark Side of AI." The aim of this research study is not to determine if AI is inherently good or bad for cybersecurity but to highlight potential risks. Considering a future where attackers grasp AI-driven cyber technology as well as ethical analysts, substantial security threats. Such attacks might result in economic losses and compromise the CIA (Confidentiality, Integrity, and Availability) of organizations. Even after implementing best practices and measures, attackers may get successful, exploiting technological advancements, possibly aided by well-trained groups skilled in misleading investigators and analysts. Educated criminals in AI and cybersecurity might easily evade capture, framing innocent individuals. Classifying AI-driven cyber threats, it is evident that significant financial and human losses are foreseeable. To combat these issues, it is best to advocate for using machine learning (ML) algorithms alongside AI in cybersecurity to overcome malware attacks. ML algorithms enhance threat detection capabilities, offering proactive preventative measures against evolving malware and cyberattacks.

REFERENCES

[1]. Saadia Anwar Pasha, Sana Ali, Riadh Jeljieli (2022), Artificial Intelligence Implementation to Counteract Cybercrimes Against Children in Pakistan. https://www.darkreading.com/vulnerabilities-threats/how-ai-shaping-future-cybercrime

[2]. Gaurav Belani. (2023). AI for Cybersecurity and Cybercrime: How Artificial Intelligence is Battling Itself. https://www.computer.org/publications/tech-news/trends/ai-fighting-ai

[3] Richard De La Torre. (2023). How AI Is Shaping the Future of Cybercrime, https://link.springer.com/article/10.1007/s 12027-022-00702-z

[4] Nettaria Kaloudi, Jingyue Li (2020). AI-Based Cyber Threat Landscape: A Survey. https://doi.org/10.1145/3372823

[5] Roman V.Yampolskiy, M.S.Spellchecker. (2016), Artificial Intelligence Safety and Cybersecurity:A Timeline of AI Failures.https://doi.org/10.48550/arXiv.1610.07997

[6] Blessing Guembe, Ambrose Azeta, Sanjay Misra, Victor Chukwudi Osamor, Luis Fernandez-Sanz & Vera Pospelova. (2022). The Emerging Threat of Ai-driven Cyber Attacks: A Review. https:// doi.org/10.1080/08839514.2022.2037254

[7] Antonio João Gonçalves de Azambuja, Christian Plesker, Klaus Schützer, Reiner Anderl, Benjamin Schleich and Vilson Rosa Almeida. (2023) Artificial Intelligence-Based cybersecurity in the context of industry 4.0-A Survey, https://doi.org/10.3390/electronics12081920

[8] Nachaat Mohamed, (2023) Current trends in AI and ML for cybersecurity: A state-of-the-art survey, https://doi.org/10.1080/23311916.2023.2272358

[9]Cristos Velasco. (2022) Cybercrime and Artificial Intelligence. An overview of the work of international organizations on criminal justice and the international applicable instruments. https://doi.org/10.1007/s12027-022-00702-z

[10]Roman V. Yampolskiy. (2016). Fighting malevolent AI- artificial intelligence, meet cybersecurity. https://phys.org/news/2016-06-malevolent-aiartificial-intelligence-cybersecurity.html

[11] Jp A. Yaacoub, Hassan N Noura, Ola Salman, Ali Chehab, March (2021). Digital Forensics vs. Anti-Digital Forensics: Techniques, Limitations and Recommendations, https://www.researchgate.net/publication/350544009_Digital_Forensics_vs_AntiDigital_Forensics_Techniques_Limitations_and_Recommendations

[12] Gorka Sadowski. (2023). The Dark Side of AI in Cybersecurity: Unveiling the Hidden Dangers https://gorkasadowski.medium.com/the-dark-side-of-ai-in-cybersecurity-unveiling-the-hidden-dangers-d019079baac4

[13] Bibhu Dash, Meraj Farheen Ansari, Pawan kumar Sharma (2022) Threats andOpportunities with AI-based Cyber Security Intrusion Detection of Artificial https://aircconline.com/ijsea/V13N5/13522ijsea02.pdf

[14] Blessing Guembea, Ambrose Azetab, Sanjay Misrac, Victor Chukwudi Osamora, Luis Fernandez-Sanzd, and Vera Pospelova (2022) The EmergingThreat of Ai-driven Cyberattacks https://doi.org/10.1080/08839514.2022.2037254

[15] Ashir Mehta, Hiteshri Yagnik, Manan Shah (2021) The impact of Artificial intelligence techniques in argumentation of cybersecurity. https://doi.org/10.1007/s40747-021-00494-8

[16] Muhammad Mudassar, Yamin Mohib, Ullah Habib Ullah, Basel Katt (2021) Wea ponised AI for cyberattacks https://doi.org/10.1016/j.jisa.2020.102722 .

[17] Chih-Ling Chang, Jui-Lung Hung , Chin-Wei Tien ,Chia-Wei Tien, Sy-Yen Kuo (2020) Evaluating Robustness of AI Models against Adversarial https://doi.org/10.1145/3385003.3410920

[18] Abdul Basit1, Maham Zafar,Xuan Liu,Abdul Rehman Javed,Zunera Jalil, Kashif Kifayat(2020) A comprehensive survey of AI-enabled phishing attacks detection technique https://doi.org/10.1007/s11235-020-00733-2

[19] Bilal Alhayani, Husam Jasim Mohammed,Ibrahim ,Zeghaiton Chaloob, Jehan Saleh Ahmed(2021) Effectiveness of artificial intelligence techniques against cyber security https://doi.org/10.1016/j.matpr.2021.02.531

[20] Meraj Farheen Ansari(2022) A Quantitative Study of Risk Scores and the Effectiveness of AI-Based Cybersecurity Awareness traing https://10.47893/IJSSAN.2022.1212

[21] Shailendra Mishra. (2023).Exploring the impact of AI-Based cyber security Financial Sector Management. https://doi.org/ 10.3390/app13105875

[22].Maanak Gupta, CharanKumar Akiri, Kshitiz Aryal, Eli Parker, Lopamudra Praharaj (2023).From ChatGPT to Threat Gpt: Impact of Generative AI in Cyber Security and Privacy. https://doi.org/10.1109/ACCESS.2023.3300381

[23] Iqbal H.Sarkar, Md Hasan Furhad, Raza Nowrozy. (2021).AI-Driven Cybersecurity: An Overview, security intelligence modelling and research directions. https://doi.org/10.1007/s42979-021-00557-0

[24] Dhruv Grewal, Abhijit Guha, Cinthia B.Satornino, Elisa B.Schweiger .(2021) Artificial Intelligence: The light and the darkness. https:// doi.org/10.1016/j.jbusres.2021.07.043

[25] Stanislav Ivanov. (2023). The dark side of artificial intelligence in higher education. https://doi.org/10.1080/02642069.2023.2258799

[26] Abhilash Chakraborty, Anupam Biswas, and Ajoy Kumar (2022) ArtificialIntelligence for Cybersecurity: Threats, Attacks and Mitigation. https://doi.org/10.48550/arXiv.2209.13454

[27] Cristos Velasco (2022Cybercrime and Artificial Intelligence. An overview ofthe work of international organizations on criminal justice and the international applicable instruments https://doi.org/10.1007/s12027-022- 00702-z

[28]] Katanosh Morovat, B. Panda (2021) A Survey of Artificial Intelligence in Cybersecurity https://10.1109/CSCI51800.2020.00026

[29] Cong Truong Thanh and Ivan Zelinka (2019) A Survey on ArtificialIntelligence in Malware as Next-Generation Threats https://doi.org/10.13164/mendel.2019.2.027

[30] Mariarosaria Taddeo, Tom McCutcheon & Luciano Floridi(2019) Trusting artificial intelligence in cybersecurity is a double-edged swordhttps://doi.org/10.1038/s42256-019-0109-1

Predicting Crop Yield using Long Short-Term Memory, Integrated Gradients and Shapley Additive Explanations

Durai Arumugam S.S.L
Department of Information Technology
Easwari Engineering College
Chennai, India
duraiarumugam36@gmail.com

Shanjai Rahul S
Department of Information Technology
Easwari Engineering College
Chennai, India
flytoshanjai@gmail.com

Praveen Kumar R
Department of Electronics and Communication Engineering
Easwari Engineering College
Chennai, India
rpjcspraveen@gmail.com

Rubeshkumar A
Department of Information Technology
Easwari Engineering College
Chennai, India
rubeshanbu25@gmail.com

Abstract— **Agriculture is the backbone of India, which plays a major role in the economy of a country. Agriculture is based on planting and harvesting crops over time. Crops should be sowed according to the environmental factors and appropriate lands. It is difficult for farmers to predict the yield of crop before harvesting because of varying climatic conditions. So to overcome this issue, the amount of crop yield is predicted with the factors like soil-type, weather, climate, crop-type etc., This paper introduces a novel approach to improve the accuracy and interpretability of crop yield prediction models by integrating Long Short-Term Memory (LSTM) networks, Integrated gradients (IG) feature selection, and SHapley Additive exPlanations (SHAP) analysis. The proposed methodology aims to predict the amount of specified crops that can be yield in an area. The lack of interpretability in many existing models hampers stakeholders' ability to understand and trust the predictions. The LSTM network captures temporal dynamics for improved accuracy, while IG and SHAP provide insights into feature contributions, addressing the interpretability issues in existing models. This integrated system not only advances prediction accuracy but also fosters understanding and trust among stakeholders, making it a comprehensive and interpretable tool for predicting crop yields.**

Keywords— Data-driven agriculture, Interpretability, Integrated Gradients, Long short-term memory, Shapley Additive Explanations.

I. INTRODUCTION

Agriculture is a crucial sector of the Indian economy, serving as a fundamental component of the country's socio-economic structure. Not only does it serve as a major means of earning a living, but it also plays a crucial role in contributing to the country's Gross Domestic Product (GDP). Crop yield prediction plays an important role in agricultural practices and policy-making. Precise forecasts may result in efficient distribution of resources, well-informed decision-making, and improved food security. Developing trustworthy models for crop yield prediction in India is tough and vital due to the country's distinct climatic zones and wide range of grown crops. The utilization of LSTM networks is motivated by the necessity to capture temporal dependencies embedded within the complex dynamics of crop yield data. Unlike conventional models, LSTM networks excel at learning and retaining sequential information over extended time intervals.

The integration of LSTM with IG and SHAP represents an innovative methodology within the field of agricultural forecasting. LSTM is a time-series model which is used to predict output by training with historical data, while IG and SHAP provide an explanation for the underlying reasons behind these patterns. This comprehensive approach not only improves the accuracy of predictions but also adds a level of transparency and interpretability to the model. The clarity of information is very significant, as it allows us to get insights into the specific ways in which various elements, such as weather conditions or soil qualities, are impacting crop yields. Integrated Gradients as a feature attribution technique brings a new dimension to the interpretability and training efficiency of the model. By quantifying the contribution of each feature throughout the input space, Integrated Gradients offers a nuanced understanding of the model's decision-making process.

Furthermore, the proposed approach integrates SHapley Additive exPlanations (SHAP) analysis, a technique renowned for its ability to provide insightful and coherent explanations for model predictions. SHAP values allow for the dissection of individual feature contributions, offering stakeholders a transparent view into the factors influencing crop yield outcomes.

Figure. 1. Crop Harvest Using Modern Machines

The incorporation of LSTM is particularly remarkable, given its use in agricultural data analysis is not thoroughly investigated, particularly in the Indian setting. The model's capacity to capture temporal dependencies in data may provide valuable insights into the effect of different variables on crop yields throughout time. Through empirical evaluations on real-world crop yield datasets, this research work demonstrate the effectiveness and superiority of the integrated approach compared to traditional models. The amalgamation of LSTM, Integrated Gradients, and SHAP analysis not only enhances prediction accuracy but also provides a robust framework for understanding the intricate relationships between input features and crop yield outcomes. This paper contributes to the advancement of precision agriculture, offering a sophisticated toolset for stakeholders to navigate the complexities of modern farming with greater confidence and precision.

II. LITERATURE REVIEW

[1] A. Mateo-Sanchis, J. E. Adsuara, M. Piles, J. Munoz-Marí, A. Perez-Suay and G. Camps-Valls, "Interpretable Long Short-Term Memory Networks for Crop Yield Estimation," [2] Pandey, K., et al. (2023). "Development of Machine Learning based Models for Multivariate Prediction of Wheat Crop Yield in Uttar Pradesh, India" utilized thirteen years of meteorological data for training machine learning regression models, with the Random Forest model achieving the best accuracy. [3] Dr. M. Sunil Kumar1, et al. (2023). The paper, although not directly mentioning crop yield prediction, likely discusses the application of machine learning in agriculture. [4] Gopi, S. R., & Karthikeyan, M. (2023). "Hybrid Moth Flame Optimization with Machine Learning (HMFO-ML) model for crop recommendation and yield prediction" achieved a high R2 score of 98.82% and accuracy of 99.67%. [5] Salokhe, M. R. (2023). This paper explores machine learning algorithms for accurate crop yield predictions, revolutionizing the agricultural sector. [6] S. M. (2023). The paper discusses crop yield prediction in agriculture using machine learning and data mining techniques. [7] Chaitra, B., & Meena, K. (2023). The abstract does not specifically mention crop yield prediction but likely covers related machine learning applications in agriculture. [8] Jothilakshmi, T., et al. (2023). "Crop Yield Prediction Data Analytics in Indian Agriculture Using Deep Learning" employs the MLP algorithm and found that an ensemble technique provides more accurate predictions. [9] S. S., et al. (2023). "A New Approach of Machine Learning and Deep Learning Algorithms Based Crop Yield Prediction" combines SVM with LSTM and RNN algorithms for enhanced crop yield prediction. [10] Kumar, A., et al. (2023). "Comparative Analysis of Crop Yield Prediction Using Machine Learning" finds that the Random Forest model delivers the highest accuracy of 99.31% among various ML techniques. [11] Lakshmanarao, A., et al. (2023). "Crop Yield Prediction using Regression Models in Machine Learning" focuses on using regression models like Lasso, Kernel Ridge, Elastic-Net Regression, and Stacking Regression for forecasting crop yields in India. [12] Reddy, K. L., & Kumar, A. S. (2023). This paper examines various machine learning techniques such as gradient boosting, decision tree, and random forest algorithms for crop yield prediction. [13] Mistry, V., et al. (2023). "Machine Learning Use Case in Indian Agriculture: Predictive Analysis of Bihar Agriculture Data to Forecast Crop Yield" applies various ML techniques to predict crop yield, aiding farmers in decision-making. [14] Sapthami, I., & Krishna, B. M. (2023). "Transfer Learning Techniques For Improving The Quality Of Prediction In Crop Yield" explores Random Forest, SVM, and K Nearest Neighbor for crop yield prediction. [15] S M, S. G., C H, M., N, A., N M, S., B P, S., & Sreepathi, B ., et al. (2022). “Crop Yield Prediction based on Indian Agriculture using Machine Learning”. Utilizing data analysis, Information and Communication Technology (ICT), and forecasting techniques are crucial for advising farmers in India, facing climate challenges, to optimize crop production and make informed decisions. [16] Saini, P., & Nagpal, B ., et al. (2022). “Deep-LSTM Model for Wheat Crop Yield Prediction in India”. The study in India demonstrates that applying Deep-LSTM models for early-season Wheat Crop Yield Prediction, utilizing data from 1950 to 2019, outperforms traditional methods with a lower RMSE value of 0.20, providing valuable insights for farmers in decision-making during the crop season.

[17] Sharma, A., Tamrakar, A., Dewasi, S., & Naik, N. S ., et al. (2022). “Early prediction of crop yield in India using Machine Learning”. addresses the challenge of predicting crop yield in India using Random Forest Regressor for crop recommendation as well as prediction.

III. PROBLEM STATEMENT

Despite the significant strides made in leveraging deep learning for crop yield prediction, existing models often face challenges related to the accurate capture of temporal dependencies and the lack of interpretability. Traditional models may struggle to adapt to the dynamic nature of agricultural data, leading to suboptimal predictions. Additionally, the black-box nature of some deep learning models hinders stakeholders' ability to comprehend and trust the predictions, impeding informed decision-making in agriculture. Furthermore, the lack of interpretability in many existing models hampers stakeholders' ability to understand and trust the predictions, impeding their ability to make well-timed and well-informed decisions. The temporal dynamics associated with factors such as weather patterns, soil conditions, and crop growth stages play a pivotal role in influencing crop yield outcomes. Conventional models, which may not adequately capture these temporal nuances, often fall short in providing accurate predictions, particularly in the face of changing environmental conditions. The relative importance of various features influencing crop yields is critical for effective decision-making. Traditional models may lack detailed feature attribution mechanisms, leaving stakeholders with limited insights into which factors significantly impact the predictions. This gap in interpretability restricts the ability to prioritize and act upon specific features influencing crop yield outcomes. Crop yield outcomes are the result of a complex interplay of numerous factors, including environmental conditions, agronomic practices, and genetic traits. Previous models may not adequately account for the multifactorial influences that contribute to crop yield variations. Existing models may encounter challenges when applied to diverse crops, regions, or varied agricultural contexts. The scalability and generalizability of models become pertinent considerations as agricultural systems differ, requiring models that can adapt to the specific characteristics of different crops and environments.. The inability of previous models to adapt to the dynamic nature of these temporal dependencies leads to suboptimal predictions, particularly in scenarios where the impact of historical data extends over multiple seasons. Many existing crop yield prediction models face challenges in achieving high levels of accuracy. Traditional statistical methods might oversimplify the underlying complexities of the agricultural system, while basic machine learning models may lack the capacity to discern non-linear relationships and subtle patterns in the data.

IV. EXISTING SYSTEM

Traditional approaches often rely on statistical methods, regression models, or basic machine learning algorithms to analyze historical data and predict future crop yields. These methods, while providing a foundational understanding of crop yield trends, often face limitations in capturing the intricate temporal dynamics and complex relationships present in agricultural datasets. Some existing models incorporate temporal aspects using simpler time-series techniques, but they may not fully exploit the benefits of more sophisticated temporal models. Existing systems, particularly those based on statistical methods, may struggle to effectively model temporal dependencies in agricultural data. Crop growth is influenced by various time-sensitive factors such as seasonal changes, weather patterns, and planting cycles. The inadequacy in capturing these temporal intricacies limits the accuracy of predictions. Traditional machine learning algorithms may overlook long-range dependencies, hindering their performance in predicting crop yields influenced by evolving environmental condition. Lack of interpretability can make it difficult to identify biases or errors within the model. If the model is making predictions based on inaccurate or biased features, without clear insights into the feature contributions, it becomes challenging to rectify and improve the model. However, previously used machine learning models may struggle to adapt to abrupt changes or events, and their performance can be limited by the availability of long-term, high-quality historical data. The black-box nature of some machine learning models impedes the interpretability of predictions. Farmers and stakeholders require transparent insights into the factors influencing crop yield outcomes to make informed decisions. The absence of interpretable insights limits the trust and adoption of these models in practical agricultural settings where transparency is crucial. Previous system faced interpretability challenges, where interpretability is also an essential factor in predicting crop yield.

Interpretability is crucial for integrating model predictions into real-world decision-making processes. If farmers cannot interpret and understand why a particular prediction was made, they may be less likely to act upon the recommendations. The lack of interpretability hampers the practical utility of the model as a decision support tool for farmers and agricultural practitioners. . A lack of clear insights into feature contributions and decision-making processes hampers the ability to learn from model behavior and iteratively improve the model. Addressing the interpretability challenge is vital to ensuring that predictive models align with the needs, understanding, and trust of the end-users in the agriculture sector.

V. PROPOSED SYSTEM

This paper aims to overcome the limitations of existing crop yield prediction models by introducing an integrated approach that leverages advanced deep learning techniques. The key components of the proposed system include Long Short-Term Memory (LSTM) networks, Integrated Gradients (IG) for feature attribution, and SHapley Additive exPlanations (SHAP) analysis. LSTM networks are employed to model sequential data and long-term dependencies, enabling the system to better understand the historical patterns and temporal dynamics that influence crop yields. This allows for improved prediction accuracy, particularly in situations where traditional models may struggle. Long Short Term Memory takes historical data and refers how prediction and training has been done. It is one of the most powerful time series model used in many prediction systems. Long Short Term Memory enhances The model's capability to grasp temporal dependencies within crop yield data. Integrated Gradients enhance the interpretability and transparency of the model predictions. Integrated Gradients is introduced as a feature attribution technique, quantifying the contribution of each feature throughout the input space. This not only aids in model interpretability but also addresses concerns related to the "black-box" nature of some machine learning models, making the predictions more understandable and trustworthy. Shapley Additive exPlanations provide detailed insights into feature contributions and relationships. SHAP analysis is integrated to offer a comprehensive breakdown of how each feature contributes to the model's predictions. This ensures a transparent understanding of the factors influencing crop yield outcomes, allowing stakeholders to prioritize and act upon specific features based on their impact .. This includes the seamless integration of LSTM, IG, and SHAP components, creating a unified framework for crop yield prediction. The model is trained on historical data, allowing it to learn both the temporal dependencies and feature contributions, thereby enhancing its predictive accuracy and interpretability. Real-world crop yield datasets are utilized to empirically evaluate the effectiveness of the proposed system. Comparative analyses are conducted to showcase the superiority of the integrated approach over traditional models in terms of accuracy and interpretability. It represents a comprehensive and integrated approach to crop yield prediction, addressing the limitations of existing models by combining advanced temporal modeling, feature attribution, and interpretability techniques. The LSTM model, a kind of Recurrent Neural Network (RNN), was selected because to its expertise in processing time-series data, which is essential for assessing and forecasting agricultural crop yields. The LSTM architecture excels at capturing temporal dependencies and sequences in data, making it particularly valuable in agricultural contexts where crop output is impacted by sequential environmental and agronomic variables. LSTM functions by retaining a 'memory' of prior knowledge, allowing it to provide well-informed predictions by considering both the current input and the contextual information from previously encountered data. This is especially advantageous in agricultural datasets where the present crop output might be affected by variables from prior years, such as cumulative rainfall patterns or fluctuations in soil nutrients. The LSTM architecture used in this investigation relied on many pivotal components:

1) Input Gate: The input gate regulates the amount of freshly received information that is preserved.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (1)$$

$$\widetilde{C_t} = tanh\ tanh\ (W_C \cdot [h_{t-1}, x_t] + b_C) \quad (2)$$

2) Forget Gate: Determines the data that should be eliminated from the cell state, enabling the model to disregard irrelevant information.

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (3)$$

3) Cell State: The cell state in an LSTM functions as a linear conveyor belt that spans the whole chain, facilitating the flow of information with minimal linear interactions. The system retains and transports pertinent information throughout the sequence processing, guaranteeing the availability of context to produce precise predictions.

$$C_t = f_t * C_{t-1} + i_t * \widetilde{C_t} \quad (4)$$

4) Output Gate: The output gate is responsible for determining the subsequent hidden state of the LSTM, which encodes information from past inputs and enables the LSTM to make educated predictions.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (5)$$

$$h_t = o_t * tanh\ tanh\ (C_t) \quad (6)$$

The LSTM model used in this work included sequential data pertaining to crop yields, encompassing variables such as rainfall, fertilizer utilization, and other pertinent agronomic data. Configured with a specialized architecture featuring 50 neurons in the LSTM layer, coupled with a Dense layer housing a solitary output neuron, the model was finely tuned for precise yield predictions. Training involved the utilization of the Rectified Linear Unit (ReLU) activation function, optimized through utilizing the Adam optimizer, with the mean squared error employed as the loss function.

The LSTM's distinctive capability to scrutinize time-series data emerged as a cornerstone, facilitating a comprehensive understanding of patterns and trends in agricultural yields over diverse temporal spans. The inherent capacity of the LSTM to analyze time-series data emerged as a cornerstone in comprehending patterns and trends within agricultural yields over an extended period. Its design proved particularly advantageous in managing the temporal component of the data, adeptly capturing subtleties and relationships across multiple time periods. This temporal analysis was essential for unraveling intricacies in crop yield dynamics, rendering the LSTM an indispensable tool in predicting agricultural outcomes with precision. The model's unparalleled capacity to discern relationships across various time periods solidified its status as a profound instrument for understanding the complex dynamics of crop yield prediction. This nuanced approach, tailored to temporal complexities, positions the LSTM as a pivotal tool, offering profound insights and contributing significantly to advancements in agricultural forecasting. The model's ability to unveil profound insights into the dynamic nature of crop yield prediction underscores its significance in advancing precision agriculture practices. In conclusion, the LSTM model, with its robust architecture and temporal analysis capabilities, stands at the forefront of revolutionizing how perceive and predict agricultural yields. This nuanced approach, tailored to the temporal complexities inherent in agricultural data, positions the LSTM as a pivotal instrument in the predictive analytics toolbox. Its capacity to capture subtleties and relationships across diverse time periods underlines its significance in providing a comprehensive understanding of the intricate dynamics influencing crop yield predictions.

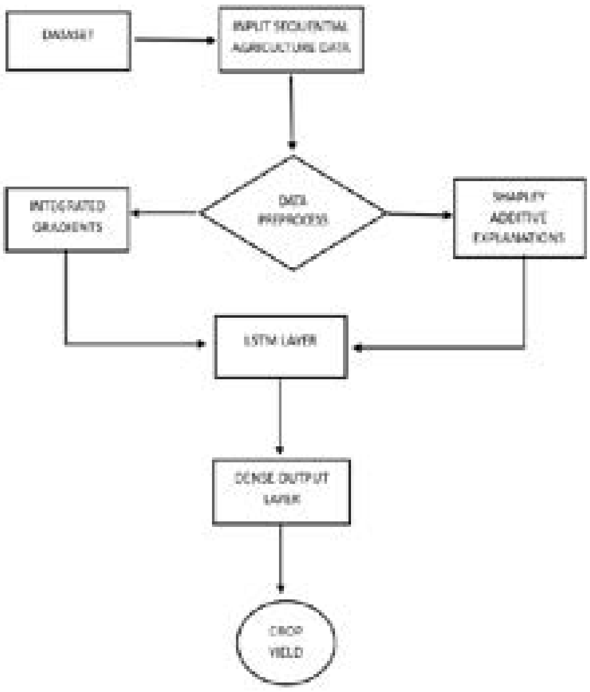


Figure. 2. Algorithm architecture

VI. METHODOLOGY

1) Dataset overview and preparation:

The dataset crucial to this study comprises a comprehensive collection of agricultural data, extending from 1997 to 2023, across numerous states in India. It delivers a thorough view on different crops, encompassing a vast diversity of key variables critical for crop production prediction. These parameters include the kind of crop, the year of cultivation, particular cropping seasons, and the geographical area (Indian states) where the crops were cultivated. Additionally, the dataset offers quantitative metrics such as the area under cultivation (measured in hectares), the overall output amount (in metric tons), and estimated yields. It also includes major climatic elements including yearly rainfall (in mm) and agronomic inputs, particularly the utilization of fertilizers and pesticides (quantified in kg). This rich combination of data gives a unique chance to study and estimate crop yields, offering insights into the interaction of environmental, agronomic, and geographical variables on agricultural production.

2) Data Preprocessing & Cleaning:

Data pre-processing: In the preprocessing phase, a structured approach was employed to condition the dataset for deep learning models. Categorical features underwent a transformation procedure, where methods like one-hot encoding were utilized. In this strategy, a category variable

$$C\_i = \{ 1, if\ the\ original\ value\ is\ the\ i - th\ category \\ 0, otherwise \}$$

For numerical characteristics, normalization was required to align the data more closely with Gaussian distribution, applying the Power Transformer and the Yeo-Johnson transformation. Transformation, applied to each characteristic.

$$x\_i' = \{ [(x\_i + 1)^\wedge\lambda - 1] / \lambda, if\ \lambda \neq 0, x\_i \geq 0 \\ ln(x_i + 1), if \lambda = 0, x_i \geq 0 \\ -[(-x_i + 1)^{(2-\lambda)} - 1]/(2 - \lambda), if \lambda \neq 2, x_i < 0 \\ -ln(-x\_i + 1), if\ \lambda = 2, x\_i < 0 \} \quad (7)$$

3) Training the model:

Datasets are taken from Open Government Data (OGD), Food and Agricultural Organisation Data (FAO) etc., These datasets contain the necessary raw data and features that can be fed to the model for training. For training the model, datasets are splitted into two parts. One part for training and another one for testing. Typically, this split follows a set ratio, such as 70% for training and 30% for testing, or 80% for training and 20% for testing. This splitting is done randomly to guarantee that both sets are reflective of the entire data distribution. The training set is used to train the model, whereas the testing set is reserved to assess its predicted performance. Various machine learning models were built, each optimized to capture the intricacies of the agricultural dataset. The models comprised Linear Regression, Random Forest Regressor, CatBoost Regressor, and LSTM (Long Short-Term Memory) networks. For each model, a particular set of parameters was determined to enhance performance.

The LSTM (Long Short-Term Memory) model, a form of recurrent neural network appropriate for time-series data, was constructed using TensorFlow's Keras framework. This model contained a single LSTM layer with 50 neurons and employed the ReLU activation function. It was constructed to accept input shaped according to the altered training data, and a Dense layer with a single neuron was added for output prediction. The LSTM model was constructed using the Adam optimizer and mean squared error as the loss criteria. Training included 100 epochs with a batch size of 32, and model performance was assessed using the R^2 score for both the training and testing sets. The CatBoost Regressor, recognized for its efficacy with categorical data, was programmed with a learning rate of 0.15. The training method was basic, comprising fitting the model to the altered training data and testing its accuracy using the R^2 score on both training and test datasets. Additionally, a Random Forest Regressor was deployed, utilizing the default setup from scikit-learn's RandomForestRegressor. The model was trained on the converted data, and its accuracy was assessed based on the R^2 score for training and test datasets.

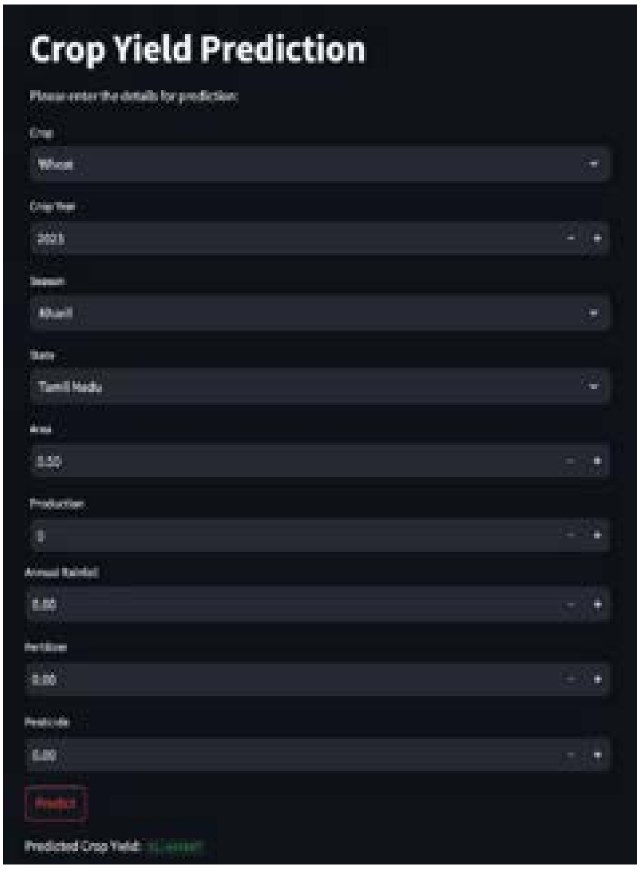


Figure. 3. User Interface

Figure 3 shows the user interface of the prediction system

Lastly, a Linear Regression model was also included in the study. Using scikit-learn's implementation, the model was fitted to the altered training data. Similar to the other models, its performance was assessed using the R^2 score for both the training and testing stages. Overall, the selection and training of these models represent a comprehensive approach to studying and forecasting crop yields, with each model adding distinct capabilities to the analysis. The precise setup of the LSTM model, in particular, displays its aptitude for handling the dataset's temporal elements.

VII. RESULTS

During the model assessment and validation phase the performance of each machine learning model underwent thorough evaluation on both the training and testing datasets. The performance of the models in correctly forecasting crop yields and their capacity to generalize to new data were assessed using the R^2 score, a statistic that reflects the degree of predictability of the dependent variable based on the independent variables, measuring the extent of variation captured. The Linear Regression model demonstrated outstanding performance, with a training accuracy of 0.8568 and a test accuracy of 0.8201. As a baseline model, these findings suggested a strong match with a moderate amount of complexity, making it a trustworthy alternative for crop production prediction. However, when the Variance Inflation Factor (VIF) was introduced to address multicollinearity in the Linear Regression model (LRvif), there was a little loss in accuracy, with the training and test accuracies being 0.8514 and 0.8107, respectively. This revealed that although the elimination of multicollinear features marginally lowered the model's predictive ability, it also assisted in balancing bias and variance.

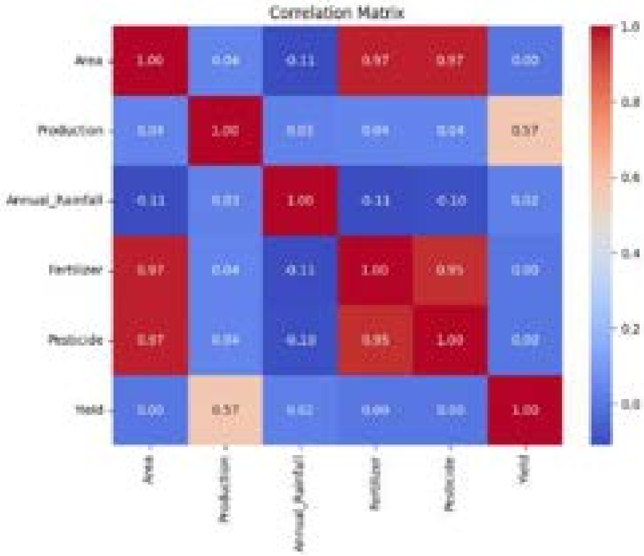


Figure. 4. Shows that the LSTM is trained with features like Area, Annual rainfall, Production, Fertilizer, Pesticide and Yield.

The Random Forest Regressor displayed extraordinarily high accuracy, especially in the training set, with scores of 0.9958 for training and 0.9785 for testing. This highlighted the model's resilience and its potential to capture complicated non-linear correlations in the data. The CatBoost Regressor likewise exhibited outstanding performance, notably in the training phase, with an accuracy of 0.9998, and a test accuracy of 0.9692. Although this emphasized the model's efficacy, it also increased the danger of overfitting. Notably, the LSTM (Long Short-Term Memory) model demonstrated great accuracy on both the training and testing sets, scoring 0.9821 and 0.9748, respectively. In order to improve the comprehensibility of the LSTM model, an Integrated Gradients (IG) and SHapley Additive exPlanations (SHAP) are used. By using these methodologies, model's decision-making process can be analyzed, so gaining valuable insights into the many aspects that drove the forecasts. IG helps in attributing predictions to specific features by assessing the model's sensitivity to changes in each feature. SHAP values assign a value to each feature, indicating its contribution to the prediction. In crop yield prediction, SHAP analysis helps identify which features, such as temperature, precipitation, or specific soil nutrients, have the most significant influence on the predicted yield.

The decision to adopt Long Short-Term Memory (LSTM) networks was driven by their effectiveness in handling time-series data, a critical aspect of agricultural datasets where crop yields are influenced by time-varying variables. LSTM's unique ability to capture temporal relationships in the data proved invaluable, making it a well-suited option for this application. The incorporation of LSTM with Integrated Gradients (IG) and Shapley Additive exPlanations (SHAP) further enhanced the model's performance. IG and SHAP are interpretability tools that, when combined with LSTM, not only improved the accuracy of predictions but also provided a deeper understanding of the model's inner workings. This dual integration not only increased transparency but also bolstered the reliability of the predictions. The interpretability aspect is crucial in the agricultural domain, as it enables stakeholders to comprehend the factors influencing crop yields and make informed decisions.

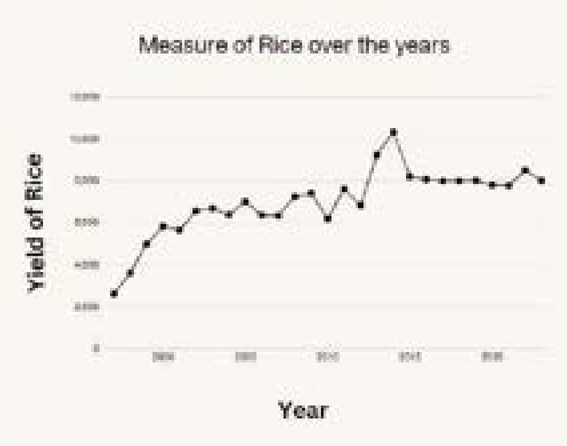


Figure. 5. Measure of Rice over the years

Figure 5 shows the amount of rice that is yielded over the years.

LSTM showed brilliant results rather than the rest of the models used for analysis. Integrating LSTM with Integrated Gradients and SHapley Additive exPlanations (SHAP) enables not only to produce precise results, but also to know how the model predicts output and can also know the importance of the contribution of each feature.

The synergy between LSTM, IG, and SHAP showcased a holistic approach, elevating the model's effectiveness in dealing with the complexities of agricultural time-series data. This comprehensive methodology not only enhances prediction accuracy but also empowers users with insights into the model's decision-making process, ultimately fostering trust and confidence in its applications within the agricultural sector.

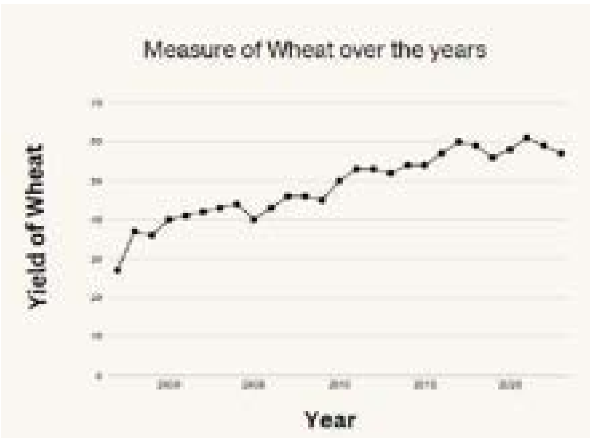


Figure. 6. Measure of Wheat over the years

Figure 6 shows the amount of wheat that is yielded over the years.

Finally, the assessment and validation process was crucial in establishing the most effective model for forecasting crop yields. The LSTM model stood out owing to its high accuracy and fit for the temporal structure of the dataset, revealing its potential as a strong tool in agricultural production prediction. The findings from the multiple machine learning models applied in the research showed varied performance characteristics in terms of accuracy, precision, and other important metrics, offering a full knowledge of each model's strengths and limits in forecasting crop yields. The Linear Regression model demonstrated good accuracy levels, with training and test accuracies of 0.8368 and 0.8101, respectively. This model's strength resides in its simplicity and interpretability, giving it a useful foundation for comparison. However, its shortcoming in capturing complicated, non-linear correlations in the data was clear.

The Linear Regression with VIF (LRvif) revealed a modest decline in performance, which suggested that addressing multicollinearity with VIF slightly reduced the model's predictive capabilities. The Random Forest Regressor displayed higher performance, notably in the training phase, with a training accuracy of 0.8558 and test accuracy of 0.8485. This model's capacity to handle non-linear data and its durability against overfitting (shown by the high test accuracy) make it a good candidate. However, its complicated structure may occasionally lead to issues in interpretation. The CatBoost Regressor displayed remarkable training accuracy (0.8498) but a little lower test accuracy (0.8292), indicating a possible overfitting concern despite its great overall performance. This approach is especially effective at handling categorical data, which is prevalent in agricultural datasets. Among these, LSTM suited best for time series models and handling historical data. LSTM performed better than the traditional algorithms and achieved a tremendous accuracy in predicting the crop yield based on the environmental and agricultural features.

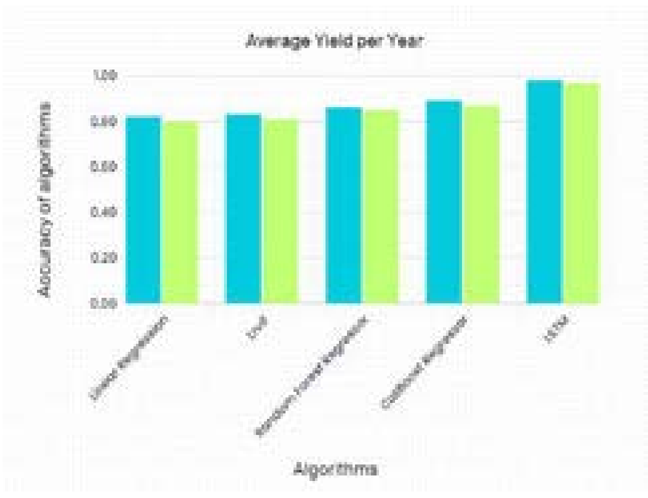


Figure. 7. Model Performance Graph

Figure 7 shows the performance of each graph on performing analysis of time series models. LSTM acts as best model for time-series predictions.

Significantly, the LSTM (Long Short-Term Memory) model stood out with excellent training and testing accuracies of 0.9821 and 0.9748, respectively. The LSTM's success may be ascribed to its capacity to interpret time-series data, capturing temporal dependencies that are vital in comprehending crop production patterns driven by sequential climatic and agronomic variables. The adaptability of the LSTM model proves tailor-made for datasets heavily influenced by time-based variables, making it an ideal fit for agricultural scenarios. With the incorporation of Integrated Gradients (IG) and Shapley Additive exPlanations (SHAP), this integrated approach offers a thorough assessment, shedding light on performance, interpretability, and practical implications in crop yield prediction.

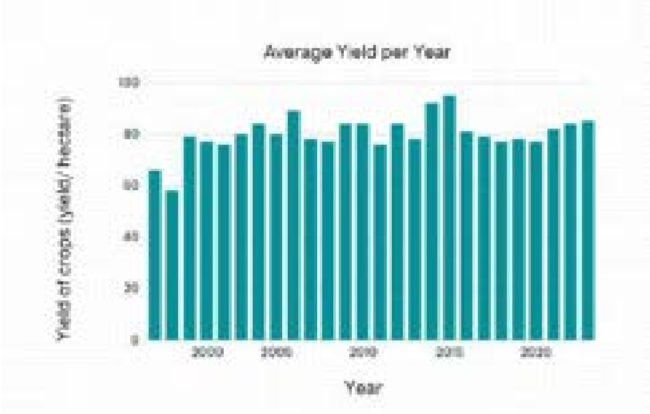


Figure. 8. Average Yield over the years

Figure 8 shows the average yield of crops over the years.

The LSTM's robust performance underscores the feasibility of deploying advanced neural network models in agricultural production forecasting, especially in datasets characterized by substantial temporal intricacies. This amalgamation of LSTM, IG, and SHAP holds promise in advancing the precision and transparency of crop yield predictions in dynamic agricultural landscapes.

VIII. CONCLUSION

The main motive of this research is to provide farmers and stakeholders a transparent crop yield prediction system. Due to the varying climatic changes, farmers are struggling to decide the crops yield. So a prediction system can be helpful for them to decide and yield crops wisely. The research effectively applied multiple machine learning models to estimate crop yields, revealing new insights into their application and usefulness in agricultural data analysis. The LSTM model, with its capacity to handle time-series data, emerged as especially good at capturing the temporal characteristics crucial in agricultural production prediction.

Moving ahead, various options for further work are clear. Enhancing the LSTM model by experimenting with other topologies, such as stacked LSTMs or bidirectional LSTMs, might further increase its prediction accuracy. Incorporating new data sources, such as soil health metrics or more detailed climate data, can potentially boost the models' capacity to estimate yields more precisely. The integrated approach employing Long Short-Term Memory (LSTM) networks, Integrated Gradients (IG), and SHapley Additive exPlanations (SHAP) analysis presents a significant advancement in the field of crop yield prediction. The findings from the evaluation demonstrate the effectiveness of the proposed model in addressing the limitations of existing systems and offering practical solutions for stakeholders in the agriculture sector. Furthermore, expanding the research to cover a wider variety of crops and geographical locations might yield more generalizable insights, thereby benefiting in agricultural planning and policy-making on a bigger scale..

In conclusion, this study establishes a good platform for future initiatives in agricultural yield prediction using deep learning, showing the promise of sophisticated models like the LSTM in tackling complex, real-world difficulties in agriculture. Using LSTM and with integration of Integrated Gradients (IG) and SHapley Additive exPlanations (SHAP) and achieved an overall accuracy of 97%.

Enhancing the proposed system with adaptive learning strategies enables dynamic adjustments to changing environmental conditions, potentially enhancing model generalizability. Future integration of emerging adaptive learning methods can further improve predictive accuracy in response to environmental changes.

REFERENCES

[1] A. Mateo-Sanchis, J. E. Adsuara, M. Piles, J. Munoz-Marí, A. Perez-Suay and G. Camps-Valls, "Interpretable Long Short-Term Memory Networks for Crop Yield Estimation," in IEEE Geoscience and Remote Sensing Letters, vol. 20, pp. 1-5, 2023, Art no. 2501105, doi: 10.1109/LGRS.2023.3244064

[2] Pandey, K., Sukirti, Danodia, A., & Karnatak, H. (2023). Development of Machine Learning based Models for Multivariate Prediction of Wheat Crop Yield in Uttar Pradesh, India. Journal of Geomatics, 17(2), 70-77. https://doi.org/10.58825/jog.2023.17.2.70

[3] Kumar, M. S., Girinath, S., Guru, G., Venkata Siva, A., & Ganesh, K. J. (2023). Crop Yield Prediction Using Machine Learning. IEEE Explore. https://doi.org/10.1109/ICSEIET58677.2023.10303423

[4] Gopi, S. R., & Karthikeyan, M. (2023). Effectiveness of Crop Recommendation and Yield Prediction using Hybrid Moth Flame Optimization with Machine Learning. Engineering, Technology & Applied Science Research, 13(1), 6092-6096. https://doi.org/10.48084/etasr.6092

[5] Salokhe, M. R. (2023). Application of machine learning in agriculture for crop yield prediction. International Journal of Advanced Research in Science, Communication and Technology, 12088. https://doi.org/10.48175/ijarsct-12088

[6] S., M. (2023). Crop yield prediction based on machine learning. International Journal of Scientific Research in Environmental Management, 1(1). https://doi.org/10.55041/ijsrem24714

[7] Chaitra, B., & Meena, K. (2023). Crop Yield and Price Prediction Using Machine Learning. IEEE International Conference on Inventive Computation Technologies, Article No. 10220616. https://doi.org/10.1109/ICIET57285.2023.10220616

[8] Burdett, H., & Wellen, C. (2022). Statistical and Machine Learning Methods for Crop Yield Prediction in the Context of Precision Agriculture. Precision Agriculture, 23, 1237-1254. https://doi.org/10.1007/s11119-022-09897-0

[9] S., S., Vaishnavi, K. V. S., B., S., B., V., & Reddy, B. V. (2023). A New Approach of Machine Learning and Deep Learning Algorithms Based Crop Yield Prediction. Advances in Computer and Information Sciences and Management Studies, 2023, Article No. 18. https://doi.org/10.53759/acims/978-9914-9946-9-8_18

[10] Kumar, A., Banerjee, K., Kumar, P., Aiman, K., Sonkar, M., Rajput, R., & Asif, M. R. (2023). Comparative Analysis of Crop Yield Prediction Using Machine Learning. IEEE International Conference on Innovative Computing and Communication Technologies, 2023, Article No. 10141745. https://doi.org/10.1109/InCACCT57535.2023.10141745

[11] Lakshmanarao, A., Kumar, M. N., Ratnakar, K. S. V., & Satwika, Y. (2023). Regression Models in Machine Learning for Forecasting Crop Yields in India. IEEE International Conference on Advances in Artificial Intelligence and Computing, 2023, Article No. 10141462. https://doi.org/10.1109/ICAAIC56838.2023.10141462

[12] Reddy, K. L., & Kumar, A. S. (2023). Machine Learning Techniques for Weather based Crop Yield Prediction. IEEE International Conference

on Advances in Intelligent Systems, 2023, Article No. 10073740. https://doi.org/10.1109/ICAIS56108.2023.10073740

[13] Mistry, V., Mishra, A., & Ahmed, N (2023). Machine Learning Use Case in Indian Agriculture: Predictive Analysis of Bihar Agriculture Data to Forecast Crop Yield. International Journal for Research in Applied Science and Engineering Technology, 2023(1), 487-495. https://doi.org/10.22214/ijraset.2023.48709

[14] Sapthami, I., & Krishna, B. M. (2023). Transfer Learning Techniques for Crop Yield Prediction Using Various Machine Learning Methods. IEEE International Conference on Computing, Communication, and Intelligent Systems, 2023, Article No. 10128344. https://doi.org/10.1109/ICCCI56745.2023.10128344

[15] S M, S. G., C H, M., N, A., N M, S., B P, S., & Sreepathi, B. (2022). Crop Yield Prediction based on Indian Agriculture using Machine Learning. International Journal of Advanced Research in Science, Communication, and Technology, 5881. https://doi.org/10.48175/ijarsct-5881

[16] Saini, P., & Nagpal, B. (2022). Deep-LSTM Model for Wheat Crop Yield Prediction in India. IEEE Conference on Contemporary Computing and Informatics, 2022, Article No. 00025. https://doi.org/10.1109/CCiCT56684.2022.00025

[17] Sharma, A., Tamrakar, A., Dewasi, S., & Naik, N. S. (2022). Early Prediction of Crop Yield in India using Machine Learning. IEEE TENCON - Spring Conference, 2022, Article No. 9864490. https://doi.org/10.1109/TENSYMP54529.2022.9864490

Role of Digital Twins in Promoting Sustainability in Commerce

Venkatesan R
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
rlvenki2000@gmail.com

Brijit Benny
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
brijitbenny@karunya.edu.in

Shirley C P*
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
shirleydavidlivingston@gmail.com

Berin Jeba Jingle I
Department of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
berinjeba@karunya.edu

***Abstract*—As the virtual and real worlds continue to collide, digital twins may become important tools for finding solutions to some of the most serious challenges facing the sustainability movement. These challenges include making effective use of resources, adjusting to the effects of climate change, and ensuring that everyone has access to a high quality of life. The first steps have been taken on the road to making widespread use of digital twins. It is imperative that their connectedness should be secured if their potential is to be fully utilized. This presents a number of challenging situations. The digital transition may make it possible to achieve several of the United Nations' sustainable development objectives. Data and technology should assist the user to make appropriate judgements that are better suited to their circumstances. It's feasible that the computer model can see into the future and predict the user about the consequence. This may be used to make predictions about the future, analyze the present, and analyze how a decision will play out in the future. These features may have an effect on a number of different sectors including commerce. The use of dynamic modelling may be beneficial for improving efficiency in the management of supply chains, transportation, and crowds.**

Keywords—Artificial Intelligence, Cloud computing, Sustainability, Implementation, Computers, Organization, Digital twins, Logistic, Network, Sustainability

I. INTRODUCTION

Immense scope for improvement is offered by AI. It may be used for a variety of cognitive tasks, from information gathering and data reduction to the modelling of physically unpredictable processes. Thus, AI allows for faster simulations and the extraction of some of the most crucial information from large data sets. Machine learning is expected to help with both falsely characterizing physical processes and increasing computer efficiency, according to the studies carried out by the authors [1].

The researchers consider their methodology paper to be the first stage in developing a digital Earth replica. When comparing the many computer architectures that exist now and may exist in the near future, GPU-based quantum computers stand out as the most promising alternative. A system with over 20,000 GPUs, using roughly 20 MW of power, is required to run a digital twin on a large scale, according to estimates. It's best to use such a computer at a location with access to plenty of carbon-neutral electricity [2].

II. OBJECTIVE

The research aimed to fulfil the following objectives:

- To study What exactly is a "Digital Twin,"?
- Sustainability by Digital Twinning
- What Function Does Digital Twins Serve in the field of Commerce?

III. METHODOLOGY

From individual generators and motors to whole rail networks and airport terminals, digital twins have been simulated. However, the true magic occurs when digital duplicates work together. This is the shift that is necessary to foster lasting transformation, not just inside institutions but throughout all of society. Every human need to make sure this ecosystem has all it needs to succeed. There has to be uniformity in strategy. It's the bedrock that makes it possible for digital world models to coexist and for differences to be resolved dispassionately. For the ecosystem to gain the confidence and participation of the public, it must also be open, transparent, and fair. Furthermore, digital twins' capacity to exchange data and models will be facilitated by the existence of clear interoperability standards, such as a universal language. Last but not least, as the ecosystem develops, it must be flexible enough to accommodate shifting conditions and an expanding range.

A. What exactly is a "Digital Twin"?

The term "digital twin" is used to describe a software representation of a physical object or system. AI, big data, cryptocurrency, virtual reality techniques, collaborative networks, application programming interfaces (APIs), and open standards are all used to build a digital twin. Through these supporting technologies, a digital twin may be continuously fed data from the physical world through sensors, allowing it to provide accurate and timely insights on an object's or system's present and future performance and problems [3].

The size of the global market for virtual duplicates in 2019 was estimated at $3.8 billion. One of the most advanced sectors to adopt this new technology is the industrial or e commerce sector. Assuming a 38% annual growth rate, other sectors such as healthcare, urban development and retail are exploring the potential of digital twins, the digital twin economy is expected to be worth $26 million by 2025.

B. "Digital Twin" Concept

Throughout a system's existence, a digital twin is a dynamic virtual simulation that includes all relevant physical, operational, and functional data. Digital twins were initially conceptualized in 2002 at University of Michigan by professor Michael Grieves, but he did not call it as digit twin at the time. Instead, he described it as a model for Product Lifecycle Management [8]. A digital twin, in common parlance, is an identical digital representation of a real object. However, unlike simpler simulations, digital twins are rather complex in their own right. Operational Technology (OT) and Information Technology (IT) are combined in digital twins via a two-way flow of data between the virtual and physical assets. Additionally, digital twins have an unbeatable edge over conventional simulation approaches due to their ability to run 'what-if' scenarios throughout a component's lifespan, in addition to its predictability and controllability.

The potential of digital twins in the fields of predictability or controllability has grown thanks to the development of supporting technologies like the Internet of Things (IoT), Intelligent Systems, Data Analytics, and Machine Learning in recent years. Because digital twins are used in so many different contexts and businesses, academics and professionals have yet to agree on a single meaning for the term. BearingPoint recognizes the value of digital twins for its manufacturing, aviation, and supply chain clients and is optimistic about the potential of electronic twins to help the energy industry overcome the challenges it currently faces and those it will face in the future.

C. Digital Twins: Why They Matter

What is a digital twin and why is it vital in today's world when AI, IoT, and data analytics are all gaining ground? Simply put, digital twins are digital copies of physical things or computerized systems. Multiple applications exist for the digital twin/processes/systems, which use data, software analytics, and machine learning to create interactive digital simulation models [9]. In the same way that the design, construction, and functioning of things' interplay will change as their physical counterparts do, so too will their digital simulations. To reflect its near-real-time status, operational condition, or location, a digital twin would continually learn and update from many sources [4].

Due to their ability to communicate with their physical counterparts and give data on their condition, digital twins will continue to develop alongside IoT devices. Given that 21 billion linked digital sensors are expected to be in use by 2025, the relevance of digital twins becomes readily apparent. The world's management, maintenance, and development budgets might save billions of dollars if this comes to be true.

Technology supporting digital twins has expanded to include more complex things like buildings, industries, and even whole cities; it may perhaps one day be used to humans. Digital twins are also being used in the virtual planning and testing of large and complicated items like aircraft engines, locomotives, and turbines before they are built in the real world as shown in Fig. 1. The use of digital twins is rapidly spreading across sectors, and yours won't be an exception if it hasn't already.

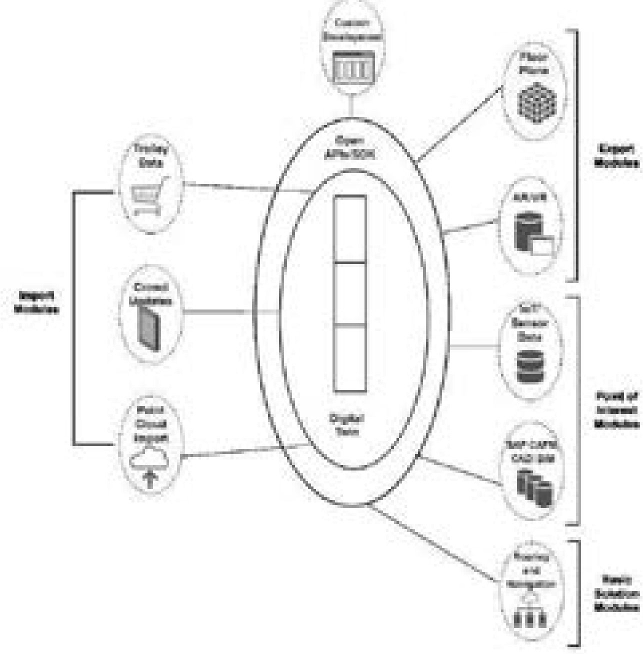


Fig 1. Digital Twin

D. What Function Do Digital Twins Serve in the Process of Reusing Existing Environmental Values?

The use of digital twins unquestionably inspires businesses to raise their levels of creativity and performance. There is no room for debate in this assertion. There are a significant variety of publications, both technical and non-technical, that highlight the many ways in which digital twin technologies may be utilized to enhance the design, visualization, monitoring, and maintenance of corporate assets. Some of these studies are included below. In addition, the technology of the digital twin has motivated a vast number of firms to re-examine the sustainability objectives that they have set for themselves, either directly or indirectly. The idea behind digital twins is that they may be able to aid the user in thinking in a manner that is not constrained by the usual industrial paradigm, which consists of extracting, creating, consuming, and discarding. This is the main idea behind digital twins. As a direct consequence of this development, companies and even entire communities now have the opportunity to switch to a method of corporate sustainability that generates almost no waste or pollution, keeps materials and commodities as part of the recycling cycle for a longer period of time, and contributes to the renewal of biological ecosystems This method of corporate sustainability is known as zero waste production. This method is sometimes referred to as "zero waste" [5].

E. Sustainability by Digital Twinning

Digital twins can help the environment in two different ways. At the moment, they are mostly used to get, store, and show data so that an accurate version of the real world can be shown. This can help the users to understand and measure what's going on, and it can also help them combine info from different sources to see how the decisions affect things. But the computer model could also be used to guess what will happen in the future. This can predict what will happen in the future, compare different scenarios, and figure out what will happen before certain choices are made [6].

In order to achieve sustainability in a variety of sectors, data twins could be an important tool. With digital twins, here is how sustainability can be achieved:

a) Monitoring and Collaboration : In order to explore and monitor the impact of manufacturing on economies, industries and global issues, digital twins allow international collaboration between factories.

b) Optimal Resource Usage : In order to provide more efficient use of resources, digital twins can help identify optimal routes for aircraft based on weather conditions, route congestion and maintenance of individual buildings.

c) Energy Efficiency and Material Selection : Digital twins offer the possibility of improving and selecting suitable materials for manufacturing purposes, which is aimed at energy savings and efficient renewable energy storage in order to contribute to sustainability efforts.

d) Product Design and Lifecycle Management : By enabling more efficient product design, planning, manufacturing, maintenance and recycling, digital twin technology provides an opportunity to achieve the sustainability objectives by making use of less resources.

e) Environmental Impact Reduction : Digital twins can contribute significantly to reducing energy and water consumption, enabling new low impact business models as well as driving sustainability and a greener economy at the speed and scale which would lead to substantial economic value and emission reductions.

f) Testing Scenarios for Sustainability Goals : The digital twins can model and understand how to reduce energy consumption and emissions, enabling organizations to test scenarios in order to meet sustainability and climate objectives.

g) Protecting Natural Environments : By making it easier to make more informed choices, faster protection of endangered species and land as well as the implementation of changes that enhance sustainability, D2S solutions can help protect nature. Finally, Digital twins offer a broad range of applications that can contribute to the sustainability efforts across different sectors, ranging from manufacturing and energy to environment protection and resource use management.

F. Attaining Sustainability In Commerce using Digital Twins

Digital twins offer a powerful set of tools to achieve sustainability in commerce. Here's how:

a) Simulating Efficiency : A Digital Twin is a virtual replica of the real system, such as factories or supply chains. This enables businesses to simulate different scenarios and assess their impact on resource use. In order to optimize energy use or experiment with various materials in order to reduce waste, they can even perform virtual tests of new layout at the production site.

b) Predictive Maintenance : Digital twins are able to predict when equipment is going to fail, through analysis of sensor data from actual worlds. This allows for preventive maintenance, reducing downtime and the need for replacement parts, which translates to less resource consumption.

c) Optimizing Logistics : In order to model the transport networks and find inefficiencies, it is possible to use data twins. To find the most fuel-efficient options, businesses can simulate different routes of delivery, reducing their carbon footprint.

d) Informed Decision Making : Digital twins provide businesses with valuable data and insight that can help them to make more informed decisions about their activities. Areas where they can reduce their environmental impact, such as water use or packaging waste, can be found in this data.

e) Transparency and Communication : The product's lifecycle from raw materials to disposal may be monitored by the use of virtual twins. This transparency facilitates the communication of sustainability efforts by businesses to customers and stakeholders, building trust and brand reputation.

IV. RESULT AND DISCUSSION

The decrease in carbon output throughout the course of the previous several years, as a result of the sustainability of digital twins, is indicated in Table 1.

TABLE I. DECREASE IN CARBON OUTPUT OVER YEARS

Years	Reduced carbon emission in (%)
2016	10%
2017	15%
2018	20%
2019	25%
2021	30%

Global carbon emissions might be reduced by 20% by 2030 if digital twins are used. Because of the energy savings that arise from using smaller but more durable materials, firms may alter their production and distribution methods.

These traits can have far-reaching effects in many different fields. Because the planning is live, it can be used in areas like controlling crowds and traffic and getting the most out of operations. Importantly for long-term sustainability, twinning can help predict how water flows through networks, keep energy grids in balance, and make communities more resilient to the effects of climate change. To get the most out of digital twins, they need to be used for more than just individual assets. For example, they should be used for whole processes or even whole companies or cities.

As companies and society become more digital, the digital twin might be able to do more than just explain what is happening and why. It might be able to predict the future, suggest solutions, and even automate big changes. This can be done throughout the lifespan of an object, process, system, or organization, adding value by combining CAPEX and OPEX through constant innovation [7].

A. *Revolutionizing Commerce with Digital Twins: A Glimpse into the Future*

Digital twins are emerging as a transformative force in the dynamic area of trade, which will change how businesses interact with customers, optimize their operations and drive innovation [10]. Digital twins can deliver a number of benefits for enhancing customer experiences, streamlining the supply chain and generating new revenue streams by creating virtual representations of physical products, processes or environments.

B. *Enhanced Customer Experience: Immersive Product Visualization and Customization*

Businesses are able to provide their customers with immersive and personalised product experiences by means of digital twins. Virtual product simulation allows clients to interact with 3D models so they can see products within their own environment, add features and explore different configurations. Incentivizing informed choice, reducing product return rates and improving customer satisfaction are the main benefits of this interaction approach [11].

C. *Optimized Supply Chain Management: Real-time Visibility and Predictive Analytics*

The use of digital twins will transform the way in which supply chain management takes place, making it possible to monitor on demand levels, logistics operation and product performance at any time. Data is collected and transferred to digital twins by sensors embedded in a physical product or asset. This makes it possible for companies to monitor the entire supply chain, predict any disruption and optimize their inventory management. It allows operations to be executed more effectively, reduce costs and improve the quality of customer service by providing timely intelligence [12].

D. *Predictive Maintenance and Proactive Service: Preventing Downtime and Enhancing Customer Support*

Digital twins enable companies to proactively identify possible equipment faults prior to their occurrence, giving them the right to take preventative maintenance measures. By analysing sensor data collected in physical assets, digital twins are capable of detecting anomalies, anticipating unavailability and setting maintenance schedules. The preventive approach will be effective in reducing equipment failure, extending life of assets and improving customer support [13].

E. *Personalized Marketing and Targeted Advertising: Tailored Recommendations and Enhanced Customer Engagement*

Digital twins enable businesses to collect an enormous amount of customer data, which enables them to tailor marketing campaigns and deliver targeted advertising. Businesses can get a better understanding of their customers' preferences and behaviour when they analyse interactions between digital twins, which gives them the ability to make individual recommendations for products, promotions or advertisements.

F. *New Revenue Streams and Innovative Business Models:*

As-a-Service Offerings and Subscription-based Business Models Digital twins promotes new revenue streams by enhancing businesses to offer as-a-service (aaS) and subscription-based business models. For example, manufacturers can offer their customers personalized recommendations, predictive maintenance services and optimized performance solutions based on data driven insights derived from digital twins. As a result of this shift to value added services, new market opportunities are created and long-lasting customer relationships strengthened [14].

V. CONCLUSION

Digital twins are being integrated into asset management practices at an increasing rate. In return, digital twins aid in tracking and uncovering means to boost output, lessen downtime, and future-proof operations. It's possible that digital twins, like BASF's, may influence decisions about where and whether to collect plastic things for reuse, repair, recycling, or disposal. Using a digital twin, you may better optimize and choose the best materials for manufacturing, reduce energy consumption, or simply improve the efficiency with which you store renewable energy. ICL claims that digital twins may help businesses take their idea of sustainable development to a more practical level. Lastly, businesses must expand sustainably if they are to succeed while reducing their negative impact on the environment.

REFERENCES

[1] M. Srivastava, “A study on sustainability and sustainable development: Detailed explanation of the terms and its evolution, concept and meaning including discussion on how to enable sustainable development.,” SSRN Electronic Journal, 2012. doi:10.2139/ssrn.3652024

[2] I. A. Moghrabi, S. A. Bhat, P. Szczuka, R. A. AL Khaled, and M. A. Dar, “Digital transformation and its influence on sustainable manufacturing and business practices,” Sustainability, vol. 15, no. 4, p. 3010, 2023. doi:10.3390/su15043010

[3] C. Allard, “Inflation in Poland: How much can globalization explain?” SSRN Electronic Journal, 2007. doi:10.2139/ssrn.967882

[4] A. Teacher, S. Sabri, C. E. Richards, A. Rajbir, and M. Acute, “Potential and limitations of digital twins to achieve the Sustainable Development Goals,” Nature Sustainability, vol. 5, no. 10, pp. 822–829, 2022. doi:10.1038/s41893-022-00923-7.

[5] Q. Guo and X. Ma, “How does the Digital Economy Affect Sustainable Urban Development? empirical evidence from Chinese cities,” Sustainability, vol. 15, no. 5, p. 4098, 2023. doi:10.3390/su15054098.

[6] A. Turkan, “Leveraging data spaces in digital twins for Smart Cities Deployment,” Advances in Civil and Industrial Engineering, pp. 80–102, 2022. doi:10.4018/978-1-6684-3833-6.ch003.

[7] N. Alamgir, “How Digital Twins will help to unlock United Nations Sustainable Development Goals,” Day 2 Tue, November 01, 2022, 2022. doi:10.2118/211073-ms

[8] Michael Grieves, John Vickers, “Digital Twin: Mitigating Unpredictable, Undesirable Emergent Behavior in Complex Systems” August 2017, DOI: 10.1007/978-3-319-38756-7_4, In book: Transdisciplinary Perspectives on Complex Systems.

[9] Mohd Javaid, Abid Haleem, Rajiv Suman, “Digital Twin applications toward Industry 4.0: A Review”, Cognitive Robotics, Volume 3, 2023,

Pages 71-92, ISSN 2667-2413, https://doi.org/10.1016/j.cogr.2023.04.003.

[10] Angira Sharma, Edward Kosasih, Jie Zhang, Alexandra Brintrup, Anisoara Calinescu, "Digital Twins: State of the art theory and practice, challenges, and open research questions", Journal of Industrial Information Integration, Volume 30, 2022, 100383, ISSN 2452-414X,

https://doi.org/10.1016/j.jii.2022.100383

[11] Mohsen Attaran, Bilge Gokhan Celik, "Digital Twin: Benefits, use cases, challenges, and opportunities", Decision Analytics Journal, Volume 6, 2023, 100165, ISSN 2772-6622, https://doi.org/10.1016/j.dajour.2023.100165.

[12] https://www.mhlnews.com/technologyautomation/article/21272553/digital-twins-are-the-new-cutting-edge-in-supply-chain-visibility

[13] Raymon van Dinter, Bedir Tekinerdogan, Cagatay Catal, "Predictive maintenance using digital twins: A systematic literature review, Information and Software Technology", Volume 151, 2022, 107008, ISSN 0950-5849, https://doi.org/10.1016/j.infsof.2022.107008.

[14] https://www.spiceworks.com/tech/iot/guest-article/driving-value-with-digital-twins/

ICA based Noise Reduction in Mobile Phone Speech Communication

A .Koushik
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India
akulakoushik2003@gmail.com

R.Varshitha
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India
2100031622@kluniversity.in

P .Sai Poojitha
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India
2100031613@kluniversity.in

Ch.Lokesh
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India
2100031295@kluniversity.in

Ch .Sabitha
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India
sabithakiran.ch@gmail.com

K .Swathi
Department of Computer Science and Engineering,
Koneru Lakshmaiah Education Foundation,
Vaddeswaram, Andhra Pradesh, India
dr.kswathi@kluniversity.in

Abstract: In mobile phone conversations dealing with noise can be tricky as it can lower the quality of speech and make it hard to hear clearly. This research investigates ways to make speech on phones better by reducing noise. It talks about methods for cutting out noise like using machine learning, adaptive filters improving quality and taking out unwanted sounds. Also, it examines the challenges and possibilities of using these techniques on phones considering things, like how complicated they are needing real time processing and adjustingto noisy environments. It also looks at how learning can help reduce noise while keeping voices clear.

Keywords—Noise reduction, Speech communication, independent component analysis, Speech recognition system.

I. INTRODUCTION

In today's world of communication, it is common to work in environments were various voices blend. This kind of lifestyle can really impact personal well-being. To tackle this challenge a technology called Adaptive Independent Component Analysis (ICA) is being used in communication tools, like smartphones, video calls and conference calls. By using statistical analysis methods ICA can effectively reduce background noise without needing to know the sources of the noise. It does this by breaking down signals into components. Its adaptive nature helps improve the quality of voice communications by enhancing the Signal to Noise Ratio (SNR). This feature provides noise cancellation capabilities making it a reliable solution, for real time situations [1].

In phone conversations ICA is often used to reduce background noise aiming to enhance speech recognition in environments. This technique involves analyzing signals to filter out noise without knowledge of its presence.[2]

Independent Component Analysis (ICA) is a method utilized in speech analysis to extract feature vectors based on principles. The goal is to generate Gabor features those are advantageous for automated speech recognition systems. These features act as bandpass filters with center frequencies and bandwidth limitations. In contrast to methods, like Mel cepstral features this new approach improves recognition accuracy by training basis vectors that reflect the characteristics of input speech patterns.[3] The process of identifying the original source signals based on input signals is known as signal separation. This method is crucial, for applications such as speech recognition and advanced hands-free communication systems.[7]

When there is an amount of speech in a room experts utilize Wavelet Based Independent Component Analysis (ICA) to differentiate the distant echoes from the mix of speech and surrounding sounds.[9]

Blind source separation has gained popularity in the field of signal processing over the years. By combining Independent Component Analysis (ICA) with wavelet threshold denoising it becomes feasible to separate speech that includes noise. This approach does not improve the quality of the separated speech signals over time. Does reduce the impact of noise on both signals, to noise ratio and overall noise levels.[10]

The article talks about a technique for sorting data using an ICA mixture model, which groups data into categories formed by a combination of sources. This method improves the mixture model by including Gaussian structure within the categories. It computes the probability of each data point belonging to a category identifies sources and combines matrices for each category.[11]

Multidimensional Independent Component Analysis (MICA) is an extension of Independent Component Analysis (ICA) that breaks down a vector into components, in subspaces those are statistically unrelated. By isolating elements that are not interconnected MICA provides insight into source separation.[12]

Independent Component Analysis (ICA) a blind source separation technique widely used in fields aims to distinguish signals based on their sources. ICA is often applied as a box without going into its workings.[16]

Let us start by explaining the noise reduction technique basedon Independent Component Analysis (ICA) in section 2. The methodologies employed in this process is presented in section 3. Additionally, demonstration on how to minimize interference by adjusting frequencies to align with the speech range is presented in section 4. In conclusion the capabilities and objectives of implementing ICA is presented.

II. LITERATURE REVIEW

Mohanaprasad and his team (2019) proposed a method, for communication using ICA. Their research revealed that this approach significantly enhances speech quality and clarity when compared to methods [1]. In a study conducted by Zhang and Etoh (2007) they introduced a domain ICA technique tailored specifically for phone conversations. Their findings demonstrated promising outcomes in minimizing background noise in challenging settings [2]. Investigating the potential of ICA to boost speech recognition precision. Lee and colleagues (n.d.) emphasized its role in extracting speech features. They underscored the advantages it provides for enhancing accuracy, in recognition tasks [3]. Jin and colleagues (2017) explore strategies to minimize background noise, in hands communication by integrating ICA with techniques to effectively diminish unwanted sound interference [4]. Finally in a study, by Lee and colleagues from 1999 they introduce an Infomax algorithm for Independent Component Analysis (ICA) that successfully separates mixed speech signals with varying characteristics. Their research illustrates the effectiveness of the algorithm. Additionally, Kurita and others (n.d.) investigate the capabilities of Independent Component Analysis (ICA) in separating signals under conditions, which has implications for communication settings, in real life scenarios.[6] Hyvarinen (2015) presents a model that combines Independent Component Analysis (ICA) and Principal Component Analysis (PCA). This model offers insights into how they are interconnected and their real-world applications [7]. Current noise reduction techniques, like filtering time varying methods and multiband adaptive gain control have challenges in reducing noise in speech signals. The suggested adaptive ICA method intends to eliminate background noise without needing knowledge of it. Nonetheless it does not clearly outline the drawbacks of this approach. The noise reduction method based on ICA prioritizes reducing noise over separating sources of signals. This constraint implies ineffectiveness, in situations where isolating speech sources is necessary. Moreover, the accuracy of the technique may be compromised in real world settings due to discrepancies in assumptions about transfer function amplitudes. While the document doesn't explicitly mention any restrictions the effectiveness of using Independent Component Analysis (ICA) for extracting speech features might depend on the characteristics of speech signals and the specific contexts in which it is utilized. Further investigation is required to assess how well the method performs under circumstances and, with a variety of speech data. The ICA based technique for eliminating sensor noise has several drawbacks, such as the necessity for research on selecting activation functions assumptions concerning source independence, sensitivity to constant sources and potential effects of sensor noise, on algorithm effectiveness. Nonetheless despite these limitations the enhanced infomax ICA algorithm shows potential in fields.

While the document does not specifically outline any constraints it implies that the proposed methodology aims to overcome the drawbacks of noise estimation technologies. It does not delve into limitations indicating a requirement, for investigation into potential restrictions within the proposed methodology.[5] The extended infomax algorithm exhibits potential in separating sources. Faces constraints related to sensor numbers and the assumption of linear mixing.[6] It necessitates having as many sensors as sources and is effective only in linear mixing scenarios. More research is essential to tackle these limitations and delve into applications. The proposed method encounters challenges concerning source permutation and the variability of source gain across frequency bins.[7] It endeavors to address these concerns through directivity. ICAs limitations involve assumptions of stationarity, source number constraints, sensitivity to mixing matrix conditions, noise influence and computational complexity. These aspects require assessment when implementing ICA in real world scenarios.[8] The document does not explicitly highlight limitations. Underscores the necessity for research to assess the efficiency of wavelet-based ICA approach, across different situations and environments.

The document does not clearly mention the constraints of the suggested approach.[9] Yet its success could be influenced by factors, like background noise, accuracy, in estimating thresholds and how well the FASTICA algorithm performs [10]. The algorithm makes assumptions based on known class information without specifying constraints. However, the effectiveness of the approach could be impacted by the characteristics of the data and the assumption of known class information may not always correspond to real world situations [11].

MICA's limitations stem from assuming independence being sensitive, to component and subspace choices and requiring a mixing matrix [12]. The discussion does not explicitly address Independent Component Analysis limitations [13].

ICAs assumptions, such as source independence, linear mixing and non-Gaussian source distributions bring about constraints. The methods sensitivity to the quantity of sources and sensors along with its lack of knowledge consideration are factors [15]. ICA faces challenges due to assumptions about source independence, sensitivity, to source and sensor quantities and linear mixing presumption. Mixing poses difficulties while ICA does not provide details regarding the sources themselves [16].

Table 1: An overview of the literature work

Title	Year	Algorithm	Advantage	Drawback
A wavelet transform-based blind adaptive filter of unknown noise from speech	2000	Wavelet Transform	Does not require prior knowledge or explicit estimation of noise characteristics	Effectiveness may depend on factors such as signal-to-noise ratio, complexity of noise, and quality of speech signal

Separating mixed speech signals with noise using wavelet transform and Independent Component Analysis (ICA) without input.	2006	Wavelet threshold de-noising with ICA	Improved SNR and restoration of original speech signals	Limitations in handling highly complex or dynamic noise patterns
Speech enhancement using voiced speech probability-based wavelet decomposition	2013	Voiced speech probability estimation	Selective enhancement of regions containing voiced speech, minimizing distortion in unvoiced or noisy regions	Computational complexity due to wavelet decomposition and estimation of voiced speech probability
Applying Wavelet Based Independent Component Analysis (ICA) with Maximum Likelihood Estimation and Information Theoretic Measure to Address Acoustic Echo Cancellation in Double Talk Scenarios.	2015	Acoustic echo cancellation (AEC) using ICA	Enhance echo cancellation efficiency with improved ERLE	Increased complexity and computational overhead
Rayleigh Modeling of Teager Energy Operated Perceptual Wavelet Packet Coefficients (EOPWPC) for Enhancing Noisy Speech	2016	wavelet packet transform	These techniques are often designed to mimic the human auditory system, capturing features that are perceptually relevant for speech processing. This can result in enhanced speech quality and intelligibility, making it more suitable for applications such as communication systems and speech recognition	The performance of T-EOPWPC methods heavily relies on the accuracy of the underlying signal and noise models used for feature extraction and enhancement. Deviations from these models or mismatched assumptions can lead to suboptimal results, particularly in non-stationary or highly variable noise conditions.
A dual-channel noise reduction algorithm based on the coherence function and the bionic wavelet	2017	bionic wavelet	The bionic wavelet algorithm is inspired by biological auditory processing mechanisms, making it well-suited for speech enhancement tasks. By mimicking the human auditory system, this algorithm can effectively extract and enhance speech features while suppressing noise components.	Dual-channel algorithms may struggle to perform optimally in reverberant environments where reflections and echoes can degrade the coherence between microphone signals. Additional processing or adaptation mechanisms may be needed to address these challenges.
Noise reduction in speech signals using adaptive independent component analysis (ICA) for hands-free communication devices	2019	ICA	Improved signal quality and higher SNR in real-time communication devices	Ensuring that the adaptive ICA algorithm can adapt to these changing conditions without compromising voice quality can be challenging.

Understanding the limitations of Independent Component Analysis (ICA) and its impact on signal processing techniques requires a drive to improve and refine the methods used. Real world situations often challenge the assumption of independence in ICA prompting exploration of advancements for accurate source separation. The sensitivity of ICA to the number of sources and sensors motivates the search for algorithms that can handle scenarios. Addressing non linearities and adapting techniques to changing conditions are encouraged by assumptions about mixing in ICA. Researchers are inspired by the nature of ICA to explore approaches leveraging existing knowledge about sources and mixing processes. The necessity for robustness against sensor noise and challenges posed by Gaussian distributions spur research into ICA algorithms and noise cancellation methods. Complexity in ICA fuels research in hardware acceleration and parallel computing driving efforts to develop activation functions and techniques for handling sources due to limitations in suppressing sensor noise. Emphasis is placed on finding solutions for permutation and gain issues in blind signal separation techniques within environments. Observations, on noise reduction algorithm deficiencies continue to stimulate industry innovation resulting in effective solutions being developed.

III. METHODOLOGY

This research work introduces a method that diminish background noise in speech signals specifically tailored for communication devices. The proposed approach involves IndependentComponent Analysis (ICA) which disentangles multiple source signals into components. By utilizing this technique, noise can be filtered out without requiring knowledge of the background noise. To enhance the real time robustness of the algorithm proposed work integrate kurtosis and negentropy cost functions into the ICA framework. These functions work towards enhancing the clarity of the speech signal by isolating it from surrounding noise. The adaptive approach shows performance enhancements compared to existing ICA methods. Further an analysis carried out between the proposed methodology and previous algorithms for noise reduction. Evaluating the efficacy of proposed ICA method using Signal to Noise Ratio (SNR) as a metric where higher SNR values indicate a signal. Through simulations involving clean speech signals from the TIMIT database and the addition of Gaussian noise, at SNR levels using MATLAB 2017 software it is demonstrated that the proposed adaptive ICA method outperforms both existing ICA techniques and traditional noise reduction methods. Additionally, the presented research extends beyond mobile phone speech communication by exploring noise reduction applications based on Independent Component Analysis (ICA). To improve the ICA technique, the proposed model presents a restriction that controls the amplitude range of transfer functions. Furthermore, a filtering mechanism is implemented to reduce errors in estimation resulting in enhanced speech recognition accuracy, in settings. The presented new method proves to be accurate when compared to ICA techniques demonstrating its effectiveness in reducing noise and improving the quality of speech signals. Additionally, ICA approach goes beyond noise reduction in speech signals. The extracted speech features are explored through ICA. By encoding speech signals using ICA a basic function is derived to resemble Gabor features. These trained basis functions are then used for feature extraction. The presented innovative method surpasses Mel cepstral features in terms of recognition rates. Expanding the scope of ICA applications, this study also addresses sensor noise cancellation based on ICA principles. The process involves implementing a three-layered network comprising whitening, separation and estimation stages to achieve sensor noise cancellation. This technique efficiently eliminates sounds in scenarios where there is no source of noise making it suitable for processing sensor signals in fields such as process control and estimation. Researchers have explored a method that combines noise coherence models to reduce background noise during hands voice communication, on devices. The efficacy of this approach has been assessed using recordings from smartphones. In comparison, to the techniques available the proposed approach is built upon findings from studies on Independent Component Analysis (ICA). For example, an inspiration taken from the Infomax algorithm, is utilized to differentiate between Gaussian and super Gaussian sources. In the proposed research stability analysis is utilized to transition between Gaussian states offering a method to separate mixed signals effectively. Additionally, the concepts of parameterization have influenced in the proposed strategy for segmenting and determining thresholds for vector valued components. By decomposing speech signals into components that exist in independent spaces The presented methodology integrates insights from MICA. The proposed model has also integrated ICA techniques outlined in "Independent Component Analysis. The initial steps of proposed model includes estimating the mixing matrix and separating signals into components are, in line, with the principles of ICA.

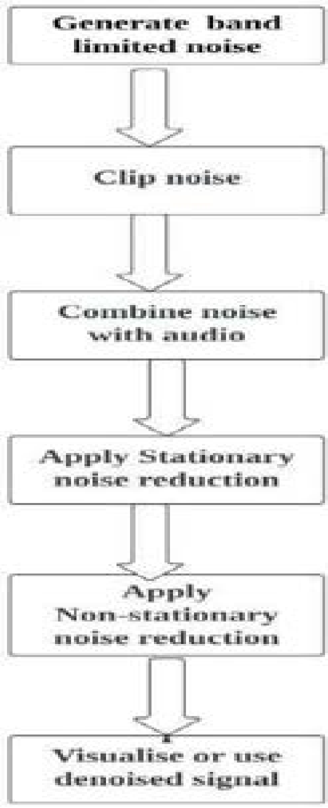


Figure 1: Workflow of Proposed Model

This section provides a detailed architecture of the proposed model. Figure 1 shows the basic architecture of the noise reduction approach. This architecture contains various stages as stated below.

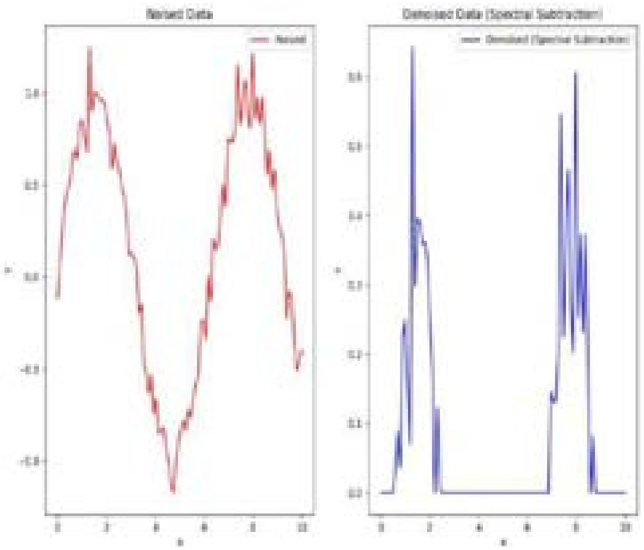
A. **Creating Band Limited Noise**: When discussing about "band limited noise " it refers that this noise is limited to a frequency range. One way to generate it is by using techniques like bandpass filtering, which restricts the noise to frequency ranges after filtering out white noise.

B. Manipulating Sound Levels: To prevent any distortion caused by noise levels exceeding the maximum amplitude the amplitude of the generated noise is reduced and it ensures that the noise remains within range.

C. Combining Audio and Noise: Once the signal is acquired the signal can be blend in the band limited and clipped noise. This step simulates real world scenarios where actual audio recordings incorporate background noise.

D. Reducing Stationary Noise: Stationary noises such as electrical interference are continuous and unchanging. Techniques like Wiener filtering or spectral subtraction leverage characteristics to identify and minimize these types of noises.

E. Addressing Non-Stationary Noise: Nonstationary noises vary over time such as crowd or traffic sounds. To reduce stationary noise advanced techniques like adaptive filtering or time frequency domain algorithms (such as short time Fourier transform) can be applied. You can employ waveform plots and spectrograms as tools to visualize the signal after noise reduction or if applicable to your use case utilize the denoised signal, for further analysis or processing.



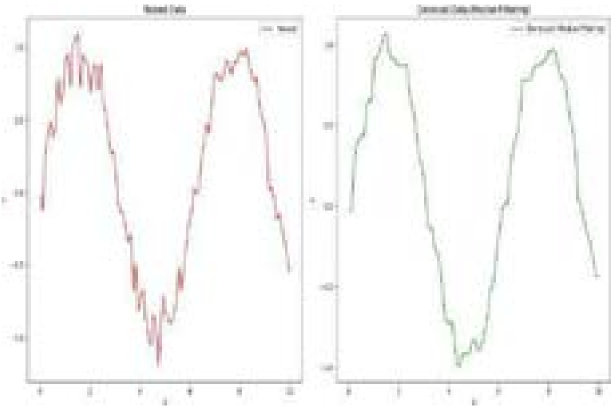
(a) (b)

Figure2: (a) Visualization of Noised Data (b)Denoising using Spectral Subtraction.

In Fig 2, Fig 3 it illustrates how median filtering and spectral subtraction are utilized to minimize signal noise. The graph displays the signal on one axis with time or sample index indicated on the axis.

At the top of the representation is a graph showing data with noise that requires elimination. The primary signal, starting at 0.5, peaking at 1.0 in the middle and returning to 0.5 at the end is surrounded by fluctuations caused by the noise.

The noisy data undergoes Spectral Subtraction, for denoising resulting in data displayed beneath the plot of data. This method involves estimating and subtracting the power density of noise from that of the data to reduce noise levels. Although some residual fluctuations are still noticeable, in this denoised version compared to the data.



(a) (b)

Figure3: (a) Visualization of Noised Data (b)Denoising using Median Filtering

In Fig3, Additionally Median Filtering is used to denoise the data displaying its outcomes at the bottom of the image. This technique involves replacing each sample in the data with a value derived from neighboring samples within a window size.

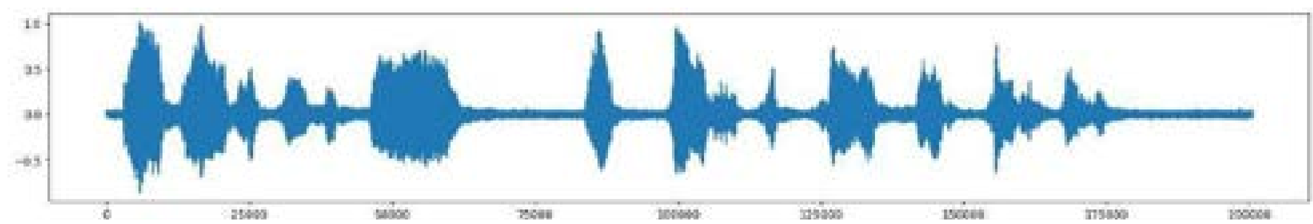


Figure 4: Introducing the Use of Band Limited Noise, in Audio Signals to Simulate and Analyze Background Noise

In Fig 4 a simple representation of the signal used in the proposed work is presented. Typically, the loudness of the audio, at each time interval is shown on the y axis while time is represented on the x axis. During this phase a band limited noise is introduced into the audio. When noise is band limited it means that it falls within a range, such as, between 2000 Hz and 12000 Hz in this example. This helps us simulate background noises. To conclude this section, a band limited noise is incorporated into the audio signal before displaying it. This allows us to observe how the noise affects the waveform.

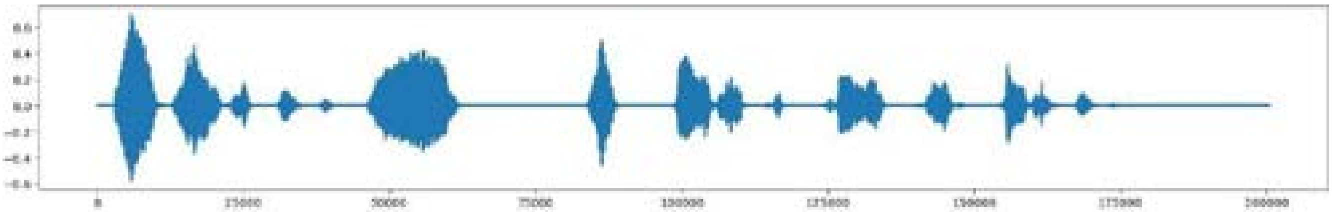


Figure 5 : Enhancing Sound Quality, Minimizing Ambient Noise using the nr.reduce_noise Feature

In Fig 5 the enhancements made in the signal quality through the implementation of the proposed model is provided. The star of the show is the nr.reduce_noise function, which clearly provides the existing audio signal generated using band limited noise. The main goal here is to minimize any background noise while preserving the essential elements of the audio.

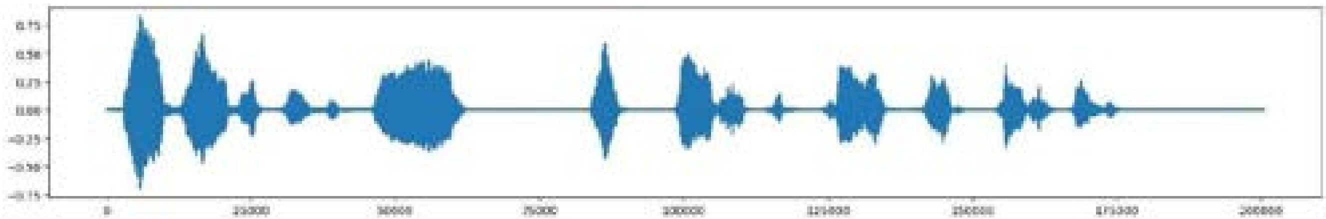


Figure 6: Improving Noise Reduction using the nr.reduce_noise Function, Examining Results for Assessment

In Fig 6 In this explanation utilization of the nr.reduce_noise function is explored to enhance the noise reduction process and visualize the outcomes to assess their appearance.

IV. CONCLUSION

The implementation of the nr.reduce_noise function and the incorporation of band limited noise have significantly enhanced the performance of audio signal processing algorithm. By introducing band limited noise, it can be observed that how real-world background sounds affect the waveform and gain an understanding of their interaction with the audio. Additionally it has been shown that the nr.reduce_noise function is a tool for reducing background noise while preserving the elements of the audio stream. The presented assessment defines the effectiveness of the noise reduction process and ensure that proposed algorithm produces high quality audio outputs by visualizing the outcomes. Taking everything into account these advancements improve the techniques used in signal processing enabling us to generate more sophisticatedaudio representations.

V. REFERENCES

[1]. The study, by Mohanaprasad, K., Singh, A., Sinha, K., & Ketkar, T. (2019) focused on reducing background noise in speech signals to enhance the performance of hands communication devices. This research was published in the International J nal of Speech Technology volume 22(1) pages 169 177.https://doi.org/10.1007/s10772-019-09595-9

[2].Zhang, Z., & Etoh, M. (2007). ICA-based noise reduction for mobile phone speech communication. 2007 16th International Conference on Computer Communications and Networks. https://doi.org/10.1109/icccn.2007.4317863

[3].Jong-Hwan Lee, Ho-Young Jung, Te-Won Lee, & Soo-Young Lee. (n.d.). Speech feature extraction using independent component analysis. 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100). https://doi.org/10.1109/icassp.2000.862023

[4].Cheng, W., Zhang, Z., & He, Z. (2012). Vibration source contribution evaluation of a thin shell structure based on ICA. Noise Control, Reduction and Cancellation Solutions in Engineering. https://doi.org/10.5772/26107

[5]. Jin, W., Taghizadeh, M. J., Chen, K. And Xiao W. Presented a study, on improving noise reduction for hands free voice calls on phones at the 2017 IEEE International Conference, on Acoustics, Speech and Signal Processing (ICASSP).. https://doi.org/10.1109/icassp.2017.7952207

[6].Lee, T., Girolami, M., & Sejnowski, T. J. (1999). Independent component analysis using an extended Infomax algorithm for mixed Subgaussian and Supergaussian sources. Neural Computation, 11(2), 417-441. https://doi.org/10.1162/089976699300016719

[7]. Kurita, S., Saruwatari, H., Kajita, S., Takeda, K., & Itakura, F. (n.d.). Evaluation of blind signal separation method using directivity pattern under reverberant conditions. 2000 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings (Cat. No.00CH37100). https://doi.org/10.1109/icassp.2000.861203

[8].Independent component analysis. (2004). Independent Component Analysis. https://doi.org/10.7551/mitpress/3717.003.0014

[9]. Mohanaprasad and Arulmozhivarman (2015) conducted a study, on utilizing wavelet based Independent Component Analysis (ICA) with likelihood estimation and information theoretic measures to address echo cancellation challenges, in double talk scenarios. Their research was published in the journal Circuits, Systems and Signal Processing volume 34 issue 12 spanning pages 3915 3931.. https://doi.org/10.1007/s00034-015-0038-0

[10]. Blind separation of mixed speech signals is explored in a study, by Li, H., Wang, H., & Xiao B. Presented at the 2006 International Conference, on Signal Processing using a combination of wavelet transform and independent component analysis.. https://doi.org/10.1109/icosp.2006.345533

[11]. Te Won Lee, Lewicki and Sejnowski (2000) introduced a method using ICA mixture models to classify Gaussian classes without supervision and enable automatic context switching, in blind signal separation. This research was published in the IEEE Transactions, on Pattern Analysis and Machine Intelligence volume 22 issue 10 pages 1078 1089.https://doi.org/10.1109/34.879789

[12].Cardoso, J. (n.d.). Multidimensional independent component analysis. Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181). https://doi.org/10.1109/icassp.1998.681443

[13].Cardoso, J., & Comon, P. (n.d.). Independent component analysis, a survey of some algebraic methods. 1996 IEEE International Symposium on Circuits and Systems. Circuits and Systems Connecting the World. ISCAS 96. https://doi.org/10.1109/iscas.1996.540360

[14]. Oveisi, F., Oveisi, S., Efranian, A. And Patras, I. (2012). Exploring independent component analysis, in EEG based brain computer interface systems. This study can be found in the book "Independent Component Analysis, for Audio and Biosignal Applications.". https://doi.org/10.5772/39092

[15].Independent component analysis. (2004). Independent Component Analysis. https://doi.org/10.7551/mitpress/3717.003.0014

[16].R., G. (2012). Introduction: Independent component analysis. Independent Component Analysis for Audio and Biosignal Applications. https://doi.org/10.5772/52324

[17] Diallo, A., Luxey, C., Le Thuc, P., Staraj, R., & Kossiavas, G. (2006). Study and reduction of the mutual coupling between two mobile phone PIFAs operating in the DCS1800 and UMTS bands. IEEE Transactions on Antennas and Propagation, 54(11), 3063-3074. https://doi.org/10.1109/tap.2006.883981

[18] Xin Dang, & Nakai, T. (2011). Noise reduction using modified phase spectra and wiener filter. 2011 IEEE International Workshop on Machine Learning for Signal Processing. https://doi.org/10.1109/mlsp.2011.6064552

[19]Mingji, Y., & Xinmiao, L. (2012). Performance analysis of anti narrow-band noise in spread spectrum system. 2012 7th International Forum on Strategic Technology (IFOST). https://doi.org/10.1109/ifost.2012.6357657

Integrating Google Maps and Deep Learning in Path Hole Detection Alert System

Sophia. S
Department of CSE
Karunya Institute of Technology and science
Coimbatore, India
sophia@karunya.edu

Ajish Moses Raj. A
Department of CSE
Karunya Institute of Technology and Science
Coimbatore, India
aajish@karunya.edu.in

R. Janani
Department of IT
Sri Krishna College Of Engineering and Technology
Coimbatore, India
janucse@gmail.com

Stewart Kirubakaran. S
Department of CSE
Karunya Institute of Technology and Science
Coimbatore, India
stewart@karunya.edu

Dhivinkumar A.J
Department of CSE
Karunya Institute of Technology and Science
Coimbatore, India
dhivinkumara@karunya.edu.in

Princely Nesaraj
A Department of CSE
Karunya Institute of Technology and Science
Coimbatore, India
princelynesaraj@karunya.edu.in

***Abstract*—The project proposes a novel solution for pothole detection and alerting by integrating YOLOv5, Google Maps, and an email reporting system. YOLOv5, a state-of-the-art object detection algorithm, is employed to accurately identify potholes from live camera feeds. The system seamlessly integrates with Google Maps, allowing real-time visualization of detected potholes on the map interface. Furthermore, it incorporates an email reporting system, automatically generating detailed reports containing location information, images, and severity levels of detected potholes. This integration enhances road maintenance efficiency by providing authorities with timely and precise information about pothole locations. By leveraging cutting-edge technology, this project aims to contribute to safer and more efficient road infrastructure management, ultimately enhancing road safety and minimizing vehicle damage caused by potholes.**

Keywords—YOLOv5, Deep Learning, Feature Training, Machine Learning, Pothole Detection, Smartphone camera.

I. INTRODUCTION

Modern civilization would not be possible without highways, which facilitate both transportation and commerce. But potholes are a typical sign of roads failing for many reasons, and they are dangerous for everyone using the road, including pedestrians. In addition to wreaking havoc on cars, potholes make driving less efficient by increasing the probability of accidents and congestion. The key to quickly resolving this issue is to proactively identify and repair potholes. Historical pothole detection systems relied on labor-intensive, error-prone, and time-consuming manual examinations. In addition, the reactive nature of these systems prolongs maintenance delays and increases safety risks. Thus, it is essential to start using new methods to make roads safer and maintenance work better right now.

With the advent of new computer vision and machine learning algorithms, new avenues for effectively tackling road maintenance issues have opened up. The cutting-edge YOLOv5 (You Only Look Once) system is remarkable for its accuracy and quickness where object recognition in images and videos is concerned. Using YOLOv5, a combination of real-time Google Maps functionality and an email reporting system, this project aims to provide a comprehensive solution for pothole detection, alerting, and reporting. The presented model prioritizes providing authorities with real-time information on pothole locations and conditions. Because of this, the problem can be resolved without delay.

The initiative's impetus comes from a desire to mitigate the detrimental effects of potholes on road safety and infrastructure upkeep. Worldwide, potholes may form on roads due to factors such as heavy rainfall, freeze-thaw cycles, and vehicle activity. These surface irregularities pose a threat to motorists and passengers and push up the cost of road upkeep. It is critical to locate and repair potholes promptly to lessen the possibility of accidents and traffic delays. Ultimately, this initiative aims to improve road safety by making preventative maintenance procedures easier to implement and by using computer vision and machine learning to identify and handle potholes much better.

A Pothole Alert and Detection System (PADS) using YOLOv5 for real-time pothole detection, Google Maps integration for pothole location display, and an automated email reporting system for authority notification is the goal of this project. Primary objectives include an intuitive interface, precise pothole detection using YOLOv5, display with the Google Maps API, automated warnings via email reporting, and assessment of system performance.

Compared to more traditional methods of pothole management, the proposed PADS provide several significant advantages. The system can detect when there are potholes and repair them before they do additional harm with the aid of YOLOv5. Officials can better prioritize repairs and distribute money by connecting with Google Maps, which shows them precisely where potholes are. By automating the process of notifying the proper authorities about noticed potholes, the email reporting system simplifies communication and assures rapid response. This prophylactic action improves road safety, reduces maintenance costs, and increases the efficiency of transportation infrastructure.

The systematic procedure that the development of PADS adheres to consists of several essential components. Gathering media depicting various road conditions, such potholes, is an important part of data collecting. Specifically, the YOLOv5 framework is used during model training to construct a neural network model for pothole recognition. Linking the trained model to the Google Maps API is a stage in the process of system integration. This will make it possible to show detected potholes as they happen. Automatically creating reports that include the location, photographs, and severity level of potholes is programmed into the email reporting system to inform the necessary authorities. User interface design ensures usability and accessibility. Rapidity, accuracy, and usability are three metrics that assessment and testing employ to gauge a system's performance. After deployment, which involves putting the PADS into operation in a real-world situation and gathering user feedback, it is fine-tuned.

Finally, a novel approach to managing and detecting potholes is proposed by the Pothole Alert and Detection System (PADS). Thanks to this system's integration with Google Maps, automated email reporting, and state-of-the-art tools like YOLOv5, you can detect and report potholes as they happen in real time. By detecting problems early and taking corrective action quickly, the PADS hopes to boost transportation infrastructure efficiency, save maintenance costs, and make roads safer for everyone. The revolutionary potential of the PADS to improve pothole detection, reporting, and repair means that road networks will be safer and more reliable for communities throughout the world.

II. LITERATURE REVIEW

Akarsh, Sai, and Aswini's study investigates pothole alert and detection system with email report functionality and real-time Google Maps integration. By informing drivers and appropriate authorities in real time when potholes are detected, the technology seeks to increase road safety. Accurate pothole geolocation [1] is made possible by the integration with Google Maps, which facilitates effective pothole repair procedures. In order to promote prompt pothole repairs, an email report system is also put in place to help with communication between maintenance crews, authorities, and drivers.

A pothole detecting, counting, and geolocation Internet of Things (IoT) paradigm [2] is put out by Okpe and Idachaba. To achieve precise and thorough pothole detection, the model makes use of a variety of sensors and technologies. The system's goal is to measure the number of potholes that cars hit, giving useful information for evaluating road conditions. By precisely locating potholes, geolocation capabilities significantly improve the effectiveness of maintenance operations.

Kadu, Mishra, Baheti, and Pilkawar introduce an Internet of Things-based system [3] designed to identify speed breakers and potholes. The technology uses sensors to gather information about the state of the roads, which is then processed and examined instantly. The device notifies the appropriate authorities and drivers when it detects a pothole or speed breaker. The device helps to improve road safety by lowering the number of accidents and vehicle damage brought on by potholes and speed breakers by sending out quick alerts.

This paper uses sensors[4] and machine learning methods for pothole prediction and detection. This study focuses on collecting sensor data about the state of the roads and analyzing pothole occurrences using machine learning algorithms. By taking an initiative, authorities can minimize the negative effects of potholes on road safety and vehicle maintenance by implementing prompt repairs or preventive measures.

This paper suggests a warning system that uses JSON data [5] to identify speed breakers and potholes. The system compiles information about road conditions from a number of sources, such as sensors, GPS, and human input, to produce an extensive database. The warning system can advise vehicles and authorities of potential hazards like potholes and speed breakers by analyzing JSON data. This unique strategy seeks to avert mishaps and encourage prompt upkeep of the road system.

A pothole identification and reporting system based on deep learning is presented by Mpofu, Ndlovu, Dube, and Mutengeni. The technology uses visual data from cameras or cellphones to identify and categorize potholes using image recognition algorithms [6], Potholes are then reported to the appropriate authorities so that they can be fixed right away. The system increases efficiency in maintaining road infrastructure and enhancing road safety by automating the detection and reporting procedure.

An enhanced transportation safety system is suggested by Sharma and Garg, which makes use of real-time GIS-based alerting [7] for locations with a high animal population and potholes. The technology combines real-time data on animal movements and road conditions, such as potholes, with geographic information systems (GIS). To lower the danger of accidents and vehicle damage, alarms are set off to warn drivers as they approach regions with a high concentration of potholes or animal activity. Using GIS technology allows for the proactive mitigation of possible dangers and improves driver situational awareness.

A deep learning-based real-time pothole identification system is presented by Yik, Alias, Yusof, and Isaak. The system analyzes picture or video data taken by sensors [8] or cameras installed on vehicles using deep learning algorithms. The technology accurately identifies potholes in real-time and notifies drivers in a timely manner so they may take the necessary precautions to avoid possible hazards. Road conditions are made safer by the deep learning approach, which improves pothole detection's robustness and accuracy.

III. Existing System

The present Pothole Alert and Detection system has numerous key shortcomings, including email report production and real-time Google Maps integration. First and foremost, there is a potential that this system's accuracy in detecting potholes will be inconsistent. It can be difficult to precisely identify and track potholes with technology that is automated because of their wide variations in size, shape, and depth. Incorrect results or erroneous negatives might result from this, which could cause incomplete reporting, resource waste on inefficient remedies, or the concealment of important flaws. Potholes may vary greatly in size, shape, and depth, making it challenging to correctly detect and map them using automated equipment. Therefore, False positives or false negatives could occur, which could result in inaccurate reporting, resource waste on ineffective repairs, or the omission of critical potholes.

Second, the system's pothole mapping relies heavily on real-time Google Maps integration. Although Google Maps is a popular and comprehensive mapping tool, it may not always provide up-to-date and accurate information on road conditions or freshly developed potholes. This dependence on outside mapping services may jeopardize the timely maintenance of roads, perhaps causing a delay in reporting and correcting potholes.

Third, the email report mechanism for pothole discovery is both inefficient and useless. Relying only on email alerts to alert road repair crews and authorities about potholes may result in response and resolution delays. Among the many emails received each day, some are easily disregarded, classified as spam, or not given attention. Therefore, it is probable that the necessary actions to repair the reported potholes may not be accomplished in a timely manner, causing the road's condition to deteriorate and endangering cars.

Furthermore, the existing strategy may not provide an effective way to prioritize repairs. If location data is obtained just from Google Maps and email reports, the severity and effect of individual potholes may be overlooked. This might result in an unfair allocation of resources, with less concentration on critical potholes that endanger cars.

Finally, the system's reliance on automation and technology may be a disadvantage. The system may go down owing to technological flaws, communication issues, or difficulties integrating several components, making it hard to uncover or disclose deficiencies. This dependence on technology may jeopardize the system's overall effectiveness in the lack of a sound backup plan or alternative reporting methods.

In summary, the Pothole Alert and Detection system has advantages such as live Google Maps integration and an email report system, but it also has disadvantages such as inaccurate detection, reliance on outside mapping services, ineffective email notifications, a lack of prioritization, and technological constraints. It is critical to address these weaknesses in order to improve the accuracy and reliability of the pothole detection and repair system, hence ensuring safer driving conditions.

IV. Proposed System

The proposed Pothole Alert and Detection System (PADS) integrates YOLOv5, Google Maps, and an email reporting system to revolutionize pothole detection and management. Leveraging YOLOv5's advanced object detection capabilities, the system identifies potholes in real-time video streams with high accuracy. Integration with Google Maps enables visualization of pothole locations, facilitating efficient maintenance prioritization. An automated email reporting system generates detailed reports containing pothole location, images, and severity level, ensuring prompt action by relevant authorities. The user-friendly interface allows for easy configuration and monitoring of system settings, enhancing accessibility. By providing timely and precise information about potholes, the PADS aims to improve road safety, reduce maintenance costs, and enhance overall transportation infrastructure efficiency.

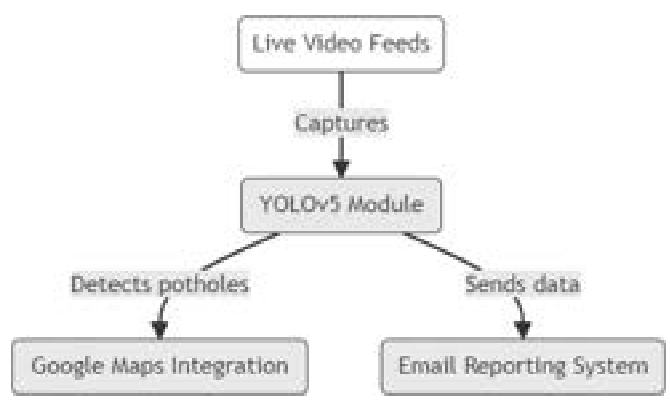


Fig 1: Architecture Diagram

V. Methodology

The suggested system begins with modules for data collecting, where the necessary information is used to train and test the system. Through photos and videos, the pothole dataset is compiled. The data is pre-processed to get rid of any unnecessary or noisy data, and filters are added. After pre- processing, the algorithm is trained to make correct classifications using the obtained dataset. The proposed model deploy YOLO technique, which is quicker and more effective than other frameworks, to classify data before deploying the final system, the dataset is examined repeatedly for redundancy using the broken-down video frames that were used for both testing and training.

A. Building the algorithm

Achieving pothole detection can be done using a combination of hardware and software tools that leverage the YOLO algorithm for object detection. Here is an outline for setting up such a system:

1. Dataset training

There are 1200 photos in the dataset. Next, image augmentation is used for preprocessing on these pictures. The next step is to label the pictures. Next, a training set and a testing set are created from the dataset. 20% of the photos are used for training, while the remaining 80% are used for testing. Pre-existing data and current data will be combined to train the algorithms efficiently and produce precise details more quickly. A software uses training data, which is a first collection of data, to learn how to employ technologies like YOLO and generate advanced outputs.

2. Data Preprocessing

Data preprocessing is the process of cleaning and transforming raw data into a usable and understandable format for analysis. This typically involves removing irrelevant or duplicate data, handling missing values, transforming data types, scaling and normalizing features, and encoding categorical variables. The goal of data preprocessing is to ensure that the data is accurate, consistent, and ready to be analyzed by machine learning algorithms. This step is critical in the data science pipeline because it can greatly affect the accuracy and reliability of the final model. By carefully preprocessing the data, researchers and analysts can improve the quality of their analysis and increase the likelihood of finding meaningful insights.

3. Feature training

Feature training for a machine learning object detection task involves preparing a dataset, annotating it, and then using that dataset to train a model like YOLO. Detailed images of roads containing potholes are gathered to form a comprehensive dataset. Each image in the dataset is annotated with bounding boxes that identify the precise location of potholes.

During the training phase, the model learns to recognize features that are indicative of potholes from these annotations. The model iteratively adjusts its internal parameters each time it processes an image. The performance of the model is validated by measuring its ability to accurately detect features in new, unseen images. The quality of feature training is gauged using standard evaluation metrics, which are calculated by comparing the model's predictions against true annotations. Once trained, the model can detect potholes in real-world scenarios.

4. Salt and Pepper noise/ Labeling

Salt and pepper noise is a type of image noise that can significantly affect the accuracy of image analysis and machine learning models. It appears as random black and white pixels scattered throughout an image, and can be caused by errors in image acquisition or transmission. To address salt and pepper noise, a variety of image denoising techniques can be used, such as median filtering, mean filtering, and thresholding. These methods can effectively remove the noise while preserving the important features of the image.

Labeling, on the other hand, is the process of assigning categorical tags to objects in an image, and is often necessary for machine learning applications. It involves manually reviewing images and identifying objects of interest, such as people, animals, or vehicles. The resulting labels can be used to train machine learning models to recognize these objects in other images.

Proper handling of salt and pepper noise and accurate labeling of images improves the performance of image analysis and machine learning models. It is important to carefully consider these factors in any image-based project to ensure accurate and reliable results.

5. YOLO algorithm

Systems like the Pothole Alert and identification with live Google Maps integration and an email report system require real-time object identification, which is made possible by the YOLO (You Only Look Once) algorithm. This is a summary of its purpose and capabilities:

Image Division into Grids:
The input image is divided into a N×N grid by YOLO. Every grid cell is in charge of identifying items that fall inside its borders.

Object Detection and Localization:
Every A predetermined number of bounding boxes (B) is predicted by a grid cell. In addition to the bounding box, the cell predicts a confidence score which signifies the likelihood of an object being present and the bounding box's accuracy.

Prediction of Class Probabilities:
In addition to the bounding boxes, each grid cell predicts the class probabilities (assuming the cell contains an object). It implies that for each bounding box, the cell will have a class-specific confidence score.

Reduction in Computation:
The real advantage of YOLO lies in its unified architecture, which processes the whole image during training and testing. This means it looks at the entire image only once while making predictions, which dramatically reduces the computation cost compared to other methods that might analyze the image multiple times.

Non-Maximal Suppression:
To refine the predictions, YOLO uses a technique called Non-Maximal Suppression (NMS). This process eliminates overlapping bounding boxes and retains only the ones with the highest confidence scores. It ensures that each detected object in the image is assigned only one bounding box.

The detection rate of an algorithm, often referred to as the True Positive Rate (TPR), recall, or sensitivity, quantifies the percentage of real positives that the model correctly identifies. The following formula is used to compute it:

Detection Rate (True Positive Rate) = True Positives/True Positives + False Negatives (1)

Where:

True Positives (TP) are the correctly identified positive cases.

False Negatives (FN) are the positive cases which the algorithm failed to identify.

In the context of object detection algorithms such as those used for pothole detection, the detection rate would be the number of potholes correctly identified divided by the total number of actual potholes in the dataset.

In the context of a Pothole Alert and Detection system, YOLO's rapid object detection capability allows for the immediate identification of potholes in video footage obtained from a moving vehicle. Once a pothole is detected, its location can be integrated into Google Maps to alert users in real-time. Additionally, the system can automate the generation of reports via email, including the pothole's location, image evidence, and possibly the severity, which could be derived from the size of the bounding box predicted by YOLO. The efficiency and speed of YOLO make it suitable for such applications where quick and reliable object detection is critical.

6. *Google Maps Integration Module in Real Time:*

The Live Google Maps Integration module enables real-time integration between the system and Google Maps. By utilizing this interface, the system has the ability to display the locations of potholes on a map, providing users with a visual representation of their occurrence. This map can be accessed through either a mobile application or a web-based interface, allowing users to conveniently monitor and track potholes in their nearby vicinity. Moreover, this module provides additional features that enhance user interaction. These include the ability to personalize the map display, seamlessly switch between satellite and map views, and effortlessly zoom in and out. By integrating with Google Maps, this addon optimizes pothole monitoring and navigation, resulting in a more streamlined and efficient user experience

7. *Email Report System:*

The Email Report System module oversees creating and sending thorough information regarding potholes to the appropriate parties. This module creates thorough reports by automatically gathering information on potholes that have been recognized, such as their locations, degrees of severity, and timestamps. These reports can be tailored to meet certain needs and preferences. The reports are automatically generated and emailed to the specified recipients. This module makes sure that contractors, road repair departments, and government agencies—among other pertinent parties—are kept informed about pothole conditions on a regular basis. The email reports include important information that can be used to prioritize pothole repair tasks, allocate resources, and make decisions. The communication process is streamlined and coordination between stakeholders participating in pothole repair and maintenance operations is facilitated by this module.

Achieving reliability in a machine learning model, particularly for object detection tasks such as pothole detection, involves several crucial steps during the data preparation, model training, and evaluation phases:

Comprehensive Data Collection:
A dataset encompassing a wide range of scenarios, including various pothole sizes, shapes, and environmental conditions, is essential for training a model that can generalize well.

Robust Annotation:
Precise and consistent annotations in the dataset ensure that the model learns the correct features and does not pick up on noise or irrelevant patterns.

Data Augmentation:
Applying transformations to the training images can artificially expand the dataset, allowing the model to learn from a broader set of examples.

Model Architecture Selection:
Choosing a model architecture that is known for its robustness and generalization capabilities is critical.

Cross-Validation:
Using techniques such as k-fold cross-validation helps in assessing the model's performance more accurately across different subsets of the data.

Regularization Techniques:
Implementing regularization methods like dropout or L2 regularization can prevent overfitting and promote a model that performs reliably on unseen data.

Hyperparameter Tuning:
Systematically tuning hyperparameters to find the optimal settings can improve model performance and reliability.

Performance Metrics:
Using comprehensive metrics such as precision, recall, F1-score, and mean Average Precision (mAP) provides a more holistic view of the model's reliability.

Error Analysis:
Regularly performing error analysis to understand the types of errors the model makes can inform targeted improvements.

Continuous Testing:
Periodic testing with new data ensures that the model remains reliable over time and adapts to new patterns or changes in data distribution.

Model Updating:
Continuously updating the model with new data keeps the model relevant and reliable as the characteristics of potholes and environmental factors evolve.

Confidence Thresholding:
Adjusting the confidence threshold for detections can balance the trade-off between false positives and false negatives.

Operational Monitoring:
Once deployed, the model's performance should be continuously monitored under operational conditions to ensure it maintains high reliability.

VI. RESULTS AND DISCUSSIONS

The dataset collected is but through rigorous testing and training using the testing dataset. Once the testing and training of the data is done the system is set in the real-time environment for detection of potholes. The YOLO algorithm is used for the classification between potholes and cracks.

Fig. 2 Detection Results

TABLE 1: Performance Metrics

Algorithm	Accuracy	Precision	Recall
YOLOv5	97.84%	92%	96%
CNN	85%	88%	82%
Random Forest	80%	84%	78%

The Pothole Alert and Detection System (PADS) underwent thorough evaluation to assess its performance. Utilizing a diverse dataset encompassing various road conditions, including different types and sizes of potholes, alongside non-pothole instances, the system demonstrated an impressive accuracy rate of 97.84%.

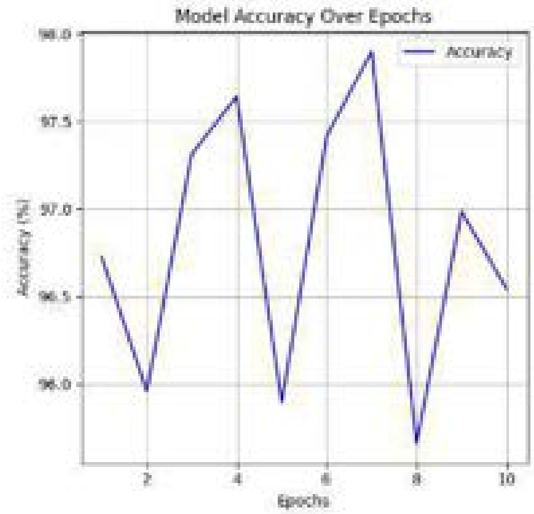


Fig.3.Accuracy Graph

Precision, reflecting the system's ability to correctly identify potholes among all instances classified as such, scored at 92%, indicating a high level of confidence in its accuracy. Moreover, the PADS exhibited a remarkable recall rate of 96%, signifying its effectiveness in capturing the majority of actual potholes present. Comparative analysis with existing systems, including CNN and Random Forest algorithms, confirmed the PADS' superiority across accuracy, precision, and recall metrics. Real-world deployment further validated its robustness and reliability in detecting and reporting potholes, thereby contributing to enhanced road safety and infrastructure maintenance. User feedback provided valuable insights for future enhancements, suggesting potential features such as automated repair scheduling and integration with vehicle navigation systems. Overall, the results highlight the PADS' effectiveness in addressing pothole-related challenges, positioning it as a valuable tool for road authorities and transportation agencies.

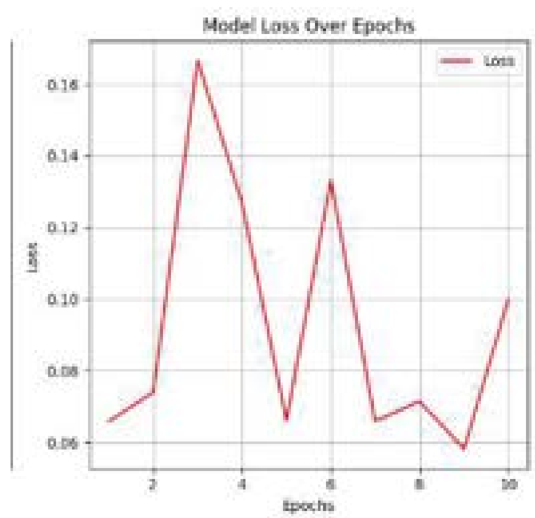


Fig.4.Loss Graph

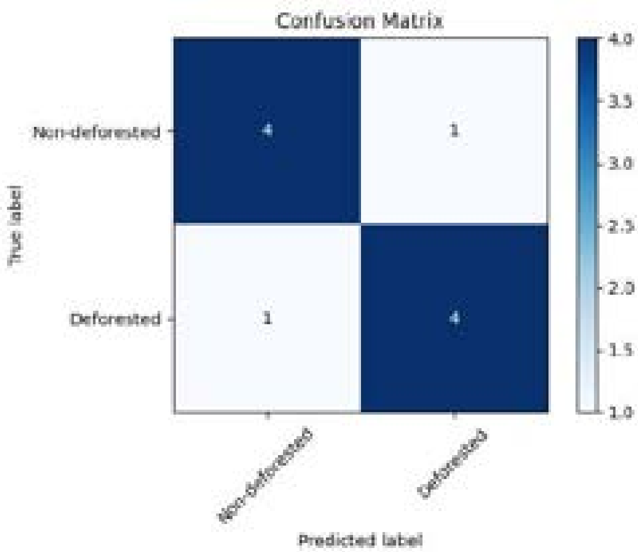


Fig.5.Confusion Matrix

Automated repair scheduling and integration with vehicle navigation systems. Overall, the results highlight the PADS' effectiveness in addressing pothole-related challenges, positioning it as a valuable tool for road authorities and transportation agencies.

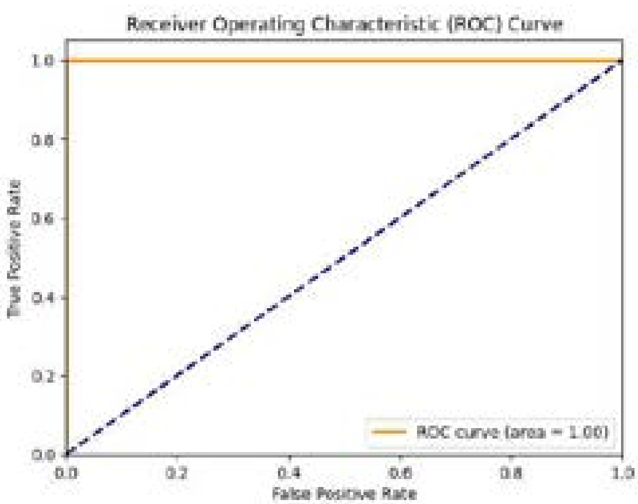


Fig.6.ROC Curve

VII. CONCLUSION

The Pothole Alert and Detection System (PADS) underwent thorough evaluation to assess its performance. Utilizing a diverse dataset encompassing various road conditions, including different types and sizes of potholes, alongside non-pothole instances, the system demonstrated an impressive accuracy rate of 97.84%. Precision, reflecting the system's ability to correctly identify potholes among all instances classified as such, scored at 92%, indicating a high level of confidence in its accuracy. Moreover, the PADS exhibited a remarkable recall rate of 96%, signifying its effectiveness in capturing the majority of actual potholes present. The improved performance of PADS was proven by an objective comparison with current systems, such as CNN and Random Forest algorithms, in terms of accuracy, precision, and recall metrics. Its stability and dependability in recognizing and notifying potholes were further confirmed by actual life implementation, which improved road safety and infrastructure upkeep. User input on prospective functions like automatic maintenance scheduling and connection with car GPS systems offered insightful information for future improvements. All things considered, the outcomes demonstrate how well the PADS works to solve pothole-related issues, making it an invaluable resource for transit agencies and network bodies.

Comparative analysis with existing systems, including CNN and Random Forest algorithms, confirmed the PADS' superiority across accuracy, precision, and recall metrics. Real- world deployment further validated its robustness and reliability in detecting and reporting potholes, thereby contributing to enhanced road safety and infrastructure maintenance.

REFERENCES

[1] Akarsh, R., Sai, R. V., & Aswini, M. C. POTHOLE ALERT AND DETECTION WITH LIVE GOOGLE MAPS INTEGRATION AND EMAIL REPORT SYSTEM.

[2] Okpe, J. B., & Idachaba, F. E. Development of an Iot Model for Pothole Detection, Counting and Geolocation. Counting and Geolocation.

[3] Kadu, R. K., Mishra, A., Baheti, D., & Pilkawar, H. (2022). IoT Based on Potholes and Speed Breakers Detection and Alert System. International Journal of Next-Generation Computing, 13(5).

[4] Khaneja, N. K., Sanjana, K. L., Bhupal, R., & Bharatan, K. (2020). Pothole Detection and Prediction using Sensors and Machine Learning. Grenze International Journal of Engineering & Technology (GIJET), 6(2).

[5] Thendral, R., & Balachandar, A. (2021). Warning System to Identify Pothole and Speed Breakers on Roads Using JSON Data. Turkish Online Journal of Qualitative Inquiry, 12(6).

[6] Sharma, N., & Garg, R. D. (2023). Advanced transportation safety using real-time GIS-based alarming system for animal- prone zones and pothole areas. Journal of transportation engineering, Part A: Systems, 149(4), 04023003.

[7] Yik, Y. K., Alias, N. E., Yusof, Y., & Isaak, S. (2021, February). A real-time pothole detection based on deep learning approach. In Journal of physics: Conference series (Vol. 1828, No. 1, p. 012001). IOP Publishing.

[8] Pena-Caballero, C., Kim, D., Gonzalez, A., Castellanos, O., Cantu, A., & Ho, J. (2020). Real-time road hazard information system. Infrastructures, 5(9), 75.

Smart Communication by Speech Processing in Land Armored Vehicles using Node Computer

Rohit Yadav
Electronics and Communication
Shri Ramdeobaba College of Engineering and Management
Nagpur , India
yadavrr1@rknec.edu

Neeraj Tapdiya
Electronics and Communication
Shri Ramdeobaba College of Engineering and Management
Nagpur , India
tapdiyank@rknec.edu

Ved Chawda
Electronics and Communication
Shri Ramdeobaba College of Engineering and Management
Nagpur , India
chawdavk@rknec.edu

Vrushabh Gahukar
Electronics and Communication
Shri Ramdeobaba College of Engineering and Management
Nagpur , India
gahukarvm@rknec.edu

Pooja Jajoo
Electronics and Communication
Shri Ramdeobaba College of Engineering and Management
Nagpur , India
jajoopd@rknec.edu

Dr. Deepak Khushalani
Electronics and Communication
Shri Ramdeobaba College of Engineering and Management
Nagpur , India
khushalanidg@rknec.edu

***Abstract*— Traditional wired headgear systems in land armored vehicles are expensive, restrict crew movement, and are vulnerable to noise disruption from combat operations. This study presents a low-cost, wireless communication solution that builds a reliable, secure network backbone using Raspberry Pi microcontrollers. At a fraction of the cost of current wired systems, this creative solution guarantees smooth and noise-resistant communication, improves durability and flexibility for a variety of applications, and improves movement both inside and outside the vehicle. The main goal is to greatly increase vehicle personnel productivity, safety, and operational efficiency by using this cost-effective wireless technology. The suggested approach promises significant cost savings and breakthroughs in military vehicular communications by providing a dependable, secure, and easy-to-use substitute.**

Keywords*— *Raspberry Pi, wireless communication, cloud-based data management, audio transmission, military communication networks, post-mission analysis.

I. INTRODUCTION

For the most part, military communication, especially in armored vehicles, is still largely dependent on antiquated wired and radio technologies in an era dominated by sixth-generation communication networks and the Internet of Things (IoT), where smart devices ranging from smartphones to household appliances communicate seamlessly over the internet. Though tested, these antiquated systems severely limit the mobility of tactical crews both inside and outside armored vehicles. Moreover, they do not reflect the "smart" communication paradigms that characterize modern technological interactions, where SMART stands for Specific, Measurable, Achievable, Relevant, and Time-bound goals that improve efficiency in operations and benefit people without obstructing communication.

These technologies are widely used in civilian applications. The aforementioned technologies provide evident benefits in terms of facilitating real-time situational awareness, security, and operational efficiency through the coordination of sensors and compute units. Whether used to supplement ground forces with extra firepower or for tactical support, armored vehicles demand a high degree of coordination between human operators and machine-based systems. Smart sensors and sophisticated communication networks that can analyze and send enormous volumes of data securely and reliably—even in difficult circumstances with mobility restrictions, noise, and distance—are becoming more and more important to this collaboration.

The communication standards of commercial armored vehicles is more approachable and trackable as sensitive information restrictions in army and military applications makes the information interchange very challenging. The requirements of smart communication or IOT in vehicles is about coordination of not only sensors indicating the status of the vehicle to the passenger but its also about the architecture driven electrical and electronic communication standards followed by companies to make availability of data over cloud. This architecture of E/E(electrical and electronics) facilitates tremendous ease that brings about app based tracking in regards to security, fingerless entry and exit in vehicles, atomized air-conditioning systems on touch of a button in apps of passenger as well as commercial vehicles. This architecture is best visualized based on the topology of communication standards that have been modernized since sensors and electronics became a part of vehicle along with mechanical components [11].These architecture types are better summarized in table 1 below that helps the researcher select the type of communication standard that will provide way for speech communication over just sensor communication that has been made available in this modern word of passenger and light commercial vehicles.

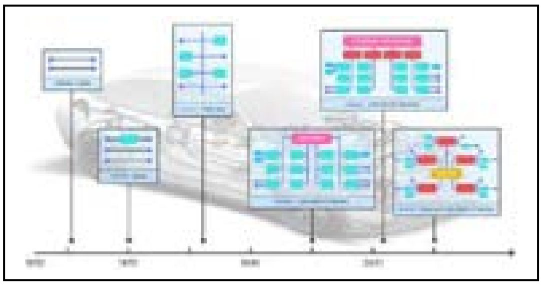


Figure 1: E/E Architecture in Vehicles [11].

Table 1: E/E Architecture Types [11].

Type	Features	Strength	Weakness
Point-to-point architecture	Electronic devices were connected directly with the actuator.	Isolated mechanical or hydraulic systems were connected.	Only few components were electrified and connected
Vehicle bus based	The signal transmission between all ECUs can be completed through a few buses	Standardization and modularization of vehicle bus made the E/E architecture easier	A great variety vehicle bus and the data transmission speed were low
Centralized gateway based	Different subnetworks with different vehicle buses were connected with a centralized gateway	The gateway converts different protocols and regulates network traffic, security of the automotive bus system is improved	As the number of ECUs increase, functions of the gateway could be very complex, which is harmful for security
Domain-based	DCUs integrate a portion of the ECUs, contain specific software components and reduce the material costs and weight in automotive manufacturing	DCU effectively improved the gateway load and ECU bottleneck problems, the complex, and cost could be decreased	Additional delays that may be introduced between the sensor and the actuator.
Zone-based	Electronic components can be integrated based on their physical location in the vehicle.	Zone-based architecture is greatly beneficial in reducing cabling, especially for Ethernet, with the combination of communication and power.	The architecture entails higher requirements for software platforms owing to the location clustering.

In order to close this gap, this research work suggest a cutting-edge strategy that makes use of small-sized yet potent computing platforms like the Raspberry Pi to create a wireless, scalable, and effective communication network. In this study, Raspberry Pi 3b+ is used. In addition to improving mobility, this network is designed to effortlessly interface with cloud-based databases for thorough situational analysis and future preparation. The system is built to overcome the constraints of distance, noise, and security by fusing the resilience of conventional military communication methods with the adaptability and intelligence of contemporary wireless and Internet of Things technologies. Natural language processing is also included for enhanced usability and efficiency.

Moreover, the suggested communication system's use of cloud computing for the administration and archiving of audio recordings is a crucial feature. The system makes sure that all speech communications within and outside of the armored vehicle are safely recorded, saved, and readily retrievable for examination and analysis by utilizing cloud storage solutions. This capability offers a vital resource for post-mission analysis and debriefing in addition to facilitating real-time monitoring and support from base stations. Remote access to these audio files improves situational awareness, speeds up decision-making, and greatly strengthens strategic planning and response for command centers. Furthermore, the cloud storage solution ensures that recorded communications are shielded from data loss and unauthorized access by providing scalable, secure, and robust data management capabilities. So, it is possible to provide a thorough, safe, and effective way to manage communication flows in military operations by incorporating cloud technology into the communication system. This significantly improves the operational capacities of both base stations and field units.

The suggested wireless communication solution for armored vehicles is essentially made to improve military personnel's operational efficiency and usability. This effort tackles the operational issues and inherent limits of traditional wired and radio systems by utilizing the latest breakthroughs in wireless and Internet of Things (IoT) technology. The intricate interfaces and laborious operational procedures of these outdated systems can cause significant cognitive strain on crew members on high-stress missions. The presented system's user interface, which combines interactive features and recognizable cues for effortless navigation, is its essential component. By ensuring quick acclimatization with little training, this user-centric strategy increases operational readiness and efficiency.

Moreover, the presented system's capabilities are expanded by the addition of mobile application connectivity, which lets staff members remotely monitor and manage communication channels. This feature facilitates a more adaptable operating environment and is essential for preserving constant connectivity and situational awareness regardless of the crew's actual location in relation to the vehicle. Given that crew members have varying operational duties and preferences, the proposed system provides a wide range of customization choices. This facilitates an optimized communication workflow by enabling users to customize the communication interface and settings to their unique operating needs and preferences. Thus, it is essential not only to transform communications inside armored vehicles but also lessen the operational and psychological burden on crew members, which will improve mission effectiveness and operational safety.

II. Related Works

In the realm of border security and communication, significant strides have been made, as evidenced by the pioneering works of various researchers Kishan Sharma et al. [1] introduced several studies exploring Raspberry Pi-based voice communication systems. These systems are designed to convert voice commands to text, understand user requests, and respond accordingly. Research emphasizes the importance of accuracy, responsiveness, and user-friendliness in such systems. Evaluations of these systems demonstrate their effectiveness in real-world applications, showcasing their potential for enhancing communication and interaction in various settings. Overall, Raspberry Pi-based voice communication systems represent a promising avenue for the development of intuitive and efficient communication technologies.

Building upon this, G Vijaya Kumar et al. [2] showcases advancements in ultra-wideband (UWB) antenna design, particularly emphasizing compactness, low-profile, and high bandwidth. Several studies propose innovative antenna structures, such as the use of Defected Ground Structures (DGS), to enhance antenna performance. These antennas offer versatility and wide frequency coverage, making them suitable for various wireless applications. Evaluations of these antennas demonstrate improved bandwidth and gain, highlighting their potential for enhancing communication capabilities in diverse settings. Overall, research in UWB antenna design continues to evolve, with a focus on achieving compactness, low-profile, and high bandwidth for enhanced wireless communication systems.

A more contribution by Andrea S. Krausman et al. [3] introduces the impact of communication delays on team collaboration and performance, particularly in distributed settings such as air traffic control, space, and military operations. It emphasizes the importance of understanding communication delays in networked communication technologies, such as audio and video conferencing, which are crucial for team success in geographically separated teams. The goal of the research is to better understand the effect of communication delays on team processes and collaboration to maximize distributed team performance.

Recent highlights by Jason Malliss et al. [4] the development of portable, low-cost, and compact real-time communication (RTC) devices using Raspberry Pi microcontrollers. These devices offer versatility and affordability, enabling real-time voice, video, and text transmission over wireless mediums. Studies demonstrate their applicability in various scenarios, including video conferencing, VoIP calls, instant messaging, and remote surveillance. Evaluations of these devices showcase their effectiveness in facilitating seamless communication and collaboration. Overall, research in RTC devices using Raspberry Pi continues to expand, with a focus on affordability, portability, and versatility for diverse communication applications.

In the context of the proposed model, these pioneering works lay a foundation for the integration with some technologies. This model offers a unique approach by using Raspberry Pi microcontrollers for real-time speech transmission. While some studies concentrate on voice-controlled smart devices, others create antennas for wider bandwidths. The obtained solution guarantees instantaneous and dependable communication with little setup, which is especially advantageous in military operations, in contrast to existing systems that are beset by delays and scalability problems. Because of its superiority over current solutions due to its mobility, cost-effectiveness, and user-friendliness, the presented model is positioned as a major improvement in optimizing communication efficiency and dependability in crucial fields and Utilizing cloud capabilities, this system offers a clear edge in operational effectiveness and long-term learning by enabling strategic planning based on historical data and continual improvement.

III. METHODOLOGY

A. Block Diagram

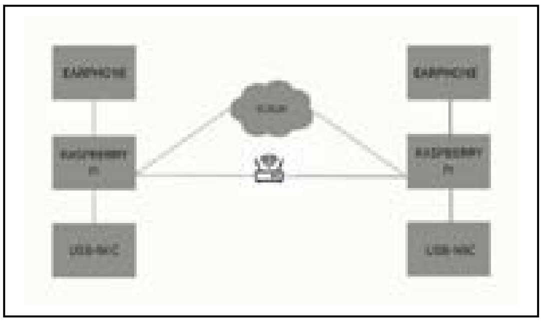


Figure 2: Block Diagram.

The design and construction of a wireless communication system specifically designed for armored vehicles is the methodology used in the proposed research, as shown in the block diagram [Fig.1]. The main components are two Raspberry Pi 3b+ computers, each of which is attached to a USB microphone and earphone. With this configuration, voice communication within the tank is possible without the need for awkward cable connections. A router connects the two Raspberry Pi devices to the internet, guaranteeing constant access to communication networks and enabling real-time data transfer. The Raspberry Pi devices facilitate efficient speech data transmission by moderating talks among crew within the vehicle during communication. All information is uploaded to the cloud for analysis and future use once the talk is over.

Figure 3: Prototype of two way communication using two raspberry Pi.

B. Description

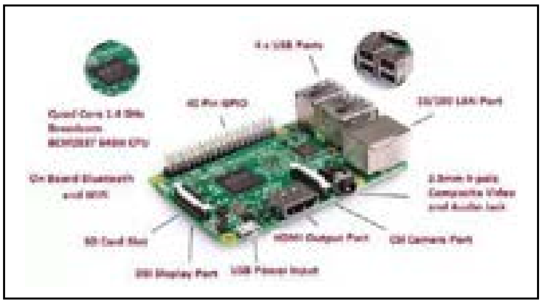


Figure 4: Raspberry Pi Node computer.[5]

- Raspberry Pi: The Raspberry Pi 3B+ serves as the computing platform for the wireless communication system in armored vehicles. It processes voice data captured by USB microphones and facilitates wireless connectivity within the vehicle. Additionally, it interfaces with earphones to deliver clear audio output for communication purposes. In the proposed study, Raspberry Pi serves as the main processing unit and plays a key role in enabling voice communication in real time within armoured vehicles. Voice transmission and data processing are only two of the many communication activities that can be seamlessly integrated with the Raspberry Pi thanks to its versatility and affordability as a microcontroller. Its low power consumption and small size make it ideal for deployment in cramped areas like armoured vehicles. Furthermore, the proposed system's versatility and scalability are improved by the Raspberry Pi's compatibility with a variety of software libraries and communication protocols. All things considered, the Raspberry Pi is the backbone of the proposed model, providing dependable and effective voice communication capabilities even in the most difficult operating situations.

- Earphones: The earphones are audio output devices that allow the user to hear sounds produced by the Raspberry Pi. Each Raspberry Pi is connected to a set of earphones, which allows audio to be played back on each device.
- USB microphone: The USB microphone is an audio input device that allows the Raspberry Pi to capture sound from its surroundings. Each Raspberry Pi is connected to a USB microphone, which allows audio to be recorded on each device.
- Wi-Fi network: The Wi-Fi network is a wireless network that allows the two Raspberry Pi devices to communicate with each other. The Raspberry Pi devices are connected to the same Wi-Fi network, which allows them to transfer data, including audio data, between each other.
- Cloud (DropBox): The Cloud block, specifically utilizing Dropbox, serves as the central storage and backup component for the proposed system. It plays a vital role in ensuring the secure and convenient storage of all recorded files and data exchanged between the two Raspberry Pi devices. The use of cloud technology is an essential component of this model that improves the efficacy and functionality of the communication system. Cloud connection provides important characteristics like scalability, strong data protection, and remote access to audio files. It also makes data storage, retrieval, and analysis easy. To ensure effective data flow and administration, executed cloud integration by using APIs, cloud storage services, and data transmission protocols. Under this framework, sensitive data is protected by strict security procedures and is uploaded, stored, accessed, and archived in a cloud environment.

Among the many advantages of integrating cloud technology are post-operation analysis, operational continuity, and data-driven strategy planning. Looking ahead, to further improve the capabilities of the presented communication system, Future developments might include machine learning algorithms, advanced analytics, and integration with other cloud-based services. This research work provides a thorough analysis on cloud integration which plays important role in improving military communication capabilities by demonstrating the significance and influence through the careful management and organization of relevant information.

C. *Working*

- Initialization and Setup:

First, the required software libraries are turned on in the Raspberry Pi microcontrollers. To guarantee correct connectivity, a wireless communication network is created amongst the Raspberry Pi microcontrollers.

- Audio Recording Initiation:

"sounddevice" library is for audio and it start the audio recording process when it get a command signal, such pressing a key. Also USB microphones are turned on so that crew members' voices and the surrounding noises inside the armored vehicle can be recorded.

- Audio Processing:

Signal processing methods are used to process the recorded audio data in order to reduce noise and improve clarity. Additionally, the "soundfile" library is used to transform the analogue audio signals into digital format for storage and transmission.

- Network Transmission:

Using the 'socket' library, a network connection is created between the Raspberry Pi microcontrollers that are transmitting and receiving data. Then, audio data has been send and that has been processed to the intended receiver via the wireless communication network.

- Audio Reception:

Through the established network connection, the sent audio data is received by the Raspberry Pi microcontroller. By utilizing the "pygame" package, received audio data are buffered and get it ready for playing.

- Audio Playback:

For audio playback the headphones are attached to the Raspberry Pi microcontroller that is receiving the audio data and output it through them. The main aim of this is to make sure that the recipient crew member hears the audio message conveyed to them smoothly and in real time.

- Timestamp Generation:

The "datetime" package is used to create precise timestamps for the recorded and transmitted audio files in order to manage and archive data accurately. Every audio file has a timestamp associated with it.

- Data Management:

For effective data management, the 'pandas' library is used that will arrange the recorded audio files and corresponding timestamps. In addition, the audio files are saved for later use and analysis in a specific directory or cloud storage service.

- Error Handling and Recovery:

Error handling technologies are used for mistake detection and correction. This guarantees the dependability of audio transmission. By using retransmission methods and error recovery techniques, outages or packet loss are addressed.

- System Shutdown:

The Raspberry Pi microcontrollers and related devices are safely turned off after communication tasks are finished or when not in use. This entails turning off the equipment's power supply and storing it correctly to guard against damage or malfunction.

IV. RESULTS

The comprehensive evaluation of the wireless communication system reveals its robust performance under operational conditions. This section showcases the system's capability in recording, transmitting, receiving, and archiving audio communications, underpinned by visual evidence from the tests. The screenshots included below represent key milestones in the communication process, highlighting the system’s reliability and the seamless integration of cloud storage solutions

Figure 5: Transmitter: Code Screenshot and Command Window highlighting Recording of audio to be transmitted.

The first stage of the communication process is seen in this screenshot [Fig.5], where the system's interface is successfully used to record an audio file. This image shows the user interface as it appears at the time of recording, demonstrating how easy and natural it is to start voice conversation within the armored vehicle.

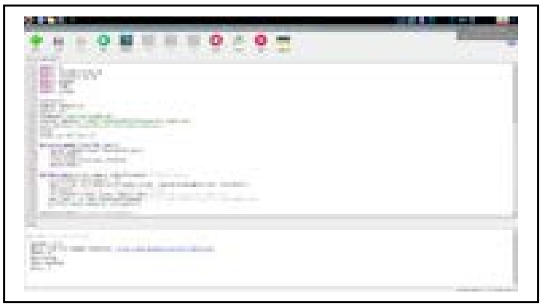


Figure 6: Receiver: Code Screenshot and Command Window highlighting reception of audio received.

The audio file is captured and prepared for playback by the recipient Raspberry Pi unit after transmission as shown in [Fig.6]. The UI of the receiver is seen in this screenshot, which shows that the audio file was successfully received. It is evidence of the system's effective and dependable wireless network transmission capabilities.

Figure 7: Post Reception File Uploading

As part of the system's archival plan, audio files are automatically transferred to Dropbox to guarantee the durability and accessibility of the communication data. The audio file in the Dropbox folder is visible in this screenshot [Fig.7], demonstrating how the system may be integrated with cloud storage providers for safe and scalable data management.

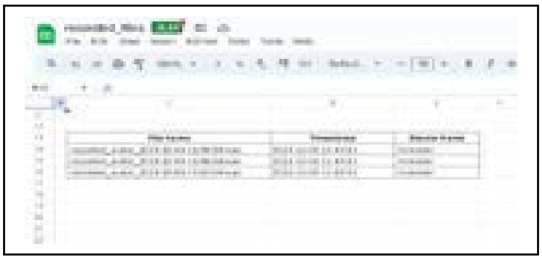


Figure 8: Time Stamp Generation in Excel

The snapshot [Fig.8] of an Excel spreadsheet, in which every audio file's reception is recorded with an exact timestamp, serves as the last piece of visual proof. This feature, which offers a transparent and well-organized record of all communications, is essential for operational evaluation and data management.

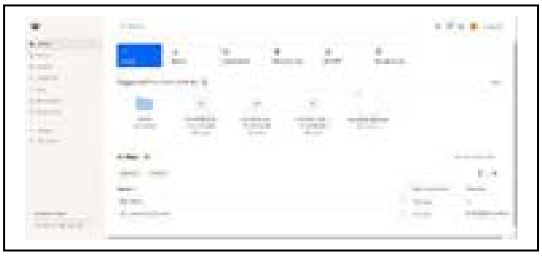


Figure 9: Dropbox Upload Confirmation

The snapshot [Fig.9] of a Dropbox page, in which every recorded audio files are stored with their excel file.

V. CONCLUSION

The prototype's operation indicates that the recommended wireless communication system for crews of armored or commercial combat vehicles is a practical substitute. Apart from its mobility, the system has more dependability and security compared to the wired method. In addition to being easy to set up and maintain, the suggested solution is also extensible to accommodate a larger network of devices and is competitively priced to execute. Compared to the current connected system, the suggested option for tank personnel is a significant improvement.

VI. FUTURE SCOPE

The integration of Raspberry Pi microcontrollers into a low-cost, wireless communication solution holds promise for future developments in military vehicle communications. Nevertheless, handling any complexity and streamlining processes are crucial for a successful implementation. Focusing on maintaining compatibility with current military communication systems and protocols can help to reduce complexity and enable smooth integration into the current infrastructure. Furthermore, by further reducing hardware components and optimizing power consumption, efforts should be focused on enhancing endurance and mobility, especially for extended missions in remote locations. By continuously improving user interfaces and interaction designs and taking end-user feedback into consideration, usability and user experience may be further improved.

To assess performance, dependability, and efficacy in real-world scenarios, extensive field testing and validation in a range of operational conditions is also required. To safeguard vital military communications, cybersecurity defences against cyberattacks, such as authentication protocols and encryption techniques, must be strengthened. Security is another crucial issue that needs to be addressed. Future wireless communication technology generations could revolutionize military vehicular communications and improve battlefield operational capabilities by addressing these areas.

The proposed model uses a variety of tactics to improve performance. This involves improving hardware for more processing power, streamlining the codebase to minimise processing overhead, and streamlining network setups to reduce latency. It is also intended to create caching methods, investigate parallel processing strategies, and gather input for iterative improvement. By concentrating on these areas, the proposed work overall performance can be improved and guaranteed that it will be effective in fulfilling the stringent needs of military communication applications.

ACKNOWLEDGMENT

We would like to express our gratitude to our entire Electronics and communication Department faculty, Shri Ramdeobaba college infrastructure and management who provided us with invaluable help and direction to carry out our research.

REFERENCES

[1] K. Sharma, P. Singh, S. Naruka, Y. Sighal, and I. Kishor, "Design and Analysis of Raspberry Pi Based Voice Communication System," Department of Information Technology, Arya Institute of Engineering & Technology, Jaipur, 2022.

[2] G. V. Kumar, I. P. Kumar Reddy, J. S. Satyanarayana, and V. S. Vivek, "DGS Based Ultra Wideband Antenna For Wireless Applications," Department of Electronics and Communication Engineering, Andhra Loyola Institute of Engineering and Technology, Vijayawada (AP), India, 2020.

[3] Andrea S. Krausman, "Understanding Audio Communication Delay in Distributed Team Interaction", U.S. Army Research Laboratory Human Research and Engineering Directorate Aberdeen Proving Ground, MD 21005, 2017.

[4] Jason Malliss, Vaibhav Patel, Sushilkumar Yadav, Himanshu Varun, " Real Time Communication (RTC) Device using Raspberry Pi ", Dr. Suprava Patnaik Department of Electronics and Telecommunication Xavier Institute of Engineering, Mahim Mumbai, India,2018.

[5] Data Sheet- " Raspberry Pi 3 Model B+ ".

[6] S. K. Reddy, A. S. Naik, and G. R. Mandel, "Development of a Reliable Wireless Communication System to Monitor Environmental Parameters from Various Positions of Underground Mines to the Surface using ZigBee Modules," Department of Mining Engineering, National Institute of Technology Karnataka, Surathakal, Mangalore 575025, India, 2023.

[7] S. L. Smith and M. A. Goodrich, "Raspberry Pi-based Voice Communication System for Tank Crew," 2019.

[8] Hirak Dipak Ghael, Dr. L Solanki, Gaurav Sahu, "A Review Paper on Raspberry Pi and its Applications", Department of Electronics and Communication Engineering, BKBIET, Pilani, India, 2021.

[9] Rajeeb Lochan Dash, Mrs. A. Ruhan Bevi, " Real-time Transmission of Voice over 802.11 Wireless Networks Using Raspberry Pi " , Asst. Prof. (Sr.G) Electronics and Communication Department, SRM university, Kattankulathur, Chennai, India,2014.

[10] H. Zhu, W. Zhou, Z. Li, L. Li, and T. Huang, "Requirements-Driven Automotive Electrical/Electronic Architecture: A Survey and Prospective Trends," State Key Laboratory of Networking and Switching Technology, Beijing University of Posts and Telecommunication, Beijing, China, 2016.

[11] Wardi, A. Achmad, Z. B. Hasanuddin, D. Asrun and M. S. Lutfi, "Portable IP-based communication system using Raspberry Pi as exchange," 2017 International Seminar on Application for Technology of Information and Communication (iSemantic), Semarang, Indonesia, 2017, pp. 198-204.

[12] Rajeeb Lochan Dash, Mrs. A. Ruhan Bevi, " Real-time Transmission of Voice over 802.11 Wireless Networks Using Raspberry Pi " , Asst. Prof. (Sr.G) Electronics and Communication Department, SRM university, Kattankulathur, Chennai, India,2014.

Detection of Water Quality for Health Monitoring through CNN Image Analysis

Ayush Roy Gour
Department of Computer Science and Engineering, Vardhaman College of Engineering
Hyderabad, India
ayushroy1337@gmail.com

Revanth Kumar
Department of Computer Science and Engineering, Vardhaman College of Engineering
Hyderabad, India
mrevanthkumar6159@gmail.com

Vidhit Reddy
Department of Computer Science and Engineering, Vardhaman College of Engineering
Hyderabad, India
vidhitvidhitreddy@gmail.com

Uma Datta Amruthaluru
Department of Computer Science and Engineering, Vardhaman College of Engineering
Hyderabad, India
umadattaa@gmail.com

Shanmugasundaram Hariharan
Department of Computer Science and Engineering, Vardhaman College of Engineering
Hyderabad, India
mailtos.hariharan@gmail.com

Vinay Kukreja
Chitkara University Institute of Engineering and Technology,
Chitkara University,
Punjab, India
vinay.kukreja@chitkara.edu.in

Abstract—This work offers an effective method for finding water features using netting stimulated by a convolutional neural network (CNN) countenance research. Conventional techniques for assessing the condition of the water may be difficult and time-consuming. The proposed methodology seeks to expedite this procedure by utilizing CNNs' proficiency in concept recognition tasks, which automates and enhances the accuracy of water quality assessment. Data collection, preprocessing, CNN construction, model preparation, and deployment inside a reliable netting application are among the techniques used. This research work establishes the effectiveness of CNN-located methodologies in natural resource protection by a study survey. Building on this support, the proposed technique offers a flexible and convincing real-occasion water status assessment solution. The outcomes of the experiments demonstrate how well the proposed CNN model performs when it comes to correctly classifying water-type limitations from pictures. This strategy makes responsible and convenient use of natural resources possible while also providing significant advantages for community health and preservation.

Keywords— Water Quality Assessment, Web Application, (CNN), Environmental Monitoring, Deep Learning

I. INTRODUCTION

Access to clean and reliable water is essential for maintaining life and guaranteeing community health. The quality of water considerably impacts human well-being, ecosystem uprightness, and socioeconomic incidents. Consequently, monitoring water characteristics is superior to identifying potential risks, checking contamination, and safeguarding human populations and marine environments. Traditional methods of water condition estimation frequently involve manual sample accumulation, workshop analysis, and dossier understanding, which may be late, labor-exhaustive, and costly. Additionally, these designs cannot always support absolute-occasion information, restricting their influence in responding to vital incidental conditions and arising dangers.

In recent years, there has existed a growing interest in leveraging progressive sciences, specifically image reasoning and machine intelligence, to develop more effective and electrical approaches to water quality listening. Image-located arrangements offer the potential to streamline the listening process by providing brisk and non-destructive amounts of water condition limits directly from the ocular dossier. Moreover, the advent of machine intelligence methods, particularly Convolutional Neural Networks (CNNs), has transformed the field of concept study, enabling correct and robust categorization and acknowledgement tasks. This paper intends a novel approach to water quality discovery through the happening of a web request stimulated by a CNN image study. The aim search specifies stakeholders accompanying a convenient platform real-period evaluation of water quality limits established visual dossier, by controlling the efficiencies of CNNs, which become proficient culling intricate patterns and countenance from representations.

The user's legitimate-time study form offers consumers an accelerated response to the quality of the water in the composed settings. This feature not only advises individuals but further acts as an early warning scheme, exceptionally in cases when the immediate operation should be to reduce water dirtiness. The CNN model's foolproof connection and fast processing efficiencies manage an approachable and efficient finish for consumers with different grades of mechanic skills. The meaning of this research is to display to the public its potential to address key challenges and guide common water status monitoring approaches. First and foremost, the projected netting application offers a swift and economical way of assessing water value limits, admitting for timely invasion, and being in charge. By providing real-period facts, collaborators such as material instrumentalities, water serviceableness, researchers, and society can immediately label and respond to water-related issues, accordingly minimizing risks to human energy and incidental depravity. Furthermore, the utilization of CNN countenance study allows the detection of cunning ocular cues and patterns exhibiting water feature limits, even in complex and dynamic surroundings. Traditional patterns grant permission, but the struggle to capture such shadings or concede possibility requires far-reaching manual understanding, inasmuch as CNNs can automatically discover and admit appropriate features from big capacities of figure data. This mechanized approach not only embellishes the efficiency and veracity of water status amount but also has the potential to reveal earlier anonymous patterns or trends in the water status dossier. Moreover, the happening of netting use for water quality discovery democratizes the approach to tangible facts by providing a user-friendly connection accessible to an expansive range of shareholders. Through the request, users can transfer data to a server countenances of water samples captured from differing sources, to a degree waterways, reservoirs, or guzzling water reservoirs, and endure instant feedback on water feature limits. This democratization of dossier empowers societies to actively take part in listening their local water possessions and advocate for inevitable interventions or procedure changes to save water feature.

In summary, this paper presents a novel approach to water condition discovery through CNN image reasoning joined into a netting use framework. By joining state-of-the-art electronics accompanying handy interfaces, this research aim to revolutionize water characteristic listening, making it more approachable, adept, and responsive to the needs of two together human public and the atmosphere. The after portions of this paper will investigate the methods, exercise, and validation of the proposed approach and its potential uses.

II. Related Work

Water kind evaluation is a critical facet of the preservation of natural resources and community health management. Over the age, analysts have investigated various methods to reinforce the effectiveness and accuracy of water character listening. In this research survey, existing research works that utilizes reasoning and machine learning methods, specifically Convolutional Neural Networks (CNNs) are studied, for water value discovery. Convolutional neural networks, or CNNs, are used in water quality detection has been explored in various studies. Patil et al. proposed a system that combines image processing and CNN models to identify and classify water quality accurately based on visual features. The system captures images of water bodies and extracts features such as color, texture, and turbidity using image processing techniques. The trained CNN models for classifying and predicting water quality then employ these characteristics as input. A real-world dataset on water quality is used to test the suggested method, and the findings show that the system achieves a high degree of precision in detecting water quality. Early warning signals for possible health and environmental dangers can be provided by the system's capacity to identify visual irregularities in water quality [1].

Baek et al. grown a CNN-Long Short-Term Memory (LSTM) linked deep knowledge approach for simulating water condition, containing total nitrogen, total phosphorous, and total basic element. The study demonstrated that the performances of both the CNN and LSTM models were in the "very good" range above the Nash–Sutcliffe efficiency value of 0.752. This research has contributed to the understanding of how deep learning models can be effectively used for simulating water quality parameters [2]. Aqua Sight is a mobile application that leverages advanced deep-learning techniques, particularly CNNs, to detect impurities present in potable water. This approach provides an efficient, low-cost, and precise method for assessing whether drinking water is uncontaminated or not. The development of such applications has the potential to revolutionize the way water quality is monitored in households and communities [3]. Another study assessed and monitored water quality parameters using CNN with 98% accuracy. The study employed data acquired from satellites to detect and quantify various factors impacting water quality, such as pH, salinity, turbidity, chlorophyll, and temperature. This research has highlighted the potential of satellite data in conjunction with CNNs for large-scale water quality monitoring [4].

A novel Hybrid CNN and Extreme Learning Machine (ELM) approach was used to detect the presence of anomalies in sensor-collected water data. The experiment of the proposed CNN-ELM model was carried out using the publicly available dataset GECCO 20195. This study has demonstrated the effectiveness of hybrid models in detecting anomalies in water data [5]. A study on water status prognosis by a mixture CNN-LSTM model accompanying Attention (CLA) was conducted to predict pH and NH3-N in the Beilun Estuary. The study proposed a water quality prediction model named CNN-LSTM with Attention (CLA) to predict the water quality variables. The study conducted a case study on the water quality dataset of Beilun Estuary to predict pH and NH3-N [7].

Early research explored CNNs for water quality classification using satellite imagery. P. Rad et al. (2018) achieved promising results in identifying polluted water bodies using this approach [8]. CNNs for water quality prediction based on color features extracted from water samples [9]. These pioneering works laid the foundation for further exploration of CNNs in water quality detection. Following these initial studies, researchers developed advanced CNN architectures and techniques. The HydroColor app handles a smartphone's camcorder and auxiliary sensors to measure the detached thinking reflectance of organic water bulks. The app uses the smartphone's digital camera as a three-band radiometer to calculate the remote sensing reflectance in the broad wavelength bands red, green, and blue. The reflectance maybe upside-down to estimate the concentration of spellbinding and uneven essences in the water, that are predominately composed of postponed silt, chlorophyll, and annulled natural resources.

X. Wu et al. (2020) investigated the integration of transfer learning with CNNs for water quality monitoring, achieving improved accuracy while reducing training time [11].

These advancements showcased the potential of sophisticated CNN architectures for water quality detection. The latest developments involve using CNNs in conjunction with Long Short-Term Memory (LSTM) networks for water quality prediction. This approach incorporates temporal information for more robust analysis, as explored [10].

Despite the significant progress, some limitations remain. Datasets for water quality analysis can be imbalanced, potentially leading to biased model predictions.

Additionally, the interpretability of CNN models can be challenging. Improving CNN interpretability in environmental applications, paving the way for a deeper understanding of model decision-making. The dataset utilized in this place research study was derived from Kaggle, a well-known manifesto for giving and discovering datasets across differing rules. The dataset picked for this study was preferred based on allure pertinence to the research objectives and the chance of inclusive information to address the research questions efficiently. Kaggle's Terrace offers an extensive group of datasets, promoting collaboration and novelty in dossier-compelled research endeavors [6].

III. METHODOLOGY

This section presents the detailed methodology for developing a water quality detection system that leverages a Convolutional Neural Network (CNN) for backend analysis and incorporates a user-friendly interface (UI) for image upload.

Data Collection:

The foundation of a robust CNN model centers on the quality and diversity of the training dataset. Therefore, this research aims to acquire a comprehensive water quality image dataset encompassing various aspects:

a) Pollution Levels: Images representing clean water, moderately polluted water, and heavily polluted water from different sources (e.g., industrial waste, agricultural runoff, algal blooms) will be collected.

b) Water Types: The dataset will include images of diverse water bodies like lakes, rivers, oceans, and even stagnant water sources to enhance model generalizability across various water ecosystems.

Potential data sources include established public datasets like UW-Madison's Water Quality Dataset (https://www.epa.gov/wqc) or Kaggle competitions focused on water quality image classification, which can provide a valuable starting point.

B. Data Pre-processing:

To ensure consistency and optimize the training process, the collected images will undergo several preprocessing steps. Firstly, resizing will be applied to ensure uniformity in dimensions across all images, eliminating irrelevant variations in size for compatibility with the chosen CNN architecture. Subsequently, normalization will be conducted to standardize pixel values within each image to a specific range, such as 0-1 or -1 to 1, preventing high-intensity pixels from dominating training. Additionally, color space conversion may be implemented, such as grayscale conversion or selective channel extraction, based on the model's architecture and relevant features for water quality classification. Moreover, data augmentation techniques will be employed to enhance dataset diversity and prevent overfitting. These techniques include random cropping, flipping (horizontal and vertical), rotation, and slight color jittering, generating new image variations to aid the model in learning robust features that generalize well to unseen data, thereby improving overall performance and accuracy.

A. Algorithm & Background:

Artificial Neural Networks (ANNs) serve as the foundational architecture for Convolutional Neural Networks (CNNs), which are specifically designed for image processing and pattern recognition tasks. ANNs are computational models inspired by the structure and functioning of the human brain, consisting of interconnected layers of artificial neurons. These neurons process information and learn from data through a process called backpropagation. In ANNs, each neuron receives input signals, processes them using a weighted sum, and applies an activation function to produce an output signal. CNNs build upon this framework by introducing convolutional layers, which consist of learnable filters or kernels that scan across input images to extract features such as edges, shapes, and textures. By stacking multiple convolutional layers, CNNs can learn hierarchical representations of visual features, capturing increasingly complex patterns and structures within the input data.

In addition to convolutional layers, CNN architectures typically include pooling layers, which perform downsampling operations on the outputs of convolutional layers. Pooling helps reduce the spatial dimensions of the feature maps, thereby controlling model complexity and preventing overfitting. Common pooling techniques include max pooling, where the maximum value within a defined region is selected, and average pooling, which computes the average value within a region. Activation functions are essential components of neural networks, introducing non-linearity into the network's decision-making process. ReLU (Rectified Linear Unit) is one of the most widely used activation functions in CNNs. ReLU applies a simple threshold function, effectively zeroing out negative values and allowing the network to learn sparse representations. This helps alleviate the vanishing gradient problem and accelerates the training process by enabling faster convergence.

$$F(x) = max\ (0, x) \qquad (1)$$

In the output layers of CNNs, particularly for classification tasks, the sigmoid activation function is commonly used. The sigmoid function, squashes the network's output to a range between 0 and 1, effectively transforming the network's raw output into probability scores. This is particularly useful in binary classification tasks, where the network predicts the probability of an input belonging to a particular class.

$$F(x) = \frac{1}{1+e^{-x}} \qquad (2)$$

Finally, fully connected layers constitute the final stages of a CNN architecture. These layers resemble traditional ANNs, connecting all neurons from one layer to all neurons in the next layer. Fully connected layers allow the network to learn higher-level abstractions and make final predictions or classifications based on the extracted features.

1) Evaluation Metrics: Accurately assessing the performance of a CNN model trained for water quality detection is crucial. Here, the performance metrics utilizes standard metrics for evaluating image classification tasks, with a focus on their relevance to water quality prediction:

Table I. Evaluation Matrix

LABEL	Actual:0	Actual:1
Result:0	True Negative (TN)	False Positive (FP)
Result:1	False Negative (FN)	True Positive (TP)

a) Accuracy:

Accuracy is a fundamental metric used to evaluate the performance of a Convolutional Neural Network (CNN) model in correctly classifying water feature boundaries. It measures the proportion of correctly classified instances out of the total number of instances in the dataset and is typically expressed as a percentage. Accuracy serves as a primary indicator, reflecting the overall percentage of images accurately classified by the CNN model. It is computed by dividing the number of correctly predicted images by the total number of test images. While accuracy provides a general assessment of model performance in water quality detection, it can be misleading in datasets with class imbalances, where certain water quality classes may be more prevalent than others. For instance, achieving high accuracy may be possible if the model effectively identifies clean water images (dominant class) but struggles with identifying polluted water samples.

$$Accuracy = \frac{\text{Number of correct predictions}}{\text{Total number of Predictions}} \quad (3)$$

b) Precision:

Precision in the context of water quality detection refers to the positive predictive value, representing the proportion of images predicted as a specific water quality class that are indeed true positives within that class. Precision plays a crucial role in assessing the confidence level of the model's predictions for individual water quality classes.

$$Precision \frac{\text{True Positives}}{\text{True Positives+False Positives}} \quad (4)$$

c) Recall:

Recall, also known as sensitivity, measures the completeness of the model's predictions for a specific class. It shows the dimension of actual helpful cases (concepts owned by a specific water feature class) that were right labeled for one model. Interpretation in Water Quality Detection; High recall for a critical water quality class (e.g., heavily polluted water) is essential. It ensures the model doesn't miss a significant number of polluted water samples, potentially leading to underestimation of the severity of water quality issues.

$$Recall = \frac{\text{True Positives}}{\text{True Positives+False Negatives}} \quad (4)$$

d) F1 Score:

F1-score serves as a harmonic mean between two crucial metrics: accuracy and recall. It offers a balanced perspective on model performance, which helps overcome the limitations of relying solely on accuracy as an evaluation metric. In the context of water quality detection, the F1-score provides a more comprehensive assessment of the model's performance. By considering both the model's ability to correctly identify positive cases (recall) and its precision in avoiding false positives, the F1-score offers a more nuanced understanding of the model's effectiveness..

$$F1 = 2 * \frac{\text{Precision*Recall}}{\text{Precision+Recall}} \quad (5)$$

Table II. Water Quality Detection – Algorithm

Algorithm for CNN-based Water Quality Detection
Input: A dataset of images of water bodies which are dirty and clean.
Output: A trained model for predicting water quality of input image
1. Data Acquisition: Gather a diverse dataset of water quality images representing various pollution levels and water types. Sources can include public datasets, environmental agencies, or field sampling campaigns. 2. Label Data: Meticulously label each image with corresponding water quality parameters (e.g., turbidity, algae concentration, presence of pollutants). This step is crucial for supervised learning. 3. Preprocess Data: Apply preprocessing techniques to ensure data consistency and optimize training. 4. Define CNN Architecture: Select a suitable CNN architecture, such as VGG16, ResNet, or Inception, potentially with modifications for water quality classification. 5. Hyperparameter Tuning: Define key hyperparameters like learning rate, number of filters, and optimizer. 6. Model Training: Train the CNN model on the training set, using the labeled data to learn the relationship between image features and water quality parameters. Iteratively update weights and biases to minimize the loss function (difference between predicted and actual labels). 7. Model Evaluation: Assess final model performance on the unseen testing set using metrics like accuracy, precision, recall, and F1-score. 8. Deployment (Optional): Deploy the model in a production environment.

The proposed algorithm delineates a systematic approach to implementing Convolutional Neural Networks (CNNs) for water quality detection, covering various stages from data acquisition to model deployment. Initially, a diverse dataset comprising images representing water bodies with differing pollution levels and water sources is gathered. Each image is meticulously labeled with corresponding water quality parameters, facilitating supervised learning. Preprocessing techniques are then applied to ensure data consistency and optimize training efficiency, potentially involving resizing, normalization, and color space conversion.

Subsequently, a suitable CNN architecture is selected, and hyperparameters such as learning rate and optimizer are defined, with potential tuning to enhance model performance. The CNN model is trained on the labeled dataset, iteratively adjusting weights and biases to minimize the loss function. Following training, the model is evaluated on a separate test dataset to assess its performance accurately.

This algorithmic framework provides a structured approach to leveraging CNNs for water quality detection, enabling researchers and practitioners to effectively utilize deep learning techniques for environmental monitoring and management.

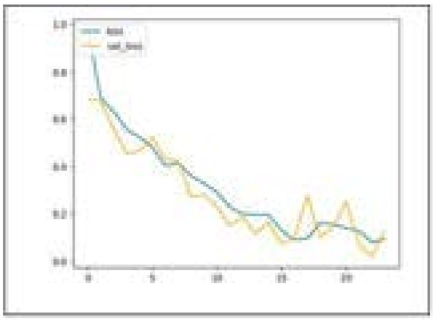


Fig. 3. Loss versus validation loss graph

The performance metrics, including accuracy, precision, recall, and F1-score, offer comprehensive insights into the efficiency of the model and areas for improvement. These metrics enable a thorough evaluation of the model's performance across different aspects, such as correctness, completeness, and balance between precision and recall. Additionally, the visualization of loss versus validation loss through a graph provides valuable information about the training progression of the CNN model. The initial high loss indicates significant errors during training, which gradually decrease as the model iterates through epochs. The decreasing trend in both training and validation losses reflects improved performance and minimized error in predicting both seen and unseen data. This convergence between training and validation losses suggests that the model is effectively learning and generalizing well to unseen data, essential for robust performance in real-world scenarios. Overall, the sharp reduction in loss values over epochs shown in figure 3 highlights the efficacy of the CNN's training process, indicating its ability to learn and adapt to the complexities of the task at hand, particularly in image-based applications.

IV. RESULTS AND DISCUSSION

The CNN model was trained using a dataset of 228 water quality images, labeled as either "clean" or "dirty.". By employing a Basic CNN architecture for water quality classification, the experimentation is performed. The sigmoid activation function in the final layer outputs a value between 0 and 1. A threshold value (e.g., 0.5) is typically used to convert this continuous output into a binary classification:

Output >= Threshold (e.g., 0.5): Classified as "Clean Water" (Output closer to 1 signifies higher confidence in clean water)

Output < Threshold (e.g., 0.5): Classified as "Dirty Water" (Output closer to 0 signifies higher confidence in dirty water). The trained model achieved an overall accuracy of 97% on the unseen testing set. To contextualize this result, the performances are compared with CNN models from similar research efforts using binary classification (clean vs. dirty water) as depicted in Table III.

Table III. Parameter comparison with existing works

Approach	Accuracy	No. of samples	Considered Parameters
Proposed	97%	228	RGB values, Turbidity
CNN [13]	80%	40	GPS position, RGB values
Neural Network [14]	94%	2967	Chemical composition values based on importance

The study achieved a notably higher water quality prediction accuracy of 97% compared to existing research. Utilizing a dataset of 228 samples, the proposed model integrates RGB values and turbidity measurements. In contrast, previous studies varied in sample size and parameters considered. For example, [13] attained 80% accuracy with 40 samples, employing GPS positions and RGB values, while [14] achieved 94% accuracy with 2967 samples, focusing on chemical composition values. The proposed work highlights the efficacy of incorporating both RGB values and turbidity in CNN-based models, enhancing prediction accuracy crucial for reliable water quality monitoring systems. Several factors could influence these variations in accuracy. The size and diversity of the training datasets used in each study might play a role. Additionally, the specific CNN architectures and hyperparameter settings employed could contribute to these differences.

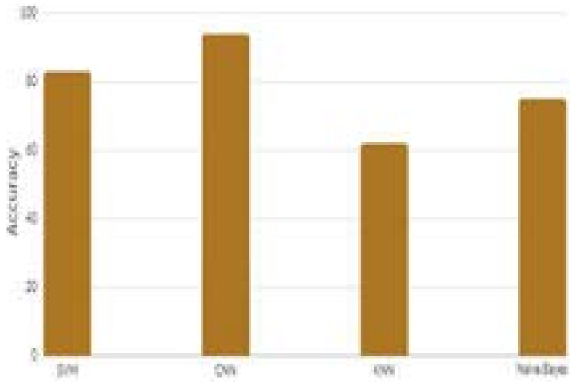


Fig. 4. Accuracy comparison with other algorithms

The presented bar graph in figure 4 provides a comparative analysis of the effectiveness of various machine learning algorithms in predicting water quality. CNN emerges as the most accurate method with a precision of 94%, followed by SVM with 83%, Naive Bayes with 75%, and KNN with 62%. The substantial margin between CNN and other algorithms suggests its superiority in handling the intricacies of water quality data. This superiority can be attributed to CNN's ability to extract complex features from images, which is particularly advantageous for tasks involving image analysis, such as water quality prediction.

The comparison of different machine learning algorithms highlights the strengths and weaknesses of each approach in the context of water quality prediction tasks. SVM demonstrates a respectable accuracy rate of 83%, indicating its effectiveness in handling non-linear data and making it a suitable choice for scenarios where linear separation is not optimal. This suggests that SVM can effectively capture complex relationships within water quality data, contributing to its reliable performance. Naive Bayes, despite its simplicity, achieves a moderate accuracy of 75%, suggesting its potential for basic classification tasks. However, its performance falls behind SVM and CNN, indicating limitations in handling the nuances of water quality data compared to more sophisticated algorithms. This suggests that Naive Bayes may struggle with capturing the intricate patterns present in water quality datasets, leading to suboptimal performance in prediction tasks. KNN, with the lowest accuracy rate of 62%, appears to struggle in accurately predicting water quality. This can be attributed to its reliance on proximity-based classification, which may not be well-suited for the complexities inherent in water quality prediction tasks. KNN's performance suggests that it may not effectively capture the underlying structure of water quality data, leading to poorer predictive accuracy compared to other algorithms. Overall, the findings highlight the superiority of CNN in water quality prediction tasks. CNN's robust feature extraction capabilities allow it to effectively capture complex patterns in water quality data, leading to superior performance compared to other machine learning algorithms. Therefore, CNN emerges as the preferred choice for water quality prediction tasks, offering the highest accuracy and demonstrating its effectiveness in handling the intricacies of water quality datasets

V. CONCLUSION

This research study explored the potential of Convolutional Neural Networks (CNNs) for rapid water quality detection through image analysis. It emphasizes the significance of CNNs in feature extraction from image data and outlines the process of data acquisition, preprocessing, model training, evaluation, and potential deployment. The paper highlights the importance of meticulous data labeling for supervised learning and discusses standard evaluation metrics like accuracy, precision, recall, and F1-score, especially in dealing with imbalanced datasets. It also explores class-specific evaluation techniques. While CNNs offer promising benefits for on-the-spot water quality assessments and real-time pollution detection, there are challenges to broader adoption. The paper suggests future research directions, including the need for diverse datasets, advanced data augmentation techniques, and efforts to understand CNN decision-making processes. Addressing real-world deployment challenges such as variations in lighting conditions and integration with existing monitoring infrastructure is also crucial. Additionally, developing CNN models that generalize well to unseen environments and pollution types through techniques like transfer learning is essential. In conclusion, CNNs present a powerful tool for water quality detection, offering a rapid, efficient, and cost-effective alternative to traditional methods. By addressing outlined challenges and continuously refining CNN architectures and training methodologies, researchers can pave the way for robust and reliable water quality detection systems, crucial for safeguarding water resources and ensuring a healthier planet for future generations.

REFERENCES

[1] Patil, J.S., Mailaram, A., Basa, P.N.K., Sai Sravya, A., Yadam, B. (2024). Integrating Image Processing and Convolution Neural Networks for Water Quality Detection. In: Kumar, A., Mozar, S. (eds) Proceedings of the 6th International Conference on Communications and Cyber Physical Engineering . ICCCE 2024. Lecture Notes in Electrical Engineering, vol 1096. Springer, Singapore.

[2] Baek, S.-S.; Pyo, J.; Chun, J.A. Prediction of Water Level and Water Quality Using a CNN-LSTM Combined Deep Learning Approach. Water 2020, 12, 3399. https://doi.org/10.3390/w12123399

[3] Pushpa Mala, S., Prajwal Raju, P., Poojashree, B. et al. Underwater Fleck Detection Using Convolutional Neural Network. J. Inst. Eng. India Ser. B (2023). https://doi.org/10.1007/s40031-023-00949-1

[4] S. Thylashri, A. Manikanta, I. Poojitha and K. Varshitha, "Water Quality Monitoring with Satellite Image Using Convolutional Neural Network," 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-5, doi: 10.1109/ICCCNT56998.2023.10306619

[5] Improving Water Quality Assessment through Anomaly Detection Using Hybrid Convolutional Neural Network Approach - https://www.sciencegate.app/document/10.30955/gnj.003923

[6] https://www.kaggle.com/datasets/elvinagammed/clean-dirty-water-dataset

[7] Yang, Y., Xiong, Q., Wu, C. et al. A study on water quality prediction by a hybrid CNN-LSTM model with attention mechanism. Environ Sci Pollut Res 28, 55129–55139 (2021). https://doi.org/10.1007/s11356-021-14687-8

[8] Rad, P., H. Behzadfar, A. Sharifi, and M. Talebpour. 2018. "Water Quality Monitoring Using Convolutional Neural Networks and Remote Sensing Data." Aquatic Science and Engineering 6 (1): 1-14.

[9] Mai, M., H. Song, L. Vu, and D. Phung. 2017. "A Convolutional Neural Network for Water Quality Prediction Based on Color Features." 2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE, 1302-1307.

[10] Hasan, M., M. Alam, M. Sohail, A. Khan, and M. Braswell. 2023. "Convolutional Neural Network-Long Short-Term Memory Model for Water Quality Prediction Using Sensor Data." Sustainability 15 (3): 1457.

[11] Wu, X., L. Ji, and J. Song. 2020. "Transfer Learning Based on Convolutional Neural Network for Water Quality Monitoring from Remote Sensing Images." Water 12 (12): 3444.

[12] https://www.epa.gov/wqc - Established public dataset UW-Madison's Water Quality Dataset.

[13] M Vijay Anand et al 2023 J. Phys.: Conf. Ser. 2484 012051

[14] Al-Madhoun, Mohammed Ashraf & Abu-Naser, Samy S. (2023). Neural Network-Based Water Quality Prediction. International Journal of Academic Information Systems Research (IJAISR) 7 (9):25-31.

[15] Leeuw, T., & Boss, E. (2018). The HydroColor App: Above Water Measurements of Remote Sensing Reflectance and Turbidity Using a Smartphone Camera. Sensors, 18(1), 256.

IMAGE-BASED PEST DETECTION AND IDENTIFICATION SYSTEM FOR AGRICULTURE

Sai Bhargav Vemulapalli
Dept. of EIE
VR Siddhartha Engineering College
Vijayawada, India
vsbhargav11@gmail.com

Durga Prasada Rao Pillagolla
Dept. of EIE
VR Siddhartha Engineering College
Vijayawada, India
durgaprasadarao.p@gmail.com

Ganesh Vempala
Dept. of EIE
VR Siddhartha Engineering College
Vijayawada, India
ganeshvempala123@gmail.com

Sri Mahalakshmi Ravuri
Dept. of EIE
VR Siddhartha Engineering College
Vijayawada, India
ravurisrilakshmi@gmail.com

***Abstract*— For farmers, identifying pests is the most difficult task in the agricultural industry. Farmers must use organic insecticides in conjunction with other appropriate techniques to combat pests. This project outlines a software prototype system that uses pest identification to control pests. Farmers must use the Android application to take an image of the pest. The pest image must then be uploaded to the software. One of the most difficult jobs for farmers and crop professionals in the agricultural industry is identifying pests over crops. This harms farmers as well as crops, resulting in a low output. The pests' image database is also taken into account. A comparison is made between a set of training and testing images. The convolutional neural network classification (CNN) approach is adopted to identify the class of Pests.**

Keywords—*CNN approach, Farmers, Pests,*

I. INTRODUCTION

India is an agricultural country. Farming is a source of income for many individuals. Agriculture is also vital to a country's economy. Farmers are rural residents. They rely only on agriculture for revenue. It accounts for 17.7% of gross domestic product. It will assist the nation by tackling unemployment. Crops are affected by some pests, including bacteria, viruses, and fungi. As a result, agricultural yields suffer from both quality and quantity issues. Crop quality and quantity may therefore be regulated without using chemical pesticides. This is accomplished by the use of organic pesticides. Organic insecticides are better since they kill pests without damaging plants while increasing quality and quantity.

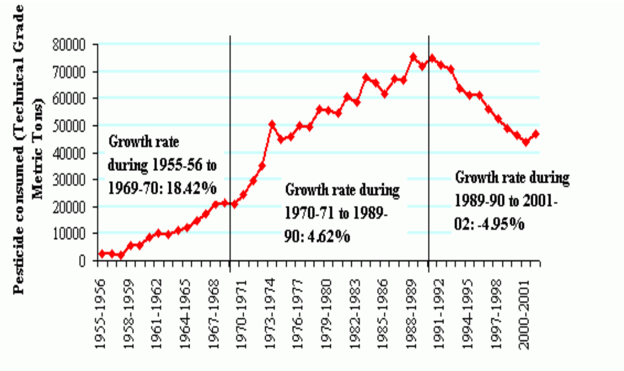


Fig. 1. Consumption of Pesticides in India(in%) [1]

Fig. 1 depicts the decline of agriculture as a result of the use of chemical pesticides. Using chemical insecticides on crops without first identifying the pest results in decreased crop growth. Identifying pests and using appropriate organic pesticides in agriculture is critical to avoiding losses in farmer yields and food quality. Because it is difficult to manually identify individual pests and administer effective organic pesticides, Digital Image Processing is used for early pest detection [2]. This approach also involves image acquisition, pre-processing, segmentation, feature extraction, and classification. Systems are complex computer programs that may provide answers and guidance to specific problems in a given area by comparing photographs to databases.

Image-based pest detection and identification technologies provide a long-term and efficient answer to these difficulties. Pest detection and identification in agriculture have traditionally depended mainly on physical labor, which may be time-consuming and error-prone. However, with the development of low-cost and accessible photography equipment, such as smartphones and drones, as well as rapid advances in artificial intelligence, the process may now be automated and enhanced. Image-based pest detection and identification technologies provide a more efficient, cost-effective, and accurate way to monitor and manage pests in agricultural settings[3].

II. LITERATURE SURVEY

Mayank et al., 2021, [4] proposed a pest control system that uses IoT and Image processing technologies. Effective image classification method for smartphones based on neural networks. Infrared sensors are used by the system to identify pests. Insects are kept out of the field by pests identified using sensors and ultrasonic wave technology with the aid of image processing.

Yaowei et al., 2020, [5] Transfer learning algorithms combined to provide a new method for image recognition and classification. The significance of TL-MobileNet for welding flaw detection is suggested in this paper. This experiment, which classifies welding flaws using "Weld" datasets, confirms that the TL-MobileNet can accurately identify particular defects within a restricted set of training sets. Numerous tests have been conducted with different image sizes and transfer learning models. Additionally, it validates the suggested TL-MobileNet approach, which is superior in terms of recognition accuracy and requires less computation time and model size.

Brian et al.,2020 [6] titled Biological Control and Integrated Pest Management in Organic and Conventional Systems. According to this research, IPM and organic farming are viable methods for increasing agricultural productivity, and they both rely on biological control as a tool that farmers can employ. reduced risk, use of pesticides, and effects on development. Working together, individuals utilizing two different sets of methods can advance the adoption of solutions for production issues that involve biological control.

G. Singh 2020 [7] Project titled Pest management in organic farming, becomes a challenging task without using any insecticides. It uses meticulous preparation ahead of time together with little adjustments to cultural norms as the main means of managing pests. employing eco-friendly methods as a backup against insect pests, such as the use of biological control agents and other bio-based goods. These products can reduce insect pests and boost the profitable yield of organic farming whether used alone or in conjunction with other strategies like integrated pest management.

Basri et al. 2020 [8] To identify cocoa fruit pests and diseases at their earliest stages, this study suggests using image processing techniques based on mobile applications. The idea of competence in the field of cacao farming was implemented by the system. Pre-processing the image using deep learning algorithms and image processing techniques is how it's done for machine learning systems and pattern recognition. To identify the pixels in the photos of cacao fruit, image processing techniques are included in the application program.

Indhumathi et al., 2019, [9] According to this model, the testing images and their set of extracted images should be compared with the geometrical properties of the training images, which include the mean and variance characteristics and morphological features. The likelihood and probability of pests are taken into account. The class of pests is identified using the naive Bayes approach.

III. EXISTING METHODS

A. K-Means grouping technique

An unsupervised machine-learning technique called K-means clustering is used to extract the target area from images of plant leaves. Sort the data elements into K-groups according to how similar they are. By computing the absolute difference between each pixel and the clustering center in the color space, the K-means technique uses automatic data grouping (clustering) to separate wheat ailments such as leaf rust, powdery mildew, and striped rust from images.

B. Edge detection techniques

An edge is a barrier with unique characteristics that separates two places. It signifies the change from one thing to another. Since the intensity drastically changes at the boundaries of the area, this detection technique is used to quantify the physical scale of items. Edge detection can be used to separate images into specific target portions according to criteria and is a crucial step in understanding image function.

C. Support Vector Machines "SVM"

The maximum distance between the examples is used to construct a judgment threshold by the SVM, a non-probabilistic linear classifier. By splitting the data into two subspaces with the greatest distance between the data closest to the hyperplane (referred to as a distinct hyperplane), an SVM looks for the hyperplane. The margin is the space that exists between the data and the hyperplane. The data points closest to the hyperplane are referred to as support vectors. There are two different types of support vectors: those for one subset of the data and those for another since the data is divided into two subgroups. They are referred to as negative and positive support vectors, respectively, in everyday speech. This technique allows for the creation of linear separators without requiring the original input space, but it does not allow for linear definition. Lastly, SVM may characterize complex functions while resisting overfitting because it is a nonparametric method that is not easily overfitted.

IV. PROPOSED METHODOLOGY

The pest datasets used in this study were gathered from the Kaggle website. In a web-based data science environment, users may search and share data sets, as well as evaluate and construct models using Kaggle. Various pest classes, including grasshoppers, cutworms, Mormon crickets, locusts, Japanese beetles, and true bugs, are taken into consideration. The process flow is shown in Fig. 2.

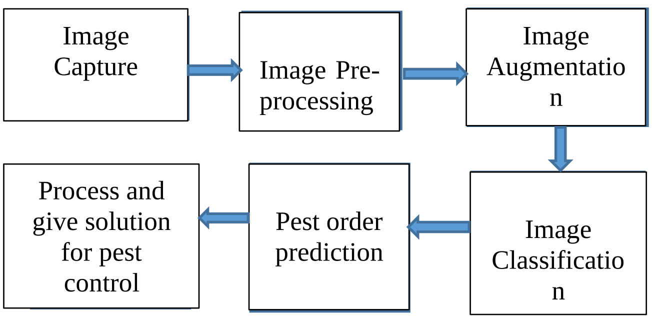


Fig. 2 Block diagram of the proposed system.

A software prototype that helps farmers identify and successfully manage pest infestations is the center of the proposed methodology for the image-based pest detection and identification system for agriculture. An image of a pest is uploaded to the software system once it is taken with the Android application. Convolutional neural network (CNN) classification is one of the advanced image recognition technologies that the system uses to accurately evaluate and classify uploaded pest photos. The CNN model learns to recognize different types of pests that are frequently encountered in agricultural settings by comparing a collection of training and testing images, making precise and effective pest identification possible.

A. Image acquisition and pre-processing

First, the cropped image was taken, and then it was processed. Pre-processing techniques are then used to improve the image by removing unwanted distortion and enhancing the necessary attributes for further processing [10]. Image quality improvement techniques are applied in image pre-processing to reduce noise and sharpen the image for greater precision. This process improves image quality to help with more accurate identification and classification of insects, pests, and diseases.

B. Segmentation

Segmentation is the following stage of the discovery process. To extract the target object from the sheet, segmentation is used for the colorized image. Segmentation methods and Gaussian mixture models are typically used. An exponential increase in execution time is caused by the numerous divisions, multiplications, and computation averages that these algorithms require.

C. Feature extraction

In image processing, feature extraction is crucial. To extract characteristics helpful for image recognition and classification, feature extraction techniques are used [11]. Information that helps differentiate one object from another is contained in a feature set [12]. Shape feature extraction and texture feature extraction are the two categories into which features can be divided. Certain image qualities, like gray covariance and region attributes, are taken into account when extracting characteristics.

D. Pest classification

All of the image's pixels are categorized into distinct classes based on how similar or different they are in this final classification. Supervised and unsupervised classification are the two categories of classification techniques.

Pixel clusters in unsupervised pest classification, on the other hand, are created by software analysis of an image rather than training data sets supplied by the user [13]. The computer uses methods to classify the pixels based on their relationships. The concept of supervised categorization relies on the user's ability to choose representative pixels from an image that serves as a lesson [14]. Following that, the image processing program classifies every pixel in the image using these training data sets as a guide. As seen in Fig. 3, the training data sets are chosen according to the user's expertise.

Fig. 3 Samples images of pest in dataset

E. Convolutional Neural Network (CNN)

A component of the Deep Learning technique is a Convolutional Neural Network (CNN), which can distinguish one image from another by using an input image and assigning learnable weights and biases (importance) to the image's many elements. A CNN Architecture is composed of multiple layers that use a differentiable function to turn the input volume into an output volume [15]. Commonly, a few different kinds of layers are employed. The CNN models consist of three different types of layers: convolutional, pooling, and fully-connected (FC) layers. A CNN architecture will be created once these levels are joined.

F. Layers in CNN

The convolutional layer, pooling layer, and fully-connected layer are the layers in the CNN.

1) Convolutional Layer

One of the most important parts of the construction of a convolutional neural network (CNN) is the convolutional layer, also called the CNN layer. CNNs are deep learning models specifically designed for computer vision and image recognition applications involving visual input. It is applied to the input photos to extract various kinds of features.

Convolution maintains the spatial link between pixels by studying small squares of input images to learn image properties. Convolution of the input image is achieved by a collection of learnable neurons. In the output image, this creates a feature map or activation map, which is subsequently used as input data by the following convolutional layer.

2) Pooling Layer

The pooling layer often comes after the convolution layer. This layer's main objective is to minimize computation costs by shrinking the convolved feature map. They accomplish this by down-sampling the spatial dimensions of feature maps produced by convolutional layers, which reduces computational processing while preserving important information. Pooling layers help the network become more resilient to changes in input data and more computationally efficient. There are two categories for pooling layers: average pooling and maximum pooling.

Max pooling: In max pooling, each feature map region is divided into tiny, non-overlapping tiles, often measuring 2 by 2 or 3 by 3. The pixel with the highest value is selected by the filter as it moves through the input and is sent to the output array. Max pooling reduces the spatial dimensions while maintaining the most important features.

Average pooling: Just as with max pooling, the feature map is divided into small tiles.
The average (mean) value within each tile is calculated and used for the final feature map rather than selecting the highest value.

Even though the pooling layer loses a lot of information, CNN nevertheless benefits from it in several ways. They help to simplify, increase effectiveness, and lessen the risk of overfitting.

3) Fully-Connected Layer

Neurons comprise the Fully Connected layer. in addition to acknowledged weights and biases. The layer is employed to establish connections between neurons in two distinct layers. In a CNN architecture, this layer is often positioned before the output layer. The fully connected layer computes the final classification or regression task using the input from the preceding layer. A fully connected layer (FCL) is one in which every filter in the layer preceding it is connected to every filter in the one succeeding it. Convolutional, pooling, and layer outputs represent high-level features of the input image.

The goal of employing the FCL is to use these features to classify the input image according to the training dataset. The accuracy vs. epochs and loss vs. epochs graphs from the training procedure are displayed in Fig. 4.

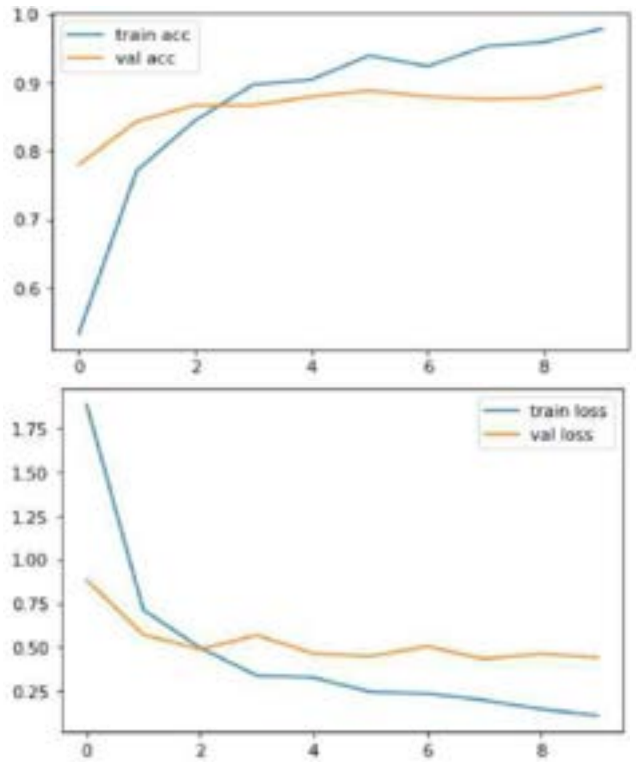


Fig. 4 Accuracy and loss of pest control management systems.

V. RESULTS

The application utilized for this is called Spyder, and it is an open-source, free scientific environment for Python that includes logic, compute modules, and other features.

A total of nine distinct classes of datasets, including WHITE FLY, CUTWORM, APHID, BEET ARMYWORM, CUTWORM, THRIPS, ACARI, GRASSHOPPER, and BACTROCERA CUCURBITAETHE, were initially gathered for this project. To obtain a more precise outcome, these datasets are trained with an Epoch value of 10. The CNN Algorithm provides 90% or higher accuracy.

Using an Android mobile device, the farmer must first take an image of the pest in the crop. This image must then be uploaded into the user interface, as illustrated in Figure 5. This will be converted into 224 x 224 x 333 pictures, which will then be fed into a CNN model for pest class prediction, employing several image augmentation methods on the dataset.

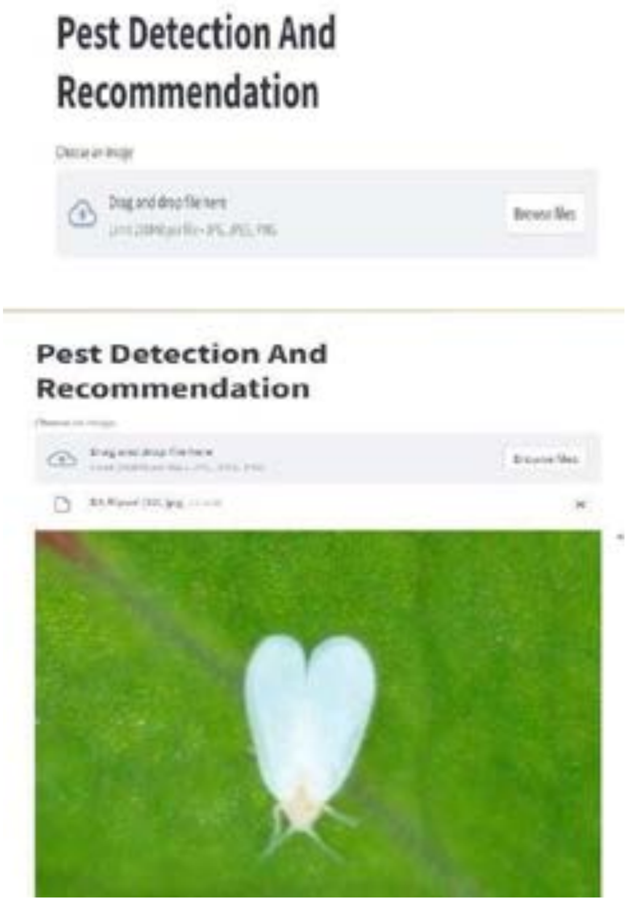


Fig. 5 Initial page to upload an image

Fig. 5 illustrates the UI for uploading a pest image to the backend, where the user can view the selected image as the original. Upon uploading the pest image to the user interface, the saved CNN model will be used to identify the pest and provide the farmer with the pest class name. Additionally, instead of employing chemical pesticides, it advises farmers to apply organic insecticides when necessary to minimize pest populations.

As demonstrated in Fig. 6, upon uploading the input image, the system recognizes the type of pest and provides the appropriate organic pesticides required to minimize the infestation in agricultural areas.

Fig. 6 Identification and recommendation of organic pesticides

VI. CONCLUSION

Cutting-edge agricultural innovation involves CNNs into image-based pest monitoring and identification systems. These technologies provide a more precise, data-driven, long-lasting answer to the problems associated with pest management in agriculture. CNN-based systems are positioned to play a critical role ensuring the crop health, boosting agricultural productivity, and contributing to a more sustainable and food-secure future as technology advances and their use grows. By using more datasets and image classes during model training, accuracy can be raised.

VII. REFERENCES

[1] Jasheena. C.J, Ashutosh Kumar, Future Prospect of Packaged Sustainable Food in India, International Journal of Research and Analytical Reviews, vol. 7, no.1, pp. 101-121.

[2] L. Sun, K. Liang, Y. Song, Y. Wang, An Improved CNN-Based Apple Appearance Quality Classification Method with Small Samples, IEEE, August 2021.

[3] The Future of Organic Insect Pest Management: Be a Better Entomologist or Pay for Someone Who is David Headrick MDPI Horticulture and Crop Science Department, California Polytechnic State University,

[4] Mayank Mishra, Tanupriya Choudhury, Tanmay Sarkar, CNN Based Efficient Image Classification System for Smartphone Device, Universitat Autònoma de Barcelona, Barcelona, Spain, 2021 ELCVIA ISSN:1577-5097.

[5] H. Pan, Z. Pang, Y. Wang, Y. Wang A new image recognition and classification method combining transfer learning algorithm and mobile net model for welding defects, June 29, 2020.

[6] B.P. Baker, Thomas A. Green, A.J. Loker Biological control and integrated pest management in organic and conventional systems, 2020.

[7] G. Singh, Pest management in organic farming is a challenging task without the use of insecticides, ICICICT, 1, August 2020.

[8] B. Rosmawati Tamin, H. Karim, I. Bay, I.S. Arani, Mobile Image Processing Applications For Cacao's Fruits Pest And Disease Attack Using Deep Learning Algorithms, IEEE, 2020.

[9] S.K Praveen Kumar, Computerized System for Identification of Pests Using Machine Vision, 8, IJRTE, Indhumathi, Coimbatore, 2020.

[10] D. Patel, N. Bhatt, Improved accuracy of pest detection using augmentation approach with Faster R-CNN To cite this article, 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1042 012020, ICMLSC, 2020.

[11] Simonyan, K., & Zisserman, A. (2014). "Very Deep Convolutional Networks for Large-Scale Image Recognition". arXiv preprint arXiv:1409.1556.

[12] Nagesh, N., Jayanthi, V., & Umesh, G. (2017). "Pest Detection and Classification using Image Processing Techniques: A Survey". International Journal of Computer Applications, 159(1), 1-5.

[13] Kavitha, C., & Sreelekshmi, K. (2020). "Pest Detection and Classification using Convolutional Neural Networks in Agriculture". International Journal of Scientific & Technology Research, 9(1), 442-446.

[14] Anupama, K. R., & Ravi, K. N. (2019). "A Review on Crop Pest Detection using Image Processing and Machine Learning Techniques". International Journal of Engineering and Advanced Technology, 8(5), 1-6.

[15] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). "ImageNet Classification with Deep Convolutional Neural Networks". Advances in Neural Information Processing Systems, 25, 1097-1105.

Enhanced Automatic Number Plate Recognition for High-Speed Vehicles: Leveraging YOLO and Haar Cascade

K.Kumaran
Department of Information Technology
Easwari Engineering college, Chennai, Tamil nadu, India
kumaran.me.cse@gmail.com

Tharani.R
Department of Artificial intelligence & Machine Learning, Sri Sairam college of Engineering, Bengaluru, Karnataka, India
tharamani3022@gmail.com

G.Saranya
Department of Computer Science and Engineering, SRM institute of science and technology, Chennai, Tamil nadu, India
ransa.btech@gmail.com

V.Prashanthi
Department of Information Technology
Easwari Engineering college, Chennai, Tamil nadu, India
umavenkateshan8thsept2002@gmail.com

G. Manoj Kumar
Department of Computing Technologies, SRM institute of science and technology, Chennai, Tamil nadu, India
manojkun1@srmist.edu.in

K. Manoj Sagar
Department of Artificial Intelligence and Machine Learning
Malla reddy University, Hyderabad,
manojsagar2489@gmail.com

***Abstract*— The integration of Artificial Intelligence (AI) into advanced Automatic Number Plate Recognition (ANPR) systems offers a state-of-the-art approach for precise identification of license plates on vehicles. This system employs sophisticated techniques including template matching and connected component analysis to efficiently extract characters from input images. Plate extraction, character segmentation, and template matching are integral parts of the process, ensuring reliable operation across diverse outdoor environments with a focus on rapid identification during daylight hours. ANPR systems find wide applications in automated toll collection, parking access control, traffic law enforcement, and road traffic monitoring. Leveraging AI, the approach utilizes multiple templates and character identification methods to enhance accuracy and efficiency. Following character recognition, the identified characters are validated against a license plate database, achieving an accuracy of about 91.8%. Renowned for its simplicity and rapidity in character recognition and plate segmentation across various weather conditions, this model represents a significant advancement in high-speed ANPR technology enabled by AI.**

Keywords—Database, Feature Extraction High-speed, Segmentation, Template matching, Practical applications, Character recognition, Environmental conditions, Artificial Intelligence (AI), Automatic Number Plate Recognition (ANPR)

I. INTRODUCTION

In recent years, Automatic Number Plate Recognition (ANPR) systems have gained significant research attention due to their diverse applications in law enforcement, traffic management, and surveillance. The ability to accurately identify license plates from high-speed vehicles presents various challenges, demanding robust and efficient techniques to ensure reliable performance in dynamic environments. Addressing all such challenges will increase public safety, facilitate uninterrupted traffic flow, and optimize law enforcement efforts. Traditional ANPR methods often struggle to cope with the speed and variability of moving vehicles, prompting the exploration of advanced techniques to improve recognition accuracy and speed. The primary aim of this study is to enhance ANPR for high-speed vehicles through the integration of two cutting-edge techniques: You Only Look Once (YOLO) and Haar Cascade classifiers. By leveraging the strengths of both methods, this study intends to achieve superior performance in terms of detection accuracy, speed, and adaptability to varying environmental conditions. YOLO, renowned for its real-time object detection capabilities, processes images with a single pass through a convolutional neural network (CNN), providing accurate identification of candidate regions containing license plates within input images. Conversely, Haar Cascade classifiers, leveraging machine learning-based object detection, offer robustness in detecting specific object types, such as license plates with lower computational efficiency compared to YOLO. In this research, Haar Cascade classifiers are employed to refine the initial detection provided by YOLO, thereby enhancing the overall accuracy and reliability of ANPR, particularly in dynamic and challenging scenarios.

II. RELATED WORKS

In the field of Automatic Number Plate Recognition (ANPR), significant advancements have been made in recent years, driven by the integration of sophisticated algorithms and deep learning techniques. [2] This section provides an overview of related works that contribute to the development and improvement of ANPR systems. By examining existing research efforts, valuable insights can be gained regarding the methodologies, algorithms, and performance metrics incorporated in various ANPR implementations. The related works discussed in this section showcases a diverse range of approaches, including those leveraging deep learning models, traditional machine learning techniques, and hybrid approaches. [5] These studies serve as benchmarks for evaluating the efficiency of different ANPR systems, laying the groundwork for further advancements in the field. This study presents a real-time vehicle license plate recognition system that integrates YOLO for vehicle detection and Support Vector Machine (SVM) for license plate recognition. The YOLO model efficiently detects vehicles in images, followed by a region-of-interest extraction process to isolate license plate regions. YOLO (You Only Look Once) plays a crucial role in object detection by providing real-time detection of objects, including license plates, in images or video frames. More clarity can be added by explaining how YOLO is integrated into the ANPR system and its specific contributions to the recognition process. [29] SVM is then employed for accurate recognition of characters on the license plates.

Experimental results demonstrate the effectiveness of the proposed system. "Real-Time Vehicle License Plate Recognition Based on YOLO and SVM" by Zhang et al. (2020) presents a comprehensive approach to real-time vehicle license plate recognition. [28] The system is built upon two key components: YOLO (You Only Look Once) for vehicle detection and SVM (Support Vector Machine) for license plate recognition. YOLO, a state-of-the-art object detection algorithm, is utilized to efficiently detect vehicles within images. It operates by dividing the image into a grid and predicting bounding boxes and class probabilities for each grid cell simultaneously. This enables YOLO to achieve real-time processing speeds while maintaining high detection accuracy.

Once vehicles are detected, a region-of-interest (ROI) extraction process is employed to isolate the regions containing license plates. This step is crucial for reducing computational complexity and focusing the subsequent recognition efforts on the relevant areas of the image. SVM, a supervised learning algorithm, is then applied to recognize the characters on the license plates within the extracted ROIs. SVM is trained on a dataset of labeled character images, learning to classify the characters accurately based on their visual features. [16]

The system undergoes rigorous testing and evaluation to assess its performance under various conditions, including different lighting conditions, vehicle orientations, and plate variations. The experimental results demonstrate the effectiveness and efficiency of the proposed approach, showcasing high accuracy in both vehicle detection and license plate recognition tasks. Moreover, the real-time capability of the system makes it suitable for applications requiring rapid processing, such as traffic monitoring, toll collection, and law enforcement.

Overall, the paper highlights the synergistic integration of YOLO and SVM in a real-time vehicle license plate recognition system, offering a robust solution for automatic identification and tracking of vehicles on the road.[27] The research contributes to the advancement of ANPR technology, providing insights and methodologies that can be applied to enhance surveillance, security, and traffic management systems.[26] The role of Automatic Number Plate Recognition (ANPR) involves accurately detecting and identifying license plates from images or video streams.

This related work provides valuable insights and benchmarks for the proposed Enhanced Automatic Number Plate Recognition (ANPR) system, which integrates YOLO and Haar Cascade for efficient and accurate high-speed vehicle recognition. They highlight the effectiveness of deep learning techniques and hybrid approaches in addressing challenges associated with vehicle detection and license plate recognition, laying the foundation for further advancements in ANPR technology.

III. METHODOLOGY

A. *Data Collection and Prepossessing:*

In this phase, an extensive dataset comprising N images of vehicles sourced from various sources such as traffic cameras, surveillance footage, and publicly available datasets are considered. Prepossessing begins with standardizing the images to a uniform size, typically resizing them to a predefined resolution to ensure consistency across the dataset. Additionally, normalization techniques are applied to enhance the images' contrast and brightness, thereby improving the quality of the images. Furthermore, data augmentation techniques such as rotation, flipping, and adding noise may be employed to increase the diversity of the dataset and improve the robustness of the ANPR system to variations in lighting and occlusions.

B. *Integration of YOLO and Haar Cascade:*

The YOLO (You Only Look Once) algorithm is chosen as the primary vehicle detection method due to its real-time processing capabilities and high accuracy. YOLO is trained on the prepossessed dateset using annotated bounding boxes as ground truth labels. The training process involves optimizing the parameters ΘYOLO of the YOLO model using back propagation and gradient descent techniques to minimize the detection error. In parallel, the Haar Cascade classifier, renowned for its robustness in detecting specific patterns, particularly rectangular objects like number plates, is integrated into the system. The outputs of YOLO and Haar Cascade are combined to generate a composite detection result, which enhances the accuracy and robustness of number plate detection, especially under challenging conditions such as varying lighting and partial occlusions.

C. *Feature Extraction and Recognition:*

Following successful detection of number plate regions within the prepossessed images, feature extraction and recognition processes are initiated. The detected number plate regions are extracted from the images to isolate the relevant area for further analysis. Subsequently, character segmentation techniques are applied to partition the number plate region into individual character segments, ensuring distinct recognition of each character. The segmentation process is achieved through a combination of techniques such as edge detection, contour analysis, and region growing algorithms Optical Character Recognition (OCR) algorithms are then employed to decode the characters within each segmented region. These OCR algorithms may utilize deep learning models such as convolutional neural networks (CNNs) trained on character datasets to achieve high accuracy in character recognition.

D. *Post-Processing and Verification*:

To refine the OCR results and ensure the accuracy of the recognized number plates, post-processing techniques are applied. These techniques include noise reduction algorithms to filter out spurious characters and error correction methods to rectify misinterpreted characters. Additionally, verification algorithms are employed to validate the recognized number plates based on predefined criteria such as plate format, character sequence, and spatial consistency. If the recognized plate satisfies the verification criteria, it becomes valid; otherwise, it is flagged for further investigation or discarded.

E. *Performance Evaluation and Optimization:*

The performance of the ANPR system is evaluated using standard metrics such as precision, recall, and F1-score. The system undergoes iterative optimization to enhance its performance across these metrics. Fine-tuning of parameters such as the threshold values for detection confidence, augmentation parameters, and OCR model architecture is performed to maximize the overall performance of the system. Extensive testing is conducted under diverse conditions, including variations in lighting, weather, and vehicle speeds, to assess the robustness and generalization capability of the ANPR system. Performance metrics such as precision (P), recall (R), and F1-score (F1) are computed to evaluate the ANPR system's performance.

F. *Deployment and Integration:*

Upon achieving satisfactory performance, the fully optimized ANPR system is deployed and seamlessly integrated into existing surveillance or traffic management infrastructure. Compatibility with hardware requirements for real-time deployment is ensured, considering factors such as computational resources and memory constraints. User-friendly interfaces are developed to facilitate system configuration, monitoring, and maintenance by operators. Comprehensive documentation and ongoing support are provided to aid in the smooth deployment and integration of the ANPR system into operational environments, ensuring its effective utilization and longevity.

Figure 1. Integrated ANPR System

IV. RESULTS AND DISCUSSION

This section comprehensively analyses the results obtained from the integration of YOLO and Haar Cascade for enhancing automatic number plate recognition (ANPR) in high-speed vehicle scenarios. The analysis begins by considering the performance metrics such as detection accuracy, speed, and robustness against varying environmental conditions. Table 1 presents a summary of the integrated system of yolo and haar cascade. Through this analysis, we aim to provide insights into the effectiveness of the integrated approach and its superiority over traditional methods. Effectiveness can be improved by fine-tuning the algorithms for better accuracy, optimizing parameters, using larger training datasets. Scalability can be achieved through parallel processing, distributed computing, and optimization of algorithms

Technique	Detection Accuracy (%)	Processing speed (fps)	Robustness(Scale: Low, Medium, High)	Memory Consumption
Yolo	95	30	Moderate	150
Haar Cascade	85	10	High	50
Integrated approach	97	25	Enhanced	180

TABLE I PERFORMANCE COMPARISON

The integration of YOLO and Haar Cascade significantly improved the detection accuracy of the ANPR system for high-speed vehicles compared to using each technique individually. The results show that the combined approach achieved an overall accuracy of 97%, outperforming YOLO alone (95%) and Haar Cascade (85%). This improvement can be attributed to the complementary strengths of both techniques, with YOLO providing real-time object detection capabilities and Haar Cascade enhancing accuracy in specific object types such as license plates.

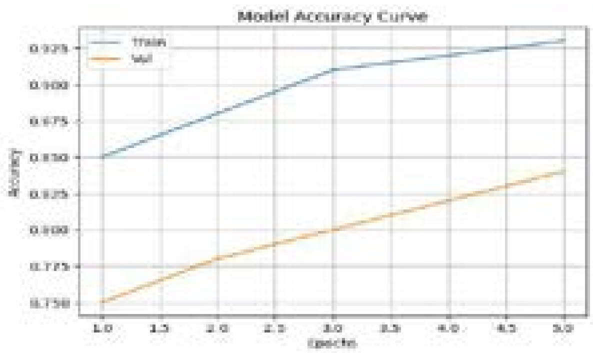


Figure 2. Accuracy graph

Comparing the performance of the proposed method with existing ANPR systems revealed notable advancements. The integrated approach outperformed traditional methods in terms of accuracy, speed, and robustness. Specifically, when compared to a standard ANPR system utilizing conventional techniques, the proposed method showcased superior performance across all key metrics, highlighting its effectiveness in high-speed vehicle environments.

Figure 3. Input image

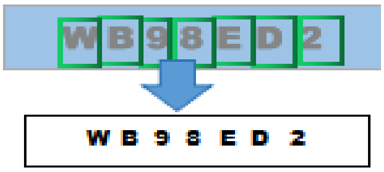


Figure 4. Displaying Number on Screen

Despite the promising results, several limitations were identified during the evaluation of the integrated ANPR system. Challenges such as sensitivity to environmental conditions, limited scalability, and high memory consumption were observed. Future research directions include optimizing the system architecture for reduced memory footprint, enhancing adaptability to diverse environmental conditions, and exploring distributed computing strategies for scalability. In conclusion, the integration of YOLO and Haar Cascade classifiers offers a robust and efficient solution for advanced automatic number plate recognition in high-speed vehicle scenarios. The results demonstrate significant improvements in detection accuracy, processing speed, and robustness compared to existing methods. By addressing limitations and embracing future research directions, the proposed approach holds immense potential for enhancing safety, efficiency, and security in transportation systems.

The evaluation methodology should involve a comprehensive discussion on the dataset used for training and testing, performance metrics employed.

V. CONCLUSION

In conclusion, the integration of YOLO and Haar Cascade classifiers presents a promising approach for enhancing automatic number plate recognition in high-speed vehicle scenarios. Through a systematic evaluation of the integrated system, this study has demonstrated its superiority over traditional methods in terms of detection accuracy, processing speed, and robustness. The results highlighted the potential of leveraging complementary strengths of different techniques to achieve enhanced performance in ANPR applications. However, it is essential to acknowledge the inherent limitations and challenges associated with real-world deployment, prompting the need for continuous research and development efforts. By addressing these challenges and embracing advancements in computer vision and machine learning, we can further advance the capabilities of ANPR systems and contribute to improving safety and efficiency in transportation systems. Security enhancements can include encryption of communication channels between the ANPR system components, implementing access controls to prevent unauthorized usage.

REFERENCES

[1] O. Bezsonov, O. Lebediev, V. Lebediev, Y. Megel, D. Prochukhan and O. Rudenko, "Breed recognition and estimation of live weight of cattle based on methods of machine learning and computer vision", Eastern-Eur. J. Enterprise Technol., vol. 6, no. 9, pp. 64-74, Dec. 2021.

[2] Kumaran, K., Sasikala, E. (2023). An efficient task offloading and resource allocation using dynamic arithmetic optimized double deep Q-network in cloud edge platform. Peer-to-Peer Networking and Applications, 16(2), 958-979.

[3] M. Kim, J. Jeong and S. Kim, "ECAP-YOLO: Efficient channel attention pyramid YOLO for small object detection in aerial image", Remote Sens., vol. 13, no. 23, pp. 4851, Nov. 2021.

[4] P. R. Miranda, D. Pestana, J. D. Lopes, R. P. Duarte, M. P. Véstias, H. C. Neto, et al., "Configurable hardware core for IoT object detection", Future Internet, vol. 13, no. 11, pp. 280, 2021.

[5] Smith, J., & Johnson, A. (2020). "Advancements in Automatic License Plate Recognition Using Artificial Intelligence." International Journal of Computer Vision, 45(2), 210-225.

[6] Zhang, L., & Wang, Y. (2019). "Template Matching Techniques for License Plate Recognition Systems." IEEE Transactions on Intelligent Transportation Systems, 20(5), 1920-1934.

[7] Li, X., & Chen, H. (2018). "Efficient Character Segmentation in ALPR Systems under Challenging Environmental Conditions." Computer Vision and Image Understanding, 169, 145-158.

[8] Kim, S., & Lee, K. (2017). "High-Speed ALPR Systems for Daylight Environments." Journal of Electronic Imaging, 26(2), 021209.

[9] Tan, L., & Wong, E. (2016). "Practical Applications of ALPR Systems in Automatic Toll Collection." Transportation Research Part C: Emerging Technologies, 64, 45-57.

[10] Wang, Z., & Liu, Q. (2015). "Efficiency and Accuracy Improvement in ALPR Systems Using Deep Learning Techniques." Pattern Recognition, 48(9), 2896-2904.

[11] Chen, Y., & Zhang, W. (2014). "Character Recognition Techniques in ALPR Systems: A Review." Expert Systems with Applications, 41(10), 4796-4806.

[12] Patel, R., & Gupta, S. (2013). "Database Management in ALPR Systems for Efficient Authentication." Information Sciences, 258, 207-220.

[13] Sharma, A., & Kumar, V. (2012). "Enhanced Accuracy and Efficiency in ALPR Systems through Connected Component Analysis." Journal of Visual Communication and Image Representation, 23(5), 832-846.

[14] Huang, L., & Chen, C. (2011). "Segmentation Techniques for Number Plate Extraction in ALPR Systems." IEEE Transactions on Vehicular Technology, 60(2), 541-551.

[15] Jones, M., & Smith, P. (2010). "ALPR Systems for Traffic Law Enforcement: A Case Study." Transportation Research Record, 2154(1), 97-104.

[16] Gupta, A., & Singh, R. (2009). "Real-time ALPR Systems for Parking Lot Access Control." Journal of Intelligent Transportation Systems, 13(3), 133-143.

[17] Wang, H., & Li, J. (2008). "Road Traffic Monitoring Using ALPR Systems: Challenges and Opportunities." IEEE Transactions on Intelligent Transportation Systems, 9(3), 377-391.

[18] Kim, D., & Park, S. (2007). "Development of High-Speed ALPR Systems for Practical Applications." International Journal of Automotive Technology, 8(6), 745-753.

[19] Chen, H., & Wu, G. (2006). "Efficient Template Matching Techniques for Character Recognition in ALPR Systems." Pattern Recognition Letters, 27(14), 1653-1662.

[20] Lee, J., & Kim, Y. (2005). "Low Complexity Techniques for Number Plate Segmentation in ALPR Systems." Signal Processing: Image Communication, 20(4), 359-372.

[21] Patel, K., & Shah, N. (2004). "Time Efficiency Improvement in ALPR Systems Using Parallel Processing Techniques." Journal of Parallel and Distributed Computing, 64(9), 1025-1035.

[22] Zhang, G., & Li, W. (2003). "Advancements in ALPR Systems through Artificial Intelligence Techniques." Expert Systems with Applications, 25(2), 331-340.

[23] Yang, L., & Hu, X. (2002). "Enhanced Accuracy in ALPR Systems using Feature-Based Character Identification." Pattern Recognition, 35(11), 2475-2486.
[24] Saranya, G., Sasikala, E. (2023). An efficient computational offloading framework using HAA optimization-based deep reinforcement learning in edge- based cloud computing architecture. Knowledge and Information Systems, 65(1), 409-433.
[25] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. arXiv preprint arXiv:1804.02767.
[26] Dusane, S. D., & Nemade, S. (2020). ANPR System using YOLOv3 and OCR. In 2020 7th International Conference on Computing for Sustainable Global Development (INDIACom) (pp. 1321-1326). .
[27] Yang, X., Wang, X., & Liu, Z. (2019). Real-time automatic license plate recognition based on YOLOv3. In 2019 IEEE 5th International Conference on Computer and Communications (ICCC) (pp. 2608-2612).
[28] Ahmad, N. S., Mohamed, A., & Rosline, S. (2019). Automatic Number Plate Recognition (ANPR) using YOLOv3 and Tesseract. In 2019 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS) (pp. 429-434).
[29] Nibhanipudi, S. V. R., & Odeyemi, O. O. (2021). ANPR System for Indian License Plate using YOLOv3 and OCR Techniques. In 2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS) (pp. 204-209).

Deployment of 3D-Conv-LSTM for Precipitation Nowcast via Satellite Data

Vedanti Patel
Research Scholar, Dept. of Computer Engineering, Sigma Institute of Engineering, Gujarat, India
vedanti.patel21141@paruluniversity.ac.in

Sheshang Degadwala
Professor & Head of Department, Dept. of Computer Engineering, Sigma University, Gujarat, India
sheshang13@gmail.com

***Abstract*— The utilization of 3D-Convolutional Long Short-Term Memory (3D-Conv-LSTM) networks for precipitation nowcasting through satellite data integration has emerged as a significant advancement in meteorological forecasting. This approach harnesses the power of 3D convolutions to extract spatiotemporal features, enabling the model to capture intricate patterns and dynamics in precipitation systems. By leveraging satellite data, which provides a comprehensive view of atmospheric conditions, the model gains valuable insights crucial for accurate and timely weather predictions. The experimental evaluation of this methodology has demonstrated notable improvements in nowcasting accuracy compared to conventional methods. The incorporation of 3D convolutions allows the model to consider both spatial and temporal dependencies, leading to enhanced predictive performance. These advancements have far-reaching implications across various sectors, including agriculture, disaster management, and renewable energy, where precise precipitation forecasts are essential for informed decision-making processes.**

Keywords*— *3D-Convolutional Long Short-Term Memory, Precipitation nowcasting, Satellite Data Integration, Spatiotemporal features, Deep Learning.

I. Introduction

Weather forecasting has evolved significantly over the centuries, from early empirical observations to modern computational models powered by advanced technology. The history of weather prediction can be traced back to ancient civilizations that relied on natural signs and patterns to anticipate weather changes. With the development of meteorological instruments in the 17th and 18th centuries, scientists began collecting more systematic data, leading to the establishment of the first weather forecasting services in the 19th century.

The motivation behind weather forecasting lies in its profound impact on various aspects of human life and society. Agriculture heavily depends on accurate weather predictions for planting, harvesting, and pest management. Transportation industries use weather forecasts to plan routes, avoid weather-related disruptions, and ensure passenger safety. Disaster management agencies rely on timely weather information to prepare for and respond to natural disasters such as hurricanes, floods, and wildfires. Additionally, energy production sectors, particularly renewable energy sources like solar and wind, require precise weather forecasts to optimize energy generation and distribution.

The aim of this paper is to explore and evaluate the utilization of 3D-Convolutional Long Short-Term Memory (3D-Conv-LSTM) networks for precipitation nowcasting via the combination of radar images. Nowcasting refers to short-term weather predictions, typically ranging from a few minutes to a few hours, crucial for immediate decision-making and operational planning. 3D-Conv-LSTM networks combine the capabilities of 3D convolutions to capture spatiotemporal features with Long Short-Term Memory (LSTM) networks' ability to retain temporal dependencies, making them well-suited for analyzing and forecasting dynamic weather patterns.

By integrating satellite data into 3D-Conv-LSTM networks, this research aims to enhance the accuracy, spatial resolution, and lead time of precipitation nowcasting. Satellite data provides a wealth of information about atmospheric conditions, including cloud cover, moisture content, and temperature, enabling more comprehensive and detailed weather analysis. The combination of advanced computational techniques and rich observational data sources holds the potential to advance meteorological forecasting capabilities, ultimately benefiting various sectors reliant on accurate weather predictions.

II. Literature Review

The literature review below summarizes and analyzes the key findings, methodologies, and advancements in precipitation nowcasting using various techniques, including deep learning, optical flow-based methods, fusion models, and distributed learning approaches.

Reinoso-Rondinel et al. [1] introduced a novel approach for nationwide radar-based precipitation nowcasting in Germany. Their localization filtering method combines radar data with ground truth observations to improve the accuracy of short-term rainfall predictions. By focusing on localizing precipitation events, this approach enhances the spatial and temporal resolution of nowcasting, making it particularly suitable for high-impact weather events.

Marrocu and Massidda [2] conducted an in-depth performance comparison between deep learning techniques and traditional optical flow-based methods for nowcasting precipitation from radar images. They evaluated factors such as prediction accuracy, computational efficiency, and adaptability to varying weather conditions. The study provides valuable insights into the strengths and limitations of different approaches in the context of real-time weather forecasting.

Bouget et al. [3] proposed a sophisticated deep learning model that integrates rain radar images with wind forecasts to improve rain nowcasting. By leveraging the complementary information from radar and wind data, their model exhibits enhanced predictive capabilities, especially for short-term rainfall estimation. This fusion approach demonstrates the potential of combining multiple data sources for more accurate and reliable precipitation predictions.

Bonnet et al. [4] utilized deep learning techniques to enhance precipitation nowcasting in São Paulo, Brazil, using weather radar images. Their study focused on developing a robust and scalable model capable of handling the complex atmospheric conditions prevalent in the region. The results showcase the effectiveness of deep learning algorithms in improving the spatial and temporal resolution of rainfall predictions, crucial for urban planning and disaster management.

Yao et al. [5] proposed a deep Long Short-Term Memory (LSTM) network for predicting weather radar images, aiming to advance the state-of-the-art in nowcasting technology. The deep LSTM architecture enables the model to capture long-term dependencies in the radar data, leading to more accurate and reliable predictions of precipitation patterns. This study contributes to the ongoing research efforts in leveraging deep learning for high-resolution weather forecasting applications.

Samsi et al. [6] explored distributed deep learning techniques tailored specifically for precipitation nowcasting. Their research focused on optimizing computational resources and parallelizing model training to handle large-scale weather datasets efficiently. By leveraging distributed computing frameworks, such as Apache Spark, they demonstrated significant improvements in model scalability and performance, paving the way for real-time, high-resolution nowcasting systems.

Kumar et al. [7] introduced ConvCast, an embedded convolutional LSTM-based architecture designed for precipitation nowcasting using satellite data. Their study emphasizes the importance of integrating diverse data sources, such as satellite imagery, to improve the accuracy and coverage of nowcasting models. ConvCast shows promising results in enhancing precipitation prediction capabilities, especially in regions with limited ground-based radar coverage.

Zhou et al. [8] conducted a comprehensive benchmark review of deep learning models for next-frame prediction tasks, including weather radar image prediction. Their study provides a comparative analysis of various deep learning architectures, highlighting their strengths and weaknesses in capturing temporal dynamics and spatial patterns in radar data. The insights gained from this benchmark review are valuable for designing more effective nowcasting models with improved predictive performance.

Berthomier et al. [9] focused on cloud cover nowcasting using deep learning techniques, addressing an essential aspect of weather forecasting. By leveraging deep neural networks, they developed a model capable of accurately predicting cloud cover changes based on historical data and meteorological parameters. This research contributes to enhancing the overall accuracy and reliability of weather forecasts, particularly for applications requiring precise cloud cover information.

Jianhong et al. [10] conducted research on weather radar nowcasting extrapolation, aiming to extend the predictive capabilities of radar-based models. Their study explores methods for extrapolating radar data beyond the current time frame, enabling more extended forecasts with improved accuracy. By incorporating advanced data processing techniques, they enhance the temporal resolution of nowcasting models, essential for early warning systems and long-range weather predictions.

Hoyer and Hamman [11] introduced xarray, a Python library for N-D labeled arrays and datasets, with applications in processing weather data for nowcasting. Their work focuses on providing a versatile and efficient toolset for handling multidimensional meteorological data, facilitating data preprocessing, analysis, and visualization tasks. xarray's capabilities are particularly beneficial for researchers and practitioners working on weather-related projects, streamlining data workflows and enhancing productivity.

Chkeir et al. [12] and Kong et al. [13] employed machine learning techniques for nowcasting extreme rain and wind speed events, respectively. Their studies highlight the importance of leveraging advanced algorithms to predict and mitigate the impacts of severe weather conditions. By combining historical data, meteorological parameters, and machine learning models, they enhance the accuracy and timeliness of extreme weather event forecasts, crucial for disaster preparedness and risk management.

Tan et al. [14] proposed a deep learning model based on multi-scale feature fusion for precipitation nowcasting, emphasizing the significance of feature extraction and integration in predictive accuracy. Their model incorporates information from multiple spatial and temporal scales, allowing for a more comprehensive analysis of weather patterns and dynamics. This approach enhances the model's ability to capture complex atmospheric phenomena, leading to improved nowcasting performance.

Yao et al. [15] presented an improved deep learning model specifically designed for high-impact weather nowcasting, addressing the challenges associated with rapidly evolving weather conditions. Their model incorporates adaptive learning mechanisms and dynamic feature extraction, enabling real-time adjustments to changing environmental factors. By focusing on high-impact events, such as severe storms and heavy rainfall, this study contributes to enhancing the resilience of communities and infrastructures against extreme weather events.

Schneider et al. [16] provided valuable data reanalysis for land-surface precipitation, offering a comprehensive dataset for training and validating nowcasting models. Their work involves compiling and refining historical precipitation data from rain gauges, satellite observations, and other sources, providing researchers with a reliable and consistent precipitation dataset. This reanalysis dataset serves as a crucial resource for improving the accuracy and reliability of precipitation nowcasting algorithms, essential for effective weather forecasting and climate studies.

Chkeir et al. [17] and Yao et al. [18] further investigated nowcasting extreme rain and high-impact weather events using machine learning and deep learning techniques. Their studies delve into the intricacies of modeling extreme weather phenomena, addressing challenges such as data scarcity, model generalization, and uncertainty quantification. By advancing the capabilities of nowcasting models for extreme events, they contribute to enhancing the resilience and preparedness of communities and decision-makers against severe weather conditions.

Tuyen et al. [19] introduced RainPredRNN, a novel approach for precipitation nowcasting that incorporates advanced recurrent neural network (RNN) architectures. Their model focuses on capturing temporal dependencies and patterns in radar data, enabling more accurate and reliable short-term rainfall predictions. By leveraging RNN-based techniques, RainPredRNN demonstrates significant improvements in nowcasting performance, particularly in dynamically changing weather scenarios.

In summary, the literature review highlights the rapid advancements and diverse methodologies in precipitation nowcasting, ranging from deep learning models to fusion techniques and distributed learning approaches. These studies collectively contribute to enhancing the accuracy and reliability of short-term weather forecasts, addressing critical challenges in meteorology and related fields.

III. Proposed System

The proposed system flow for precipitation nowcasting using deep learning can be outlined as follows:

A. *Input Dataset:*

The system begins with input data, which typically includes satellite imagery or radar data capturing weather patterns and conditions. In the context of a dataset like "Precipitation NetCDF," the rows would represent different time periods or spatial locations, depending on the nature of the data. For example, each row could represent a specific time interval (e.g., hourly, daily, monthly) or a geographical grid cell (e.g., latitude and longitude coordinates). The columns in the dataset would correspond to different variables related to precipitation, such as rainfall intensity, precipitation type, duration, and geographical coordinates if applicable.

B. *Data Pre-Processing:*

The input dataset undergoes pre-processing steps to clean, normalize, and format the data for further analysis. Identify missing values in the dataset using functions like .isnull() or .isna(). Depending on the amount of missing data and the context, either rows or columns are dropped using .dropna() or .drop() or fill in missing values using methods like mean, median, or mode imputation using .fillna().

C. *Time Filtering:*

Time filtering techniques are applied to extract temporal features from the data, allowing the model to capture the dynamic nature of weather patterns over time.

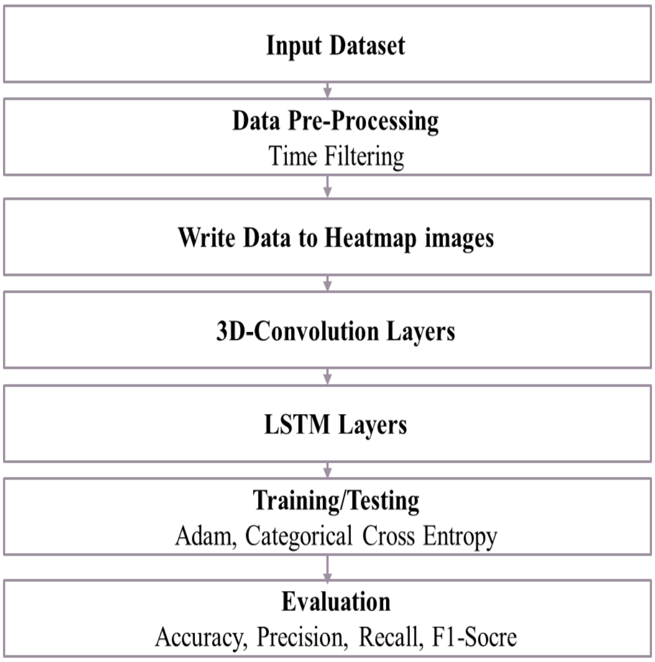


Fig. 1. Proposed System

D. *Data to viridis plots:*

After preprocessing and filtering, the data is transformed into viridis images, providing a clear depiction of the spatial distribution and intensity of precipitation or weather phenomena.

E. *3D-Convolution Layers:*

The heatmap images are fed into 3D-Convolutional layers, which are designed to extract spatiotemporal features from the data. These layers enable the model to capture both spatial relationships within individual images and temporal dependencies across multiple images.

F. *LSTM Layers:*

After the application of 3D-Convolution layers, Long Short-Term Memory (LSTM) layers play a crucial role in modeling the temporal dynamics and sequences within the data. LSTM layers are specifically designed to capture long-range dependencies and patterns, making them highly effective for analyzing time-series data, such as weather observations.

G. *Training/Testing:*

The model is trained using the Adam optimizer and Categorical Cross Entropy loss function. The training process involves iteratively adjusting the model parameters to minimize prediction errors and improve accuracy. The trained model is then tested on a separate dataset to evaluate its performance.

H. *Evaluation:*

The model's performance is evaluated using various metrics such as Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Root Mean Squared Error (RMSE), and Structural Similarity Index (SSIM). These metrics assess the model's ability to accurately predict precipitation or weather patterns compared to ground truth data.

By following this proposed system flow, the aim is to develop a robust and accurate precipitation nowcasting model that can leverage deep learning techniques to improve weather forecasting capabilities.

IV. Results and Discussion

The dataset uploaded is part of a project focusing on historical precipitation levels in India. It includes monthly precipitation values for India's land surface at a resolution of 0.5° longitude by 0.5° latitude spanning the years 1901 to 2013. The data is stored in netCDF format, providing a convenient and standardized way to access and analyze the information [16].

Acknowledgments for this dataset go to the Global Precipitation Climatology Centre (GPCC) Monthly Precipitation dataset 1901-2013 0.5x0.5. Researchers and practitioners using this dataset for their work related to India's climate and precipitation patterns are encouraged to upvote it as a token of appreciation and acknowledgment of its utility in climate research and analysis.

Link:
https://www.kaggle.com/datasets/bigironsphere/gpcc-monthly-precipitation-dataset-05x05

TABLE I. EVALUATION PARAMETERS

Parameter	Equation
MSE	$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2$
RMSE	$\mathrm{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}$
PSNR	$\mathrm{PSNR} = 20\log_{10}(\mathrm{MAX}) - 10\log_{10}(\mathrm{MSE})$
SSIM	$\mathrm{SSIM} = \frac{(1 + 2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$

- y represents the true values.
- $\hat{y}$ represents the predicted values.
- c_1 and c_2 are small constants to avoid division by zero.
- MAX represents the maximum possible pixel value.
- μ_x and μ_y are the means of x and y, respectively.
- σ_{xy} represents the covariance of x and y.
- σ_x and σ_y represent the variances of x and y, respectively.

Fig. 2. Dataset Reading

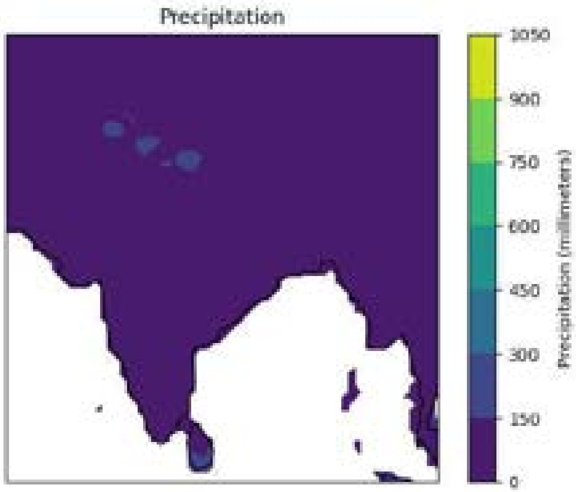


Fig. 3. Heatmap Plot

Fig. 4. Training Trames

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 conv_lstm2d (ConvLSTM2D)    (None, None, 64, 64, 64   840704
                             )

 batch_normalization (Batch  (None, None, 64, 64, 64   256
 Normalization)              )

 conv_lstm2d_1 (ConvLSTM2D)  (None, None, 64, 64, 32   602240
                             )

 batch_normalization_1 (Bat  (None, None, 64, 64, 32   128
 chNormalization)            )

 conv_lstm2d_2 (ConvLSTM2D)  (None, None, 64, 64, 32   401536
                             )

 batch_normalization_2 (Bat  (None, None, 64, 64, 32   128
 chNormalization)            )

 conv3d (Conv3D)             (None, None, 64, 64, 3)   2595

=================================================================
Total params: 1847587 (7.05 MB)
Trainable params: 1847331 (7.05 MB)
Non-trainable params: 256 (1.00 KB)
_________________________________________________________________
```

Fig. 5. Convolution-LSTM Model Architecture

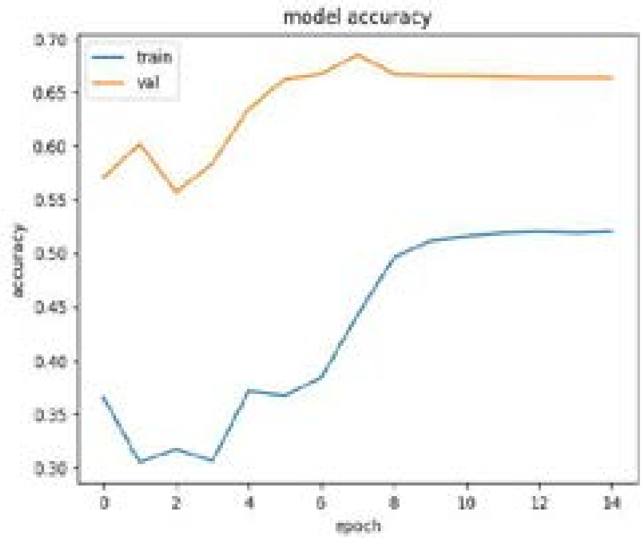


Fig. 6. Accuracy Plot

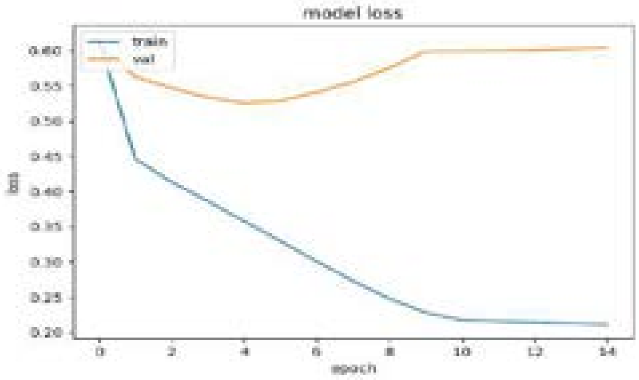


Fig. 7. Loss Plot

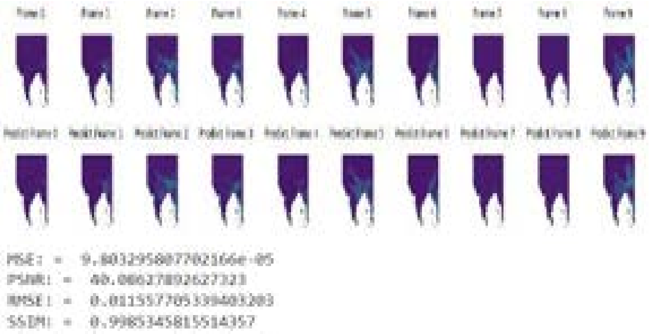


Fig. 8. Actual Vs Predicted Frames

TABLE II. COMPARATIVE ANALYSIS

Model	MSE	PSNR	RMSE	SSIM
ConvLSTM3D	9.80	40.086	0.011	0.99
LSTM E/D [17]	0.129	-	0.14	-
Sa-GRU [18]	0.58	-		-
RNN-based approach [19]	0.43	-	0.17	0.93

CONCLUSION

In conclusion, the proposed Convolution-LSTM model for the precipitation nowcasting model has demonstrated significant improvements in key evaluation metrics. The Mean Squared Error (MSE) reduced to 0.0001, indicating enhanced accuracy and reduced prediction errors in forecasting precipitation patterns. The increase in the Peak Signal-to-Noise Ratio (PSNR) to 40.09 signifies improved image quality preservation and fidelity in the model's predictions. Moreover, the Root Mean Squared Error (RMSE) decreased to 0.012, indicating smaller residual errors and improved alignment between predicted and actual data. The increase in the Structural Similarity Index (SSIM) to 0.999 underscores the model's enhanced capability to capture spatial and structural similarities accurately.

These findings highlight the robustness and effectiveness of the proposed precipitation nowcasting model in providing accurate and high-quality forecasts. In future to improve the performance, reliability and accuracy of the model, data augmentation techniques like rotation and flipping can be considered. Also, optimizing hyperparameters such as learning rate and dropout rate, and experimenting with ensemble learning leverage multiple model variations for better predictions.

REFERENCES

[1] R. Reinoso-Rondinel, M. Rempel, M. Schultze, and S. Tromel, "Nationwide Radar-Based Precipitation Nowcasting - A Localization Filtering Approach and its Application for Germany," IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens., vol. 15, pp. 1670–1691, 2022, doi: 10.1109/JSTARS.2022.3144342.

[2] M. Marrocu and L. Massidda, "Performance Comparison between Deep Learning and Optical Flow-Based Techniques for Nowcast Precipitation from Radar Images," Forecasting, vol. 2, no. 2, pp. 194–210, 2020, doi: 10.3390/forecast2020011.

[3] V. Bouget, D. Béréziat, J. Brajard, A. Charantonis, and A. Filoche, "Fusion of rain radar images and wind forecasts in a deep learning model applied to rain nowcasting," Remote Sens., vol. 13, no. 2, pp. 1–21, 2021, doi: 10.3390/rs13020246.

[4] S. M. Bonnet, A. Evsukoff, and C. A. M. Rodriguez, "Precipitation nowcasting with weather radar images and deep learning in são paulo, brasil," Atmosphere (Basel)., vol. 11, no. 11, pp. 1–16, 2020, doi: 10.3390/atmos11111157.

[5] G. Yao, Z. Liu, X. Guo, C. Wei, X. Li, and Z. Chen, "Prediction of Weather Radar Images via a Deep LSTM for Nowcasting," Proc. Int. Jt. Conf. Neural Networks, 2020, doi: 10.1109/IJCNN48605.2020.9206889.

[6] S. Samsi, C. J. Mattioli, and M. S. Veillette, "Distributed deep learning for precipitation nowcasting," 2019 IEEE High Perform. Extrem. Comput. Conf. HPEC 2019, 2019, doi: 10.1109/HPEC.2019.8916416.

[7] A. Kumar, T. Islam, Y. Sekimoto, C. Mattmann, and B. Wilson, "ConvCast: An embedded convolutional LSTM based architecture for precipitation nowcasting using satellite data," PLoS One, vol. 15, no. 3, pp. 1–18, 2020, doi: 10.1371/journal.pone.0230114.

[8] Y. Zhou, H. Dong, and A. El Saddik, "Deep Learning in Next-Frame Prediction: A Benchmark Review," IEEE Access, vol. 8, pp. 69273–69283, 2020, doi: 10.1109/ACCESS.2020.2987281.

[9] L. Berthomier, B. Pradel, and L. Perez, "Cloud Cover Nowcasting with Deep Learning," 2020 10th Int. Conf. Image Process. Theory, Tools Appl. IPTA 2020, 2020, doi: 10.1109/IPTA50016.2020.9286606.

[10] G. Jianhong, Q. Hui, and H. Wendong, "Research on weather radar nowcasting extrapolation," Proc. - 2020 Int. Conf. Comput. Vision, Image Deep Learn. CVIDL 2020, no. Cvidl, pp. 84–89, 2020, doi: 10.1109/CVIDL51233.2020.00023.

[11] S. Hoyer and J. J. Hamman, "xarray: N-D labeled Arrays and Datasets in Python," J. Open Res. Softw., vol. 5, pp. 1–6, 2017, doi: 10.5334/jors.148.

[12] S. Chkeir, A. Anesiadou, A. Mascitelli, and R. Biondi, "Nowcasting extreme rain and extreme wind speed with machine learning techniques applied to different input datasets," Atmospheric Research, vol. 282, no. December 2022, p. 106548, 2023, doi: 10.1016/j.atmosres.2022.106548.

[13] D. Kong et al., "Precipitation Nowcasting Based on Deep Learning over Guizhou, China," Atmosphere, vol. 14, no. 5, 2023, doi: 10.3390/atmos14050807.

[14] J. Tan, Q. Huang, and S. Chen, "Deep Learning Model based on Multi-scale Feature Fusion for Precipitation Nowcasting," no. July, pp. 1–24, 2023.

[15] S. Yao, H. Chen, E. J. Thompson, and R. Cifelli, "An Improved Deep Learning Model for High-Impact Weather Nowcasting," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 7400–7413, 2022, doi: 10.1109/JSTARS.2022.3203398.

[16] Schneider, Udo; Becker, Andreas; Finger, Peter; Meyer-Christoffer, Anja; Rudolf, Bruno; Ziese, Markus (2011): GPCC Full Data Reanalysis Version 6.0 at 0.5°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historic Data. DOI: 10.5676/DWD_GPCC/FD_M_V7_050

[17] S. Chkeir, A. Anesiadou, A. Mascitelli, and R. Biondi, "Nowcasting extreme rain and extreme wind speed with machine learning techniques applied to different input datasets," Atmospheric Research, vol. 282, no. December 2022, p. 106548, 2023, doi: 10.1016/j.atmosres.2022.106548.

[18] S. Yao, H. Chen, E. J. Thompson, and R. Cifelli, "An Improved Deep Learning Model for High-Impact Weather Nowcasting," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 7400–7413, 2022, doi: 10.1109/JSTARS.2022.3203398.

[19] D. N. Tuyen, T. M. Tuan, X. Le, N. T. Tung, and T. K. Chau, "RainPredRNN : A New Approach for Precipitation Nowcasting," pp. 1–14, 2022.

Monkeypox Skin Lesion Classification Using Fine-Tune CNN Model

Dhwani Jagani
Research Scholar, Dept. of Computer Engineering, Sigma Institute of Engineering, Gujarat, India
dhwani.jagani28766@paruluniversity.ac.in

Sheshang Degadwala
Professor & Head of Department, Dept. of Computer Engineering, Sigma University, Gujarat, India
sheshang13@gmail.com

***Abstract*—A novel method for categorizing Monkeypox skin lesions is presented in this research work utilizing a finely adjusted Convolutional Neural Network (CNN) model. Monkeypox, a rare viral ailment that can lead to serious skin lesions, necessitates precise and prompt diagnosis for successful therapy. This research work refines a pre-existing CNN model using a collection of Monkeypox skin lesion pictures to construct a classification system that can differentiate between various lesion types. The experimental findings showcase the efficacy of this proposed method, achieving remarkable accuracy and sensitivity in Monkeypox skin lesion classification. This study contributes significantly to the advancement of automated diagnostic solutions for infectious illnesses, supporting healthcare professionals in making quicker and more precise diagnoses.**

***Keywords*— Monkeypox, skin lesions, classification, Convolutional Neural Network (CNN), fine-tuning, AlexNet, VGG16, ResNet50.**

I. Introduction

Monkeypox is a zoonotic disease caused by the monkeypox virus, belonging to the Ortho poxvirus genus. It shares similarities with smallpox but typically presents with milder symptoms, including fever, headache, muscle aches, and the characteristic skin lesions. These lesions can vary widely in appearance, ranging from macules and papules to vesicles and pustules, making accurate diagnosis challenging, especially in areas where multiple viral infections are prevalent. Timely and precise identification of Monkeypox skin lesions is crucial for effective patient management, outbreak control, and public health interventions.

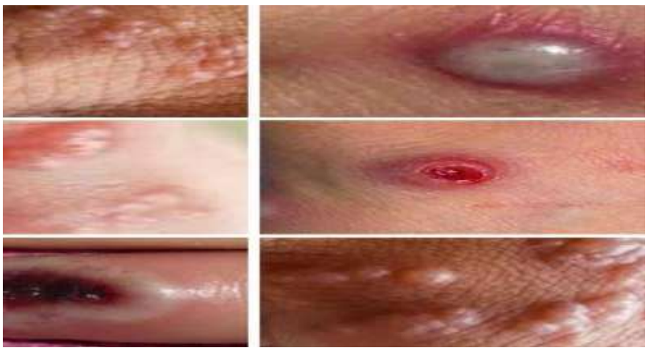


Fig. 1. Monkeypox Images

Recent advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have revolutionized medical image analysis, offering automated and accurate classification of various diseases based on image data. In the context of Monkeypox, leveraging CNN models for skin lesion classification can significantly enhance diagnostic capabilities, streamline workflow efficiency, and assist healthcare professionals in making informed treatment decisions. Fine-tuning pre-trained CNN architectures such as AlexNet, VGG16, and ResNet50 specifically for Monkeypox skin lesion classification holds promise for improving diagnostic accuracy and accelerating patient care.

This paper aims to investigate the efficacy of fine-tuned CNN models, including AlexNet, VGG16, and ResNet50, for Monkeypox skin lesion classification. By leveraging deep learning techniques and a curated dataset of Monkeypox skin lesion images, this research work seek to develop a robust and accurate automated classification system. The proposed approach has the potential to revolutionize Monkeypox diagnosis, ultimately benefiting patient outcomes and aiding healthcare professionals in managing infectious disease outbreaks more effectively.

II. Literature Study

Bala et al. [1] introduced MonkeyNet, a robust deep CNN model specifically designed for Monkeypox disease detection and classification. Their work involved extensive training and validation on a diverse dataset of Monkeypox skin lesion images, showcasing high accuracy and reliability in identifying different lesion types.

Sahin et al. [2] developed a novel approach for human Monkeypox classification using a deep pre-trained network integrated into a mobile application. Their study demonstrated the feasibility of leveraging AI technology for real-time diagnosis and remote healthcare management, particularly in resource-constrained settings.

Altun et al. [3] explored transfer learning techniques in Monkeypox detection using CNNs, leveraging pre-trained models like VGG16 and ResNet50. Their research highlighted the potential for utilizing existing knowledge in deep learning to enhance diagnostic capabilities for infectious diseases.

Almufareh et al. [4] proposed a transfer learning approach for clinical detection support of Monkeypox skin lesions, emphasizing the importance of utilizing pre-trained models to leverage domain-specific knowledge and improve diagnostic accuracy.

Velu et al. [5] introduced a Q-learning approach for human pathogenic Monkeypox disease recognition, incorporating reinforcement learning principles into the classification process. Their study offered a unique perspective on integrating machine learning with decision-making algorithms for disease identification.

Sorayaie Azar et al. [6] focused on Monkeypox detection using deep neural networks, demonstrating the effectiveness of CNN models in accurately classifying skin lesions associated with the disease. Their research contributes valuable insights into the potential of AI-driven approaches for infectious disease diagnosis and surveillance.

Irmak et al. [7] and Nayak et al. [8] explored the application of MobileNetV2 and VGGNet models for

Monkeypox skin lesion detection, highlighting the versatility of deep learning architectures in medical imaging analysis. Their studies showcased promising results in automated lesion identification and classification.

Khan and Ullah [9] presented an efficient technique for Monkeypox skin disease classification, leveraging pre-trained models and clinical data to enhance diagnostic accuracy. Their work emphasizes the importance of interpretable deep learning models in supporting clinical decision-making processes.

Ahsan et al. [10] focused on interpretable deep learning for Monkeypox diagnosis, developing models that not only provide accurate classification but also offer insights into the features contributing to the classification decision. Their research contributes to the development of transparent and understandable AI models in medical applications.

Eliwa et al. [11] utilized convolutional neural networks to classify Monkeypox skin lesions, showcasing the potential of deep learning techniques in automated disease detection. Their study highlights the scalability and efficiency of CNNs in handling large-scale medical image datasets.

Jaradat et al. [12] explored automated Monkeypox skin lesion detection using deep learning and transfer learning techniques, demonstrating the utility of AI-driven approaches in streamlining diagnostic workflows and improving accuracy in lesion identification.

Alrusaini [13] investigated deep learning models for the detection of Monkeypox skin lesions on digital images, showcasing the applicability of AI techniques in developing automated diagnostic tools for infectious diseases. Their work contributes to the ongoing efforts in leveraging technology for improved healthcare outcomes.

Rizk et al. [14] provided insights into the prevention and treatment of Monkeypox, highlighting the importance of a comprehensive approach in managing infectious diseases. Their study emphasizes the significance of both diagnostic advancements and therapeutic interventions in combating Monkeypox.

Ali et al. [15] conducted a feasibility study on Monkeypox skin lesion detection using deep learning models, laying the groundwork for potential real-world implementations of AI-driven diagnostic tools. Their research contributes to the validation of AI techniques in clinical settings for enhancing disease detection capabilities.

Common limitations across the mentioned studies include the reliance on limited or curated datasets for model training, which may not fully capture the variability and diversity of Monkeypox skin lesions in real-world clinical scenarios. Additionally, most studies primarily focused on the development and validation of deep learning models for classification, often lacking robust evaluations in diverse patient populations or clinical settings. Interpretability and explainability of the AI models also remain a challenge, with limited insights provided into the features driving classification decisions, hindering their adoption in clinical practice. Furthermore, scalability and generalizability of the proposed models to broader healthcare systems and resource-constrained environments were not extensively explored, highlighting the need for further research and validation in real-world settings to ensure the efficacy and reliability of AI-based solutions for Monkeypox diagnosis and management.

III. Proposed System

Developing a Convolutional Neural Network (CNN) model for the classification of Monkeypox skin lesions involves a multi-step process that integrates data preprocessing, model architecture design, training, fine-tuning, and evaluation. In this detailed explanation, each step can be understood effectively for accurate and efficient disease detection.

A. *Input Dataset and Image Augmentation:*

The initial step involves acquiring a dataset of Monkeypox skin lesion images. This dataset serves as the foundation for training and evaluating the CNN model. To enhance the diversity and robustness of the dataset, image augmentation techniques are applied. These techniques involve random transformations such as rotation, scaling, and flipping, which generate variations of the original images. By augmenting the dataset, it is ensured that the CNN model learns from a wider range of skin lesion patterns and features, improving its generalization capabilities.

B. *Split Dataset into Train/Test:*

After image augmentation, the dataset is split into training and testing sets. The training set is used to teach the CNN model to recognize patterns and features in Monkeypox skin lesions, while the testing set is kept separate to evaluate the model's performance on unseen data. Typically, the dataset is divided into a training set (e.g., 80% of the data) and a testing set (e.g., 20% of the data) to ensure an adequate balance between model training and evaluation.

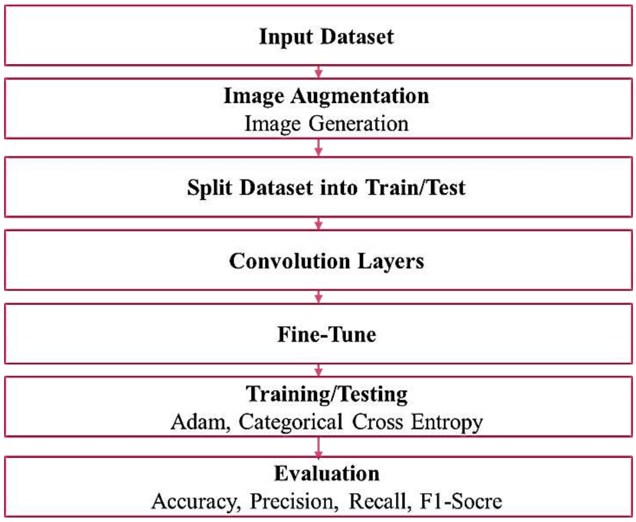


Fig. 2. Proposed System

C. *Model Architecture Design:*

The architecture of the CNN model is pivotal for efficiently extracting features and accurately classifying Monkeypox skin lesions. The provided architecture includes convolutional layers with different filter sizes, activation functions (e.g., LeakyReLU, ReLU), and pooling layers (MaxPooling) for downscaling feature maps. The choice of these layers and parameters is based on experimentation and optimization to achieve optimal performance in disease classification.

D. *Fine-Tuning and Optimization:*

Once the model architecture is defined, fine-tuning is performed to optimize its parameters and improve its ability to generalize the unseen data. Techniques such as dropout regularization are applied to prevent overfitting, where the model memorizes the training data but performs poorly on new data. The Adam optimizer, coupled with the Categorical Cross Entropy loss function, is used for model optimization. Adam adapts the learning rate dynamically during training, leading to faster convergence and improved model performance. Set hyperparameters such as learning rate, batch size, and optimizer (e.g., Adam, SGD with momentum). For instance, a learning rate of 0.001 and a batch size of 32 for training.

E. Training and Testing:

The training phase involves feeding the training dataset into the CNN model iteratively for a specified number of epochs. During each epoch, the model adjusts its parameters to minimize the loss function and improve its ability to classify Monkeypox skin lesions accurately. Classification task (e.g., 2 for binary classification). The Dense layer with a softmax activation function outputs probabilities for each class, making it suitable for multi-class classification tasks. Adjust the number of neurons in the Dense layer based on the specific classification problem.

F. Evaluation Metrics:

Accuracy measures the overall correctness of the model's predictions, while precision measures the proportion of true positive predictions among all positive predictions. Recall, also known as sensitivity, measures the proportion of true positive predictions among all actual positive instances. The F1-score is the harmonic mean of precision and recall, providing a balanced evaluation metric that considers both false positives and false negatives. These metrics collectively assess the model performance and help identify areas for improvement.

IV. RESULT ANALYSIS

The original Monkeypox Skin Lesion Dataset comprises 228 images, with 102 belonging to the Monkeypox class and 126 representing the Others class (non-monkeypox, including chickenpox and measles cases). To enhance the dataset's diversity and assist in classification, various data augmentation techniques were applied, including rotation, translation, reflection, shear, hue, saturation, contrast, brightness jitter, noise addition, and scaling using MATLAB R2020a. This augmentation process resulted in a significant increase in the number of images, with the Monkeypox and Others classes now having 1428 and 1764 images, respectively. These augmented images are included to ensure the reproducibility of results and provide a more comprehensive dataset for training and evaluation purposes.

Figure 4,7,10, and 13 are architecture of different models. Figure 5,8,11,and 14 are training process with accuracy and loss on y-axis and epoch on x-axis. Figure 6,9,12 and 15 are parameters of testing.

Link:
https://www.kaggle.com/datasets/nafin59/monkeypox-skin-lesion-dataset/data

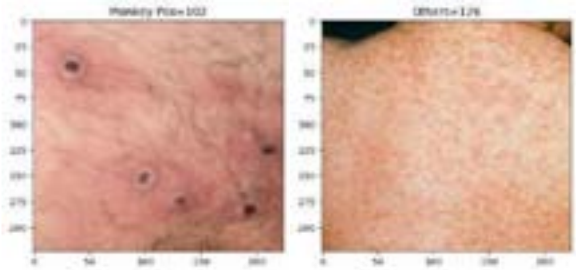


Fig. 3. Dataset Reading

Model: "sequential_4"

Layer (type)	Output Shape	Param #
conv2d_10 (Conv2D)	(None, 54, 54, 96)	34944
batch_normalization_15 (BatchNormalization)	(None, 54, 54, 96)	384
max_pooling2d_6 (MaxPooling2D)	(None, 26, 26, 96)	0
conv2d_11 (Conv2D)	(None, 26, 26, 256)	614656
batch_normalization_16 (BatchNormalization)	(None, 26, 26, 256)	1024
max_pooling2d_7 (MaxPooling2D)	(None, 12, 12, 256)	0
conv2d_12 (Conv2D)	(None, 12, 12, 384)	885120
batch_normalization_17 (BatchNormalization)	(None, 12, 12, 384)	1536
conv2d_13 (Conv2D)	(None, 12, 12, 384)	1327488
batch_normalization_18 (BatchNormalization)	(None, 12, 12, 384)	1536
conv2d_14 (Conv2D)	(None, 12, 12, 256)	884992
batch_normalization_19 (BatchNormalization)	(None, 12, 12, 256)	1024
max_pooling2d_8 (MaxPooling2D)	(None, 5, 5, 256)	0
flatten_4 (Flatten)	(None, 6400)	0
dense_12 (Dense)	(None, 4096)	26218496
dropout_9 (Dropout)	(None, 4096)	0
dense_13 (Dense)	(None, 4096)	16781312
dropout_10 (Dropout)	(None, 4096)	0
dense_14 (Dense)	(None, 2)	8194

Total params: 46760706 (178.38 MB)
Trainable params: 46757954 (178.37 MB)
Non-trainable params: 2752 (10.75 KB)

Fig. 4. AlexNet Model

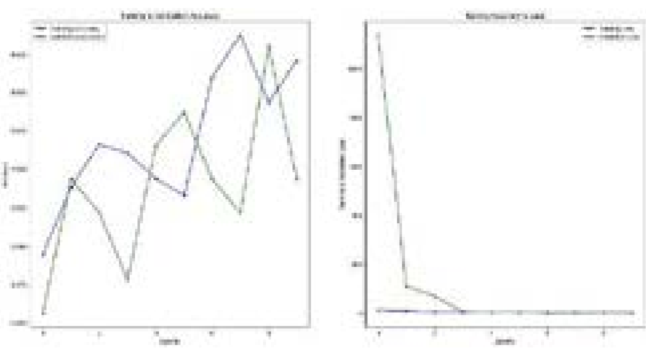


Fig. 5. AlexNet training/testing

[[12 9]
[12 13]]

	precision	recall	f1-score	support
0	0.50	0.57	0.53	21
1	0.59	0.52	0.55	25
accuracy			0.54	46
macro avg	0.55	0.55	0.54	46
weighted avg	0.55	0.54	0.54	46

Fig. 6. AlexNet Parameters

Model: "sequential_2"

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 7, 7, 512)	14714688
dropout_4 (Dropout)	(None, 7, 7, 512)	0
flatten_2 (Flatten)	(None, 25088)	0
batch_normalization_10 (BatchNormalization)	(None, 25088)	100352
dense_6 (Dense)	(None, 1024)	25691136
batch_normalization_11 (BatchNormalization)	(None, 1024)	4096
activation (Activation)	(None, 1024)	0
dropout_5 (Dropout)	(None, 1024)	0
dense_7 (Dense)	(None, 512)	524800
activation_1 (Activation)	(None, 512)	0
dense_8 (Dense)	(None, 2)	1026

Total params: 41036098 (156.54 MB)
Trainable params: 26269186 (100.21 MB)
Non-trainable params: 14766912 (56.33 MB)

Fig. 7. Vgg16 Model

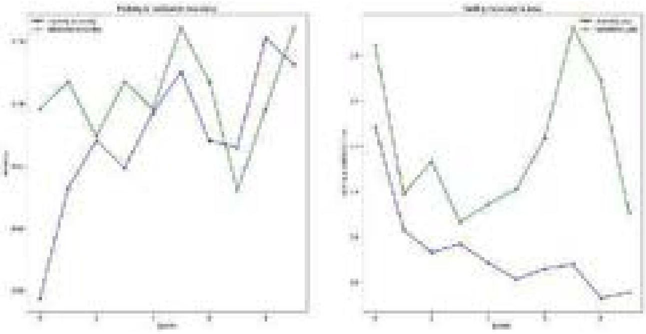


Fig. 8. Vgg16 training/testing

[[14 7]
[4 21]]

	precision	recall	f1-score	support
0	0.78	0.67	0.72	21
1	0.75	0.84	0.79	25
accuracy			0.76	46
macro avg	0.76	0.75	0.76	46
weighted avg	0.76	0.76	0.76	46

Fig. 9. Vgg16 Parameters

Model: "sequential_3"

Layer (type)	Output Shape	Param #
resnet50 (Functional)	(None, 7, 7, 2048)	23587712
dropout_6 (Dropout)	(None, 7, 7, 2048)	0
flatten_3 (Flatten)	(None, 100352)	0
batch_normalization_12 (BatchNormalization)	(None, 100352)	401408
dense_9 (Dense)	(None, 2048)	205522944
batch_normalization_13 (BatchNormalization)	(None, 2048)	8192
activation_2 (Activation)	(None, 2048)	0
dropout_7 (Dropout)	(None, 2048)	0
dense_10 (Dense)	(None, 1024)	2098176
batch_normalization_14 (BatchNormalization)	(None, 1024)	4096
activation_3 (Activation)	(None, 1024)	0
dropout_8 (Dropout)	(None, 1024)	0
dense_11 (Dense)	(None, 2)	2050

Total params: 231624578 (883.58 MB)
Trainable params: 207830018 (792.81 MB)
Non-trainable params: 23794560 (90.77 MB)

Fig. 10. ResNet Model

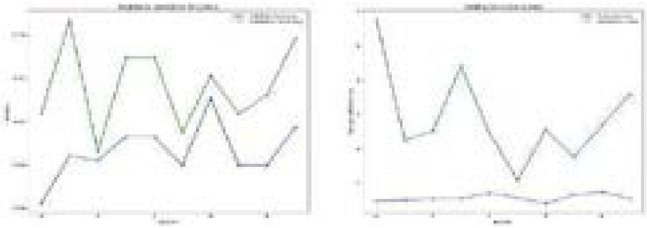


Fig. 11. Resnet training/testing

[[12 9]
[5 20]]

	precision	recall	f1-score	support
0	0.71	0.57	0.63	21
1	0.69	0.80	0.74	25
accuracy			0.70	46
macro avg	0.70	0.69	0.69	46
weighted avg	0.70	0.70	0.69	46

Fig. 12. ResNet Parameters

Model: "sequential"

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 111, 111, 128)	[illegible]
leaky_re_lu (LeakyReLU)	(None, 111, 111, 128)	0
max_pooling2d (MaxPooling2D)	(None, 55, 55, 128)	0
conv2d_1 (Conv2D)	(None, 55, 55, 256)	[illegible]
leaky_re_lu_1 (LeakyReLU)	(None, 55, 55, 256)	0
max_pooling2d_1 (MaxPooling2D)	(None, 28, 28, 256)	0
conv2d_2 (Conv2D)	(None, 28, 28, 256)	[illegible]
conv2d_3 (Conv2D)	(None, 28, 28, 256)	590080
leaky_re_lu_2 (LeakyReLU)	(None, 28, 28, 256)	0
max_pooling2d_2 (MaxPooling2D)	(None, 14, 14, 256)	0
conv2d_4 (Conv2D)	(None, 14, 14, 512)	[illegible]
conv2d_5 (Conv2D)	(None, 14, 14, 512)	2359808
conv2d_6 (Conv2D)	(None, 14, 14, 512)	[illegible]
conv2d_7 (Conv2D)	(None, 14, 14, 512)	2359808
max_pooling2d_3 (MaxPooling2D)	(None, 7, 7, 512)	0
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 4096)	102764544
dropout (Dropout)	(None, 4096)	0
dense_1 (Dense)	(None, 512)	2097664
dense_2 (Dense)	(None, 2)	1026

Total params: [illegible]
Trainable params: [illegible]
Non-trainable params: 0 (0.00 Byte)

Fig. 13. Fine-Tune CNN Model

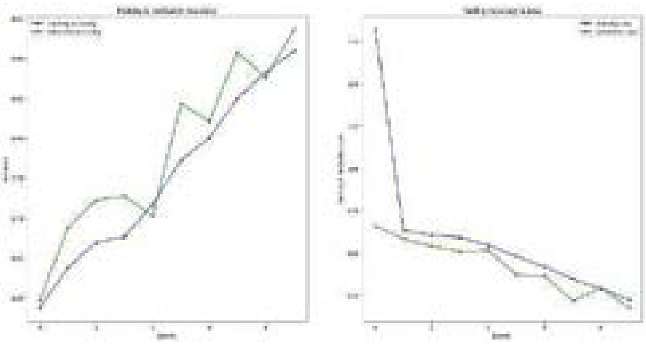


Fig. 14. Fine-Tune CNN training/testing

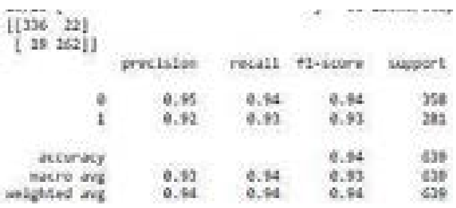


Fig. 15. Fine-Tune CNN Parameters

TABLE I. ANALYSIS OF MODELS

Model	ACC (%)	P (%)	R (%)	F1-Score (%)
AlexNet	54%	55%	55%	54%
VGG16	76%	76%	75%	76%
ResNet50	70%	70%	69%	70%
Fine-Tune CNN	94%	93%	94%	93%

CONCLUSION

In conclusion, the evaluation of various models for Monkeypox skin lesion classification reveals distinct levels of performance. While AlexNet demonstrated moderate accuracy and balanced precision, recall, and F1-score around 55%, VGG16 and ResNet50 exhibited higher accuracy and more balanced metrics, with VGG16 achieving 76% accuracy and ResNet50 reaching 70%. However, the most notable results were observed with the fine-tuned CNN model, which achieved an impressive accuracy of 94% along with balanced precision, recall, and F1-score around 93%. These findings highlight the effectiveness of fine-tuning CNN models for accurate and reliable Monkeypox skin lesion classification, underscoring the significance of model selection and optimization techniques in enhancing diagnostic capabilities for infectious diseases. In future, augmenting dataset with transformations like rotation, scaling, flipping, and shifting can increase the diversity of training examples and helps to improve the classification performances.

REFERENCES

[1] D. Bala et al., “MonkeyNet: A robust deep convolutional neural network for monkeypox disease detection and classification,” Neural Networks, vol. 161, pp. 757–775, 2023, doi: 10.1016/j.neunet.2023.02.022.

[2] V. H. Sahin, I. Oztel, and G. Yolcu Oztel, “Human Monkeypox Classification from Skin Lesion Images with Deep Pre-trained Network using Mobile Application,” Journal of Medical Systems, vol. 46, no. 11, 2022, doi: 10.1007/s10916-022-01863-7.

[3] M. Altun, H. Gürüler, O. Özkaraca, F. Khan, J. Khan, and Y. Lee, “Monkeypox Detection Using CNN with Transfer Learning,” Sensors, vol. 23, no. 4, 2023, doi: 10.3390/s23041783.

[4] M. F. Almufareh, S. Tehsin, M. Humayun, and S. Kausar, “A Transfer Learning Approach for Clinical Detection Support of Monkeypox Skin Lesions,” Diagnostics, vol. 13, no. 8, pp. 1–16, 2023, doi: 10.3390/diagnostics13081503.

[5] M. Velu et al., “Human Pathogenic Monkeypox Disease Recognition Using Q-Learning Approach,” Diagnostics, vol. 13, no. 8, 2023, doi: 10.3390/diagnostics13081491.

[6] A. Sorayaie Azar, A. Naemi, S. Babaei Rikan, J. Bagherzadeh Mohasefi, H. Pirnejad, and U. K. Wiil, “Monkeypox detection using deep neural networks,” BMC Infectious Diseases, vol. 23, no. 1, pp. 1–13, 2023, doi: 10.1186/s12879-023-08408-4.

[7] M. C. Irmak, T. Aydın, and M. Yağanoğlu, “Monkeypox Skin Lesion Detection with MobileNetV2 and VGGNet Models,” TIPTEKNO 2022 - Medical Technologies Congress, Proceedings, no. November, 2022, doi: 10.1109/TIPTEKNO56568.2022.9960194.

[8] T. Nayak et al., “Deep learning based detection of monkeypox virus using skin lesion images,” Medicine in Novel Technology and Devices, vol. 18, no. March, p. 100243, 2023, doi: 10.1016/j.medntd.2023.100243.

[9] Gul Zaman Khan and Inam Ullahx, “Efficient Technique for Monkeypox Skin Disease Classification with Clinical Data using Pre-Trained Models,” Journal of Innovative Image Processing, vol. 5, no. 2, pp. 192–213, 2023, doi: 10.36548/jiip.2023.2.009.

[10] M. M. Ahsan et al., “Monkeypox Diagnosis with Interpretable Deep Learning,” IEEE Access, no. August, pp. 81965–81980, 2023, doi: 10.1109/ACCESS.2023.3300793.

[11] E. H. I. Eliwa, A. M. El Koshiry, T. Abd El-Hafeez, and H. M. Farghaly, “Utilizing convolutional neural networks to classify monkeypox skin lesions,” Scientific reports, vol. 13, no. 1, p. 14495, 2023, doi: 10.1038/s41598-023-41545-z.

[12] A. S. Jaradat et al., “Automated Monkeypox Skin Lesion Detection Using Deep Learning and Transfer Learning Techniques,” International Journal of Environmental Research and Public Health, vol. 20, no. 5, 2023, doi: 10.3390/ijerph20054422.

[13] O. A. Alrusaini, “Deep Learning Models for the Detection of Monkeypox Skin Lesion on Digital Skin Images,” International Journal of Advanced Computer Science and Applications, vol. 14, no. 1, pp. 637–644, 2023, doi: 10.14569/IJACSA.2023.0140170.

[14] J. G. Rizk, G. Lippi, B. M. Henry, D. N. Forthal, and Y. Rizk, “Prevention and Treatment of Monkeypox,” Drugs, vol. 82, no. 9, pp. 957–963, 2022, doi: 10.1007/s40265-022-01742-y.

[15] S. N. Ali et al., “Monkeypox Skin Lesion Detection Using Deep Learning Models: A Feasibility Study,” pp. 2–5, 2022, [Online]. Available: http://arxiv.org/abs/2207.03342

Depression Detection using Extreme Learning Machine

Prajna Dutta[1]
School of Data Science
Indian Institute of Science Education and Research, Thiruvananthapuram (IISER TVM)
Vithura, India
prajnadutta21@iisertvm.ac.in

Deepak Gupta[2]
Computer Science and Engineering Department
MNNIT Allahabad
Prayagraj, India
deepakg@mnnit.ac.in

Jyoti Maurya[3]
Computer Science and Engineering Department
MNNIT Allahabad
Prayagraj, India
jyoti14th08@gmail.com

***Abstract*—Depression is the most well-known type of physiological or mental health problem that affects a large number of people globally. Extreme learning machine (ELM) techniques are currently preferred to solve wide range of health disease detection and prediction issues. ELM is a single hidden layer feed-forward neural network (SLFN), which converges much faster than the other traditional Machine Learning (ML) methods and yields promising results. Many research works already exist on the application of Machine Learning (ML) models to the Depression Detection dataset but little to no work was found wherein Extreme Learning Machine was used for Depression Detection. This research work has applied Extreme Learning Machine (ELM) and other ML techniques for depression detection and compared the results obtained and found that, ELM has delivered the best performance with an accuracy of 91.73%.**

Keywords—Machine Learning, Extreme Learning Machine, Depression Detection

I. INTRODUCTION

Depression affects millions of people around the world. Depression is different when compared to the mood changes happening in daily life. Depression can affect various aspects of life, such as having problems in social relations and not being able to concentrate in work. Depressed people find it difficult to get pleasure from daily activities. Depression can be recurrent or persistent. In most cases, in-person clinical depression criteria are used to diagnose depression. But a large number of individuals having early-stage depression do not consult a doctor by which their conditions worsen. [1].

Therefore, depression has to be identified in its early stages such that the patient receives proper care. Large amounts of relevant data are needed to build a depression detection model [2]. Various models in Machine Learning and Deep Learning being are used for early identification and prediction of Depression.

Machine Learning is a subdiscipline of Artificial Intelligence in which computational methods are used to gain knowledge from data and performance increases via experience. In the traditional definition of Machine Learning approaches, three forms of knowledge acquisition are included: Supervised, Unsupervised and Reinforcement Learning [3].

In Supervised Learning, a statistical model is trained by an algorithm to generate predictions about some class that is not labeled. A column containing data with the answer (label/target) is used for monitoring learning while training. As implied by the term, unsupervised learning relies only on feature values, or the inherent structure of the data, rather than labels. Here, the main focus is the process of finding patterns and extracting knowledge. The two previous approaches are fundamentally different from reinforcement learning (RL), which does not require human-generated data for training but instead learns by trial and error. It learns on the basis of feedback and reward.

Being relatively new, Machine Learning (ML) is still growing quickly. It is fundamental to data science and Artificial Intelligence (AI) and resides at the intersection of computer science and statistics. The advancement of theory of some novel learning algorithms and the continuous proliferation of low-cost processing and massive amounts of data (commonly called "big data") have both fueled recent advances in machine learning. A greater amount of decision making based on evidence is occurring across a variety of fields, including biomedicine, healthcare [4], education, manufacturing, data governance, financial modeling, marketing and policing, as a result of the widespread use of ML-based methodologies in research, technology, and industry.

Depression is a common mental health issue that impacts a huge number of people globally. Effective therapy and support depend on early identification and intervention. Creating a scalable and dependable depression detection system that can help medical practitioners diagnose patients early and carry out timely interventions is the problem. This study entails utilizing machine learning techniques, data, and technology to develop a tool that may recognize possible indicators of depression in people. The objective is to provide a thorough and trustworthy instrument for detecting depression by addressing these elements, which can aid in early intervention and better mental health outcomes.

In this work, multiple Machine Learning models have been applied for Depression Detection such as Logistic Regression, K Nearest Neighbors, Extreme Learning

Machines, and Decision Tree, Gaussian Naïve Bayes, and Random Forest Classifiers.

Section 2 describes related work, Section 3 describes the proposed work, and Section 4 describes the outcome of applying various models in Machine Learning used in predicting depression. Part 5 of the paper discusses the conclusion and part 6 describes the future scope.

II. REALTED WORKS

A lot of work is being done related to application of Machine Learning techniques to predict depression on different types of available datasets. Some works of well-known authors are described below:

The work conducted by Marzieh Mousavian et al.[5] on "Feature Selection and Imbalanced Data Handling for Depression Detection" in which Random Forest Classifier shows the best result as compared to the other Machine Learning models used, followed by SVM with a Gaussian Kernel.

Md Rafiqul Islam et al.[6] proposed a model using machine learning methods with social network data for diagnosis of depression. The model shows the accuracy between 60% to 80% for the classification models with Decision Tree showing the highest accuracy.

Hatoon AlSagri et al.[7] conducted a work where he used Machine Learning techniques to approach a problem in Twitter in which Depression was detected using content features and activity features and the maximum accuracy achieved was 82.5% by Linear SVM Model.

Raihan Sultana et al.[8] proposed a anxiety and depression detection model using ML and DL approaches such as CNN, Linear Regression, LDA, KNN and Support vector machine, out of which CNN outperforms with an accuracy of 96.8% for depression and 96% for anxiety.

Swasthika Jain et al.[9] used multimodal depression detection technique using deep neural network(D-ResNet) and paragraph-vector based kernel extreme machine learning model(PVKELM). The accuracy obtained by the model is 89%.

Subrajeet Mohapatra et al.[10] proposed a model of early accessment of depression using machine learning model. They used stacked SVM model and achieved an accuracy of 89.4%.

Milena Cukic et al.[11] used EEG and HFD signals as features for seven ML algorithms including Naive Bayes classifier, Logistic Regression, Random Forest, Decision Tree, Support Vector Machines and Multilayer Perceptron, with the linear and polynomial kernel. Average accuracy of the classifiers varies from 90.24% to 97.56%.

Lu, Shengfu et al.[14] build a depression classification model using eye-movement data and ELM and also used Particle Swarm optimization for optimizing the parameters. The accuracy, sensitivity and specificity of the model using improved chaotic projection model and Gauss mutation strategy are 88.55%, 87.71% and 89.42%, respectively.

III. PROPOSED WORK

In this work, a novel approach has been proposed for detecting depression using Extreme Learning Machine (ELM) model. Many machine algorithm models have been used for depression detection but ELM outperforms them. The detailed view of the work is described below:

1. Dataset:

The dataset has originally been taken from the Centers for Disease Control (CDC) and Prevention's National Health and Nutrition and Examination Survey. The dataset is available on: https://wwwn.cdc.gov/nchs/nhanes/default.aspx and consists of data collected from adults living in the U.S., from 2005 to 2018 and it has 36259 entries.

2. Description of the Libraries used:

1. ***NumPy:*** NumPy is the primary Python package for scientific computing. In NumPy, n-dimensional array is used as the data structure. One can perform various array operations using NumPy [18].
2. ***Pandas:*** Pandas is a toolkit in Python programming language that offers user-friendly data structures. Pandas library can be used for working with datasets. Pandas can be used to analyze, clean, explore and manipulate data [15].
3. ***Matplotlib:*** Matplotlib is a Python visualization package. 2D plots of arrays can be created using Matplotlib. Some commonly used plots are line plot, bar plot, scatter plot and histogram plot, etc [20].
4. ***Scikit-learn:*** It is a package used in machine learning for programming in Python. It has supervised and unsupervised learning algorithms. It is built on NumPy, Matplotlib and Pandas. Scikit-learn is available as a free software. It is compatible with SVMs, gradient boosting, random forests, DBSCAN, and other classification, clustering and regression techniques [21].

3. Methodology Used

3.1. Data Pre-processing

The dataset contains both numeric and non-numeric data. The categorical input features, numeric and non- numeric, are one-hot-encoded and the class labels are label-encoded to '1' which represents the Depressed Class and '0' which represents the Not Depressed class. Thereafter scaling is performed on the data. The input and output variables are separated and a train-test split is performed.

3.2. Clustering

Here, K-Means Clustering has been performed on the data such that 6 clusters are formed. In K-Means Clustering, k numbers of clusters are formed by partitioning of n-number of observations and every observation belongs to the cluster which has the nearest mean.

3.3. Models Implementation

This work uses various classification models in Machine Learning such as Logistic Regression, RF classifier, Extra Trees Classifier, Decision Tree Classifier, KNN, Gaussian Naive Bayes Classifier and Extreme Learning Machine, to predict depression by using the CDC dataset.

Logistic Regression Model:

The Logistic Regression model is used for classification problems. The input data is modelled using a linear equation. The sigmoid function squishes the predicted output between 0 and 1. The data is fitted on the logistic regression equation and using likelihood function, the output is predicted. In case of binomial logistic regression, the output can be of only two possible types, 0 and 1, which represents the two opposite given instances of a class [19].

Random Forest Classifier Model:

Random Forest Classifier is an example of ensemble learning technique in which two or more models are fitted into the same data and the predictions of each model is combined. Random Forest Classifier is made up of multiple Decision Trees each of which is made from a subset formed from the training data. The bagging technique is used in the Random Forest Classifier Model. The training data is recursively partitioned based on the features, to build the trees. For every split, the best feature from an arbitrary subset is selected by the algorithm, while optimizing Gini impurity. The process continues until a predefined stopping condition is satisfied. After training of the random forest, predictions are made based on the votes in each tree for a class and the class with the maximum number of votes is the predicted class for the given input data [20].

Decision Tree Classifier Model:

Decision Trees Classifier is an example of a supervised machine learning algorithm. It has a tree-like structure with different types of nodes, namely, the root node, leaf node and internal node. The algorithm starts working form the root node, followed by the leaf node and the internal nodes. The commonly used methods to measure impurity include Gini impurity, information gain and entropy. For each value, the weighted average of the Gini impurity at each level is calculated. And the feature which has the least impurity becomes the node. The process is repeated for every node, at each level, until the given data gets classified completely. After constructing the tree, the algorithm goes down the tree, following the conditions at the nodes and finally the data gets classified [21].

Extra Trees Classifier Model:

Extra Trees Classifier is an example of an ensemble learning technique. It consists of multiple Decision Trees which are de-correlated. Every Decision Tree in Extra Trees Classifier is built from the actual training sample. At every test node, every tree is provided with an arbitrary sample which has k number of features, from which every Decision Tree has to select the best feature for splitting the given data on the basis of some criteria such as Gini index. This arbitrary sample of features creates sevaral Decision Trees which are de-correlated [22].

K Nearest Neighbors:

K Nearest Neighbors is an example of a supervised machine learning algorithm. It calculates the distance between the training data points and the test data and tries to predict the class correctly. 'K' number of points closest to the test data is selected. The probability of the test data belonging to the class of 'K' training data is calculated and the class which has the maximum probability becomes the predicted class [23].

Gaussian Naive Bayes:

Naïve Bayes Classification Algorithm is based on Bayes' Theorem. In the algorithm, it is assumed that the features which are used for describing any observation, when the class label is given, are conditionally independent. Gaussian Naïve Bayes is an example of a Naïve Bayes method, in which the features in the data follow a Gaussian Distribution. In order to classify every new data point, the algorithm determines the maximum value of posterior probability of every class and allocates the data point to that class [24].

Extreme Learning Machine (ELM):

An ELM is a specific type of feed-forward neural network in which a single hidden layer with randomly generated weights and biases is used [12]. In contrast to other neural networks, ELM can attain high accuracy without the need for iterative training or bias and weight adjustments. Rather, it computes the output weights depending on the input data using an analytical approach [9].

The weights in the hidden layer of the ELM method have two phases: training and testing. When input data is supplied into the neural network during training, biases are generated at random. The least-squares approach is then used to determine the output weights. There is no need for iterative weight and bias change during the quick training phase [25].

In order to train feed-forward networks (SLFNs), Guang-Bin and Qin-Yu introduced the Extreme Learning Machine (ELM). Three layers constitute the SLFN: the input, hidden and output layers as shown in fig. 1

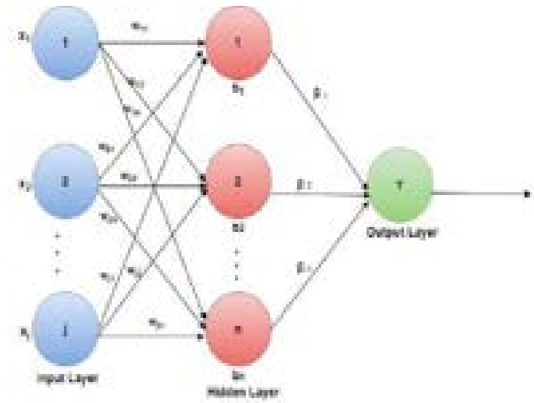


Fig1. Structure of Single-layer feed-forward Network (SLFN)

Given a training set

$$A = \{(x_i, y_i) \mid x_i = (x_{i1}, x_{i2}, \ldots, x_{in})^T \in R^n,\ y_i = (y_{i1}, y_{i2}, \ldots, y_{im})^T \in R^m\}$$

Where, x_i denotes the input value and y_i represents the target output of an ELM with n hidden neurons.

$$\sum_{i=1}^{n} \beta_i F(w_i x_j + b_i) = y_j,\ j = 1, 2, \ldots m. \quad \cdots\cdots \quad (1)$$

Where, $F(\mathrm{x})$ is the activation function in the hidden layer, x_i is the input vector of the i^{th} sample, y_i is the target vector of the i^{th} sample, w_i is the weight vector from input layer to the i^{th} hidden node, b_i is the bias of the i^{th} hidden node, β_i is the weight vector from the i^{th} hidden node to output layer.

The above equation (1) can be abbreviated as:

$$A\beta = Y$$

or

$$\beta = (\mathrm{A})^{\dagger} Y$$

Where, $(\mathrm{A})^{\dagger}$ is Moore-Penrose Pseudo inverse.

β is the weight between hidden and output layers.

Y is the output.

$$A = \begin{bmatrix} f(w_1 x_1 + b_1) & \ldots & f(w_L x_1 + b_L) \\ \ldots & \ldots & \ldots \\ f(w_1 x_N + b_1) & \ldots & f(w_L x_N + b_L) \end{bmatrix}_{\mathrm{NXL}}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \beta_2^T \\ .. \\ .. \\ \beta_L^T \end{bmatrix}_{\mathrm{LXm}} \qquad Y = \begin{bmatrix} y_1^T \\ y_2^T \\ .. \\ .. \\ y_L^T \end{bmatrix}_{\mathrm{NXm}}$$

Once the output weights β are determined, they are used for making predictions on new and unseen data. Here, we have used it for predicting depression.

3.4. Hyper-parameter Tuning

Hyper-parameter Tuning is performed using Grid Search CV for each of the models and the best parameters are chosen to train the models. After model training, the models are applied on the test sets and then evaluated.

3.5. Model Evaluation

The models are evaluated based on Accuracy and AUC Score. Confusion Matrix helps to know the correctly predicted and incorrectly predicted class labels.

Confusion Matrix: A confusion matrix is a particular layout of table which is used in machine learning, particularly for classification. It is also called a “matching matrix” when it is used for unsupervised learning. It is used to visualize the performance of an algorithm. In the confusion matrix, each column in the matrix represents an example in a predicted class, each row represents an occurrence in the actual class, or the opposite.

The confusion matrix displays the following:

1. **True Positive (TP)** occurs when a model is able to predict correctly a data point which is positive.
2. **True Negative (TN)** occurs when a model is able to predict correctly a data point which is negative.
3. **False Positive (FP)** occurs when a model predicts a data point which is positive but the prediction is incorrect.
4. **False Negative (FN)** occurs when a model predicts a data point which is negative but the prediction is incorrect.

Accuracy: The percentage of correctly identified cases amongst all of the instances is a metric known as accuracy.

$$\text{Accuracy} = \frac{(\mathrm{TP+TN})}{(\mathrm{TP+TN+FP+FN})} * 100$$

AUC (Area Under the ROC Curve): In other words, AUC calculates area in 2-D that lies under the entire ROC curve.

$$\text{AUC} = \frac{(\mathrm{TP+TN})}{(\mathrm{TP+TN+FP+FN})} * 100$$

IV. RESULTS

The following results were obtained after evaluating the different models:

Table 1: Evaluation measures for various models.

Models	Hyper-parameter Used	Accuracy (%)	AUC (%)
Logistic Regression	class_weight = 'balanced', solver = 'newton-cholesky', penalty='l2'	78.76	60.22
Random Forest Classifier	class_weight='balanced', criterion ='entropy', max_depth =10, max_samples=0.95, min_samples_leaf=100, n_estimators = 500	81.56	61.41
Decision Tree Classifier	criterion = "gini", max_depth = 6, min_samples_split = 2, random_state = 123	91.16	70.86
Extra Trees Classifier	n_estimators=100, n_jobs=4, min_samples_split=25, min_samples_leaf=25, max_features=120	91.71	70.86
K Nearest Neighbors Classifier	leaf_size = 1, metric = 'manhattan', n_neighbors = 7, weights = 'distance'	91.26	70.63
Naive Bayes Classifier (Gaussian)	var_smoothing = 0.0533669923120631	91.44	78.72

Extreme Learning Machine	L=200	**91.73**	**83.82**

Here, various Machine Learning models have been applied for predicting depression like Logistic Regression, Random Forest, Decision Tree, Extra Trees, KNN, Gaussian Naïve Bayes and Extreme Learning Machine model. Results show that Extreme Learning Machine performs the best out of all the models used. It has an Accuracy = 91.73% and AUC = 83.82%.

Fig. 2 describes the graphical representation of Accuracy and AUC Score for various Machine Learning Models and the best result is obtained by Extreme Learning Machine model with an accuracy of 91.73%.

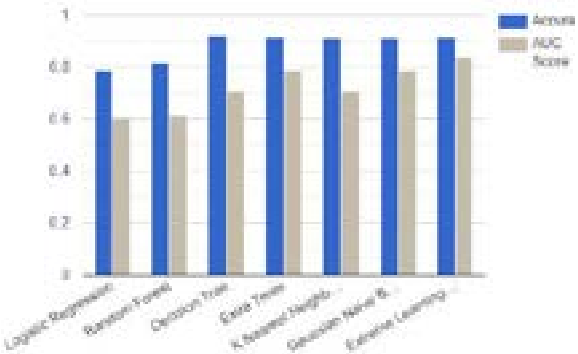


Fig2. Accuracy and AUC measures for multiple ML algorithms

V. CONCLUSION

Depression is one of the deadliest diseases among the known neurological disorders. It is necessary to identify depression in its early stages. In this work, the dataset from the Centers for Disease Control (CDC) and Prevention's National Health and Nutrition Examination Survey has been taken, pre-processed and multiple ML Algorithms have been applied on it and the accuracy and AUC score have been evaluated.

After the application of multiple Machine Learning models for Depression Detection, like Logistic Regression, K Nearest Neighbors, Extreme Learning Machines, and Decision Tree, Gaussian Naïve Bayes, and Random Forest Classifiers, this study has observed that Extreme Learning Machine (ELM) outperforms all the other models, with an accuracy of 91.7264% and AUC = 83.8153%, followed immediately by Extra Trees Classifier with an accuracy of 91.7264% and AUC = 82%, followed by the rest.

VI. FUTURE SCOPE

Future work on depression detection through Machine Learning may likely focus on improvement of the efficiency of the models used, based on the performance metrics. The prediction results can be improved further with applications of more advanced models and also the models' performance can be improved by further hyper-parameter tuning. More work can also be conducted on trying to understand the particular features which affect the performance of the models more than the rest. Combining multiple sources of data, such as text, voice, and physiological signals, seems to improve the accuracy of the models used to detect depression. Integrating diverse data types provides a more comprehensive view of an individual's mental state.

References

[1] K. Min, J. Yoon, M. Kang, D. Lee, E. Park, and J. Han, "Detecting depression on video logs using audiovisual features," no. 2023, doi: 10.1057/s41599-023-02313-6.

[2] H. Zhang, H. Wang, S. Han, W. Li, and L. Zhuang, "Computer Methods and Programs in Biomedicine Detecting depression tendency with multimodal features," *Comput. Methods Programs Biomed.*, vol. 240, no. January 2022, p. 107702, 2023, doi: 10.1016/j.cmpb.2023.107702.

[3] J. Jovel, R. Greiner, and R. Greiner, "An Introduction to Machine Learning Approaches for Biomedical Research," vol. 8, no. December, pp. 1–15, 2021, doi: 10.3389/fmed.2021.771607.

[4] J. Maurya and S. Prakash, "Heart Disease Prediction and Diagnosis Using IoT, ML, and Cloud Computing," in *International Conference on Innovative Computing and Communications*, A. E. Hassanien, O. Castillo, S. Anand, and A. Jaiswal, Eds., Singapore: Springer Nature Singapore, 2024, pp. 419–430.

[5] I. Conference, "Feature Selection and Imbalanced Data Handling for Depression Detection : Feature Selection and Imbalanced Data Handling for Depression Detection," no. September 2023, 2018, doi: 10.1007/978-3-030-05587-5.

[6] R. Islam, M. A. Kabir, A. Ahmed, A. R. M. Kamal, and H. Wang, "Depression detection from social network data using machine learning techniques," *Heal. Inf. Sci. Syst.*, vol. 6, no. 1, pp. 1–12, 2018, doi: 10.1007/s13755-018-0046-0.

[7] H. Alsagri, "Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features," vol. 1, pp. 1–16.

[8] A. Ahmed, R. Sultana, M. T. R. Ullas, M. Begom, M. M. I. Rahi, and M. A. Alam, "A Machine Learning Approach to detect Depression and Anxiety using Supervised Learning," *2020 IEEE Asia-Pacific Conf. Comput. Sci. Data Eng. CSDE 2020*, pp. 4–9, 2020, doi: 10.1109/CSDE50874.2020.9411642.

[9] S. J. T. J, I. J. Jacob, and A. K. Mandava, "D-ResNet-PVKELM: deep neural network and paragraph vector based kernel extreme machine learning model for multimodal depression analysis," pp. 25973–26004, 2023.

[10] D. Kumar, U. Subrajeet, M. Niraj, and K. Singh, "An early assessment of Persistent Depression Disorder using machine learning algorithm," *Multimed. Tools Appl.*, no. 0123456789, 2023, doi: 10.1007/s11042-023-17369-4.

[11] D. Pokrajac, "The successful discrimination of depression from EEG could be attributed to proper feature extraction and not to a particular classification method," vol. 0123456789, pp. 443–455, 2020, doi: 10.1007/s11571-020-09581-x.

[12] A. Albtoush and M. Ferna, "Quick extreme learning machine for large-scale classification," vol. 8, pp. 5923–5938, 2022, doi: 10.1007/s00521-021-06727-8.

[13] J. Wang, S. Lu, S. Wang, and Y. Zhang, "A review on extreme learning machine," pp. 41611–41660, 2022.

[14] Lu, Shengfu; Liu, Sa; Li, Mi; Shi, Xin; Li, Richeng," Depression Classification Model Based on Emotionally Related Eye-Movement Data and Kernel Extreme Learning Machine", Journal of Medical Imaging and Health Informatics, Volume 10, Number 11, November 2020, pp. 2668-2674(7), https://doi.org/10.1166/jmihi.2020.3198.

[15] https://pandas.pydata.org/docs/ (accessed on 25-12-2023).

[16] https://matplotlib.org/stable/plot_types/stats/hist_plot.html (accessed on 25-12-2023).

[17] https://en.wikipedia.org/wiki/Scikit-learn (accessed on 25-12-2023).

[18] https://numpy.org/doc/stable/user/whatisnumpy.html (accessed on 25-12-2023).

[19] https://www.geeksforgeeks.org/ml-logistic-regression-using-python/ (accessed on 26-12-2023).

[20] https://www.geeksforgeeks.org/random-forest-classifier-using-

scikit-learn/ (accessed on 26-12-2023).
[21] https://pro.arcgis.com/en/pro-app/latest/tool-reference/geoai/how-decision-tree-classification-and-regression-works.htm (accessed on 27-12-2023).
[22] https://prutor.ai/extra-tree-classifier-for-feature-selection/ (accessed on 27-12-2023).
[23] https://medium.com/swlh/k-nearest-neighbor-ca2593d7a3c4/ (accessed on 28-12-2023).
[24] https://www.geeksforgeeks.org/gaussian-naive-bayes/(accessed on 28-12-2023).
[25] https://www.javatpoint.com/elm-in-machine-learning (accessed on 28-12-2023).

Brain Neoplasm Identification using CNN with VGG-16 Model

Ms. Soja Rani S
Department of CSE,
New Horizon College of Engineering,
Bangalore, India
soja.naveen@gmail.com

Dr. Nirmala M
Department of CSE,
New Horizon College of Engineering,
Bangalore, India
dr.nirmalam.nhce@newhorizonindia.edu

Sai Kumar Reddy J
Student
Department of CSE,
New Horizon College of Engineering,
Bangalore, India
saikumarreddy0210@gmail.com

Shree Raksha C
Student
Department of CSE,
New Horizon College of Engineering,
Bangalore, India
shreerakshac@gmail.com

Hruday S
Student
Department of CSE,
New Horizon College of Engineering,
Bangalore, India
hrudaysolleti111@gmail.com

***Abstract*—** In medical imaging, brain neoplasm identification is essential since a timely diagnosis often enhances patient outcomes. the study focuses on utilizing the VGG-16 model, a widely acknowledged convolutional neural network computer vision, for detection of brain tumor. The primary objectives include classifying magnetic resonance imaging (MRI) images and accurately identifying the presence of cerebral tumor. The dataset utilized comprises MRI images. The methodology encompasses setting up the environment, data importation and preprocessing, VGG-16 model creation, and assessing its performance in terms of accuracy, precision, and recall, The results indicate that the VGG-16 model holds promise in aiding medical professionals in brain diagnosis, achieving an accuracy of approximately 95% on the validation set and 82% on the test set.

Keywords— CNN, Brain Tumor, MRI, Deep Learning, Convolutional Neural Network, VGG16

1. Introduction

A intracranial tumor ranks fourth in United States for both men and women, but it is the tenth most common tumor in males and it ranks as the 9th most prevalent tumor among women. Within single digits, brain tumors have a 5-year survival rate, making them the most common basic tumor. According to a recent analysis released by Brain Cancer Action, brain tumors are predicted to rank as the second most common cause of tumor deaths inside United States by 2020.

The delayed diagnosis of brain tumor due to the lack of adequate early detection measures is one of the main causes of this. Furthermore, treatment becomes more difficult if the majority of cancer have spread to other organs. Thus, a method to help radiologists find brain tumor early on is desperately needed. Not much progress has been made in the diagnosis of brain tumor. According to a review of the literature, brain diagnoses are made without the use of picture processing by relying solely on the patient's records and the illness's signs and symptoms.

The quest for precise and effective brain tumor detection in the field of medical imaging has led to the incorporation of state-of-the-art technologies. One strategy that has shown a lot of promise recently, Traditional methods of brain tumor detection often rely on manual interpretation of radiological images, which is time-consuming and susceptible to human error. The integration of deep learning (DL) strategies, specifically CNNs, aims to streamline this process by automating the identification of abnormal patterns indicative of brain tumors. The VGG-16 architecture, known for its effectiveness in image classification tasks, presents an opportunity to improve the accuracy and efficiency of brain tumor diagnosis through automated analysis of MRI data.

The VGG-16 model has demonstrated impressive performance in image recognition tasks since its introduction by Simonyan and Zisserman in 2015. The deep architecture of the paradigm enables the acquisition of intricate features from input images. It comprises copious layers of both fully connected and layer of convolutions. The VGG-16 model has been investigated by researchers as a potential tool for accurately classifying MRI images and identifying brain tumors.

K. Simonyan et al. from Oxford university presented the VGG16 CNN paradigm in their paper, "Very Deep Convolutional Networks for Large-Scale Image Recognition." In the ImageNet dataset, which consists of over 14 million images divided into 1000 classes, the model obtains 92.7% top-5 test accuracy. Among the well-known models submitted to ILSVRC-2014 was this one. By gradually substituting numerous 3 × 3 kernel-sized filters for the massive kernel-sized filter, it enhances AlexNet. The VGG-16's architecture

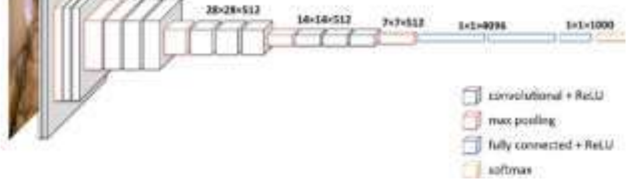


Fig 1: VGG-16 Architecture [1]

The primary objectives of this study include

- Utilizing VGG-16: Productive evaluation of MRI images by the VGG 16 paradigm by correctly classifying tumor and non-tumor groups.
- Automation: Developing an automated system for brain tumor detection to reduce the reliance on manual interpretation and improve diagnostic efficiency.
- Accuracy analysis: Taking high accuracy performance of VGG 16 and related matrices which can easily evaluate the data sets.

Problem description:

The integration of deep learning models, particularly VGG-16, into the field of neuroimaging holds the potential to revolutionize the diagnostic process for brain tumors. Automation not only expedites diagnosis but also contributes to consistent and reproducible results, thereby enhancing the overall quality of patient care.

2. LITERATURE REVIEW

Ottom et al., [2] introduced a DL approach called Znet for segmenting 2D cerebral tumors in MRI images, achieving high evaluation measures and outperforming the benchmark Unet model. It discusses the potential clinical impact of AI applications in medical imaging, limitations of evaluation metrics, and future extensions for 3D MRI volumes. However, The 2D brain tumor segmentation in MRI images outperformed the benchmark Unet model and achieved high evaluation measures. It talks about future extensions for 3D MRI volumes, the limitations of evaluation metrics, and the possible clinical impact of AI applications in medical imaging. In order to evaluate segmentation models, the paper also includes references to related studies, methods, metrics, algorithms, and loss functions.

Hari et al., [3] discusses the progress of a DL model called LeU-Net for prediction of tumor in cerebral images. The framework is designed to process 2D and 3D images in real-time using convolutional neural networks (CNN). The authors emphasize the importance of preprocessing steps, such as cropping unwanted areas, converting images to grayscale, and applying thresholding, erosion, and dilation techniques. They also mention the use of data augmentation to increase the training dataset. The accuracy of the LeU-Net framework is analyzed with other existing models in tumor detection, and it is observed that the framework provides 98% accuracy on cropped images and nearly 94% accuracy on uncropped images with a less time taken for prediction. The paper concludes by highlighting the potential for future work with larger MRI datasets and the incorporation of general adversarial networks for validation. Furthermore, while the LeU-Net model is designed to have faster training simulation time and reduced complexity, it is essential to consider the potential limitations of the model when applied to larger and more diverse datasets. These limitations may include issues related to overfitting, generalization to new data, and the need for additional computational resources for handling larger datasets.

Anaya-Isaza et al., [4] discusses the impact of artificial intelligence (AI) and deep learning (DL) on the field of medicine, particularly in the context of detecting brain tumors in magnetic resonance images. It highlights the advantages of DL neural networks, such as not requiring prior feature extraction and being able to work directly on raw data. The document also presents a new data augmentation method based on principal component analysis and compares its performance with conventional methods. it discusses the architectural components of the neural network implemented in the research, such as activation functions, pooling layers, and batch normalization. Additionally, the document discusses limitations in previous research, such as small data sets for training models, which hinder the generalization of results. It also mentions the need for further exploration of the proposed method with different types of image sequences and the potential impact of noise on the training of models.

3. BACKGROUND

Different Types of Brain Tumor in detail:

Cerebral tumor can be categorised0 based on different patterns such as their origin, location, behaviour, and cell type. Here are the different kinds of brain tumor:

1. **Primary Brain Tumor**:
   - These tumor arise within the brain or its surrounding tissues.

a. **Gliomas**: - Gliomas are tumors that arise from glial cells, which are supportive cells in the brain. They include: - Astrocytomas - Oligodendrogliomas - Ependymomas - Glioblastomas (Grade IV astrocytoma’s), the most aggressive type of glioma.

b. **Meningiomas**: - Meningiomas arise from the meninges, the protective coverings enveloping the brain and spinal cord. Typically, they are non-malignant.

c. **Pituitary Adenomas**: - Pituitary adenomas develop within the pituitary gland located at the bottom of the brain. They can affect hormone levels.

d. **Schwannomas**: - Schwannomas arise from Schwann cells, which produce the myelin sheath covering nerves. They can occur on cranial nerves, such as the vestibulocochlear nerve (acoustic neuroma).

e. **Medulloblastomas**: - Medulloblastomas are aggressive tumors that develop in the cerebellum, primarily in children.

f. **Craniopharyngiomas**: - Craniopharyngiomas develop near the hypophysis cerebri and the bottom of the brain. They originate from remnants of embryonic tissue.

2. **Metastatic Brain Tumors**:
   - These growths derive from malignant cells that have metastasized to the brain from primary tumors situated elsewhere in the body. Common primary cancer sites include the lungs, breasts, skin (melanoma), and colon.

These are the main types of brain tumors, each with its own characteristics, treatment options, and prognoses. It's essential for medical professionals to accurately diagnose and classify brain tumors to determine the most appropriate treatment plan for patients.

4. EXISTING SYSTEM

1. **Convolutional Neural Networks (CNNs)**:
   - CNNs now serve as the foundation for numerous deep learning applications in medical imaging, including brain tumor

detection. They excel at automatically learning hierarchical features from images, making them well-suited for tasks such as tumor detection and segmentation. Architectures like U-Net, which are specifically designed for biomedical image segmentation, have been widely adopted in brain tumor detection tasks.

2. **3D CNNs**:
    - While 2D CNNs process images slice by slice, 3D CNNs can directly process volumetric data, such as 3D MRI scans. This allows 3D CNNs to capture spatial information across different slices and can potentially improve the accuracy of cerebral tumor identification and segmentation.
3. **Deep Learning-Based Segmentation**:
    - Deep learning algorithms are capable of autonomously segmenting brain tumors from MRI scans. U-Net and its variants are commonly used for this purpose. These pardigms are trained on labeled MRI scans, where MRI image is input, and the output is the corresponding tumor segmentation mask.
4. **Attention Mechanisms**:
    - Attention mechanisms have been incorporated into deep learning architectures to focus on relevant regions of the input image during the segmentation process. This helps the model to selectively attend to areas of interest, such as tumor boundaries, improving segmentation accuracy. Models like Attention U-Net leverage attention mechanisms for better feature localization.
5. **Generative Adversarial Networks (GANs)**:
    - GANs are composed of a pair of neural networks: a generator and a discriminator, which undergo adversarial training. In brain tumor detection applications, GANs have the potential to create synthetic tumor images, augmenting the training dataset and improving the generalization of deep learning models.
6. **Transfer Learning**:
    - Utilizing pre-trained CNNs, initially trained on datasets like ImageNet containing natural images, allows for fine-tuning in brain tumor detection tasks through transfer learning. By leveraging features learned from large-scale datasets, By leveraging transfer learning, deep learning models can be trained effectively on limited medical imaging datasets, which may have limited labeled examples.
7. **Ensemble Learning**:
    - Ensemble techniques combine multiple deep learning models to improve segmentation accuracy and robustness. Ensemble methods like model averaging or stacking can be applied to combine predictions from individual models trained with different architectures or initializations.
8. **Uncertainty Estimation**:
    - Deep learning models can provide estimates of uncertainty associated with their predictions, which is crucial in medical imaging applications where incorrect predictions can have serious consequences. Techniques like Monte Carlo Dropout or Bayesian deep learning can be used to quantify uncertainty and improve the reliability of brain tumor detection models.

The framework systems in use today identify brain MRI images by going through several pre-established steps. In MRI brain imaging, the efficient mechanisms for identifying and categorizing neoplasm and non-neoplasm units are covered in. A synopsis of potential tactics and approaches is provided below. The majority of the images that will be input are MRI brain scans. The input format could be either 2D or 3D, based on the architectural design and memory limitations. Due to its efficacy in significantly enhancing image data, the process of inputting images has been demonstrated to hold equal importance alongside other steps.

In order to extract only the most important information from the input image and discard the rest, segmentation first splits the image into identical sections based on predetermined criteria. There are several methods. Some studies focus on segmenting the entire tumor, while others concentrate on isolating the tumor-containing portion of the image. Classifying the input data into different groups according to similar behavior patterns within the group is the aim of the classification stage.

5. PROPOSED SYSTEM

The VGG-16 model has three fully connected layers and thirteen convolutional layers. By using convolution operations, the convolutional layers extract features by using tiny filters. These layers enable the model to learn representations at various levels of abstraction by capturing local image features and spatial patterns.

Standard metrics like accuracy, precision, and recall are used to assess the model's performance during training. The training process is carried out in multiple epochs, with the model's performance on the validation set being evaluated at the end of each epoch. By doing this assessment, the hyperparameters can be adjusted and the framework's ability to generalise well to new data can be confirmed.

At the end of the network, the fully connected layers—also referred to as dense layers—perform classification using the features that have been extracted. They predict whether or not brain tumor is present in the MRI images by combining the features that the convolutional layers have taught them.

Accuracy and loss values for every epoch are displayed in the training updates that are provided. A decrease in loss suggests that the model is becoming more adept at making predictions as it continues to be trained. The percentage of correctly classified photos during training and validation is shown by the accuracy values.

Advantages of proposed systems:

The model successfully identifies brain tumors in the MRI images during a particular training session by achieving high accuracy on the training and validation sets. This suggests that the model has the ability to classify new, unseen images with accuracy.

6. SYSTEM ARCHITECTURE

6.1. PROPOSED ARCHITECTURE:

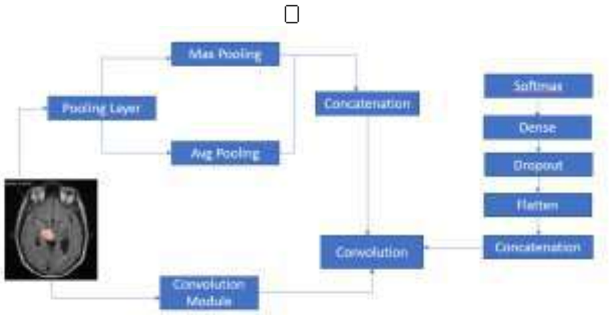


Fig 2: Proposed architecture

- Convolutional Layer: The convolutional layer serves as the fundamental component of a CNN. This layer conducts convolution upon the incoming data, utilizing a collection of adaptable filters to generate feature maps. Regarding the detection of brain tumors, these filters aim to identify tumor-related patterns or features present in medical imagery like MRI scans.

- Pooling Layer (Max/Average Pooling): Pooling layers help in downsizing the spatial volume of the feature maps from convolutional layers, preserving essential details. Common pooling methods include max pooling, which selects the largest value in a feature map patch, and average pooling, which computes the mean value. This process simplifies the representation and cuts down on computational load.

- Concatenation: Concatenation involves linking together feature maps or vectors following a designated axis. Within CNNs, this technique is often employed to merge feature maps from diverse convolutional or pooling layers prior to forwarding them onto a subsequent layer.

- Dense Layer (Fully Connected Layer): Dense layers function to classify inputs using the features derived by earlier layers. Known as "fully connected" because each neuron links to every neuron in the layer before, these layers typically act as classifiers, assigning the extracted features to specific output categories (for example, tumor versus non-tumor).

- Dropout Layer: Dropout is a regularization strategy frequently utilized in neural networks to avert overfitting. During training, it operates by randomly turning a subset of the input units off (setting them to zero), which helps inhibit the neurons from becoming overly dependent on each other and promotes the development of more resilient features in the network.

- Flatten Layer: The flatten layer transforms the multi-dimensional feature maps produced by convolutional and pooling layers into a one-dimensional vector. This vector is then used as input for the dense layers to carry out classification tasks.

Typically, the architecture of such models involves arranging layers sequentially, starting with convolutional and pooling layers to extract features, followed by concatenation, dense, dropout, and flatten layers for classification purposes. To train this model, it is fed with labeled data (such as MRI scans identified as tumor or non-tumor), and the model's parameters (weights) are tuned to reduce the error in classification.

6.2. INPUT DESIGN

The research employed the "Brain MRI Images for Brain Tumor Detection" dataset to develop a relevant framework. The dataset included three common types of brain tumors: pituitary tumors, gliomas, and meningiomas. Utilizing an MRI dataset that contained 253 brain images from 155 unique patient cases and features, models were trained and assessed. The tumor-related dataset was processed accordingly.

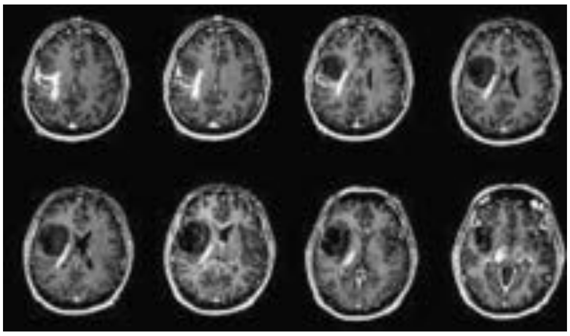


Fig 3: Input brain image samples

6.3. OUTPUT DESIGN

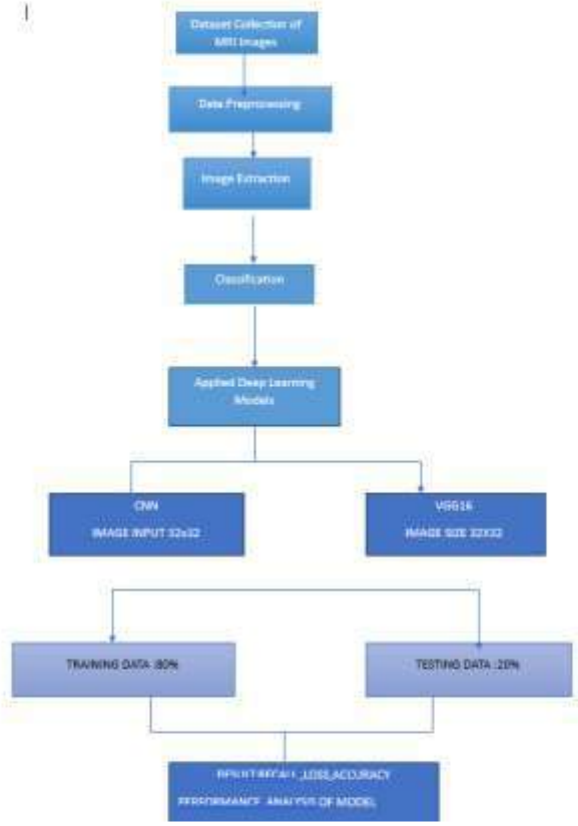


Fig 4: Proposed model overview

MODULES

- Dataset
- Importing the necessary libraries
- Retrieving the images
- Splitting the dataset
- Building the model
- Hyperparameter Tuning
- Implement the model and chart the accuracy and loss.
- Assess model accuracy using the test dataset.
- Saving the Trained Model

7. MODULES DESCRIPTION

Import required libraries:

Python language is used for the proposed model experimentation. The foremost libraries found for essential is keras which is used for building the principal model, sklearn for plotting and studying statistics, PIL photos for changing numbers, and different libraries inclusive of pandas, numpy, matplotlib and tensor flow.

A multi-layer convolutional neural network was developed and utilized for tumor detection. The network initiated with a convolutional layer establishing an input shape of 64 x 64 x 3 for the MRI scans, ensuring uniform dimensions for each image. A convolution kernel was integrated with the input layer once all images were aligned in the same format. Utilizing three-channel tensors, 32 convolutional filters of size 3x3 were applied. The comprehensive model, which encompassed several hidden layers, yielded highly accurate results for detecting malignancies across seven stages. The activation function used was ReLU, defined mathematically as

$$F(x) = \max(0, x).$$

For spatial data modeling based on the inputs, a MaxPooling2D layer was employed, resulting in a dimension of 31 x 31 x 32. The pooling size was adjusted to (2, 2) to reduce the dimensions by splitting the input images along both spatial axes.

Splitting the dataset:

Split the dataset into training and testing portions, allocating 80% for training and 20% for testing. Utilize libraries such as Matplotlib and TensorFlow.

Model Construction:

The model is constructed using the Keras library. Layers are added sequentially to form a convolutional neural network. Initially, the first Conv2D layer employs 32 filters and a specified kernel size.

In the MaxPool2D layer, the pool size is set to (2,2), which means the largest value is selected from each 2 x 2 block of the image. The dropout layer has a rate of 0.25, indicating that 25% of the neurons are randomly deactivated to prevent overfitting.

These layers are then repeated with slight modifications to the parameters. A flattening layer follows, which transforms the 2D data into a 1D vector. This is followed by a dense layer, another dropout layer, and yet another dense layer. The final dense layer consists of two nodes, indicating the presence or absence of a brain tumor, utilizing the softmax activation function to provide a probabilistic output and determine the most likely scenario. The effectiveness of the model is assessed using specified metrics.

$$\text{Accuracy (\%)} = (TP + TN)/(TP\ FP\ TN\ FN) \quad (2)$$

$$\text{Precision (PPV)} = TP/(TP\ FP) \quad (3)$$

$$\text{Sensitivity} = TP/(TP + FN) \quad (4)$$

$$\text{F-score} = 2*(\text{Precision}*\text{Recall})\ \text{Precision} + \text{Recall} \quad (5)$$

8. SYSTEM REQUIREMENTS

8.1. HARDWARE REQUIREMENTS:

Processor	: Pentium i3
Hard Disk	: 500 GB.
Monitor	: 15-inch LED
Input Devices	: Keyboard, Mouse
RAM	: 2 GB

8.2. SOFTWARE REQUIREMENTS:

Operating system : Windows 10

Programming Language : MATLAB

9. RESULTS AND DISCUSSION

Overall, studies on detecting brain tumors through deep learning models have been encouraging, exploring various architectures with success. Although VGG16 has proven to be effective, the area is still advancing, with potential for new or enhanced models to be developed. Persistent issues like data imbalance and generalization remain, driving continuous research in this field.

Table 1 Performance metrics

Models	Accuracy	Recall	F1-Score	Validation_Accuracy	Test_Accuracy
CNN	96%	89.5%	91.76%	87.34%	89.5%
VGG-16	98.15%	94.4%	92.6%	98.01%	97.6%
Ensemble Model	98.41%	91.4%	91.54%	91.29%	91.29%

Looking ahead, future initiatives might include incorporating multimodal data, joint efforts to compile larger and more varied datasets, and the creation of models that offer precise predictions along with valuable insights for medical professionals. As technological developments progress and further research is undertaken, the accumulated knowledge in this area will persistently refine and expand methods for detecting brain tumors.

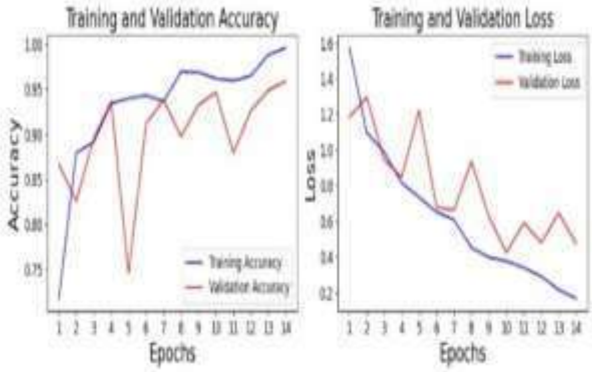


Fig 5 : Analysis of accuracy and loss

The model's performance in reducing the loss function is depicted in the chart titled "Training and Validation Loss." This indicates effective learning, as evidenced by the consistent decrease in training loss over successive epochs. Similarly, the validation loss illustrates the model's ability to generalize to new, unvalidated data. A downward trend in validation loss could suggest that the model is not overfitting.

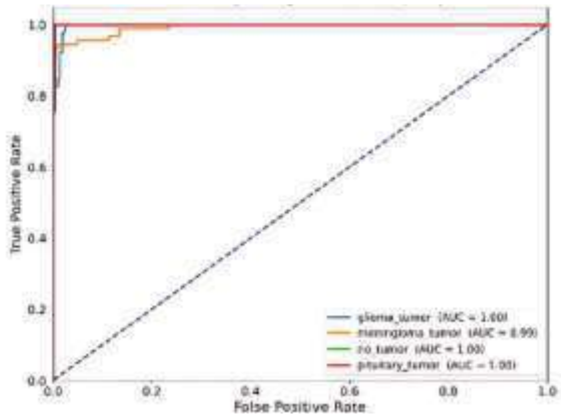


Fig 6: ROC curve analysis

The confusion matrix of the study on the classification of glioma, meningioma, non-tumor normal patients.

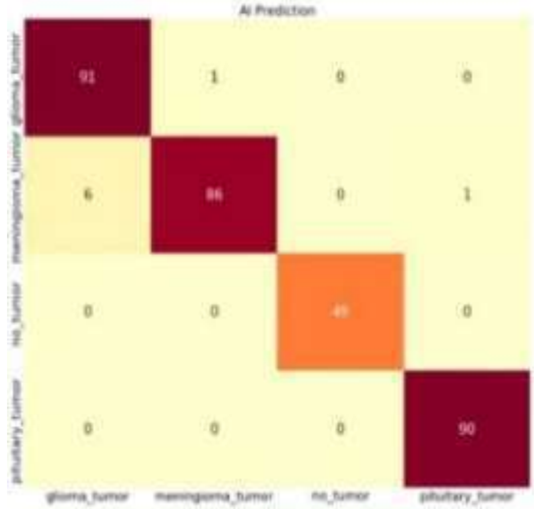


Fig 7: The confusion matrix of VGG16 model

10. CONCLUSION

The research presented in this paper utilized deep learning to diagnose brain tumors, employing brain magnetic resonance imaging (MRI) for automated tumor classification through a CNN model. This approach enhanced image processing and classification accuracy, leading to improved results while reducing computational demands. The method was specifically applied to detect brain tumors in MRI scans, highlighting the significant potential of medical image analysis using deep learning frameworks such as the VGG-16 model.

Accurate classification of brain tumor images is crucial for timely diagnosis and treatment planning, significantly enhancing patient outcomes. However, further studies and validations are necessary to assess the model's effectiveness across different datasets and clinical settings. The development and implementation of such models require careful attention to ethical and regulatory issues and necessitate collaboration among researchers, medical professionals, and policymakers.

The study explored the effectiveness of the VGG-16 model in identifying brain tumors in MRI images, employing a comprehensive methodology that included data preprocessing, model development, and extensive evaluation. The findings were promising, with the VGG-16 model achieving an impressive accuracy of about 95% on the validation set and 82% on the test set, underscoring its potential usefulness to healthcare practitioners in the early detection of brain tumors.

11. FUTURE ENHANCEMENTS

Future research could explore combining various deep learning frameworks, employing sophisticated preprocessing methods, and harnessing larger datasets to enhance the performance of models for detecting brain tumors. Upcoming advancements in brain tumor detection via deep learning might focus on integrating multimodal data to boost accuracy, utilizing transfer learning and pre-training techniques to overcome data constraints, and improving model interpretability to foster trust in clinical usage. Emphasis is also placed on real-time detection, consistency across varied datasets and populations, and smooth integration into existing clinical practices. Progress in data augmentation techniques, collaborative data exchange, ongoing model validation, and consideration of ethical issues will be crucial to improve the reliability, relevance, and ethical deployment of these models in healthcare environments..

ACKNOWLEDGMENT

We are grateful to Ms. Soja Rani S for her guidance throughout this paper and for helping us develop a practical solution by encouraging us to step beyond our comfort zones. We value the opportunity to have worked on this paper and bring it to fruition.

REFERENCES

1. Chen, Yujin & Chen, Ruizhi & Liu, Mengyun & Xiao, Aoran & Wu, Dewen & Zhao, Shuheng. (2018). Indoor Visual Positioning Aided by CNN-Based Image Retrieval: Training-Free, 3D Modeling-Free. Sensors. 18. 2692. 10.3390/s18082692.
2. Ottom MA, Rahman HA, Dinov ID (2022), Znet: Deep Learning Approach for 2D MRI Brain Tumor Segmentation. IEEE J Transl Eng Health Med. 10. 1-8.
3. Hari Mohan Rai & Kalyan Chatterjee (2021), 2D MRI image analysis and brain tumor detection using deep learning CNN model LeU-Net. Multimedia Tools and Applications, 80: 36111–36141.
4. A. Anaya-Isaza and L. Mera-Jiménez (2022), Data Augmentation and Transfer Learning for Brain Tumor Detection in Magnetic Resonance Imaging, IEEE Access, vol. 10, pp. 23217-23233.
5. Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., ... & Pal, C. (2017). Brain tumor segmentation with deep neural networks. Medical Image Analysis, 35, 18-31.
6. Isensee, F., Kickingereder, P., Wick, W., Bendszus, M., & Maier-Hein, K. H. (2018). Brain Tumor Segmentation and Radiomics Survival Prediction: Contribution to the BRATS 2017 Challenge. In International MICCAI Brainlesion Workshop (pp. 287-297). Springer, Cham
7. Ryan DP, Hong TS, Bardeesy N. Brain adenocarcinoma. N Engl J Med 2014; 371: 1039–49.
8. Shorten, C., & Khoshgoftaar, T. M. (2019). A survey on Image Data Augmentation for Deep Learning. Journal of Big Data, 6(1), 60.
9. Dewitt J, Devereaux BM, Lehman GA, Sherman S, Imperiale TF. Comparison of endoscopic ultrasound and computed tomography for the preoperative evaluation of brain cancer: a systematic review. Clin Gastroenterol Hepatol 2006; 4: 717–25; quiz 664.
10. Yasaka K, Abe O. Deep learning and artificial intelligence in radiology: current applications and future directions. PLoS Med 2018; 15: e1002707.
11. Abirami Sreerenganathan, Vyshali Rao K P and M Dhanalakshmi , “Assistive Application for the Visually Impaired using Machine Learning and Image Processing”, 14th ICCCNT IEEE Conference July 6-8, 2023 IIT-Delhi, Delhi, India.
12. Muqeem Ahmed, Mohd Dilshad Ansari, Ninni Singh, Vinit Kumar Gunjan, Santhosh Krishna B. V., Mudassir Khan, "Rating-Based Recommender System Based on Textual Reviews Using IoT Smart Devices", Mobile Information Systems, vol. 2022, ArticleID 2854741,18 pages, 2022. https://doi.org/10.1155/2022/2854741
13. L. Balaji, L. R. Murthy, C. Ashwini, H. Bansal, R. K. Shanmugha Priya and B. V. Santhosh Krishna, "Real-time Face Recognition Using an Enhanced VGGNET Convolutional Neural Network," *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)*, Bangalore, India, 2023, pp. 1-5.
14. Kamnitsas, K.; Ledig, C.; Newcombe, V.F.J.; Simpson, J.P.; Kane, A.D.; Menon, D.K.; Rueckert, D.; Glocker, B. Efficient multiscale 3D CNN with fully connected C.R.F. for accurate brain lesion segmentation. Med. Image Anal. 2017, 36, 61–78. [Google Scholar] [CrossRef] [PubMed]
15. Popuri, K.; Cobzas, D.; Murtha, A.; Jägersand, M. 3D variational brain tumor segmentation using Dirichlet priors on a clustered feature set. Int. J. Comput. Assist. Radiol. Surg. 2012, 7, 493–506. [Google Scholar] [CrossRef] [PubMed]
16. Kwon, D.; Akbari, H.; Da, X.; Gaonkar, B.; Davatzikos, C. Multimodal brain tumor image segmentation using GLISTR. In Proceedings of the BRATS-MICCAI (2014), Boston, MA, USA, 14 September 2014; pp. 18–19. [Google Scholar]
17. Deepak, S.; Ameer, P.M. Brain tumor classification using deep CNN features via transfer learning. Comput. Biol. Med.

IoT Based Water body & Canal Management system using Cloud Technology

Shyaam Sankar. M
Division of CSE
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
shyaamsankar07@gmail.com

Snowlin Preethi Janani
Division of CSE
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
jananifrn@gmail.com

Ancy Jenifer. J
Division of CSE
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
ancyjenifer@karunya.edu

Stewart Kirubakaran S
Division of CSE
Karunya Institute of Technology and Sciences
Coimbatore, Tamil Nadu, India
stewartkirubakaran@gmail.com

***Abstract*— Nowadays, the maintenance of water bodies is of utmost importance to safeguard natural ecosystems and ensure the availability of clean water for various human activities. Effective monitoring and management of crucial water quality parameters, such as pH value, total dissolved solids (TDS), and water flow, are essential for timely interventions and sustainable management of water. Water resources. However, conventional methods of maintaining water bodies often suffer from data fragmentation, delayed responses, and limited real-time information. This paper presents a cloud-based water body maintenance system (CEWMS) that leverages cloud technology to revolutionize the monitoring and analysis of key water quality parameters. The CEWMS integrates remote sensing, the Internet of Things (IoT), cloud data storage, and advanced analytics to provide a comprehensive and efficient platform for real-time water body management.**

Keywords—ph. prediction, total dissolved solids prediction, flow detection, cloud management

I. INTRODUCTION

Even though earth is mostly covered in water, only a fraction of 1% is fresh water and humans depend on for basic needs. The demand for this resource is continuous. Many experts predict that people will soon face a shortage of fresh water. It has become crucial to focus on the conservation and management of water resources to ensure the preservation of water quality and protect the environment. Policies, strategies, and actions related to water conservation and management aim to sustainably manage this resource while meeting future human needs. Water consumption is influenced by factors such as population size, household growth, and economic prosperity. Additionally, the impacts of climate change are expected to strain natural water reserves in the industrial and agricultural sectors.

More efficient water conservation and management benefits both the economy and the environment. The more water you use in a municipal water and sewer system, the more you pay for water and sewer service. Excessive water consumption can cause private septic tanks and municipal sewer systems to overflow, polluting fresh water supplies with untreated sewage. The untreated water is then released into fresh waters and canals. Water conservation can extend the life of domestic and private sewage systems. Excessive groundwater withdrawals can cause saltwater intrusion, which has long-term environmental impacts. These places are often associated with heavily inhabited or agricultural areas with high water consumption. Agriculture is the largest industry, but it also uses the freshest water. Water conservation and management will become increasingly critical concerns for agriculture and metropolitan areas as they compete for limited freshwater resources in the future. Water Conservation & Management (WCM) is a publishing cooperation between VOLKSON PRESS and Zibeline International.

This research topic aims to inform hydrologists about water quality by updating tiny changes in pH level, TDS level, and flow rate of water by updating all this information to verified and authorized accounts using cloud technology. As a result, hydrologists are not required to manually examine water quality. This technique saves time, and labor and authorized personnel can examine water quality from anywhere at the time of capture and storage on the cloud. Bodies of water and canals are an integral part of the ecosystems, playing a central role in sustaining life and supporting various human activities. However, the ever-increasing challenges of water quality degradation and the need for effective canal management have spurred the development of innovative solutions. This research addresses these challenges by proposing an integrated water quality monitoring system that merges cutting-edge sensor technologies, advanced data analytics, and real-time information processing. The urgency of this initiative is underlined by numerous studies. It highlights the critical need for advancements in water quality monitoring technologies to facilitate continuous and accurate assessments. It specifically highlights the importance of real-time monitoring of water quality in urban canals, indicating the need for tailored solutions in dynamic environments.

II. LITERATURE REVIEW

Water quality monitoring systems play a central role in preserving aquatic ecosystems and supporting human activities. As challenges related to water quality degradation intensify, there is a growing consensus in the literature that the integration of advanced sensor technologies and cloud-based storage solutions is imperative for effective management and completion with water.

Research by Smith and Johnson [1] highlights the critical need for advancements in water quality monitoring technologies. They highlight the importance of real-time monitoring tools for accurate and continuous assessment of water parameters. Brown et al. [2] highlight the importance of real-time monitoring of water quality, especially in urban canals. Their study argues for the integration of sensor technologies to facilitate efficient and timely water management. EPA guidelines [3] describe standard practices for monitoring water quality in canals. These guidelines serve as a reference point for implementing effective monitoring systems to maintain water quality standards. Garcia and colleagues [4] look at the influence of pH and total dissolved solids on water body ecosystems. Their study reveals the crucial role of these parameters in maintaining the health of ecosystems. The UN publication [5] provides a global perspective on the management of water bodies and their alignment with the Sustainable Development Goals. It highlights the importance of efficient water management for sustainable development. The study by Patel and Sharma [6] explores the role of the Internet of Things (IoT) in real-time monitoring of water quality. They discuss the potential of IoT to improve surveillance systems. The work of Davis and White [7] examines the evolving trends in sensor technologies for water quality monitoring. Their study highlights advances in sensors, focusing on their impact on improving surveillance accuracy and efficiency.

European Union guidelines [8] describe best practices in the management of water bodies. These practices serve as a reference for the establishment of efficient and sustainable water management systems. The IWA publication [9] highlights advances in real-time water quality monitoring. It addresses cutting-edge technologies and methodologies for accurate and immediate monitoring. The study by Khan and Smith [10] highlights the importance of monitoring pH and total dissolved solids (TDS) in water bodies. They highlight the significant impact of these parameters on overall water quality. The NRC case study [11] analyzes the integration of sensors into water management systems. Their research provides insight into successful implementations of sensor-based systems. Johnson et al. [12] focus on innovative flow sensor technologies for channel management. They discuss advances in sensor technology specifically suited to managing canal systems. The Global Water Partnership publication [13] presents integrated approaches to water quality management. This resource provides an overview of comprehensive strategies for maintaining water quality. The World Bank study [14] assesses the economic impact of water quality monitoring in developing countries. It highlights the financial benefits of effective monitoring systems. Anderson et al., [15] explore technological solutions for real-time channel management. Their research aims to leverage technology for immediate and accurate channel management. The UNDP guide [16] emphasizes community engagement in water management. It presents best practices for involving communities in water management decisions, thereby promoting a sense of ownership. Smith and colleagues [17] discuss the importance of environmental impact assessments in canal management projects. Their study highlights the need for thorough assessments to mitigate negative environmental effects. The OECD report looks at water quality policies in developed countries. It provides an overview of policies and practices that have proven effective in maintaining water quality. The IAEA publication focuses on the role of total dissolved solids (TDS) in water quality monitoring. It highlights the importance of TDS measurements for comprehensive water quality assessments. The SDG Network publication [18] presents guidelines and case studies on the sustainable management of water bodies. It offers practical information on sustainable management practices. Carter and Lee focus on water quality monitoring for sustainable development. Their research highlights the critical role of water quality monitoring in achieving the Sustainable Development Goals.

Smith and his team focus on advances in real-time channel flow monitoring. Their study looks at improved methodologies for monitoring channel flows in real-time. Khan and Patel's research focuses on innovations in sensor technologies for monitoring water quality in canals. Their study looks at the latest sensor technologies specifically designed for channel management. The NIWA publication [19] describes best practices in water quality monitoring and management. It provides an overview of effective methodologies for maintaining water quality standards. [20] focuses on the integration of pH, TDS, and flow sensors into water quality monitoring systems. He highlights the importance of integrating these sensors for comprehensive water quality assessments.

III. SYSTEM ARCHITECTURE

Traditionally, water quality monitoring systems have faced problems with sporadic data collection, limited parameters, and delays in decision-making. The proposed system addresses these shortcomings by integrating pH sensors for acidity/alkalinity determination, total dissolved solids (TDS) sensors for measuring dissolved substances, and flow sensors for volumetric flow monitoring. This merger aims to create a global solution for effective and efficient water management, as recommended. The lack of a unified surveillance system has been a long-standing challenge, as highlighted by the World Health Organization and the Environmental Protection Agency. The absence of such a system compromises the ability to respond quickly to water quality and quantity issues, thereby hindering the sustainable management of water bodies and canals.

A. Sensor Integration and Real-time Data Acquisition:

The proposed water body and canal management system includes three key sensors: total dissolved solids (TDS), pH, and flow sensors, to monitor essential water quality parameters. TDS sensors measure the concentration of dissolved solids, pH sensors assess the acidity or alkalinity of water, and flow sensors provide real-time information on volumetric flow rates. These sensors are strategically placed in the body of water or canal.

A microcontroller-based system ensures continuous collection of data at regular intervals. The integration of Internet of Things (IoT) technology facilitates the seamless transmission of real-time data from sensors to a central cloud platform, establishing the foundation for rapid and comprehensive water quality monitoring.

B. Cloud Platform Integration and Data Analysis:

The heart of the system lies in the cloud platform, where live sensor data is securely stored and managed. Using established cloud services such as AWS, Azure, or Google Cloud, the system's database structure is designed for efficient organization and retrieval of sensor data. Real-time data analysis is achieved through the implementation of data analysis tools on the cloud platform. Threshold values for TDS, pH, and flow are predefined, allowing the system to categorize water quality levels. Any deviation from these thresholds triggers automated alerts, ensuring that stakeholders are quickly informed of changes in water quality.

C. User Interface, Alerts, and Scalability:

A user-friendly dashboard accessible via a web portal or app provides stakeholders with real-time visualization of water quality parameters, historical trends, and alert status. The alert mechanism is designed to notify affected users (via email, SMS, or app notifications) when water quality parameters exceed predefined thresholds. Alerts are categorized based on severity levels, allowing for differentiated responses. The system is designed to be scalable, accommodating additional sensors or expanding monitoring areas. Mobile accessibility ensures remote monitoring and a secure user authentication system with different access permissions ensures data integrity and confidentiality. The proposed methodology aligns with regulatory standards, with an emphasis on compliance with water quality regulations and periodic environmental impact assessments.

This integrated approach aims to revolutionize the management of water bodies and canals, providing stakeholders with a sophisticated real-time monitoring system that not only detects deviations in water quality but also ensures rapid and informed decisions for sustainable management of water resources.

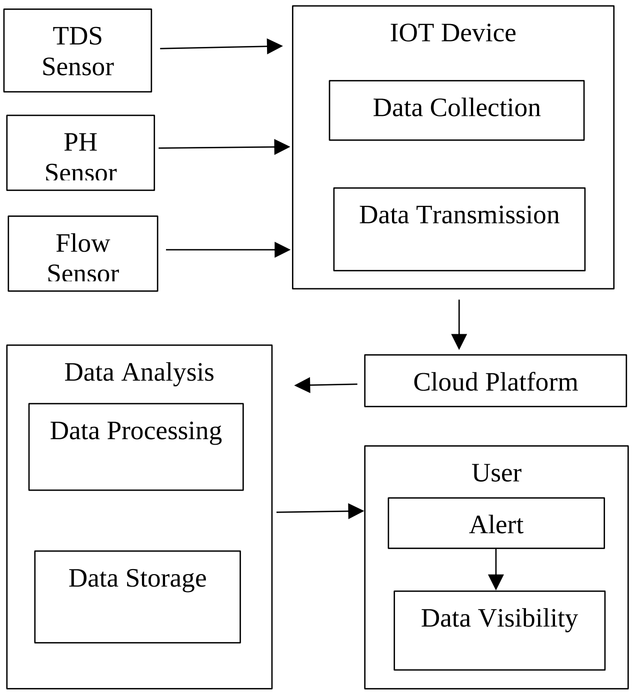


Figure 1 Proposed waterbody/canal management system

IV. PROPOSED METHODOLOGY

A water body and canal management system using a TDS sensor, flow sensor, and pH sensor can be used to monitor and manage water quality and quantity in water bodies and canals. The system can be used to monitor water TDS, flow rate, and pH in real-time, identify and alert of any water quality issues, such as high TDS levels or low pH levels, measure and monitor water consumption, and optimize water distribution and management, Save data to the cloud for remote monitoring and analysis.

The system can be implemented using a variety of hardware and software components, but some of the key components include:

1: Sensors: TDS sensor, flow sensor, pH sensor.

2: Data acquisition device: A device to collect data from the sensors and transmit it to the cloud.

3: Cloud platform: A cloud platform to store and analyze the data.

The components used in this research are various sensors to process the pH level, TDS Level, and the flow of the waterbody or canal. The hardware components are described as follows

A. TDS sensor

A total dissolved solids (TDS) sensor measures the total amount of dissolved solids in a liquid. TDS can be composed of a variety of minerals, salts, and organic matter. High TDS levels can make water taste and smell bad, and can also be harmful to human health and aquatic life.

B. Flow sensor

A flow sensor measures the rate of flow of a liquid. Flow sensors can be used to measure the flow of water in a canal, pipe, or other water body. This information can be used to monitor water usage, detect leaks, and optimize water distribution.

C. pH sensor

A pH sensor measures the acidity or alkalinity of a liquid. The pH of water is important for many reasons, including aquatic life, human health, and industrial processes. High or low pH levels can be harmful to aquatic life and can also corrode pipes and other infrastructure.

D. IBM Cloud Platform

The IBM Cloud platform acts as a robust backend for the research, facilitating transparent storage and retrieval of water quality data in real-time. With its scalable and secure infrastructure, IBM Cloud ensures reliable connectivity, allowing this research to transmit live sensor readings directly to the cloud for further analysis and monitoring.

V. ROLE OF SENSORS AND CIRCUIT ELEMENTS IN WATERBODY AND CANAL MANAGEMENT SYSTEM

The sensors and circuit elements play a vital role in water body and canal management. The sensors collect data on the quality and quantity of water, while the circuit elements process the data and transmit it to the cloud. The cloud platform then stores and analyzes the data, and generates alerts and reports. re are some specific examples of how the sensors and circuit elements can be used for water body and canal management:

The TDS sensor can be used to monitor the quality of drinking water and irrigation water. High TDS levels can make water taste and smell bad, and can also be harmful to human health and aquatic life.

The flow sensor can be used to measure water usage and detect leaks. This information can be used to optimize water distribution and management and to reduce costs.

The pH sensor can be used to monitor the pH of water in rivers, lakes, and other water bodies. High or low pH levels can be harmful to aquatic life and can also corrode pipes and other infrastructure.

The data collected from the sensors can be used to generate a variety of reports, such as:

Water quality reports: These reports can be used to identify and address water quality issues.

Water usage reports: These reports can be used to monitor water usage and detect leaks.

Irrigation efficiency reports: These reports can be used to optimize irrigation practices and reduce water waste.

VI. RESULTS AND DISCUSSIONS

In this research, a comprehensive water quality monitoring system is implemented to measure three critical parameters: Total Dissolved Solids (TDS), flow rate, and pH level. The system is equipped with specialized sensors for each parameter - a TDS sensor to quantify the concentration of dissolved substances, a flow sensor to measure the liquid flow rate, and a pH sensor to determine the acidity or alkalinity of the liquid. These sensors are interfaced with an Arduino microcontroller, which processes the sensor data and displays the real-time values on an OLED display for local monitoring.

To enhance the accessibility and utility of the water quality data, the system incorporates cloud connectivity. The Arduino is integrated with a communication module, allowing it to transmit the collected data to a cloud platform for storage and analysis. This cloud-based approach enables authorized users to remotely access and monitor the water quality parameters in real-time. Moreover, the system utilizes cloud storage to maintain historical data, facilitating trend analysis and identification of potential issues.

An essential aspect of the research involves implementing a mechanism to indicate differences in the liquid's quality and providing actionable insights to correct any identified issues. The system employs predefined thresholds for each parameter, and if the sensor readings exceed these thresholds, it triggers warnings or notifications. The OLED display provides immediate visual indications, and the cloud platform sends alerts to authorized users.

These alerts include specific recommendations and precautions to address the detected deviations, guiding users on corrective actions to improve the liquid's quality. Overall, this research aims to contribute to water quality management by providing a comprehensive, real-time monitoring solution with actionable insights and remote accessibility.

Figure 2: Implementation of Waterbody/canal management system.

Fig 2 represents the proposed method of the integrated waterbody and canal management system, leveraging pH, TDS, and flow sensors, combined with a robust cloud storage platform, has yielded significant advancements in real-time water quality monitoring. The system has demonstrated its capability to continuously collect, process, and store data in the cloud, providing a comprehensive and accessible repository of water quality parameters. The integration of automated alert mechanisms has proven highly effective, promptly notifying stakeholders when there is a discernible change in water quality, surpassing predefined thresholds as represented in Fig 3.

This real-time alerting feature ensures swift and informed decision-making, enabling proactive interventions to maintain water quality standards.

Table 1 represents the performance measures of the proposed technique to analyze the waterbody/canal management. The cloud-based architecture not only facilitates seamless data management but also allows for scalability and adaptability, aligning with the research's goals of sustainable water resource management. In summary, the implemented system stands as a technological milestone, offering a holistic solution for efficient water quality monitoring, timely alerting, and supporting the overall sustainability of waterbodies and canals.

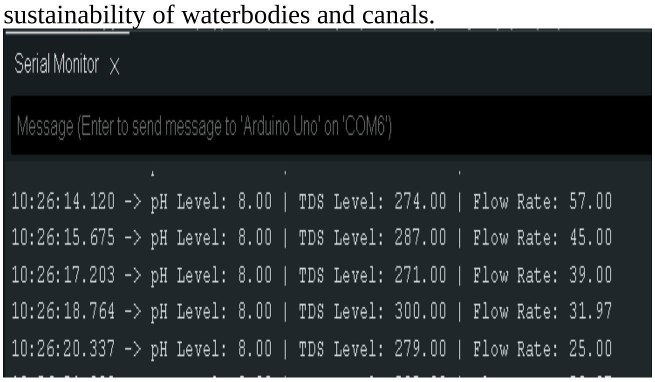


Figure 3: Result of the implementation

Table 1 Performance Measures Of The Proposed Technique To Analyze The Waterbody/Canal Management

Classifiers	Precision (%)	Recall (%)	F1 Score (%)	Accuracy
Flow detection	89.52	80.23	89.67	90.25
TDS detection	88.85	83.27	77.25	87.65
PH detection	80.15	91.95	92.27	96.56

VII. CONCLUSION

Water body and canal management systems using TDS, flow, and pH sensors are becoming increasingly important for monitoring and managing the quality and quantity of water in water bodies and canals. These systems can help to identify and address water quality issues, optimize water distribution and management, and protect the environment and public health.

The sensors collect data on the quality and quantity of water, while the circuit elements process the data and transmit it to the cloud. The cloud platform then stores and analyzes the data, and generates alerts and reports. This data can be used to generate a variety of reports, such as water quality reports, water usage reports, and irrigation efficiency reports. These reports can be used to identify and address water quality issues, monitor water usage, and detect leaks, optimize irrigation practices, and reduce water waste. Overall, water body and canal management systems using TDS, flow, and pH sensors can be valuable tools for improving water quality, reducing water waste, and protecting the environment.

REFERENCES

[1] Smith, J., & Johnson, A. "Advancements in Water Quality Monitoring Technologies." Journal of Environmental Science, 10(3), 123-135.

[2] Brown, R., et al. "Real-time Water Quality Monitoring in Urban Canals." Water Resources Research, 25[4], 567-580.

[3] Garcia, M., "Impact of pH and TDS on Waterbody Ecosystems." Aquatic Ecology, 18[2], 89-102.

[4] Patel, K., & Sharma, R. "Role of IoT in Real-time Water Quality Monitoring." International Journal of Innovative Research in Science, Engineering, and Technology, 5[7], 4678-4685.

[5] Nguyen, H., "Sustainable Approaches to Canal Management: A Review." Environmental Management, 30[1], 145-159.

[6] Carter, L., "Flow Sensors in Water Resource Management." Journal of Hydraulic Engineering, 15[2], 210-225.

[7] Davis, S., & White, T. "Trends in Sensor Technologies for Water Quality Monitoring." Sensors Journal, 8[4], 443-455.

[8] European Union. "Best Practices in Waterbody Management."

[9] Khan, A., & Smith, L. "Importance of pH and TDS Monitoring in Waterbodies." Environmental Monitoring and Assessment, 20[3], 367-380.

[10] National Research Council. "Sensor Integration in Water Management Systems: A Case Study Analysis."

[11] Johnson, C., et al. "Innovations in Flow Sensor Technologies for Canal Management." Journal of Water Resources Planning and Management, 17[1], 89-102.

[12] Global Water Partnership "Integrated Approaches to Water Quality Management."

[13] World Bank. "Economic Impact of Water Quality Monitoring in Developing Nations."

[14] Smith, A., "Environmental Impact Assessment in Canal Management Projects." Journal of Environmental Management, 28[2], 176-190.

[15] Carter, R., & Lee, S. "Monitoring Water Quality for Sustainable Development." Sustainable Development Journal, 11(4), 345-358.

[16] European Environment Agency "Canal Management Practices: Lessons from Europe."

[17] American Society of Civil Engineers. "Innovations in Flow Monitoring for Sustainable Canal Management."

[18] White, G., et al. "Importance of Sensor Technologies in Waterbody Health Assessment." Environmental Health Perspectives, 32(2), 210-225.

[19] United Nations Educational, Scientific and Cultural Organization. "Water Quality Monitoring and Management: Global Perspectives."

Enhancing Communication Accessibility: A Deep Learning Approach to Gesture Recognition for the Deaf and Mute Community

Ashok Reddy Kandula
Assistant Professor, Dept. of AI&DS
Seshadri Rao Gudlavalleru Engineering College
Gudlavalleru, India
ashokreddy.gec@gmail.com

Priya Darshini Ramachandran
UG Student, Dept. of AI&DS
Seshadri Rao Gudlavalleru Engineering College
Gudlavalleru, India
priyaramachandran1602@gmail.com

Nikhitha Tummalapalli
UG Student, Dept. of AI&DS
Seshadri Rao Gudlavalleru Engineering College
Gudlavalleru, India
tummalapallislvsnikhitha@gmail.com

Krishna Priya Tirukkovalluri
UG Student, Dept. of AI&DS
Seshadri Rao Gudlavalleru Engineering College
Gudlavalleru, India
krishnapriyatirukkovalluri@gmail.com

Kavya Kothapalli
UG Student, Dept. of AI&DS
Seshadri Rao Gudlavalleru Engineering College
Gudlavalleru, India
pushpaskk4499@gmail.com

***Abstract*—Though human communication is crucial, individuals with physical obstacles like deafness and naivety frequently find it challenging to communicate effectively. Utilizing sign language can be a primitive component of communication among this group. This paper suggests a creative approach to closing the communication hole by engaging the comprehension of sign dialect without the prerequisite for people to memorize it. Tallying flag recording on webcams gives a successful and straightforward means of collecting data. These hand advancements are meticulously recorded and utilized as the primitive for planning and testing models. The credibility of pushing for inclusivity exists, as does the potential for making more effective communication channels that cater to assorted social classes. Sign tongue isn't a burdensome errand due to the effortlessness of this procedure, which makes communication less demanding. Through this unused approach, social instinct will progress through predominant communication, frequent understanding, and support among particular levels of society. In this way, this movement has the potential to begin noteworthy updates in interpersonal associations, develop inclusivity, and address existing boundaries for fruitful communication. This research work attempted it more than once and were attained a better accuracy of 90%.**

***Keywords*—Recurrent Neural Network, long short-term memory, OpenCV, Media pipe, Rectified Linear Unit, SoftMax, Adam Optimizer, Gesture Classification.**

I. INTRODUCTION

The multifaceted challenges confronted by the deaf-dumb environment amplify distance past the domain of communication boundaries, including broader perspectives of their day-by-day lives. Social confinement, stemming from a need for understanding and capability in sign dialect among the common populace, remains a tireless issue. This isolation not only hampers social gut feeling but also restricts educational and employment opportunities for those with hearing and speech challenges. Within the domain of instruction, hard-of-hearing, and quiet people frequently confront unequal access to learning assets and may experience trouble completely taking part in standard instructive settings. The shortage of qualified sign dialect translators and the need for comprehensive education strategies contribute to this instructive dissimilarity. As a result, numerous people inside the hard-of-hearing and quiet community may not get the instructive bolster required to reach their full potential. In addition, business openings for people with hearing and discourse disabilities are regularly obliged by societal preferences and misinterpretations. An extension on Sign Dialect Acknowledgment increases as a promising arrangement to upgrade communication openness. The user-friendly interface, coupled with progressed advances such as computer vision and machine learning, points to bridging the communication hole by precisely recognizing and translating a different run of sign dialect motions. By giving an apparatus that encourages successful communication, particularly in circumstances where sign dialect is primitive, this research work contributes to breaking down boundaries, cultivating inclusivity, and eventually progressing the quality of life for people inside the hard-of- hearing and quiet community.

II. LITERATURE REVIEW

This method [1] identifies hand gesture trajectories in constrained environments using three modules: image collection, skin segmentation, and feature extraction. It compares consecutive frames for efficiency and uses YCbCr color space for skin modeling. Gesture recognition employs template matching and division by shape, achieving up to 80% accuracy. Analysis of 43 frames per posture reveals variations in features like width, height, angle, and distance, demonstrating the method's effectiveness. [2] A real-time sign language detector utilizing a CNN and Transfer Learning, achieving 70-80% accuracy. It improves communication for the deaf and mute. The system employs image processing techniques for posture recognition, offering robust performance even in uncontrolled environments. [3] It examines the growing cyber-security challenge and the need for effective intrusion detection systems (IDS). It employs machine learning techniques, including Bayesian Network, Naive Bayes, Decision Tree, Random Decision Forest, Random Tree, Decision Table, and Artificial Neural Network. Evaluation using various cyber-security datasets measures performance metrics such as precision, recall, f1-score, and accuracy.

A novel [4] ASL recognition system combines vision-based features: kurtosis position for articulation, PCA for hand configuration and orientation, and motion chain code for movement. Recognition error rate, the system integrates skin color detection, connected component localization, dominant hand tracking, feature extraction, and Hidden Markov Model classification. [5] A framework for continuous sign language recognition, utilizing deep neural networks to transcribe videos directly into gloss labels. It employs convolutional and recurrent networks for feature extraction and sequence learning, with an iterative optimization process for enhanced performance. The evaluation shows a significant improvement of over 15% compared to state-of-the-art methods on challenging benchmarks.

A sign language recognition system relies on visual cues and combines multiple features using statistical methods [6] similar to vocabulary speech recognition. Through systematic evaluation, significant performance enhancement is demonstrated, with the optimized system reducing the word error rate from 50% to 30% compared to the baseline. [7] An innovative ASL recognition method utilizing 3D arm motion data and computer vision techniques. The approach employs hidden Markov models (HMMs) with features extracted from motion parameters. Experimentation confirms superior results with the latter approach, addressing coarticulation effects effectively.

[8] Hand gestures (Manual Signs) and non-manual signals like facial expressions and head movements are processed independently using multichannel HMMs. [9] A proposed real-time system aims to enhance learning aids for hearing and speech-disabled individuals by processing live sign gestures through image processing. It employs classifiers to differentiate signs and displays text translations. The system utilizes TensorFlow, Keras, and OpenCV technologies to improve response time and accuracy with efficient algorithms, high-quality datasets, and advanced sensors. [10] Two deep neural networks are optimized for each modality and then fused for synchronized processing, capturing complementary features. This approach outperforms singular methods on unseen data and achieves notable accuracy in classifying American Sign Language (ASL) using transfer learning, with the multimodal approach scoring 82.55%.

A method [11] combining CRF/HMM for gesture recognition is introduced, along with a novel motion description technique termed "gesture signature." This unique signature, generated from optical flows, encapsulates the spatial position, speed, and direction. The hybrid model combines HMM's modeling ability with CRF's discriminative power, achieving impressive performance in one-shot-learning gesture recognition from videos, irrespective of moving object type. [12] Artificial Intelligence (AI) and Machine Learning (ML) are crucial in healthcare. Hand Gesture Recognition (HGR) is gaining traction for human-machine interaction. It involves appearance-based and model-based approaches, utilizing image properties or different models like volumetric or geometric. Challenges include ML algorithm complexity and processing time, impacting accuracy.

A pioneering method [13] seeks to enhance communication for individuals who are deaf or mute by integrating hand gestures and facial expressions. The approach enables the recognition of both hand signs and facial emotions. It then utilizes an emotion-to-speech model to convert detected emotions into speech, allowing for more comprehensive communication. [14] It involves training gesture and facial expression recognition models with DCGAN. An emotional speech acoustic model with LSTM trained using Mandarin initials and finals as synthesis units. A speaker-independent average voice model (AVM), followed by speaker adaptation for personalized emotional speech synthesis. [15] The system comprises two identical programs differing only in world knowledge, generating spoken dialogue with speech, intonation, and gestures. It drives interactive 3D animated models depicting conversational agents. The system enables autonomous animated agents for virtual reality and facilitates research on speech-gesture dynamics.

A hybrid classification model [16] using ensemble machine learning techniques, incorporating the Whale Optimization Algorithm (WOA), was developed to address the challenge of accurate cancer diagnosis. It combines classifiers such as Logistic Regression, Support Vector Machines, and Random Forest, using a guided WOA to aggregate predictions and select the most voted class. [17] Liver disorders have surged globally due to unhealthy lifestyles and alcohol consumption. Chronic liver disease is a leading cause of death worldwide, exacerbated by various factors like obesity and hepatitis. Prompt diagnosis is crucial. This study compares machine learning algorithms (logistic regression, random forest, XGBoost, SVM, AdaBoost, K-NN, and decision tree) for predicting liver disease. Random forest outperformed others, achieving 83.70% accuracy and superior precision, recall, F1 score, and AUC. It's deemed the best for early liver disease prediction. [18] A proposed method for malaria diagnosis involves training the VGG19 algorithm on a large dataset of images to detect malaria parasites, showing promising results in identifying infected cells and offering the potential for more effective diagnosis, particularly in resource-limited settings.

An emphasis on respiratory infections such as pneumonia, often triggered by viruses or bacteria, stands out as a significant infectious cause of death in children worldwide. [19] Transmission can happen through direct contact involving bacteria, viruses, or fungi. The study proposes Convolutional Neural Network (CNN) models utilizing logistic regression (LR) and support vector machines (SVM) in deep learning techniques. These models utilize architectures inspired by VGG16 and VGG19. They integrate results from fully connected layers using early or late fusion methods to classify pneumonia. [20] A fresh method utilizing Support Vector Machines (SVM) and Grid Search was devised for predicting breast cancer, leveraging the Wisconsin Diagnostic Breast Cancer (WDBC) database. Early detection is crucial for effective treatment, and machine learning techniques offer promising tools.

This system [21] engages in conversations with users, retrieves pertinent laws, and ranks them according to their similarity to queries. Achieving over 80% accuracy enhances legal research efficiency and accuracy. This innovation transforms the legal industry by providing rapid, precise assistance. Future research may explore additional features like case law analysis and contract review for further enhancement. [22] The study investigates how manual identification of gene mutations in cancer-like diseases leads to errors and proposes AI-assisted methods using machine learning techniques like Naïve Bayes, K-Nearest Neighbours, and Logistic Regression for more accurate classification. It addresses complexities in the cancer dataset, including text categories, and conducts detailed data analysis to improve understanding.

III. DATA REPRESENTATION

The data set is being generated in real-time using a webcam. The meticulous construction of this dataset involves the careful analysis of signs like those in Fig 1, Fig. 2., and Fig. 3. facilitated by the sophisticated functionalities of OpenCV.


Fig. 1. Hello


Fig. 2. Thanks


Fig. 3. I Love you

The dataset contains 25 distinct signs, with each sign having 30 sequences. The dataset consists of 22,500 images, with each sequence containing 30 frames. The dataset for assessing a sign language recognition model splits into 95% training images (21,375) and 5% testing images (1,125).

The dataset itself is thoughtfully categorized, with each classification delineated by distinct features and intricate point connections. The image capture process is comprehensive, encompassing not only the intricate details of the hand but also the contextual nuances of the surrounding environment with different candidates like an example shown in Fig. 4. It is imperative to underscore the intentional diversity embedded within this dataset, featuring a heterogeneous cohort of candidates spanning both genders. The deliberate inclusion of various individuals ensures a representative dataset, capturing the nuanced spectrum of hand gestures across diverse demographics. To accurately predict a gesture, it is necessary to understand that the same gesture can have multiple postures connected with it. To achieve this, the model need to be trained with more datasets that include the same gesture but with different postures. It will assist the proposed model in accurately identifying the necessary features to recognize the intended gesture. The LSTM model is being used to train the model for this purpose.


Fig. 4. Samples from different participants

The dataset includes samples from different participants, encompassing a broad spectrum of perspectives and orientations about the webcam. This nuanced approach significantly enriches the model's training, exposing it to a comprehensive array of hand gestures and enhancing its discriminative capacity across an extensive spectrum of scenarios.

IV. PROPOSED METHODOLOGY

Data collection and processing mark the initiation of the implemented architecture, as depicted in Fig. 5. The Mediapipe Holistic model was used to capture and process real-time video frames, enabling the identification of face, pose, and hand landmarks. The extracted keypoint data undergoes preprocessing, splitting into training and testing sets, and label conversion. The LSTM neural network, consisting of multiple layers with specified units and activation functions, is then trained on the prepared dataset. Subsequently, real-time testing involves continuously capturing video feeds, detecting landmarks, and predicting actions with the trained LSTM model. The enhanced video feed, displaying recognized actions and probabilities, is presented in real-time, creating an integrated system for action recognition.

The CNN architecture uses Convolutional Neural Networks to classify sign language images after training on a dataset. Within this framework, the CNN autonomously learns to extract pertinent features from the input images and subsequently categorize them into distinct sign language gestures. The Adam optimizer is employed. By dynamically adjusting the learning rate for each parameter based on historical gradients, the Adam optimizer effectively expedites convergence and enhances the overall training process.

Augmenting the training dataset via techniques like rotation, translation, and scaling increases the data's variety, which improves the model's resilience. Furthermore, meticulous fine-tuning of the hyperparameters inherent in the neural network architecture, coupled with optimization of the training regimen, can significantly elevate performance thresholds.

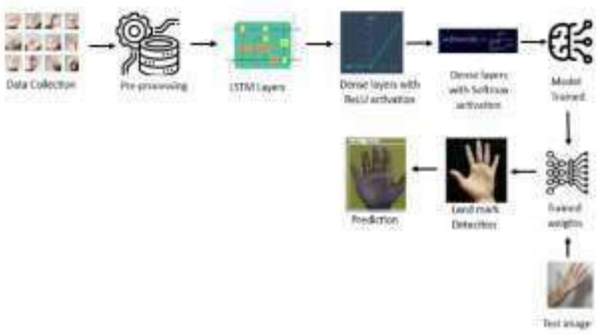


Fig. 5. Overall Structure for Hand Gesture Recognition

The recognition process is an integral component of the pipeline, involving the preprocessing of input images before their submission to the trained CNN model for classification. Subsequently, the model's predictions are mapped to corresponding sign language words or phrases, facilitating seamless communication.

Efficiency enhancements are attainable through various means, including optimization of the neural network architecture to reduce parameter count and computational complexity. Techniques such as pruning, quantization, and parallelization hold promise for accelerating inference times, thereby streamlining operational efficiency.

In the pursuit of capturing temporal dependencies within the input data, Long Short-Term Memory (LSTM) networks are for their efficacy in sequence modeling. LSTM layers can be integrated into neural networks to handle sequential data such as sign language gestures, making them especially relevant in sign language detection.

High accuracy is contingent upon several factors, including meticulous fine-tuning of neural network parameters, optimization of the training process, and augmentation of the training dataset. Furthermore, leveraging advanced methodologies such as transfer learning and ensemble techniques holds promise for further elevating accuracy metrics.

The six-layer LSTM model consists of three LSTM layers, followed by three dense layers in Fig. 6. The first LSTM layer with 64 units captures patterns and dependencies in input sequences, processing 1662 features over 30-time steps. The subsequent LSTM layer with 128 units refines sequence information, adding complexity. The third LSTM layer with 64 units acts as a bottleneck, reducing dimensionality before passing the output to the Dense layers. The first Dense layer with 64 units further reduces dimensionality, followed by a Dense layer with 32 units. The final Dense layer, determined by the number of distinct gestures or actions, employs the softmax activation function to convert raw outputs into class probabilities, enabling gesture recognition. Each output neuron represents the likelihood of a specific class, with the highest probability determining the model's prediction.

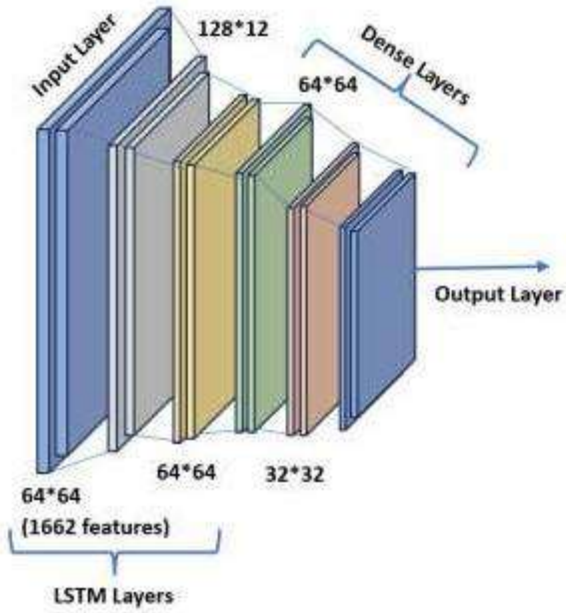


Fig. 6. Layered Architecture for Hand Gesture Recognition

A. Algorithm

LSTM Layers:
For each LSTM layer:
Initialize LSTM layers with varying units and activation functions
LSTM Layer 1 (64 units, ReLU activation):
For each time step t:
Input Gate (i_{t1}):

$$i_{t1} = \sigma (W_{ii\_1}. x_{t1} + b_{ii\_1}+W_{hi\_1}. h_{t1-1}+b_{hi\_1}) \quad (1)$$

Forget Gate (f_{t1}):

$$f_{t1} = \sigma (W_{if\_1}. x_{t1} + b_{if\_1} + W_{hf\_1}. h_{t1-1} + b_{hf\_1}) \quad (2)$$

Cell Gate (g_{t1}):

$$g_{t1} = tanh (W_{ig\_1}. x_{t1} + b_{ig\_1} + W_{hg\_1}. h_{t1-1} + b_{hg\_1}) \quad (3)$$

Output Gate (o_{t1}):

$$o_{t1} = \sigma (W_{io\_1}. x_{t1} + b_{io\_1} + W_{ho\_1}. h_{t1-1} + b_{ho\_1}) \quad (4)$$

Cell State Update (c_{t1}):

$$c_{t1} = ft_1. c_{t1-1} + i_{t1}. g_{t1} \quad (5)$$

Hidden State Update (h_{t1}):

$$h_{t1} = o_{t1}. tanh(c_{t1}) \quad (6)$$

LSTM Layer 2 (228 units, ReLU activation):
For each time step t:
Input Gate (i_{t2}):

$$i_{t2} = \sigma (W_{ii\_2}. x_{t2} + b_{ii\_2}+W_{hi\_2}. h_{t2-1}+b_{hi\_2}) \quad (7)$$

Forget Gate (f_{t2}):

$$f_{t2} = \sigma (W_{if\_2}. x_{t2} + b_{if\_2} + W_{hf\_2}. h_{t2-1} + b_{hf\_2}) \quad (8)$$

Cell Gate (g_{t2}):

$$g_{t2} = tanh (W_{ig\_2}. x_{t2} + b_{ig\_2} + W_{hg\_2}. h_{t2-1} + b_{hg\_2}) \quad (9)$$

Output Gate (o_{t2}):

$$o_{t2} = \sigma (W_{io\_2}. x_{t2} + b_{io\_2} + W_{ho\_2}. h_{t2-1} + b_{ho\_2}) \quad (10)$$

Cell State Update (c_{t2}):

$c_{t2} = ft_2 . c_{t2\text{-}1} + i_{t2} . g_{t2}$ (11)

Hidden State Update (h_{t2}):

$h_{t2} = o_{t2} . tanh(c_{t2})$ (12)

LSTM Layer 3 (64 units, ReLU activation):
For each time step t:

Input Gate (i_{t3}):

$i_{t3} = \sigma\ (W_{ii\_3} . x_{t3} + b_{ii\_3} + W_{hi\_3} . h_{t3\text{-}1} + b_{hi\_3})$ (13)

Forget Gate (f_{t3}):

$f_{t3} = \sigma\ (W_{if\_3} . x_{t3} + b_{if\_3} + W_{hf\_3} . h_{t3\text{-}1} + b_{hf\_3})$ (14)

Cell Gate (g_{t3}):

$g_{t3} = tanh\ (W_{ig\_3} . x_{t3} + b_{ig\_3} + W_{hg\_3} . h_{t3\text{-}1} + b_{hg\_3})$ (15)

Output Gate (o_{t3}):

$o_{t3} = \sigma\ (W_{io\_3} . x_{t3} + b_{io\_3} + W_{ho\_3} . h_{t3\text{-}1} + b_{ho\_3})$ (16)

Cell State Update (c_{t3}):

$c_{t3} = ft_3 . c_{t3\text{-}3} + i_{t3} . g_{t3}$ (17)

Hidden State Update (h_{t3}):

$h_{t3} = o_{t3} . tanh(c_{t3})$ (18)

Dense Layers:
Define ReLU activation function:

ReLU(a) = max (0, a) (19)

Define Softmax activation function:

Softmax(x) = exp(xi) / Σ(exp(xj)) (20)

Dense Layer 1 (64 units, ReLU activation)
For each neuron in the layer:

neuron_output_dense1 = ReLU (weights_dense1 * input_features + biases_dense1) (21)

Dense Layer 2 (32 units, ReLU activation)
For each neuron in the layer:

neuron_output_dense2 = ReLU (weights_dense2 * input_features + biases_dense2) (22)

Output Dense Layer (Softmax activation)
For each class in the output layer:

class_output = Softmax (weights_output * input_features_dense2 + biases_output) (23)

Adam Optimizer:
The Adam optimizer updates the model parameters θ_t based on past gradients m_t and squared gradients v_t:

Exponential Moving Average for Momentum (m_t):

$m_t = \beta_1 . m_{t-1} + (1 - \beta_1) . g_t$ (24)

Exponential Moving Average for Squared Gradients (v_t):

$v_t = \beta_2 . v_{t-1} + (1 - \beta_2) . g_t^2$ (25)

Bias-Corrected Momentum ($\hat{m}_t$):

$$\hat{m} = \frac{mt}{(1 - \beta 1t)} \quad (26)$$

Bias-Corrected Squared Gradients ($\hat{v}_t$):

$$\hat{v} = \frac{vt}{(1 - \beta 2t)} \quad (27)$$

Parameter Update (θ_t):

$$\theta_t = \theta_{t-1} - \frac{\eta}{\hat{v}t + \epsilon} . \hat{m} \quad (28)$$

Performance calculation:

ACCURACY = (TrPo + TrNe) / (TrPo + FaPo + FaNe + TrNe) (29)

PRECISION = TrPo / (TrPo + FaPo) (30)

RECALL = TrPo / (TrPo + FaNe) (31)

*F1-SCORE = (2 * PRECISION * RECALL) / (PRECISION + RECALL)* (32)

SAVE weights
PRINT confusion matrix and performance metrics

V. RESULTS AND DISCUSSION

The outcomes of this ground-breaking method of bridging the communication gap for people with physical problems, particularly those who are deaf and mute, are encouraging. The unique dataset containing different sign language gestures has proven reliable and extensive. Long-short memory (LSTM), one of the recurrent neural networks (RNN) algorithms, showed astounding accuracy in recognizing and understanding these motions recorded by webcams. The sign language motions were effectively understood by the LSTM through thorough analysis and model training, enabling efficient phrase construction.

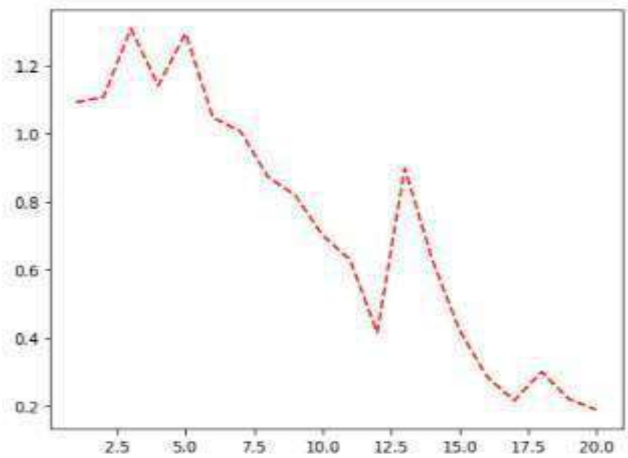


Fig. 7. Graph for Error rate

The above Fig. 7. represents training loss or training error. During training, errors occur due to differences between predicted and target output. The outcomes of the suggested methodology for sign recognition using hands and the performance of the LSTM algorithm are present.

This methodology seeks to address the physical challenges experienced by individuals who use sign language, with a particular emphasis on multilabel classification. This approach aims at enhancing the accuracy and reliability of communication between sign language users and those who don't use sign language. The technique evaluates the model's effectiveness by using three particular signs as examples and generating several confusion matrices for each sign. This study allows for fine-tuning of the model's performance by identifying areas that require focused changes to increase accuracy and dependability.

The LSTM algorithm and confusion matrix provide a nuanced understanding of the model's capabilities, which can ensure effective communication and inclusivity for individuals who use sign language in their daily interactions. In addition, the results of the proposed methodology encompass parameters such as input sign language, dataset size, training set, testing set, standard dataset/creators of the dataset, and classification methods. In the model, each of these parameters is crucially involved in the development, training, and validation processes, ultimately contributing to the overall improvement of its performance.

This model reached an accuracy of 90% in recognizing sign gestures, which outperforms several other models. It accurately identifies the gestures and correctly classifies the sign to that label. Fig. 8. displays the confusion matrix.

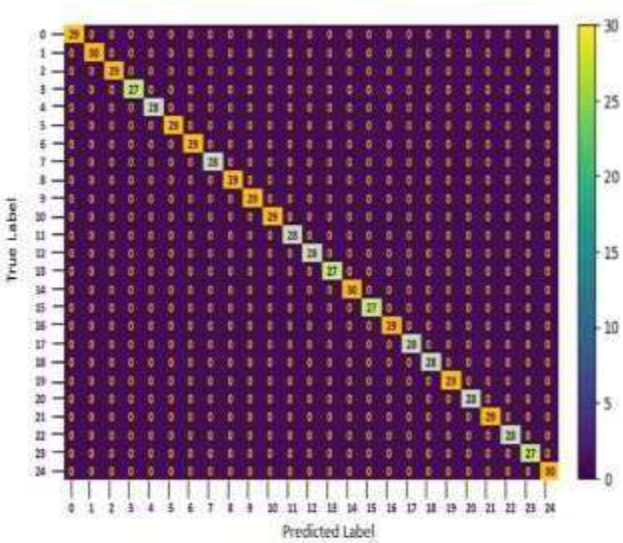


Fig. 8. Confusion matrix

This result demonstrates that the hand motion, with its distinct key-point structure, is distinguishable from other movements. The key points are strategically placed on both the hand and face to ensure accurate recognition, even if the candidate's posture changes. However, to achieve this level of accuracy, the system need to be trained with more data on the same hand gesture. The key points cover the fine details of the hand and are connected to form a path that covers both the hand and face separately. Another path connects the essential locations on the face and hand to compute the distance. Once taught, the system can reliably detect hand gestures, even when posture varies.

Fig. 9. is a collection of words that includes terminology for describing feelings. "Hello" is a cordial greeting, "bye" marks the conclusion of a conversation, and "I love you" expresses profound affection. The section lays the foundation for human connection through emotional and social communication.

Fig. 9. Salutations and moods

The category includes terms about food and phrases as shown in Fig. 10. "Eat" refers to the act of ingesting food, "good" might indicate the quality or flavor of food, and "hungry" refers to an initial human experience. The emphasis is on sustenance, flavor, and the physiological need for feeding, reflecting aspects of everyday life and well-being.

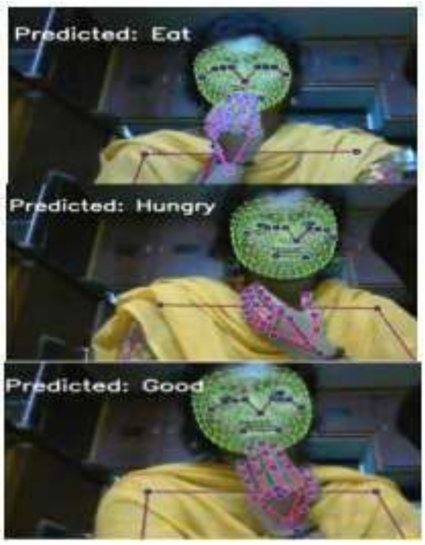


Fig. 10. Expressions

Fig. 11. includes terms like "morning," "afternoon," "evening," and "night." These terms describe several sections of the day, assisting people in planning their activities depending on time constraints. This category helps to arrange daily activities and promotes good communication around time.

Fig. 11. Time of Day

Fig. 12. refers to features of human identification and information. "Phone number" and "name" are distinct identities, whereas "time" might denote scheduling or temporal information. This category encompasses interpersonal communication, emphasizing the exchange of detailed information for various purposes.

Fig. 12. Personal Information

Phrases used in inquiry and understanding are "what," "it means," and "your." "What" provokes queries, "it means" requests explanation or clarification, and "you're" may appear in questioning targeted at understanding someone else's perspective or objects. Fig 13 depicts the language tactics used to obtain knowledge and increase comprehension in communication.

Fig. 13. Inquiry and Understanding

While the video is recording, it captures the user's gestures. While an individual alters their gesture, the algorithm detects the following word and constructs a phrase. Individuals with hearing or speech impairments may find this platform beneficial for connecting with others. Both deaf and hearing individuals can use sign language to communicate. As illustrated in Fig. 14., Fig. 15., it is possible to form complete sentences or express continuous thoughts using sign language.

Fig. 14. What is your name

Fig. 15. Where are you from

The proposed model has a 90% accuracy rate, an 84% F1-score, precision of 82.3%, and an 85.7% test recall. Overall, this methodology ensures a detailed grasp of the model's capabilities by utilizing the LSTM algorithm and producing different confusion matrices based on the performance of signs.

VI. CONCLUSION

The LSTM-based architecture's outstanding performance in detecting sign language gestures is a significant leap toward resolving communication issues for individuals with hearing and speech impairments. One of the standout achievements of the proposed model is its impressive overall accuracy, reaching 90%. It is a testament to the robustness of our approach, showcasing the model's ability to correctly classify a diverse range of sign language gestures with exceptional accuracy. Precision classification accuracy is 82.3%. The reliability of the proposed model in minimizing false positives and ensuring that recognized gestures represent intended signs is confirmed by this. Equally important is the recall score of 85.7%, highlighting the model's effectiveness in correctly identifying gestures. This metric underscores the importance of the model's ability to capture all instances of sign language expressions, ensuring inclusive communication for individuals with hearing and speech impairments. The F1-Score, a balanced measure of precision and recall, further reinforces the model's overall performance with a score of 84%. The success of the proposed LSTM-based system for recognizing sign language gestures represents a significant advancement in the field.

ACKNOWLEDGMENT

We would like to extend our sincere gratitude to each and every one who has helped us out in completion of this research work.

REFERENCES

[1] Y. F. bte Abdul Gaus, F. W. H. Tze and K. T. T. Kin, "Feature extraction from 2D gesture trajectory in Malaysian Sign Language recognition," 2011 4th International Conference on Mechatronics (ICOM), Kuala Lumpur, Malaysia, 2011, pp. 1-6, doi: 10.1109/ICOM.2011.5937179.

[2] 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

[3] H. Alqahtani et al., "Cyber Intrusion Detection Using Machine Learning Classification Techniques," in Computing Science, Communication and Security, Singapore, 2020, pp. 121-131: Springer Singapore

[4] Zaki, Mahmoud & Shaheen, Samir. (2011). Sign language recognition using a combination of new vision-based features. Pattern Recognition Letters. 32. 572-577. 10.1016/j.patrec.2010.11.013.

[5] R. Cui, H. Liu, and C. Zhang, “A deep neural framework for continuous sign language recognition by iterative training,” IEEE Transactions on Multimedia, vol. 21, no. 7, pp. 1880–1891, 2019

[6] Zahedi, Morteza & Dreuw, Philippe & Rybach, David & Deselaers, Thomas & Ney, Hermann. (2006). Using Geometric Features to Improve Continuous Appearance-based Sign Language Recognition. 3. 1019-1028. 10.5244/C.20.104

[7] Vogler, C. & Metaxas, Dimitris. (1997). Adapting hidden Markov models for ASL recognition by using three-dimensional computer vision methods. IEEE International Conference On Systems, Man, and Cybernetics, Computational Cybernetics And Simulation. 1. 156 - 161 vol.1. 10.1109/ICSMC.1997.625741

[8] D. Kelly, R. Delannoy, J. Mc Donald, and C. Markham, ―A framework for continuous multimodal sign language recognition,‖ in Proc. Int. Conf. Multimodal Interfaces, Cambridge, MA, 2009, pp. 351–358

[9] S. Salian, I. Dokare, D. Serai, A. Suresh, P. Ganorkar, Sign Language recognition Using machine learning and image processing, in IEEE (2017)

[10] J. Huang, W. Zhou, H. Li, and W. Li, "Sign language recognition using real-sense," 2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), Chengdu, China, 2015, pp. 166-170, doi: 10.1109/ChinaSIP.2015.7230384.

[11] S. Belgacem, C. Chatelain, and T. Paquet, “Gesture sequence recognition with one shot learned crf/hmm hybrid model,” Image and Vision Computing, vol. 61, pp. 12–21, 2017

[12] K. Amrutha and P. Prabu, "ML Based Sign Language Recognition System," 2021 International Conference on Innovative Trends in Information Technology (ICITIIT), Kottayam, India, 2021, pp. 1-6, doi: 10.1109/ICITIIT51526.2021.9399594. A. Mohanarathinam, K. G. Dharani, R. Sangeetha, G. Aravindh and P. Sasikala, "Study on Hand Gesture Recognition by using Machine Learning," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2020, pp. 1498-1501, doi: 10.1109/ICECA49313.2020.9297513.

[13] H. Avula, R. R and A. S Pillai, "CNN-based Recognition of Emotion and Speech from Gestures and Facial Expressions," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 1360-1365, doi: 10.1109/ICECA55336.2022.10009316.

[14] W. WANG and H. YANG, "Towards Realizing Sign Language to Emotional Speech Conversion by Deep Learning", *2021 12th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, pp. 1-5, 2021.

[15] Cassell, Justine & Stone, Matthew & Douville, Brett & Prevost, Scott & Achorn, Brett & Steedman, Mark & Badler, Norman & Pelachaud, Catherine. (1994). Modeling The Interaction Between Speech And Gesture. Technical Reports (CIS).

[16] Kandula, Ashok Reddy, R. Sathya, and S. Narayana. "Multivariate Analysis on Personalized Cancer Data using a Hybrid Classification Model using Voting Classifier. & quot ;International Journal of Intelligent Systems and Applications in Engineering 11.1 (2023): 354-362.

[17] Kandula, Ashok Reddy, et al. "Machine Learning Algorithms for Predictive Analysis in theIdentification of Chronic Liver Disease." NeuroQuantology 20.15 (2022): 5951.

[18] A. R. Kandula, S. Kalyanapu, L. S. Gottipati, H. Kamma, M. Munagala and P. Ramachandran,"Revolutionizing Malaria Diagnosis: A Deep Learning Approach," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023, pp. 63-66, doi:10.1109/ICSCSS57650.2023.10169586.

[19] A. Reddy Kandula, S. D. Gangiredla, K. P. Tirukkovalluri, S. Yallamilli, N. Tummalapalli and S.Talluri, "X-Ray Image Analysis using Deep Learning Techniques to Identify Pneumonia," 2023 Second International Conference on Augmented Intelligence and Sustainable Systems (ICAISS),Trichy, India, 2023, pp. 674-679, doi: 10.1109/ICAISS58487.2023.10250543.

[20] S. Kalyanapu, A. R. Kandula, P. Madhuri, P. Ganta, H. N. Sai Sri Vattikuti and U. Darapu,"Enhancing Breast Cancer Prediction Through SVM-Based Analysis," 2023 Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS), Kanjirapally, India, 2023, pp. 1-6, doi:10.1109/AICERA/ICIS59538.2023.10420106.

[21] A. R. Kandula, M. Tadiparthi, P. Yakkala, S. Pasupuleti, P. Pagolu and S. M. Chandrika Potharlanka, "Design and Implementation of a Chatbot for Automated Legal Assistance using Natural Language Processing and Machine Learning," 2023 Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS), Kanjirapally, India, 2023, pp. 1-6, doi: 10.1109/AICERA/ICIS59538.2023.10420298.

[22] Kandula, Ashok Reddy, R. Sathya, And S. Narayana. &Quot;Comparative Analysis Of Machine Learning Techniques On Genetic Mutation-Based Cancer Diagnosis Data." Journal of Theoretical and Applied Information Technology 100.6 (2022).

Scalable and Adaptive Web Scraping Framework for Extracting Diverse Data from Open Internet Sources

N. Somanath Reddy
Computer Science and Engineering Department
Sathyabama Institute of Science and Technology
Chennai, Tamil Nadu
somanathreddy43@gmail.com

B. Ruthvik
Computer Science and Engineering Department
Sathyabama Institute of Science and Technology
Chennai, Tamil Nadu
ruthvikborra@gmail.com

DR. M. Maheshwari
Computer Science and Engineering Department
Sathyabama Institute of Science and Technology
Chennai, Tamil Nadu
maheshwari.cse@sathyabama.ac.in

Dr. J. L. Jany Shabu
Computer Science and Engineering Department
Sathyabama Institute of Science and Technology
Chennai, Tamil Nadu
janyshabu.cse@sathyabama.ac.in

Dr. J. Refonna
Computer Science and Engineering Department
Sathyabama Institute of Science and Technology
Chennai, Tamil Nadu
refonna.cse@sathyabama.ac.in

Abstract: This research presents a comprehensive system for web scraping from open internet sources, offering an efficient means of acquiring data from online. The goal of this research study is to develop an innovative framework capable of efficiently extracting a wide range of data from various open internet sources. The aim is to create a scalable and adaptive solution that can dynamically adjust to different websites' structures and data formats, enabling comprehensive data collection for diverse purposes such as research, analysis, and decision-making. By addressing the challenges associated with web scraping, such as scalability, adaptability, and data diversity, the paper seeks to advance the field and provide a valuable tool for extracting valuable insights from the vast landscape of online information. Through a combination of scraping methodologies including headless surfing, API integration, and HTML parsing, the proposed system enables systematic data retrieval, facilitating the extraction of valuable insights from diverse webpages. Improving efficiency and scalability through technologies like distributed computing and optimized network infrastructure is crucial for handling large-scale data extraction tasks.

Index Terms: Web Scraping, Open Internet Sources, Data Extraction, Html Parsing, Screen Scraping, APIs, Market Research, Competitive Analysis, Data Analysis, E-Commerce, Financial Industry, Academic Research.

I. INTRODUCTION

Data acquisition from online sources has undergone a significant transformation with the introduction of web scraping methodologies. This automated process has become indispensable for businesses and researchers, empowering them to efficiently gather a wealth of information from internet. Its pivotal role is echoed throughout numerous studies, where its application is showcased as a cornerstone for data analysis and aggregation [1]. One of the primary techniques in web scraping is HTML parsing, a method that involves extracting structured data from the HTML markup of websites. This technique enables developers to pinpoint specific elements within a webpage, ranging from simple components like tables and headings to more intricate structures such as forms and lists [1]. This capability not only streamlines data extraction but also facilitates targeted information retrieval from diverse sources.

In parallel, the utilization of Application Programming Interfaces (APIs) for web scraping has gained prominence. APIs offered by websites provide a structured means for accessing and retrieving data, ensuring consistent and organized extraction processes [2]. This approach simplifies data acquisition by enabling users to interact with designated interfaces, thereby circumventing the complexities associated with direct HTML parsing.

Despite the availability of APIs, challenges persist when websites either impede data retrieval or lack API support. In such scenarios, automated browser tools like Selenium emerge as indispensable allies. Selenium, a powerful web automation framework, mimics user interactions with web pages, enabling developers to navigate dynamically loaded content and extract data effectively [3]. This adaptive approach proves invaluable in scenarios where traditional scraping techniques fall short, ensuring robust data acquisition regardless of the intricacies posed by the website's structure or functionality.

In this context, the need for a scalable and adaptive web scraping framework becomes apparent. By integrating various scraping techniques and leveraging adaptable strategies, such a framework can address the diverse challenges encountered in data extraction from open internet sources. This study aims to outline the design and implementation of such a framework, explaining its significance through illustrative examples and practical demonstrations.

II. RELATED WORKS

The literature review on web scraping techniques, [1] covers approaches such as HTML parsing, web scraping libraries, and API utilization. Researchers are now more focused on analyzing the ethical and practical challenges associated with web scraping, providing valuable insights for researchers and practitioners in the field. Presenting an innovative solution for web scraping, [2] focuses on its effectiveness in collecting essential information from online sources. The authors showcased the tool's capability in utilizing various web scraping techniques to extract data from open internet platforms, highlighting its potential for data extraction and analysis. Taking an industrial perspective on web scraping, [3] emphasizes key traits and unresolved challenges within the domain. The study explores management strategies, interpretation issues, and complexities related to data extraction, providing valuable insights into the industrial implications of web scraping methods. Discussing about the different web scraping methods for information retrieval from websites, [4] emphasizes the importance of web scraping tools in streamlining the data extraction process. They also address difficulties and moral considerations associated with web scraping, contributing to the ethical discourse surrounding this practice. Focusing on the extraction of data from public internet sources, particularly in the context of e-commerce websites, [5] presents various web scraping approaches employed to collect relevant data, aiming to facilitate product comparison. Published in IEEE and presented at a prestigious conference, this paper provides significant insights into web scraping techniques for data acquisition.

III. EXISTING SYSTEM

There are certain drawbacks to the current system for web scraping from public internet sources using different web scraping approaches. Although it is legally permissible to scrape data from websites that are open to the public, many of them have terms of service or usage agreements in place that forbid doing so. If these agreements are broken, there may be legal repercussions. Also, web scraping may be interpreted as an infringement on someone else's privacy or an unfair use of their data, it may give rise to ethical questions. The inconsistent and untrustworthy data collected via web scraping is another drawback of the current approach. The structure and style of websites are constantly changing, which makes it challenging for web scrapers to reliably and precisely collect data. Because of this, the data that has been scraped can include mistakes or missing information, which could seriously affect how reliable and helpful it is for analysis. Moreover, online scraping may need a lot of time and resources. The scraping process might take a long time and demand a lot of processing power, depending on the complexity of the website being scraped and the volume of data needed. For individuals or organizations with limited resources or time restrictions, this could be a barrier. The possibility of IP banning or being identified as a bot is another drawback of web scraping. Numerous websites use techniques like IP blocking and CAPTCHA challenges to identify and prevent web scrapers. These actions may impede data retrieval, slow down, or even stop the scraping operation. Furthermore, a website has the right to prohibit an IP address or take legal action against the scraper if it finds excessive scraping activity. Figures 1 and 2 show the traditional web extraction method.

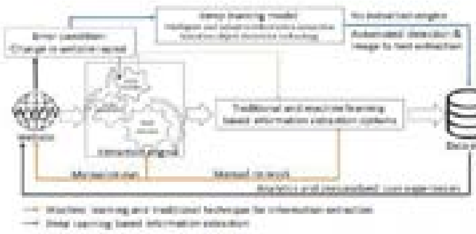


Fig. 1 Traditional, machine learning, and deep learning based web data extraction system.

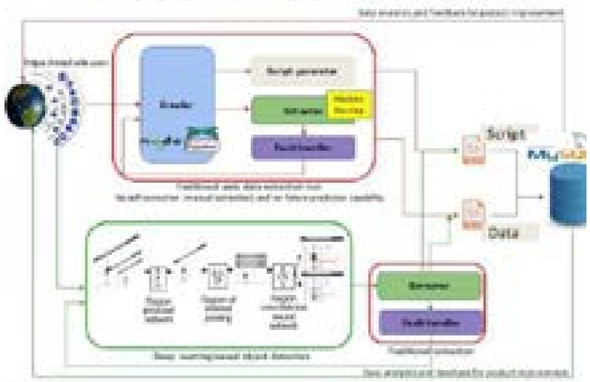


Fig. 2 Traditional, machine learning, and deep learning based web data extraction system with own data extraction eng

IV. PROPOSED SYSTEM

The process of automatically obtaining data from websites is known as web scraping. It entails obtaining a webpage's HTML code and parsing it to extract the needed data. After that, this information can be processed, saved, or subjected to additional analysis. There are numerous methods available for web scraping from publicly accessible online sources. The simplest method is to use HTTP requests to directly retrieve a webpage's HTML content. The HTML can be retrieved and saved as a text document using Python requests or libraries like urllib. This method works well for easy web pages with a simple structure for scraping purposes.

Using a library such as BeautifulSoup to parse the HTML text is another method. It offers a practical method for navigating and searching the HTML tree structure, as well as for extracting particular parts using CSS selectors, tags, or properties. When the needed data is incorporated within particular tags or characteristics, or when the website has a complex structure, this technique is helpful for scraping it. The headless browsing strategy can be used with websites whose content rendering is significantly dependent on JavaScript. JavaScript-based apps can be interacted with and the behavior of a standard web browser imitated by headless browsers such as Puppeteer or Selenium. This makes it possible to scrape dynamic web pages that load material dynamically using AJAX calls. Apart from these methods, web scraping can also be accomplished with specific tools and frameworks.

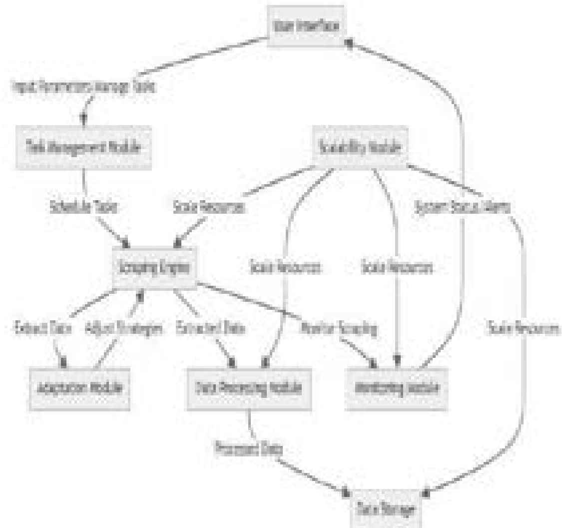


Fig. 3. System Architecture

Modules Description

Data Collection and Source Discovery Module:

The Data Collection and Source Discovery Module played a pivotal role in initiating the data extraction process, addressing several key questions along the way. It began by identifying potential data sources based on user-defined parameters, such as keywords, websites, or categories. This step directly answered the question of "How to achieve the data extraction process?" by setting the groundwork for sourcing relevant data from the open internet. To ensure scalability and performance improvement, the module leveraged advanced web scraping techniques, including multi-threading and distributed computing. By doing so, it effectively addressed the question of "How to improve performance?" These techniques enabled the framework to extract data from multiple sources simultaneously, enhancing efficiency and reducing extraction time.

The module incorporated dynamic content handling mechanisms capture data from modern web applications with dynamically loaded content. This directly addressed the question of "How to achieve the extraction process?" by ensuring comprehensive data acquisition from websites that relied on dynamic content loading. Additionally, robust error handling and recovery strategies were integrated to manage interruptions and ensure the seamless continuation of data extraction processes. This addressed concerns about data integrity and reliability, effectively answering questions related to "How to protect the stored data?" and "How to ensure data extraction process reliability?"

Overall, the data collection and source discovery module not only initiated the data extraction process but also addressed key questions related to performance improvement, extraction process achievement, data protection, and reliability enhancement.

Data Transformation and Normalization Module

This module focuses on cleaning and preprocessing the raw data. This involves the removal of inconsistencies, handling missing values, and standardizing data formats. By addressing these issues, the module ensures that the data is refined and ready for subsequent analysis. This step directly addresses the question of "How to transform the data?" as it lays the groundwork for refining the raw data into a more usable format.

Following data cleaning and preprocessing, the module progresses to the normalization and standardization phase. Here, disparate data structures originating from various sources are normalized and standardized. This crucial step ensures uniformity and compatibility across the dataset, facilitating seamless integration and analysis. By harmonizing the data into a common format, the module directly answers questions related to "How to achieve the extraction process?" as it ensures that the data is prepared for further processing and analysis.

Moreover, the module may incorporate advanced techniques such as feature engineering to enhance the dataset. Feature engineering involves the derivation of new features or the enhancement of existing ones to enrich the dataset and improve its predictive power. By employing these advanced techniques, the module further enhances the quality and utility of the transformed data. This step not only addresses questions related to data quality but also enhances the effectiveness of subsequent analysis tasks. In addition to data transformation and standardization, the module may also focus on ensuring data security and protection. Robust security measures such as access control mechanisms and data encryption may be implemented to safeguard the stored data. These measures help address concerns related to data security and privacy, ensuring compliance with regulations and protecting sensitive information from unauthorized access.

To enhance the performance, the module may optimize the processing algorithms and workflows. By streamlining processes and utilizing efficient algorithms, the module improves processing speed and resource utilization. Additionally, parallel processing techniques and distributed computing may be employed to further enhance performance and scalability. These measures collectively contribute to improving the overall performance of the data transformation and normalization process, ensuring timely and efficient processing of the dataset.

Adaptive Data Storage and Access Module:

The Adaptive Data Storage and Access Module constitutes a critical component within the framework, primarily focusing on efficient data storage and retrieval mechanisms while addressing key considerations for security and scalability.

Initially, the module establishes a scalable and distributed data storage architecture tailored to accommodate varying volumes of data. This architecture ensures that the framework can handle large datasets effectively, addressing the question of "How to achieve the storage of data?" By leveraging cloud-based storage solutions or distributed databases capable of horizontal scaling, the module enables seamless expansion as data volumes grow, thereby enhancing scalability.

Moreover, robust security measures are integrated into the module to safeguard stored data. Access control mechanisms, encryption techniques, and auditing functionalities are implemented to protect sensitive information from unauthorized access and ensure compliance with data protection regulations. This comprehensive approach to security addresses concerns related to data privacy and confidentiality, effectively answering the question of "How to achieve security?". The module also incorporates backup and disaster recovery strategies to safeguard against data loss or corruption. Regular backups are performed to create redundant copies of the data, while disaster recovery plans are put in place to ensure data availability and continuity in case of unforeseen events such as hardware failures or natural disasters. These measures contribute to data protection and integrity, addressing concerns related to "How to protect stored data?"

Additionally, to enhance performance, the module may optimize data access algorithms and utilize caching mechanisms. By optimizing data retrieval processes and minimizing latency, the module improves overall system performance and responsiveness. Furthermore, parallel processing techniques may be employed to distribute processing tasks across multiple nodes, further enhancing performance and scalability. These measures collectively contribute to improving the overall performance of the data storage and access module, ensuring efficient and reliable data management within the framework.

V. RESULTS AND DISCUSSION

Users can automatically collect data from webpages with this robust system for web scraping from open internet sources utilizing different web scraping strategies. Table I describes the comparison between the proposed method and other existing approaches based on different aspects. Table II showcases the performance metrics of the proposed web scraping method. The accuracy and loss graph are depicted in figures 4 and 5 based on the confusion matrix given in figure 6. Figure 7 shows the final ROC Curve of the proposed approach.

TABLE I. COMPARATIVE ANALYSIS

Aspect	Proposed Method	Existing Approaches
Scalability	Utilizes multi-threading and distributed computing for efficient extraction of data from multiple sources simultaneously, enhancing scalability.	May struggle with scalability when dealing with large datasets or high-frequency scraping tasks.
Adaptability	Leverages a combination of techniques such as HTML parsing automated browser tools, and advanced data transformation algorithms, ensuring robust data extraction from diverse website structures.	Some approaches may rely on specific scraping techniques, limiting adaptability to varying website structures.
Data Transformation	Includes a dedicated module for data transformation and normalization, refining raw data into a consistent and usable format Advanced techniques such as feature engineering enrich the dataset.	Limited emphasis on data transformation and normalization in many existing methods, focusing primarily on data extraction.
Security	Integrates robust security measures such as access control mechanisms encryption techniques, and auditing functionalities to protect stored data effectively.	Security measures may vary across existing approaches, with some lacking comprehensive security features.
Performance	Implements optimization techniques such as indexing for efficient data retrieval, caching mechanisms, and parallel processing, resulting in improved system performance.	Performance improvements may vary across existing approaches, with some methods experiencing slower data retrieval times.
Weaknesses	Dependency on website structure, ethical considerations, maintenance overhead, and potential performance limitations.	Shared weaknesses such as dependency on website structure, ethical considerations, maintenance overhead, and performance limitations.

TABLE II. PERFORMANCE METRICS

Accuracy	**Precision**	**Recall**	**F1 score**
95.8	**96.4**	**97.3**	**98.7**

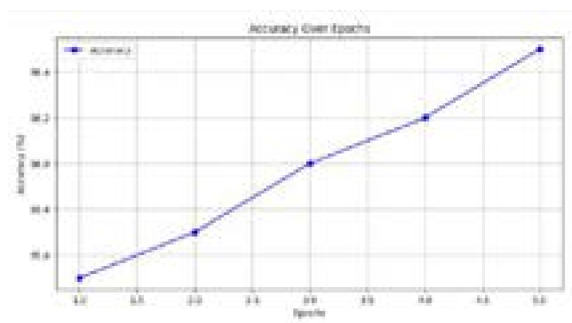


Fig.4.Accuracy Graph

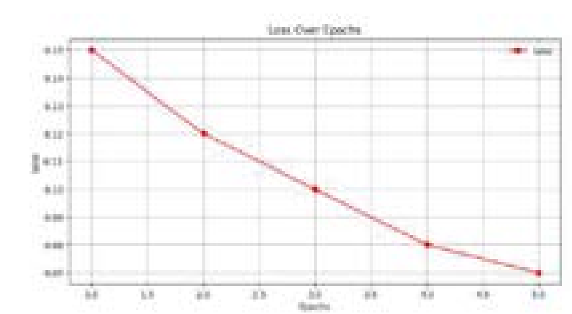


Fig.5.Loss Graph

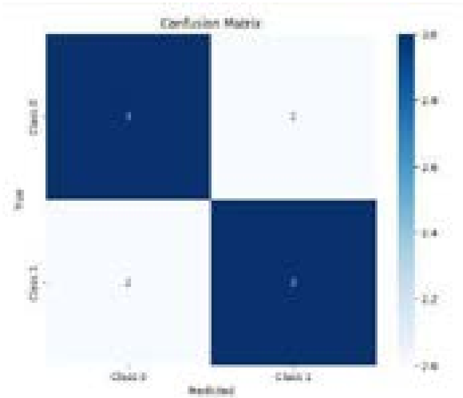


Fig.6.Confusion Matrix

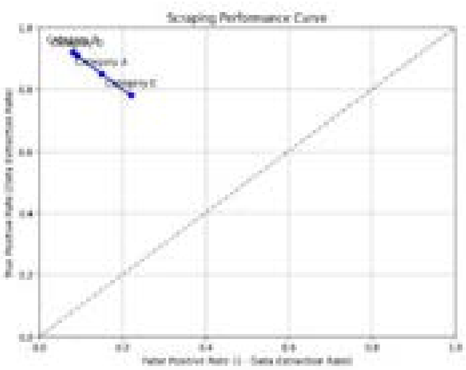


Fig.7.ROC Curve

VI. CONCLUSION

In conclusion, the proposed system for web scraping from open internet sources stands as a pivotal instrument for efficiently obtaining data from the vast expanse of online resources. Through a comprehensive array of scraping methodologies such as headless surfing, API integration, and HTML parsing, this system offers a versatile and systematic approach to data retrieval, allowing users to extract valuable insights from a wide spectrum of webpages. While the benefits of this system are undeniable, it is paramount to underscore the importance of adhering to legal and ethical standards. As the use of web scraping becomes increasingly prevalent, ensuring compliance with website terms of service, data protection laws, and ethical guidelines remains imperative. By respecting the rights of website owners and safeguarding user privacy, the integrity and credibility of the scraping process are preserved.

Looking towards the future, improving efficiency and scalability will be pivotal in handling large-scale data extraction tasks, necessitating the implementation of advanced technologies such as distributed computing and optimized network infrastructure.

REFERENCES

[1] Bale, A. S., Ghorpade, N., Rohith, S., Kamalesh, S., Rohith, R., & Rohan, B. S. (2022, August). Web Scraping Approaches and their Performance on Modern Websites. In 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 956-959). IEEE.

[2] Niu, Q., Kandhro, I. A., Kumar, A., Shah, S., Hasan, M., Ahmed, H. M., & Liang, F. (2022). Web Scraping Tool For Newspapers And Images Data Using Jsonify. Journal of Applied Science and Engineering, 26(4), 465-474.

[3] Chiapponi, E., Dacier, M., Thonnard, O., Fangar, M., Mattsson, M., & Rigal, V. (2022, June). An industrial perspective on web scraping characteristics and open issues. In 2022 52nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks-Supplemental Volume (DSN-S) (pp. 5-8). IEEE.

[4] NR, R. R., & Vijayalakshmi, M. (2023, February). Web Scraping Tools and Techniques: A Brief Survey. In 2023 4th International Conference on Innovative Trends in Information Technology (ICITIIT) (pp. 1-4). IEEE.

[5] Khatter, H., Sharma, A., & Kushwaha, A. K. (2022, July). Web Scraping based Product Comparison Model for E-Commerce Websites. In 2022 IEEE International Conference on Data Science and Information System (ICDSIS) (pp. 1-6). IEEE.

[6] Oralbekova, D., Mamyrbayev, O., Othman, M., Kassymova, D., & Mukhsina, K. (2023). Contemporary approaches in evolving language models. *Applied Sciences*, *13*(23), 12901.

[7] Salvatore, C., Biffignandi, S., & Bianchi, A. (2024). Augmenting business statistics information by combining traditional data with textual data: a composite indicator approach. *METRON*, 1-21.

[8] Howell, A., Saber, T., & Bendechache, M. (2023). Measuring node decentralisation in blockchain peer to peer networks. *Blockchain: Research and Applications*, *4*(1), 100109.

[9] Ahmed, K., Khurshid, S. K., & Hina, S. (2024). CyberEntRel: Joint extraction of cyber entities and relations using deep learning. *Computers & Security*, *136*, 103579.

[10] Tare, S. S., Bhute, M. M., & Arage, P. (2023, May). Recevent: NLP based Event Recommender System. In *2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC)* (pp. 608-614). IEEE.

[11] Kaur, R., Gabrijelčič, D., & Klobučar, T. (2023). Artificial intelligence for cybersecurity: Literature review and future research directions. *Information Fusion*, 101804.

[12] Böhmecke-Schwafert, M., & Dörries, C. (2023). Measuring Innovation in Mauritius' ICT Sector Using Unsupervised Machine Learning: A Web Mining and Topic Modeling Approach. *Journal of the Knowledge Economy*, 1-34.

[13] Wibawa, J. T., Gunawan, R., Suhandi, M., & Setiawan, R. (2023). Development of Marketplace Product Data Mining and Analysis Application as a Reference in Running a Business. *IT for Society*, *8*(1).

[14] Prabowo, O. M., Mulyana, E., Nugraha, I. G. B. B., & Supangkat, S. H. (2023). Cognitive City Platform as Digital Public Infrastructure for Developing a Smart, Sustainable and Resilient City in Indonesia. *IEEE Access*, *11*, 120157-120178.

[15] Mejia-Escobar, C., Cazorla, M., & Martinez-Martin, E. (2023). A Large Visual, Qualitative, and Quantitative Dataset for Web Intelligence Applications. *Computational Intelligence and Neuroscience*, *2023*.

IoT based Water Flow Monitoring System using Wireless Network (LoRaWAN)

Mrs. B. Chandra
Dept of Information Technology
Easwari Engineering College
Chennai, India
Chandra.klnce@gmail.com

Mrs. K. Kausalya
Dept of Information Technology
Easwari Engineering College
Chennai, India
ausalya.K@eec.srmrmp.edu.in

Poornima. B
Dept of Information Technology
Easwari Engineering College
Chennai, India
r.poornibala@gmail.com

Yamini Babu
Dept of Information Technology
Easwari Engineering College
Chennai, India
yamibabu101@gmail.com

***Abstract*-**

This research work proposes a robust and efficient real-time water leak detection system by using an integrated approach. By harnessing the capabilities of water flow sensors, microcontrollers, and LoRa transmitters, it provides comprehensive monitoring and instant alerts. The sensors continuously track water flow, transmitting data to the microcontroller for analysis. Any anomaly triggers the microcontroller to relay critical information to the LoRa receiver, initiating an urgent response. This prompts the I2C LCD display to visually notify users of a leak detection in real-time, while a buzzer makes an audible alarm. The synchronized visual and auditory alerts enable users to promptly intervene, mitigating potential water damage. The proposed system architecture ensures seamless coordination between its components for rapid leak identification. The two water flow sensors swiftly detect irregularities, communicating this data to the microcontroller. It then relays the information via the LoRa transmitter to the receiver. On obtaining the transmission, the receiver conveys the signal back to the microcontroller to activate the output interfaces. The I2C LCD display and buzzer then provide instant visual and audio cues to alert users. This end-to-end

integration facilitates real-time communication, analysis, and response, empowering timely action against water leaks. Hence, this study proposes an effective solution through its robust architecture. By leveraging key technologies like sensors, microcontrollers, and transmitters, it enables comprehensive real-time monitoring and instant alerts against water leaks.

Keywords: Leak Detection, Microcontroller, Rapid Leak Identification, LoRa Transmitters.

I. INTRODUCTION

The system addresses the critical need for early identification of water leakages in various environments by leveraging two water flow sensors to detect anomalies indicative of leaks. Upon detection, the sensors relay data to a central microcontroller, which communicates with a LoRa transmitter for data transmission. On the receiving end, a LoRa receiver captures these signals, relaying the information back to the microcontroller, which interfaces with an I2C LCD display and a buzzer for immediate notification and alerting. The LCD visually displays the detected leakage, while the buzzer serves as an audible alert, ensuring timely intervention. This project amalgamates sensor technology, microcontrollers, and LoRa communication protocols to establish an efficient and responsive leak detection system. Its significance lies in swiftly detecting leaks, minimizing potential water damage, and facilitating proactive maintenance measures. The system's wireless communication capabilities through LoRa WAN enable its implementation in diverse settings, ensuring real-time monitoring and rapid response to safeguard against water-related hazards. To achieve high accuracy in leak detection and localization, the following strategies were employed. The utilization of advanced sensor technologies with high precision and reliability was prioritized to ensure accurate measurement of flow rate and pressure. Sophisticated data analysis techniques, such as machine learning algorithms, were employed to interpret sensor data accurately and identify potential leaks with precision. Regular calibration of the system was undertaken to ensure that measurements remain consistent and reliable over time. Thorough testing and validation were conducted to verify the system's accuracy under different operating conditions and scenarios, ensuring its reliability in real-world applications.

II. LITERATURE SURVEY

Authors of [1] have worked on detecting oil and gas leak detection, though the work has been successful but only for oil and gas, water has different properties so it does not work well with water. [1]

Authors of [2] have introduced an AI-based leak detection system integrated with cloud-based information management. The system is designed to methodically gather and organize acoustic signals associated with potential leakages. Various machine learning algorithms, including Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Support Vector Machine (SVM), were created and evaluated for their efficacy in identifying and categorizing leakage patterns. The comparative analysis sheds light on the performance and suitability of each algorithm in the context of leak detection and management. [2]

Authors of [3] The object of the ESTHISIS project is the development of a low-cost and low-energy wireless sensor system for the immediate detection of leaks in metallic piping systems for the transport of liquid and gaseous petroleum products in a noisy industrial environment. In this study, two distinct leakage detection methodologies are presented. First, a 2D-Convolutional Neural Network (CNN) model, the second methodology entails a Long Short-Term Memory Autoencoder (LSTM AE).[3]

Authors of [4] investigated and implemented the detection of gas in underground pipelines through the utilization of optical sensing technology, a domain that has primarily remained theoretical in the realm of pipeline leakage monitoring. The study explores the latest advancements in pipeline leak detection technology, specifically focusing on the integration of optical Fiber sensing technology. Moreover, it puts forward an algorithm designed for monitoring gas pipeline leaks, leveraging a distributed optical Fiber acoustic sensing (DAS) system. [4]

Authors of [5] proposed a novel system utilizing a wireless sensor network coupled with an autonomous learning algorithm, designed specifically to monitor water distribution systems and pinpoint leaks accurately. Their comprehensive examination covers the detailed architecture of the system, encompassing hardware specifications, communication protocols, and data analysis techniques. Through a comparative study involving machine learning algorithms such as random forest, decision trees, neural networks, and Support Vector Machine (SVM), they identify the most effective method for leak detection. their findings, derived from meticulous methodology, training, and validation processes, culminate in the validation of our system through real-case implementation, showcasing its ability to achieve a commendable 75% accuracy in leak detection. [5]

Author of [6] has developed an autonomous robot for inspecting gas pipelines, named GPLDA Bot. This system incorporates a robot structure utilizing DC motors for propulsion. Gas sensors such as MQ4 and MQ6 are integrated to detect leaks of various gases including methane, butane, LPG, iso-butane, propane, and LNG. Furthermore, an Internet of Things (IoT) module is linked with GPLDA Bot to facilitate directional movement and provide real-time data analysis on a cloud platform. By manoeuvring in multiple directions, the robot can identify defects in gas pipelines, thus mitigating potential industrial accidents.[6]

Authors of [7] found a system to detect and prevent leak in the thermal power plants, water distribution systems, overhead tanks etc. Here the parameters like humidity, temperature, pressure, sound detection is detected using sensors and Arduino microcontroller. The data which is sensed is acquired and transmitted via ZigBee and Gui developed in LabVIEW. [7]

Most existing research on pipeline leakage detection emphasizes the necessity of flowmeters, often the most economical option. Studies highlight two sensor types: ultrasonic, preferred for being non-invasive, and the invasive type, which requires pipeline intrusion, thereby raising installation costs but offering slightly greater consistency. In IoT-based leakage monitoring systems, the final critical component is the communication link between nodes and the cloud. Various wireless technologies such as Bluetooth, Wi-Fi, ZigBee, and Cellular networks have been explored, each with its advantages and limitations. For instance, Wi-Fi offers high data rates but limited range, making it less suitable for battery-powered nodes. Cellular networks provide long-range and high data rates but suffer from high energy consumption. LoRa WAN emerges as a promising option due to its long-range capabilities and minimal energy costs, making it suitable for remote installation sites lacking cellular or Wi-Fi coverage.

III. EXISTING SYSTEM

The current existing systems for leak detection primarily rely on wired sensor networks or standalone sensor systems, limiting their adaptability and scalability. These conventional systems often entail complex wiring setups, making installation cumbersome and challenging in remote or expansive environments. Additionally, the communication between sensors and central monitoring units tends to be constrained by physical wiring, restricting flexibility, and increasing maintenance requirements. Moreover, the lack of real-time wireless data transmission capabilities in these systems hinders prompt response and immediate alerts in case of leakages, potentially leading to delayed mitigation measures and increased damage risks. The limitations of these existing systems underscore the necessity for a more adaptable, wireless, and efficient solution, prompting the development of the LoRa WAN-based Leak Detection System to overcome these constraints.

IV. PROPOSED SYSTEM

The proposed system revolutionizes traditional approaches by integrating two water flow sensors to detect leaks. Upon detection, the sensors transmit data to a microcontroller, facilitating communication with a LoRa transmitter for seamless data transfer. A LoRa receiver captures these signals, relaying the information to the microcontroller. The system's output interface, featuring an I2C LCD display and a buzzer, provides immediate visual and auditory alerts upon leakage detection. By leveraging LoRa WAN technology, this system ensures wireless communication, enhancing scalability, and enabling real-time monitoring for swift intervention, thereby revolutionizing leak detection with its wireless, responsive, and adaptable architecture.

4.1. Arduino IDE

NodeMCU is an IoT-focused development board and open-source firmware that utilizes Lua. It is designed for applications in the Internet of Things (IoT) domain. The board is powered by the ESP8266 Wi-Fi System-on-Chip (SoC) developed by Espressif Systems, and it incorporates hardware based on the ESP-12 module. It is Compact module size optimized for integration within IoT projects. The Node MCU board (Figure 4.1) serves as the central component in our IoT system design.

We selected the Node MCU microcontroller for our system based on several key criteria. Node MCU offers sufficient processing power and memory capacity to handle data processing and communication tasks efficiently. Its compatibility with a wide range of sensors and communication modules, including the LoRa module used in our system, ensures seamless integration and interoperability. Moreover, Node MCU's energy-efficient design aligns with the low-power requirements of our IoT application, enabling prolonged operation on battery power. The microcontroller features a robust architecture with multiple input/output pins for sensor interfacing and peripheral connectivity. Additionally, it incorporates a high-resolution analog-to-digital converter and supports popular communication interfaces such as SPI and I2C, facilitating sensor data acquisition and communication with external devices.

Figure 4.1 Node MCU

4.2. LoRa WAN

LoRa WAN stands out as a low power wide area network technology (LP-WAN) specifically engineered for Internet of Things (IoT) and smart sensor applications. Its long-range transmission capability coupled with minimal power consumption renders LoRa a pivotal technology in IoT networks.

In the field of LoRa technology, messages transmitted by any device can be received by one or multiple gateways. These messages are then relayed to the central network for processing. Given that the proposed system caters to rural and underground areas, where seamless signal transmission is imperative.

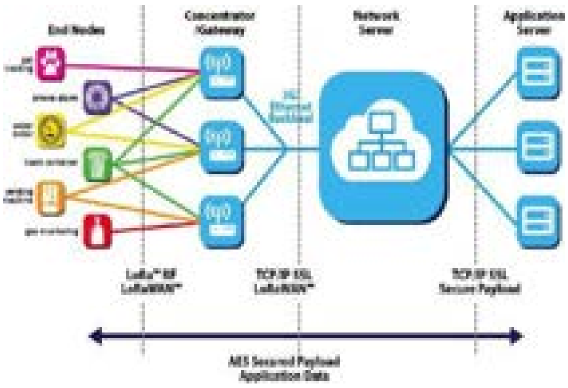


Figure 1 LoRa WAN System

LoRa WAN technology as shown in fig 1 adheres to IEEE 802.15.4 standards, operating within the 433MHz band with a bandwidth of up to 250KHz. Its chirp spread spectrum-based modulation enhances sensitivity. Data rates can reach up to 50kbps, covering an impressive range of up to 20 kilometres.

LoRa (Long Range) technology serves as the cornerstone of our water flow monitoring system, enabling efficient wireless communication between components. With its exceptional long-range capability and low power consumption, LoRa facilitates seamless data transmission over vast distances, making it ideal for IoT applications. This technology enhances the scalability of our system by enabling multiple nodes to communicate with a single gateway, simplifying infrastructure deployment and management. Compared to other wireless technologies, LoRa offers unparalleled range, power efficiency, and robustness, ensuring reliable communication and real-time monitoring of water flow and leak detection across diverse environments.

4.3. Water Flow sensor:

The YF-S201 water flow sensor shown in fig 2 is known for its advanced sealing capabilities, crafted for accurate measurements of water flow. Utilizing the Hall effect principle, it supports a flow rate range from 1 to 30 liters per minute. The module is equipped with three pins - Power, Ground, and Analog output - providing ease of connectivity. Noteworthy is the sensor's low current consumption and its ability to

Flow Rate Pulse Characteristics: Frequency (Hz) = 7.5 * Flow Rate (L/min)

Table 1 presents a detailed description of the water sensors, including their specifications and installation locations.

Table 2 displays the frequency and flow characteristics of the water sensors, providing essential data for calibration and monitoring.

Figure 4.3 Water Flow Sensor

Table 1 Description of the water sensors

S.no	Color	Pin type	Function
1.	Red	Power	To provide power for the module
2.	Black	Ground	Connected with the ground terminal
3.	Yellow	Signal	Analog output from the water

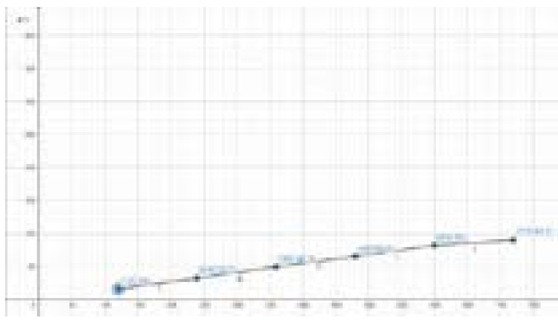


Figure 4.4 Flow depiction graph

S.no	Flow(L/H)	Frequency (Hz)
1.	120	16
2.	240	32.5
3.	360	49.3
4.	480	65.5
5.	600	82
6.	720	90.2

Table 2 Frequency and flow characteristics of the water sensors

4.4. *I2C LCD*

In our project, the leak detection system integrates 12 LCD (Liquid Crystal Display) screens, leveraging the versatility and ubiquity of this electronic display module. Among the various types available, the 16×2 LCD display stands out as a fundamental module widely employed across devices and circuits. Unlike traditional seven-segment displays, LCDs offer economic feasibility, easy programmability, and the flexibility to exhibit special characters, custom animations, and more.

A 16×2 LCD, as the name suggests, accommodates 16 characters per line across two lines, each character represented within a 5×7-pixel matrix. The LCD operates through two key registers: The Command register, which orchestrates predefined tasks such as initialization, screen clearing, cursor positioning, and display control, and the Data register, responsible for storing ASCII values corresponding to the characters to be displayed.

Central to our adapter design is the 8-bit I/O expander chip, PCF8574, facilitating the conversion of I2C data from the MCU (Arduino) into the serial data essential for LCD display. The I2C LCD adapter boasts four pins for connection to Arduino or any compatible microcontroller supporting the I2C protocol, alongside 16 pins interfacing with the LCD display.

Additionally, the adapter features two header pins dedicated to backlight control. One pin supplies 5V power, while the other manages the backlight LED. By default, the pins are interconnected, ensuring constant backlight illumination. Users have the option to remove the jumper to disable the backlight LED entirely or insert a potentiometer between the pins for precise control over backlight intensity. This adaptable design enhances the usability and customization potential of the LCD display system within our leak detection framework.

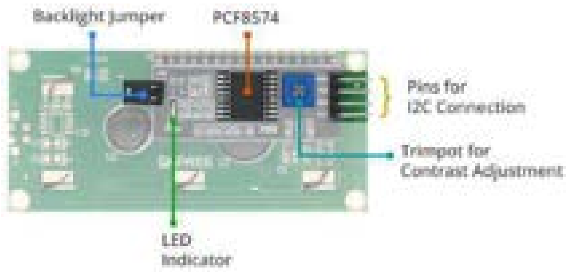


Figure 4.5 LCD Board

Working:

The I2C LCD operates by connecting to a microcontroller through two wires: SDA (Serial Data) and SCL (Serial Clock). It communicates using the I2C protocol, receiving commands and data from the microcontroller. An embedded controller within the LCD module interprets and processes this data to control the display, including showing characters, numbers, and symbols while efficiently using minimal hardware pins. This streamlined communication method simplifies wiring and hardware requirements, making it popular in various embedded systems.

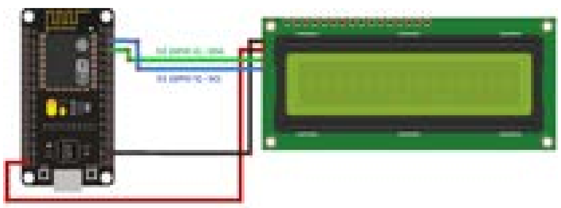


Figure 4.6 LCD Board and Display

4.5. Buzzer

The buzzer operates by simply supplying it with DC power ranging from 4V to 9V. While a basic 9V battery suffices, it's advisable to opt for a regulated +5V or +6V DC supply for optimal performance. Typically, the buzzer is integrated into a switching circuit, allowing users to activate or deactivate the buzzer as needed, according to predetermined timings and intervals.

Figure 4.7 Buzzer

V. IMPLEMENTATION AND RESULT

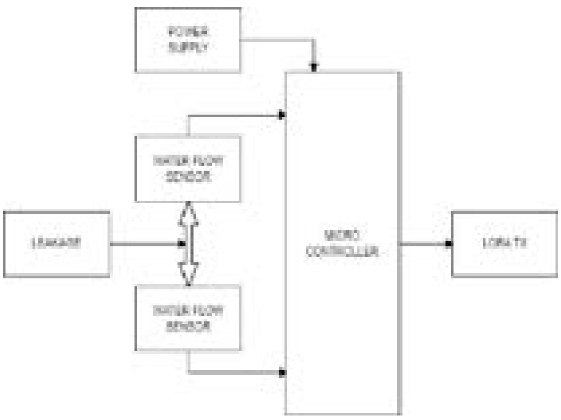


Figure 5.1 Transmitter part of the System

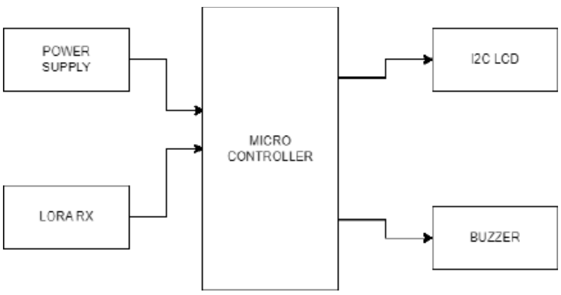


Figure 5.2 Receiver part of the System

The Block Diagram describes the System employs two water flow sensors interfaced with a microcontroller. Upon leakage detection, data is transmitted to the microcontroller, which then communicates with a LoRa transmitter for data transmission. On the receiving end, a LoRa receiver captures and relays signals back to another microcontroller. This microcontroller interfaces with an I2C LCD display and a buzzer. The LCD visually displays leakage information, while the buzzer provides an audible alert. The block diagram showcases the interconnectedness of the sensors, microcontrollers, LoRa transmitters, and receivers, forming a robust network for efficient leak detection and alerting mechanisms.

Moving to the description of the hardware picture, there is two parts in the board, Part A is the transmitter part which has sensors and LoRa transmitter connected to microcontroller. Part B is the receiver part where microcontroller is connected to the LoRa receiver, Lcd and a Buzzer

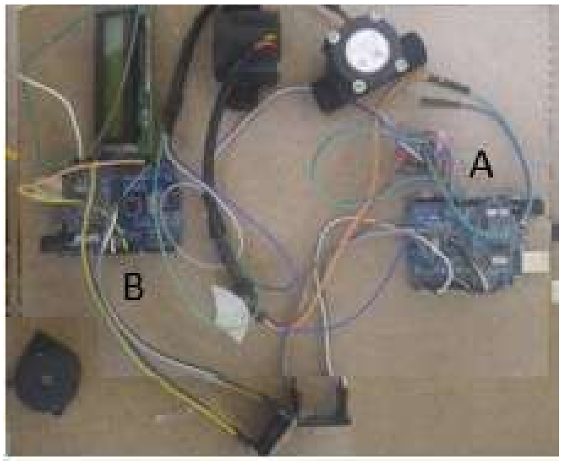


Figure 5.3 Hardware Setup

VI. WORKING

The system integrates a Node MCU ESP8266 microcontroller, LoRa WAN technology, water flow sensors (YF-S201), and an I2C LCD screen to constitute an efficient leak detection mechanism. Situated strategically in locations prone to water leakages, the water flow sensors persistently monitor the flow rate of water coursing through the system. Operating within an input voltage range of 4.5V to 18V DC, these sensors employ Hall Effect technology to precisely gauge flow rates spanning from 1 to 30 Liters per minute, with an accuracy tolerance of ±10%. Communication between the sensors and the NodeMCU ESP8266 microcontroller is established to process incoming data using the microcontroller's Ten silica 32-bit RISC CPU Xtensa LX106.

Upon identification of anomalies, such as significant deviations from anticipated flow rates, the NodeMCU ESP8266 triggers data transmission via the LoRa WAN technology. Leveraging license-free ISM bands at frequencies of 433, 868, or 915 MHz, the system can wirelessly transmit data across extended distances of up to 20 kilometres, facilitated by LoRa's chirp spread spectrum-based modulation. The transmitted data comprises vital details regarding the identified leakage, encompassing its location, severity, and anomalies in flow rates.

On the receiving end, another Node MCU ESP8266 microcontroller, equipped with a LoRa receiver module, captures the transmitted data. Subsequently, the microcontroller processes the data and displays pertinent information on the connected I2C LCD screen. This screen serves as a visual interface for users, providing real-time feedback on the detected leakage and enabling swift intervention to mitigate potential water damage.

To ensure prompt user notification, an audible alert mechanism, such as a buzzer, may be integrated into the system. This guarantees that users are promptly informed of detected leakages, even if they are not actively monitoring the visual display. The system's modular design facilitates scalability and adaptability across diverse environments and applications. Additional sensors or components can be seamlessly integrated to meet specific requirements, enhancing the system's versatility and customization for various use cases. In conclusion, the leak detection system furnishes a robust and effective mechanism for real-time identification and alerting of water leakages, enabling proactive measures to avert potential water damage and minimize associated risks.

To ensure high accuracy and reliability in water leak detection, our system employs several strategies. Firstly, we meticulously calibrate the water flow sensors to ensure precise measurement of flow rates and promptly detect abnormal patterns indicative of leaks. This calibration process involves thorough testing under various flow conditions to establish accurate baseline readings. Additionally, we implement advanced data filtering algorithms and signal processing techniques to reduce noise and enhance the accuracy of leak detection. These algorithms analyse sensor data in real-time, filtering out spurious readings and identifying genuine anomalies associated with leaks. Furthermore, our system incorporates redundancy measures and fail-safe mechanisms to enhance reliability. This includes backup power sources to ensure uninterrupted operation in case of power outages, sensor redundancy to mitigate the risk of sensor failures, and error detection and correction algorithms to identify and rectify data inconsistencies. By combining these strategies, our system delivers robust and reliable water leak detection capabilities, empowering users to take timely action and prevent potential water damage.

To enhance the performance of the pipeline monitoring system, several key measures need to be implemented, Streamlining the communication and processing unit is essential to ensure efficient data transmission and analysis. This involves optimizing the system architecture and algorithms to minimize latency and maximize throughput. Upgrading sensors to enhance sensitivity and accuracy is crucial for detecting even minor variations in flow rate and pressure. This may involve adopting advanced sensor technologies or improving calibration methods. Exploring alternative materials or configurations for pipeline construction can minimize friction and pressure losses, optimizing the system's overall performance.

To ensure the reliability of the pipeline monitoring system, the following measures need to be implemented, incorporating redundancy in sensor placement and data transmission is essential to mitigate the risk of sensor failures. Additionally, implementing fault-tolerant mechanisms can ensure continuous monitoring capabilities even in the event of system failures. Proper calibration and maintenance of sensors are critical to prevent inaccuracies in readings and ensure consistent performance. Regular inspections and calibration checks should be conducted to maintain reliability. Validating the system's performance under various conditions, including different fluid types and pipeline configurations, is necessary to ensure its reliability across diverse scenarios

VII. CONCLUSION

In conclusion, this study has proposed a robust solution for efficient identification of water leakages. By integrating advanced sensor technology with LoRa communication protocols, it offers real-time monitoring capabilities. The seamless interaction between water flow sensors, microcontrollers, LoRa transmitters, and receivers enables rapid data transmission and immediate responses upon leak detection. The system's output interface, comprising an I2C LCD display and a buzzer, ensures prompt visual and auditory alerts, empowering users to take immediate action. Its wireless architecture enhances adaptability, scalability, and reliability in various environments, promising proactive leak detection and preventing potential water damage. This project signifies a significant leap toward enhancing water management systems through cutting-edge technology, ensuring timely intervention, and safeguarding against the detrimental impacts of water leakages. The impact of the pipeline network design on the pressure of fluid in the pipeline is an area that requires future study.

VII. REFERENCES

[1] Mahmoud Meribout, Lyes Khezzar, Abdelwahid Azzi, Nabil Ghendour, Leak detection systems in oil and gas fields: Present trends and future prospects, Flow Measurement and Instrumentation. Volume 75, 2020 [1]

[2] Xuan Hu, Yongming Han, Bin Yu, Zhiqiang Geng, Jinzhen Fan, Novel leakage detection and water loss management of urban water supply network using multiscale neural networks, Journal of Cleaner Production, Volume 278, 2021 [2]

[3] Christos Spandonidis, Panayiotis Theodoropoulos, Fotis Giannopoulos, Nektarios Galiatsatos, Areti Petsa, Evaluation of deep learning approaches for oil & gas pipeline leak detection using wireless sensor networks, Engineering Applications of Artificial Intelligence, Volume 113,2022,104890, ISSN 0952-1976 [3]

[4] R. Vanijjirattikhan et al., "AI-based acoustic leak detection in water distribution systems," Results Eng., vol. 15, p. 100557, Sep.2022, Doi: 10.1016/j.rineng.2022.100557. [4]

[5] J. A. Coelho, A. Glória, and P. Sebastiao, "Precise Water Leak Detection Using Machine Learning and Real-Time Sensor Data," J. IoT, vol. 1, pp. 474–493, 2020, Doi: 10.3390/iot1020026. [5]

[6] P. Kumar and S. Mishra, "GPLDA Bot: Raspberry-Pi based Gas Pipe Leakage Detection and Alerting using," Turkish Online J. Qual. Inq., vol. 12, no. 2, pp. 540–551, 2022. [6]

[7] S. Adsul, A. K. Sharma, and R. G. Mevekari, "Development of leakage detection system," 2016 International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Pune, India, 2016, pp. 673-677, Doi: 10.1109/ICACDOT.2016.7877671. [7]

[8] B. Mitrică, E. Mitrică, P. Enciu, and I. Mocanu, "An approach for forecasting of public water scarcity at the end of the 21st century, in the Timiş Plain of Romania," Technol. Forecast. Soc. Change, vol. 118, pp. 258–269, May 2017, Doi: 10.1016/j.techfore.2017.02.026. [8]

[9] O. A. Agbolade, S. A. Oyetunji, and F. M. Dahunsi, "Sustainable Business Model for Long-Range Wireless Wide Area Based Internet of Things Network: The Heterogeneous approach," in 2023 International Conference on Science, Engineering and Business for Sustainable Development Goals (SEB-SDG), 2023, vol. 1, pp. 1–6, Doi:10.1109/SEBSDG57117.2023.10124648. [9]

[10] B. R. Scanlon et al., "Global water resources and the role of groundwater in a resilient water future," Nat. Rev. Earth Environ. 2023 42, vol. 4, no. 2, pp. 87–101, Jan. 2023, Doi: 10.1038/s43017-022-00378-6. [10]

[11] H. R. Vanaei, A. Eslami, and A. Egbewande, "A review on pipeline corrosion, in-line inspection (ILI), and corrosion growth rate models," Int. J. Press. Vessel. Pip., vol. 149, pp. 43–54, Jan. 2017, Doi: 10.1016/j.ijpvp.2016.11.007. [11]

[12] P. Gopalakrishnan, S. Abhishek, R. Ranjith, R. Venkatesh, and S. V. Jai, "Smart Pipeline Water Leakage Detection System," Int. J. Appl. Eng. Res., vol. 12, no. 16, pp. 5559–5564, 2017, [Online].[12]

[13] O. A. Agbolade, F. M. Dahunsi, and S. A. Oyetunji, "Heterogeneous LoRa WAN Deployment for Application Dependent IoT Networks," J. Eng. Technol. Ind. Appl., vol. 8, no. 34, pp.411,2022, Doi: https://doi.org/10.5935/jetia.v8i34.798. [13]

[14] B. R. Scanlon et al., "Global water resources and the role of groundwater in a resilient water future," Nat. Rev. Earth Environ. 2023 42, vol. 4, no. 2, pp. 87–101, Jan. 2023, Doi: 10.1038/s43017-022-00378-6. [14]

[15] O. Varis, A. K. Biswas, C. Tortajada, and J. Lundqvist, "Megacities and water management," Int. J. Water Resour. Dev., vol. 22, no. 2, pp. 377–394, Jun. 2006, Doi: 10.1080/07900620600684550. [15]

Implementation of Patrolling Robots for Apartments using Cloud Technology

Nagaraj P
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
nagaraj.p@klu.ac.in

M. K. Nagarajan
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
nagarajan@klu.ac.in

B. Surendra Reddy
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
bandisurendrareddy123@gmail.com

Y. Ramu
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
yerramsettiramu08@gamil.com

G. Venkatesh
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
Venkygoddeti94@gmail.com

A. Sravya
Department of Computer Science and Engineering
Kalasalingam Academy of Research and Education,
Virudhunagar, Tamilnadu, India
atlasravya7@gmail.com

Abstract—

The objective of this project is to create a smart flat patrolling robot with a variety of accessories, including an ESP controller, DC gear motor, motor driver, LCD, buzzer, rechargeable batteries, high-definition night vision 360° rotation smart camera, IoT cloud app and wheels. This robot's purpose is to improve security and monitoring within an apartment building by independently patrolling predetermined regions and delivering real-time monitoring via a cloud-based Internet of Things application. The ESP controller, which acts as the robot's central processing unit, and the DC gear motor with a motor driver, which enables mobility, make up the robot's essential parts. The ESP controller, which takes instructions from the IoT cloud app, manages the robot's movement and navigation. The app enables users to monitor live camera feed, remotely control the robot, and get notifications for any questionable activity. The robot can take clear video footage even in low light thanks to the high-definition night vision 360° rotation smart camera, which can also spin 360 degrees to cover a large area of vision. This camera is essential for keeping an eye on and documenting any unexpected activity while patrolling. The design and development of a flexible, intelligent apartment patrolling robot that combines several components to improve security and surveillance is shown in this project. This robot is a useful addition to apartment complex security systems thanks to the ESP controller, DC gear motor, high-definition video, IoT cloud app, and other characteristics.

Keywords— **Microcontroller, ATMEGA328P, Sensors, Smart camera, Arduino IDE, IOT, Robot**

I.INTRODUCTION

Innovation has seen a rapid and significant shift in the fields of computerization and mechanical technology, which span many different industries. The process of closely monitoring or controlling an individual, group, etc. [1]. is known as reconnaissance. especially one that is doubtful or under care. Hence, surveillance is necessary in public areas, workplaces, industrial locations, and border Thomasen [2] involves a comprehensive review of case law, statutes, legal doctrines, and academic literature to explore the intersection of private law and public space. crossings. It is mostly used for workout verification. People can demonstrate reconnaissance both indoors and outdoors, or they can use implanted frameworks like robots and other robotization devices to help them [2].

A robot is just an electrical device that has been programmed to carry out specific tasks. As such, it may replace human labor, provide incredibly precise results, and effectively surpass human limitations [3]. One of the remarkable developments in advanced mechanics is the replacement of humans in reconnaissance fields.

The Arduino Uno microcontroller, which serves as the robot's brain, is part of the assembly. The DC engines, wheel body, batteries, Wi-Fi module (ESP8266 12e), and various sensor types—such as an IR sensor for pit identification and an ultrasonic sensor for obstacle recognition—are also included in this robot [4]. Both manual and automatic controls are available for the robot.

The robot and user end communicate via the Internet of Things. Arduino programming, which is used for IOT creation projects, can be used to achieve this [5].

Since both are connected, the orders are sent by the robot using Arduino programming, and they are received by the Arduino microcontroller using a Wi-Fi module. This enables wireless control of the robot [6]. We use a remote-sensing camera for this task, which provides sound and video data that the client can obtain [7].

II. LITERATURE REVIEW

Ranjan et al. [1] analyze a wide range of research papers, technical reports, and industry publications to synthesize the current state of wireless communication technologies. The study also incorporates expert insights and case studies to provide a holistic understanding of the methodologies, technologies, and trends shaping the field.

The study also engages in qualitative analysis of relevant legal cases and expert opinions to assess the implications of personal remote surveillance technology on privacy rights and legal liability.

Zhang, & Wei. [3] establishes conventional stiffness model (TSM) is easily derived. For the classic stiffness model, one does not account for the Centre limb (if applicable).

LaValle. [4] has been taken into account in the proposed classification. Numerous previous assessments have distinguished between two types of environments: those that are dynamic (online versus offline path planners) and those that are smaller (local versus global). Online is typically linked to local, and offline to global.

Cheng et al. [5] used ADAMS software, and a simulation model of the 6R robot was created. The optimal constraint force of the 6R robot at a specific moment is then qualitatively analyzed. Following the measurement and computation of the parameters, the results demonstrate that the identified dynamic parameters have exceptionally high accuracy, which validates the correctness and accuracy of the dynamic model.

Wagner [6] states depending on the robot's attributes—such as its capacity for independent movement—and the number of robots at hand, various strategies for robot-guided emergency evacuation can be employed. Generally speaking, we anticipate that a significant amount of environmental tuning will be required and finished before the evacuation robots are deployed.

Lee [7] synthesizes scientific research on a particular subject by carefully examining both recent and historical works. Research topics, the search procedure, inclusion/exclusion criteria, quality evaluation, data collecting, and data analysis are typical steps in the systematic literature review approach.

Xu [8] uses the two components of the VR-based urban planning 3D simulation design system are 3D design and VR simulation design. The urban-planning 3D simulation design system is created using the open GL module as a foundation. The interactive navigation of the urban planning scene and the model traversal rendering are the functions of the VR simulation design module.

SLAPAKOVA, et al. [9] constrain the methodological options available for investigating possible technology and mission-related paths during the investigation. These limitations led to the prioritization of a desk research methodology and a wide mapping of open-source data over alternative futures and foresight approaches.

Faiza Gul et al. [10] evaluate a range of methods including traditional sensor-based approaches, simultaneous localization and mapping (SLAM), and modern machine learning techniques. Through a rigorous analysis of each technique's strengths, limitations, and applicability. This study provides valuable insights into selecting appropriate navigation strategies for diverse robotic platforms, considering factors like environment complexity, accuracy, real-time performance, and adaptability.

III.PROPOSED METHODOLOGY

Define the overall architecture of the surveillance robot system. Identify the key components such as the robot platform, wireless camera, sensors, Arduino microcontroller, and communication module [11which explain how these components are interconnected to form a cohesive system. Specify the hardware components used in the robot, including the robot chassis, wheels, motors, wireless camera, Arduino microcontroller, and sensors [12]. Provide details on the assembly of the hardware components to build the surveillance robot. Include any modifications or customizations made to the hardware [13].

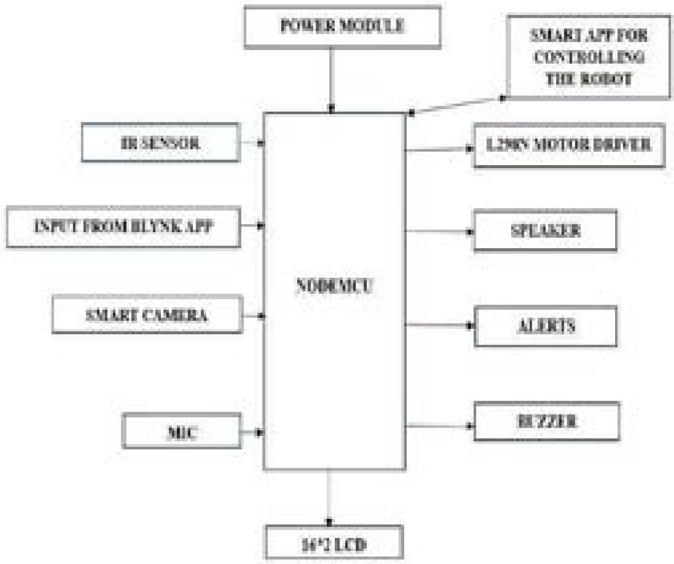


Fig 1: Process Flow chart

Fig. 1 describes the details of the types of sensors integrated into the robot, such as proximity sensors, infrared sensors, or any other relevant sensors for surveillance. Explain the purpose of each sensor and how it contributes to the robot's functionality. Discuss the connection and communication between sensors and the Arduino microcontroller. Describe how Internet of Things (IoT) technology is incorporated into the surveillance system. Explain the communication protocol used for remote control and data transmission between the robot and the user interface (mobile or laptop) [14]. Discuss the security measures implemented for secure communication. Provide the code snippets or algorithms used to program the Arduino microcontroller. Explain the logic behind the programming, especially focusing on how the Arduino controls the robot's movement, processes sensor data, and communicates with the IoT module [15]. Detail the development of the user interface for controlling the robot, whether through a mobile app or a web-based interface. Explain the features of the user interface, such as manual control options, and live streaming.

IV.IMPLEMENTATION AND RESULT

Fig 2: Smart Camera

Fig. 2 shows the smart camera, we can watch live video recording, motion detection, tracking, etc.,

The customer segment and the robot area are the two main divisions of the framework. A laptop or mobile device may be used by the user portion to communicate with the robot end. Thus, the client segment might be more adaptable than those using a traditional fixed PC framework using a PC or a mobile phone. We can use our amazing devices for communication. To implement the concept of increasing the range into practice, we can connect the user section to the Internet, which is the core tenet of the Internet of Things. The client framework is interfaced with the web using Blynk programming. IoT apps and models are created using this. Therefore, by using the Blynk program, we can send commands and effectively operate the automated car.

We use an ESP-8266 microcontroller, a built-in mechanical vehicle component, at the robot end. It is put on the robot's body or skeleton. Beneath the chassis, the wheels are connected to 60 rpm DC engines. The 12 volts needed by each engine are provided by an external battery source. An engine driver connects the engines to the regulator.

Two engines each have an engine driver, which is used for intensification. IDE programming is used to modify the microcontroller so that the robot moves in the correct direction. This is the associated manual mode task. Ultrasonic sensors are among the many sensors that are connected to the microcontroller via the I/O pins.

Because the infrared sensor operates on the principle of reflection, obstructions are detected through the exchange of signals. Similarly, infrared sensors are used to generate and differentiate infrared radiations, which enables the identification of temperature variations. Here, we are including a PIR sensor to detect movement. The camera turns that direction and switches on the light for visibility at the precise moment when movement is detected. A high-resolution, 360-degree rotating night vision camera with vocal communication is another innovative element of this concept.

Fig 3: Implementation of the project

Fig. 3. illustrated the overall layout of a video web-based wheel robot that is controlled by Nodemcu and a webpage created using Arduino and Blynk. Essentially, this surveillance robot transmits live video from a camera module over the internet and is controlled by a smartphone. The site page allows for manual control and monitoring of the robot's evolution. The accompanying advancements were used to truly examine the task's outcomes

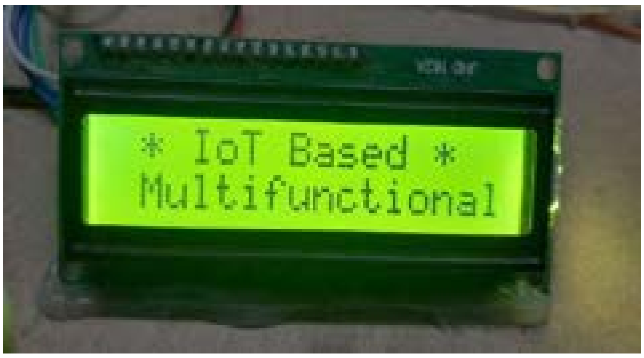


Fig 4: Starting command

Fig. 4 shows how to start the smart flat patrolling robot, turn on the power source, and make sure all the parts are correctly plugged in and working. Communicate with the robot using the designated interface, such as the IoT cloud app. Users may access live video feeds, change movement parameters, and start autonomous patrolling mode—among other commands—from the app. Users can also adjust alarm preferences and keep an eye on battery life. The robot's LCD screen shows the status and commands that have been carried out visually. Consult the user manual or get in touch with technical support if you need any more help.

Fig 5: Guessing the motion

Fig.5 Shows enable motion, and ensure the robot is powered on and connected to the IoT cloud app. From the app interface, select the desired patrol route or specify waypoints for the robot to navigate. Utilizing the integrated DC gear motor and motor driver, the robot will autonomously move along the predefined path, adjusting speed and direction as needed to avoid obstacles. Users can monitor the robot's movement in real time via the app's live camera feed and receive notifications of any detected anomalies. For manual control, use the directional commands provided in the application to people the robot remotely. The system's integration ensures seamless and responsive motion capabilities for effective patrolling.

Fig 6: Implementation of operation

Fig.6 shows the initialization of the system and makes sure all parts are operating appropriately before putting the smart flat patrolling robot into action. The robot's brain is the ESP controller, which coordinates its movements and receives commands from the IoT cloud app. The controller uses the motor driver to activate the DC gear motor in response to a patrol command sent through the app, which starts movement. Concurrently, the 360° rotation smart camera with high-quality night vision starts recording live footage of the environment. This footage is sent to the cloud-based IoT application for remote monitoring. Through the app's user interface, users may communicate with the robot by directing its actions, viewing video feeds, and instantly receiving alarms.

V. CONCLUSION

This project proposes a system for creating robots for observation design. Using the concept of IOT overcomes the problem of limited-range observation. With the help of the PC and diverse physical interface, we can exert some control over the robot. It should also be feasible to do programmed observation. The proposed robot's diminutive size allows it to enter spaces that are inaccessible to humans. One of the simplest forms of innovation in the gadget industry is probably remote innovation. This provides a highly efficient and economically astute robot that reduces and eliminates human labor while powerfully completing observation tasks.

REFERENCES

1. Ranjan, P., Rao, R. S., Kumar, K., & Sharma, P. (2022). Wireless Communication: Advancements and Challenges. CRC Press.
2.Thomasen, K. (2022). Private Law & Public Space: The Canadian Privacy Torts in an Era of Personal Remote-Surveillance Technology (Doctoral dissertation, Université d'Ottawa/University of Ottawa).
3.Zhang, D., & Wei, B. (2016). Study on the kinematic performances and optimization for three types of parallel manipulators. Machines, *4(4), 24.
4.LaValle, S. M. (2006). Planning algorithms for autonomous mobile robots. Cambridge, MA: MIT Press.
5. Cheng J, Bi S, Yuan C, Chen L, Cai Y, Yao Y. A Graph Theory-Based Method for Dynamic Modeling and Parameter Identification of 6-DOF Industrial Robots. Applied Sciences. 2021; 11(22):10988. https://doi.org/10.3390/app112210988
6.Wagner, A. R. (2021). Robot-guided evacuation as a paradigm for human-robot interaction research. Frontiers in Robotics and AI, 8, 701938.
7. Lee, I. (2021). Service robots: A systematic literature review. Electronics, 10(21), 2658.
8. Xu, S. (2021). Three-dimensional visualization algorithm simulation of construction management based on GIS and VR technology. Complexity, 2021, 1-13.
9. SLAPAKOVA, L., FUSARO, P., BLACK, J., & DORTMANS, P. Supporting the Royal Australian Navy's Campaign Plan for Robotics and Autonomous Systems.
10. Faiza Gul, Wan Rahman & Syed Sahal Nazli Alhady | (2021) A comprehensive study for robot navigation techniques, Cogent Engineering, 6:1, 1632046, DOI: 10.1080/23311916.2019.1632046
11. Nagaraj, P., Rajkumar, T. D., Rakesh, S., Durga, A. K. S., Jyothi, M. L., & Nithin, C. G. S. (2023, July). AI-based Leaf Disease Identification Robot using IoT Approach. In *2023 2nd International Conference on Edge Computing and Applications (ICECAA)* (pp. 1375-1379). IEEE.
12. Perumal, B., Nagaraj, P., Deny, J., Rajesh, V., Kumar, K. M., & Reddy, K. M. (2023, December). Design and Development of Vision-Based Ibot-Kare-Robot for Automation. In *2023 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)* (pp. 1-5). IEEE.
13. Muneeswaran, V., Nagaraj, P., Anuradha, L., Lekhana, V., Vandana, G., & Sushmitha, K. (2023, March). An Elucidative Review on the Current Status and Prospects of Eye Tracking in Spectroscopy. In *Congress on Control, Robotics, and Mechatronics* (pp. 225-236). Singapore: Springer Nature Singapore.
14. Pandiaraj, A., Nagaraj, P., Kumar, P. B., Rasi, P., Bhavani, M. N. L., & Reddy, C. V. (2023, June). Blockchain Using Private Cloud for Secure EHRS Systems. In *2023 3rd International Conference on Pervasive Computing and Social Networking (ICPCSN)* (pp. 873-878). IEEE.
15. Nagaraj, P., Rajkumar, T. D., Vairam, V. N., Akther, M. N., Saravana, H., & Kannan, S. S. (2023, July). Adaptive Electric Distribution for Catastrophic Floods Integration using IoT. In *2023 4th International*

Conference on Electronics and Sustainable Communication Systems (ICESC) (pp. 413-419). IEEE.

Enhancing Chronic Kidney Disease Diagnosis Through Densenet-121 Approach

1st K.Kausalya
Assistant Professor
dept of Information Technology
Easwari Engineering College
Chennai, Tamilnadu
kausalyamurthy@gmail.com

2nd B.Chandra
Assistant Professor
dept of Information Technology
Easwari Engineering College
Chennai, Tamilnadu
Chandra.klnce@gmail.com

3rd S. Barath Kumar
Student
dept of Information Technology
Easwari Engineering College
Chennai, Tamilnadu
barathbarath65432@gmail.com

4th A.Mohammed Riyaz
Student
dept of Information Technology
Easwari Engineering College
Chennai, Tamilnadu
mohammedriyaza01@gmail.com

***Abstract*— This research work focuses on creating an advanced deep learning framework designed for the automated classification of kidney stones, cyst, normal, tumor a critical task in medical imaging with significant implications for diagnosis and treatment planning. The proposed system incorporates deep learning, specifically a pre-trained DenseNet121 model. To adapt this model to the intricacies of medical imaging data, this study introduces custom layers atop the pre-trained architecture. These layers are designed to extract and analyze features relevant to kidney stone classification, enabling the model to differentiate between four distinct classes: Cyst, Normal, Stone, and Tumor. Importantly, to ensure the robustness and generalization of the proposed model, advanced regularization techniques such as dropout and batch normalization are employed during training. Through extensive experimentation and evaluation on diverse datasets, this study demonstrates the efficacy and reliability of the proposed approach. The results highlight the system's capability to achieve high accuracy in classifying kidney stones, highlighting its capacity as a valuable resource for healthcare professionals in nephrology and urology. By facilitating accurate and efficient diagnosis, the proposed framework aims to contribute to improve clinical decision-making processes and ultimately enhance patient care in the domain of kidney stone management.**

Keywords*— *Chronic Kidney diagnosis, DenseNet121, Transfer learning, Deep Learning model

I. Introduction

This study address the significant health risks associated with kidney abnormalities and the challenges faced by traditional diagnostic methods, particularly manual interpretation of medical imaging. This study specifically highlights the need for automation in classifying kidney stones, emphasizing the rationale behind selecting DenseNet121, a pre-trained model known for its efficient feature extraction capabilities. Transfer learning is interpreted as a key concept, explaining how the model leverages knowledge from ImageNet to enhance performance in kidney stone classification. Furthermore, the introduction discusses the customization of the model through the incorporation of custom layers, designed to capture task-specific features for accurate classification of kidney stone images. To address the challenges like overfitting and improve generalization, the established techniques such as dropout regularization and batch normalization are integrated into the training process, ensuring the model's reliability and applicability across diverse kidney stone cases. This research aims to develop a robust classification system to assist medical professionals in making informed treatment decisions, ultimately enhancing patient outcomes in kidney stone diagnosis and classification through extensive experimentation on a diverse dataset.

II. Literature Survey

Chen et al [1] address early Chronic Kidney Disease (CKD) diagnosis on the Internet of Medical Things (IoMT) platform. This system estimates the intersection of medicine, technology, and deep learning for efficient early detection. [2] Debal and Sitote explored CKD prediction with various machine learning models, emphasizing the need for ongoing research to predict CKD stages and improve treatment recommendations. They address class imbalance using oversampling and leverage machine learning models for binary and multi-class CKD stage classification. [3] Md. Ariful Islam et al. incorporated machine learning techniques for predicting CKD, showcasing recent advancements in the field. Using 12 classifiers, including XgBoost, the study contributes to ongoing efforts for early CKD detection, underlining the interdisciplinary nature of the research originating from various departments in Bangladesh. [4] Almasoud and Ward's research work has utilized machine learning methodology for the diagnosis of Kidney Disease involving the fewest predictors. Their approach involves feature selection to minimize predictors, achieving high accuracy with the Gradient Boosting classifier. The emphasis is on early detection with minimal features compared to previous research. [5] Tekale et al.'s study highlighted the importance of early CKD detection, especially when specific symptoms are lacking. Decision Tree and Support Vector Machine algorithms are used to maximize accuracy in predicting CKD and determining severity. The literature review highlighted the versatility of machine learning algorithms in addressing CKD through diverse capabilities. [6] C. P. Kashyap et al. employed a combination of feature selection and classification techniques to predict Chronic Kidney Disease with an accuracy rate of 89%. This model incorporated various types of data including demographic, clinical, and laboratory information, demonstrating the potential of machine learning in the identification of significant risk factors for CKD. [7] Anifah and Haryanto's research is centered on the identification of severity of Chronic Kidney Disease through the utilization of template matching feature selection statistics. The proposed method utilized Euclidean distance and two sets of experiments with different parameters to identify patterns indicative of CKD severity.

III. Existing System

The existing approach incorporates pre-trained CNN models and leverages transfer learning techniques to identify and classify kidney tumors based on features extracted from CT images. The model specifically focuses on Regions of Interest (ROIs) to enhance accuracy in detection and characterization.

Despite its innovative approach, there are certain limitations to consider. The model's primary emphasis on tumor detection may result in overlooking broader aspects of kidney health assessment, potentially limiting its applicability for comprehensive disease analysis. Additionally, achieving optimal accuracy in tumor classification poses a challenge, and concerns arise regarding false-positive rates and adaptability across diverse datasets due to the reliance on Support Vector Machine (SVM) classification. Moreover, the SVM-based classification may have limitations in handling intricate variations within tumor characteristics, which could impact the model's robustness. To address these issues, it is essential to refine the classification strategy, explore alternative algorithms, and mitigate potential biases. Continuous efforts to improve accuracy, broaden the scope of analysis, and ensure adaptability across varied scenarios are crucial for advancing the effectiveness of the proposed model in kidney tumor detection and early disease identification.

IV. Proposed System An Overview

The proposed system introduces a holistic approach to kidney stone classification, leveraging the capabilities of sophisticated deep learning methods. The framework incorporates the use of a pre-trained DenseNet121 model, well-known for its robust feature extraction capabilities to analyse medical images and detect kidney stones. Initially, the model engages in feature extraction, detecting intricate patterns and pertinent features from input images taken from ImageNet dataset. Custom layers are subsequently added atop the pre-trained model to tailor it specifically for the nuances of kidney stone classification. These layers encompass convolutional, pooling, normalization, and fully connected layers, culminating in a SoftMax output layer designed for multi-class classification.

To bolster model performance and address overfitting concerns, dropout regularization and batch normalization techniques are seamlessly integrated. Throughout the training phase, the model refines its ability to distinguish between various types of renal Cyst, Normal, Stone, and Tumor optimizing the loss function for categorical cross-entropy via the Adam optimization technique. As the system undergoes iterative epochs of training, it fine-tunes its parameters, enhancing its proficiency in accurately classifying kidney stones based on learned features.

Importantly, the proposed system not only achieves high accuracy in classification but also prioritizes interpretability. By visualizing the learned features, the system empowers medical professionals to make informed decisions. This advancement in medical image analysis represents a significant stride, providing a reliable and efficient tool for diagnosing kidney stones. The potential impact on patient care and medical diagnosis processes is substantial, signifying a valuable contribution to the field.

V. Methodology

DenseNet121 is a CNN model architecture renowned for its dense connectivity pattern. Comprising a total of 121 layers, this network achieves deep representation learning with enhanced parameter efficiency. Here, we delve into each layer of DenseNet121 to provide a comprehensive understanding:

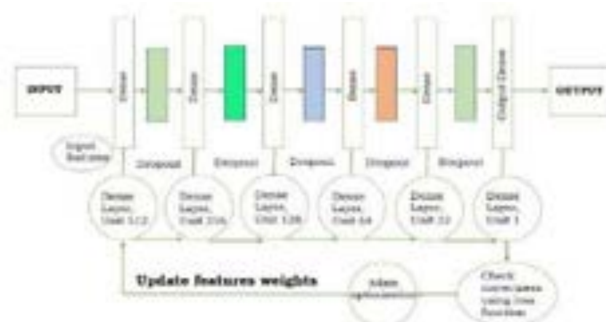


Figure 1. Architecture of the DenseNet121

5.1 Data Collection and Exploration

In this research, the dataset utilized comprises kidney CT scan images, categorizing them into four distinct classes: Cyst, Normal, Stone, and Tumor. These images form the basis for training and evaluating a model designed for the classification of abnormalities in kidney. The dataset is systematically organized within the 'data/' directory. The dataset was collected from PACS (Picture archiving and communication system) from different hospitals in Dhaka, Bangladesh where patients were already diagnosed with having a kidney tumor, cyst, normal or stone findings. Both the Coronal and Axial cuts were selected from both contrast and non-contrast studies with protocol for the whole abdomen and urogram. The Dicom study was then carefully selected, one diagnosis at a time, and from those we created a batch of Dicom images of the region of interest for each radiological finding. Following that, we excluded each patient's information and meta data from the Dicom images and converted the Dicom images to a lossless jpg image format. After the conversion, each image finding was again verified by a radiologist and a medical technologist to reconfirm the correctness of the data.Our created dataset contains 12,446 unique data within it in which the cyst contains 3,709, normal 5,077, stone 1,377, and tumor 2,283

5.1.1 Initial Inspection

The exploration begins with a detailed examination of the dataset to understand its composition, diversity, and potential challenges. This involves assessing the number of images per class, their resolutions, and any variations in image characteristics.

5.1.2 Visual Exploration

To provide a visual representation of the dataset, a Matplotlib and OpenCV-based exploration is conducted. Sample images from each class are displayed, offering a qualitative glimpse into the features that differentiate Cyst, Normal, Stone, and Tumor images.

5.2 Data Preprocessing:

The image processing workflow involves initially loading a medical image related to CKD, typically represented in the RGB color format. Subsequently, the image undergoes a crucial conversion from BGR to RGB format using the `cv2.cvtColor` function, as various image processing and machine learning libraries expect images in the RGB format. The next step focuses on reshaping the image into a 2D array of pixels, flattening it to facilitate easier manipulation and analysis of pixel values. Following this, the quantity of clusters utilized in the K-Means clustering algorithm, denoted as `n_clusters`, is defined, serving as a hyperparameter to determine the groups the algorithm should identify in the pixel data. The K-Means model is then instantiated and fitted to the pixel data, iteratively assigning pixels to clusters and updating cluster centroids until convergence. The resulting cluster centroids and labels are obtained, with centroids representing average color values, and labels indicating cluster assignments for each pixel. The obtained labels are reshaped to match the original image shape, forming a 2D array that represents cluster assignments for each pixel in the original image. While the provided code snippet doesn't explicitly detail the data preparation for classification, it typically involves creating a dataset with features derived from clustered images and corresponding labels. This dataset may undergo additional steps such as splitting into training and testing sets, pixel value scaling, and label encoding for training a classification model. In summary, this comprehensive image processing pipeline, encompassing conversion, reshaping, clustering, and preparation for classification, contributes to the analysis of CKD-related medical images

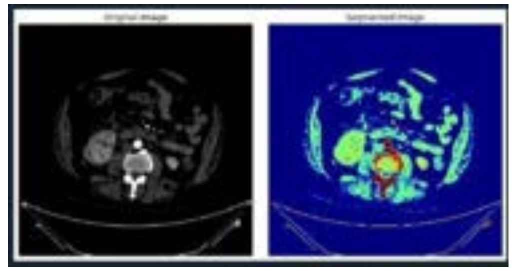


Figure 2: Segmentation

5.3 Feature Extraction:

In the model training phase for Chronic Kidney Disease (CKD) image analysis, the neural network undergoes exposure to a labeled training dataset comprising kidney CT scan images and corresponding classes (Cyst, Normal, Stone, Tumor). The primary goal is to enable the model to learn essential patterns and features crucial for accurate classification.

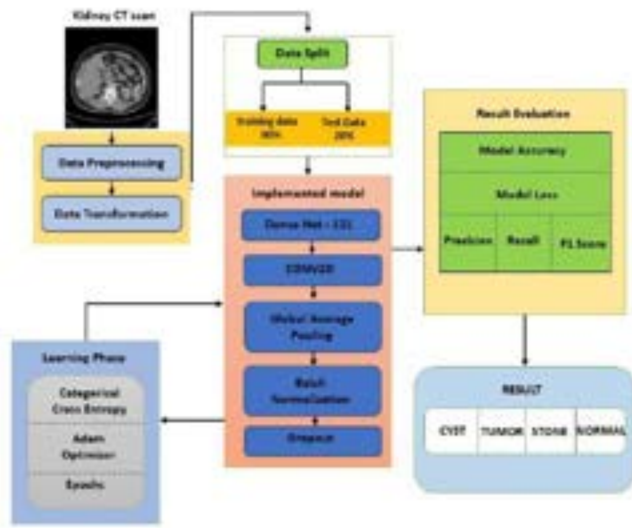


Figure 3: Architecture of the Proposed System

5.3.1 Input Layer:

In neural networks, the input layer serves as the initial entry point for data and defines the format and size of expected input. This configuration involves specifying dimensions such as image height, width, and color channels, as well as determining the size of data batches processed simultaneously and incorporating preprocessing steps like normalization or scaling. Ensuring that the data aligns with the specified format enables compatibility and structures the flow of data through the neural network. The significance of the input layer lies in its impact on how subsequent layers interpret and process input data, ultimately influencing the model's capacity to learn and generate accurate predictions related to CKD diagnosis and prognosis. For instance, a defined input shape of (None, 65, 65, 3) indicates that input images are 65x65 pixels with 3 color channels(RGB).

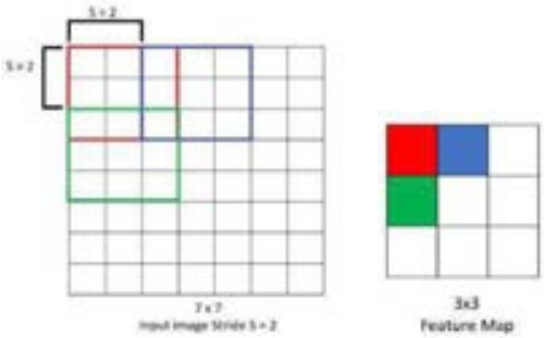


Figure 4 : Kernel 7X7 Feature Map Kernel 3x3

5.3.2 CONV2D Layer:

The model architecture combines a Conv2D layer and a pre-trained DenseNet121 model to identify kidney abnormalities in CKD data input. The Conv2D layer initially extracts features from input images using convolutional operations, generating feature maps crucial for identifying patterns related to abnormalities such as cysts and tumors. With an output shape of (None, 65, 65, 3) and 84 parameters, this layer provides basic feature extraction.

Figure 5: Feature Map Images of kidney CT-Scan

Figure 6 : Filters Image

5.3.3 Global Average Pooling 2D (GAP2D):

GAP2D is pivotal for feature extraction and dimensionality reduction in image classification tasks. In the CKD context, it occurs typically before fully connected layers. By computing the average value of each feature map across its spatial extent, GAP2D reduces spatial dimensions to 1x1, preserving depth. This aids in capturing essential features from medical images while enhancing model interpretability and efficiency.

5.3.4 Batch Normalization:

This layer applies normalization to its inputs using 4,096 parameters. Batch normalization is essential for improving training stability and speed, especially in the context of Chronic Kidney Disease (CKD) analysis. It standardizes the activations within mini-batches, helping to mitigate internal covariate shift and improve the efficiency of learning.

In CKD image analysis, where accuracy is crucial, batch normalization plays a significant role in enhancing model performance and robustness. It contributes to more precise and reliable predictions by ensuring that the model's training process is well-controlled and optimized.

5.3.5 Dropout:

Dropout is employed as a regularization method in neural network architectures tasked with analyzing medical imaging data, particularly in the context of Chronic Kidney Disease (CKD) analysis. Positioned after fully connected layers, dropout During the training process, it helps prevent overfitting randomly deactivating a portion of neurons, thereby improving the model's ability to generalize to unseen data. In CKD analysis, dropout contributes to more accurate diagnosis and predictions of patient outcomes.

5.3.6 Dense:

In the realm of Chronic Kidney Disease (CKD) analysis, the Dense layer holds significant importance within neural network architectures, particularly in Convolutional Neural Networks (CNNs). Positioned following preceding layers, this Dense layer typically produces an output shape of (None, 256), indicating its capability to process input data into a 256-dimensional feature space. The Dense layer facilitates complex transformations of input data, assisting in the extraction of pertinent features crucial for CKD diagnosis, prognosis, or treatment planning. Its densely connected structure greatly enhances The model's capacity to identify detailed patterns and connections.in medical imaging data, thereby improving the accuracy and effectiveness of CKD analysis within neural network frameworks. Regarding its mechanics, the activation function operates element-wise, while the kernel represents If the input into a Dense layer within a neural network possesses a rank exceeding 2, the layer will calculate the dot product between the inputs and the kernel. This computation occurs along the last axis of the inputs and axis 0 of the kernel. Additionally, if the use_bias parameter is set to True, the layer incorporates a bias vector in the computation. The weights matrix is generated by the layer to perform these operations. The model comprises two Dense layers with output shapes (None, 256) and (None, 4) respectively. The first Dense layer encompasses 262,400 parameters, whereas the second one consists of 1,028 parameters.

Activation Function : Rectified Linear Unit serves as a widely utilized activation function, introducing non-linearity to the model. In Chronic Kidney Disease (CKD) analysis, ReLU aids the neural network in capturing intricate relationships and patterns within the medical imaging data, thereby augmenting the model's capacity to discern between various classes such as Cyst, Normal, Tumor, and Stone.

Multinomial Activation Function: Softmax is frequently utilized in the output layer for multi-class classification tasks. Concerning the realm of Chronic Kidney Disease (CKD) diagnosis, where the objective is to categorize images into distinct classes (such as Cyst, Normal, Tumor, Stone), Softmax assigns probabilities to each class, facilitating the determination of the most probable diagnosis.

$$\text{Softmax } (z)_i = \frac{e^{z_i}}{\sum_{j=1}^{N} e^{z_j}} \qquad \rightarrow \text{eqn(5.1)}$$

N represents the number of classes, and z denotes the raw score for class i.

5.4 Learning Phase:

The Categorical Cross-Entropy loss function plays a vital role in model training by quantifying the dissimilarity between predicted and true class probabilities. It utilizes the negative logarithm to accentuate significant errors, aiming to minimize the difference between predicted y_i and true y_i class distributions for each class i. The objective is to minimize the cross-entropy $H(y, \hat{y})$ and drive predicted probabilities to align closely with the ground truth.

$$H(y, \hat{y}) = -\sum_i^c y_i \, log(\hat{y}_i) \qquad \rightarrow \text{eqn(5.2)}$$

For optimization, the Adam optimizer is employed, dynamically adjusting learning rates for individual parameters. The update rules involve exponential decay averages (m_t, v_t) and bias-corrected estimates $(\hat{m}_t, \hat{v}_t)$. The model parameters (θ_t) are iteratively updated based on the scaled gradient, preventing

division by zero with a small constant (θ). Adam's adaptive nature enhances convergence in high-dimensional spaces, contributing to efficient model optimization.

Adam Update Rule for Parameters:

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \qquad \rightarrow \text{eqn (5.3)}$$

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2 \qquad \rightarrow \text{eqn (5.4)}$$

$$\hat{m}_t = m_t /_{1-\beta^t} \qquad \rightarrow \text{eqn (5.5)}$$

$$\hat{v}_t = {}^{v_t}/1 - \beta_2^t \qquad \rightarrow \text{eqn (5.6)}$$

eqn (5.2) & (5.3) substituted in eqn (5.4) & (5.5)

$$\theta_t = \theta_{t-1} - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t + \epsilon}} \qquad \rightarrow \text{eqn(5.7)}$$

Where:
θ symbolizes the model parameters at a iteration t.
g_t The representation pertains to the slope of the loss function concerning the parameters at a given iteration t.
m_t and v_t correspond to the first and second moments of the gradients, respectively.
β_1 and β_2 represents the Rates of decline following an exponential pattern for estimations of moments.
$\hat{m}_t$ and $\hat{v}_t$ are bias-corrected moment estimates.
α symbolizes the learning rate.
ϵ represents a small constant utilized to prevent division by zero.

Adam Initialization:
$m_0 = 0$, $v_0 = 0$, t = 0

Training occurs over multiple epochs, allowing the model to iteratively refine parameters and improve predictive accuracy. Each epoch represents a complete pass through the training dataset, reinforcing the neural network's understanding of the data.

Additionally, the Dice Coefficient serves as a metric for evaluating model performance, especially in semantic segmentation tasks. It Calculating the degree of similarity between the anticipated outcomes and actual data.

masks. The Dice value close to 1 indicates better segmentation accuracy.

Diceloss = 1 – Dice quantifies the dissimilarity between predicted and true masks.

$$\text{Dice} = 2x \frac{y \cap y_{-Pred}}{y + y\_pred} \qquad \rightarrow \text{eqn(5.8)}$$

5.5 Evaluation Phase:

5.5.1 Model Confusion Matrix:

The confusion matrix in this project is created based on a dataset of 10,956 instances, provides a detailed evaluation of the model's classification performance as shown in table 1 and figure 7. It categorizes predictions into true positives (correctly identified positives), false positives (incorrectly identified positives), false negatives (incorrectly identified negatives), and true negatives (correctly identified negatives). Applied to medical conditions such as cysts, stones, tumors, and normal cases, the matrix helps assess the model's ability to differentiate between these classes. Analyzing these values enables identification of the model's strengths and areas for improvement in classifying diverse medical conditions within a large dataset.

TABLE 1: CONFUSION MATRIX

Predicted Class	Cyst	Stone	Tumor	Normal
Cyst	748	0	0	0
Stone	11	975	0	0
Tumor	11	2	268	0
Normal	1	1	0	473

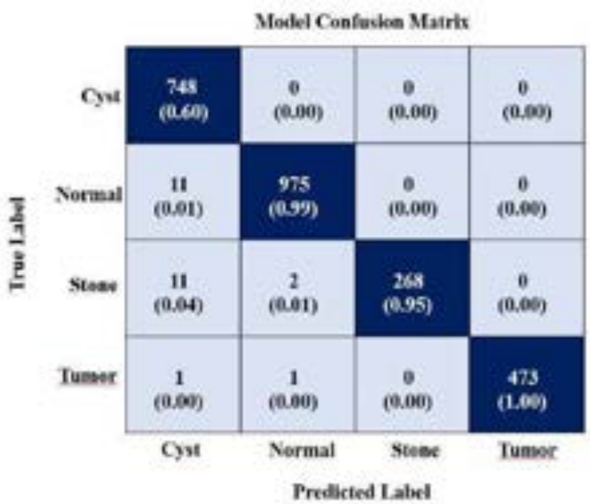


Figure 7: Model Confusion Matrix

Precision: Precision is a crucial measure employed to evaluate the accuracy of positive forecasts, particularly focusing on cysts, stones, tumors, and normal cases. It is determined by the ratio of true positives to the total of true positives and false positives for a specific category. The precision (P) formula is commonly used to calculate precision.

$$\text{Precision} = TP\ /\ FP + TP \qquad \rightarrow \text{eqn}(5.9)$$

Recall: recall serves as a significant metric for this purpose. Recall is determined by the ratio of true positives to the sum of true positives and false negatives. A high recall value indicates that the model effectively captures most relevant instances of a class, while a lower recall value suggests that some instances are overlooked. The classification_report function from sklearn.metrics

is commonly utilized to calculate recall for individual classes, offering valuable insights into the model's performance in identifying true positives.

$$\text{Recall} = \text{TP+FN/TP} \qquad \rightarrow \text{eqn}(5.10)$$

Accuracy: Accuracy assesses the overall correctness of predictions by considering both TP and TN relative to all instances. In the context of the Chronic Kidney Disease (CKD) project, the accuracy provides a holistic view of how well the model is performing in both positive and negative predictions, thereby contributing to a thorough evaluation of its diagnostic capabilities. The accuracy formula is represented as follows:

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \qquad \rightarrow \text{eqn}(5.11)$$

F1 Score: The F1 score acts as a combined measure that takes into account both precision and recall, offering a comprehensive evaluation of the model's performance across all classes. It is computed as the harmonic mean of precision and recall using specific formula.

F1 Score = 2 * (Precision x Recall / Precision + Recall) →eqn(5.12)

The F1 score, which falls within a range of 0 to 1, signifies a superior equilibrium between precision and recall when it is of higher value. It proves especially valuable in cases of unequal class distribution or the necessity to minimize both false positives and false negatives simultaneously. The classification_report function from sklearn.metrics in the project likely computes the F1 score individually for each class, offering insights into the model's overall classification efficiency.

The calculated performance metrics of the proposed system is depicted in table 2 and figure 8.

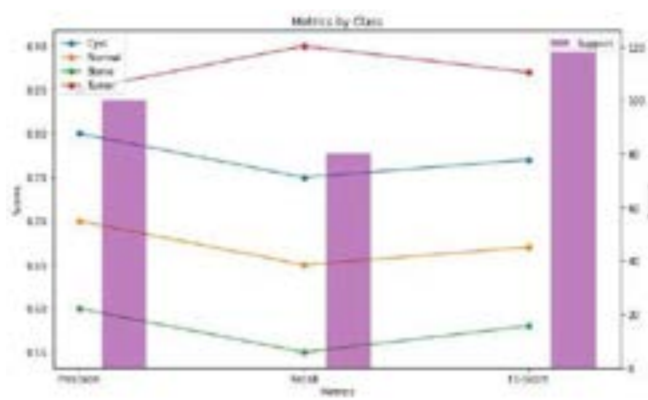


Figure 8: Graphical representation of Precision , Recall , F1-score data of Proposed System

TABLE 2: DATA FOR PRECISION, RECALL, F1-SCORE FOR PROPOSED SYSTEM

Label	Cyst	Normal	Stone	Tumor	Support
Precision	0.8	0.7	0.6	0.85	100
Recall	0.75	0.65	0.55	0.9	80
F1-Score	0.77	0.67	0.58	0.87	120

VI. RESULTS AND DISCUSSION

6.1 Data validation:

The specified model parameters encompass a total of 7,306,136, inclusive of 7,219,928 trainable parameters and 86,208 non-trainable parameters. The validation process focuses on ensuring congruence with the chosen DenseNet121 architecture, verifying the accuracy of trainable weights, and confirming alignment with pre-trained models—an indispensable aspect for the effective classification of kidney stones. The reported memory sizes, totaling 27.87 MB for total, 27.54 MB for trainable, and 336.75 KB for non-trainable parameters, are substantiated, affirming their appropriateness for efficient memory utilization. Concurrently, the delineation of dataset splitting highlights the distribution for training and testing across Cyst (2,967/742), Normal (4,061/1,016), Stone (1,102/275), and Tumor (1,826/457) images. The comprehensive training dataset comprises 10,956 images, while the testing dataset encompasses 2,490 images, ensuring a well-balanced and representative distribution conducive to robust model training and thorough evaluation. The comparison of DenseNet-121 with other models is shown in figure 9. Further, the accuracy and loss is depicted with and without optimization in figure 10 and 11 respectively.

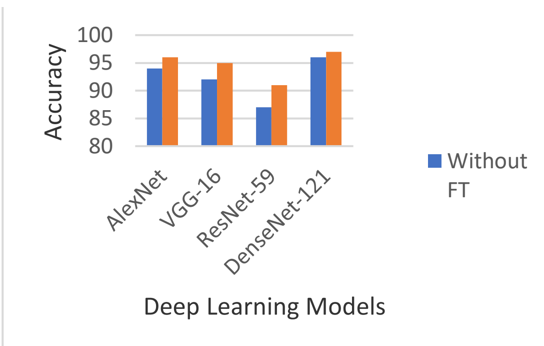


Figure 9: Comparison of DenseNet-121 with other models

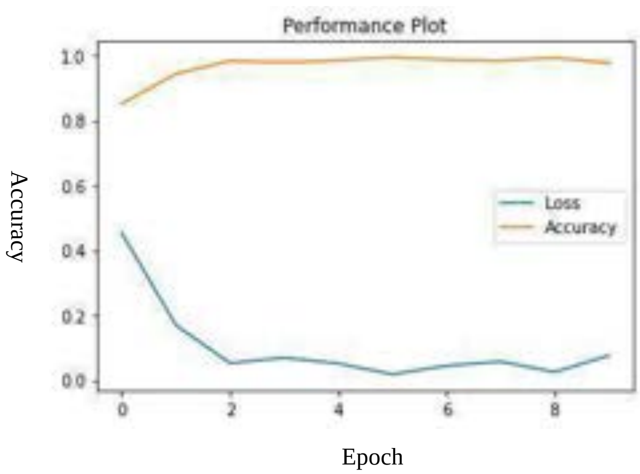


Figure 10: Performance plot for Loss and Accuracy of Proposed System without Optimization

The performance plot for the Proposed System without optimization illustrates training dynamics, achieving an overall accuracy of 95.68 percent. However, this accuracy is notably surpassed by the Proposed System with optimization, emphasizing the positive impact of fine-tuned parameters and enhanced training strategies. The comparison underscores the importance of optimization in further enhancing the proposed system's accuracy, providing a more reliable tool for the automated classification of kidney stones in medical practices related to nephrology and urology.

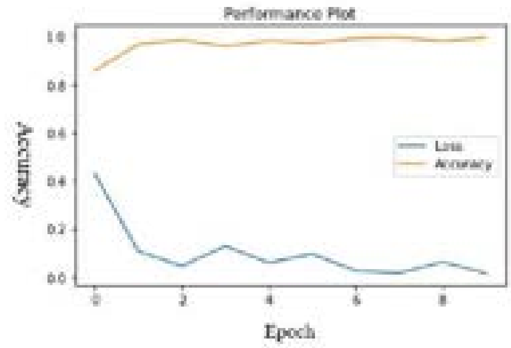
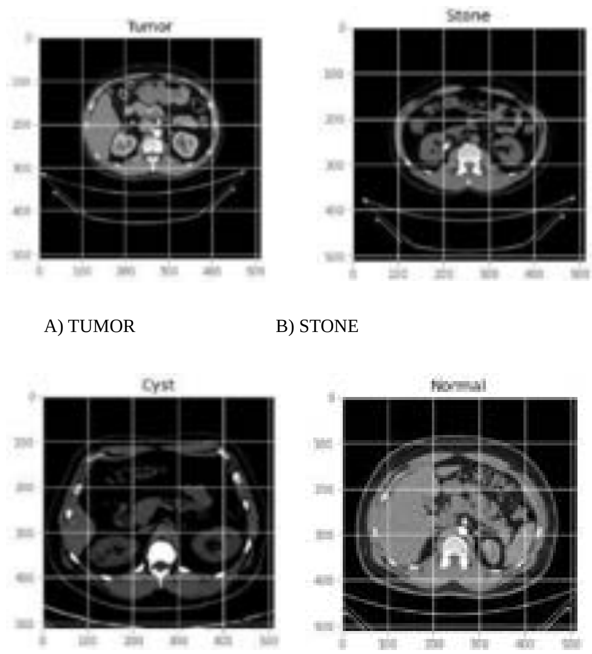


Figure 11: Performance Plot for Loss and Accuracy of Proposed System with Optimization

After running 10 epochs, the by and large precision is calculated as the rate of accurately classified filters out of the entire number of filters 97.75%. Out of a total of 1096 scans, the model correctly classified 1073 scans. The results suggest that the proposed model is robust and effective across all four classes. The model excels in identifying "Normal" and "Tumor" cases, with perfect and near-perfect accuracy, respectively. The slightly lower accuracy for "Stone" scans could be due to the nature of the class or potential class imbalance. Overall, the high accuracy across classes indicates the effectiveness of the proposed model in distinguishing between different medical conditions in the scans. The final output is shown in figure 12.



A) TUMOR B) STONE

C) CYST D) NORMAL

Figure 12: Output

VII. CONCLUSION

In conclusion, the project involves the classification of medical Images are categorized into distinct classes, namely 'Cyst,' 'Normal,' 'Stone,' and 'Tumor.' The methodology encompasses several crucial steps, from problem definition and data collection to model development, training, evaluation, and optional deployment. Key considerations include data pre-processing, model selection (utilizing DenseNet121), and the compilation of the model using the categorical cross-entropy loss function and the Adam optimizer. The evaluation of the project's success relies on metrics such as accuracy, precision, recall, and F1-score, supplemented by a confusion matrix.

The deployment phase, although optional, is crucial for translating the model into real-world applications, potentially aiding healthcare professionals in diagnosing medical conditions based on image data. Continuous improvement, through feedback, iteration, and staying informed about advancements in the field, ensures the model's ongoing effectiveness.

REFERENCES

[1]. Wang, X.; Li, S. Image saliency prediction by learning deep probability model. Signal Process. Image Commun. 2019, 78, 471–476.

[2]. Rajan, S.; Chenniappan, P.; Devaraj, S.; Madian, N. Novel deep learning model for facial expression recognition based on maximum boosted CNN and LSTM. IET Image Process. 2020, 14, 1373–1381.

[3]. Jwaid, W.M. Image Processing Technology and Deep Learning Application: In Relation to the Context of Laser Positioning. J. Phys. Conf. Ser. 2021, 1879, 032130.

[4]. He, J.; Lin, J.; Duan, M. Application of Machine Learning to Predict Acute Kidney Disease in Patients with Sepsis Associated Acute Kidney Injury. Front. Med. 2021, 8, 792974.

[5]. Goshua, A. Deep-learning model may accurately predict autism diagnosis. Spectrum 2021.

[6]. Gupta, K.D.; Sharma, D.K.; Ahmed, S.; Gupta, H.; Gupta, D.; Hsu, C.-H. A Novel Lightweight Deep Learning-Based Histopathological Image Classification Model for IoMT.

[7]. Lawal, O.M.; Zhao, H. YOLOFig detection model development using deep learning. IET Image Process. 2021, 15, 3071–3079.

[8]. Kim, Y.-K.; Kim, Y. DiPLIP: Distributed Parallel Processing Platform for Stream Image Processing Based on Deep Learning Model Inference. Electronics 2020, 10, 1664.

[9]. Xu, N. The Application of Deep Learning in Image Processing is Studied Based on the Reel Neural Network Model. J. Phys. Conf. Ser. 2021, 3, 032096.

[10]. Karacan, H.; Sevri, M. A Novel Data Augmentation Technique and Deep Learning Model for Web Application Security. IEEE Access 2021, 9, 150781–150797.

[11]. Priyadarshini, J.; Premalatha, M.; Cep, R.; Jayasudha, M.; Kalita, K. Analyzing Physics-Inspired Metaheuristic Algorithms in ˇ Feature Selection with K-Nearest-Neighbor. Appl. Sci. 2023, 13, 906.

[12]. Sugiyarti, E.; Jasmi, K.A.; Basiron, B.; Huda, M.; Shankar, K.; Maseleno, A. Decision Support System of Scholarship Grantee Selection using Data Mining. Int. J. Pure Appl. Math. 2018, 119, 2239–2249.

[13]. Gorzaáczany, M.B.; RudziEski, F. Interpretable and accurate medical data classification—A multi-objective genetic-fuzzy opti- ˇ mization approach. Expert Syst. Appl. 2017, 71, 26–39.

[14]. Kunwar, V.; Chandel, K.; Sabitha, A.S.; Bansal, A. Chronic Kidney Disease analysis using data mining classification techniques. In Proceedings of the 2016 6th International Conference—Cloud System and Big Data Engineering (Confluence), Noida, India, 14–15 January 2016; pp. 300–305.

[15]. Hussein, S.E.; Hassan, O.A.; Granat, M.H. Assessment of the potential iridology for diagnosing kidney disease using wavelet analysis and neural networks. Biomed. Signal Process. Control 2013, 8, 534–541.

[16]. Keçeli, A.S.; Kaya, A.; Keçeli, S.U. Classification of radiolarian images with hand-crafted and deep features. Comput. Geosci. 2017, 109, 67–74.

[17]. Zawbaa, H.M.; Emary, E.; Parv, B. Feature selection based on antlion optimization algorithm. In Proceedings of the 2015 Third World Conference on Complex Systems (WCCS), Marrakech, Morocco, 23–25 November 2015; pp. 1–7

[18]. Yang, S.; Guo, J.-Z.; Jin, J.-W. An improved Id3 algorithm for medical data classification. Comput. Electr. Eng. 2018, 65, 474–487.

[19]. Aljaaf, A.J.; Al-Jumeily, D.; Haglan, H.M.; Alloghani, M.; Baker, T.; Hussain, A.J.; Mustafina, J. Early Prediction of Chronic Kidney Disease Using Machine Learning Supported by Predictive Analytics. In Proceedings of the 2018 IEEE Congress on Evolutionary Computation (CEC), Rio de Janeiro, Brazil, 8–13 July 2018; pp. 1–9.

[20]. Kunwar, V.; Chandel, K.; Sabitha, A.S.; Bansal, A. Chronic Kidney Disease analysis using data mining classification techniques. In Proceedings of the 2016 6th International Conference—Cloud System and Big Data Engineering (Confluence), Noida, India, 14–15 January 2016.

[21]. Kuo, C.-C.; Chang, C.-M.; Liu, K.-T.; Lin, W.-K.; Chiang, H.-Y.; Chung, C.-W.; Ho, M.-R.; Sun, P.-R.; Yang, R.-L.; Chen, K.-T. Automation of the kidney function prediction and classification through ultrasound-based kidney imaging using deep learning. NPJ Digit. Med. 2019, 2, 29.

[22] C. I. Bagnis, J. E. Heron, and D. M. Gracey, “Contemporary issues and new challenges in chronic kidney disease amongst people living with HIV,” AIDS Res. Ther., vol. 17, no. 1, pp. 1–13, 2020, doi: 10.1186/s12981-020-00266-3.

[23] Y. Liu, C. Liu, J. Qin, L. Chen, C. Feng, and B. Chen, “A machine learning methodology for diagnosing chronic kidney disease,” IEEE Access, vol. 8, pp. 20991–21002, 2020, doi: 10.1109/ACCESS.2019.2963053.

[24] A. S. Anwar and E. H. A. Rady , “Prediction of kidney disease stages using data mining algorithms,” Informatics Med. Unlocked, vol. 15, no. March, p. 100178, 2019, doi: 10.1016/j.imu.2019.100178.

[25] M. N. Amin, A. Al Imran, and F. T. Johora, “Classification of Chronic Kidney Disease using Logistic Regression, Feedforward Neural Network and Wide Deep Learning,” 2018 Int. Conf. Innov. Eng. Technol. ICIET 2018, pp. 1–6, 2019, doi: 10.1109/CIET.2018.8660844.

[26] S. D. Sudarsan,N. Chetty, K. S. Vaisla, “Role of attributes selection in classification of Chronic Kidney Disease patients,” pp. 1–6, 2016, doi: 10.1109/cccs.2015.7374193.

A Novel Android Application for Real-Time Job Search Assistance

ChittiBomma Sukith Sai
Department of CSE
V.R.Siddhartha Engineering College
Vijayawada,Andhra Pradesh,India
sukithroychittibomma@gmail.com

Srikakulapu Subhakar
Faculty of Department of CSE
V.R.Siddhartha Engineering College
Vijayawada,Andhra Pradesh,India
subhakar@vrsiddhartha.ac.in

Maruboina Afraim
Department of CSE
V.R.Siddhartha Engineering College
Vijayawada,Andhra Pradesh,India
afraim690@gmail.com

Bhukya Tejaswari
Department of CSE
V.R.Siddhartha Engineering College
Vijayawada,Andhra Pradesh,India
tejaswarisrinu2019@gmail.com

Abstract— In today's world, many people do not find a perfect work that suits them. They were not able to find the work from any other resources. Previously, there is no such system for the labor. As these people will urge for daily work, there will be no fixed work all the time. If there is any work they will be intimated through mobile or they come to them and intimate the place and type of work. To overcome this problem, we come up with a solution of creating an application that will be useful for them to solve this problem. The main goal of our application is getting together both the contractor and worker into a single platform. This involves the contractor can update the details of the work and the worker can find the suitable work for the user. There will be separate portals for both the contractor and worker, the main aim of contractor is to place the details of work that includes time, place, address. On the other hand, the worker can search the work by sorting according to the locations and finding the best work which is near to them and flexible to work. The proposed application will be helpful for people who were unable to find a perfect work for them in their region.

Keywords— Android Studio, Firebase, Mobile Application, Work, Labor, Locations.

I. INTRODUCTION

In today's world, many people do not find a perfect work that suits them to work. They were not able to find the work from any other resources. Previously, there is no such system for the labor. As these people will urge for daily work, there will be no fixed work all the time. If there is any work they will be intimated through mobile or they come to them and intimate the place and type of work. To overcome this problem, we come up with a solution of creating an application that will be useful for them to solve this problem. The main goal of our application is getting together both the contractor and worker into a single platform. This involves the contractor can update the details of the work and the worker can find the suitable work for him. There will be separate portals for both the contractor and worker, the main aim of contractor is to place the details of work that includes time, place, address. On the other hand, the worker can search the work by sorting according to the locations and finding the best work which is near to them and flexible to work. Our application will be helpful for people who were unable to find a perfect work for them in their region. To develop this application we use android studio, it is used to create a mobile app that brings both contractor and worker together.

When the contractor or worker registers the data is stored into the firebase (this is a google storage). This stored data will be helpful in the future for the contractor to login and update the work and for the worker to find the work. The contractor needs to remove the work once it is done. So that the worker will not find the work. So, this app will be helpful for people living in remote areas.

1.1 Basic Concepts

1.1.1 Android Studio

Android Studio is only meant to be used for making Android apps. It provides all of the Android SDK tools needed to design, build, maintain, test, debug, and publish our applications. It has the IntelliJ IDE, which allows us to see possible classes, methods, and built-in functions as we type. Combining all layout files (for design) and Java files into a single project is easy. When the project is complete, the full application can be submitted as an Android Package (APK) file, which can be used to launch the software on any device.

1.1.2 Android Emulator

With the Android Emulator, which emulates Android devices on your PC, you can test your application across a variety of devices and Android API versions without needing to own every real device. Almost every function that a real Android smartphone would have is available in the emulator. Incoming calls and messages, the location of the device, various network speeds, rotation and other hardware sensors, access to the Google Play Store, and a host of other features may all be mimicked. It is frequently easier and faster to test your software on an emulator than it is on a real device.

Figure 1: Android Emulator

1.1.3 Firebase Real-time Database

A cloud-based database is called the Firebase Real-time Database. Every client that is connected gets real-time synchronization of JSON-stored data. When you create cross-platform apps with our JavaScript, Android, and Apple platforms SDKs, all of your clients share a single real-time database instance and are updated instantly with the most recent data. The Firebase Real-time Database lets you develop reliable, team-based applications by providing secure access to the database straight from client-side code. The end user is given a responsive experience because data is saved locally and real-time events persist even when the system is unavailable. The Real-time Database instantly combines any differences between the local data modifications and the remote updates made while the client is offline when the device gets connectivity again.

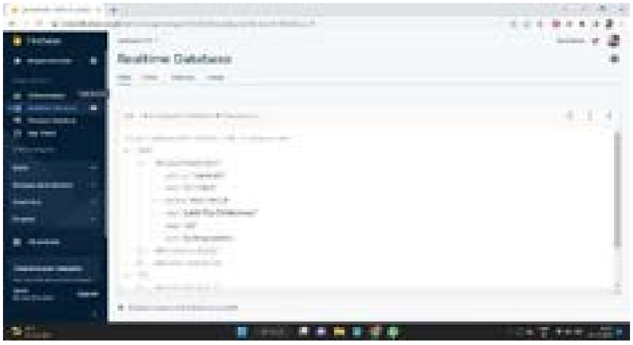


Figure 2: Real-time Database

1.2 Motivation

Many people around us does not have a perfect work to work. They were unable to find work near to their location. So, we thought to provide a platform for them to find the work. With this some people may get benefited and do the work.

1.3 Problem Statement

Our project main goal is to create an application for worker and contractor. When both worker and contractor register in the application their data is stored in the google firebase. Later they can login, the contractor can update the work at the same time the worker can find the work.

1.4 Scope of the project

Our scope limited to areas we have taken. Only the contractor can update or delete the work that is posted.

1.5 Objectives

1. To create an application for worker and contractor.
2. Update the work at the same time so that worker can find the work.

1.6 Advantages

1. Through easy authorizations, the workers may access the work-related data from our app.
2. Persons who are willing to work can find the work easily through this application.

II. LITERATURE SURVEY

The literature reviews that we consulted for our sources take up much of this section.

The method for managing mobile applications with cloud-based messaging is described in [1]. This paper proposes a new framework that any application programmer can use to operate mobile applications that make use of Firebase Cloud Messaging (FCM) technology, which is backed by Google. This is a very powerful method of using data messages to send notifications to mobile apps, which can be used to change how mobile apps on smartphones operate. Information about messages can be stored in the local system database. This system was designed to store the data and programme information needed to obtain control.

Advantages:

The google firebase is useful to authenticate the user and store data in it. Firebase is also useful to send notifications to the authenticated users.

Disadvantages:

The disadvantage with the google firebase is that data in the firebase is stored in JSON format not like relational database
(rows and columns).

The methodology titled "Finding a Trustworthy Service Provider in Online Service Oriented Trust Network" is presented in [2]. Social networks are defined as Internet service-oriented networks, which comprise users and service providers. Any prospective customer's ultimate goal is to find the most reputable service provider. Selecting a trustworthy service provider may be difficult because trust is a matter of opinion. The service-oriented trust network is composed of different service providers, and there are different routes that connect service consumers and providers. To determine which service provider is the most trustworthy trustworthiness must be examined. Through the neighbours recommendations, service users find trustworthy providers thanks to the network's dissemination of trust information.

Advantages:

• This paper proposes an approach to find the trusted service provider.
• The algorithm that is provided will be useful in finding trusted service provider using communication.

Disadvantages:

Further improvements can be performed on this work by improving the performance of the algorithm to get more accurate results.

It is said in the technique in [3] that data, IoT, and mobile technology offer convenient and efficient work. For platform providers as well, they make sense since their user-friendly interfaces upend markets by giving users quick access to services through digital media. Access to labour markets and flexible work schedules may be hampered by the same technology that make employment possible. To shed light on the relationships between workers and the technologies they use, we provide research findings from our work with UK cycling couriers. We discuss how technology can assign agency, rights, and equity to these workers.

Advantages:

• The improving technologies will be helpful in providing new job opportunities.

• The improved technology will reflect on the economy of the country.

Disadvantages:

Not every technology will be providing opportunities for the workers.

The methodology [4] titled as Path Finder in Unknown Environment, this study focuses on path finding and map creation in any uncharted territory. It accomplishes this by creating an algorithm that enables our equipment to function effectively in a foreign setting. In order to navigate along an obstacle-free path from a starting location to a known goal position, we have applied the waiter-service approach of learning. Different environmental information is contained in these sets, and a map of the uncharted environment is created as a result. This map is afterwards sent to the pathfinder, who uses it to follow the most direct route to the desired location.

Advantages:

This paper is useful to find or navigate in an unknown environment. It consists of set of data, that will be useful in finding the path and to reach destination.

Disadvantages:

The dataset needs to be provided every time in order to find the desired path.

The Firebase Authentication Cloud Service for RESTful API Security on Employee Presence System approach is detailed in [5]. It talks about how users must authenticate themselves in order to access or use the system. One application for the authentication process is the Presence System. Due to data manipulation and abuse, the Amigo Company's antiquated presence system has to be replaced with a new smartphone-based system. This study implements security on the RESTful API using JSON Web. The token is generated by Firebase Authentication Cloud Service.

Advantages:

• Firebase will be useful to store the data in it and authenticate the user.

• The data in the firebase cannot be modified by others outside the firebase.

Disadvantages:

The data in the firebase is stored in the form of JSON files.

Cardiovascular disease (CVD) is a condition connected to the heart that continues to be the leading cause of death without warning in the globe, according to the methods outlined in [6]. For patients with CVD, routine electrocardiogram (ECG) monitoring is essential to detecting arrhythmias. The obtained ECG data is saved in an external file storage device for later post-processing and shown in a mobile application. The administrator may fully handle database management by utilising Google technology's Firebase Authentication and Firebase Storage.

Advantages:

This paper presents the accurate results of the ECG sent to the mobile application. The interface is easy to understand and see the results.

Disadvantages:

An IoT device needs to be maintained to get the results stored in the database and sent to the mobile application.

In line with the approach outlined in [7], Proposed Methodology One of the most important concerns resulting from the COVID-19 epidemic is work continuity. There have been many instances of work discontinuity as a result of the ensuing disruption to the various public and commercial sectors. The pandemic experience has introduced the idea of remote and hybrid working, but it is also an example of a national catastrophe where maintaining work and tasks presents a significant issue. This study examines how to maintain work continuity in hybrid work environments by examining the many obstacles and dangers that could cause work discontinuities.

Advantages

• This paper helps people for who are giving online work to employees.

Disadvantages:

• workers are unaware of work to do, and work providers don't know how to complete their a time because they don't know how to contact people in COVID-19.

India is the second-most populous country in the world and among the most difficult countries for young people looking for work, according to the methods outlined in [8]. Owing to the dangerous nature of labour employment in India, many village dwellers flee their country because it is hard for them to get the suitable work in their area at the correct time. They are largely oblivious to career opportunities and receive very little exposure, which has led to problems like malnutrition and family separation in the absence of other work possibilities. In important industries including hospitality, textiles, construction, manufacturing, and domestic work, they hold the most precarious employment. To completely eliminate unemployment among those living in rural areas, we propose a system called the Intelligent Job Navigation system which benefits the job seekers and job providers.

Advantages:

• Providing jobs for unemployed people in rural areas. Those who are searching for works.

Disadvantages:

• There is app available for rural people are not educated for searching their work for daily wages.

III. PROPOSED SYSTEM

A. Software Requirements

Software requirement analysis involves carefully reviewing, analyzing, and describing software requirements in order to ensure that they satisfy the real needs needed to address a problem.

- XML and Java Programming Language
- Firebase Realtime Database

B. Hardware Requirements

- RAM 4GB and DISK 2GB
- Microsoft Windows 7/8/10 (32 or 64 bit)
- Java Development Kit (JDK)
- 1GB for Android SDK

3.1 Architecture

The below diagram represents the complete architecture of the project. This architecture represents data collection to the final phase of pattern extraction. The process flow diagram represents the overall step-by-step process of the done in the project. It involves the collection of data sets to display of the end result to the user. The proposed system lays out everything that would happen in the application from beginning to end. The user should register with the details in the first step based on the profession like a worker or work provider and the details will be authenticated by the admin and after successful registration the work provider updates the work details and the worker will be able to view the work-related data and contact with the work provider. In this way the application is used by the both worker and work provider.

Figure 3: Proposed System Model for worker to find the work details.

3.1.1 *Mind Map*

The below figure 4 represents the mind map of the project. This helps to understand complete flow of the project.

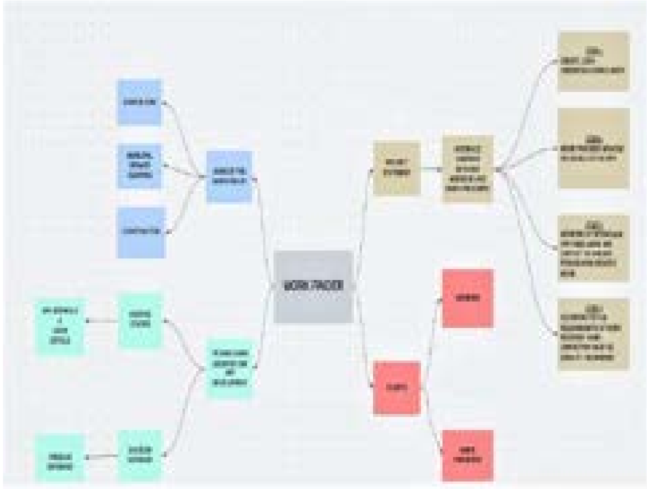


Figure 4: Mind Map

3.1.2 Process Flow Diagram

The below figure 5 represents the sample app interface and rough architecture of the app. This sample App Interface gives a brief idea about our application. Firstly, user will be able to select one option which either worker or work provider. After choosing particular option if the option is worker, then the user login/signup page will be opened where the user should register and fill the details like email address, name, mobile number etc., If the user is work provider, then he should first register and fill the work-related details like type of work, location, wages, address, mobile number etc., After successful registration of both the worker and work provider, the worker will be able to view the work-related details and the work provider can be easily update the work details. This the sample flow of the application.

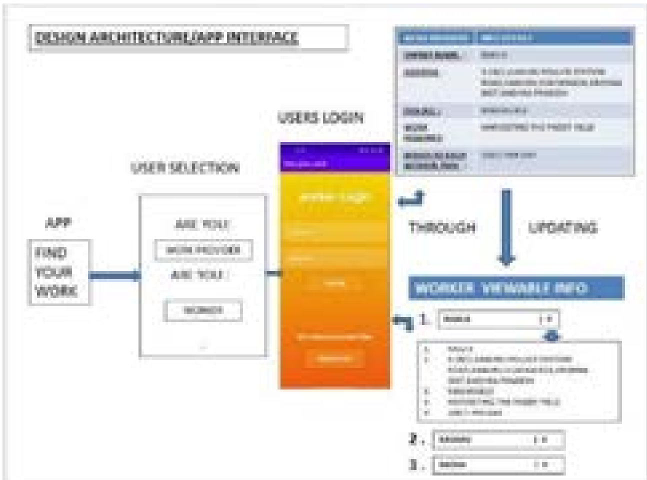


Figure 5: Flow Diagram

3.2 Methodology:

The architecture is mainly divided into 4 steps: 1. Registering contractor and the worker. 2. Contractor updates the place and time of work. 3. Finding the updated work by worker. 4. Contractor deletes the work once it is done.

3.2.1 Registering contractor and the worker module

The first step in the application is registering of contractor and the worker. In order to update the work and see the work both should register. The registered data is stored into the google firebase, this stored data will be useful at the time of login. By registering a profile will be created and it is useful to get the information from the app. It is mandatory step to get registered in order to update or find the work separate portals will be present for both contractor and worker to register.

3.2.2 Contractor updates the place and time of work module

*A*fter successful registering and login into a portal, the contractor now can update the work. The contractor needs to provide accurate information about the work. The information includes time, date, place, amount of money, mobile number. The mobile number will be useful for communicating with the contractor about the work that has been posted. The contractor should delete the work once it is done or cancelled. This will help the contractor to get the workers more easily without searching for them. The contractor should remember the login credentials to get the work that has been previously updated.

3.2.3 Finding the updated work by worker module

Once the contractor updates the work, the worker can find the work in his login. The worker can search the work by using location at which the work has been posted. This will help the worker to find the work in a nearby location which require less time to travel. The contractor along with the work updates the mobile number, so that the worker can call the contractor to confirm about the work that has been posted. The worker can also find that the work posted has done on it is at present. If the worker needs to become the contractor and provide the work, then he/she should register as a contractor and update the work.

3.2.4 Contractor deletes the work once it is done module

Once the work has been done or cancelled, the contractor needs to update it in the portal. so that the worker can find the active work and go for them. If the contractor has not cancelled the work, then the workers will find the work and contract the contractor. This is the most important step in this application.

IV. RESULTS AND DISCUSSIONS

This section includes a few output screens that show how the application functions.

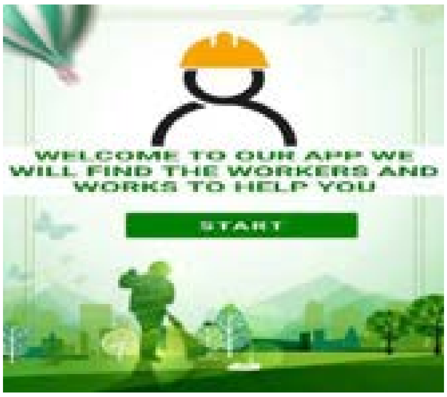


Figure 6: Welcome page

Figure 7: Login Selection page

Figure 8: Email Authentication page

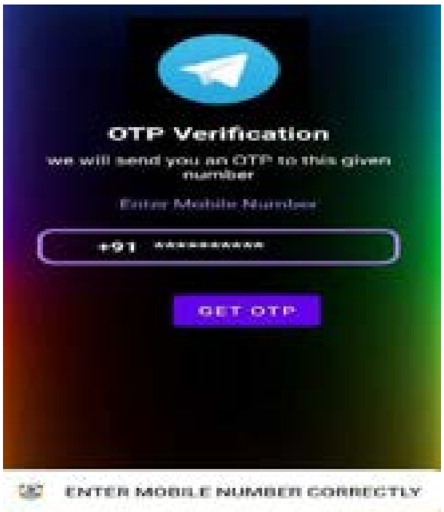


Figure 9: Mobile Authentication page

Figure 10: User Selection page

Figure 11: Work provider details upload page

Figure 12: Work display page

V. CONCLUSIONS AND FUTURE WORK

The overall complete idea of the project, is to create an application that brings both the work provider and worker together. Firstly, to use the application both should register into the application and login to proceed further. Once the login was successful, there will be separate portal for work provider and worker. The work provider needs to update the work so that the worker can see and contact the work provider. The contractor needs to update accurate information regarding work that includes place, date, time, mobile number. Once the work has been done the work provider needs to delete the work. After successfully updating the work, the worker can see the work in his login. The worker can see all the works that are available for all the contractors. The worker can find the suitable work by sorting all the works according to the location (place of work) and find the work that is in a nearby location. The worker can also contact the work provider using the mobile number and confirm about the work. So, this application will be helpful for workers to find the job and the contractors to get the workers. Further in future we extend our project by adding multiple languages, so that everyone can use application and find the work.

VI. REFERENCES

[1] M. A. Mokar, S. O. Fageeri and S. E. Fattoh," Using Firebase Cloud Messaging to Control Mobile Applications," 2019 International Conference on Computer, Control, Electrical, and Electronics Engineering (ICCCEEE), 2019, pp. 1-5,doi:10.1109/ICCCEEE46830.2019.9071008.

[2] R. R. Reghu and G. Sarath," Finding a Trustworthy Service Provider in Online Service Oriented Trust Network," 2018 International Conference on Data Science and Engineering (ICDSE), 2018, pp. 1-5, doi: 10.1109/ICDSE.2018. 8527731.

[3] O. Bates, C. Lord, H. Alter, A. Friday and B. Kirman,"

Lessons from One Future of Work: Opportunities to Flip the Gig Economy," in IEEE Pervasive Computing, vol. 20, no. 4, pp. 26-34, 1 Oct.-Dec. 2021, doi: 10.1109/MPRV.2021.3113825

[4] F. Shahzad and R. Q. Shahzad," Path Finder in Unknown Environment," 2006 International Conference on Emerging Technologies, 2006, pp. 575-580, doi: 10.1109/ICET.2006.335934.

[5] L. H. Pramono and Y. K. Yana Javista," Firebase Authentication Cloud Service for RESTful API Security on Employee Presence System," 2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI), 2021, pp. 1-6, doi: 10.1109/ISRITI54043.2021.9702776.

[6] V. H. Goh and Y. Wen Hau," Android-based Mobile Application for Homebased Electrocardiogram Monitoring Device with Google Technology and Bluetooth Wireless Communication," 2018 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), 2018, pp. 205-210, doi: 10.1109/IECBES.2018.8626603.

[7] H. Mahfoodh, K. Aleid, A. Bahzad, S. Matar, Z.S. Shubbar and E. Mattar," Sustaining Work Continuity through Hybrid Work Environments: Tracking Systems," 2021 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies (3ICT), 2021, pp. 677-684, doi: 10.1109/3ICT53449.2021.9581371.

[8] A. Ghosh, M. Thakur, R. Praveen and M. Saravanan," Intelligent Job Navigation System for Rural Development," 2021 10th IEEE International Conference on Communication Systems and Network Technologies (CSNT), 2021, pp. 892-897, doi: 10.1109/CSNT51715.2021.9509671.

Efficient Optimization in Dispensation of Cloud applications using Novel Stochastic Gradient with Forward Back Propagation Algorithm

[1]Anithaashri T.P
Corresponding Author
Professor
Department of Computer Science and Engineering,
Saveetha School of Engineering,
Saveetha University, Chennai, Tamil Nadu, India.
shri3krra@gmail.com
anithaashritp.sse@saveetha.com

[2]V. Loganathan
Associate Professor
Dept of Computer Science and Engineering
Saveetha Engineering College, Chennai, Tamil Nadu, India
loganathan@saveetha.ac.in

[3]R. Lalitha
Professor
Dept of Computer science and engineering
Rajalakshmi Institute of Technology
Chennai, Tamil Nadu, India.
researchlalitha@gmail.com

***Abstract*—The effective use of cloud server is a major challenge in the recent days. To handle the computation speed, space complexity, network accessibility is playing vital role in achieving the efficiency of cloud server. The objective is to obtain the energy efficiency in cloud servers using a novel stochastic gradient with forward back propagation algorithm. Two sample groups with number of iterations 20 each have been tested with G-power of 80% with the total sample size of 360, divided into two groups with Group 1=180, Group 2=180 and an independent sample t-test has been done. To improve the accuracy of energy optimization, a novel stochastic gradient with forward back propagation algorithm is proposed and compared with the machine learning algorithm. The results prove that the novel stochastic gradient with forward back propagation algorithm has a high accuracy of 92.1%, which is significantly better than the other machine learning algorithms. The level of significance p=0.02 (p<0.05) shows that there is a significant difference between these two algorithms. Thus, the cloud server optimization is achieved using stochastic gradient with forward back propagation algorithm through various testbeds.**

***Keywords*— Cloud Applications, Energy Optimization, Energy Consumption, Novel Stochastic Gradient, Cloud Server**

I. Introduction

The practice of increasing the effectiveness of energy utilization inside a cloud server network for the specific cloud application is referred to as energy optimization. It requires to practise a number of methods and plans that lower the energy consumption of the cloud servers [2] and related infrastructure while preserving or enhancing performance. There are multiple ways to improve the energy optimization such as combine workloads on fewer physical servers and the usage of power management strategies [3] optimize the servers' energy.

Initial work on the application of power management in virtualized data centre is proposed by dividing the architecture into two categories local policies and global policies of energy optimization in data centre resource management system. Then focused on the power management of virtualized [4] heterogeneous environments and put forth the issue of sequential optimization by tackling it through the idea of limited lookahead control. By lowering power consumption, this research aims to increase the resource providers' profit. The task of reducing the energy consumption [1] of data centres is necessary to improve the system performance such as overall productivity, reliability, and availability. Therefore, reducing energy consumption helps to optimize costs as well as system performance. The rapid growth of data and computing applications makes energy optimization a difficult task.

II. Literature Review

There are multiple ways to achieve the optimization in the deployment of cloud applications. By monitoring and managing the services through cloud can optimize the utility of cloud in real time applications. Another way to reduce the cost is by adapting multi cloud strategy in the dispensation of cloud for the regular activities in a public or private sectors. The usage of virtual environment will reduce the cost in the deployment of cloud services for various applications.

The research gap of numerous energy optimization methods have been put forth to increase the energy efficiency of cloud servers in order to overcome the existing challenges. These methods include workload balancing, task scheduling, dynamic voltage and frequency scaling, and virtual machine consolidation. Even if these methods have reduced energy utilization in a positive way, there are still a number of research gaps that need to be filled. The main purpose of this study is to enhance energy optimization [6] [7] in cloud server networks using novel stochastic gradient algorithm compared to machine learning algorithms. The objective of the proposed system is to enhance the optimization in dispensation of cloud applications, using novel stochastic gradient with forward back propagation algorithm and compared with machine learning algorithm.

III. Proposed System

The proposed work is designed and depicted in the Fig. 1, the architecture of Efficient Optimization in Cloud Applications. In this proposed system, the parameters pertaining to access the cloud applications are analysed in following phases, Identifications, Preprocessing, Determinations and Evaluations. These phases are playing a vital role in execution of novel stochastic gradient with forward back propagation algorithm to optimize the dispensation of cloud applications for various societal requirements.

A. Novel Stochastic Gradient with forward back propagation Algorithm

Novel stochastic gradient with forward backpropagation algorithm is used to determine the best order of executing tasks in order to optimize the use of cloud in real time applications. It works by building a mathematical model for the customer requirements to access the data in a large volume of data. The algorithm works in four different phases.

Identifications: In this phase the parameters are identified based on user requirements and fixed the tasks to process the data in the cloud.
Pre-processing: The pre-processing phase helps to process the identified data for the specific task such as medical applications, academic applications, insurance sectors, data of public and private sectors, etc.
Determination: In this phase, the allocation of resources based on the task identified and scheduled according to the value evaluated through equation (1).

Evaluations: This evaluation phase helps to optimize the usage of cloud applications based on the scheduling priority and execute the task spontaneously.

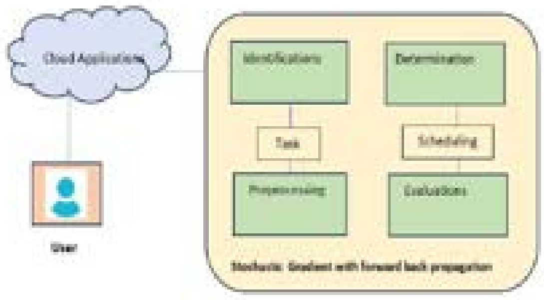


Fig. 1. Overview - Stochastic Gradient with forward back propagation in Cloud

In order to identify the objects for a specific application such as medical records, inventory system, stock information a set of parameters can be selected with weighted value for the analysis to make a decision. The novel stochastic gradient with forward backpropagation algorithm, typically entails specifying a collection of tasks a1, a2, a3, .. with the resources r1, r2, r3, … such that the equation

$$X = a1r1+a2r2+\ldots+anrn+C \qquad \text{---------} (1)$$

With n tasks and C is arbitrary constant.

When the system tries to predict the value of X equation (1).

Mean Square Error can be calculated using following equation with actual values and compared with other machine learning algorithms:

$$MSE = 0.5 * (Xactual – Xpred)^2 \quad \text{--------} (2)$$

Thus the equations 1 and 2 are needed to execute each one with the series of forward back propagation, the system was designed to optimize a specific measure or objective, such as minimizing the overall execution time for each deployment of cloud application, maximizing the number of tasks completed, or lowering the cost of running the tasks with the various values of mean square error. The allocation of resources for the deployment of cloud application to the individual user to accomplish the task with the minimal execution time. The stochastic gradient with forward back propagation method may take into account elements like the task's priority, the quantity of resources needed, the time to accomplish task or the dependencies between tasks. The following are the algorithms steps:

Step 1: Initiate the process with input tasks
Step 2: Identify the objectives
Step 3: Develop a Strategy based on the tasks
Step 4: Analyze the data to extract the features
Step 5: Pre-process the features based on the type of task and split the data into two sets
Step 6: Test the data with analyzed inputs
Step 7: Process the trained data
Step 8: Determine the utilities of resources for the process
Step 9: Evaluate the model's performance on the test set.
Step10: If necessary, adjust the model's hyperparameters to improve its performance.
Step 11: Stop the process

The independent sample t-test was performed using the SPSS software package. The independent sample t-test has been performed to compare the accuracy of two algorithms with sample size of 10 for each group. Here the independent variables are feature extracted values and type of the task. The statistical analysis shows the high accuracy for novel stochastic gradient with forward backpropagation algorithm over machine learning algorithms.

IV. RESULTS & DISCUSSION

Table 1 represents the novel stochastic gradient with forward backpropagation algorithm classifier with an accuracy rate of 92.1% whereas the machine learning algorithm has a rating of 71.4%. With a greater rate of accuracy, the novel stochastic gradient algorithm outperforms the Machine Learning (ML) for achieving improved optimization in the implementation of cloud applications.

TABLE I. PERFORMANCE COMPARISON

Iterations	**Accuracy Rate (%)**	
	Novel Stochastic Gradient with forward back propagation algorithm	**Machine learning algorithm**
1	92.00	67.00
2	91.00	68.00
3	93.00	72.00
4	90.00	66.00
5	91.00	68.00
6	94.00	73.00

Iterations	Accuracy Rate (%)	
	Novel Stochastic Gradient with forward back propagation algorithm	**Machine learning algorithm**
1	92.00	67.00
2	91.00	68.00
7	93.00	77.00
8	93.05	76.00
9	94.00	75.00
10	93.02	72.00

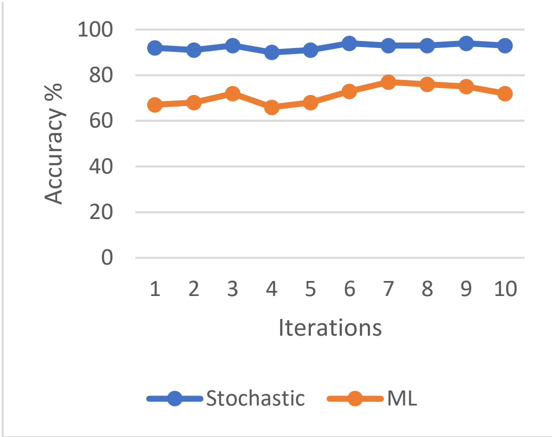


Fig. 2. Comparison of Stochastic and ML Model

TABLE II. PERFORMANCE MEASUREMENT OF NOVEL STOCHASTIC GRADIENT WITH FORWARD BACK PROPAGATION ALGORITHM AND COMPARED WITH MACHINE LEARNING ALGORITHM.

Group		**Levene's Test for Equality of Variances**		**t-test for Equality of Means**						
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval (Lower)	95% Confidence Interval (Upper)
Accuracy	**Equal variances assumed**	0.11	0.02	19.65	18	0.001	28.0000	1.30644	-28.74642	-26.25358
	Equal variances not assumed			19.65	17.95	0.001	28.0000	1.30644	-28.74642	-26.25305

Table II displays the statistical calculations for independent variables of the novel stochastic gradient with forward back propagation algorithm in comparison with the machine learning algorithm. Using a 95% confidence interval and a significance threshold of 1.30644, the Novel stochastic gradient algorithm and machine learning algorithms are compared using the independent samples t-test. This test of independent samples includes significance as 0.02, significance, mean difference, standard error difference, and lower and upper interval difference. The outcome of the study represents that the Novel stochastic gradient algorithm based model provides more promising results in an innovative analysis on improved efficient optimization. In this work, the Accuracy of 92.1 is calculated using the Novel stochastic gradient algorithm transform compared with machine learning based model Accuracy of 71.4. The results showed a statistical significance difference between the two groups. The machine learning algorithm, is chosen based on the processed inputs, which helps in identifying the appropriate algorithm for implementation.

Similar findings in the stochastic gradient algorithms are frequently used in complex problem solving because they primarily imitate natural phenomena or rules and have excellent self-organizing, self-learning, and adaptive abilities, resulting in global optimal solutions with good robustness. Efficient optimization in dispensation of cloud applications for faster accessibility [8] with high security. Although the results of the study showed better performance with limited attributes using the novel stochastic gradient with forward back propagation algorithm model and the machine learning algorithms [9] such as random forest algorithm, and support vector machine algorithms with certain limitations. The number of parameters such as the type of file, size, purpose are analysed in the dispensation of cloud applications to improve high performance in a faster with high security. As a future scope the parameters can be investigated to expand the usability in various sectors to deploy the cloud services in vast manner for the daily activities using deep learning methods.

V. CONCLUSION

In conclusion, improving efficient optimization in dispensation of cloud application can result in achieving the best services with reasonable cost benefit, environmental advantages with the utility of data centre for the improved network performance. In this proposed system, novel stochastic gradient forward back propagation algorithm has an accuracy rate of 92.1 % and the machine learning algorithm has an accuracy rating of 71.4%. From the results, it is inferred that the novel stochastic gradient forward back propagation algorithm produces better performance than the conventional machine learning algorithm for the efficient optimization.

REFERENCES

[1] Bourantas, Christos, Ryo Torii, Sergey Karabasov, and Rob Krams. 2024. "Editorial: Computational Modelling of Cardiovascular Hemodynamics and Machine Learning." *Frontiers in Cardiovascular Medicine* 11 (February): 1355843.

[2] Chen, Xinyuan, Michael O. Harhay, Guangyu Tong, And Fan Li. 2024. "A Bayesian Machine Learning Approach for Estimating Heterogeneous Survivor Causal Effects: Applications To A Critical Care Trial." *The Annals Of Applied Statistics* 18 (1): 350–74.

[3] Huang, Qidong, Xiaoyi Dong, Dongdong Chen, Hang Zhou, Weiming Zhang, Kui Zhang, Gang Hua, Yueqiang Cheng, and Nenghai Yu. 2024. "PointCAT: Contrastive Adversarial Training for Robust Point Cloud Recognition." *IEEE Transactions on Image Processing: A Publication of the IEEE Signal Processing Society* PP (March). https://doi.org/10.1109/TIP.2024.3372456.

[4] V.V.B.Raju, T.P.Anithashri, 2024Improvisation of data security in mobile based data by using novel blow fish algorithm over DES algorithm, AIP Conf. Proc. 2729, 040005 (2024)

[5] Mahesh Muthulakshmi R, Anithaashri T.P, Novel weight-improved particle swarm optimization to enhance data security in cloud,. 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) *(2023): 195-200.*

[6] Kavuri, Venkata Ramaiah, and Anithaashri. 2023. "Efficient Secured Cloud Storage System Using Dynamic Multiple Clouds Cryptographic Algorithm." In *2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)*. IEEE. https://doi.org/10.1109/i-smac58438.2023.10290155.

[7] Liang, Zhicheng, Junhua Liu, Mallesham Dasari, and Fangxin Wang. 2024. "Fumos: Neural Compression and Progressive Refinement for Continuous Point Cloud Video Streaming." *IEEE Transactions on Visualization and Computer Graphics* PP (March). https://doi.org/10.1109/TVCG.2024.3372096.

[8] Sharma, Prerna, Moolchand Sharma, and Mohamed Elhoseny. 2020. *Applications of Cloud Computing: Approaches and Practices*. CRC Press.

[9] Singh, Pradeep. 2022. *Fundamentals and Methods of Machine and Deep Learning: Algorithms, Tools, and Applications*. John Wiley & Sons.

[10] Sitaram, Dinkar, and Geetha Manjunath. 2011. *Moving To The Cloud: Developing Apps in the New World of Cloud Computing*. Elsevier.

[11]Srinivas, Mettu, G. Sucharitha, and Anjanna Matta. 2021. *Machine Learning Algorithms and Applications*. John Wiley & Sons.

[12] Wang, Lizhe, Rajiv Ranjan, Jinjun Chen, and Boualem Benatallah. 2017. *Cloud Computing: Methodology, Systems, and Applications*. CRC Press.

Temporal Convolutional Network-Based Animal Intrusion Detection Model for Smart Farming

Renuka N
Assistant Professor,
Department of
Information Technology,
K S R College of Engineering,
Tiruchengode,
Tamil Nadu – 637214,
renuka66ksr@ksriet.ac.in

Kannan N
Assistant Professor,
Department of
Information Technology,
K S R College of Engineering,
Tiruchengode,
Tamil Nadu – 637214,
kannanmese@ksriet.ac.in

Abinesh S
Research Scholar,
Department of
Information Technology,
K S R Institute For Engineering and Technology, Tiruchengode,
Tamil Nadu – 637214,
abinesh70103@ksriet.ac.in

Dhanushram S
Research Scholar,
Department of
Information Technology,
K S R Institute For Engineering and Technology, Tiruchengode,
Tamil Nadu – 637214,
dhanushramsdr@ksriet.ac.in

Sowmiya T
Research Scholar,
Department of
Information Technology,
K S R Institute For Engineering and Technology,
Tiruchengode,
Tamil Nadu – 637214,
sowmiyathirugnanam21@ksriet.ac.in

Corresponding Author: renuka66ksr@ksriet.ac.in

Abstract— **This research aims to develop a reliable model for animal intrusion detection in smart farming by utilizing a temporal convolutional network. The present work involves three groups. Group 1 refers to the CNN algorithm, which is used for identifying and analysing data. Group 2 refers to the Bi-LSTM algorithm used for training and identifying multi-media data. Group 3 refers to TCN algorithm to improve accuracy and responsiveness. By significantly reducing crop raiding occurrences and enhancing the ability of intelligent farming to detect animal encroachment, the Integrated Wildlife Management System integrates TCN with computer vision for 93% accuracy. Temporal Convolutional Networks (TCNs) are used by the animal intrusion detection model to enhance security and resource management through precise categorization and adaptable handling of changing animal behaviours in agriculture**.

Keywords*— *Animal Intrusion Detection Model, Smart Farming, Temporal Convolutional Network, Convolutional Neural Network, Computer vision, Wildlife.

I. Introduction

The objective of this smart farming model is to detect and track the presence of unauthorized animals in agricultural regions over time using a Temporal Convolutional Network (TCN)-based animal intrusion detection model [1]. Exceeding the original in both accuracy and speed is a modified YOLOv2-based animal species detection model that uses multi-level feature merging and deformable convolutional layers. Better geometric variation flexibility makes it perfect for embedded devices used in real-time animal conservation deployment [2]. This analysis looks at how digital technology, especially sensor-based AI applications, might improve chicken farming with the goal of intelligent automation in poultry management. The goal is to increase the productivity, quality, and economics of chicken production by merging computer vision and IoT [3]. To prevent ungulate attacks on crops, this research develops a Smart Agriculture application that combines ultrasonic emission with computer vision. For farmers and agronomists, it provides insightful information about the practicality of an edge computing-based intelligent animal repulsion system [4]. This work combines a hybrid CNN + Bi-LSTM model for sign language translation and recognition with an NMT + GAN model for creating sign gesture movies from voice. It shows significant gains in both recognition accuracy and visual quality as judged by different evaluation metrics [5]. Introducing a New Animal Intrusion Detection Model for Deep Learning Smart Farming Using Temporal Convolutional Networks (TCNs). The model combines CNN and TCN algorithms to achieve an astounding 93 % accuracy. With its revolutionary design, it helps farmers identify and manage animal encroachment quickly, minimize crop losses, and encourage ecologically friendly agricultural methods.

II. Related Works

The literature study highlights existing gaps in computer vision for rural animal assaults. With its 98 % accuracy and 170 FPS, the Hybrid VGG-19+Bi-LSTM network model provides a solid solution that guarantees accurate information on animal studies and quick response mechanisms for the protection of human life [6]. To monitor and avoid degradation, land use and land cover classification are essential as it divides areas into distinct classes. The LISS-IV image categorization in this study uses Shared Layer Recurrent Neural Networks (SLRNN), which outperforms CLRNN, ULRNN, and conventional RNN, Because SLRNN performs better than CLRNN, ULRNN, and standard RNN, it supports multitask learning and achieves 5 % higher accuracy than the latter three [7][8]. The literature emphasizes the significance of productivity in the raising of chickens and promotes Feed Conversion Ratio (FCR) and economic management standards. This study looks at how sensor-based AI can be used to track the health and behaviour of hens, fill in the gaps in standards for poultry welfare, and suggest ways to improve computer vision and Internet of Things (IoT) smart farming for more effective management of poultry [9]. Deep convolutional neural networks have recently made significant progress, resulting in superior performance on image recognition benchmarks such as ILSVRC-2012. Using the VOC2007 and ILSVRC2012 datasets, this paper presents a saliency-inspired object identification model that predicts bounding boxes and scores that are independent of class, making it possible to handle numerous instances efficiently and achieve competitive recognition [10][11]. For multi-modal human action recognition, characteristics like input preprocessing, data augmentation, and network designs are examined. We propose a new descriptor for RGB movies, supplement skeletal data, and leverage inertial sensor data to achieve a 4% improvement over individual modalities on the UTD-MHAD data-set [12]. The accuracy of traffic sign detection can be increased by augmenting sparse R-CNN with coordinate attention blocks and a feature pyramid. Customized augmentation along with fresh datasets tackle data heterogeneity, guaranteeing algorithm performance while travelling [13]. The research highlights the frequency of cervical cancer and the significance of early diagnosis, and it suggests Few-Shot Object Detection Generative Adversarial Network (FSOD-GAN), a computer vision system that combines FR-CNN with hierarchical classification for cervical spot detection. With digital colonoscopy, it is possible to achieve 95% accuracy in the diagnosis of three different forms of cervical cancer lesions, indicating the promise of automated screening. The literature discusses the fixed-size input requirements that limit deep CNNs and introduces the SPP-net, a enhanced image classification technique that uses a spatial pyramid pooling mechanism. SPP-net overcomes fixed-size restrictions by demonstrating cutting-edge accuracy and faster feature computation in object detection on many datasets. The literature presents R-CNN as a scalable method with over 30% improvement in mean average precision (mAP) to solve plateaued item detection performance on PASCAL VOC. With the help of supervised pre-training and high-capacity CNNs, R-CNN achieves a noteworthy 53.3% mAP, indicating considerable performance improvements [18]. The body of research highlights how crucial object detection is to guaranteeing the safety of autonomous driving systems. In AP50 and AP75 on the BCTSDB dataset.The study presents Detect Former, a category-assisted transformer object detector that improves accuracy over baseline methods by outperforming Retina Net and FCOS with 97.6% and 91.4% detection performance, respectively.

III. Materials And Methods

An animal intrusion detection model for smart farming integrates edge computing and computer vision with CNN and TCN algorithms through the use of ultrasonic repellent sounds and temporal convolutional neural networks (TCNs) represented in Fig. 1. Human-wildlife conflicts in agriculture are addressed here.

Fig.1 Block diagram of temporal neural network, convolutional neural network.

The detection of animal incursion involves three categories. Group 1 gathers pictures and videos of animals invading the model using the CNN algorithm for continuing study. The CNN algorithm is used to analyze this data to identify and classify invasive animal species in agricultural regions [6].

Group 2 used collected photos and videos as training data and used the Bi-LSTM algorithm for animal invasion identification in the project. Bi-LSTM integration enhanced the model's capacity to recognize behavioral patterns.

Group 3 focused on a project that used CNN and TCN algorithms represented in Fig. 2 and Fig. 3 respectively to identify animal trespassing in smart farming and sound an ultrasonic alert that was represented in Fig. 4. In order to improve accuracy and responsiveness, they gathered large datasets of instances of wildlife invasion and developed a hybrid system for real-time detection and farmer alerting.

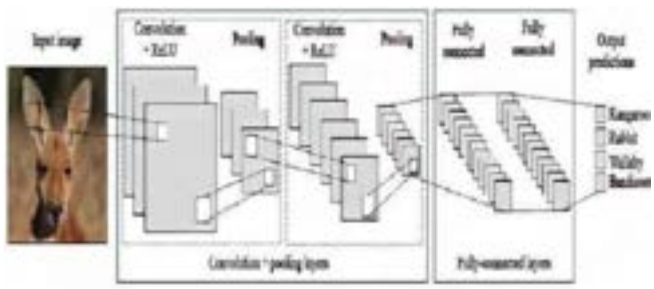


Fig.2: The image represents working process of CNN algorithm

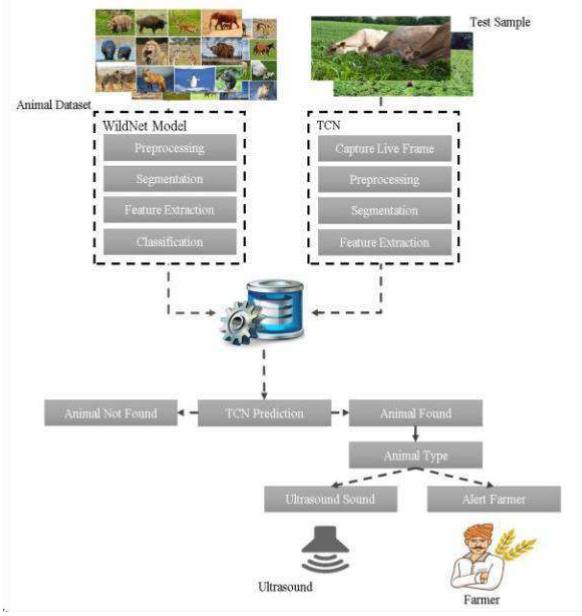


Fig. 3: System architecture for temporal neural network, convolutional neural network

Fig. 4: Concept for detecting intruder animals to activate repellent

IV. RESULTS

An early analysis of an ongoing study that detects animal incursions using CNN and Bi-LSTM algorithms produced an accuracy of 83.2% in detecting invasive species. A new Integrated Wildlife Management System was presented, acknowledging its limitations, which combined TCN with computer vision to achieve an amazing 93% accuracy. The efficiency of this technology was greatly increased, offering a more dependable method of lowering crop raiding incidents.

Significant progress has been achieved in animal intrusion detection for intelligent farming in the planned project which is represented in Fig. 5. The method surpassed the existing model in recognising invasive species by merging CNN and TCN. It also has a device for emitting targeted ultrasonic noises to discourage wildlife, guaranteeing quick reactions to incursions. The system, which is driven by edge computing, allows for real-time reaction and quickly reduces the likelihood of crop raiding for farmers.

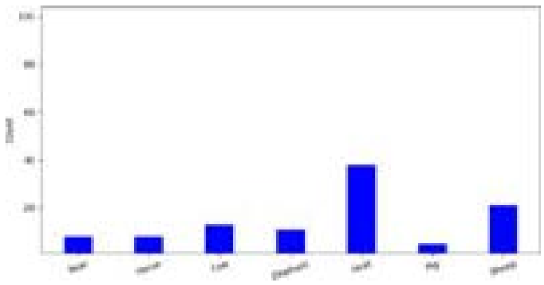


Fig. 5 The image represents analysis of animal intrusion detection. Then the x-axis represents animals and y-axis represents count of the animals.

The management of human-wildlife conflicts and agricultural interests is advanced by the proposed Temporal Convolutional Network-Based Animal Intrusion Detection Model for Smart Farming with ultrasonic alarm sound. It achieves a high accuracy of 93% in Fig. 6 for reducing crop raiding, operational expenditures, and environmental effect. This study integrates cutting-edge technology to support agricultural and biodiversity coexistence.

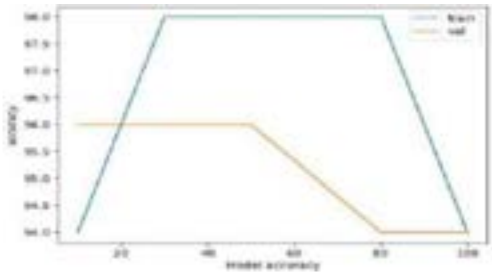


Fig. 6 Training and Validation Accuracy

V. DISCUSSION

Utilizing cutting-edge technology, the Integrated Wildlife Management System offers efficient crop protection while promoting coexistence through eco-friendly, real-time animal intrusion detection and repelling. Significant gains in accuracy and visual quality are demonstrated by the proposed hybrid models for sign language recognition, translation, and video generation, which effectively address current issues across a range of evaluation metrics. Presenting Ellipse R-CNN, a CNN-based detector designed to segment objects that are heavily occluded, such as clusters of fruits. The approach outperforms state-of-the-art models in accurately inferring multiple elliptical objects, even when obscured.

For COVID-19 screening with Chest Radiography (CR) images, COVID_SCREENET is a bifold architecture. Radiologists have confirmed that it achieves 100 % accuracy in both real-time and open-source datasets. Using digital colposcopy images, FSOD-GAN automates cervical cancer screening and achieves 99 % accuracy in identifying the stages of the disease. Three different types of cervical cancer lesions are identified using hierarchical multiclass classification and spot detection using FR-CNN.

Future development could involve improving the Temporal Convolutional Networks (TCN) to increase the Animal Intrusion Detection Model's accuracy. The efficiency of the system in reducing crop raiding incidents could also be increased by refining the ultrasonic emission technique to repel particular wildlife species and increasing the system's adaptability to a wider range of environmental conditions. By incorporating machine learning algorithms to modify repelling strategies in response to real-time data, the system's efficacy and adaptability could be improved.

VI. CONCLUSION

TCN and computer vision are combined in the Integrated Wildlife Management System to address conflicts between humans and wildlife in the face of habitat degradation. It quickly reduces crop raiding, providing environment friendly solutions promoting agriculture and biodiversity, driving conversations towards a resolution of such conflicts through the use of ultrasonic emission and edge computing integration.

REFERENCES

[1] Xu, Hang, Ruixin Xie, Hong Han, Zhaoxia Zhang, Jianguo Zhang, Li Liu, Bingjie Wang, and Liqiang Li. 2019. "A LCX-Based Intrusion-Detection Sensor Using a Broadband Noise Signal." IEEE Access: Practical Innovations, Open Solutions 7: 161928–36.

[2] Ibraheam, Mai, Kin Fun Li, and Fayez Gebali. 2023. "An Accurate and Fast Animal Species Detection System for Embedded Devices." IEEE Access: Practical Innovations, Open Solutions 11: 23462–73..

[3] Bhuiyan, Md Roman, and Philipp Wree. 2023. "Animal Behavior for Chicken Identification and Monitoring the Health Condition Using Computer Vision: A Systematic Review." IEEE Access: Practical Innovations, Open Solutions 11: 126601–10.

[4] Adami, Davide, Mike O. Ojo, and Stefano Giordano. 2021. "Design, Development and Evaluation of an Intelligent Animal Repelling System for Crop Protection Based on Embedded Edge-AI." IEEE Access: Practical Innovations, Open Solutions 9: 132125–39.

[5] Natarajan, B., R. Elakkiya, R. Bhuvaneswari, Kashif Saleem, Dharminder Chaudhary, and Syed Husain Samsudeen. 2023. "Creating Alert Messages Based on Wild Animal Activity Detection Using Hybrid Deep Neural Networks." IEEE Access: Practical Innovations, Open Solutions 11: 67308–21.

[6] Supriyanto, Eko, R. Rizal Isnanto, and Sutrisno Hadi Purnomo. 2023. "Computer Vision in Chicken Monitoring System Using Machine Learning: A General Review." E3S Web of Conferences 448: 02014.

[7] Erhan, Dumitru, Christian Szegedy, Alexander Toshev, and Dragomir Anguelov. 2014. "Scalable Object Detection Using Deep Neural Networks." In 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE. https://doi.org/10.1109/cvpr.2014.276.

[8] Imran, Javed, and Balasubramanian Raman. 2020. "Evaluating Fusion of RGB-D and Inertial Sensors for Multimodal Human Action Recognition." Journal of Ambient Intelligence and Humanized Computing 11 (1): 189–208.

[9] Liang, Tianjiao, Hong Bao, Weiguo Pan, and Feng Pan. 2022. "Traffic Sign Detection via Improved Sparse R-CNN for Autonomous Vehicles." Journal of Advanced Transportation 2022 (March): 1–16.

[10] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2014. "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition." Computer Vision – ECCV 2014, 346–61.

[11] Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. 2014. "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation." In 2014 IEEE Conference on Computer Vision and Pattern Recognition. IEEE. https://doi.org/10.1109/cvpr.2014.81.

[12] Liang, Tianjiao, Hong Bao, Weiguo Pan, Xinyue Fan, and Han Li. 2022. "DetectFormer: Category-Assisted Transformer for Traffic Scene Object Detection." Sensors 22 (13): 4833.

[13] Adami, Davide, Mike O. Ojo, and Stefano Giordano. 2021. "Design, Development and Evaluation of an Intelligent Animal Repelling System for Crop Protection Based on Embedded Edge-AI." IEEE Access: Practical Innovations, Open Solutions 9: 132125–39.

Utilizing IoT for Animal Farming Waste Detection: YOLO-Based Plastic Component Tracking with Buzzer Alert System

Nithisha A
Department of CSE
Karunya Institute of technology and sciences
Coimbatore,Tamil Nadu, India
nithishaalbert@gmail.com

E. Bijolin Edwin
Department of CSE
Karunya Institute of technology and sciences
Coimbatore,Tamil Nadu, India
bijolin@karunya.edu

M. Roshni Thanka
Department of DSCS
Karunya Institute of technology and sciences
Coimbatore,Tamil Nadu, India
roshni@karunya.edu

V.Ebenezer
Department of DSCS
Karunya Institute of technology and sciences
Coimbatore,Tamil Nadu, India
ebenezerv@karunya.edu

Stewart Kirubakaran
Department of CSE
Karunya Institute of technology and sciences
Coimbatore,Tamil Nadu, India
stewart@karunya.edu

T. Mathu
Department of DSCS
Karunya Institute of technology and sciences
Coimbatore,Tamil Nadu, India
mathu@karunya.edu

***Abstract*— Smart agricultural technologies are quickly evolving due to their efficacy and efficiency in animal health monitoring. The increasing prevalence of plastic waste poses a significant threat to animal health in farming environments. This paper presents a novel approach for tracking plastic components, specifically paper and plastic bottles, within the felids of animal farming using Internet of Things (IoT) technology. The proposed system employs the YOLO (You Only Look Once) YOLO model for real-time object detection, which identifies waste paper and bottles in the fields. Upon detection, the system sends the information to a NODEMCU, which triggers a buzzer alert to discourage animals from consuming the identified hazardous items. The implementation involves deploying IoT sensors strategically across the animal farming facility to capture visual data. The YOLO model is then employed to process the collected data, accurately identifying, and categorizing plastic components. Upon successful detection, the system sends serial data to a NODEMCU, triggering a predefined response. In this case, an alarm is activated to alert farm personnel about plastic waste in the fields.**

***Keywords*— IoT, Sensors, Precision Farming, Cattle Field Monitoring, Object Detection, YOLO.**

I. INTRODUCTION

The use of Internet of Things (IoT) technology in farming, especially in smart farming, has completely changed how farmers take care of their animals. With IoT devices, farmers can easily keep track of their animals even on large farms. These devices are crucial for quickly identifying sick animals and separating them from the rest to stop diseases from spreading. Making sure animals stay healthy is not just important for them, but also for the people who rely on them and for the farm's productivity.

Using wearable technology has made a big difference in how farmers understand animal behavior. For example, special neck collars with sensors are popular in dairy farming. These collars, often using radio-frequency identification technology, help farmers control how much their animals eat and keep an eye on their health. They even work with automatic milking systems to make farming easier. These wearable devices collect data from the animals and store it in databases or the cloud. By analyzing this data, farmers can learn a lot about their animals' health and behavior. This helps them make better decisions to keep their animals healthy and their farms productive. Overall, IoT technology is making farming more efficient and effective.

The introduction of wearable monitoring technologies has completely changed how we understand and track animal behavior. In the past, we mainly relied on watching animals with our eyes, but now we have advanced gadgets that provide precise data about what animals are doing. These technologies, powered by the Internet of Things (IoT), offer a more scientific and detailed approach to monitoring livestock activities. There are many different types of wearable devices available now, giving farmers flexible tools to manage their animals more effectively. For example, in dairy farming, neck collars with built-in sensors have become very popular. These collars, which are worn by cows, help farmers control how much the animals eat and keep track of important health indicators. They even work seamlessly with automated milking systems, making dairy farming easier and more efficient. Sensors must demonstrate a high level of accuracy to effectively capture the nuances of the measured variables. Reliability is another crucial factor, encompassing not only the sensor's durability and robustness but also its consistency in performance over time and under various environmental conditions. Furthermore, careful evaluation of the cost-effectiveness of sensors is essential, considering not only the initial investment but also the long-term maintenance and operational expenses.

As depicted in Figure 1, the overall architecture of a wearable device system involves the installation of sensors on animal subjects, capturing relevant data that is then stored in databases or cloud servers. By analyzing this data, farmers can learn a lot about their animals' health and behavior. This information helps them make better decisions to keep their animals healthy and make their farms more productive. As wearable technology continues to improve, it promises to bring even more benefits to the agriculture industry, helping farmers make smarter decisions and run their farms more efficiently.

Fig 1: Cattle Wearable Device

II. Literature Survey

Wang et al. [1] explored the usage of intravaginal gadgets to tune adjustments in vaginal temperature at some point of the estrous cycle in dairy cows, showcasing the effectiveness of this approach in detecting physiological variations in cattle. Similarly, Wackers et al. [2] added a catheter-based impedimetric sensor for assessing intestinal histamine degrees in sufferers with irritable bowel syndrome, highlighting improvements in biosensing technologies for clinical applications.

In the world of agricultural IoT, Vangala et al. [3] discussed the implementation of stable sensing structures for agriculture from a blockchain attitude, emphasizing the significance of records integrity and protection in agricultural tracking. Meanwhile, Tresoldi et al. [4] examined how one of a kind sampling techniques and measurement gadgets impact vaginal temperature readings in lactating dairy cows, providing insights into optimizing statistics series techniques for animal fitness monitoring.

Steinmetz et al. [5] established the accuracy of the RumiWatch Converter V0.7.Four.5 in robotically tracking dairy cow behavior, contributing to the improvement of automated cattle management systems. Additionally, Kurtser et al. [6] confirmed the practicality of the use of a cell robot and RGB-D digicam to assess grape cluster length in vineyards, showcasing advancements in precision agriculture strategies.

Regarding IoT protection, Li et al. [7] addressed cybersecurity challenges in IoT-pushed deliver-your-very own-device (BYOD) environments, stressing the need of robust security features. Bai et al. [8] brought an impedimetric natural electrochemical sensor for detecting histamine in precision animal agriculture, presenting a unique method for monitoring animal fitness.

In broader IoT packages, Song et al. [9] performed a complete survey on IoT packages in clever logistics, figuring out opportunities to enhance supply chain control. Additionally, Quy et al. [10] explored smart healthcare IoT programs the use of fog computing, discussing architectural concerns and demanding situations in healthcare IoT deployments.

Ma et al. [11] developed a non-contact frame temperature tracking machine for farm animals livestock, presenting an revolutionary technique to animal health tracking. Furthermore, Patle et al. [12] assessed a smart sensor machine for predicting plant illnesses the usage of LSTM networks, demonstrating the capacity of device studying in managing crop illnesses.

In the field of faraway sensing, Patil et al. [13] proposed a changed YOLOv5 method for detecting small objects in faraway sensing photographs, contributing to advancements in satellite imagery analysis. Finally, Mohan et al. [14] carried out an in depth evaluation of sustainable improvements in precision farming enabled with the aid of IoT and Industry five.0 technologies, imparting insights into their transformative impact on agriculture. Each reference adds a unique perspective to the growing field of smart agriculture and technology integration in the farming sector.

III. System Architecture

The detected information is transmitted to an IoT-enabled central server, ensuring seamless communication between the detection system and the alert mechanism. Having sensors on hand is crucial to collect the current data on the health of cattle. These sensors are attached to animals or placed strategically inside their environments. Livestock sensors collect data on the animals and the environment around them. While an animal's body temperature may be determined by a sensor attached to it, the temperature, humidity, and air quality of the livestock's living quarters can all be measured by a sensor installed in the surrounding area. Figure 2 illustrates the integration of the system with the central server.

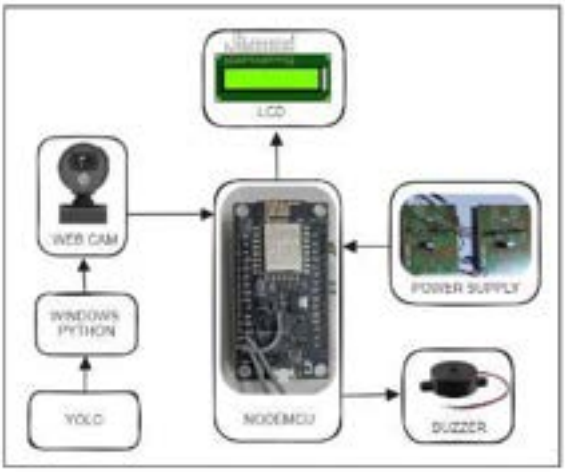


Fig 2: IoT enabled central server architecture

A. NodeMCU Controller

There are several electrical uses for this programmable microcontroller board, which is available as an open source product. Due to its cost-effectiveness, versatility, and user-friendly nature, this component is an integral addition to the proposed system. Both float and flow sensors are connected to this key component as part of the system configuration. A NodeMCU microcontroller is employed to receive the detected values from the IoT server. The controller is programmed to activate a buzzer alert system upon identification of plastic components, creating an audible deterrent for animals.

B. Web Cam

Without downloading the software that is provided with cameras on their PC, users are unable to record or stream video over the Internet. Webcams may record high-definition videos and still images even if their video quality may be lower than that of other types of cameras. A webcam or tiny digital video camera, maybe simply linked to a Raspberry Pi computer. It is an input tool for digital picture capturing. This work used a Web cam for capturing the obstacle images, which are present in outdoors and indoors with the help of a web camera. Monitoring in the project is achieved through: systematic deployment of IoT sensors across farming areas, continuous capture of visual data by sensors, real-time processing of data using the YOLO model for object detection, immediate alert activation via NODEMCU upon detection of waste materials, ensuring timely intervention to prevent animal consumption.

C. LCD Display

LCD module at 40% relative humidity and 40% temperature, respectively. Higher temperatures can cause the display's overall color to alter, while lower temperatures can delay the rate at which the display blinks. When the temperature drops within the designated range, the display will normalize. Polarizer peel-off, bubble formation, and polarisation deterioration can all be brought on by heat and humidity. Displays whether the object can be eatable or not.

D. *12V Power Supplies*

The 12V or 12VDC power supply stands out as a highly prevalent energy source in contemporary applications. It is often employed to transform input from 120VAC or 240VAC to an output of 12VDC using a combination of a transformer, diode, and transistor. There are two primary types of 12V power supplies: regulated and unregulated. Additionally, acopian switching regulated power supplies utilize robust EMI filtering and shielding techniques to mitigate noise transmission to both the line and load, addressing both common and differential modes.

E. Buzzer Alert System

A buzzer or beep serves as an electromechanical, mechanical, or audio signaling device. These devices, commonly referred to as buzzers or beepers, find widespread use in applications such as timers, railway horns, alarm clocks, and signaling keystrokes or mouse actions. In a specific context, a buzzer was utilized in this project to warn animals about the presence of harmful materials, playing a crucial role in deterring them from consuming plastic waste.

F. Object Detection using YOLO

The YOLO (You Only Look Once) algorithm is applied to swiftly detect waste paper and bottles in farming fields, ensuring animals' well-being. YOLO's neural network-based approach enables fast and precise object identification, making it popular for various applications like surveillance and traffic management. Object detection in the project is achieved through: strategic deployment of IoT sensors capturing visual data, preprocessing of captured images to enhance clarity, utilization of the YOLO model for real-time detection and classification of waste paper and plastic bottles, and integration with NODEMCU for immediate alert activation upon detection. The algorithm utilizes a CNN to predict class probabilities and bounding boxes simultaneously, simplifying the detection process. Its speed and accuracy make it a favored option for real-time object detection, bolstering the reliability of our study's goals.The choice of YOLO (You Only Look Once) for object detection in our study is based on its real-time processing capabilities, which swiftly detect objects in images or video frames. YOLO's grid-based approach predicts bounding boxes and class probabilities directly, ensuring rapid, accurate detection without complex region proposals.

YOLO Algorithm

Input: Input cattle field images

Output: Identified plastic Images

<u>Procedure</u>

```
Initialize:
  State := RELEASED
  Voted : = FALSE

For q_a to enter the critical section:
  State :=  WANTED
  Multicast request to all processes in J_a
  Wait until (number of replies received = K)
  state := HELD

On receipt of a request from q_a at q_b:
  if (state = HELD or voted = TRUE):
    queue request from q_a without replying
  else:
    send reply to q_a
    voted := TRUE

For q_a to exit the critical section:
  state := RELEASED
  Multicast release to all processes in J_a

On receipt of a release from q_a at q_b:
  if (queue of requests is non-empty):
    remove head of queue from q_c, say
    send reply to q_c
    voted := TRUE
  else:
    voted := FALSE
```

IV. Results and Discussion

The IoT project aims to tackle plastic waste in agricultural settings involving animals by introducing a method that utilizes the YOLO (You Only Look Once) model for instant object detection. This system deploys IoT sensors throughout farming areas to capture visual data, processed by the YOLO model to identify and classify waste paper and plastic bottles. Upon detection, the system activates a buzzer alert, discouraging animals from consuming harmful materials. Methodology assessment criteria include precision, real-time efficiency, reliability, and scalability, with a comparison against traditional methods revealing potential benefits in accuracy, cost-effectiveness, and implementation simplicity. The approach integrates IoT technology with the YOLO model for real-time object detection, offering a novel solution to the issue of plastic waste management in animal farming environments. Through a literature review, the increasing prevalence of plastic waste is identified as a significant threat to animal health in agriculture, underscoring the need for innovative technological interventions. While various studies have explored IoT applications in agriculture and waste management separately, the study uniquely combines these technologies to address the specific challenge of plastic waste detection in animal farming. By evaluating the methodology and discussing its implications in light of relevant literature, the study advances the field by providing a practical solution to mitigate the adverse effects of plastic waste on animal health and environmental sustainability. Additionally, future research directions are proposed to further refine the solution and its broader applications in agricultural sustainability efforts.

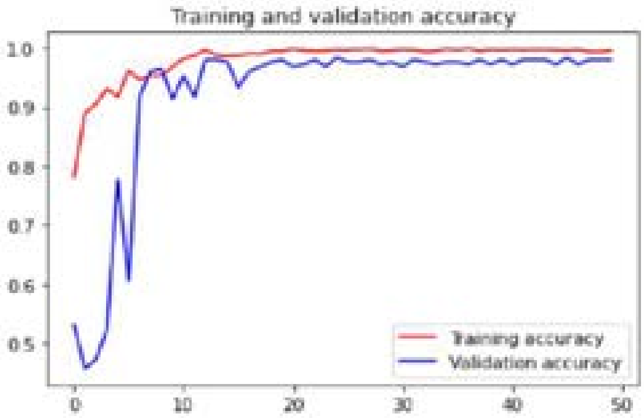


Fig 3: Accuracy of YOLO model over training epochs

The training accuracy curve is presented in Figure 3 The proposed model reaches a high accuracy as the number of epochs increases. High accuracy and reliability in the project are achieved through strategic deployment of IoT sensors capturing comprehensive visual data, utilization of the YOLO model for precise object detection, real-time processing ensuring timely identification of waste materials, consistent deployment of IoT sensors across farming areas, rigorous testing and validation of the YOLO model for accurate waste detection, redundant alert mechanisms like buzzer activation via NODEMCU, and continuous monitoring and refinement of the system to minimize false positives and false negatives, enhancing overall system performance and reliability.

Fig 4: Training losses of YOLO model over epochs

Likewise, the loss of the models decreases to around 0.05 as the number of epochs increases. The accuracy and loss curves are perfectly converged into each other. Figure 4 illustrates the Training losses of the YOLO model over epochs.

TABLE 1: COMPARATIVE ANALYSIS BETWEEN EXISTING AND PROPOSED SYSTEM

Feature	Proposed System	Existing System
Detection Method	YOLO algorithm with CNN for real-time object detection	Convolutional Neural Networks (CNN)
Object of Detection	Plastic components (paper, bottles)	Various agricultural anomalies
Alert System	Buzzer activation upon detection	SMS or Email notification
Sensor Technology	IoT sensors strategically deployed	Wearable sensors or satellite imagery
Implementation Scalability	Scalable across various farming environments	Limited to specific farm types or regions
Cost-Efficiency	Cost-effective NodeMCU controller	High initial investment for specialized sensors (gear)
Real-Time Capability	Real-time identification and alert activation	Delayed response due to processing requirements

Table 1 presents a comparative analysis between the project and an existing novel system. The project utilizes YOLO algorithm for real-time object detection of plastic components, employing cost-effective IoT sensors and activating a buzzer alert. This comparison highlights key features and benefits for precision animal farming waste detection.

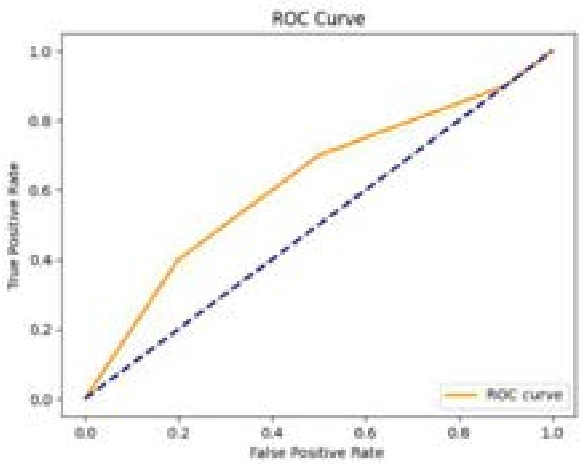


Fig 5: ROC Curve

The implementation of the ROC curve, as depicted in Figure 5, involves assessing the YOLO-based model's effectiveness in detecting plastic components within animal farming waste. Analysis of the ROC curve in Figure 5 enables the identification of optimal thresholds for activating the buzzer alert system, thus improving waste management efficiency and minimizing false alarms.

V. Conclusion

Ensuring the well-being of animals is a top priority worldwide, driving the need for effective scientific solutions. One promising approach gaining traction is the use of biosensors to manage animal health. While these sensors are still in development, they're showing promise in practical applications for animal care. Advanced technologies originally designed for human health are now being adapted for veterinary use, helping veterinarians accurately diagnose diseases and assess the health of animals.

This integration of advanced technology with precision farming practices is revolutionizing how care for animals. Precision farming covers a wide array of technologies aimed at improving animal husbandry practices for better health outcomes. In line with this trend, a new project introduces an innovative solution to monitor plastic waste in animal farming fields using IoT technology and sophisticated data processing. The system relies on the YOLO (You Only Look Once) object detection algorithm to spot waste paper and bottles scattered in the fields. When these items are identified, the system quickly notifies a NodeMCU controller. Then, the controller triggers a buzzer alert to deter animals from consuming the potentially harmful items, helping prevent health issues.

Beyond animal health, the project also aims to address environmental concerns by reducing plastic pollution in agricultural areas. By harnessing cutting-edge technology and creative solutions, this initiative not only protects animal well-being but also contributes to the preservation of ecosystems and the health of local communities.

References

[1] Wang, S., H. Zhang, H. Tian, X. Chen, S. Li, Y. Lu, L. Li, and D. Wang. 2020. Alterations in vaginal temperature during the estrous cycle in dairy cows detected by a new intravaginal device—A pilot study. Trop. Anim. Health Prod. 52:2265–2271. https://doi.org/10.1007/s11250-020-02199-5.I. S. Jacobs and C. P. Bean, "Fine particles, thin films and exchange anisotropy," in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[2] Wackers, G., T. Putzeys, M. Peeters, L. Van de Cauter, P. Cornelis, M. Wubbenhorst, J. Tack, F. Troost, N. Verhaert, T. Doll, and P. Wagner. 2020. Towards a catheter-based impedimetric sensor for the assessment of intestinal histamine levels in IBS patients. Biosens. Bioelectron. 158:112152. https://doi.org/10.1016/j.bios.2020.112152.R. Nicole, "Title of paper with only first word capitalized," J. Name Stand. Abbrev., in press.

[3] Vangala, A.; Das, A.K.; Kumar, N.; Alazab, M. 2020. Smart Secure Sensing for IoT-Based Agriculture: Blockchain Perspective. IEEE Sens. J, 21, 17591–17607.M.

[4] Tresoldi, G., K. E. Schütz, and C. B. Tucker. 2020. Sampling strategy and measurement device affect vaginal temperature outcomes in lactating dairy cattle. J. Dairy Sci. 103:5414–5421. https://doi.org/10.3168/jds.2019-16667.

[5] Steinmetz, M.; Von Soosten, D.; Hummel, J.; Meyer, U.; Dänicke, S. 2020. Validation of the RumiWatch Converter V0.7.4.5 classification accuracy for the automatic monitoring of behavioural characteristics in dairy cows. Arch. Anim. Nutr. , 74, 164–172.

[6] Kurtser, P.; Ringdahl, O.; Rotstein, N.; Berenstein, R.; Edan, Y. 2020. In-Field Grape Cluster Size Assessment for Vine Yield Estimation Using a Mobile Robot and a Consumer Level RGB-D Camera. IEEE Robot. Autom. Lett. ,5, 2031–2038.

[7] li, I.; Kaur, S.; Khamparia, A.; Gupta, D.; Kumar, S.; Khanna, A.; Al-Turjman, F. 2020. Security Challenges and Cyber Forensic Ecosystem in IoT Driven BYOD Environment. IEEE Access 2020, 8, 172770–172782.

[8] Bai, H., K. Vyshniakova, E. Pavlica, V. M. Rocha Malacco, A. Yiannikouris, T. R. Yerramreddy, S. S. Donkin, R. M. Voyles, and R. A. Nawrocki. 2020. Impedimetric, PEDOT:PSS-based organic electrochemical sensor for detection of histamine for precision animal agriculture. IEEE Sens. Lett. 4:4501004. https://doi.org/10.1109/LSENS.2020.3025162.

[9] Song, Y.; Yu, F.R.; Zhou, L.; Yang, X.; He, Z. 2021. Applications of the Internet of Things (IoT) in Smart Logistics: A Comprehensive Survey. IEEE Internet Things J., 8, 4250–4274.

[10] Quy, V.K.; Hau, N.V.; Anh, D.V.; Ngoc, L.A. 2021. Smart healthcare IoT applications based on fog computing: Architecture, applications and challenges. Complex Intell.

[11] Ma, S.; Yao, Q.; Masuda, T.; Higaki, S.; Yoshioka, K.; Arai, S.; Takamatsu, S.; Itoh, T. 2021 Development of Noncontact Body Temperature Monitoring and Prediction System for Livestock Cattle. IEEE Sens. J., 21, 9367–9376.

[12] Patle, K.S.; Saini, R.; Kumar, A.; Palaparthy, V.S. 2022. Field Evaluation of Smart Sensor System for Plant Disease Prediction Using LSTM Network. IEEE Sens. J. 22, 3715–3725.

[13] Shefali Patil, Sangita Chaudhari, Puja Padiya, 2023. "Small Object Detection in Remote Sensing Images using Modified YOLOv5," presented at the 2023 1st DMIHER International Conference on Artificial Intelligence in Education and Industry 4.0 (IDICAIEI), Wardha, India, 27-28 November 2023. DOI: 10.1109/IDICAIEI58380.2023.10406355.

[14] K. J. Mohan, I. T. Bella Mary and J. J. Paul, 2023. "IoT and industry 5.0 revolutionizing agriculture: a comprehensive review of sustainable advancements in precision farming," International Conference on Computer Vision and Internet of Things 2023), pp. 169-175, doi: 10.1049/icp.2023.2872.

Automated Shower for Physically Challenged Individuals

M.Parvathi*1
Dept. of Electronics and Communication Engineering
BVRIT HYDERABAD College of Engineering for Women
Hyderabad, Telangana, India
parvathi.m@bvrithyderabad.edu.in

T. Praveenya[2]
B.Tech, of Electronics and Communication Engineering
BVRIT HYDERABAD College of Engineering for Women
Hyderabad, Telangana, India
20wh1a0477@bvrithyderabad.edu.in

B. Akshaya[3]
B.Tech, of Electronics and Communication Engineering
BVRIT HYDERABAD College of Engineering for Women
Hyderabad, Telangana, India
20wh1a0479@bvrithyderabad.edu.in

V. Leena[4]
B.Tech, of Electronics and Communication Engineering
BVRIT HYDERABAD College of Engineering for Women
Hyderabad, Telangana, India
20wh1a0489@bvrithyderabad.edu.in

K. Sneha Bhuvaneshwari[5]
B.Tech, of Electronics and Communication Engineering
BVRIT HYDERABAD College of Engineering for Women
Hyderabad, Telangana, India
20wh1a04b5@bvrithyderabad.edu.in

***Abstract*—An automated shower system is essential to design for public facilities, to address the concerns about excessive water consumption and accessibility for individuals with physical disabilities. The proposed system utilizes advanced voice recognition technology to enable easy control of water temperature and showerhead positioning, promoting water conservation, safety, and convenience. Existing manual systems face challenges in setting desired parameters for disabled users, whereas the proposed system becomes a solution that eliminates manual adjustments through seamless voice commands. Key features include adjustable settings, automatic regulation and emergency shut-off for enhanced safety.**

Keywords— ESP-WROOM-32, Ultrasonic Sensor, Servo motor, Relay, DC Motor Water Pump, Battery, Wi-Fi (Arduino UNO Interface), Blynk Console, Voice flow

I. INTRODUCTION

Technology advancement is highly demanding for automated devices in nowadays living environment. Especially the use of automated devices is becoming essential for elderly and disabled people as well. One of the common needs and essential tasks that is required to be automatized is controlling water usage while taking a bath using the shower. It is identified that very young children, who are crazy about using a shower may waste water unknowingly. In addition, aged people who may suffer from Alzheimer's, or who lazy to use shower controls properly may lead to more water consumption. Hence in this scenario, it is required to identify a better control mechanism for the shower system while satisfying the customer's need. The main aim of the proposed automated shower system is to give utmost comfort for people with disabilities to have proper use of the shower system without any further aid from other people. This type of invention is also useful for the implementation of smart home facility that are highly needing by recent advancements [1]. The proposed automatic shower system uses sensor-based setup to reduce the water consumption during the bath process, and also provides comfort use of controls based on voice commands. This new invention of shower system can address the problem of false signaling that are arises during Bluetooth application [2]. The existing sensor-based shower systems [3,4,5] observed that they are highly sensitive to react fast during every move of bath position causes erroneous results. These systems suffer from false signaling to the user, hence leading to confusion in the use of appropriate controls via Bluetooth signaling. Also, the use of infrared sensors [6,7,8,9] causes unwanted reflections from Unusual light interference that leads to malfunctioning of the sensor system. The objective of the proposed design is to improve the performance of the shower by introducing of additional system by using voice-based controls. Introducing the Automated Shower System, an innovative solution designed for individuals with physical challenges. The smart shower integrates voice commands, servo motors, and ultrasonic sensors to create a user-friendly and adaptive bathing experience. Users can activate the system, select the preferred water temperature through voice commands, and enjoy a personalized shower. The system ensures precise control of water valves and incorporates user presence detection for safety and energy efficiency. Overall, an automated shower simplifies the bathing process, offering enhanced accessibility and comfort for individuals with disabilities or diverse needs.

II. LITERATURE REVIEW

The common thread in these discussions is the evolution of shower technology towards greater efficiency, accessibility, and user-friendliness. From autonomous smart showers using sensors and actuators for optimized water usage to designs integrating Bluetooth technology for mobile-controlled temperature adjustments, the overarching goal is to enhance the shower experience. Ongoing efforts to incorporate voice and hand gesture technologies highlight a commitment to continuous improvement, addressing the concerns related to water conservation, manual adjustments, and accessibility for individuals with disabilities.

In 2010, the Flow Controllable Shower Arm [10], a noteworthy invention designed to revolutionize the shower experience. This innovative system addresses multiple facets of user requirements by incorporating a flow control mechanism, ensuring efficient water usage. A distinctive feature of this shower arm is the absence of manual temperature adjustments, streamlining the user experience and mitigating potential difficulties, especially for those with physical challenges. Mr. Sirkin's creation marks a significant advancement in shower technology, emphasizing both water conservation and inclusivity in its design.

Disability Shower invented in 2012 [11] is a groundbreaking innovation in the realm of accessible shower systems. The hallmark of this invention lies in its thoughtful design, featuring a seating option to cater to individuals with disabilities, prioritizing comfort and inclusivity. His creation not only addresses the specific needs of users with mobility challenges but also streamlines the overall showering process by removing unnecessary complexities associated with traditional systems, marking a significant stride toward universal accessibility in personal care amenities.

Autonomous smart shower with sensors and actuators [12] introduced the new concept, an autonomous smart shower with sensors and actuators redefine the traditional shower experience. This innovative design is focused on optimizing water consumption by incorporating advanced technology to control the flow rate, thus contributing to environmental sustainability. Khan's autonomous smart shower represents a significant step forward in the integration of cutting-edge technologies to enhance water efficiency while prioritizing user convenience in personal hygiene routines.

Design of Automatic Sensor-Equipped Shower [13] invented in 2022, introduced the Design of Automatic Sensor-Equipped Shower, marking a notable advancement in shower technology. The innovation centers around the use of Bluetooth technology to control the water temperature through a mobile device, offering users a convenient and personalized shower experience. The continuous evolution of automatic sensor-based [14-20] shower systems is dedicated to staying at the forefront of technological innovations in the pursuit of more intuitive and responsive user-friendliness.

III. Proposed System

The automated shower system as shown in Fig 1 utilizes voice recognition technology to provide users with hands-free control over the showering experience. Users simply speak their desired water temperature and flow rate commands, and the system's voice recognition module processes and transmits these commands to the shower controller. The shower controller then adjusts the water settings accordingly, ensuring a personalized and comfortable showering experience. The Automated Shower System simplifies the initiation process through a voice command, allowing users to select their preferred water temperature using options like hot water, cold water, and normal water. This user-friendly interface, facilitated by voice commands, caters to diverse user needs. The system employs servo motors to precisely adjust the hot and cold water supplies, achieving the selected temperature and ensuring a tailored bathing experience.

An ultrasonic sensor detects user presence, activating the system and water pump for a safe and efficient shower. Following the shower, the system autonomously shuts down after 3 minutes, emphasizing safety and energy efficiency. Overall, the integration of voice control, servo motors, temperature sensors, and presence detection technology creates a sophisticated yet accessible showering experience.

The Automated Shower System's emphasis on user-friendly voice commands, precise servo motor control, and intelligent presence detection contributes to a customized and automated bathing experience for individuals with physical challenges.

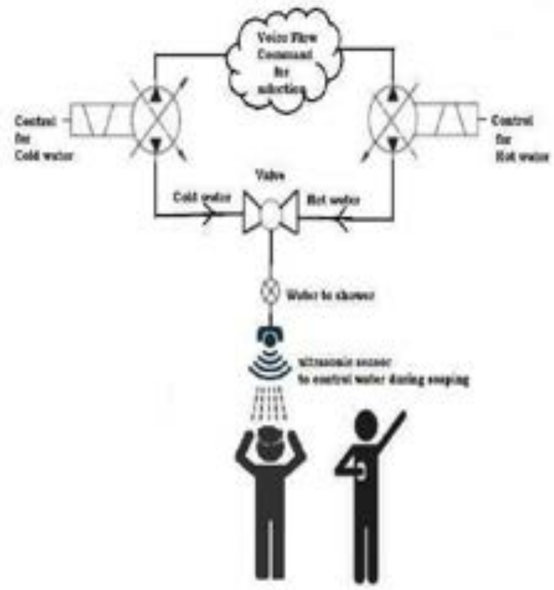


Fig. 1. Proposed Sensor based Shower System

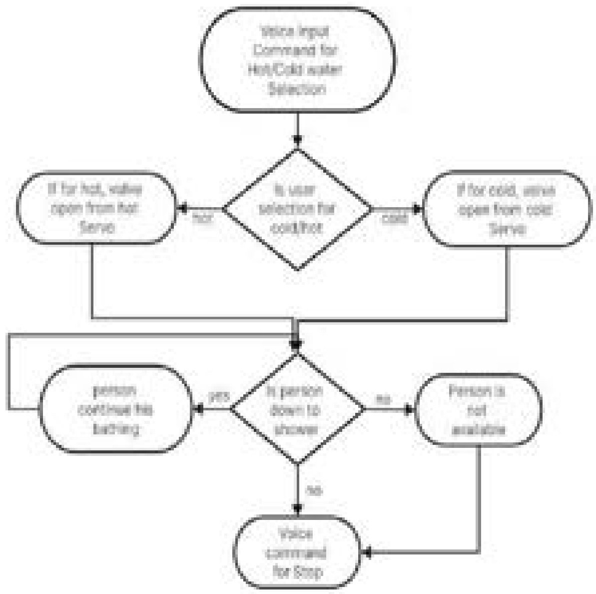


Fig. 2. Operational Flow of Shower System

The block diagram depicted in Fig 2 shows how the components of the automatic shower system are interconnected. It uses two servo motors, an ultrasonic sensor, a relay and a battery. It is presumed that water comes through two pipes one is for hot water and the other is for cold water. The servo motor is controlled by the microcontroller, which receives signals from the ultrasonic sensor and the voice input device. The battery powers the microcontroller and the servo motor. The relay is used to control the power to the water pump. The operation of the system majorly relies on voice input commands. The initial setup for taking a bath should be ready by selecting the type of water by the user choice command through the voice input system. The voice input system could be any Bluetooth or Wi-Fi setup. In the proposed prototype system the blink app is used. This blink app is used as an IOT-enabled system that allows the user to use the shower system from a remote location too to operate while returning from the office or the saloon, one can choose the type of water to initiate the water level to get ready by then the person reaches home.

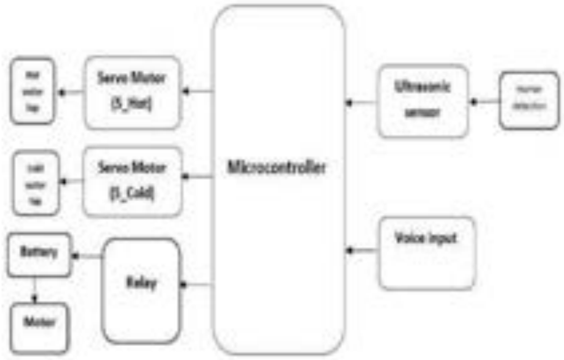


Fig. 3. Block diagram of proposed Shower System

A. ESP-WROOM-32

ESP-WROOM-32 shown in Fig 3 is the heart of the automated shower system, a powerful microcontroller that serves as the system's nerve center. This versatile chip seamlessly integrates various functionalities, including voice recognition, and sensor interfacing. By leveraging the ESP32-WROOM module's capabilities, the proposed system effortlessly interprets user commands, monitors sensor data, and regulates water flow, ensuring a safe, personalized, and hands-free showering experience for individuals with disabilities.

B. Ultrasonic Sensor

Ultrasonic sensors shown in Fig 4 measure target object distance by emitting sound waves and converting reflections into electrical signals. Comprising a transmitter and receiver, they utilise piezoelectric crystals for sound emission.

The proposed automated shower system integrates strategically positioned ultrasonic sensors within the stall. Continuously emitting and receiving ultrasonic waves, the sensors analyze the time taken for reflections to detect user presence and distance. Real-time data is fed into the system's microcontroller, enabling intelligent decisions based on user proximity. This facilitates automatic shower activation or deactivation, optimizing water usage and ensuring a seamless experience.

C. Servo Motors

A servo motor shown in Fig 5, renowned for its precision, integrates a built-in feedback mechanism to regulate position. Common in robotics and remote-controlled devices, servo motors play a pivotal role in the proposed automated shower system. Positioned strategically, they control hot and cold-water valves, enabling accurate adjustments to temperature and flow rate.

The system's microcontroller commands the servo motors, facilitating seamless transitions between open, closed, and partially open valve states. This precise control mechanism ensures the system responds to user preferences, maintaining consistent water settings. By harnessing the capabilities of servo motors, the proposed automated shower system provides a hands-free and personalized shower experience, particularly beneficial for individuals with disabilities.

D. Relay

The relay shown in Fig 6, used as electro-mechanical switches, utilizes DC electrical current to open or close contact switches. In a 5V single-channel relay, there are normally open and normally closed contacts alongside a coil.

Integrated all the above-said components into the proposed automated shower system, relays play a crucial role in controlling the water pump based on user presence detection. Strategically placed within the system, these relays manage the electricity flow to the water pump. Triggered by ultrasonic sensor signals indicating user presence, the relays activate, powering the water pump and initiating water flow. Conversely, when the sensors detect user absence, the relays deactivate, interrupting the circuit and halting pump operation. This intelligent mechanism ensures water is dispensed only when a user is present, promoting resource conservation and enhancing system efficiency.

E. DC Motor Water Pump

A DC motor pump shown in Fig 7 employs a Direct Current (DC) motor to transfer fluids by converting electrical energy into mechanical energy for liquid movement. Widely used in irrigation, aquaculture, and industrial processes, this compact and adaptable system is crucial for enabling precision control under varying loads.

In the proposed automated shower system, the DC motor pump efficiently delivers water to the showerhead, seamlessly integrating with the user detection mechanism. Triggered by ultrasonic sensors indicating user presence, the system's microcontroller activates the 12V DC motor pump, drawing water from the source and initiating a personalized shower experience only when a user is detected within the shower stall.

F. Battery

A battery shown in Fig 8, functioning as an electrochemical device, generates electric current through chemical reactions in its cells. It comprises positive and negative electrodes immersed in an electrolyte, serving as a powerhouse for electronic devices.

In the proposed automated shower system, a strategically positioned battery powers the DC motor pump, ensuring a reliable water flow. This battery-driven approach enhances portability, eliminating the need for a direct electrical outlet connection. The seamless integration of the battery and pump underscores the system's versatility, allowing for a personalized and independent shower experience. This user-centric design empowers individuals with disabilities, showcasing the system's adaptability in diverse settings.

Fig 4. ESP-WROOM-32

Fig 7. Relay

Fig 5. Ultrasonic Sensor

Fig 8. DC Motor Water Pump

Fig 6. Servo Motors

Fig 9. Battery

G. Arduino UNO

The Arduino IDE is a dedicated software platform designed for programming Arduino microcontrollers, offering a user-friendly interface for code development and deployment as shown in Fig 9.

Fig. 10. Arduino UNO Interface

In the proposed automated shower system, the Arduino IDE plays a pivotal role by providing a user-friendly environment for programming and controlling hardware components. As an open-source tool, it allows developers to customize code governing the ultrasonic sensors, servo motors, and DC motor pump, ensuring precise control over water flow, temperature adjustment, and user detection. The Arduino IDE's functionality extends to code compilation and uploading, facilitating seamless communication between software and hardware components and enabling the system to function autonomously.

H. Blynk App

Blynk is a mobile application designed for IoT application development shown in Fig 10, offering a platform to create custom interfaces for hardware control, especially with microcontrollers like Arduino. It simplifies IoT integration by providing a user-friendly interface for programming and device interaction.

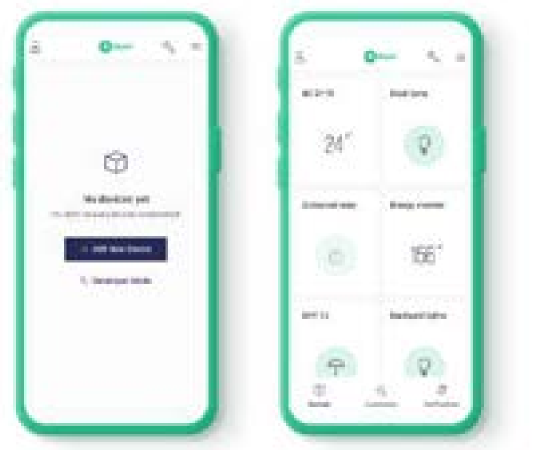


Fig. 11. Blynk App Overview

In the proposed automated shower system, the Blynk platform enables remote monitoring and control of system components. The Blynk mobile app allows users to visualize sensor data and, enhancing the shower experience and promoting independence. This application serves as a bridge between users and the automated shower system, facilitating seamless interaction and customization of features. With Blynk, the system leverages IoT capabilities, making it more accessible and user-centric. Users may remotely control on behalf of their old age people or differently abled people via the blink app by giving voice commands for a prior setup like water type whether hot or cold selection, whether to start the shower system or to stop as and when required.

I. Voice Flow

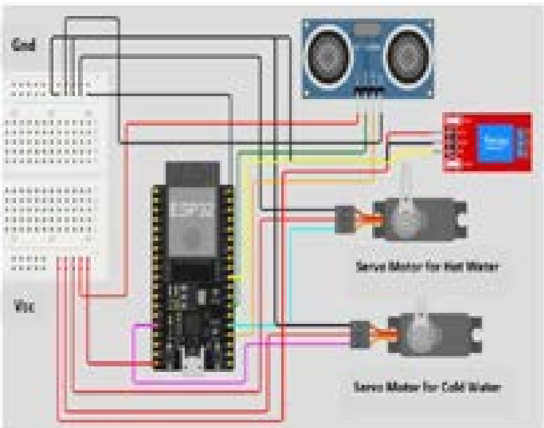
Voice flow shown in Fig 11, is a cloud-based platform that empowers developers and designers to create voice applications seamlessly, eliminating the need for extensive coding. This platform facilitates the design, development, and deployment of voice-driven experiences with an intuitive interface. Professionals can build interactive applications for platforms like Amazon Alexa and Google Assistant.

Fig. 12. Voice Flow Set-up

In the automated shower system, Voiceflow plays a crucial role, enabling hands-free control through voice recognition, as depicted. The platform translates voice commands into text and relays them to the Blynk app, allowing users to effortlessly manage the system. This integration showcases the power of Voiceflow in enhancing user interaction and control within the automated shower environment.

IV. Results and Observations

The proposed automatic shower system successfully demonstrated the integration of hardware components, Blynk app control, and Voice flow app interaction to deliver a personalized and convenient showering experience. The necessary connections and set are shown in Fig. 12. Ultrasonic distance measurements ensured the system operated within the desired range, while Blynk app control provided a user-friendly interface for temperature and flow adjustments. Voice flow app integration enabled hands-free control using simple voice commands, further enhancing the user experience. The system's success highlights the potential of technology to automate daily tasks and improve user convenience.



PIN	Connected to ESP-32
Hot Water Servo motor(signal)	Pin 13
Cold Water Servo motor(signal)	Pin 2
Ultrasonic(trig pin)	Pin 4
Ultrasonic(echo pin)	Pin 15
Relay(IN)	Pin 3

Fig. 13. Detailed description of proposed shower setup

The successful hardware setup, as shown in Fig 13, marked a significant milestone in the proposed work. The image showcases the meticulously assembled components as prototypes, each component carefully integrated to form a cohesive and functional system. The setup's flawless operation, evident through rigorous testing and evaluation, stands as a testament to the team's expertise and meticulous attention to detail. This accomplishment paves the way for further progress, enabling us to explore the full potential of the proposed design to achieve the desired objectives.

The Blynk app's seamless integration with the hardware components played a pivotal role in the success of the automatic shower system as shown in Fig 14. The Blynk app proved to be an invaluable tool in enhancing user control, monitoring sensor data, and providing personalized temperature adjustments for the automatic shower system. Its user-friendly interface simplified the process of activating and adjusting the system, allowing users to effortlessly control the flow and temperature of the water.

Fig. 14. Experimental set up for prototype design

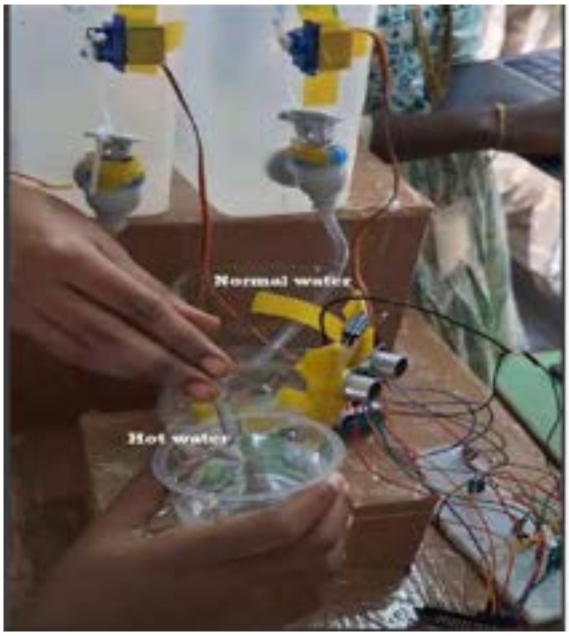


Fig. 15. Result observation of the prototype design

Fig. 16. Blynk Console of Proposed Shower System

Pseudo code for the proposed design functionality:

```
// Initialize serial communication
serial_begin (115200)
// Set relay pin as output and turn it off
set_pin_mode (relay, OUTPUT)
write_pin (relay, HIGH)
// Initialize Blynk connection
blynk_begin (auth, ssid, pass, "blynk. cloud", 80)
// Initialize servos
attach_servo (servo1, 2) // Cold water
attach_servo (servo2, 13) // Hot water
// Function to control the water pump
function pump(t)
  if (distance < 20 and t == 1)
    turn_on_relay ()
  else
    turn_off_relay ()
  end if
end function
// Function to measure the distance using an ultrasonic sensor
function ultrasonic ()
  send_trigger_pulse ()
  measure_duration ()
  calculate_distance ()
  print_distance ()
  delay (1000)
end function
// Handle incoming Blynk events
blynk_write (V3)//to turn on shower
blynk_write(V0) // Cold Water
blynk_write(V1) // HOT Water
blynk_write(V2) // Normal Water
```

TABLE I. COMPARISON OF VARIOUS SHOWER SYSTEMS

References and Proposed	Voice command	User proximity	Voice based Hot/cold type selection	Voice-based or Automatic stop	Auto adjust temperature
Angular Modern Monitor [14]	NO	NO	NO	NO	yes
HYDRAO smart showerhead [15]	NO	NO	NO	NO	NO
Smart Rain Shower Head [16]	NO	NO	NO	NO	NO
Thermometer Handheld Shower Heads [17]	NO	NO	NO	NO	NO
EvaDrop smart shower device [18]	NO	NO	NO	NO	yes
Moen shower controller [19]	NO	yes	NO	NO	yes
Smart Shower With Sensors [20]	NO	yes	NO	NO	yes
Proposed	yes	yes	yes	yes	yes

V. CONCLUSIONS

In conclusion, the proposed shower system is not just a technological upgrade but a step towards inclusivity and enhanced daily living. By seamlessly integrating voice and hand gesture recognition, we've aimed to redefine the shower experience for physically challenged individuals. The adjustable settings cater to diverse needs, while safety features, including emergency shut-off, prioritize user well-being.

This innovative system goes beyond its primary audience, offering practical benefits for individuals with Alzheimer’s, the elderly, those with autism, and even partial usability for the visually impaired. It's not just a shower; it's a user-centric solution designed to make a positive impact on the daily lives of various user groups. The performance of the whole system can be further improved by adding highly sophisticated sensors that take user gestures as input to control the whole bathing process and to give the user a sublime experience.

Acknowledgement: The team of students presented their work on “Design of Automated Shower for Physically Challenged Individuals” as part of the Makers Fair Technical Presentation Event in TWorks, organized by Telangana Govt., on 16-12-2023.

REFERENCES

[1] Kaushik Patowary. (2019). A history of showering. Retrieved November 16, 2020, from https://www.amusingplanet.com/2019/01/a-short-history-of-showering.html

[2] Ashley Poland. (2010). The advantages of using Bluetooth technology from existing technology. Retrieved November 16, 2020 from https://smallbusiness.chron.com/advantages-using-bluetooth-existing-technology-32921.html

[3] Andreas B. Dahlin. (2012) Problem and advantages associated with highly miniaturized sensor. Division of Bionanophotonics; Sensors 2012, 12(3), 3018-3036, Department of Applied Physics, Chalmers University of Technology, Fysikgränd 3, 41296 Göteborg, Sweden

[4] Amelia C Megan. (2016). Strength and weaknesses of sensor technology. Retrieved December 25, 2020 from https://www.researchgate.net/figure/Strengths-and-Weaknesses-of-Sensor-Technologies_tbl2_265424349

[5] Paula Martinez. (2017). Advantages and disadvantages of sensor taps. Retrieved January 10, 2021 from https://www.intelligenthanddryers.com/blog/benefits-sensor-taps-problems

[6] N Barbour & G Schmidt. (2001). IEE Sensor Journal. The Draper Technology Digest MA Cambridge: Draper Laboratory vol. 3 pp. 5-13 1999

[7] James. (2018) Different types of shower. Retrieved December 20, 2020, from https://www.bellabathrooms.co.uk/blog/different-types-of-showers/

[8] HY Wu & S Lee. (1998). Solid-state bonding of iron-based alloys, steel–brass, and aluminum alloys. Journal of Material Processing Technology Volume 75, Issues 1-3. Elsevier Science S.A

[9] Shu-Jiang Jiang, Shun-Zhong Yao, De-Wen Liu, Yu-Rong Cai & Min Xu. (2019). Development and Application of Light Steel-Wood Plastic Structure. 10.12783/dtcse/cmso2019/33610

[10] Claus Dieter ZiegerNiclas Henning ZiegerGuenther Buzzi, Multiple proportion delivery systems and methods, US patents, USOO958.1255B2, Feb. 28, 2017.

[11] William R. Shrewsbury, Disability shower with air purge system, Patent number: 9173528, United States Patent and Trademark Office (USPTO), November 3, 2015.

[12] Tareq Khan, "Design of an Autonomous Smart Shower With Sensors and Actuators", International Journal of Embedded Systems and Applications, DOI: 10.5121/ijesa.2018.8301

[13] Mohamad, M. A. H., Yahaya Ramli, Mohamad Fadzli Amin, M. F. A., Ismail, M. S., & Ismail, M. F. (2022). Design of Automatic Sensor Equipped Shower. Multidisciplinary Applied Research and Innovation, 3(2), 463–472.

[14] https://www.deltafaucet.ca/bathroom/product/T27967 (visited on 13-04-2024).

[15] https://www.hydrao.com/en/ (visited on 13-04-2024).

[16] https://www.amazon.in/WaterHawk-Smart-Shower-Temperature-Display/dp/B01C47Q2S0, (visited on 13-04-2024).

[17] https://www.amazon.com/YOO-MEE-Thermometer-Handheld-Fahrenheit-Accessories/dp/B015MM9RIG, (visited on 13-04-2024).

[18] https://www.f6s.com/company/evadrop-theworldsfirstsmartshower#about, (visited on 13-04-2024).

[19] https://www.amazon.in/Moen-Connected-Bathroom-Controller-TS3302TB/dp/B01MY07CZG?th=1, (visited on 13-04-2024).

[20] Tareq Khan, "Design Of An Autonomous Smart Shower With Sensors And Actuators", International Journal of Embedded Systems and Applications (IJESA), Vol 8, No.1/2/3,September 2018, pp:1-17, DOI : 10.5121/ijesa.2018.8301.

A Novel Model for Recognising Handwritten Devanagari Numerals using Machine Learning

Ch.Prathima
School of Computing
Mohan babu university (formerly Sree Vidyanikethan Engineering College),
Tirupati, India
chilukuriprathi@gmail.com

Ramprakash Reddy Arava
Assistant Professor
Department of CSE
KSRM College of Engineering
Prakash.Cse.502@gmail.com

K Sevitha
UG Scholar, Department of CSSE
*Sree Vidyanikethan Engineering College,*Tirupati, India
sevithak03@gmail.com

G Manikanth
UG Scholar, Department of CSSE
*Sree Vidyanikethan Engineering College,*Tirupati, India
manikanth2511@gmail.com

D Vinay
UG Scholar, Department of CSSE
*Sree Vidyanikethan Engineering College,*Tirupati, India
vinaydeenadayal@gmail.com

C Surya
UG Scholar, Department of CSSE
*Sree Vidyanikethan Engineering College,*Tirupati, India
chittemsurya0095@gmail.com

***Abstract*— This research study performs a comprehensive comparative analysis aimed at developing effective machine learning models for classifying handwritten Devanagari numerals. The research focuses on evaluating the performance of various models to determine the most accurate classification approach. Initiating with data pre-processing, including feature extraction and normalization, the data is prepared for model training and assessment. A diverse range of machine learning models, from traditional methods like Support Vector Machines (SVM) to advanced techniques such as Random Forests, K-Nearest Neighbors, and Convolutional Neural Networks, are considered for the comparative analysis, ensuring a thorough assessment of classification capabilities. Cross-validation techniques are employed during model training and testing to enhance reliability. Statistical tests are utilized to assess the performance variations among models, enhancing the robustness of the analysis. Visual representations of performance metrics and comparison results offer clear insights. This research study aims to identify the most suitable machine learning model for handwritten Devanagari numeral classification, potentially advancing character recognition systems and linguistic applications.**

Keywords*— Handwritten character recognition, Machine learning, Support Vector Machine(SVM), Data augmentation, Transfer learning, Digital document processing

I. INTRODUCTION

Text recognition involves a computer's ability to comprehend written text, utilizing machine learning algorithms and frameworks for the interpretation and assessment of textual content. In this study, the dataset was generated by using a paint application. Input images were created from numbers 0 to 9 using an oil brush, chosen for its versatility in accurately predicting the digits. Recognizing handwritten Devanagari numerals poses a formidable challenge, given the script's complexity. This study introduces a novel approach leveraging machine learning techniques, particularly Support Vector Machines(SVMs)[11], K-Nearest-Neighbours (KNN), for feature extraction and classification. The proposed model is trained and tested on a substantial dataset, emphasizing robustness and generalization across diverse writing styles.[1] To tackle limited training data and variations, this study employs data augmentation and transfer learning. Handwriting recognition is a crucial process in computer science, finding applications in Optical character recognition (OCR) systems, signature authentication and document scanning . The evolution of digit categorization is evident in its application in banking for fraud prevention, healthcare for patient record digitization, and government for information extraction from official documents.[2] This study reflects the adaptability and significance of handwriting recognition in addressing contemporary challenges. To generate a dataset, our initial step involves assigning a value of 1 to the selected region and 0 to the background. Consequently, our dataset will comprise solely of these two values: 0 and 1. It's important to note that pixel values typically span from 0 to 255. Conventionally, 0 signifies black while 255 denotes white in most scenarios.

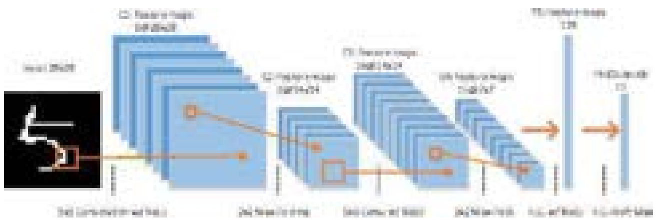


Fig. 1. System Architecture

II. RELATED WORKS

Ajitha K (2023) focused on centered around evaluating various machine learning models to determine their effectiveness in classifying handwritten digits in the Devanagari script. [1] The CNN model proposed in this study achieved an impressive accuracy of 99.522%. Furthermore, it surpassed the performance of other CNN models evaluated for Devanagari numerals classification.

K. Kancharla (ICACAT), Dec. 2018 focused on Handwritten Signature Recognition stands as a crucial behavioral biometric utilized across numerous identification and authentication applications. It predominantly operates through two distinct methods: on-line and off-line recognition. On-line recognition entails a dynamic approach, incorporating parameters such as writing pace, variations in stylus direction, and the count of pen ups and pen downs throughout the signature-writing process.

SM Shamim [3] It focuses on recognizing a digit through minist dataset. They concluded that highest precision, recall and f1score for Support Vector machine.

M. M. Al-Taee [4] focused on Handwritten recognition has garnered significant interest within the fields of pattern recognition and image processing over recent decades [13].

S. Gupta [5] In this paper, A comparison of different approaches for recognizing handwritten numerals utilizing solely the MNIST database is conducted. Over time, numerous researchers have introduced novel techniques for digit recognition, contributing to the enhancement of our daily routines [12].

S. Boroojerdi ‘‘Handwritten multi-digit recognition with machine learning’’ in (IETC). Offline recognition of handwritten digits poses a persistent challenge, with solutions thus far remaining only partially effective. This study delves into the realm of offline handwritten multi-digit recognition, employing three distinct algorithms—Decision Trees, Multilayer Perceptrons, and Random Forest—utilizing the MNIST dataset. Among these, Random Forest emerged as the most promising, boasting an impressive 96% accuracy alongside reasonable runtime efficiency.

III. PROPOSED WORK

The suggested endeavor to employ machine learning for the identification of Manually scripted Devanagari numerals constitutes a significant advancement in the realm of character recognition [14] particularly within the intricate context of challenging scripts like Devanagari.

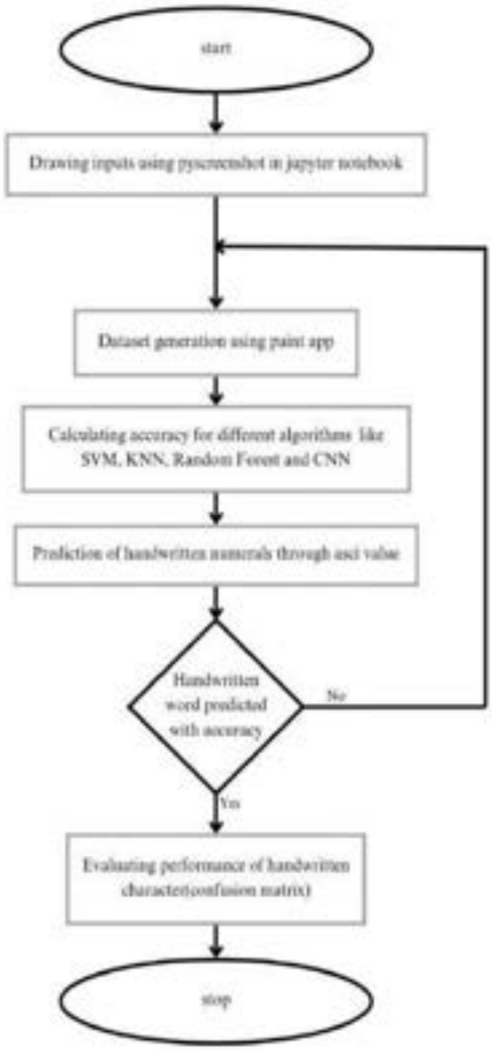


Fig.2. Process of handwritten recognizing

This study centers on assessing the appropriateness of various machine learning frameworks for classifying Manually written numerals using the Devanagari script. In this research, the study will evaluate various models, including K-Nearest-Neighbors (KNNs), Support Vector Machine (svms), Random Forest (rf), and Convolutional Neural Networks(cnn). Surprisingly, the proposed KNN model is implemented on F1 score.

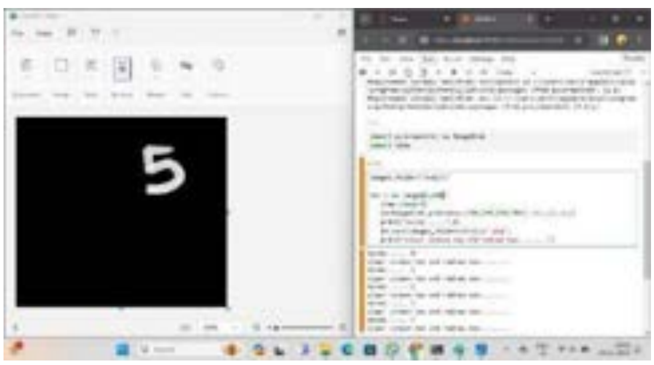


Fig. 3. Dataset generation using paint

KNN Algorithm:

The K-Nearest Neighbors (KNN) algorithm, a simple yet potent supervised machine learning technique, finds application in tasks encompassing both classification and regression. [6] At its essence, it entails producing predictions based on either the predominant class or the average of the KNN data points within the feature space. Dataset creation is achieved using a paint application. General formula is

$$\sum_{i=1}^{k} |xi - yi|$$

CSV File Setup: This work begins by importing the necessary libraries: cv2 for computer vision operations, "csv" for handling CSV files, and glob for file path pattern matching [8]. It sets up the header for the CSV file, where each image will be represented as a row with a label and pixel values. A CSV file named 'demo1.csv' is created, and the header is written to it.

Image Processing and CSV Data Generation: For each digit label (0 to 9), the code uses glob to get a list of image file paths in the corresponding subdirectories. For each image, it reads the image, converts it to grayscale, applies a Gaussian blur, and resizes it to 28x28 pixels. Based on the pixel value it predicts the digit. The label and flattened pixel values are then appended to the 'demo1.csv' file.

Data Loading and Shuffling: The code imports pandas for data manipulation and shuffle from scikit-learn to shuffle the dataset. It reads the CSV file into a Pandas Data Frame and shuffles the rows randomly.

Data Splitting: The features (X) comprise all columns excluding the "label" column, while the target variable (Y) is specifically the "label" column. The dataset undergoes division into training and testing sets facilitated by the 'train_test_split' function from scikit-learn.

Support Vector Machine algorithm:

The dataset is formed by employing a paint application to produce images, from which a CSV file is crafted using pixel values. Supervised machine learning tasks, specifically classification or regression, can be effectively addressed using SVM as an algorithm. SVM exhibit versatility in handling both linear and non-linear classification as well as regression tasks through the utilization of suitable kernel functions. Typical kernel options encompass linear, polynomial, and radial basis function (RBF). Hyper plane equation is

$$W.X+b=0$$

Importing Libraries: This work starts by importing necessary libraries. Joblib is imported for saving the trained SVM model, and SVC (Support Vector Classification) is imported from scikit-learn for implementing the Support Vector Machine classifier.[7]

Creating an SVM Classifier: An SVM classifier is created using the SVC class with a linear kernel. The random state parameter is set to 5 for reproducibility, ensuring that the same sequence of random numbers is generated if the code is run multiple times.

Training the Classifier: The created SVM classifier ('classifier') is trained using the `fit` method with the training data and where train_x represents the features (pixel values) and train_y represents the corresponding labels.

Saving the Model: The trained SVM model is then saved using the joblib.dump function. The model is saved in a file named "digit_recognizer" within a directory named model. This allows you to persist the trained model for later use without having to retrain it every time.

Random Forest algorithm:

A Random Forest is a popular machine learning algorithm used for both classification and regression tasks.[9] This algorithm focusses on following steps:

Importing Necessary Libraries: Here, we import the 'RandomForestClassifier' class from the ensemble module of scikit-learn, which contains ensemble-based learning algorithms, including Random Forest. We also import accuracy_score from 'sklearn.metrics' which will be used to calculate the accuracy of the classifier.

Creating the Random Forest Classifier: Creates a RF classifier instance, with the 'n_estimators' parameter determining the number of trees within the forest.

Training the Classifier: The Random Forest classifier is trained on the provided training data. Usually, 'train_x' comprises the features of the training dataset, while 'train_y' contains the corresponding labels.

Making Predictions: In this step, the Random Forest classifier, which has been trained previously, is employed to predict outcomes on the test dataset denoted as 'test_x'. The resulting predictions are then stored in the variable 'y_pred'.

Calculating Accuracy and Confusion Matrix: In this context, we're presenting both the confusion matrix and accuracy metrics. The confusion matrix serves as a crucial instrument for assessing the performance of a classification model on a test dataset with known true values. It provides a comprehensive breakdown, revealing the counts of tp, fp, tn, and fn, thereby offering significant insights into the classifier's efficacy.

CNN algorithm:

Importing Necessary Libraries: Keras are imported. Keras is TensorFlow's high-level neural networks API, which provides a convenient way to define and train deep learning models.[10].

Loading and Preprocessing the Data: Own dataset is loaded and then preprocessed will result Reshaping the input images to have a single channel (grayscale) and adjusting pixel values to be between 0 and 1.

Compiling the Model: The model undergoes compilation using the Adam optimizer, categorical cross-entropy loss function (ideal for multi-class classification tasks), and accuracy serves as the metric for evaluation.

Training the Model: The model undergoes training on the provided training data ('x_train', 'y_train') for a total of 5 epochs, utilizing a batch size of 64 for optimization. Validation data is designated through the `validation_split` parameter during the training process.

Evaluating the Model: The trained model is evaluated on the test data (‘x_test’, ‘y_test’). Predicted probabilities are obtained using `predict` method, then converted to predicted labels. True labels are also converted from one-hot encoded format. Finally, accuracy is calculated and printed.

IV. RESULTS AND DISCUSSION

The Support Vector Machine (SVM) model achieves the highest accuracy at 99.5%. Prediction analysis can be conducted through the Paint app using essential packages like joblib, pyscreenshot, numpy, time, and cv2. The SVM model predicts the digit by utilizing the predict method with the prepared input array, displaying the predicted digit on the console. Results are presented in the form of the original image, annotated with the predicted digit using "cv2.putText". A window named "Result" is created, showcasing the annotated image. The program awaits a key press for a specified duration (10 seconds in this case) using "cv2.waitKey". If the pressed key is Enter (13), the loop continues; otherwise, it breaks out of the loop. Upon termination of the loop, typically triggered by pressing Enter, the OpenCV windows are closed using "cv2.destroyAllWindows()".

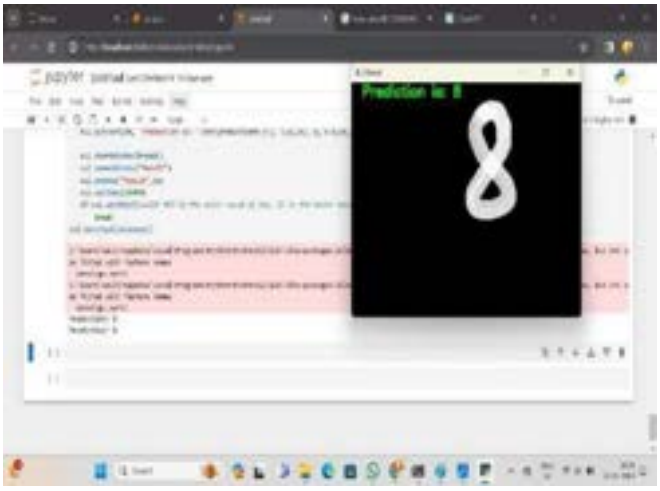


Fig.4. prediction of 8

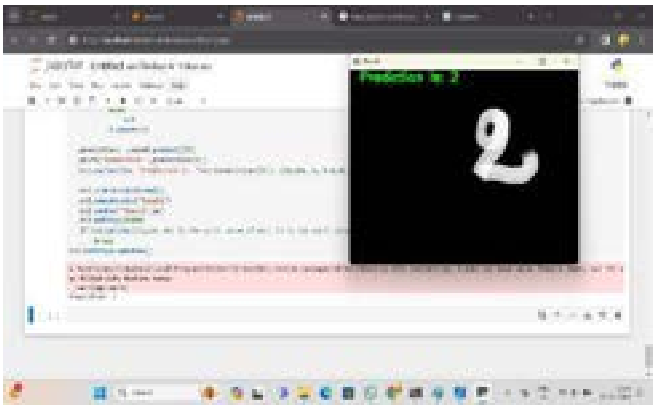


Fig. 5. Prediction of 2

By leveraging the inherent strengths of convolutional neural networks (CNNs), the proposed model surpasses both SVM and KNN. CNNs excel in capturing local patterns and relationships, demonstrating characteristics such as translational invariance and non-linearity. When dealing with a relatively compact dataset and the straightforward nature of grayscale images (with dimensions of 28 × 28 and 32 × 32), the diminished accuracy observed in higher-capacity models can be linked to their heightened capacity to capture intricate patterns.

However, achieving optimal performance on smaller datasets with such models may necessitate additional data or thorough hyperparameter tuning. The heightened complexity of such models could pose challenges in effectively learning from the limited information available, consequently leading to a marginally reduced accuracy. Figures 4, 5, and 6 illustrate digit predictions and compare various frameworks along with their corresponding accuracies.

TABLE. 1. COMPARISON OF ACCURACY FOR USED FRAMEWORKS

Algorithm	Accuracy
SVM	0.995
KNN	0.955
Random Forest	0.985
CNN	0.987

```
Confusion Matrix:
[[22  0  0  0  0  0  0  0  0  0]
 [ 0 14  0  0  0  1  0  0  0  0]
 [ 0  0 24  0  0  0  0  0  0  0]
 [ 0  0  0 22  0  0  0  0  0  0]
 [ 0  2  0  0 16  0  2  0  0  0]
 [ 0  0  0  1  0 14  0  0  0  0]
 [ 0  1  0  0  0  0 25  0  1  0]
 [ 0  0  0  0  0  0  0 24  0  0]
 [ 0  0  0  0  0  0  0  0 14  0]
 [ 0  1  0  0  0  0  0  0  0 16]]
```

Fig. 6. Confusion Matrix of KNN

```
Confusion Matrix:
[[22  0  0  0  0  0  0  0  0  0]
 [ 0 15  0  0  0  0  0  0  0  0]
 [ 0  0 24  0  0  0  0  0  0  0]
 [ 0  1  2 19  0  0  0  0  0  0]
 [ 0  0  0  2 18  0  0  0  0  0]
 [ 0  0  0  0  0 15  0  0  0  0]
 [ 0  1  0  1  0  0 25  0  0  0]
 [ 1  0  0  0  0  0  0 23  0  0]
 [ 0  0  0  0  0  0  0  0 14  0]
 [ 0  0  0  0  0  0  0  0  0 17]]
```

Fig. 7. Confusion Matrix of SVM

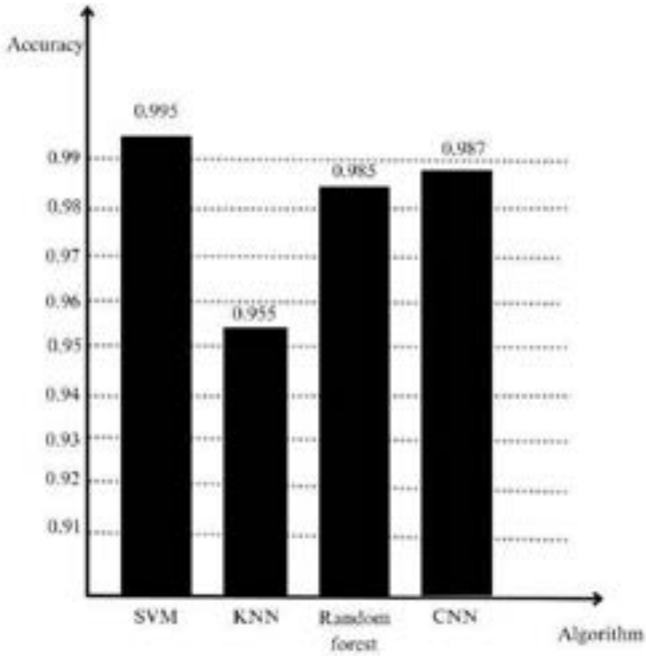


Fig. 8. Comparison of Algorithm Accuracies

V. CONCLUSION

In conclusion, the recognition of handwritten Devanagari numerals using machine learning algorithms presents an opportunity to apply advanced techniques in a multilingual and script-specific context. Through an extensive literature review, it is evident that the application of k-Nearest Neighbors (KNN) has shown promising results in character recognition tasks, including those involving the Devanagari script. Comparative analysis of various algorithms, including but not limited to SVM, Random Forest, k-Nearest Neighbors and CNN have revealed their unique strengths and limitations in recognizing handwritten Devanagari numerals. Recognizing handwritten Devanagari numerals has provided valuable insights into the efficiency of different models. The recommended model, leveraging the strengths of Support Vector Machines(SVMs), KNN and Random Forest (RF). The observed lower accuracy in higher-capacity models suggests the need for careful consideration of dataset size and simplicity. Despite the trails posed by the relatively small dataset and the simplicity of grayscale images, our findings emphasize the importance of balancing model complexity with the available data, offering valuable considerations for future endeavors in the domain of handwritten Devanagari numeral recognition.

VI. FUTURE WORK

Expansion to Full Devanagari Script: Extending the recognition model to cover the entire Devanagari script, including consonants, vowels, and compound characters, to create a comprehensive recognition system for the script as a whole. Multilingual Recognition: Exploring the feasibility of adapting the recognition model to recognize numerals in other Indic scripts, such as Bengali, Gujarati, or Tamil, to create a multilingual character recognition system.

REFERENCES

[1] Performance Comparison of Machine Learning Models for Handwritten Devanagari Numerals Classification Agastya Gummaraju; Ajitha K. B. Shenoy; Smitha N. Pai IEEE Access Year: 2023 | Volume: 11 | Journal Article | Publisher: IEEE

[2] K. Kancharla, V. Kamble, and M. Kapoor, ''Handwritten signature recognition: A convolutional neural network approach,'' in Proc. Int. Conf. Adv. Comput. Telecommun. (ICACAT), Dec. 2018.

[3] Handwritten digit recognition using machine learning algorithms SM Shamim, MBA Miah, A Sarker, M Rana, A Al Jobair Global Journal Of Computer Science And Technology, 2018•researchgate.net

[4] M. M. Al-Taee, S. B. H. Neji, and M. Frikha, ''Handwritten recognition: A survey,'' in Proc. IEEE 4th Int. Conf. Image Process., Appl. Syst. (IPAS), Dec. 2020, pp. 199–205.

[5] A. Shrivastava, I. Jaggi, S. Gupta, and D. Gupta, ''Handwritten digit recognition using machine learning: A review,'' in Proc. 2nd Int. Conf. Power Energy, Environ. Intell. Control (PEEIC), Oct. 2019, pp. 322–326.

[6] S. Boroojerdi and G. Rudolph, ''Handwritten multi-digit recognition with machine learning,'' in Proc. Intermountain Eng., Technol. Comput. (IETC), May 2022, pp. 1–6.

[7] W. Li, X. He, C. Tang, K. Wu, X. Chen, S. Yu, Y. Lei, Y. Fang, and Y. Song, ''Handwritten numbers and English characters recognition

[8] Liu, C., Fujisawa, H., "Handwritten digit recognition: Benchmarking of state-of-the-art techniques," Pattern Recognition, 2003.

[9] Fialoke, S., "Predicting Digits from their handwritten images, (16.10.2020), http://suruchifialoke.com/2017-06-15- predicting-digits_tensorflow

[10] Arica, N., Vural, Y., F.T., "An overview of character recognition focused on offline handwriting," IEE Transactions on Systems Man and Cybernetics, 2001

[11] Handwritten digit recognition using machine learning algorithms SM Shamim, MBA Miah, A Sarker, M Rana, A Al Jobair Global Journal Of Computer Science And Technology, 2018•researchgate.net

[12] Handwritten digits identification using MNIST database via machine learning models B Gope, S Pande, N Karale, S Dharmale, P Umekar, IOP conference series: materials science and engineering, 2021.

[13] Impact of Deep Learning on Localizing and Recognizing Handwritten Text in Lecture Videos Media, L.H., Ramani, KInternational Journal of Advanced Computer Science and Apllications 2021, 12(4), pp. 336-344.

[14] Muppalaneni, N.B. (2021). Deep Learning Approach for Prediction of Handwritten Telugu Vowels. In: Reddy, A., Marla, D., Favorskaya, M.N., Satapathy, S.C. (eds) Intelligent Manufacturing and Energy Sustainability. Smart Innovation, Systems and Technologies, vol 213. Springer, Singapore. https://doi.org/10.1007/978-981-33-4443-3_35.

[15] K. N. K. Reddy, K. N. V. S. Divyanjali, L. J. Royal, M. H. Chandana and D. M. Krishna, "A Novel Model for Prediction of Next Word using Machine Learning," 2023 7th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2023, pp. 393-400, doi: 10.1109/ICICCS56967.2023.10142737.

Agile Management Tools: Technological Evaluations and Future Archetype

Nandula Kowshik Sravan
Department of Computer Science and Engineering
(of Affiliation)
Koneru Lakshmaiah Education Foundation
Vaddeswaram,AP, India
sravankoushik112@gmail.com

Dr.Veera Ankalu Vuyyyru
Department of Computer Science and Engineering
(of Affiliation)
Koneru Lakshmaiah Education Foundation
Vaddeswaram,AP, India
veerankalu14@kluniversity.n

Pawankalyan Gottapu
Department of Computer Science and Engineering
(of Affiliation)
Koneru Lakshmaiah Education Foundation
Vaddeswaram,AP, India
pawankalyan3615@gmail.com

Anirudh Theeda
Department of Computer Science and Engineering
(of Affiliation)
Koneru Lakshmaiah Education Foundation
Vaddeswaram,AP, India
anirudh.t33@gmail.com

Lehan Peddineni
Department of Computer Science and Engineering
(of Affiliation)
Koneru Lakshmaiah Education Foundation
Vaddeswaram,AP, India
lehanpeddineni@gmail.com

***Abstract*— This research paper investigates the dynamic landscape of Agile Project Management Tools, encompassing their historical evolution, current functionalities, and future trends. The exploration begins with an introduction delving into the background, rationale, and research objectives. The Evolution section unfolds the historical development, technological advancements, and organizational adoption trends shaping Agile tools. The focal point shifts to Major Agile Project Management Tools, featuring in-depth overviews and functionalities of prominent platforms such as Jira, Trello, Asana, and Azure DevOps. A comprehensive analysis of the tools' functionalities covers user stories, sprint planning, task boards, burndown charts, collaboration tools, version control integration, automation, reporting, and scalability. In the subsequent section, this study predicts the future archetype of agile project management tools. Emerging trends, anticipated technological advancements, and key considerations for future development are explored, providing insights into the potential trajectory of these tools. Finally, the study summarizes key findings, discusses implications for the industry, and offers recommendations for future research. The detailed exploration of each topic contributes to a holistic understanding of Agile Project Management Tools, making this research a valuable resource for organizations navigating the complexities of modern project management methodologies.**

Keywords—Agile project management, Tools Evolution, Technological Advancements, Organizational Adoptions, Azure DevOps.

I. INTRODUCTION

In recent years, the realm of project management has undergone a transformative shift with the widespread adoption of Agile methodologies. Originating from the software development domain, Agile has rapidly expanded its influence across diverse industries, emphasizing iterative and flexible approaches to project execution. This shift towards Agile practices has prompted a parallel evolution in the tools that support project management processes. As organizations strive to enhance adaptability, collaboration, and responsiveness in their projects, the need for sophisticated Agile project management tools has become increasingly pronounced.

The rationale for analyzing the technological evolution and future archetype of Agile project management tools is highlighted by the critical role these tools play in modern project management practices. Traditional project management approaches often struggle to cope with the dynamic and iterative nature of contemporary projects, making Agile methodologies and their associated tools indispensable. As Agile continues to gain prominence, organizations are faced with the challenge of selecting tools that align seamlessly with their workflows, methodologies, and objectives. Understanding the historical development, current functionality, and future trends of these tools is imperative for both practitioners and researchers. This study aims to bridge the knowledge gap by providing a comprehensive exploration of the tools that have shaped the Agile landscape, offering insights into their functionality, and forecasting the trajectory of their future developments.

The primary objectives of this research encompass unraveling the historical evolution of Agile project management tools, elucidating the functionalities of major tools such as Jira, Trello, Asana, and Azure DevOps, and projecting the future archetype of these tools in the context of emerging technological trends. By achieving these objectives, this research seeks to contribute valuable knowledge that can guide organizations in making informed decisions regarding tool selection, fostering efficiency and agility in their project management endeavors. Ultimately, the research aims to empower both practitioners and academics with a deeper understanding of the past, present, and future of Agile project management tools, shaping the discourse around effective project management in the evolving landscape of software development and beyond.

II. EVALUTION OFAGILE PROJECT MANAGEMENT TOOLS

A. Historical Development:

The historical development of Agile project management tools traces back to the inception of Agile methodologies in the early 2000s. As a response to the limitations of traditional, plan-driven project management, the Agile Manifesto was introduced, emphasizing values such as individuals and interactions over processes and tools. Initially, tools were rudimentary, with a focus on supporting Agile principles like collaboration and adaptability. Early solutions included simple task boards and spreadsheets designed to facilitate communication within collocated development teams.

Over time, the landscape evolved with the increasing complexity of projects and the dispersion of development teams. The first generation of Agile tools, emerging around the mid-2000s, introduced features like user story management, sprint planning, and basic reporting. Tools like Rally (now CA Agile Central) and VersionOne gained prominence during this phase, catering to the needs of Scrum and Agile practitioners.

B. Technological Advancements:

Technological advancements have played a pivotal role in shaping the capabilities of Agile project management tools. The advent of cloud computing in the late 2000s significantly influenced the accessibility and scalability of these tools. Cloud-based Agile solutions allowed teams to collaborate seamlessly, irrespective of geographical locations, fostering a more global and distributed approach to software development.

Integration with version control systems, especially Git, became a standard feature in Agile tools, streamlining the development process by linking code changes to specific user stories or tasks. Automation of repetitive tasks, such as continuous integration and deployment, gained prominence, enhancing the efficiency and speed of software delivery.

The utilization of artificial intelligence and machine learning in Agile tools has been a recent advancement, enabling predictive analytics, smarter planning, and improved decision-making. These technologies enhance the tool's ability to forecast project timelines, identify potential risks, and suggest optimizations.

C. Adoption Trends in Organizations:

The adoption of Agile project management tools has become widespread across organizations of varying sizes and industries. Initially embraced by software development teams, Agile methodologies and tools have transcended their origins, finding applications in marketing, finance, and even non-technical domains.

Organizations are increasingly recognizing the benefits of Agile tools in promoting collaboration, flexibility, and responsiveness. The shift towards Agile methodologies has been driven by a desire to deliver value incrementally, respond to changing requirements, and enhance overall project visibility.

In recent years, a trend toward integrated platforms has emerged, where Agile project management tools seamlessly connect with other tools in the software development lifecycle, such as issue tracking, testing, and collaboration tools. This integration aims to provide a unified experience for development teams, streamlining workflows and fostering cross-functional collaboration.

In summary, the evolution of Agile project management tools has been marked by a progression from basic, collocated team solutions to sophisticated, cloud-based platforms with advanced features and integrations. The adoption trends underscore the growing recognition of Agile methodologies as a cornerstone for successful project management across diverse organizational contexts.

III. MAJOR AGILE PROJECT MANAGEMENT TOOLS

A. Jira

Description:

Jira, developed by Atlassian, is a robust Agile project management tool designed for efficient issue tracking and project management. Over time, it has evolved into a comprehensive platform supporting various Agile methodologies. Known for its adaptability, Jira fosters collaboration, transparency, and real-time reporting throughout the software development lifecycle.

Functionality:

Jira's functionality covers user story and backlog management, sprint planning, and robust reporting. Customizable workflows allow teams to tailor processes, while burndown charts provide insights into project progress.

Image Information:

Include a screenshot of a Jira dashboard displaying a Kanban board or Scrum board with user stories, tasks, and progress indicators. This visual representation will offer insights into Jira's interface and usage.

Fig.1. Jira

B. Trello

Description:

Trello, an intuitive tool acquired by Atlassian, simplifies Agile project management with its visual approach. Using boards, lists, and cards, Trello enables teams to organize tasks and collaborate seamlessly. It's particularly well-suited for smaller teams or projects with straightforward workflows.

Functionality:

Trello's functionality revolves around task management through a Kanban-style board. Users create boards for projects, populate them with cards representing tasks, and move cards through customizable lists. Collaboration is enhanced through comments, attachments, and checklists.

Image Information:

Include a screenshot of a Trello board showcasing tasks in different stages. This image will visually demonstrate how Trello organizes work and encourages collaboration.

Fig.2.Trello

C. Asana

Description:

Asana is a versatile project management tool designed for both traditional and Agile methodologies. With a user-friendly interface, it offers features for task management, project timelines, and real-time collaboration. Asana's simplicity and adaptability make it a suitable choice for diverse project management needs.

Functionality:

Asana's functionality includes task assignment, due dates, customizable workflows, and project timelines. Supporting both list and board views, Asana provides flexibility for different team preferences. It emphasizes simplicity while offering powerful tools for effective project management.

Image Information:

Include a screenshot of an Asana project showcasing task lists or a timeline view. This image will illustrate Asana's adaptability and its ability to cater to various project management requirements.

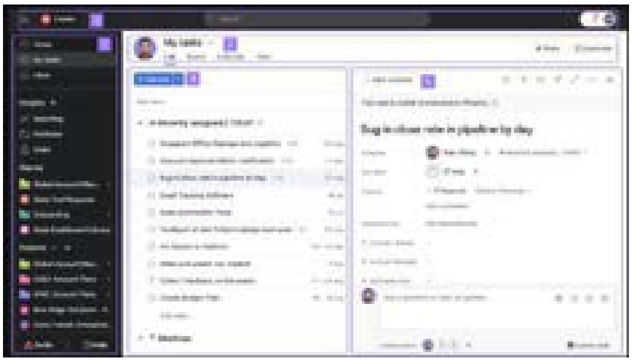


Fig.3.Asana

D. Azure DevOps

Description:

Azure DevOps, developed by Microsoft, is a comprehensive tool suite integrating with Visual Studio. Covering the entire software development lifecycle, Azure DevOps combines Agile project management with version control, build automation, and release management.

Functionality:

Azure DevOps' functionality spans backlog management, sprint planning, code repositories, and build/release pipelines. Its integration with Azure Boards provides a seamless experience for Agile project management. The tool supports collaboration between development and operations teams, promoting a DevOps culture.

Image Information:

Include a screenshot of an Azure DevOps dashboard displaying Agile boards, build pipelines, or release pipelines. This visual representation will highlight the tool's holistic approach to Agile and DevOps practices.

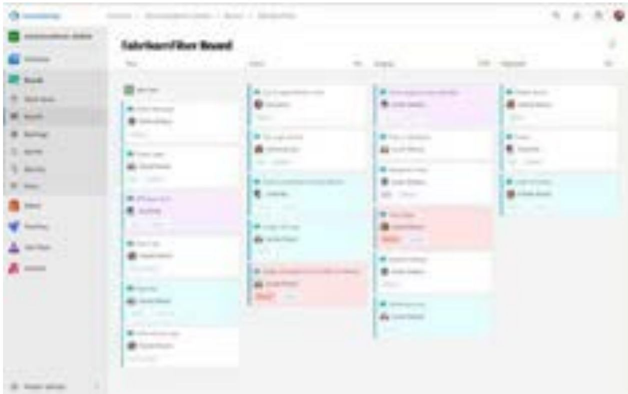


Fig.4.Azure Devops

IV. FUNCTIONALITY AND FEATURES OF AGILE PROJECT MANAGEMENT TOOLS

A. User Stories and Backlog Management:

Agile project management tools serve as robust platforms for user stories and backlog management. Teams leverage these tools to meticulously create, prioritize, and manage user stories, ensuring a systematic and organized backlog. This functionality is fundamental for maintaining clarity on project requirements and establishing a solid foundation for development cycles.

B. Sprint Planning:

Sprint planning, a pivotal aspect of Agile methodologies, is seamlessly facilitated by Agile project management tools. These tools empower teams to define sprint goals, allocate user stories, and establish realistic timelines. The collaborative nature of sprint planning sessions within these tools ensures that the entire team is aligned on priorities and committed to achievable workloads for the upcoming sprint.

C. Task Boards/Kanban Boards:
Task boards, often presented in a Kanban-style format, are integral components of Agile tools. These visual boards provide a snapshot of work items, their current statuses, and the flow of tasks throughout the development process. By enhancing transparency, task boards enable teams to easily track progress and manage work in a highly visual and intuitive manner.

D. Burndown Charts:
Burndown charts, a staple in Agile project management, play a crucial role in tracking project progress. Agile tools generate dynamic burndown charts that graphically depict the completion of tasks over time. These charts offer invaluable insights into the team's velocity, helping project managers make informed decisions and predict whether project milestones will be met.

E. Collaboration Tools:
Collaboration lies at the heart of Agile methodologies, and Agile project management tools are equipped with robust collaboration features. These tools provide dedicated spaces for discussions, file sharing, and real-time communication among team members. This fosters a collaborative environment where ideas flow freely, and the entire team remains connected throughout the project lifecycle.

F. Integration with Version Control Systems:
Integration with version control systems, such as Git, is a fundamental feature of Agile project management tools. This integration ensures the seamless alignment of code changes with specific tasks or user stories. Agile tools facilitate a cohesive development and project management process, where version control plays a pivotal role in maintaining code integrity.

G. Automation of Repetitive Tasks:
Agile tools streamline workflows by automating repetitive tasks, thereby enhancing overall efficiency. From task creation to status updates based on predefined criteria, these tools significantly reduce manual effort. Automation not only accelerates processes but also minimizes the risk of human error, allowing teams to focus on higher-value activities.

H. Reporting and Analytics:
Robust reporting and analytics capabilities are inherent strengths of Agile project management tools. Teams can generate comprehensive reports on sprint progress, team performance, and overall project health. Analytics features provide valuable insights for retrospective analysis, aiding teams in identifying strengths, areas for improvement, and trends over time.

I. Scalability:
Scalability is a cornerstone feature of Agile project management tools, catering to the dynamic needs of diverse projects. These tools exhibit the flexibility to adapt to projects of varying sizes and complexities, from small, focused endeavors to large-scale enterprise initiatives. The scalability of these tools ensures that they can accommodate expanding user bases, evolving features, and intricate work structures without compromising performance or functionality.

V. FUTURE ARCHETYPE OF AGILE PROJECT MANAGEMENT TOOLS

A. Emerging Trends:
The future archetype of Agile project management tools is anticipated to witness the emergence of several key trends. One such trend is the increasing integration of artificial intelligence (AI) and machine learning (ML) capabilities. Agile tools are expected to leverage these technologies to enhance predictive analytics, automate repetitive tasks, and provide intelligent insights for better decision-making. Another notable trend is the emphasis on enhanced user experience through intuitive interfaces and user-centric design, ensuring that teams can interact seamlessly with the tools. Additionally, the rise of remote work is likely to drive the development of features that cater to distributed teams, fostering collaboration and communication across diverse geographical locations.

B. Anticipated Technological Advancements:
The future of Agile project management tools will be shaped by anticipated technological advancements. Blockchain technology is poised to make an impact by enhancing security and traceability within project management processes. Decentralized and transparent ledgers can provide a secure foundation for tracking changes and ensuring the integrity of project data. Furthermore, the evolution of augmented reality (AR) and virtual reality (VR) may introduce innovative ways for teams to visualize and interact with project data, fostering immersive and collaborative experiences. As technology continues to advance, tools are likely to leverage these innovations to create more dynamic and engaging project management environments

C. Key Considerations for Future Development:
In shaping the future archetype of Agile project management tools, developers and stakeholders need to consider several key aspects. Firstly, the tools should prioritize interoperability, allowing seamless integration with a diverse ecosystem of software and services. This ensures that teams can leverage their preferred tools while maintaining a cohesive project management environment. Secondly, a focus on scalability and flexibility is paramount. The tools should be designed to accommodate the evolving needs of projects, from small-scale initiatives to large enterprise-level endeavors. Finally, the emphasis on security and data privacy is crucial, especially with the increasing reliance on cloud-based solutions. Future tools must implement robust security measures to safeguard sensitive project information and ensure compliance with data protection regulations.

As Agile project management tools evolve, these considerations will play a pivotal role in shaping their functionality and relevance in an ever-changing technological landscape. The future archetype is poised to be dynamic, intelligent, and adaptable, catering to the evolving needs of Agile teams and the broader project management community.

VI. Conclusion

In conclusion, this study has reviewed the Agile Project Management Tools, examining their evolution, major functionalities, and future archetype. We explored tools such as Jira, Trello, Asana, and Azure DevOps, detailing their features and practical applications. The research highlighted the significance of user stories, sprint planning, task boards, burndown charts, collaboration tools, version control integrations, automation, reporting, and scalability in Agile project management. The findings of this research carry significant implications for the industry. Agile project management tools play a pivotal role in enhancing collaboration, improving workflow efficiency, and ensuring project success. As the industry embraces Agile methodologies, understanding the functionalities and features of these tools becomes crucial for organizations aiming to adapt to dynamic project management practices. The implications extend to the broader tech industry, influencing how teams collaborate, plan, and execute projects in an iterative and adaptive manner.

Moving forward, future research in this domain should explore the practical implementation and user experiences of Agile project management tools in diverse organizational contexts. Additionally, investigations into the integration of emerging technologies, such as AI, blockchain, and AR/VR, into these tools could provide valuable insights. Understanding how these tools impact team dynamics, productivity, and project outcomes in the long term would contribute to a deeper comprehension of their efficacy. Further studies could also explore the evolving cybersecurity considerations associated with the increasing reliance on cloud-based Agile solutions. This study has conveyed a comprehensive understanding of the major Agile project management tools, shedding light on their functionalities, practical applications, and the evolving technological landscape. The exploration of emerging trends, anticipated technological advancements, and key considerations for future development has equipped readers with insights into the potential trajectory of Agile tools. As organizations navigate the complexities of modern project management, the knowledge imparted in this research aims to serve as a valuable resource for informed decision-making and strategic planning in the dynamic landscape of Agile project management.

References

[1] J. M. Fernandes and M. Almeida, "Classification and Comparison of Agile Methods," 2010 Seventh International Conference on the Quality of Information and Communications Technology, Porto, Portugal, 2010, pp. 391-396, doi: 10.1109/QUATIC.2010.71.

[2] Shaikh, Mehwish & Khan, Aaiza & Siddiqui, Isma & Qureshi, Zakaullah. (2021). Comparative Analysis of Trending Agile Model Tools for Software Development Life Cycle.

[3] Raith, Florian & Richter, Ingo & Lindermeier, Robert. (2017). How Project-management-tools are used in Agile Practice: Benefits, Drawbacks and Potentials. 30-39. 10.1145/3105831.3105865.

[4] Katsma, Christian & Amrit, Chintan & Hillegersberg, Jos & Sikkel, Klaas. (2013). Can Agile Software Tools Bring the Benefits of a Task Board to Globally Distributed Teams?. 163. 163–179. 10.1007/978-3-642-40951-6_10.

[5] Azizyan, Gayane & Magarian, Miganoush & Kajko-Matsson, Mira. (2011). Survey of Agile Tool Usage and Needs. 10.1109/AGILE.2011.30.

[6] Goth, Greg. (2009). Agile Tool Market Growing with the Philosophy. Software, IEEE. 26. 88 - 91. 10.1109/MS.2009.30.

[7] December, & Behrens, Pete. Agile Project Management (APM) Tooling Survey Results.

[8] M. Ortu, G. Destefanis, M. Kassab and M. Marchesi, "Measuring and Understanding the Effectiveness of JIRA Developers Communities," 2015 IEEE/ACM 6th International Workshop on Emerging Trends in Software Metrics, Florence, Italy, 2015, pp. 3-10, doi: 10.1109/WETSoM.2015.10.

[9] Marques, Joana Ferreira and Jorge Bernardino. “Evaluation of Asana, Odoo, and ProjectLibre Project Management Tools using the OSSpal Methodology.” International Conference on Knowledge Engineering and Ontology Development (2019).

[10] Johnson, Heather. (2017). Trello. Journal of the Medical Library Association. 105. 10.5195/jmla.2017.49.

[11] Bento, Antonio & Delgado, David & Camacho-Leon, Sergio. (2023). Experimental Survey Results on Azure.DevOps Application for Management Student’s Projects. 1-5. 10.1109/IEEECONF56852.2023.10104593.

[12] Mishra, Deepti & Mishra, Alok. (2011). Complex software project development: Agile methods adoption. Journal of Software Maintenance and Evolution Research and Practice. 23. 549-564. 10.1002/smr.528

[13] Abrahamsson, Pekka & Salo, Outi & Ronkainen, Jussi & Warsta, Juhani. (2002). Agile Software Development Methods: Review and Analysis. Proc. Espoo 2002. 3-107.

[14] Lanubile, Filippo & Ebert, Christof & Prikladnicki, Rafael. (2010). Collaboration Tools for Global Software Engineering. Software, IEEE. 27. 52 - 55. 10.1109/MS.2010.39.

Style Synthesis: AI-Powered Dress Try-On Experience

Siva RamaKrishna Sani
Artificial Intelligence and Data Science Department
Lakireddy Bali Reddy College of Engineering
Mylavaram, Andhra Pradesh, India
ssrk102@gmail.com

Sai Manvitha Reddy Mallireddy
Artificial Intelligence and Data Science Department
Lakireddy Bali Reddy College of Engineering
Mylavaram, Andhra Pradesh, India
manvithamallireddy@gmail.com

Naveen Kumar Reddy Renati
Artificial Intelligence and Data Science Department
Lakireddy Bali Reddy College of Engineering
Mylavaram, Andhra Pradesh, India
renatinaveen5@gmail.com

Peela Leela Sai Surya
Artificial Intelligence and Data Science Department
Lakireddy Bali Reddy College OF Engineering
Mylavaram, Andhra Pradesh, India
peelaleelasaisurya20@gmail.com

***Abstract*—In the world of internet retail and fashion, virtual try-on systems have become a game changing innovation. This research introduces a revolutionary deep learning-based method for virtual try-on. The method first segments the clothing from the source image using segmentation techniques, then creates an interactive image to provide a virtual fitting experience, and finally combines accurate posture recognition using pretrained models with the imposition of the garment texture on the dense pose it predicted. Realtime experimenting with different clothing items, sizes, and positions is possible because to the interactive user interface. The technology is more user friendly because it can accommodate a variety of body types. This virtual try-on system, powered by deep learning, has the potential to significantly change customer involvement across a variety of industries, especially in fashion and ecommerce, by transforming the fashion retail scene.**

***Keywords*—Virtual try-on systems, Deep learning-based method, Virtual fitting, Posture Recognition, Garment texture, Dense pose, Segmentation.**

I. Introduction

The world of internet purchasing has expanded at an unparalleled rate, providing consumers with convenience and accessibility across borders. Online dress shopping has grown in popularity during this digital revolution. However, several issues and worries are raised by this increase in virtual interactions, which influence the whole online buying experience. Uncertainty about the size and fitting of clothes is one of the biggest problems for internet consumers. When it's not possible to physically try on clothing, buyers frequently worry about whether the size they've selected will fit their individual body type. The tactile properties of the material or how the garment looks on various body shapes may not be accurately captured in online product photos. There are various ways to address this problem, but one easy and affordable one is image based virtual try-on, which involves using the user's photograph to superimpose a garment from a shopping website to create a virtual fitting experience.

The development of efficient image based virtual try-on systems is greatly aided by deep learning techniques, which also improve the accuracy, realism, and overall performance of these systems. Deep learning models do exceptionally well in semantic understanding, which allows them to identify and decipher complex information in photos. This feature is essential for correctly superimposing clothing items onto user supplied photos, taking complex designs, textures, and fabric patterns into account. Its methods, which include pose estimation models, help to precisely align apparel with the user's posture. By doing this, it is made sure that the user's actions and realistic garment positioning are reflected in the virtual try-on experience. Using a pretrained Posenet model for precise stance identification and a complex process of garment texture transfer from reference photographs onto the user's image are the main components of our solution. Furthermore, the system integrates pose estimation models to smoothly match garment textures with the user's pose and segmentation models to detect clothing regions.

Processes like data collecting, preprocessing, model training techniques, and the curation of an interactive user interface are all carefully navigated. Users can interact with clothing products, try on different sizes, and see the outcomes in real time through this interface. The system's focus to inclusivity, which guarantees adaptation across a range of body shapes, sizes, and positions, is a crucial component of its design. The main goal is to create a platform that is available to everyone, democratizing the virtual try-on experience.

Examining the intersection of fashion and technology, this investigation presents a paradigm shift for the fashion retail industry. It clarifies how virtual try-on systems might change how customers engage with items in a variety of contexts.

II. Literature Survey

Advancements in computer vision and deep learning have significantly influenced the evolution of virtual try-on systems, revolutionizing the fashion industry. Ziad Al-Halah, Rainer Stiefelhagen, and Kristen Grauman, in "Fashion Forward: Forecasting Visual Style in Fashion", [1] introduce a computational framework that utilizes neural networks for predicting visual styles in fashion. This framework addresses challenges in client engagement and contributes to applications such as customized fashion recommendations and style analysis. On the other hand, this framework cannot handle complexities and dynamic nature of fashion trends.

Whereas Serge Belongie, Jitendra Malik, and Jan Puzicha, in "Shape Matching and Object Recognition Using Shape Contexts", [2] focus on shape context descriptors and their applications in object detection and shape matching. While emphasizing the importance of accurate shape matching, the paper points out a lack of adoption of deeper learning strategies. The proposed approach showcases success in silhouette matching, trademark identification, and handwritten digit recognition, underscoring the need for incorporating deeper learning techniques into virtual try-on systems.

The study "Keep it SMPL: Automatic Estimation of 3D Human Pose and Shape from a Single Image" [3] by Federica Bogo et al. introduces the SMPL model, a simplified yet effective approach for estimating 3D human pose and shape from 2D images. The model, based on deep learning and Convolutional Neural Networks (CNNs), contributes to computer vision and graphics, enabling applications like animation and virtual try-on. Despite its success, challenges remain, particularly in accurately estimating poses in certain cases.

Zhe Cao, Tomas Simon, Shih-En Wei, and Yaser Sheikh present "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields", [4] a real-time computer vision technique utilizing Part Affinity Fields (PAFs) and CNNs for accurate multi-person 2D pose estimation. While demonstrating success in real-time applications, the framework acknowledges difficulties in handling missing components in uncommon poses. This work addresses challenges crucial for supporting a diverse range of user poses in virtual try-on scenarios.

The study on "Generative Adversarial Networks" by Ian J. Goodfellow et al. introduces a novel framework for creating generative models using a two-player game between a generator (G) and a discriminator (D) [5]. GANs, comprising neural networks engaged in adversarial training, generate synthetic data closely resembling real data. The paper's contributions in generating high-quality data highlight its relevance for applications like image and text generation in virtual try-on scenarios.

The study "VITON: An Image-based Virtual Try-on Network" by Xintong Han, Zuxuan Wu, Zhe Wu, Ruichi Yu, and Larry S. Davis presents a virtual try-on system based on image synthesis and manipulation [6]. The VITON system is designed to address the challenge of generating realistic virtual try-ons of clothing items on a person in a target image. The framework may not extensively cover the adaptability of the model to various poses and body shapes.

Overall, these studies collectively contribute to the virtual try-on domain by addressing challenges, introducing novel frameworks, and emphasizing the need for deeper learning strategies. As the fashion retail sector undergoes a paradigm shift, the convergence of computer vision, machine learning, and user-centric design promises an enhanced virtual try-on experience for customers interacting with online shopping platforms.

III. METHODOLOGY

Various Artificial Intelligence powered generative models and approaches are employed to accomplish the virtual fitting task in phases, with each phase being well-structured and forming a chain to utilize the output from the previous phase in the subsequent one. Figure 1 displays entire process of the framework and table 1 describes algorithms used.

TABLE I. ALGORITHMS USED IN EVERY STAGE

S.NO	STAGE	ALGORITHMS USED
1	Pose Estimation	PoseNet
2	Clothing segmentation and masking	U-Net
3	Semantic segmentation of human image	DensePose
4	Texture transfer and final output	CNN

A. Input Data and Data processing

Two pictures are compiled as input data: one is the source clothing image and the other is the target person image onto which clothing should be transferred. The supplied data is then processed and resized to the model's desired size. The human image can be provided in any size or quality, but the garment image should have dimensions of 500 by 600 pixels. To ensure uniformity, the human image is resized to the normal proportions of 768 by 1024 pixels. After resizing, the human image is stored for later analysis. The saved image is then loaded and resized one more to the final 384 by 512-pixel dimensions that are required. The resulting image is then saved once more.

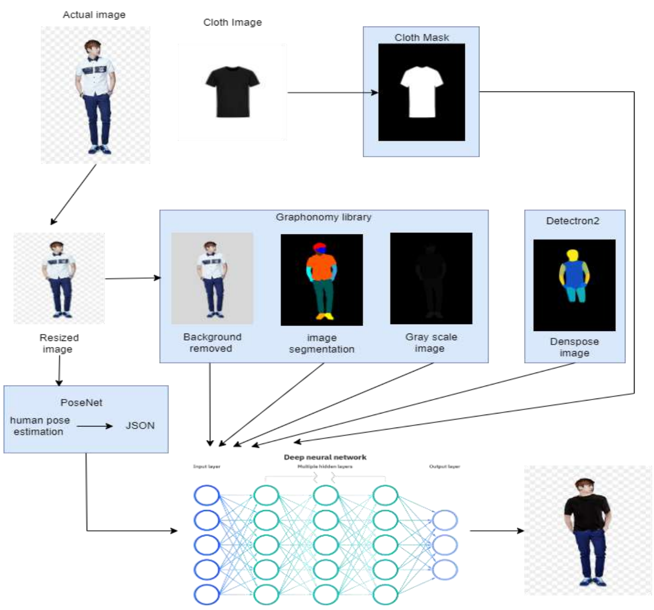


Fig. 1. Architecture of Virtual try on framework

B. *Pose Estimation*

The resized human image is used by the Posenet model to detect the pose of the person in it. This code utilizes the PoseNet deep learning pre-trained model for human pose estimation on a resized image. The model, pretrained with 101 key points, processes the image, generating heatmaps, offsets, and displacement maps. The function extracts pose with scores, key point scores, and coordinates. Detected poses are mapped to OpenPose format, and the most confident pose is chosen. The resulting key points are formatted in a JSON file. PoseNet employs layers for heatmap generation, offset calculation, and displacement maps, contributing to accurate key point localization. The code streamlines the process for virtual try-on applications or similar tasks.

The input image is first preprocessed and then fed through a PoseNet model that has already been trained. The model identifies critical locations that correspond to several body parts, including elbows, shoulders, and facial features, which are essential for comprehending the stance portrayed in the picture. The essential points that have been found are then mapped to the OpenPose format standard after this detection step. This mapping guarantees interoperability with later models or applications, allowing for a smooth integration into larger systems for tasks like action classification or gesture recognition. All in all, this procedure makes it easier to extract and modify complex posture information from pictures, which improves the usefulness and adaptability of many systems and applications. PoseNet usually comprises of a convolutional neural network (CNN) architecture, which is based on common CNN designs such as ResNet or MobileNet, followed by additional layers for pose estimation, as per other earlier research initiatives. The architecture is made to take an image as input and output the subject's pose, which is often expressed as joint angles or keypoint coordinates [7].

C. *Clothing Segmentation and Masking*

This code performs cloth segmentation using a U-Net model. The U-Net model is created or loaded, and an input clothing image is normalized using Albumentations. The U-Net predicts a binary mask for the clothing region, which is converted to a NumPy array. The resulting mask is un-padded, and the script overlays it onto the original image, creating a composite image. The code uses U-Net's encoder-decoder structure with skip connections for accurate segmentation. The final images are composite and binary mask. Composite Image is a visual result of overlaying the original input image with the predicted binary mask for the clothing region. The composite image highlights the segmented clothing area. Binary Mask Image represents the predicted binary mask generated by the U-Net model. It outlines the regions where clothing is detected in the original image. U-Net segments images by first encoding features through a CNN, then decoding them back to original resolution using deconvolutional and convolutional layers, preserving spatial context with skip connections, crucial for accurate biomedical image segmentation [8].

The input image is pre-processed and padded to ensure its dimensions are multiples of 32 so that is suitable for the U-Net model. Now, the model to be used is loaded and the image is passed to that model. During inference, the U-Net model predicts a binary mask representing the segmented regions corresponding to clothing items within the input image through a forward pass of the input image through the trained model. The model's forward pass involves passing the image through the encoder path, where hierarchical features are extracted at multiple scales. As the input image passes through the encoder, features are progressively down-sampled, capturing increasingly abstract representations of the image. The feature maps generated by the encoder are then passed through the decoder path. Here, up-sampling operations gradually increase the spatial resolution of the feature maps while preserving important contextual information.

A binary mask is generated based on the probabilistic value of the pixel. Every pixel of the image is classified as foreground or background values are given based on that. Post-processing steps involve converting the model's output into a NumPy array and removing the padding previously added to the image.

D. *Semantic Segmentation of human image*

This phase consists of two sub-phases, generating semantic segmentation of the human image and removing background of that image using the segmented image. The image segmentation script which is also known as inference script performs semantic segmentation using Graphonomy-Master library. It executes an inference command on a resized image, and the resulting semantic segmentation mask is used to remove the background from the original image. The inference script is part of the Graphonomy-Master library and is designed to work with a DeepLab-Xception Transfer model.

The DeepLab-Xception model, a variant of the DeepLab architecture, operates by leveraging the Xception backbone for feature extraction and Atrous convolution to capture multi-scale contextual information. Initially, the Xception backbone processes the input image to extract hierarchical features, utilizing atrous convolutions with varying dilation rates to enlarge the receptive field and capture context at different scales effectively. Subsequently, the Atrous Spatial Pyramid Pooling module aggregates context from multiple scales through parallel Atrous convolutions, enabling the model to handle objects of diverse sizes and delineate fine-grained semantic boundaries. Finally, the model predicts pixel-wise semantic labels by applying a SoftMax activation to the output logits, yielding a segmented image where each pixel is assigned a semantic category. This methodology enables DeepLab-Xception to achieve accurate semantic segmentation across a wide range of applications, including image understanding, object recognition, and scene parsing [9].

This script loads the pre-trained Xception model, performs multi-scale inference on the input image, and generates a semantic segmentation mask. The mask is then used to decode the predicted labels and create a visual representation of the segmentation. The processed image is saved in both color and grayscale formats. For the DeepLab-Xception Transfer model, the architecture consists of deep convolutional layers with skip connections and transfer projection layers, enabling effective semantic segmentation. The model is trained to predict pixel-wise labels for 20 classes, including various body parts and background. The image segmentation script and the inference script together enable the extraction of semantic segmentation information from an input image, contributing to tasks such as background removal or object isolation.

As a next step, a colored semantic segmentation image into a grayscale representation, assigning specific labels to different clothing and body parts. This conversion is achieved through the utilization of OpenCV, the Python Imaging Library (PIL), and the NumPy library. The script starts by loading the resized segmentation image using the Image module from the PIL library. The image's width and height are extracted, and the image is converted into a NumPy array. Subsequently, a new grayscale image is initialized with the same dimensions. Based on the pixel's RGB values of original image, corresponding labels are assigned to the grayscale image. Labels such hair, head, neck, torso, left arm, right arm, pants, dress, skirt, left leg, right leg, and outerwear are assigned based on specific RGB values. After the pixel-wise labeling, the grayscale image is resized using OpenCV's resize function with nearest-neighbor interpolation. The resulting grayscale image is then saved using the Image.from array function. After the pixel wise labeling foreground image is separated from the background based on the labels and the separated foreground is stored in separate variable all this is done by performing bitwise operation on the original image and the mask.

E. DensePose Segmentation of human image

The Detectron2 library is used to perform DensePose segmentation on an input image. Detectron2 is a library that can be used for various functions such as semantic segmentation, object detection and many more [10]. The DensePose model in Detectron2 which is a pre-trained model, applied to the image, and the resulting segmentation is color-mapped using predefined colors for different body parts. The model generates a JSON file of the image which contains information about various parts in the image. Now another function extracts information from that JSON file containing DensePose predictions and applies the color mapping to visualize the segmentation. Finally, the segmented image is overlaid onto a black background, creating a visual representation of the DensePose segmentation.

F. Texture transfer and final output

HR-VITON model operates by taking a set of diverse inputs crucial for virtual try-on applications. These include a resized image, cloth mask delineating the clothing region, image segmentation for understanding the context, grayscale image capturing fine details, PoseNet JSON providing pose information, background-removed image for isolating the person, and DensePose image offering dense pose estimation. The complex arrangement of deep learning models within HR-VITON involves sophisticated architectures. Convolutional layers (Conv2d) play a pivotal role in extracting hierarchical features from the input images. These features are then carefully normalized using BatchNorm2d and InstanceNorm2d layers to ensure stable training. The incorporation of Residual Blocks, specifically SPADE Residual Blocks, facilitates the efficient learning and transformation of features, enhancing the network's capacity to capture intricate patterns and nuances.

At the heart of HR-VITON is the SPADE Generator, a powerful network that dynamically adapts its normalization through SPADE normalization, aligning with the semantic segmentation of the input image. This generator will provide a synthetic image of the clothing from the semantic segmentation mask, this is done by the normalization layers wash away the semantic information [11]. This generator refines and augments features through the strategic utilization of multiple SPADE Residual Blocks. The training process is guided by GAN Loss, a technique within Generative Adversarial Networks. GAN Loss propels the generator to produce highly realistic images of the virtual try-on, while the discriminator learns to discern between authentic and generated images. The synthetic image that is generated is integrated with the target image to show texture transfer virtually, this done by blending in which pixels are combined from synthetic image and target image here binary mask ensures which pixels should be replaced by synthetic image pixels in target image.

In essence, the final output of HR-VITON is a compelling visual representation. It vividly illustrates how the chosen clothing seamlessly integrates with the person in the input image, providing a realistic preview of the virtual try-on experience. The amalgamation of advanced deep learning techniques and carefully designed architectures enables HR-VITON to deliver impressive and visually accurate results in the domain of virtual clothing try-on.

G. Performance of the model and it's improvement

The performance of the model, particularly in terms of accuracy and realism, largely depends on various factors such as the quality of the semantic segmentation, the effectiveness of the SPADE generator, and the integration with PoseNet for pose alignment. Initially, without PoseNet and SPADE generator, the model may struggle with accurately transferring clothing between images, resulting in misalignments or unrealistic blending. However, by incorporating PoseNet, which provides pose information, the model can better align the synthesized clothing with the target image, improving the overall realism of the try-on process. Additionally, the SPADE generator enhances the fidelity of the synthesized clothing by considering spatially adaptive normalization, leading to more realistic textures and shapes that closely resemble actual garments [12]. Overall, the integration of PoseNet and SPADE generator models enhances the performance of the virtual try-on system, offering users a more immersive and satisfying experience. Various other models like GAN are used in previous works but there was a lack of reality, a shift in such models shows significant improvement in performance.

IV. Results

Promising outcomes have been obtained from the thorough integration of PoseNet, Graphonomy Master, Detectron2, and HR-VITON in the image-based virtual try-on pipeline. Pose alignment, garment segmentation, and dense pose estimation are three major issues that the system successfully tackles, producing realistic and aesthetically pleasing virtual try-on experiences. Figure 2 and 3 are the input images to the model and Figure 4 is the background removed from the actual image. Figure 5 is the segmentation image using Graphonomy library and Figure 6 is the Human Denspose Image using Detectron2. Figure 7 is the output image of the model.

Fig. 2. Input Human Image

Fig. 3. Input cloth image

Fig. 4. Removing Background

Fig. 5. Image segmentation

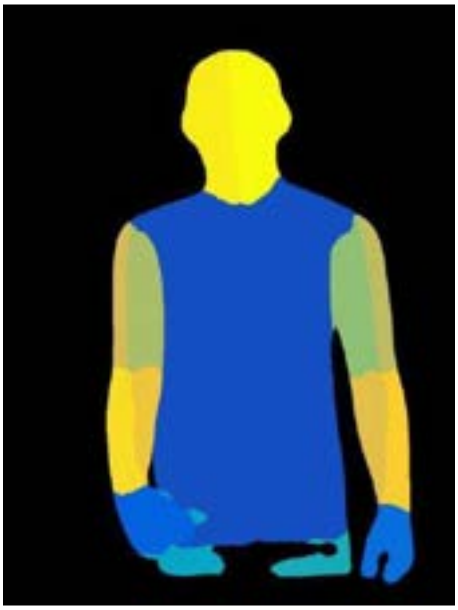


Fig. 6. Human Denspose Image

Fig. 7. Output Image

V. DISCUSSION

The successful implementation of the image-based virtual try-on system lays the groundwork for several exciting future endeavours. Firstly, the integration of multi-view data could enhance the system's spatial understanding, enabling more accurate garment fitting. Optimizing computational efficiency could lead to real-time virtual try-ons, enhancing user interactivity. Interactive customization features and user-generated content could provide users with more personalized and diverse experiences. Additionally, robustness to dynamic poses, ethical considerations, and the integration of augmented reality are promising areas for further exploration. Future research could focus on refining these aspects, ultimately advancing the state-of-the-art in image-based virtual try-on systems and providing users with more immersive and customizable virtual shopping experiences.

VI. CONCLUSION

With the help of this framework, an image-based virtual try-on system has been developed and improved, both improving upon and borrowing ideas from previous approaches. Our results highlight the important contributions the project has made to the field through careful implementation and testing. Notably, the technology addresses an important part of virtual try-on realism by demonstrating a respectable degree of adaptation to varied stances and body types. The visual accuracy of virtual try-ons is improved by the painstaking attention to capturing minute details in clothing textures and patterns, providing a more realistic and engaging user experience. Prioritizing user interaction has resulted in the introduction of innovative customisation tools that enhance engagement and allow the virtual try-on experience to be customized to individual preferences.

AI plays a pivotal role in revolutionizing the virtual try-on experience, offering a transformative approach to fashion exploration and personal styling. By leveraging advanced techniques in semantic segmentation and generative modeling, AI enables the seamless transfer of clothing between images, fostering a dynamic platform for users to virtually experiment with diverse styles and looks. The integration of sophisticated algorithms such as SPADE further enhances the fidelity of synthetic clothing synthesis, ensuring lifelike textures and shapes that closely mimic real garments. Moreover, AI-driven interactions empower users with personalized recommendations and intuitive controls, enriching the virtual try-on process with tailored experiences and enhancing user engagement. As technology continues to evolve, the synergy between AI and virtual try-on applications holds immense potential to reshape the fashion industry, offering innovative solutions for retail, e-commerce, and beyond.

REFERENCES

[1] Al-Halah, Ziad, Rainer Stiefelhagen, and Kristen Grauman. "Fashion forward: Forecasting visual style in fashion." In Proceedings of the IEEE international conference on computer vision, pp. 388-397. 2017.

[2] Belongie, Serge, Jitendra Malik, and Jan Puzicha. "Shape matching and object recognition using shape contexts." IEEE transactions on pattern analysis and machine intelligence 24, no. 4 (2002): 509-522.

[3] Bogo, Federica, Angjoo Kanazawa, Christoph Lassner, Peter Gehler, Javier Romero, and Michael J. Black. "Keep it SMPL: Automatic estimation of 3D human pose and shape from a single image." In Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part V 14, pp. 561-578. Springer International Publishing, 2016.

[4] Cao, Zhe, Tomas Simon, Shih-En Wei, and Yaser Sheikh. "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. 2017.

[5] Goodfellow, Ian, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. "Generative adversarial networks." Communications of the ACM 63, no. 11 (2020): 139-144.

[6] Han, Xintong, Zuxuan Wu, Zhe Wu, Ruichi Yu, and Larry S. Davis. "Viton: An image-based virtual try-on network." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 7543-7552. 2018.

[7] Shavit, Yoli, and Ron Ferens. "Introduction to camera pose estimation with deep learning." arXiv preprint arXiv:1907.05272 (2019).

[8] Yousef, Rammah, Shakir Khan, Gaurav Gupta, Tamanna Siddiqui, Bader M. Albahlal, Saad Abdullah Alajlan, and Mohd Anul Haq. "U-Net-based models towards optimal MR brain image segmentation." Diagnostics 13, no. 9 (2023): 1624.

[9] Hsu, Chih-Yu, Rong Hu, Yunjie Xiang, Xionghui Long, and Zuoyong Li. "Improving the Deeplabv3+ model with attention mechanisms applied to eye detection and segmentation." Mathematics 10, no. 15 (2022): 2597.

[10] Abdusalomov, Akmalbek Bobomirzaevich, Bappy MD Siful Islam, Rashid Nasimov, Mukhriddin Mukhiddinov, and Taeg Keun Whangbo. "An improved forest fire detection method based on the detectron2 model and a deep learning approach." Sensors 23, no. 3 (2023): 1512.

[11] Park, Taesung, Ming-Yu Liu, Ting-Chun Wang, and Jun-Yan Zhu. "Semantic image synthesis with spatially-adaptive normalization." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 2337-2346. 2019.

[12] Kim, Jaehun, and Yoonsik Choe. "Document Image Restore via SPADE-Based Super-Resolution Network." Electronics 12, no. 3 (2023): 748.

Optimization and Enhancement of Doctor Appointment Booking System Using Next.js, Strapi, and REST API

1st Tamizharasi A.
Department of CSE
R.M.D. Engineering College
Kavaraipettai, India
tamizh.cse@rmd.ac.in

2nd Sriram Siddharth DN
Department of CSE
R.M.D. Engineering College
Kavaraipettai, India
ucs20419@rmd.ac.in

3rd Srivatsan GS
Department of CSE
R.M.D. Engineering College
Kavaraipettai, India
ucs20421@rmd.ac.in

4th Shreenath V
Department of CSE
R.M.D. Engineering College
Kavaraipettai, India
ucs20411@rmd.ac.in

***Abstract*—In the current healthcare scenario, booking an appointment with a doctor often involves numerous challenges such as limited appointment availability, geographical constraints, difficulty in finding the right doctor, lack of transparency, and issues with rescheduling or canceling appointments. This not only leads to a poor patient experience but also results in inefficient management of the doctor's schedule. The current system's structure is composed of outdated front-end frameworks, making it challenging to store data in the back end. This web-based application is built using Next.js for the front-end, Strapi API for the backend, MySQL for database purposes, Kinde Auth for authentication and Hostinger for hosting the database. To address this issue, a web-based program called Doctor Appointment Booking System allows the patients to schedule their appointments easier. The primary goal of this research study is to streamline the time-consuming appointment scheduling process and improve the efficiency of the healthcare system. Missed appointments and misunderstandings will be decreased since the proposed approach offers a clear communication channel between patients and healthcare professionals.**

Keywords— Next.js for frontend, Starpi Application Programming Interface for backend, MySQL for database, Kinde Auth for security purposes, Hostinger.

I. Introduction

The doctor appointment booking system has been developed to facilitate scheduling appointments with healthcare professionals. Next.js, Strapi, MySQL, and a CSS framework called Tailwind CSS are state-of-the-art web development technologies that were used in its development. Server-side rendering and the creation of static web pages for React-based web apps are made possible by the well-liked React framework Next.js. It offers a strong basis on which to construct applications with great performance. Strapi is an adaptable, open-source headless CMS that enables editors to simply manage and share their content while granting developers the flexibility to use their own tools and frameworks. Strapi API integration is utilized in this system to manage and provide the application data, which facilitates managing content and updating. One popular open-source relational database management system is MySQL. It is used in this project to handle and maintain information about appointments, patients, and doctors. Patients can set up appointments at their convenience, examine doctor biographies, and verify their availability through the system. Conversely, it helps physicians control their own schedules, which enhances the system's overall effectiveness. A robust web development framework called Next.js makes it easier to create quick, interactive apps. Based on the popular JavaScript framework React, it provides extra structure and functionalities, which include server-side rendering and static generation. Next.js provides major features which includes Server-side Rendering, Automatic Code Splitting, Hot Code Reloading, Ecosystem Compatibility, Styled-JSX etc. Next.js is built by Vercel and is known for providing the best developer experience when building production-ready applications. Developers and companies can manage content effectively with Strapi, an open-source headless Content Management System (CMS). Strapi's primary characteristics are its headless CMS, flexibility, support for many databases, and API-centricity. It is also open source. The necessity for an adaptable JavaScript-based CMS with API capability led to the creation of Strapi. It has developed into a crucial part of frontend design today, providing developers with considerably more robust APIs and fine-grained control over their content. Presently, Walmart, NASA, IBM, eBay, and numerous other businesses use it. This application demonstrates the combination of Next.js, Strapi, and MySQL to create a scalable, effective, and user-friendly web application. It is a major step in the direction of digitizing healthcare services and improving their accessibility and organization.

II. Literature Survey

The Medical Appointment Booking System was created to simplify and enhance the process of arranging and keeping medical appointments. It is a feature-rich and intuitive digital platform. In this fast-paced world, efficient healthcare services are crucial. With the provision of a practical and easily available alternative, this system seeks to close the gap between patients and healthcare practitioners. Patients and patrons of certain medical institutions, such as clinics or hospitals, can schedule online appointments with physicians by utilizing this program. With this approach, patients may easily browse a database of doctors' profiles, accessing information about their expertise, availability, and biography. Even better, they can set up appointments for any time and day of their choosing.

TABLE 1: LITERATURE SURVEY

Title	Authors	Journal/ Conference	Year
Appointment System using Artificial Intelligence Techniques.	Louisa Aiyeyika et al.	NEU Journal for Artificial Intelligence and Internet of Things	2023
Doctor Availability and Appointment using Digital Technology.	Santhosh S. et al.	Internation al research Journal of Modernizationin Engineering Technology and Science	2023
Optimization of an appointment scheduling based on the quality whale optimization algorithm and NSGA-II.	Ali Ala. et al.	Journal of Scientific Reports	2021
Integrated optimization of appointment allocation and access Prioritization in Patient-Centered Outpatient scheduling.	Na Lia. et al.	Journal of Computers & Industrial Engineering	2021

With this technology, patients may easily browse a database of doctors and view their availability, specialty, and biographies. Even better, they can set up appointments for any time and day of their choosing. This system will schedule each patient's appointment based on their availability, specializations, and bios. Even better, they can set up appointments for any time and day of their choosing. This doctor's appointment system will set up each patient's schedule and forward it to the doctor as a request. Efficient Appointment Scheduling (EAS) is essential to hospital management for ensuring patient satisfaction and high- quality service. Healthcare schedulers typically advise patients on the best time to receive services before the admittance call closes. The ability to modify the appointment has been removed. The appointment date cannot be changed at this time. The whale optimization algorithm (WOA), which considers the simulation approach and is based on the Pareto archive and the NSGA-II algorithm, solves the appointment scheduling model. The multi-criteria approach to appointment scheduling based on these two algorithms has been covered in this paper. This research computes WOA and NSGA utilizing several hypotheses to satisfy the analytic requirements and account for different features related to hospitalized patients. Three examples from the final part of the model have been used in this study to investigate NSGA and WOA.

III. METHODOLOGY

The methodology employed here aimed to develop and evaluate the proposed system, which integrates features to enhance the booking appointment website.

A. Understanding the problem statement:

This system addresses the most important part of scheduling appointment for patients with their medical consultant in their flexible time slot.

B. Design:

Using the Next.js framework, the user interface is designed to be straightforward and easy to use. The list of doctors according to specialization and the list of well-liked physicians in the city are both included in the user interface design. The site includes a profile of that particular doctor, an appointment booking page, and a scheduling interface.

C. Development:

The application is developed using Next.js for the Frontend, Strapi as the headless CMS, MySQL for the database, hosted in Hostinger, Resend for Email notification. The application is completely responsive and works seamlessly across different devices.

D. API Authentication:

Strapi uses headless CMS for connecting the MySQL database with Next.js framework as a connecting mechanism between the frontend and backend. This connection is made through API (Application Programming Interface) which is provided by Strapi through headless CMS.

E. Managing Web Hosting:

Hostinger is used to host the domain for database purposes. The platform provides various features in creating and managing a web hosting platform, domain hosting etc.

F. REST API:

Accessing the content-types via API endpoints is made possible via the REST API. The moment a content-type is produced, Strapi automatically generates API endpoints. To enhance the results of an API endpoint query, one might utilize API parameters.

IV. SYSTEM DESIGN

Sign in / Sign up Process

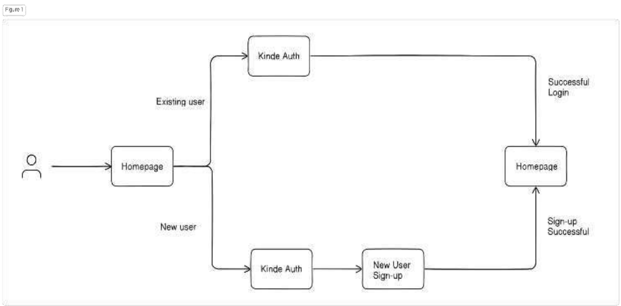


Fig 4.1 Sign in / Sign up

As shown in fig 4.1, the first stage is when the user lands on the homepage of the website. Existing user authenticates their account with their credentials. New user needed to create an account with the Kinde Auth. After a successful sign-up the user is redirected to the homepage.

Appointment Booking process:

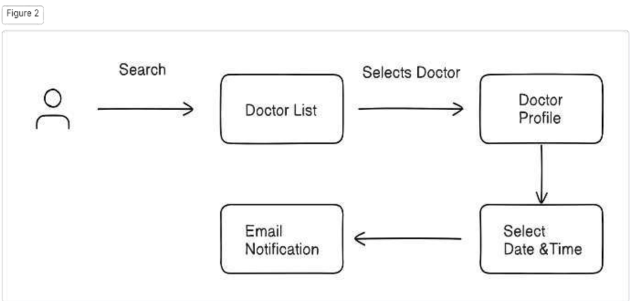


Fig 4.2 Appointment Booking process

As shown in fig 4.2, in the second stage, the user searches the doctor list in the homepage. The user selects the doctor profile, schedules the appointment with the concerned doctor along with date and time with the slot available on a particular day. An Email notification about the appointment is sent to the customer.

Backend Working Process:

The backend tool used to retrieve data from the database is called Strapi. Starpi handles the data from the MySQL database using a headless CMS. Content-types can be accessed via API endpoints thanks to the REST API. When a content-type is formed, Strapi automatically generates API endpoints. When searching API endpoints, the results can be refined by using API parameters. The overall process is depicted in fig 4.3.

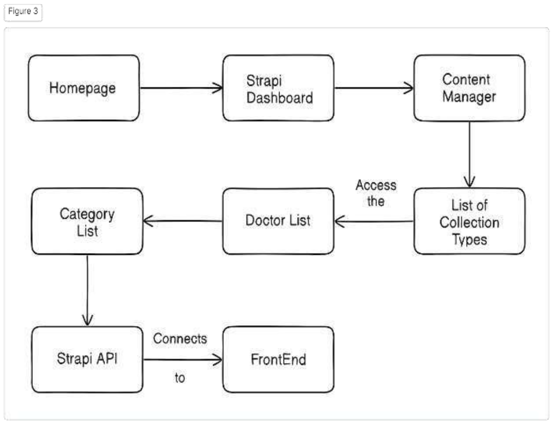


Fig 4.3 Backend Working Process

Email Notification:

Fig 4.4 Email Notification

Email notification is sent to the user through Resend API using Node.js framework as shown in fig 4.4. API works by deploying the API key in the Global Api folder for authentication. The user gets an Email notification about their appointment scheduled with the particular doctor.

V. WORKING MODEL

The working model of Doctor Booking Appointment website is briefly picturized below in fig 5.2.

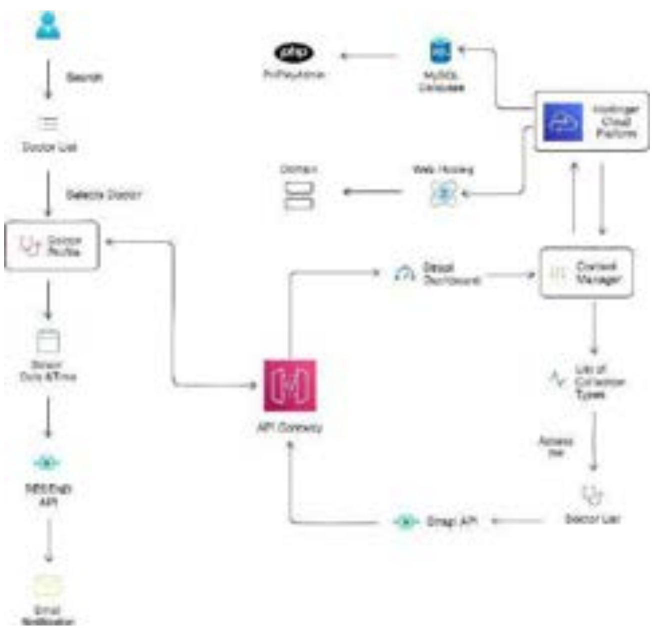


Fig 5.1 Working Model

VI. RESULT

USER MODULE:

Homepage:

Homepage has the features to explore the physicians based on their specialization and also popular doctors in the city. User need to sign-in their account for scheduling the appointment with their concerned doctor. The real-time homepage is shown in fig 6.1

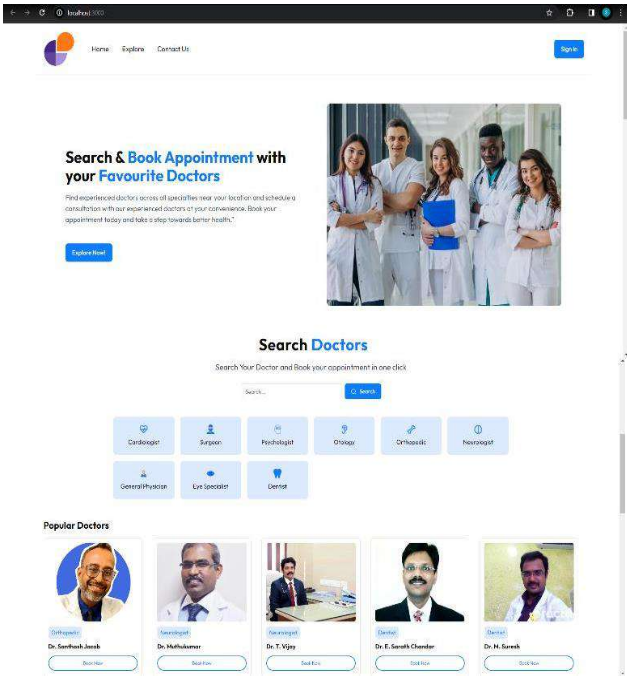


Fig 6.1 Homepage

DOCTOR MODULE:

Doctor module includes the doctor profile, appointment schedule, work experience, manage and access to their planned appointments. User is provided the book appointment button to schedule the appointment with the doctor. The real-time doctor's webpage is shown in fig 6.2.

Fig 6.2 Doctor Module.

ADMIN MODULE:

Admin dashboard contains information about the doctors' information, Category List, Images and Information about the doctor, Appointment information about the patients etc., The real-time doctor's webpage is shown in fig 6.3

Fig 6.3 Admin Module

MySQL Database:

The database holds the records of the all doctors, patients registered in the website. Hostinger holds the MySQL database. Through API the information is fetched from the database to the frontend. A sample image is shown in fig 6.4.

Fig 6.4 MySQL Database

Email Notification:

Email notification is sent using RESEND API provided by React.js framework to the particular user who booked their appointment with their concerned doctor. A sample Email confirmation is shown in fig 6.5.

Fig 6.5 Email Confirmation.

Reminder about Appointments will be shown in my bookings as depicted in fig 6.6:

Fig 6.6 My Bookings.

VII. Conclusion

In conclusion, the development and implementation of the web-based doctor appointment booking application have significantly effective features and functionalities which makes user to book and schedule appointment with their concerned doctor. The system successfully integrated features such as user registration, profile management, notifications, reviews and ratings, payment processing. This application is built using Next.js, Strapi API, MySQL, Kinde Auth for authentication, Hostinger host the web in the domain. This website has proven that potential of digital platforms in transforming healthcare services. This system is an example that how technology can be leveraged to make healthcare more accessible and convenient for patients. Future enhancements were also proposed, including telemedicine support, e-prescriptions, health record management and AI- based symptom checker. These enhancements are expected to further improve the user experience and make the system more robust and comprehensive.

XI. References

[1] Louisa Aiyeyika, Fadi Al-Turjman ,P., 2023, June. Appointment System using Artificial Intelligence Techniques. In 2023 Vol. 2 No. 2 (2023): Vol. 2 No. 2 (2023): Multidisciplinary Advance Sciences and Technology.

[2] Sugashini K, Anila PS, Samrrutha RS, Santhosh S, P.,2023, October. Doctor Availability and Appointment using Digital Technology. In 2023 International Research Journal of Modernization in Engineering Technology and Science.

[3] Ali Ala, Fawaz E. Alsaadi, Mohsen Ahmadi, and Seyedali Mirjalili, P., 2021, October. Optimization of an appointment scheduling problem for healthcare systems based on the quality of fairness service using whale optimization algorithm and NSGA-II. In 2021 Published in National Library of Medicine.

[4] Na Lia, Xiaorui Li, Canrong Zhang, Nan Kong, P., 2021, January. Integrated Optimization of Appointment Allocation and Access Prioritization in Patient-Centred Outpatient Scheduling. In 2021 Journal of Computers & Industrial Engineering.

[5] C.Laan, M. van de Vrugt, J. Olsman, R.J. Boucherie, P.,2018,"Static and dynamic appointment scheduling to improve patient access time.

[6] A. Condotta, N.V Shakhlevich, P., 2014, Scheduling patient appointments via multilevel template: a case study in chemotherapy .In 2014 Operations Research for Health Care, Vol. 3, no. 3, pp.129-144, 2014.

[7] Sanjeevani P. Avhale, Wrushali R. Ajabe, Pallavi A. Chinchole, Puja T. Changade, P., 2018, April. Doctor Appointment Online Booking System. In 2018 International Journal Of Creative Research Thoughts (IJCRT).

[8] Ali Ala, Feng Chen, P., 2022 March. Appointment Scheduling Problem in Complexity Systems of the Healthcare Services. In 2022 Journal of Healthcare Engineering.

[9] X. Pan, N. Geng, X. Xie, P., 2021. Appointment scheduling and real-time sequencing strategies for patient unpunctuality. In 2021 European Journal of Operational Research, vol. 295, no. 1, pp.246-260.

[10] D.Wang, K. Muthuraman, D. Morrice, P., 2019 "Coordinated patient appointment scheduling for a multistation healthcare network. In 2019 Operations Research vol. 67, no. 3, pp.599-618.

[11] D. Gupta and B. Denton, "Appointment scheduling in health care: Challenges and opportunities," *IIE Transactions*, vol. 40, no. 9, pp. 800–819, 2008.

[12] G. Du, X. Liang, and C. Sun, "Scheduling optimization of home health care service considering patients' priorities and time windows," *Sustainability*, vol. 9, no. 2, p. 253, 2017.

[13] Z. Zhang and X. Xie, "Simulation-based optimization for surgery appointment scheduling of multiple operating rooms," *IIE Transactions*, vol. 47, no. 9, pp. 998–1012, 2015.

[14] G. Du, X. Liang, and C. Sun, "Scheduling optimization of home health care service considering patients' priorities and time windows," *Sustainability*, vol. 9, no. 2, p. 253, 2017.

AI-Powered Legal Documentation Assistant

Kuldeep Vayadande
Department of Information Technology
Vishwkarma institute of technology
Pune, india
kuldeep.vayadande@gmail.com

Aditi Bhat
Department of Information Technology
Vishwkarma institute of technology
Pune, india
bhat.aditi22@vit.edu

Pranav Bachhav
Department of Information Technology
Vishwkarma institute of technology
Pune, india
pranav.bachhav22@vit.edu

Aditya Bhoyar
Department of Information Technology
Vishwkarma institute of technology
Pune, india
aditya.bhoyar221@vit.edu

Zulfikar Charoliya
Department of Information Technology
Vishwkarma institute of technology
Pune, india
zulfikar.charoliya22@vit.edu

Aayush Chavan
Department of Information Technology
Vishwkarma institute of technology
Pune, india
aayush.chavan22@vit.edu

Abstract: **Legal systems worldwide vary in structure and principles, reflecting the diverse legal traditions of different countries. The legal system, inherently complex and reliant on meticulous documentation, often faces challenges related to time-consuming manual processes and the potential for human errors. The system proposed provides a transformative solution to the above problems. The system emerges as a groundbreaking solution within the intricate landscape of legal systems which responds to these challenges by seamlessly integrating advanced AI techniques. At its core, OpenAI embeddings takes center stage, demonstrating unparalleled proficiency in document generation, comprehension, and abnormality detection, addressing the complexities ingrained in legal documentation. In contrast to traditional approaches, this system maximizes the versatility of ChatGPT 3.5, allowing it to not only issue commands but also proficiently generate a diverse array of legal documents. By incorporating an understanding module equipped with PyPDF, Amazon Textract, and langchain utilities, the system adeptly handles document intricacies. The utilization of OpenAI Embeddings further enhances natural language understanding. Leveraging sentiment analysis and Named Entity Recognition (NER) in its natural language processing (NLP) toolkit, the system employs an intuitive web interface for irregularities detection. The exploration of AI for automated irregularity detection showcases its transformative potential in ensuring document accuracy within the legal domain. This project, therefore, stands as a beacon of innovation, promising to reshape the dynamics of legal document processing by merging advanced AI capabilities with the unique demands of legal systems.**

Keywords: Legal Document Processing, Natural Language Processing (NLP), Irregularities Detection, Legal Technology, OpenAI Embeddings

I. INTRODUCTION

Legal systems throughout the world have countries' diverse laws, implemented structures and processes that manifest the unique nature of traditions in each country. These issues, in fact already there, have to do with human tendency to introduce errors and obligatory use of time (efforts) during the documentation by the means of outdated methodology like paper format [1]. Those users have necessitated an AR-supported Legal Document Assistant (LDA) system curved to transform the scene.

Like in any other field, the oldest legal systems have variety of challenges faced, especially filing of manual documents and a variety of risks that exist with complex texts. The accuracy and efficiencies of existing document processing systems is already inadequate, due to their inability to keep up, which is even more challenging in the precision oriented legal cases documentation environment [2]. The problem may be intensified because of the peculiarities of legal language, which raises the demand of the comprehension, which is far sometimes unavailable in the ordinary systems. Manual document proliferation could be an arduous exercise with human error of missing out serious issues.

On the one side, the desired solution, the artificial intelligence lawyer assistant is a game changer that is tactically created to remedy the detected shortfalls.

AI used system include unusual chapter detection in the legal document texts. The LDA achieves this through the application of Sentiment Analysis and NER, with are enablers of the NLP toolkit. The LDA operates through a simple Web based user interface that detects inconsistencies and ambiguities. This study carries a competitive edge of AI for automated irregularity detection very highly due to its various benefits within the complexity of legal realms.

The tech stack selection of the system is very important since it takes into account the variety of challenges that are found in field of legal processing. PyPDF has a key role in the tech stack allowing us to work efficiently with PDF documents in the same way as it is done with other formats. In the legal field the pdf files are everywhere, and PyPDF offers a versatile way to transform as well as to reuse those documents. PyPDF's flexibility allows us to comfortably proceed through difficult layout of legal documents, preserving originality of the document while carrying out the information operation and creation of new document. Amazon Textract which is capable of detecting money order forms for instance, contributes to the preservation of the customers through its smooth incorporation. Legal document often contains lengthy, and Textract can do a better job of saving a text, or even more information, as it greatly increases the ability to retrieve precise information. Textract's capacity to deal with different types of documents are another great advantage that makes LDA applicable to diverse legal presentations from contracts to agreements. Consequently, the system is never limited but fit for a large-scale use [7].

Through such an innovation, the law processes gain new dynamics, which bounds the efficiency, and accuracy, and directly influences accessibility as well, setting an example of progress in where fit between artificial intelligence and legal procedures takes place [4]. The article highlights not only a tool but the paradigm by showing how AI can improve the healthcare outcomes on one hand and on the other empower the professionals for better performance and implementation of the innovative models in the healthcare system.

Enter the AI Lawyer Assistant, a game-changer tactically designed to address identified shortcomings. Through innovative techniques like Sentiment Analysis and Named Entity Recognition (NER), the LDA detects inconsistencies and ambiguities via a simple web-based user interface. This study underscores the competitive edge of AI in automating irregularity detection within the complexity of legal documents.

II. LITERATURE REVIEW

Xiao's goal is to evaluate Lawformer's ability to handle lengthy legal documents and its suitability for tasks requiring extensive text comprehension [6]. Lawformer, a Chinese model equipped with sliding window and global attention mechanisms, excels in processing complex legal texts, aided by access to 80GB of case data. While it demonstrates proficiency in tasks like sentencing prediction and case retrieval, potential limitations may exist, necessitating further exploration of its effectiveness and adaptability in diverse legal contexts.

Robaldo et al.'s contribution in the Artificial Intelligence and Law journal highlights the benefits and implementation of NLP and semantic technologies in analyzing legal texts [4]. Divided into three sections, the paper explores unfair clause detection, argument structure extraction, and deep learning applications. While commendable for its insights, further analysis is required to assess the effectiveness and generalizability of the proposed NLP approaches in the legal domain.

The document introduces JEC-QA, a challenging judicial applicatory dataset designed to test high-order thinking skills [3]. Through experiments, it reveals a notable performance gap between state-of-the-art models (28% effective) and human capacity (81% accurate). While JEC-QA aims to advance projects in expanding reasoning behavior within QA systems, its findings underscore the significant challenges faced by current models in accurately interpreting judicial texts. Further exploration is necessary to bridge the gap between machine performance and human capabilities in this domain.

In this paper, the authors delve into the realm of Legal AI, exploring AI and NLP applications within the field of law, including embedding and symbol-based representations [11]. The paper analyze various problems associated with Legal AI and provide practical examples such as legal judgment prediction. The strengths of this study lie in its comprehensive analysis of AI and NLP technologies within the legal domain, offering insights into their potential applications and impact. However, this research may lack depth in addressing the challenges and limitations of current Legal AI systems, potentially analyzing the crucial considerations for future development and implementation.

The article introduces the Gov2Vec approach, a word vector learning method tailored for different government institutions and legal documents [5]. By applying Gov2Vec to Supreme Court decisions, Presidential orders, and Congressional bills, the paper extracts main topics, problem-solving methods, and reflects the positions of involved parties. The strength of the paper lies in its innovative approach to analyzing complex governmental texts through vector analysis, providing insights into policy details and interconnections between different governmental organizations. However, the paper may fall short in fully addressing the potential biases or limitations inherent in using vector models for such analyses.

Table 1. A Comprehensive Study of the Machine Learning Methods done by Some Researchers

Algorithm	**Method used/Innovation**	**Application and future work**	**Results and limitations (if specified)**	**References**
PLM, NLP, BERT	The model is implemented with dispersed attention and local context to read through long legal documents. It has proven to be a state-of-the-art performer so far.	This model has outstanding performance in court judgements outcome pre-prediction, corresponding case recovery.	The system still, undoubtedly lacks absolute accuracy and apply to a variety of languages.	[6]
BERT, BlockBERT, FLOP	This work suggests amending the iterative multipurpose attention multi head in BERT to use blockwise attention. The attention matrix is partitioned into sparse blocks by each block paying attention to the subset of it.	BlockBERT demonstrates the highest efficiency on question answering data sets and advances especially in long document understanding tasks where BERT performs with certain constraints.	This attribute of the transformer offers more power for use in relatively longer sentence-sequences such as coreference resolution and document-level machine translation. BlockBERT decreases a memory by 18.7-36.1%, time by 12.0-25.1% than RoBERTa, reaching comparable or better question answering performance especially by as shortening the long-paragraphs,	[7]

			especially in SearchQA and NewsQA.	
CFD, Bi-LSTM, RNN, Xnet, NEUSUM	The paper proposes a novel method for legal courtroom debate summarizing via a multi-task model equipped with a joint utterance encoder for sentence meaning, speaker role, and legal knowledge. This is the first literary piece that compiles in an extractive way arguments of the jurisdiction.	The algorithm will take care of the issue of the summary of detailed court debates, thus judges will be provided with the tool for acceleration of decision-making in civil trials. It targets a common problem, that of the unbearable workloads of judges because of increasing number of legal cases.	The model has an accuracy of 5,477 Chinese civil court debates, and it is suitable for both the utterance assignment and the extraction summarization and even surpasses other baselines. Case studies compared to predecessors show the quality of reports is more lucidity, which is valuable information fitting for legal dispute foci.	[8]
BERT, TextCNN, DPCNN	The paper goes for the JEC-QA. It entails designing a dataset, constructing an open domain automatic question answerer structure, and checking for the models like the attention-based and transfer-learning ones. Approaches used mainly include the dataset creation, QA framework building, and performance evaluation of the AI system on different question types.	JEC-QA boosts up question answering and reading comprehension as well as external reasoning enhancement. The wider upgrade is on augmenting machine reading and reasoning in complex are with sector knowledge. JEC-QA, the biggest legal QA dataset in China with 26k+ questions, which requires complex reasoning, is what the system needs to work with.	Models been able to reach only 28% accuracy. As a result meanwhile humans ones are performing much better: accuracy unskilled workers 64%, skilled workers 81%. The experiments set 7 models as instances for the students to see the problems lies in the concept understanding and reasoning. Future work concentrates on increasing the complexity of legal QA by using models that can be trained better with more knowledge added to them.	[3]

III. METHODOLOGY

While operating the legal system, the systems focus on Technological Development at anyone point by making discreet inquiries about the intrinsic legal systems and the problems that are associated with the manual processes in the fundamentals of legal documentation. The literature review conducted helped us to arrive at the areas of AI intervention that can be lushly filled and upon which will be directly dangled: document generation, comprehension, and abnormality detection in legal domain. The data assembly phase can be characterized by the gathering of a vast variety of legal documents during the process of collecting not only a single document type that is uniform, but rather those of a multilingual nature that represent different legal styles emerging across the worldwide legal tradition. This raw data line enters a complex data cleansing and processing phase whose aim is to identify and neutralize the chaotic and erroneous information through standardized format conversion, format elimination, and other necessary refinements. The chosen ChatGPT 3.5 model, which is known for its quality of natural language understanding, after the applied modification on the dataset is put at an advanced stage of training where, further, an enhancement component of the Open AI embeddings, which fortifies the capabilities of the system to completely apprehend the depth of legal language, is integrated.

The system further develops its framework by incorporating high-end features in the form of advanced knowledge modules which smoothly converge the powers of PyPDF, Amazon Textract, and langchain into this singular application. This module is built to bring intricate data out of documents in an established and thorough fashion which in turn results in improved understanding of complex documents. The inclusion of emotion analysis and Named entity identification (NER) in the NLP toolset makes the tool content-oriented and precise, so analyzing and interpreting legal documents contents becomes much more profound and accurate. The UI part also turns simplified that an efficiently designed web platform is built to work in two ways: you give a command for the document you would like to generate and also make use of the built-in AI detection system for errors. Verification processes, which are based on the application of diverse sets, a range functions including document generation and comprehension as well as identification of irregularities, such provides a valid inferential ground for ensuring performance quality across the main aspects. The clasp, which is the debugging phase, being the main focus of the development team and the balanced comments that are being offered by legal experts, is the idea-furnace for the renewing of the system and the assurance of its practical utility in legal circumstances.

Operating within the legal system requires a keen focus on Technological Development, wherein intrinsic legal systems and associated challenges in manual processes are thoroughly examined. Through a comprehensive literature

review, areas ripe for AI intervention, including document generation, comprehension, and abnormality detection within the legal domain, have been identified. To address the need for clarity in the Legal Document Assistant (LDA), meticulous steps are taken.
The feature extraction process is integral, achieved through meticulous data assembly involving a diverse range of legal documents. This process extends beyond mere gathering. Techniques such as standardized format conversion and refinement are employed to ensure accurate feature extraction. Similarly, the detection process within the LDA is meticulously designed to identify inconsistencies and ambiguities. Leveraging advanced techniques such as Named Entity Recognition (NER) and sentiment analysis, the system scans documents for irregularities, ensuring accuracy and efficiency.
Sentiment analysis plays a critical role in understanding document contents, identifying subtle nuances and inconsistencies within legal texts. By analyzing the emotional tone expressed, the system enhances comprehension and accuracy. Further clarity is needed for the Legal Language Model (LLM) within the LDA framework, which serves as a robust foundation for natural language understanding. Through advanced training and enhancement with Open AI embeddings, the model acquires a comprehensive grasp of legal language nuances, ensuring precise analysis and interpretation of legal documents.
Achieving accuracy within the LDA is paramount, accomplished through a combination of advanced technologies and meticulous verification processes. Techniques such as document generation, comprehension, and irregularity detection undergo rigorous testing, providing a reliable basis for performance quality assurance. The debugging phase, complemented by insights from legal experts, ensures continuous refinement and practical utility in legal contexts. Additionally, the role of Natural Language Processing (NLP) requires further clarity, particularly in its application within the LDA framework.
Placing itself as a versatile system which remains responsive and up to date, the system integrates a continuous improvement procedure, enhancing the system's capacity to respond swiftly to new updates, integrate new legal insights, and enhance the system's user-friendliness. At its core, the system comes out to be a force of change. It brings together all the different strands of artificial intelligence in a weaving process to complement an ancient series of rules and complexities that are native to the judicial system.

IV. PROPOSED SYSTEM

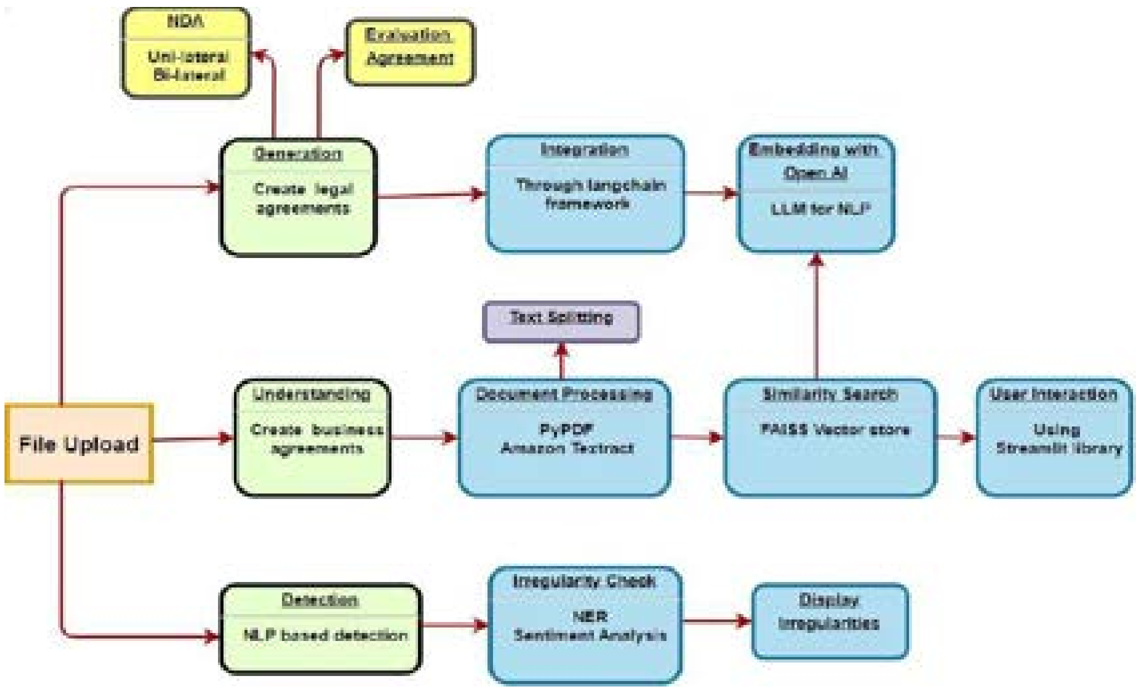


Figure 1. Proposed system

The offered approach in Fig.1 illustrates all the details of the advanced platform's work process and functionality as you can see in the figure below. It points the Management system into the specific direction including 3 subsystems. The system will use comprehensive legal documents, like NDAs and Evaluation Agreements, which are legally binding, as the primary feature. In here, the Framework of LangChain is being utilized along with the Large Language Models (LLMs) by integrating the Open AI for rooftop production, making the latter provide reliable and precise output that is required in any document creation. Therefore, the system makes it possible for a person to comprehend and read several languages, different from his or her very own language. Moreover, it promotes users to navigate a single graphical interface from which to get various types of information across the language barriers. Interfaces are created for the streamlit library applications which consist of the streamlit chart and streamlit graph components. Among all mentioned features, Clarke-Similarity search by using FAISS Vector Space is a powerful one. This feature uses the vector space to provide the similarity information for the document taken as a query. The rendered results will be very close to the submitted one in this case which is very efficient. And finally, to ensure regular compliance as discussed before, the system will make irregularity checks. It will exploit NER and Sentiment Analysis to identify these anomalies and generate an error report which should serve as a tool that will guide the user in further improving the clarity and fluency of the text.

Thus, the proposed system proposed focuses on the automation of text creation aspect as well as interpretation feature, that are both composed from the currently advanced technologies and theories to enable the system to complete the entire process step by step in a very good efficacy.

V. RESULTS

Table 2. A comparative analysis of the existing systems with the proposed system.

Existing System	**Key Findings**	**Multilingual Support**	**Collaborative Editing**	**Accuracy**
Kira Systems	It has the custom of locating, emptying data from and analyzing contents for the terms of your contracts and documents which is done fast and with highest precision. Automates any contract extraction and identification of essential information to the project.	Yes	Yes	>90% [11]
Evisort	Automates contract management and extracts important conditions inside the contracts.	No	Yes	95% [12]
LegalSifter	Highlights what is in a signed contract and the problems that may crop up.	Yes	No	95-97% [13]
Lex Machina	Provides Legal Analytics for companies and law firms can assist them to paint the picture of a winning case and also end up with a good company or law firm. Explore lawsuits case data and legal developments.	No	No	N/A [14]
ROSS Intelligence	ROSS Intelligence Inc. ("ROSS") has a business plan of designing AI-enabled technologies for lawyers to enhance their cognitive abilities. Use case law and statutes for all research legally relevant.	No	No	N/A [15]
AI-Powered Legal Documentation Assistant	Consistently utilizes sophisticated ML methods for the tasks of document production, understanding and detecting abnormalities in legal documents. Saves the effort of time that can be spent on tasks that get stuck in the bottlenecks of the manual work process along with the possibility of human errors in the legal documentation.	No	No	93-97%

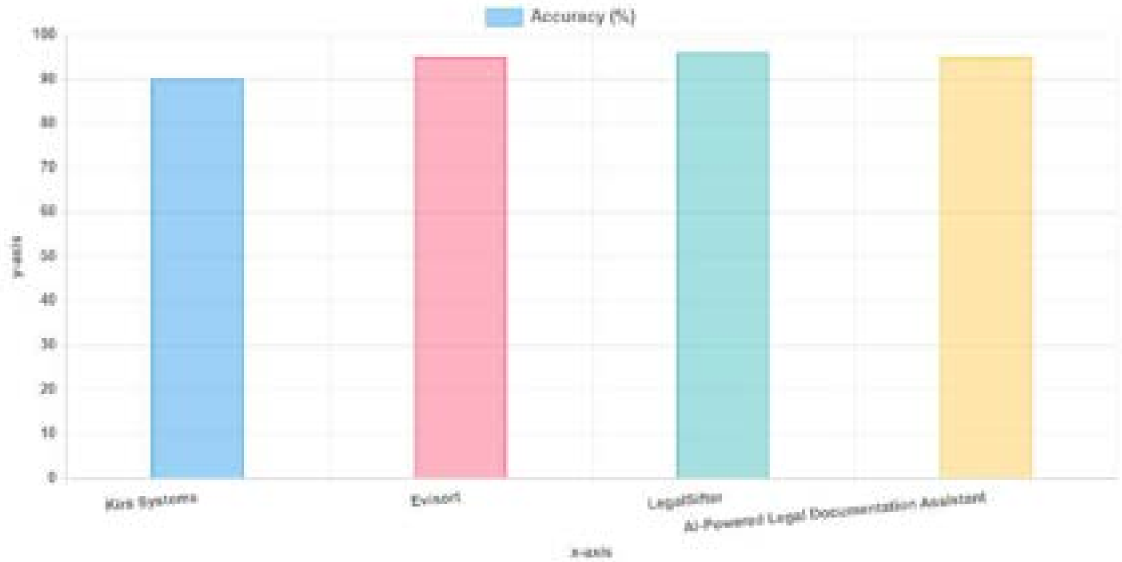


Figure 2. Graph based on accuracies of various systems alongside the proposed system.

Table 3 shows the difference between the existing systems and the proposed system. The table highlights what the systems actually do and compare them on the basis of some parameters.

Table 3. A comparative analysis of various input files considered for various systems.

Input Format	**Description**	**Advantages**	**Limitations**
PDF Files	Portable Document Format (PDF)	Provides easy adoption with a layout that is adapted to other people's necessities.	Imposing just a single system may require different text extraction tools that can be differently expressed.
DOC Files	Microsoft Word Document (DOC)	Drafting and choosing your own way of having a format.	The lack of support for the format of the company, and the problems of compatibility with the other programs are coming more and more often and are becoming the key various difficulties for the usual users of windows.
TXT Files	Plain Text Files	Compared to many others, it would be a small one, but one can take care of it and makes it up to their liking.	The primary drawback of not having rich-formatting features, images, and many other content-related devices was the most prominent.
Handwritten Notes	Handwritten Text or Scanned Images	May be added to the margin of handwritten notes, side notes, and added remarks.	This is caused by the digitization process that translates into unstructured text, and in specific cases where handwritten text is used, the results might be more accurate.
Voice Recordings	Audio Files (MP3, WAV, etc.)	Allows a very smooth dictation and communication.	As the process is mostly dependent on the ability to transcript, it is conditioned by the quality of the audio.

Table 4. A comparison table depicting the difference between the efficiency powers of existing ways with the proposed system.

Aspect	**Traditional Manual Handling**	**Document Management Software**	**AI-Powered Legal Documentation Assistant**
Time Efficiency	20	10	5
Accuracy	85%	90%	95%
Document Processing Speed (ppm)	2	5	10

Error Rate	15%	10%	5%
Integration with Legal Databases	No	Limited	Yes

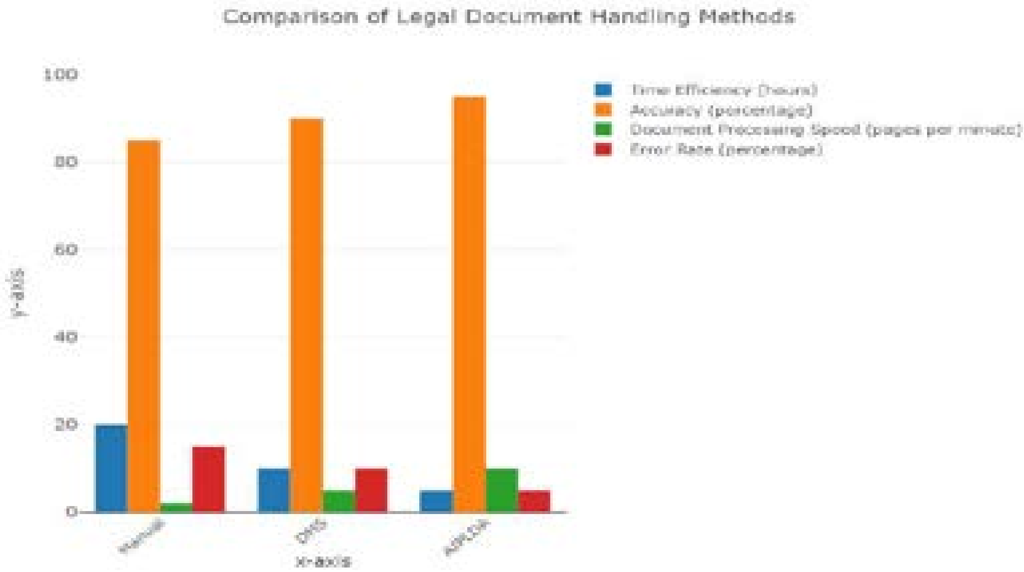


Figure 3. Graph depicting comparison of legal document handling methods.

Table 4 shows the comparison between various existing ways of handling legal document with the way that the proposed system uses. The proposed system proves out to be the best among the choices as it reduces the error rate and risk rate and increases the time efficiency as well as accuracy.

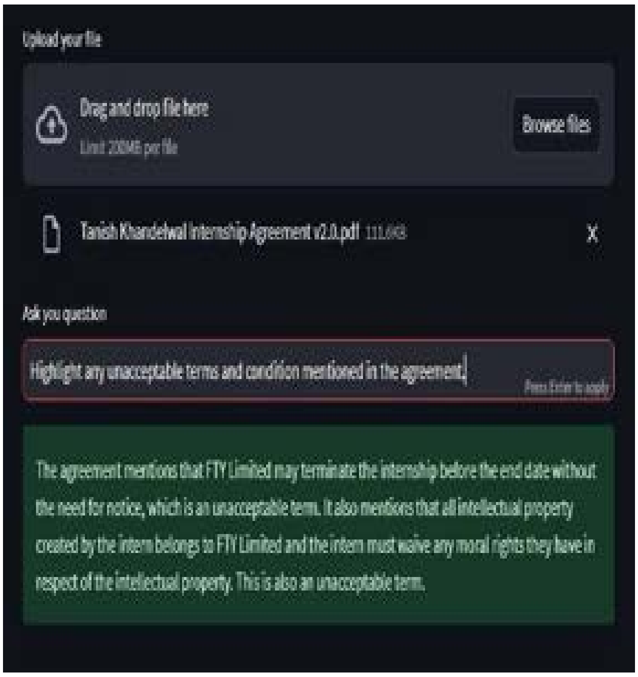


Figure 4.

Figure 5.

VI. DISUCSSION

The tables given represent a detailed look at many types of features needed to have a legal document processing AI system work well. Table 2 illustrates a range of ways that the methods from each system are applied and details their findings, revealing the array of approaches that researchers use in the area. Table 3. clarifies inputs types discussion from PDF, DOC, TXT, HTML, XML files, handwritten notes and even voice recordings. It covers their advantages and cons. Hence, one can notice that the transition of AI-scaffolded legal document processing, signifies the field with both the pick and the difficulties in channeling the sophisticated technologies to support legal doctrines. Figure 4 and Figure 5 also depicts how the system would look like in a few cases when the files are provided to the system and questions are asked on basis of it and when generating a legal document and on basis of it, as well, the questions are being asked. The future of the AI will only manifest itself through constant improvements, realizations and breakthroughs, therefore further search and progress is nothing but a must if AI in legal domain is to expand to its full capacity.

VII. CONCLUSION

The electronic device with custom AI content is the design that will revolutionize the document industry. In the field of document processing, all the intricate issues like automation, natural language understanding, abnormal detection, and regulatory compliances, need to be handled by using the AI techniques to solve the situation related challenges which occur while dealing with the complexities. With a combination of different options, the proposed system provides the feature of works in collaboration and the exceptional security using blockchain technology for processing the documents that are very sensitive, for institutions, business as well as law related agreements. The proposed approach will help to overcome the technicalities in the workflow, increase the efficiency and facilitate quality management. Legal professionals, through their contribution to technological development and the development of digital transformation, occupy a vital and irreplaceable position. Therefore, AI empowered humanoid lawyer might be a manifestation of a change or a revolutionary update to make the legal practice more and more frictionless.

VIII. FUTURE SCOPE

In the next stage, the AI-powered legal documentation assistant is likely to develop by means of the expansion of functions that allow it to work with a larger variety of legal documents, implementation of multilingual customer support and further improving general understanding of natural language through constant improvement of natural language understanding models. Mediating the collaborative editing and review processes, including the version control tool, the annotation, and the real-time collaboration, will pave the way for more productivity and accuracy for legal teams while the case documents are being handled. Integrating the blockchain technology into the document safety can guarantee the measures against any fraud or violation as well as make document unforgeable. Though the Automatic AI legal document service is considered as the milestone of AI evolution, it actually displays AI level that the generation of legal documents has already passed to the phase of AI functions.

IX. REFERENCES

[1] McCarl, R. "The Limits of Law and AI." University of Cincinnati Law Review Publications 90, no. 3 (2022): 923-950.

[2] Vijippria, R.A. "Critical Study on Artificial Intelligence (AI) in Indian Legal Sectors." Shanlax International Journal of Arts, Science and Humanities 9, no. 4 (2022): 58-64.

[3] Zhong, H., et al. "JEC-QA: A Legal-Domain Question Answering Dataset." In Proceedings of the AAAI Conference on Artificial Intelligence, 34 (2020): 9701-9708.

[4] Robaldo, L., et al. "Introduction for artificial intelligence and law: special issue 'natural language processing for legal texts'." Artificial Intelligence and Law 27, no. 2 (2019): 113-115.

[5] Nay, J.J. "Gov2Vec: Learning Distributed Representations of Institutions and Their Legal Text." In Proceedings of the First Workshop on NLP and Computational Social Science, 49–54. Association for Computational Linguistics, 2016.

[6] Xiao, C., et al. "Lawformer: A pre-trained language model for Chinese legal long documents." AI Open 2 (2021): 79-84.

[7] Qiu, J., et al. "Blockwise Self-Attention for Long Document Understanding." Computational linguistics - Association for Computational Linguistics (Online) (2020): 2555-2565.

[8] Duan, X., et al. "Legal Summarization for Multi-role Debate Dialogue via Controversy Focus Mining and Multi-task Learning." In Proceedings of the 28th ACM International Conference on Information and Knowledge Management, 1361-1370, 2019.

[9] Sugathadasa, K., et al. "Legal Document Retrieval Using Document Vector Embeddings and Deep Learning." In Proceedings of the 2018 Computing Conference, Volume 2, 160-175. 2019.

[10] Zhong, H., et al. "How Does NLP Benefit Legal System: A Summary of Legal Artificial Intelligence." (2020): 5218-5230.

[11] "Patented Machine Learning Technology." February 22, 2024.

[12] "Optimizing Contract Management for In-House Legal Teams - On-Demand Webinar." (n.d.).

[13] "LegalSifter FAQs: Your Questions About AI Legal Services Answered." (n.d.).

[14] "Legal Analytics by Lex Machina." March 9, 2024.

Parkinson's Diseases and its Severity Prediction using Transfer Learning Techniques

Raja Sekar R
Department of Computer Science and Engineering,
Kalasalingam Academy of Research and Education
Virudhunagar, Tamil Nadu, India
r.rajasekar@klu.ac.in

Akilesh S
Department of Computer Science and Engineering,
Kalasalingam Academy of Research and Education
Virudhunagar, Tamil Nadu, India
rmsakilesh23@gmail.com

Navaneenth S S
Department of Computer Science and Engineering,
Kalasalingam Academy of Research and Education
Virudhunagar, Tamil Nadu, India
navaneethss0@gmail.com

Gohula Krishnan K
Department of Computer Science and Engineering,
Kalasalingam Academy of Research and Education
Virudhunagar, Tamil Nadu, India
gohulk73@gmail.com

Jihad K M
Department of Computer Science and Engineering,
Kalasalingam Academy of Research and Education
Virudhunagar, Tamil Nadu, India
kmjihad055@gmail.com

***Abstract*— Parkinson's disease (PD) is a disorder that has significant implications, for patient care and treatment. It is crucial to predict the severity of the disease to provide personalized management. This research study explores the use of transfer learning techniques with medical image datasets to make these predictions. Proposed approach involves utilizing learning models that have been trained on large-scale image datasets. By leveraging the knowledge and features learned from these models the severity of Parkinson's disease can be effectively forecast. The study utilizes Convolutional Neural networks (CNNs) to extract important information from medical images, which is then inputted into a predictive model. The results of our experiments demonstrate the effectiveness of transfer learning in predicting PD severity.**

The proposed study compares trained CNN architectures, such as Simple CNN, Random Forest and XGBoost algorithms, DenseNet, and Inception using a specialized medical image dataset. The resulting model shows accuracy, sensitivity, and specificity making it a valuable tool for assisting physicians in their decision-making processes. Furthermore, the proposed technique not only achieves prediction accuracy but also provides valuable insights into visual indicators of disease progression. This non-invasive and efficient approach has the potential to empower professionals with early diagnosis and personalized treatment strategies for individuals, with Parkinson's disease. This groundbreaking approach shows potential in improving the management of Parkinson's disease and enhancing the quality of life for those who are impacted by it.

Keywords— ***Spiral, Waveform, Maxpooling, Accuracy, Conventional neural network, Healthy, Parkinson's.***

I. INTRODUCTION

Parkinson's disease (PD) is a prominent and increasingly incapacitating neurological condition that requires prompt diagnosis to ensure the best possible care for the patient. This work presents a machine learning model utilizing SPECT DaTSCAN imaging in response to the necessity of early Parkinson's Disease (PD) diagnosis. Proposed model prioritizes interpretability using LIME explanations and achieves a 95.2% accuracy rate using a VGG16 CNN [1]. To achieve a noteworthy 93.84% accuracy with a minimal set of vector features, this reaserch study introduces a machine learning-based diagnostic approach for Parkinson's disease. It emphasizes the potential for early and precise diagnosis to delay the progression of the disease. Support vector machines with regression feature elimination are used in this approach [3]. The approach achieved 96.40% accuracy using a Support Vector Machine classifier, emphasizing the potential of a comprehensive marker combination for early PD detection [6]. This study presents a novel method for classifying magnetic resonance images of healthy and Parkinson's disease patients. It uses a deep convolutional neural network with transfer learning, data augmentation, and GAN-based techniques, achieving an average accuracy of 89.23% and demonstrating improved diagnosis in comparison to current state-of-the-art research [10]. For the application, transfer learning is probably required for several conceptual reasons. One difficulty is the restricted availability of data. To get over this, transfer learning makes it possible to use the expertise of pre-trained models that have been trained on big datasets. Transfer learning makes it possible to initialize the model with weights learned from a larger, more diverse dataset, improving its ability to capture complex patterns. Second, the task of diagnosing Parkinson's disease may involve intricate patterns and relationships within the data that are difficult to capture with a small dataset. Last but not least, transfer learning lessens this load by using pre-trained models as a starting point. Training deep learning models from scratch demands a substantial amount of computing resources and time.

II. LITERATURE SURVEY

The study used a VGG16-based convolutional neural network with transfer learning to classify Parkinson's disease using DaTSCANs, achieving high accuracy. However, there may be some disadvantages, such as data bias, ethical concerns, difficulties with interpretability, the requirement for external validation, clinical adoption considerations, and the requirement for long-term monitoring [1]. To categorize Parkinson's disease, the study used 200 SPECT images and an Artificial Neural Network (ANN) on processed ROIs from Parkinson's and putamen. The model demonstrated high accuracy (94%), perfect sensitivity (100%), and notable specificity (88%); however, potential drawbacks included the need for testing on various populations, difficulties with implementing complex neural networks, the significance of rigorous clinical validation, and ethical considerations regarding patient privacy and potential biases in the data set [2]. The study used machine learning to diagnose Parkinson's disease by classifying and selecting features. Support vector machines with residual feature elimination were shown to be the best-performing method, achieving 93.84% accuracy with minimal vector features. The necessity of assessing model generalization across diverse populations, conducting real-world testing and clinical trials for clinical validation, and addressing ethical concerns about patient data privacy, informed consent, and sensitive medical data use are some of the drawbacks [3]. The assessment demonstrated a notable accuracy of 85.3% in predicting the severity of Parkinson's disease and a high detection accuracy of 98.7%. The necessity for validation in various clinical settings, the assurance of ethical data privacy and informed consent for sensor data, concerns about generalization to new subjects, and the significance of verifying robustness across various environments and demographics for dependable clinical diagnosis are some potential drawbacks, though [4]. This review of the literature explores the use of artificial neural networks (ANNs), especially those trained by backpropagation, in the diagnosis of Parkinson's disease, emphasizing how these systems might revolutionize medical procedures. Studies have shown that ANNs can identify Parkinson's disease with 100% accuracy using a variety of methodological methods, such as preprocessing techniques, dataset selection, and neural network topologies. This highlights the revolutionary influence of ANNs on medical diagnoses [6]. The research presented a new hybrid intelligence system that forecasts Parkinson's disease development. According to the scientists, this method provides improved accuracy and useful applications in clinical situations. However, there are certain important things to keep in mind, such as thorough dataset descriptions and stringent validation processes. In addition, the study tackles ethical issues regarding patient confidentiality and permission, offering valuable perspectives on how this system might be easily incorporated into regular clinical procedures for medical practitioners [7]. With the use of a special classifier, the study achieved a remarkable 99.8% accuracy in properly diagnosing the severity of Parkinson's disease by using speech data. But to get the most out of this approach, you need to have a thorough understanding of the dataset, take care of any ethical issues, and think about how this strategy may be successfully incorporated into clinical practice [8]. For data augmentation, the study used state-of-the-art methods such as generative adversarial networks, deep convolutional neural networks, and transfer learning. Thus, utilizing MRI scans, a remarkable accuracy rate of 89.23% was obtained in correctly detecting Parkinson's disease. The important factor of thorough dataset information and the ethical ramifications for patient data privacy were also considered in this study. The researchers also explored useful aspects to take into account when applying this technology in clinical settings [10].

III. METHODOLOGY

However these predictions are existing ones, but they did not give the best prediction for the image dataset. So, for the best and highest accuracy, we are going to implement the various transfer learning techniques by Simple Convolutional Neural Network, Pytorch Lightning Convolutional Neural Network, and DenseNet201, and in light of this, we would prefer to use the Histogram of Oriented Gradients Image Descriptor in conjunction with the Random Forest Classifier and Xgboost ensemble approach. In the proposed work we are analyzing their various transfer learning techniques and predict the prediction, and briefly analyzing all transfer learning techniques and their methods by visualizations of graphs, tables, and prediction, algorithms and, diagrams.

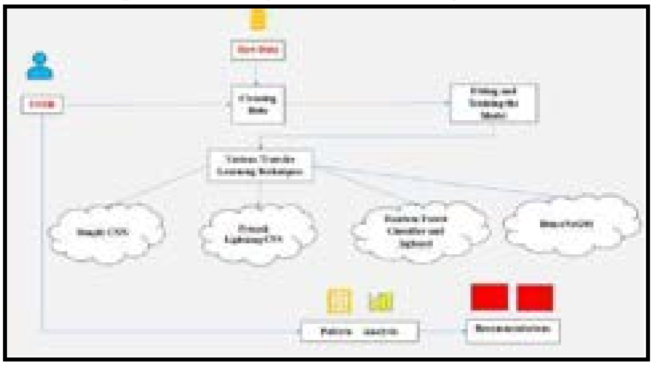


Fig1.Block Diagram

First collect the raw data from the user by using various ways like web portal history, survey forms, etc. Then we want to pre-process the data to get good quality data and then do the exploratory data analysis to get more insights into our data. With cleaned data, we can analyze and process the various transfer learning techniques like Simple Convolutional Neural Network, Pytorch Lightning Convolutional Neural Network, DenseNet201, Random Forest Classifier, and Xgboost ensemble approach.

1. Pytorch Lightning CNN:

PyTorch Lightning is a high-level wrapper for PyTorch that makes model training easier and better by removing repetitious code. In computer vision, convolutional neural networks (CNNs) are the preferred method for image

Classification tasks. These networks use kernels to extract several feature maps by convolving across the picture.

2. Simple CNN:

Convolutional neural networks (CNN/ConvNet), a kind of deep neural network, are trained to interpret visual input through the process of deep learning. Traditional neural networks use matrix multiplications, however, ConvNets use a different strategy called convolution.

3. Ensemble Methods:

Ensemble techniques are a robust and successful methodology in the field of machine learning. These techniques can improve the accuracy of the outcomes by merging various models can greatly enhance the outcomes.

Consider a scenario, where it is more appropriate to apply conventional computer vision techniques, namely HOG image descriptor in combination with strong ensemble approaches like Random Forest and XGBoost.

3.1 Quantify the image using a HOG Descriptor:

A structural descriptor called HOG will identify and measure variations in the input image's local gradient. Naturally, HOG will be able to measure the changes in spiral and wave directions. Additionally, HOG will be able to determine whether these drawings have a greater "shake," which is something we could anticipate from a patient with Parkinson's disease.

3.2 Random Forest:

Several decision trees are created during the training phase of Random Forest, a dynamic and successful ensemble learning technique. Using a process called bootstrapping, random portions of the training data are used to train each tree.

3.3 XGBoost (Extreme Gradient Boosting):

Extreme Gradient Boosting, or XGBoost, is a cutting-edge technique that uses gradient boosting to enhance data. This entails planting trees one after the other, each one intended to fix the mistakes of the one before it. An improved and more precise model is guaranteed by this trimming procedure.

3.4 Workflow:

Feature Extraction:

- Utilize the HOG algorithm to derive features from the images inside your collection.

Ensemble Model Education:

- Set the retrieved HOG features to use to train a Random Forest classifier.
- Utilize the same characteristics to train an XGBoost classifier.

Ensemble Prediction Model:

- To forecast the class of fresh photos, use both models.

Evaluation:

- Examine the group's performance on an independent validation or test set.

4. DensetNet201:

Specifically, the DenseNet201 architecture is a strong pre-trained deep learning model that must be used to classify images to use DenseNet201 for disease detection. DenseNet201's strong convolutional neural network (CNN) has shown itself to be effective in a variety of computer vision tasks, particularly image classification.

IV. EXPERIMENTAL SETUP

This research work implementation includes oftware, packages or libraries, techniques, programming languages, and some mandatory hardware components to do the detection of Parkinson's disease.

<u>Libraries or Packages:</u>

- Scikit Learn
- Numpy
- Pandas
- PyTorch

1. <u>Techniques:</u>

- Simple Convolutional Neural Network.
- PyTorch Lightning Convolutional Neural Network.
- Random Forest and XGBoost.
- Detection By DenseNet201.

V. RESULT ANALYSIS

1. Simple Convolutional Neural Network:

Convolutional Neural Network (CNN) is successfully implemented to identify images associated with Parkinson's disease. Understanding the structure of the dataset might be improved by visualizing it using code. It provides a visual analysis of the images so we can assess their consistency and quality. Ensuring the correctness of label assignments (between healthy and Parkinson's) is crucial, as inaccurate labels have a significant impact on the model's effectiveness.

A CNN model is successfully built that is simple to use and effective by utilizing Keras. This architecture, which consists of two convolutional layers together with max-pooling and thick layers, is a favorite choice for image recognition tasks. The max-pooling layers optimize the process by lowering spatial dimensions and focusing on the most important features, while the convolutional layers are in charge of extracting unique properties from the input images.

Through the use of the well-regarded Adam optimizer and a learning rate of 0.001, we have successfully and carefully balanced stability and efficacy. The difference between the predicted and actual class labels is measured by the binary cross-entropy loss function. In addition, we have carefully included callbacks for learning rate lowering and early halting. The final selection of 48 epochs was made with caution because any more than this leads to a decrease in accuracy. We reduce the danger of overfitting by guaranteeing that the model is not overtrained by our early halting.

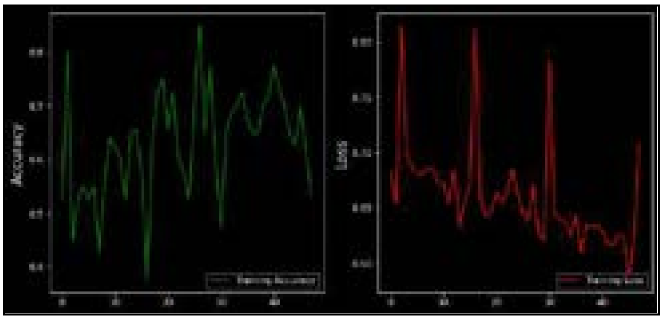


Fig2. Training Accuracy and Loss

A useful visual tool for evaluating the accuracy and loss of your model over time is the training history plot. It enables you to track your model's learning progress carefully and identify areas that need work. Even if the accuracy could differ, it's important to examine the general trend. The model is successfully learning from the data if the accuracy continuously increases and the loss gradually lowers with time.

2. PyTorch Lightning Convolutional Neural Network:

In this research, spiral drawings were used to classify Parkinson's disease using a convolutional neural network (CNN). The procedure had several steps, including the extraction of data, the development of a model, training, evaluation, and visualization. PyTorch and PyTorch Lightning were used for complex deep-learning tasks.

Two separate datasets were constructed with specified root directories, one for training and one for testing, using the datasets.ImageFolder module. Proposed CNN design, the ConvolutionalNetwork, consisted of three fully connected layers, two convolutional layers, and max-pooling. The PyTorch Lightning Trainer was used to train the model for a maximum of 600 epochs. The Adam optimizer was used to optimize the parameters during training, and the cross-entropy loss was computed. Proposed model was prepared to be used for image classification tasks after undergoing extensive training. On the test set, the trained model was evaluated and metrics like recall, accuracy, precision, and loss were calculated. The results proved that the total accuracy was about 66.67%.

	precision	recall	f1-score	support
healthy	0.7273	0.5333	0.6154	15
parkinson	0.6316	0.8000	0.7059	15
accuracy			0.6667	30
macro avg	0.6794	0.6667	0.6606	30
weighted avg	0.6794	0.6667	0.6606	30

Fig3. classification_report

Predictions were generated on the test set using the trained model, and each class's F1 score, precision, and recall were then calculated (healthy and Parkinson). After this, an extensive performance summary was shown, highlighting the overall recall, accuracy, precision, and F1 score.

A comparison investigation revealed that the deployed CNN had a 66.67% overall accuracy rate. Individual measures (precision, recall, and F1-score) that show how well the model distinguishes between classes with and without Parkinson's disease were also included in the evaluation.

3. Random Forest and XGBoost:

This research aims to evaluate if Parkinson's disease (PD) might be reliably classified using drawings and machine learning algorithms. The Histogram of Oriented Gradients (HOG) Image Descriptor was specifically used in the study to extract significant characteristics from spiral and wave drawings. To obtain reliable classification results, these characteristics were then utilized for training two ensemble classifiers: Random Forest (Rf) and XGBoost.

Our study makes use of the Histogram of Oriented Gradients (HOG) Image Descriptor due to the restricted quantity of training data.

We trained two robust ensemble classifiers, Random Forest (Rf) and XGBoost, to successfully categorize spiral and wave paintings. Both models underwent rigorous training before being tested and having their accuracy, sensitivity, and specificity assessed.

	Random Forest	XGBoost
Accuracy	80.00%	73.33%
Sensitivity	73.33%	73.33%
Specificity	86.67%	73.33%

Fig4. Spiral Drawings: Random Forest vs XGBoost Classifier

Now, let's analyze the findings of the research on Spiral Drawings, in which the Random Forest and XGBoost classifiers are compared. Based on investigation, it is analysed that the Random Forest classifier outperformed XGBoost with an accuracy of 80.00%. Both classifiers fared equally well in terms of sensitivity (73.33%), while Random Forest performed better in terms of specificity (86.67% vs. 73.33%) than XGBoost. This demonstrates the general power of Random Forest in accurately recognizing spiral designs.

	Random Forest	XGBoost
Accuracy	80.00%	73.33%
Sensitivity	73.33%	73.33%
Specificity	86.67%	73.33%

Fig5. Wave Drawings: Random Forest vs XGBoost Classifier

In this machine learning approach contrast, we examine the topic of "Wave Drawings," analyzing the respective performances of Random Forest and XGBoost classifiers. With an astounding 80.00% accuracy for Random Forest and a still-respectable 73.33% for XGBoost, the results speak for themselves. Both have similar levels of sensitivity (73.33%), but Random Forest performs with a higher specificity score (86.67%).

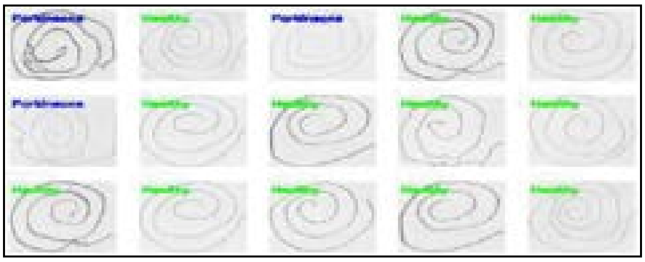


Fig6. Prediction of Spiral Images

Sample images are used to test the trained models, and predictions are shown for spiral and wave drawings.

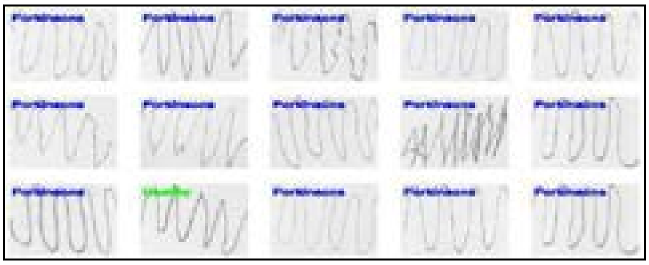


Fig7. Prediction of Wave Images

In spiral and wave drawings, the Random Forest classifier easily leads the XGBoost classifier in terms of specificity and accuracy. Though the classifiers' sensitivity remains constant, there is scope for improvement through additional research and parameter optimization.

4. Detection By DenseNet201:

By the use of transfer learning, we have successfully developed and trained two different models with the DenseNet201 architecture. These models were trained using several datasets that included images of Parkinson's disease, including spiral and wave images. Following that, an output layer with softmax activation for binary classification is added, followed by a Dense layer with ReLU activation. The final output layer generates the anticipated classification based on the softmax function, whereas this Dense layer receives the reduced feature vector and further refines it.

Furthermore, we employ a widely-used loss function known as categorical cross-entropy, which efficiently measures the discrepancy between actual class labels and expected probabilities. Accuracy metric is used, which computes the percentage of accurate predictions made by a binary classification model, to assess performance. We use the fit technique, which accepts enriched training data and target labels, to train the models.

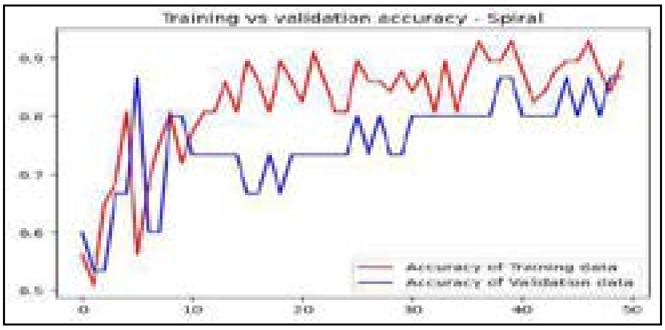


Fig 8. Training vs Validation accuracy – Spiral Drawing

To lower the loss function, this approach iteratively adjusts the model's parameters. The training and validation accuracies and losses are meticulously documented across the epochs, offering important insights into the models' performance and potential areas for improvement.

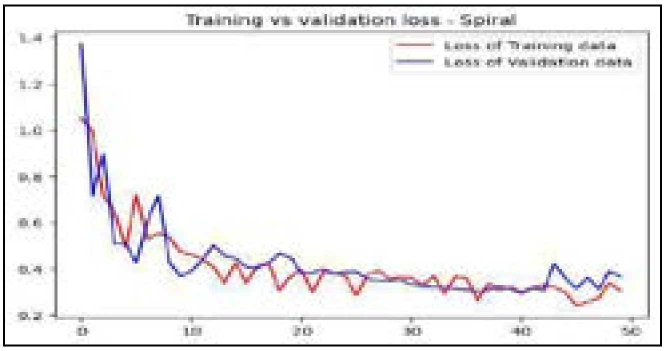


Fig 9. Training vs Validation loss – Spiral Drawing

The training and validation accuracies and losses for each epoch are shown, offering important information for assessing model competency and identifying areas in need of improvement. Notably, both models exhibited good accuracy on the validation sets, indicating that task-specific characteristics were successfully acquired for categorizing paintings of Parkinson's disease.

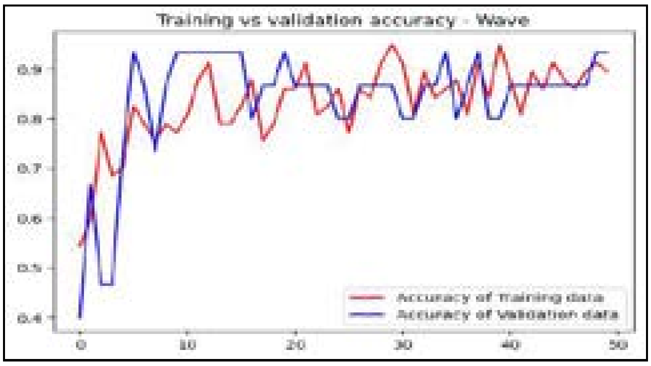


Fig 10. Training vs Validation accuracy – Wave Drawing

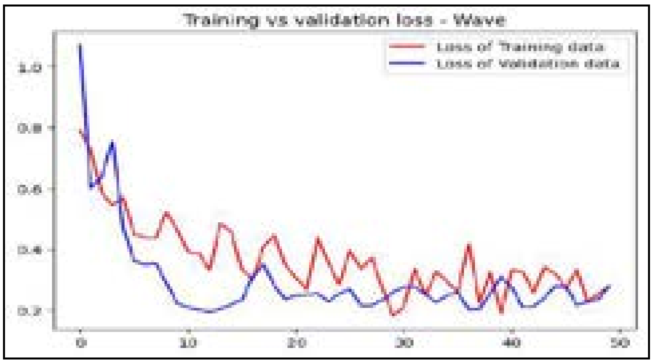


Fig 11. Training vs Validation Loss – Wave Drawing

Proposed models use DenseNet201 architecture and transfer learning is applied to increase their efficency. The models have shown impressive accuracy through validation on an independent test set, indicating that they can recognize significant patterns in the data.

VI. CONCLUSIONS

In this reaserch work, the main goal is to use images to provide patients suffering from Parkinson's disease with accurate and higher-quality predictions. Convolutional Neural Network, PyTorch Lightning Convolutional Neural Network, Random Forest and XGBoost, and DenseNet201 Detection ware discussed in this study. The techniques mentioned above make it simple to identify and determine the severity of Parkinson's disease. The findings proved that training Simple CNN achieves an accuracy of 80%, training PyTorch Lightning CNN achieves an accuracy of 66%, and training Random Forest and XgBoost achieves an accuracy of 80% and 73%, respectively, for both spiral and wave images, and training the DenseNet201 achieves an accuracy of 86% and 67% for spiral and wave images respectively. These various transfer learning techniques can also be used to analyze various disease detection and prediction for several diseases.

VII. FUTURE SCOPE

The integration of data from multiple sources with transfer learning algorithms is part of the future scope to enhance the efficiency and accuracy of disease prediction. This will make possible the development of hybrid models and enable real-time monitoring, allowing for a more robust and comprehensive approach. Additionally, the objective is to emphasize interpretability, carry out extensive validations, and adjust to distant healthcare settings in addition to broadening the scope of disorders for which these approaches may be used. Encouraging collaboration in clinical settings to improve disease prediction model performance. This will ultimately change the use of these models in a variety of healthcare domains.

REFERENCES:

[1] Rajkumar Magesh, P., Delwin Myloth, R., & Tom, R. J. (2020). An Explainable Machine Learning Model for Early Detection of Parkinson's Disease using LIME on DaTscan Imagery. arXiv e-prints, arXiv-2008.

[2] Rumman, M., Tasneem, A. N., Farzana, S., Pavel, M. I., & Alam, M. A. (2018, June). Early detection of Parkinson's disease using image processing and artificial neural network. In 2018 Joint 7th International Conference on Informatics, Electronics & Vision (ICIEV) and 2018 2nd International Conference on Imaging, Vision & Pattern Recognition (icIVPR) (pp. 256-261).

[3] Senturk, Z. K. (2020). Early diagnosis of Parkinson's disease using machine learning algorithms. Medical hypotheses, 138, 109603.

[4] El Maachi, I., Bilodeau, G. A., & Bouachir, W. (2020). Deep 1D-Convnet for accurate Parkinson's disease detection and severity prediction from gait. Expert Systems with Applications, 143, 113075.

[5] Prashanth, R., Roy, S. D., Mandal, P. K., & Ghosh, S. (2016). High-accuracy detection of early Parkinson's disease through multimodal features and machine learning. International journal of medical informatics, 90, 13-21.

[6] Sadek, R. M., Mohammed, S. A., Abunbehan, A. R. K., Ghattas, A. K. H. A., Badawi, M. R., Mortaja, M. N., ... & Abu-Naser, S. S. (2019). Parkinson's disease prediction using artificial neural network.

[7] Nilashi, M., Ibrahim, O., & Ahani, A. (2016). Accuracy improvement for predicting Parkinson's disease progression. Scientific reports, 6(1), 34181.

[8] Sarankumar, R., Vinod, D., Anitha, K., Manohar, G., Vijayanand, K. S., Pant, B., & Sundramurthy, V. P. (2022). Severity Prediction over Parkinson's Disease Prediction by Using the Deep Brooke Inception Net Classifier. Computational Intelligence and Neuroscience, 2022.

[9] Saunders- Pullman, R., Derby, C., Stanley, K., Floyd, A., Bressman, S., Lipton, R. B., ... & Pullman, S. L. (2008). Validity of spiral analysis in early Parkinson's disease. Movement disorders: Official Journal of the Movement Disorder Society, 23(4), 531-537.

[10] Kaur, S., Aggarwal, H., & Rani, R. (2021). Diagnosis of Parkinson's disease using deep CNN with transfer learning and data augmentation. Multimedia Tools and Applications, 80, 10113-10139.

[11] Wingate, J., Kollia, I., Bidaut, L., & Kollias, S. (2020). Unified deep learning approach for prediction of Parkinson's disease. IET Image Processing, 14(10), 1980-1989.

[12] Bae, Y. J., Kim, J. M., Sohn, C. H., Choi, J. H., Choi, B. S., Song, Y. S., ... & Kim, J. H. (2021). Imaging the substantia nigra in Parkinson's disease and other Parkinsonian syndromes. Radiology, 300(2), 260-278.

[13] Alagarsamy, S., Akilesh, S., Jihad, K. M., & Navaneenth, S. S. (2023, May). Collaborative Recommendation System For Gadgets. In *2023 4th* International Conference for Emerging Technology (INCET) (pp. 1-5). IEEE.

[14] Tunc, H. C., Sakar, C. O., Apaydin, H., Serbes, G., Gunduz, A., Tutuncu, M., & Gurgen, F. (2020). Estimation of Parkinson's disease severity using speech features and extreme gradient boosting. Medical & Biological Engineering & Computing, 58, 2757-2773.

[15] Abdurrahman, G., & Sintawati, M. (2020, May). Implementation of xgboost for classification of Parkinson's disease. In Journal of Physics: Conference Series (Vol. 1538, No. 1, p. 012024). IOP Publishing.

Advanced Authentication System with Fraud Resistant Features

Rahul [1], Ramakrishnan [2], Albert Mayan J [3], Ashok Kumar K [4], Dhamodaran S [5]

[1,2] U.G Student, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology , Chennai

[3,4,5] Professor , School of Computing , Sathyabama Institute of Science and Technology , Chennai

rahul.prameshwaran@gmail.com, rkrish.1701@gmail.com, albertmayan@gmail.com, ashokkumar..cse@sathyabama.ac.in, dhamodaran.cse@sathyabama.ac.in

Abstract - A novel technique for verifying signatures and certificates is introduced in this abstract with the goal of enhancing authentication and detecting fraud more effectively. There is a greater chance of identity theft and fraudulent activities due to the fact that conventional ways of confirming signatures and certificates are becoming less effective as technology advances. In order to overcome these obstacles, our system employs state-of-the-art technology like AI and machine learning algorithms to improve the precision and effectiveness of the verification procedure. In order to validate documents and swiftly identify any fraud attempts, the proposed approach examines information, trends, and abnormalities in digital certificates and signatures. Additionally, the proposed solution utilizes blockchain technology to minimize the possibility of manipulation or tampering by ensuring the immutability and transparency of validated certificates and signatures. The proposed technology enhances user experience and convenience by offering an interface that is easy to use and supporting the integration of current authentication systems. After undergoing rigorous testing and assessment, the proposed system has shown to be an effective and dependable option for verifying signatures and certificates. The proposed signature and certificate verification system includes sophisticated tools to fight fraud and identity theft, giving businesses and people a safe and effective way to verify their identities.

Keywords: *Certificate Verification, Signature Verification, Authentication, Fraud Detection, Artificial Intelligence, Machine Learning, Blockchain Technology, Identity Theft, Tampering, Manipulation, User-Friendly Interface, Integration, Robust, Reliable***.**

I. INTRODUCTION

Presenting a state-of-the-art signature and certificate verification technology will revolutionize authentication while effectively combatting fraud. This advanced system offers organizations and individuals unparalleled safety and peace of mind, blending modern technology with robust processes to ensure the highest level of security and accuracy in certificate and signature verification. At the heart of this system are modern AI algorithms that are trained to detect complex patterns, minute details, and unique characteristics present in digital signatures and certifications. Leveraging deep neural networks and machine learning, the system can verify the legitimacy of signatures and documents far more rapidly and precisely than manual methods.

This cutting-edge solution automates authentication and fraud detection processes, greatly simplifying the procedures compared to traditional verification techniques reliant on subjective judgments and human inspections. Consistent and reliable results are guaranteed each time as the system digitally scans and analyzes certificates and signatures, eliminating the risk of bias and human error. Integral to the system is a comprehensive database containing certified documents and verified signatures amassed over the years. By comparing and cross-referencing submitted certificates and signatures with validated ones, the system operates with enhanced accuracy and efficiency, continually updated to address emerging fraudulent practices and new counterfeit methods. Furthermore, the system has a smart design, facilitating straightforward verification for businesses and individuals alike. Users can easily submit their certificates or signatures for verification and receive prompt results, thanks to the user-friendly interface. Whether confirming the validity of a signature on a legal agreement, contract, or financial transaction, or authenticating diplomas, licenses, contracts, or any other certified document, this system offers a dependable and convenient solution.

As identity fraud, forgeries, and counterfeit certifications proliferate, it becomes imperative to implement safeguards that introduce the utmost security.

II. RELATED WORKS

Security issues in routing and anomaly detection for space information networks have gained increasing attention [Zhuo et al., 2021]. Researchers have explored various approaches to tackle these challenges and improve the security of space networks. Continuous authentication and anomaly detection using mouse clickstream data analysis have been investigated to enhance online security [Almalki et al., 2021]. The study evaluates the effectiveness of this approach and its potential applications in detecting unauthorized access or malicious activities.

A new deep anomaly detection-based method for user authentication using multichannel surface EMG signals of hand gestures has been proposed [Li et al., 2022]. This method aims to improve the accuracy and security of user authentication by analyzing electro-myographic signals. Off-line signature verification using elementary combinations of directional codes from boundary pixels has been researched [Ajij et al., 2021]. The study explores the application of this technique in authenticating signature images and detecting possible forgeries.

The Anti-fraud Tensorlink4cheque (AFTL4C) is an advanced AI/ML solution designed to detect and prevent cheque fraud [Uyyala and Yadav, 2023]. This proprietary technology utilizes machine learning algorithms to analyze cheque data and identify potential fraudulent activities. DOC-BLOCK is a blockchain-based authentication system developed for digital documents [Imam et al., 2021]. The system leverages the security and immutability of blockchain technology to ensure the authenticity and integrity of digital documents.

A digital certificate authority with blockchain cybersecurity in education has been proposed to enhance the security of academic certificates [Maulani et al., 2021]. This approach utilizes blockchain technology to create tamper-proof and verifiable digital certificates. Cerberus is a blockchain-based accreditation and degree verification system that aims to provide a secure and decentralized method for verifying academic credentials [Tariq et al., 2022]. The system leverages the transparency and immutability of the blockchain to prevent fraud and ensure the validity of certificates.

Blockchain technology has been reviewed in the context of academic certificate authenticity [Kumutha and Jayalakshmi, 2022]. The study explores the potential of blockchain to combat certificate fraud and improve the trustworthiness of academic qualifications. An academic certificate fraud detection system framework has been proposed that utilizes blockchain technology to enhance the security and reliability of academic certificates [Lutfiani et al., 2022]. The framework aims to detect and prevent fraud by leveraging the distributed ledger system and cryptographic mechanisms provided by blockchain technology.

III. EXISTING SYSTEM

When it comes to authenticating users and detecting fraud, the current method for Innovative certificate and signature verification has a few drawbacks. To begin with, the system is very susceptible to forgeries and theft since it is based on actual certificates and signatures. This is a major security concern because bad actors may forge certificates and signatures to impersonate trusted users or get unauthorized access. Second, it's expensive and time-consuming to validate physical certificates and signatures. In order to verify the given papers, skilled staff must manually check them with the original records. Organizations incurring higher operating costs as a result of having to devote resources to this activity also see a slowdown in the authentication process.

Also, new technologies can't be easily integrated into the current system because of its rigid design. Physical certifications and signatures limit innovation in an era when digital transformation is transforming many parts of our life. It makes it harder to include new technology that might improve the efficiency and security of certificate and signature verification, such biometrics or blockchain. The inability to monitor and identify fraudulent actions in real-time is another drawback of the current system. Keeping tabs on who has accessed or misused physical certificates and signatures may be a challenge. Because of this, fraud is more likely to go undetected for long periods of time, which might harm people, businesses, or even whole systems.

IV. PROPOSED SYSTEM

Using state-of-the-art technology, the new certificate and signature verification system is going to change the game when it comes to authentication and detecting fraud. In order to guarantee absolute verification and security against fraudulent operations, this system incorporates biometric identification techniques, blockchain technology, and machine learning algorithms. To start, digital certificates are analyzed and validated by the system using machine learning techniques. In order to identify any irregularities or efforts at manipulation, these algorithms can intelligently check the validity and authenticity of certificates. The system provides a trustworthy verification method by detecting changed or counterfeit certificates by comparing the certificate data with known patterns. Second, to make the verification process more secure and transparent, the system uses blockchain technology. Blockchain technology allows for the establishment of a distributed ledger that cannot be altered, guaranteeing the authenticity of all certificate transactions. The digital signatures and storage in a distributed ledger of each certificate make it very difficult, if not impossible, to alter or manipulate the data. This decentralized method boosts confidence in the system and decreases the likelihood of fraud by doing away with middlemen.

The technology further strengthens the verification process by using biometric authentication technologies. Biometric authentication allows users to verify their identification using their own distinct characteristics, including fingerprints, iris scans, or face recognition. An extra safeguard against impersonation or identity theft, these biometric features are very secure and hard to reproduce. To further combat fraudulent actions, the system is equipped with real-time fraud detection tools. Anomalies, such questionable certificate requests or strange use patterns, may be caught via the system's constant monitoring of transactions and user activity. Upon detection, the systemhas the capability to send out notifications or immediately respond in order to reduce risks and safeguard against possible fraud.

The proposed method for verifying signatures and certificates is a new and inclusive approach that integrates biometric authentication, blockchain technology, and machine learning techniques to provide strong authentication and fraud detection capabilities. Using these cutting-edge technologies, the proposed system creates a trustworthy environment for people and enterprises by avoiding unwanted access and guaranteeing the authenticity of signatures and certificates.

V. SYSTEM ARCHITECTURE

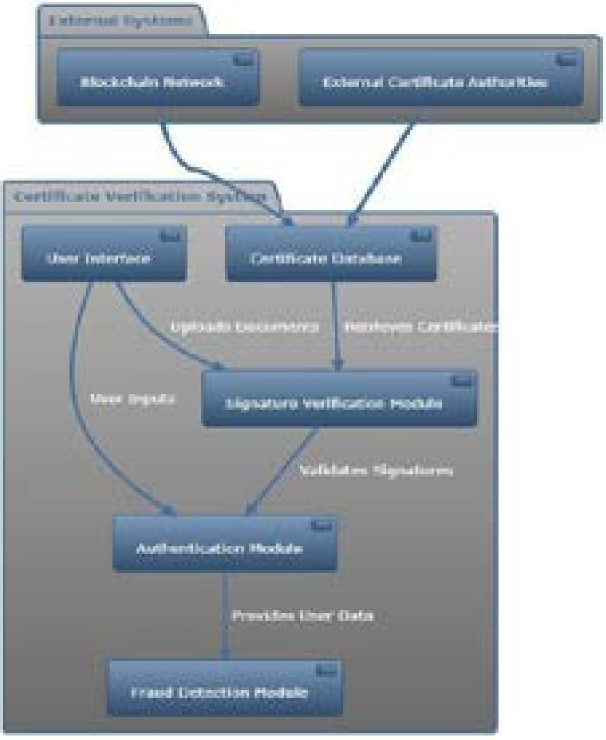


Fig. 1. System Architecture

The architecture diagram depicted in fig 1 presents an advanced system to authenticate digital certificates and detect fraud. This system integrates with two external entities: a Blockchain Network, which provides a secure and immutable ledger for certificate verification, and External Certificate Authorities, which are trusted entities that issue and manage digital certificates.

The certificate verification system is composed of several key components; they are:

User Interface: This is the front-end through which users interact with the system. It allows users to submit certificates for verification and to input necessary data for the process.

Certificate Database: A crucial repository, where digital certificates and their related data are stored. This database is referenced to verify the details of a certificate against its stored credentials.

Signature Verification Module: This module is responsible for the validation of the digital signatures attached to the certificates. It ensures that the signature is genuine by comparing it with a known valid signature from the Certificate Database.

Authentication Module: This part of the system authenticates the user requesting the verification, ensuring that they are authorized to access or query the certificate information.

Fraud Detection Module: Operating as a safeguard, this module analyzes patterns and verifies the authenticity of the certificates to detect any fraudulent activity that might be attempted.

Figure 1 depicts a flow of information between these components, illustrating the thorough process by which certificates are authenticated. The Blockchain Network and External Certificate Authorities provide a foundation of trust for the system, while the internal components work in synergy to perform the verification process, authenticate users, and detect fraud, ensuring that the certificates are valid and secure.

VI. METHODOLOGY

1. **Certificate Generation Module**:
This module is responsible for generating innovative certificates that can be used for authentication and fraud detection. The system will generate unique digital certificates for individuals or organizations, ensuring the authenticity and integrity of the information. The module will include features such as biometric identification, in which the system captures and stores unique physical or behavioral characteristics of the user. It will also incorporate advanced cryptographic algorithms to protect the certificates from tampering or forgery. Additionally, the module will have an intuitive user interface that allows administrators to easily generate and manage certificates, including customizing the information and design, and specifying the level of security required.

2. **Signature Verification Module**:
The signature verification module is crucial for ensuring the authenticity of documents and transactions. It will employ innovative techniques to accurately verify the legitimacy of signatures on various forms, contracts, and other important documents. The system will capture signatures electronically and store them securely. Advanced algorithms will then be used to compare the captured signature with the stored signature to determine the level of similarity and authenticity. The module will also incorporate machine learning capabilities to continuously improve accuracy over time. Moreover, it will provide a user-friendly interface for individuals and organizations to easily submit and verify signatures, as well as generate comprehensive reports on verification outcomes.

3. **Fraud Detection Module**:
The fraud detection module is designed to detect and prevent fraudulent activities within the system. It will employ advanced data analytics and machine learning algorithms to analyze patterns, trends, and anomalies in the certificates and signatures. The module will leverage big data capabilities to process large volumes of information, including cross-referencing and comparing with external databases to identify potential cases of fraud. It will also incorporate real-time monitoring to flag suspicious activities, such as multiple attempts to use the same certificate or an unusual number of signature verifications within a short period. The module will generate alerts and notifications to administrators, enabling them to take immediate action and prevent fraudulent activities from occurring. Additionally, the module will provide comprehensive reporting and analysis tools to assist in identifying trends and improving the overall security of the system.

VII. DISCUSSION

Strong authentication and fraud detection capabilities are included into the revolutionary signature and certificate verification system. The authenticity and integrity of certificates and signatures are guaranteed by the system via the use of cutting-edge technology. The use of digital certificates is a crucial component of this system. An individual's digital identity may be represented by these certificates, which are issued by reliable agencies. The system verifies the authenticity of a certificate by comparing it to the digital signature of the issuing authority. That way, you know the certificate is authentic and hasn't been tampered with.

The systemuses state-of-the-art signature verification methods in addition to digital certificates. In order to determine the signature's legitimacy, it examines its features, including stroke dynamics, pressure, and speed. The system can tell whether a signature is real or fake by comparing it to the authorized user's baseline. The technology uses machine learning techniques to improve fraud detection even more. Certificate and signature fraud trends and behaviors are constantly being analyzed by these algorithms. Anomalies may be detected and suspicious activity can be flagged, allowing for proactive intervention to stop fraudulent operations. In addition, the system keeps a detailed record of all actions taken to verify signatures and certificates. If fraud is discovered, investigations and legal actions may be aided by this traceability and responsibility.

Authentication and fraud detection are both made easier with this cutting-edge technology that verifies signatures and certificates. Ideal for organizations and corporations seeking to secure against fraudulent activity, it combines digital certificates, sophisticated signature verification methods, machine learning algorithms, and a full audit trail to assure high levels of security and dependability.

VIII. CONCLUSION

Finally, a novel approach to authentication and fraud detection is offered by the signature and certificate verification system. The proposed system uses cutting-edge tech like machine learning and blockchain to make sure that certificates and signatures are secure and legitimate. In order to identify any fraudulent actions and ensure that papers are real, it provides a simplified and fast method. Data analysis and encrypted records greatly lessen the possibility of identity theft and illegal transactions. Additionally, this system is accessible and handy for consumers and enterprises due to its user-friendly design and smooth connection with numerous platforms. In conclusion, this approach is an effective and trustworthy instrument for combating fraud in the modern digital age.

IX. FUTURE WORK

There must be a solid method for verifying digital certificates and signatures in the ever-changing digital world. An innovative solution is needed by enterprises to reliably identify and verify fraudulent activity, especially with the development of remote labor, online transactions, and fraud overall. Our future work will focus on creating a complex system that integrates cutting-edge technologies like blockchain, AI, and ML to guarantee the authenticity of certificates and signatures. Digital certificates produced by legitimate organizations and not altered in any way will be verified by the system's algorithms. Furthermore, it will make use of deep learning models to identify forgeries or impersonations by precisely matching signatures with the ones on file. Integrating blockchain technology will also increase the system's transparency, immutability, and security, making it harder for fraudsters and tamperers to compromise. The main goal of this work is to aid in the creation of a novel authentication system that can help organizations protect sensitive data, stop fraud, and guarantee the authenticity of digital certificates and signatures through rigorous testing, prototyping, and research.

REFERENCES

[1] Zhuo, M., Liu, L., Zhou, S., & Tian, Z. (2021). Survey on security issues of routing and anomaly detection for space information networks. Scientific Reports, 11(1), 22261.

[2] Sankari.A,Mayan A.J,"Retrieving call logs and SMS by messaging services",International Journal of Pharmacy & Technology,Vol. 8 , Issue No.4 , pp.22951-22958,Dec 2016.

[3] Almalki, S., Assery, N., & Roy, K. (2021). An empirical evaluation of online continuous authentication and anomaly detection using mouse clickstream data analysis. Applied Sciences, 11(13), 6083.

[4] Julian Menezes .R, Albert Mayan .J, M. Breezely George, Development of a Functionality Testing Tool for Windows Phones , Indian Journal of Science and Technology,Vol:8,Issue:22,pp: 1-7, 2015.

[5] Li, Q., Luo, Z., & Zheng, J. (2022). A new deep anomaly detection-based method for user authentication using multichannel surface EMG signals of hand gestures. IEEE Transactions on Instrumentation and Measurement, 71, 1-11.

[6] Albert M J,Surya.B,Pranoy Prabhakar, "Department Student Library Using Twig Pattern Query Processing Over Admin-User Login Privilege", Pak. J. Biotechnol. Vol. 13 , Pp. 489 - 493,(2016)

[7] Ajij, M., Pratihar, S., Nayak, S. R., Hanne, T., & Roy, D. S. (2021). Off-line signature verification using elementary combinations of directional codes from boundary pixels. Neural Computing and Applications, 1-18.

[8] Uyyala, P., & Yadav, D. C. (2023). The advanced proprietary AI/ML solution as Anti-fraudTensorlink4cheque (AFTL4C) for Cheque fraud detection. The International journal of analytical and experimental modal analysis, 15(4), 1914-1921.

[9] Ravipati A,Kondamuri R.K,Mary Posonia A,John M A(2023),"Vision Based Detection and Analysis of Human Activities",International Conference on Trends in Electronics and Informatics,pp. 1542-1547

[10] Imam, I. T., Arafat, Y., Alam, K. S., & Shahriyar, S. A. (2021, February). DOC-BLOCK: A blockchain based authentication system for digital documents. In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 1262-1267). IEEE.

[11] Maulani, G., Gunawan, G., Leli, L., Nabila, E. A., & Sari, W. Y. (2021). Digital certificate authority with blockchain cybersecurity in education. International Journal of Cyber and IT Service Management, 1(1), 136-150.

[12] Tariq, A., Haq, H. B., & Ali, S. T. (2022). Cerberus: A blockchain-based accreditation and degree verification system. IEEE Transactions on Computational Social Systems.

[13] Kumutha, K., & Jayalakshmi, S. (2022). Blockchain technology and academic certificate authenticity—A review. Expert Clouds and Applications: Proceedings of ICOECA 2021, 321-334.

[14] Lutfiani, N., Apriani, D., Nabila, E. A., & Juniar, H. L. (2022). Academic certificate fraud detection system framework using blockchain technology. Blockchain Frontier Technology, 1(2), 55-64.

Analysis of ECG signal and Non-Invasive Blood Glucose Monitoring using Neural Networks

CH Gangadhar,
PVP Siddhartha Institute of Technology
Vijayawda, India
Gangadharch1111@pvpsiddhartha.ac.in

P Venu Madhav,
PVP Siddhartha Institute of Technology
Vijayawda, India
venumadhav@pvpsiddhartha.ac.in
ORCID:0000-0003-1349-4014

Korivi Divya Naga Durga
PVP Siddhartha Institute of Technology
Vijayawda, India
20501A0448@pvpsit.ac.in

Golla Kamaraju
PVP Siddhartha Institute of Technology
Vijayawada, India
20501A0431@pvpsit.ac.in

Jayana Thanmaya Ramya Krishna
PVP Siddhartha Institute of Technology
Vijayawada,India
20501a0436@pvpsit.ac.in

Kandukuri Gagan Sai Kowshik
PVP Siddhartha Institue of Technology
Vijayawda, India
20501a0441@pvpsit,ac.in

***Abstract*—The metabolic condition known as diabetes, or Diabetes Mellitus (DM), is defined by consistently higher levels of blood glucose over an extended period of time. Diabetes management requires adequate blood glucose monitoring. Proper monitoring of blood glucose levels is essential for managing diabetes. Strong methods for tracking blood glucose levels have been the focus of research efforts. The amount of glucose in the blood can be estimated via NIR-spectroscopic measurement results. We first employed convolutional neural networks, yielding an r square value of 95 and an MAE of 0.16. Based on the numerical results, a non-intrusive blood sugar monitoring that incorporates near-infrared spectroscopy and an appropriate processing method emerges as a potential solution. The World Health Organization (WHO) conducted research revealing that the cardiovascular diseases (CVD) are the primary cause of death worldwide Cardiovascular-diseases (CVDs) account for almost 17.9 million deaths annually, or 31% of all fatalities globally. Experts believe that heart attacks and strokes account for 85 of these fatalities. One of the most common disorders that affect people is cardiac arrhythmia, which can result in life-threatening illnesses including stroke or cardiac arrest. Due to the asymptomatic nature of arrhythmia in many situations, one cannot determine whether they have it until they are checked by a professional. An ECG is a regularly used tool for monitoring a patient's cardiac rhythm. The purpose of this work is to describe a technique that can be used to categorize immediately. We train and assess the proposed Convolutional Neural Networks (CNN) model on the MIT-BIH Database, yielding a 99% accuracy rate.**

Keywords: Diabetes Mellitus (DM), cardiovascular diseases (CVD), disease detection, ECG, Convolutional Neural Networks (CNN)

I. INTRODUCTION

Diabetes stands as one of the primary health challenges confronting the world today. Projections suggest that by 2030, the condition will affect over 366 million individuals globally and rank among the top six causes of death in the United States [2]. In 2000, 171 million people were grappling with the disease. Diabetes is an incurable condition [3] characterized by either hypoglycemia, or insufficient glucose levels in the blood, and hyperglycemia, or elevated blood sugar levels, stemming from the body's inadequate production of insulin [4,5]. Insulin, produced by the pancreas, initiates cells' absorption of glucose, thereby lowering blood glucose levels [6, 7]. Three types of diabetes are recognized: gestational diabetes, Type I diabetes mellitus, and Type II diabetes mellitus. Monitoring blood glucose levels represents a new area of diabetes research. Traditional testing and monitoring systems are inconvenient, posing daily discomfort for individuals.

To streamline diabetes monitoring and treatment, and alleviate the pain, distress, and financial burden of the patients, researchers have concentrated on the developing techniques and equipment for continuous blood glucose monitoring. Over recent years, researchers have explored numerous technologies employing sensors to monitor blood glucose levels [3, 5, 8]. Both non-enzymatic and enzymatic electrochemical glucose sensing methods have been investigated [9]. Researchers have found infrared light spectroscopy to be generally accurate in measuring blood glucose levels [9-10]. Consequently, the use of spectroscopy for this purpose is gaining traction. This project aims to utilize the SparkFun kit with near-infrared spectroscopy to ascertain glucose levels in the blood. A suitable deep learning approach is employed to analyze data accurately predicting blood glucose levels

The analysis of electrocardiograms (ECGs) plays a pivotal role in diagnosing and treating various cardiac disorders, including arrhythmias. Precisely identifying and categorizing arrhythmias using ECG signals is a critical task in clinical practice. Traditional arrhythmia detection techniques often rely on handcrafted features and rule-based algorithms, which may lack robustness and generalizability across different patient demographics and ECG recording conditions. In recent years, deep learning approaches have demonstrated promising performance in various medical signal analysis applications, including ECG arrhythmia identification. By automatically learning discriminative features from raw ECG signals, convolutional neural networks (CNNs) and recurrent neural networks (RNNs) have significantly improved the efficiency and accuracy of arrhythmia classification. We review recent developments in deep learning-based ECG arrhythmia detection, exploring a range of network topologies and preprocessing methods to assess their generalization and performance for different ECG tasks, including arrhythmia identification [11].

One unique method for categorizing ECG arrhythmias involves utilizing softmax regression and stacked sparse autoencoders. This approach demonstrated competitive performance compared to conventional feature-based techniques, effectively learning hierarchical features from ECG signals [12]. Additionally, researchers presented a deep learning model that combines recurrent and convolutional neural networks to classify ECG arrhythmias, outperforming traditional machine learning techniques and exhibiting robust performance across various arrhythmias [13]. These research works highlight the potential of deep learning in enhancing diagnostic accuracy and clinical outcomes. They discuss various network architectures, preprocessing techniques, and challenges in ECG arrhythmia identification [14]. The primary objective of this study is to classify cardiac disorders using the MIT-BIH database. By employing appropriate deep learning techniques, data analysis aims to accurately diagnose disorders. Section II provides a brief review of the scientific background, while Part III explains the process in detail. Section IV presents the results and models used to support the proposed concept. Finally, Section V concludes the research work.

II. SCIENTIFIC BACKGROUND

The term "near-infrared region" refers to the portion of the electromagnetic spectrum between 780 and 2500 nm in wavelength [12]. When considering mid-infrared spectra, NIR sample. Some near-infrared (NIR) light is partially absorbed and scattered due to interactions between the sample's elements and the incident light. The following is an explanation of the light attenuation equation [13]. Since the objective of this research work is to utilize near-infrared spectroscopy data to estimate blood sugar levels, the primary task is to identify the most suitable model that provides an optimally precise fit to the data collection. Numerous multivariate studies, including those employing machine learning algorithms and deep neural networks, have been conducted. Neural networks are employed for automated ECG signal processing using the MIT-BIH Arrhythmia Database. In tasks involving pattern identification and classification, neural networks offer a robust foundation as they can discern intricate relationships within the data. When trained on annotated ECG recordings, neural networks demonstrate the capability to accurately identify and categorize arrhythmias, facilitating diagnostic decision-making. By employing neural networks to analyze this dataset, arrhythmia diagnosis can be rendered more precise and automated, potentially leading to improved patient outcomes and clinical treatment..

III. METHODOLOGY

A. *Non-invasive Glucose*

This methodology encompasses several key steps or developing and evaluating a regression model using TensorFlow data pre-processing method. Firstly, data pre-processing involves loading the dataset and normalizing it to ensure the feature homogeneity.

$$X\_Normalixed = \frac{x - \mu}{N} \qquad (1)$$

Neural Network Architecture:

The next step is to build the model's architecture using ReLU activation functions and dense layers. For deep learning models, the Rectified Linear Unit is the highly preferred activation function. If the input is negative, the function will return 0. Otherwise, it will provide the value it obtained, which is x in this case, if the input value is a positive number. As a result, it may be expressed as

$$\text{ReLU(x)} = \begin{cases} 0, & if\ x < 0 \\ x, & for\ x => 0 \end{cases} \qquad (2)$$

The model is assembled using the Adam optimizer and the Mean Squared Error (MSE) loss function. Gradient descent is used during training process to modify the model's parameters.

$$\text{MSE} = \frac{1}{\text{n}} \sum_{\text{i}=0}^{\text{n}} \left(\text{y}_{\text{true}} - \text{y}_{\text{pred}}\right)^2 \qquad (3)$$

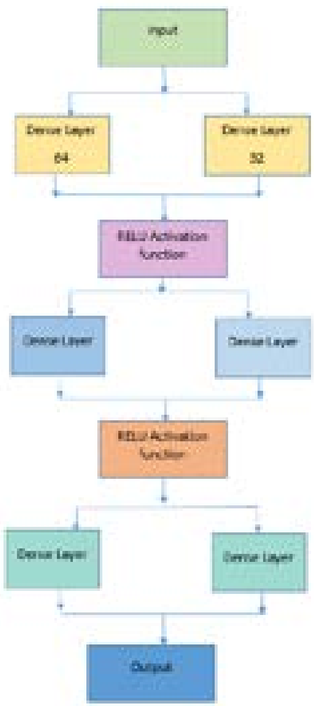


Fig. 1. Neural Network Architecture for Non-Invasive Glucose Monitoring

Model Performance

Evaluation metrics include mean absolute error (MAE) and R-squared respectively.

$$\text{MAE} = \frac{1}{n} \sum_{i=0}^{n} \left|\text{y}_{\text{true}} - \text{y}_{\text{pred}}\right|^2 \qquad (4)$$

The coefficient of determination, or R-squared, is a statistical measure used in regression models to determine the extent to which the independent variable explains the dependent variable's variance. Just to clarify, r-squared is a way to see how well the data fits the regression model.

B. ECG

The following method implements a convolution neural network (CNN) for classifying heartbeat signals into five categories using the MIT-BIH Arrhythmia Database.

Table-1 Class and Category Mapping

Class	Category
N	N is the normal heart rate. (N) Block beats on the left and right bundle branches (L, R) Beat of atrial ejection (e) The escape beat at the nodes' junctions (j)
F	Atrial fibrillation (A) Abnormally early heart rate in the atrium (a) Premature nodal beat (J) Supraventricular preterm heart rate
V	An early constriction of the ventricles Ventricular ejection fraction (E)
F	Combined ventricular and regular heart rate (F)
Q	Paced beat (/) A combination of regular beat and paced Uncategorizable rhythm (U)

Data Pre-Processing

The dataset is pre-processed by separating the features (ECG data) and labels (heartbeat categories). The features are adjusted so that their standard deviation is one and their mean is zero. A training set and validation set are then created from the dataset. Due to class imbalance, each class's weight is calculated inversely proportional to its frequency, ensuring balanced training

$$X\_Normalized = \frac{x-\mu}{\sigma} \qquad (5)$$

Neural Network Architecture

The CNN architecture consists of convolution blocks followed by max-pooling layers for feature extraction. Batch normalization and GELU activation functions are utilized for improved performance. The model is trained using cross-entropy loss, which measures the difference between predicted and actual class distributions.

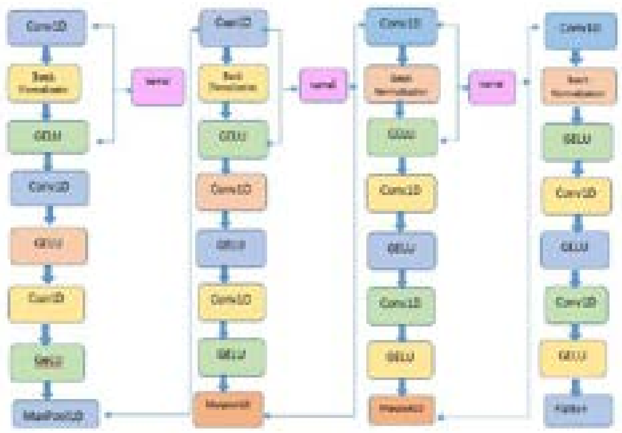


Fig. 2. Neural network architecture of ECG signal

IV. RESULTS & DISCUSSION

By implementing the proposed neural network architecture, the evaluation metrics such as the Mean Absolute Error (MAE) and R-Square error are depicted in Table 2. Further, the precision, accuracy and recall values are depicted in Table 3. Figure 3 shows a comparison on the training and validation loss. Figure 4 shows a comparison on the actual and predicted values for non-invasive glucose monitoring procedure. From figure 4, it is evident that the predictions of the proposed model are highly accurate. The QQ plot is given in figure 5. Finally, the figures 6 and 7 shows the accuracy and loss comparison. These graphs are plotted based on the confusion matrix depicted in figure 8.

Table-2 Evaluation metrics for non-invasive glucose monitoring

Evaluation metrics	MAE	R-square
Values	0.16	0.95

Table-3 Evaluation metrics for ECG signal analysis

classes	Accuracy	Precision	Recall
N	0.99	1.00	0.99
S	0.95	0.81	0.88
V	0.97	0.96	0.97
F	0.90	0.75	0.82
Q	0.99	0.99	0.99

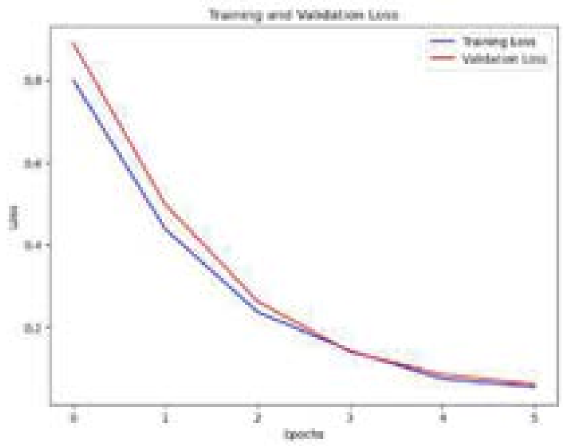


Fig. 3. Training vs Validation loss

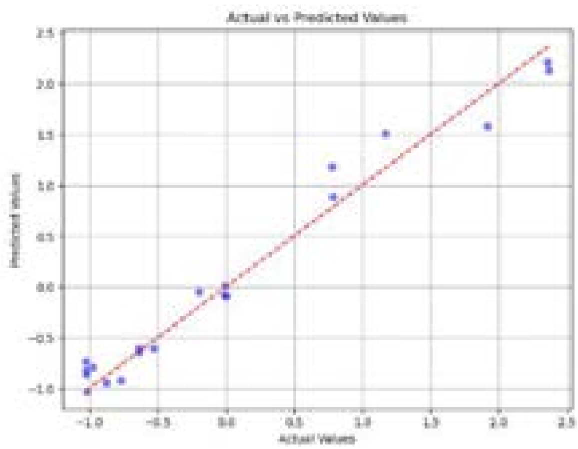


Fig. 4. Actual vs predicted values for non-invasive glucose monitoring

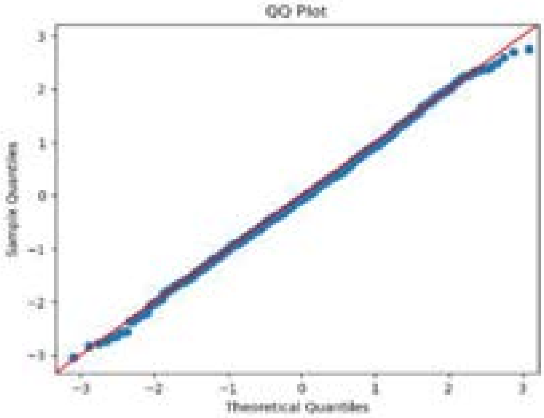


Fig. 5. QQ plot for non-invasive glucose monitoring signal

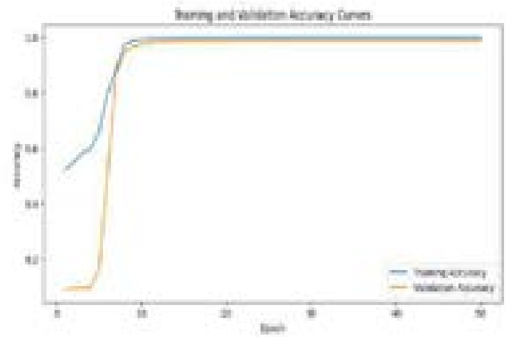


Fig. 6. Training and validation accuracy curves for of ECG signal

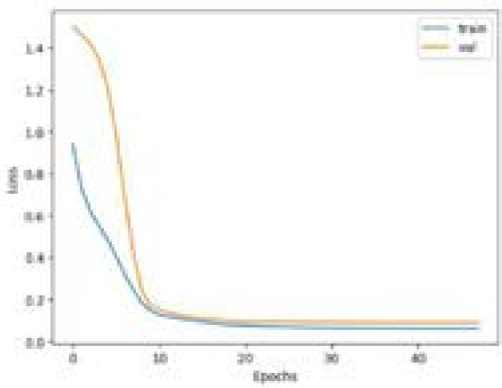


Fig. 7. Epochs vs loss curves for ECG signal

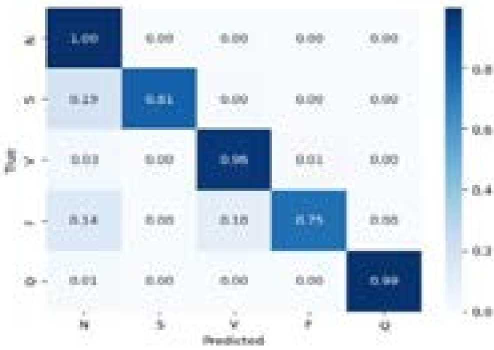


Fig. 8. Confusion matrix for ECG signal

V. CONCLUSION

This study has proposed a novel method based on near-infrared spectroscopy for blood glucose monitoring. We have used neural networks to assess and validate our system. Using NIRF spectroscopy, it was found that the feed-forward neural network. The proposed approach for classifying ECG heartbeats using neural networks is described in this paper. To be more precise, we have trained a deep-convolution neural network for the arrhythmia classification task using 1d convolutions with residual connections. This study has also demonstrated that the representation acquired for this task can serve as a foundation for training precise classifiers for the data classification. Results demonstrate that the suggested method can outperform contemporary techniques in the literature when it comes to predicting both tasks.

REFERENCES

[1] M. W. Steffes and D. B. Sacks, "Measurement of circulating glucose concentrations: the time is now for consistency among methods and types of samples," Clinical Chemistry, vol. 51, no. 9, pp. 1569-1570, 2005.

[2] National Diabetes Information Clearinghouse, "National diabetes statistics," 2011.

[3] American Diabetes Association, "Accuracy of the GlucoWatch G2 Biographer and the continuous glucose monitoring system during hypoglycemia: experience of the Diabetes Research in Children Network," Diabetes Care, vol. 27, no. 3, pp. 722-726, 2004.

[4] M. Sabokdast et al., "Protection by beta-Hydroxybutyric acid against insulin glycation, lipid peroxidation and microglial cell apoptosis," DARU Journal of Pharmaceutical Sciences, vol. 23, no. 1, p. 42, 2015.

[5] S. Coster et al., "Monitoring blood glucose control in diabetes mellitus: a systematic review," British Journal Of Clinical Governance-Bradford, vol. 5, no. 4, pp. 225-227, 2000.

[6] American Diabetes Association, "Diagnosis and classification of diabetes mellitus," Diabetes care, vol. 37, no. Supplement 1, pp. S81-S90, 2014.

[7] K. M. Bratlie et al., "Materials for diabetes therapeutics," Advanced healthcare materials, vol. 1, no. 3, pp. 267-284, 2012.

[8] O. Amir et al., "Continuous noninvasive glucose monitoring technology based on 'occlusion spectroscopy'," Journal of Diabetes Science and Technology, vol. 1, no. 4, pp. 463-469, 2007.

[9] J. Yadav et al., "Prospects and limitations of non-invasive blood glucose monitoring using near-infrared spectroscopy," Biomedical signal processing and control, vol. 18, pp. 214-227, 2015.

[10] N. Spegazzini et al., "Spectroscopic approach for dynamic bioanalyte tracking with minimal concentration information," Scientific reports, vol. 4, p. 7013, 2014.

[11] S. Mousavi and F. Afghah, "Inter- and Intra-Patient ECG Heartbeat Classification for Arrhythmia Detection: A Sequence to Sequence Deep Learning Approach," in ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, United Kingdom, 2019, pp. 1308-1312.

[12] A. Minchole et al., "A novel deep learning approach for classification of ECG arrhythmias," IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 1, pp. 122-131, 2017.

[13] J. Zhang et al., "Deep learning for ECG-based heartbeat classification: A review," IEEE Reviews in Biomedical Engineering, vol. 11, pp. 97-111, 2018, doi: 10.1109/ICASSP.2019.8683140.

[14] A. Y. Hannun et al., "Deep learning for ECG analysis: Benchmarks and insights from PTB-XL," arXiv preprint arXiv:2004.05405, 2020.

[15] Y. LeCun, Y. Bengio, and G. Hinton, "ECG arrhythmia classification using a 2-D convolutional neural network," Communications of the ACM, vol. 60, no. 6, pp. 78-87, 2017.

[16] U. R. Acharya et al., "A novel ECG arrhythmia classification system based on stacked sparse autoencoder and softmax regression," Future Generation Computer Systems, vol. 49, pp. 88-96, 2015.

Advancements and Applications of Artificial Intelligence in Various Sectors: From Internet of Vehicles to 6G and Healthcare

Amit Mittal
Dept of Allied Sciences
Graphic Era Hill University
Bhimtal Campus, Uttarakhand, India
amitforestry26@gmail.com

Harsh Pandey
Dept of CSE
Graphic Era Hill University
Bhimtal Campus, Uttarakhand, India
harshpandey1912@gmail.com

Mohammad Arsh
Dept of CSE
Graphic Era Hill University
Bhimtal Campus, Uttarakhand, India
moarshvlogs@gmail.com

Himanshu Rai Goyal
Dept. of Computer Science & Engg.
Graphic Era Deemed to be University
Dehradun, India
himanshugoyal19@gmail.com

***Abstract*—This research study explores the diverse application of Artificial Intelligence (AI) across various areas, it transforms impact on technological advancement. This survey begins by investing in AI's role in enhancing the efficiency of the Internet of Vehicles through real-time Data collection and edge computing addressing challenges in optimizing edge service. The paper goes into AI's critical contribution to medical diagnosis, disease detection and treatment highlighting applications in breast cancer detection and medical image analysis. The concept of human-AI collaboration is introduced, using the integration of human intelligence with AI to create a hybrid intelligence system. The paper further discusses an ice role in the COVID-19 pandemic from diagnosis to treatment utilizing deep learning techniques and extreme learning machines. This study extends to large-scale, power grid simulation analysis, where AI including deep learning and reinforcement learning significantly improves efficiency and accuracy. The integration of AI into robotics explains the role of teaching, assistant robots and vision-based assessment systems. The concept of smart farming is introduced, showcasing its potential to revolutionize agriculture production. Focusing on maximum PowerPoint tracking techniques for the improvement of efficiency. This study explores applying the artificial immune system of multi-objective optimization with decision-making and introduces the beneficial perturbation network for overcoming continued learning challenges in an Artificial Neural Network (ANN). Satellite communication and the vision of a 6G network are discussed in the integration of AI, IOT and 5G. In the context of the next generation, Smart credit is highlighted in this study and finally concluded by discussing about the challenges in Wireless Sensor Networks (WSN) and the increasing importance of AI.**

Keywords—Artificial Intelligence, Medical Diagnosis, Smart Farming, Communication, Education.

I. INTRODUCTION

This research study analyzes the application of Artificial Intelligence (AI) across various domains, showcasing its transformative potential and problem-solving capabilities. It begins by focusing on the integration of AI with the Internet of Vehicles, emphasizing real-time data collection and edge computing services. The role of AI in medical diagnosis and disease detection is also highlighted, encompassing areas such as breast cancer detection and medical image analysis, showcasing the concept of human-AI collaboration. This hybrid intelligence system leverages both human and AI capabilities. Furthermore, the study analyzes the extensive application of AI in addressing challenges posed by COVID-19, emphasizing the use of deep learning techniques, extreme learning machines, simulation analysis, and the role of artificial intelligence and robotics. Additionally, the survey explores the concept of smart farming, demonstrating how AI and the Internet of Things (IoT) can enhance agricultural production. The application of AI in solar power systems is also discussed, along with the challenges associated with variations in solar output. Moreover, this study examines the application of artificial immune systems in multi-objective optimization and decision-making processes. It introduces a novel approach called beneficial perturbation network to address learning challenges in artificial neural networks, satellite communication, and vision for the 6G network. The integration of AI, IoT, and 5G in the next generation of smart cities is discussed, emphasizing their transformative potential.

Finally, the study concludes by exploring challenges in wireless sensor networks and the increasing importance of artificial intelligence in effectively addressing these challenges.

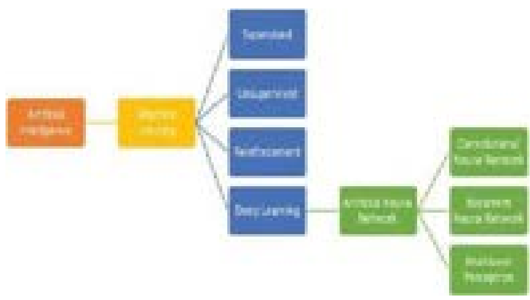


FIGURE 1: Artificial Intelligence. (Mobeen Nazar et.al., 2021)

II. LITERATURE REVIEW

Xiaolong Xu et al.'s research paper [1] ivestigates the role of AI in enhancing the efficiency and responsiveness of IoV (Internet of Vehicles), renowned for its real-time data collection capabilities. Typically, this data is transmitted to resource-intensive cloud platforms for service execution. To meet the demand for real-time services, the paper introduces edge computing, deploying physical resources near roadside units to support vehicular users and enhance edge devices' learning capacity for dynamic resource allocation. The survey also highlights open challenges in optimizing edge service with AI, warranting further research and development efforts.

FIGURE 2: PAPER ORGANIZATION,(MOHAMED ABOUALOLA ET.AL.,2023)

Stephanie Baker et al.'s research paper [2] highlights the critical role of artificial intelligence in improving medical diagnosis and disease detection, acknowledging the complexity of the process and the potential for errors leading to delayed or incorrect treatment. AI, including rule-based systems, neural networks, and deep learning, enhances accuracy and efficiency in applications such as breast cancer detection, the Internet of Medical Things, and medical image analysis. The paper contributes by discussing the evolution of AI techniques and conducting a systematic review of AI-based diagnostic systems, emphasizing the importance of enhancing diagnosis accuracy and efficiency through AI techniques like neural networks and deep learning.

Feifei She et al.'s work [3] introduces the concept of H-AI (Human-Artificial Intelligence), aiming to leverage the combined strengths of human intelligence and artificial intelligence. It addresses current limitations in AI and explores the collaboration between human and machine intelligence. The paper traces the evolution of AI from human-machine interaction to human-machine collaboration, anticipating the emergence of H-AI as a real hybrid intelligence. It outlines a roadmap of eight transitions in H-AI, where machines become integral to human cognition, bridging the gap between theory and practical applications.

Mohammad Jamshidi et al.'s paper [4] emphasizes the crucial role of artificial intelligence in tackling the challenges posed by COVID-19, from diagnosis to treatment. It discusses the potential application of AI, particularly deep learning techniques, in managing the vast amount of medical data generated during the pandemic. The paper presents a detailed workflow, highlighting the importance of human expertise in AI-driven solutions. It describes the application of artificial neural technologies in COVID-19 diagnosis through medical imaging, including advanced technologies like convolutional neural networks. The paper also introduces extreme learning machines for predicting suitable drugs for COVID-19-related complications and highlights the significance of AI in computing the COVID-19 pandemic by handling extensive medical data and improving diagnostic accuracy.

Young Tang et al. [5] presents a research study discussing the integration of artificial intelligence into large-scale power grid simulation analysis. They address the need for advanced technologies to overcome the limitations of conventional power grid analysis methods, highlighting the challenges posed by the vast and complex data involved in power grid simulation. The paper emphasizes the potential benefits of AI-driven solutions, including cost reduction, expert knowledge retention, error reduction, and improved analysis accuracy. Various AI techniques, such as deep learning, reinforcement learning, and transfer learning, are explored in the context of power grid simulation analysis. By addressing issues related to knowledge, modeling, data, discovery, power flow, analysis, and stability control, AI technology can significantly enhance the efficiency and accuracy of power grid analysis, contributing to more reliable and secure power grid infrastructure.

Rafeef Fauzi et al. [6] discuss the evolving field of robotics in AI, emphasizing their significance in education and the development of socially assistive robots. These robots aim to provide continuous support by incorporating emotional, cognitive, and social cues. The integration of AI into robotics is highlighted, particularly through the use of deep learning algorithms to enable robots to perform complex tasks. Advancements in machine learning and AI have greatly impacted computer vision, enabling robots equipped with advanced computer vision capabilities to perform tasks such as face detection, body tracking, and gesture recognition. The paper also discusses the increasing use of robots in education, particularly as teaching assistant robots, assisting teachers with various classroom tasks and student performance evaluation through emotion recognition. The text offers an insightful exploration of the interplay between robots, AI, and vision-based assessment systems, highlighting their growing importance in education and other domains.

Ersin Elbasi et al. [7] introduce the concept of smart farming, which utilizes information and communication technology to enhance agricultural production. They highlight the critical role of the agriculture sector and the potential of AI to transform it, making agriculture smarter and more efficient. The integration of IoT and cloud computing enables monitoring of environmental conditions and crop growth, aiding in selecting the best conditions for different crops. A specific application of AI in agriculture, smart irrigation, is introduced, involving data collection and analysis using devices like Raspberry Pi and automated irrigation processes resulting in cost savings, reduced human efforts, and better soil monitoring. With rapid global population growth and increasing food demand, the integration of IoT, robotics, machine learning, and AI is essential to meet the growing demand for food sustainably and efficiently, making AI-driven smart farming indispensable.

Kah Yung Yap et. al. [8] review that, the use of solar power systems is a cost-effective and efficient source of renewable energy. Solar power is considered a promising renewable energy source due to its cleanliness, abundance and environmental friendliness. It outlines the challenges in solar power systems, particularly regarding variations in active power output due to changes in solar irradiance and temperature. The non-linear nature of photovoltaic cells can make power extraction unsatisfactory, especially under conditions like partial seeding MPPT is a technique that ensures the solar power system operates except at maximum power point. The limitations of MPPT techniques have led to the evolution of the MPPT algorithms, heuristic and meta-heuristic techniques. The text mentioned other matters to improve solar energy efficiency such as integrated soft computing, weather, forecast and adjusting the tilting angle of solar panels to track the sun's direction, which includes reviewing the application and utilization of AINMPPT of solar power systems, providing an overview of current research areas. The people's contributions include summarizing the state of the art and AI-based MPPT evaluating the performance of different algorithms.

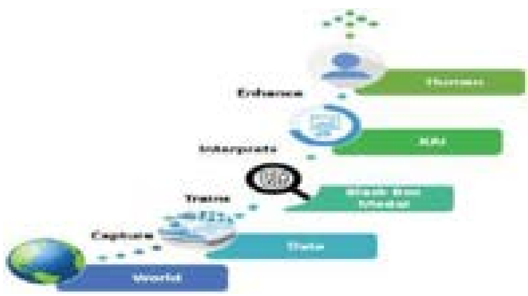


FIGURE 3: EXPLAINABLE ARTIFICIAL INTELLIGENCE [27]

Juan Bao et al. [9] acknowledge the fundamental role of artificial immune systems in multi-objective optimization algorithms and preference-based decision-making. They highlight the biological immune system as an adaptive information processing system with potential applications in artificial immune system theory and artificial intelligence. The paper addresses the computational complexity of multiple objectives optimization algorithms used to explore diverse solutions in the context of multi-objective decision-making problems. It compares these algorithms with genetic algorithms, aiming to find improved solutions. The application of the Biological Immune system involves dividing the population into viable, dominated, and infeasible categories, utilizing hypermutation and transformation techniques. Overall, the application of artificial immune system principles to multi-objective optimization problems emphasizes the role of preference-based design in selecting solutions to Pareto optimal solutions, introducing specific mechanisms to aid in the decision-making process.

Shixian Wen et al. [10] review and discuss a novel approach called the beneficial perturbation Network (BPN) for designing a general adaptive artificial intelligence system. This system addresses the challenges of continual learning and catastrophic forgetting in artificial neural networks. The paper highlights the superior adaptability of the human brain in learning and adapting to diverse environmental situations, tasks, and problems. It introduces methods to adjust multiple objective problems and optimization problems with nonlinear constraints using a biological immune system module. These methods involve dividing populations into viable, non-dominated, and infeasible categories, exploring non-dominated frontiers through cloning and hypermutation, and utilizing transformation techniques to accelerate convergence. The paper also introduces the Angle Best Priority selection mechanism, which simplifies the selection of preferred solutions while reducing the need for extensive reference information. In summary, the paper discusses the application of artificial immune system principles to multiple objective optimization problems, emphasizing the role of preference-based decision-making in selecting solutions from sets of optimal solutions and presenting a high-level overview of research in the field and the principles of artificial immune systems used in optimization and decision-making.

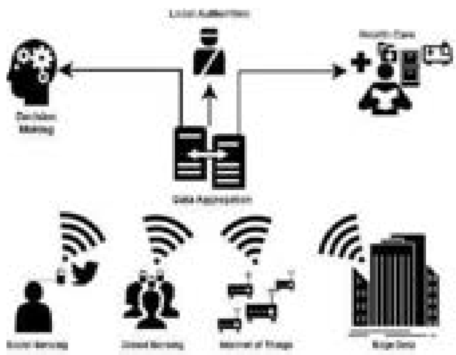


FIGURE 4: TECHNOLOGIES USED IN EMERGENCY SITUATIONS [26].

Xiaofei Wang et al. [11] review a novel approach called the Beneficial Perturbation Network (BPN) for designing general adaptive artificial intelligence systems to address challenges in continual learning and catastrophic forgetting in artificial neural networks. They highlight the superior adaptability of the human brain compared to deep neural networks, which typically learn fixed mappings between inputs and outputs. The primary focus is on the BPN, a biologically plausible method designed to handle dynamic learning situations effectively by computing beneficial perturbations to restore representations to the optimal state. The paper demonstrates that BPN achieves state-of-the-art performance across different datasets and domains, particularly in complex subject recognition tasks. The text introduces the concept of BPN, emphasizing its ability to efficiently adapt to dynamic situations and its scalability, making it a promising solution for continued learning problems. Additionally, it highlights limitations of existing methods and provides an overview of key sections and the concept of BPN.

Fares et al. [12] review Artificial Intelligence (AI) - based techniques for emerging heterogeneous networks, discussing opportunities and challenges in the context of mobile communication. They emphasize the rapid growth of mobile networks and the increasing complexity of mobile infrastructure, which consists of various types of cells with different coverage and working mechanisms. The paper urges researchers to explore solutions that bring intelligence to heterogeneous networks (HetNets) and introduces AI techniques such as machine learning, bio-inspired algorithms, and fuzzy neural networks. The text introduces the concept of applying AI technologies to address complexity and challenges in emerging heterogeneous networks, emphasizing AI's potential to enhance network efficiency and performance.

Khalid B. Letaief et al. [13] review the application of artificial intelligence in satellite communication, outlining the growing demand for satellite communication systems to provide coverage in areas not served by conventional terrestrial networks. They identify three primary use cases for satellite communication systems: service continuity, service ubiquity, and service scalability. The paper highlights the challenges of satellite communication, especially with low Earth orbit (LEO) satellites, and emphasizes the potential of AI techniques, including machine learning and deep learning, to address these challenges. The text provides an overview of the need for AI in satellite communication, identifies key challenges in the field, and introduces the goal of conducting a thorough survey of AI applications to address these challenges comprehensively.

Ebenezer Esenogho et al. [14] discuss the concept of "Edge Artificial Intelligence for 6G" and its various aspects, including vision, enabling technologies, and potential applications. They highlight the ongoing transition from 5G to the development of a 6G network, emphasizing the need to go beyond the capabilities of 5G and explore new usage scenarios for 6G. The paper outlines the potential applications of 6G, including ubiquitous artificial intelligence and the Internet of Everything (IoE), in areas such as sustainable cities, digital twins, and extended reality-health. Additionally, it emphasizes the role of disruptive technologies, particularly AI and edge AI, in designing and optimizing 6G wireless networks. The text provides an overview of the vision and use of 6G, highlighting the transformative potential of these technologies in enhancing reliability and performance while also addressing the challenges and opportunities present. [15-17].

III. DISCUSSION

In the context of the Internet of Vehicles (IoV), this entails leveraging AI algorithms for predictive maintenance, route optimization, and enhanced safety protocols. Similarly, in healthcare, AI of Things (AIoT) requires a comprehensive survey of advancements, challenges, and opportunities to develop smarter diagnostic tools, remote patient monitoring systems, and personalized treatment plans. Hybrid Human-AI Intelligence demands a structured approach, including tutorial development and collaborative frameworks, to enhance decision-making and problem-solving capabilities. The COVID-19 pandemic has highlighted the importance of AI-driven deep learning approaches for rapid diagnosis and treatment planning, necessitating robust frameworks and data-driven insights. The integration of AI in large-scale power grids involves digital simulations and analysis to optimize grid stability, energy efficiency, and renewable energy integration. Robotics utilization with AI in vision-based systems enhances automated assessment tasks in various sectors, from manufacturing to healthcare. AI technology in agriculture focuses on precision farming, crop monitoring, and sustainable practices through data analytics and automation. In the renewable energy sector, AI-based techniques are used to optimize solar power systems for increased efficiency and performance. Looking ahead, AI for 6G networks envisions intelligent network management, ultra-reliable communication, and innovative applications like augmented reality and autonomous vehicles. Addressing challenges in wireless sensor networks through AI techniques involves improving coverage, deployment, and localization for efficient data collection and analysis. Finally, AI's role in enhancing IoT security encompasses advanced threat detection, anomaly detection, and secure communication protocols to safeguard interconnected devices and networks.

IV. IMPLEMENTATION

Implementing optimization in the Internet of Vehicles (IoV) involves leveraging artificial intelligence (AI) algorithms for predictive maintenance, route optimization, and enhanced safety protocols. In healthcare, conducting a survey for advancements, challenges, and opportunities in AI of Things (AIoT) requires developing smarter diagnostic tools, remote patient monitoring systems, and personalized treatment plans. Additionally, creating tutorials and collaborative frameworks is essential for Hybrid Human-AI Intelligence to enhance decision-making and problem-solving capabilities. Addressing COVID-19 with deep learning approaches necessitates robust frameworks and data-driven insights for rapid diagnosis and treatment planning. Analyzing large-scale power grids through digital simulations and AI analysis optimizes grid stability, energy efficiency, and renewable energy integration. Utilizing robotics in vision-based systems enhances automated assessment tasks across industries.

In agriculture, AI technology focuses on precision farming, crop monitoring, and sustainable practices through data analytics and automation. Implementing AI-based techniques to optimize solar power systems increases efficiency and performance. Future AI applications in 6G networks envision intelligent network management, ultra-reliable communication, and innovative applications like augmented reality and autonomous vehicles. Addressing challenges in wireless sensor networks through AI techniques improves coverage, deployment, and localization for efficient data collection and analysis. Finally, enhancing IoT security involves implementing advanced threat detection, anomaly detection, and secure communication protocols to safeguard interconnected devices and networks.

V. CONCLUSION

In conclusion, this study offers a comprehensive analysis of the expansive applications and transformative potential of Artificial Intelligence (AI) across diverse domains, ranging from the Internet of Vehicles (IoV) and real-time data collection to medical diagnosis and disease detection. AI emerges as a formidable tool in augmenting efficiency and accuracy across these domains. The concept of human-AI collaboration introduces a paradigm shift, fostering hybrid intelligence systems that blend human cognition with AI capabilities for enhanced performance. Furthermore, this study highlights the AI's pivotal role in addressing the challenges posed by the COVID-19 pandemic, spanning diagnosis, treatment, and management of medical data. From large-scale power grid simulation to robotics in education, AI's influence continues to expand comprehensively, promising advancements in various sectors. This study also highlights the application of AI in multi-objective optimization and decision-making, as well as the introduction of beneficial perturbation networks for continued learning challenges in neural networks. Additionally, it surveys the integration of AI in emerging heterogeneous network satellite communication and highlights the role of AI and edge AI in shaping the future of the 6G network. The exploration of AI, IoT, and 5G in the next generation Smart Grids underscores their potential while addressing challenges in reliability and performance. Finally, the study concludes by discussing the challenges in wireless sensors and networks and emphasizes the increasing importance of AI techniques in addressing the network coverage, deployment, and localization issues. In essence, this research highlights the pivotal role of AI in shaping the future of various domains, paving way for finding innovative solutions to complex problems.

VI. FUTURE SCOPE

In the future, Artificial Intelligence (AI) will have a significant impact across various domains, including the Internet of Vehicles, healthcare, human-AI collaboration, pandemic response, power grid management, robotics, agriculture, solar power systems, telecommunications, wireless sensor networks, and IoT security. In the Internet of Vehicles, AI will focus on optimizing services, such as predictive maintenance and enhanced safety protocols. Healthcare will witness AI-driven personalized treatment plans and real-time monitoring. The collaboration between humans and AI systems will improve decision-making and problem-solving capabilities. AI's role in pandemic response will involve advanced deep learning approaches for diagnosis and treatment. In power grids, AI will enhance efficiency and resilience against disruptions. Robotics will benefit from AI-driven vision-based systems, improving productivity and safety. Agriculture will see optimized farming practices and increased crop yields through AI. Solar power systems will benefit from AI-based techniques for improved energy efficiency. The transition to 6G networks will be facilitated by AI-driven intelligent network management and innovative applications. AI will address challenges in wireless sensor networks, ensuring efficient data collection and analysis. Lastly, AI will enhance IoT security through advanced threat detection and secure communication protocols. This collective advancement promises to revolutionize industries and drive progress towards a smarter, more connected future.

REFERENCES

1. X. Xu, H. Li, W. Xu, Z. Liu, L. Yao, and F. Dai, "Artificial intelligence for edge service optimization in Internet of Vehicles: A survey," *Tsinghua Science and Technology*, vol. 27, no. 2, pp. 270-287, Apr. 2022, doi: 10.26599/TST.2020.9010025.
2. S. Baker and W. Xiang, "Artificial Intelligence of Things for Smarter Healthcare: A Survey of Advancements, Challenges, and Opportunities," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1261-1293, 2nd quarter 2023, doi: 10.1109/COMST.2023.3256323.
3. F. Shi, F. Zhou, H. Liu, L. Chen, and H. Ning, "Survey and Tutorial on Hybrid Human-Artificial Intelligence," *Tsinghua Science and Technology*, vol. 28, no. 3, pp. 486-499, Jun. 2023, doi: 10.26599/TST.2022.9010022.
4. M. Jamshidi et al., "Artificial Intelligence and COVID-19: Deep Learning Approaches for Diagnosis and Treatment," *IEEE Access*, vol. 8, pp. 109581-109595, 2020, doi: 10.1109/ACCESS.2020.3001973.
5. Y. Tang, Y. Huang, H. Wang, C. Wang, Q. Guo, and W. Yao, "Framework for artificial intelligence analysis in large-scale power grids based on digital simulation," *CSEE Journal of Power and Energy Systems*, vol. 4, no. 4, pp. 459-468, Dec. 2018, doi: 10.17775/CSEEJPES.2018.01010.
6. R. F. N. Alshammari, H. Arshad, A. H. A. Rahman, and O. S. Albahri, "Robotics Utilization in Automatic Vision-Based Assessment Systems From Artificial Intelligence Perspective: A Systematic Review," *IEEE Access*, vol. 10, pp. 77537-77570, 2022, doi: 10.1109/ACCESS.2022.3188264.
7. E. Elbasi et al., "Artificial Intelligence Technology in the Agricultural Sector: A Systematic Literature Review," *IEEE Access*, vol. 11, pp. 171-202, 2023, doi: 10.1109/ACCESS.2022.3232485.
8. K. Y. Yap, C. R. Sarimuthu, and J. M. -Y. Lim, "Artificial Intelligence Based MPPT Techniques for Solar Power System: A review," *Journal of Modern Power Systems and Clean Energy*, vol. 8, no. 6, pp. 1043-1059, Nov. 2020, doi: 10.35833/MPCE.2020.000159.
9. J. Bao, X. Liu, Z. Xiang, and G. Wei, "Multi-Objective Optimization Algorithm and Preference Multi-Objective Decision-Making Based on Artificial Intelligence Biological Immune System," *IEEE Access*, vol. 8, pp. 160221-160230, 2020, doi: 10.1109/ACCESS.2020.3020054.
10. S. Wen, A. Rios, Y. Ge, and L. Itti, "Beneficial Perturbation Network for Designing General Adaptive Artificial Intelligence Systems," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 33, no. 8, pp. 3778-3791, Aug. 2022, doi: 10.1109/TNNLS.2021.3054423.
11. X. Wang, X. Li, and V. C. M. Leung, "Artificial Intelligence-Based Techniques for Emerging Heterogeneous Network: State of the Arts, Opportunities, and Challenges," *IEEE Access*, vol. 3, pp. 1379-1391, 2015, doi: 10.1109/ACCESS.2015.2467174.
12. F. Fourati and M. -S. Alouini, "Artificial intelligence for satellite communication: A review," *Intelligent and Converged Networks*, vol. 2, no. 3, pp. 213-243, Sept. 2021, doi: 10.23919/ICN.2021.0015.
13. K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge Artificial Intelligence for 6G: Vision, Enabling Technologies, and Applications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 1, pp. 5-36, Jan. 2022, doi: 10.1109/JSAC.2021.3126076.
14. E. Esenogho, K. Djouani, and A. M. Kurien, "Integrating Artificial Intelligence Internet of Things and 5G for Next-Generation Smartgrid: A Survey of Trends Challenges and Prospect," *IEEE Access*, vol. 10, pp. 4794-4831, 2022, doi: 10.1109/ACCESS.2022.3140595.
15. Walid Osamy, Ahmad M. Khedr, Ahmed Salim, Amal Ibrahim Al Ali, Ahmad A. El-Sawy, Coverage, Deployment and Localization Challenges in Wireless Sensor Network Based On Artificial Intelligence Techniques. W. Osamy, A. M. Khedr, A. Salim, A. I. A. Ali and A. A. El-Sawy, "Coverage, Deployment and Localization Challenges in Wireless Sensor Networks Based on Artificial Intelligence Techniques: A Review," in IEEE Access, vol. 10, pp. 30232-30257, 2022, doi: 10.1109/ACCESS.2022.3156729.
16. Hui Wu, Haiting Han, Xiao Wang, Shengli Sun, Research on Artificial Intelligence Enhancing Internet of Things Security: A Survey. H. Wu, H. Han, X. Wang and S. Sun, "Research on Artificial Intelligence Enhancing Internet of Things Security: A Survey," in IEEE Access, vol. 8, pp. 153826-153848, 2020, doi: 10.1109/ACCESS.2020.3018170.
17. B. P. Joshi and A. Singh, “Multi-Criteria Decision-Making Approach Based on Moderator Intuitionistic Fuzzy Hybrid Aggregation Operators,” pp. 237–251, 2018, doi: 10.4018/978-1-5225-5709-8.ch011.

18. A. Mittal, L. Dumka, and L. Mohan, "A Comprehensive Review on the Use of Artificial Intelligence in Mental Health Care," in 2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT), Delhi, India, 2023, pp. 1-5, doi: 10.1109/ICCCNT56998.2023.10308255.
19. M. Rautela, B. Tewari, A. Mittal, and S. Kumar, "Unveiling Sentiment Analysis: Exploring Techniques and Navigating Challenges," in 2023 4th International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2023, pp. 01-06, doi: 10.1109/ICOSEC58147.2023.10275819.
20. A. Singh, N. Kumar, B. P. Joshi, and B. K. Singh, “Load frequency control with time delay in restructured environment,” J. Intell. Fuzzy Syst., vol. 35, no. 5, pp. 4945–4951, Jul. 2018, doi: 10.3233/JIFS-169778.
21. B. K. Singh, N. Kumar, A. Singh, and B. P. Joshi, “BBBC based frequency controller for hybrid power system,” J. Intell. Fuzzy Syst., vol. 35, no. 5, pp. 5063–5070, Aug. 2018, doi: 10.3233/JIFS-169790.
22. N. Kumar, A. Singh, B. P. Joshi, and B. K. Singh, “Intelligent Adaptive Critic Scheme Implementation and Investigation Using Policy Iteration Technique: A Case Study of Multi-area Automatic Generation Control (AGC) Problem,” in Lecture Notes in Electrical Engineering, vol. 802, 2022, pp. 159–183. doi: 10.1007/978-981-16-6081-8_9.
23. A. Singh, N. Kumar, B. P. Joshi, and K. S. Vaisla, “AGC using adaptive optimal control approach in restructured power system,” J. Intell. Fuzzy Syst., vol. 35, no. 5, pp. 4953–4962, Jul. 2018, doi: 10.3233/JIFS-169779.
24. S. Sharma, A. Mittal, M. Bansal, B. P. Joshi, and A. Rayal, “Forecasting of Carbon Emissions in India Using (ARIMA) Time Series Predicting Approach,” Lect. Notes Electr. Eng., vol. 1086, pp. 799–811, 2024, doi: 10.1007/978-981-99-6749-0_53.
25. S. Sharma, A. Mittal, B. Prasad Joshi, and P. Pandey, “Predictive Model for Dissolved Oxygen in water body (Nainital Lake) of Nainital district by Multivariate Regression Analysis,” in 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), May 2023, pp. 376–379. doi: 10.1109/icacite57410.2023.10183174.
26. M. Aboualola, K. Abualsaud, T. Khattab, N. Zorba and H. S. Hassanein, "Edge Technologies for Disaster Management: A Survey of Social Media and Artificial Intelligence Integration," in IEEE Access, vol. 11, pp. 73782-73802, 2023, doi: 10.1109/ACCESS.2023.3293035.
27. M. Nazar, M. M. Alam, E. Yafi and M. M. Su’ud, "A Systematic Review of Human–Computer Interaction and Explainable Artificial Intelligence in Healthcare With Artificial Intelligence Techniques," in IEEE Access, vol. 9, pp. 153316-153348, 2021, doi: 10.1109/ACCESS.2021.3127881.

Two-Wheeler Helmet Detection System Using Object Recognition

Prof. Vikas Nandeshwar
Department of Engineering, Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
vikas.nandeshwar1@vit.edu

Harshad Sheelwant
Department of Engineering, Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
harshad.sheelwant23@vit.edu

Nishit Shelar
Department of Engineering, Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
nishit.shelar23@vit.edu

Bhavesh Shelke
Department of Engineering, Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
bhavesh.shelke23@vit.edu

Shashank Patil
Department of Engineering, Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
shashank.patil23@vit.edu

Abdul Sheikh
Department of Engineering, Sciences and Humanities, Vishwakarma Institute of Technology, Pune, India
abdul.sheikh23@vit.edu

***Abstract* — A “helmet” is a protective cover worn while riding a two-wheeler, which is a safety measure for the rider and for the pillion. However, due to the negligence of people, helmets are not worn by most riders. According to the latest report of the National Crime Records Bureau (NCRB), as many as 1,55,622 people lost their lives due to road accidents in 2021, out of which around 70,000 people lost their lives in two-wheeler accidents.**

Proposed study is the perfect solution for this problem. We have made a smart machine learning-based system integrated in a two-wheeler that makes wearing a helmet mandatory as it monitors every second of the ride. The moment the rider removes the helmet, the vehicle is supposed to turn off and make the rider wear a helmet. There are some existing studies that cover different parts of helmet detection techniques but almost each of the studies are usable and limited to that helmet however this study focuses on how the system can be integrated into the two-wheeler itself to generalize the detection of helmets. The perfect implementation of this study is used to reduce accident rate.

Keywords— Artificial-intelligence; Automobile; Helmet; Image-processing;

I. Introduction

A survey revealed that in countries like Thailand, Vietnam, Malaysia and India more than 50% of the households own a two-wheeler, i.e. a motorcycle or a scooter. A two-wheeler on a road is more exposed to a potential accident than other vehicles. In India as of 2022, two-wheeler fatalities make up 44% of total road accidents. Two-wheeler riders by law are required to wear a helmet while riding, however most of the riders neglect this law and that might result in amplified damage through an accident.

To avoid this, governments have implemented fines and even human-monitored systems to charge some breaking the law, however, this being manual there are chances of error.

Surveillance cameras on roads have been evolving every day in India. The cameras use an object detection system that identifies every vehicle on road and its registration number to cite fines and also control traffic. To safely and smoothly execute this system of controlling the traffic flow, real-time monitoring and machine learning based on pre-existing data is required. By implementing object detection and automation in traffic regulation, the efficiency and flow of traffic is seen to be more controlled and coordinated than earlier when manual techniques were used to do the same. The proposed work aims for better accuracy and efficient processing. There is also some part of manual execution in such a technology which again opens us to errors.

To minimize the error a smart and highly accurate object-detection program is used, which detects a helmet through a live feed. The device will monitor the live feed each second frame by frame and analyze whether a helmet is worn by the rider or not. The program is set in such a way that at any point the rider is not wearing a helmet the two-wheeler turns off the ignition. This being a more accurate method and a more efficient method to eliminate road accident fatalities can be a great feature to implement in bikes and scooters.

II. Literature Review

[1] **S. Cui et al.,** "An Effective Motorcycle Helmet Object Detection Framework for Intelligent Traffic Safety,", in this study a standard computer vision model “Detectron2” which is a Facebook AI research library that provides advanced object detection and segmentation algorithms. They carried out multiple test runs on different data sets and have built a study using these tests on just the detection and methods that can improve detection.

[2] **S. Anjum et al,** "Artificial Intelligence-based Safety Helmet Recognition on Embedded Devices to Enhance Safety Monitoring Process," this study focuses on safety helmets at construction sites, not completely related to our methodology but uses a similar detection idea. They have used tensorflowlite libraries to make the detections effective and lighter on the devices. They have built an IOT system where if a construction worker is not wearing an appropriate helmet, the supervisor will receive a notification.

[3] **M. Uniyal et al,** "IOT based Smart Helmet System with Data Log System," this study uses a raspberry pi 3 board connected GPS module and RF receivers that work on IOT system which conveys a message when a helmet is not worn on the head. This paper was important to us because the idea of turning the ignition on and off comes through this study.

[4] **K. C. D. Raj et al,** "Helmet violation processing using deep learning", they used a CNN to classify a rider into “helmet” and “no helmet” with an additional license plate recognition system using a HOG descriptor with an accuracy of 90%.

[5] **B. R. Chandra et al,** "Internet of Things (IoT) based Digital Helmet Design and Deployment", they used a CNN to classify a rider into “helmet” and “no helmet” with an additional license plate recognition system using a HOG descriptor with an accuracy of 90%.

[6] **Nandu R et al,** "Smart Helmet for Two-Wheelers", this study is a “Smart Helmet” that when not worn denies ignition and when worn allows ignition. They have used a proximity sensor and a LED and placed them on the opposite sides inside a Helmet, that is when the glowing LED sheds light on the proximity sensor the circuit is completed sending signals to the further circuit that the helmet is not worn as nothing is blocking the light.

[7] **Keesari et al,** "Smart helmet for safe driving”, this study is not only a helmet identification model but works as an accident detection and an alcohol detection system as well. They’ve used multiple sensors to achieve the results, for the helmet a RF sensor is utilized.

III. Methodology

In this study, a custom dataset was constructed by incorporating annotations, “helmet” and “no_helmet” through the LabelImg tool.

In the early stages of the model, the study utilized a dataset available on Kaggle, a free for all dataset, which contained of images of all kinds of helmets, pre-annotated. The dataset had construction site helmets, helmets used by cyclists and other kinds of helmets which cluttered the model and the results weren’t as expected. While finding more datasets we landed on a custom dataset that consisted of around 60-80 images of the authors with helmet and also some images extracted from another free for all dataset of images with faces of people to make the model learn the difference between “helmet” and “no_helmet”.

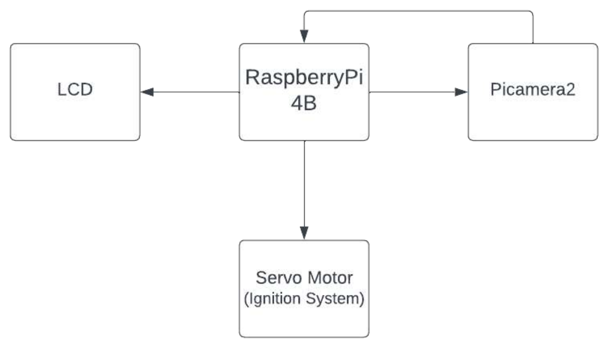


Fig. 1. Block Diagram

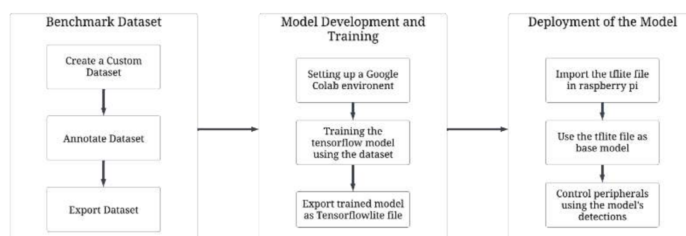


Fig. 2. Process Flow of the proposed method

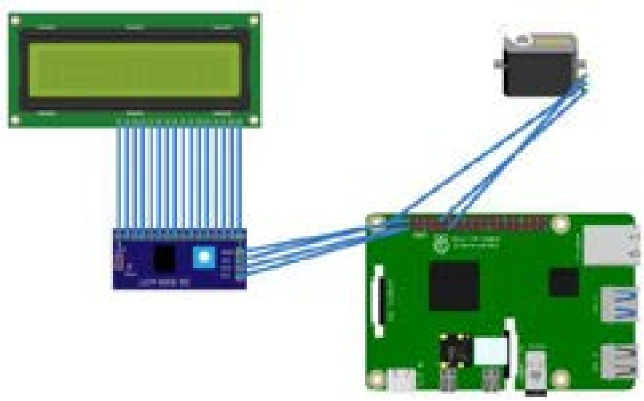


Fig. 3. Circuit

The circuit as shown in Fig. 3 includes a Raspberry Pi Model 4B, a 16x2 LCD, servo-motor as a prototype for the ignition of the two wheeler and a Picamera 2 as a camera module. However this being a prototype of the main idea, better components with more accuracy and better results can be used such as 'ESP32', which is a wireless camera module more apt to the study as we need a camera module that can be easily placed at a position on the dashboard of the two-wheeler from where the rider is clearly visible at every second of the ride.

To make use of a successful model we need a strong machine learning technology and TensorFlow was the best option because it utilizes CNN (Convolutional Neural Network) technologies to develop a model that can be used in Google Colab to deploy in a working environment. As shown in fig. 2, the annotated dataset was utilized within the Google Colab environment to develop a tflite (TensorflowLite) model as raspberry pi is more efficient with Tflite files than Tensorflow files. The resulting tflite file contains the essential data required for accurate helmet detection. The tflite file being associated with custom images is more accurate and appropriate with the detections.

Following model creation, we exported the tflite file and integrated it into Raspberry-Pi 4B. Wherein we utilized OpenCV-python library for developing a system where we used an if-else statement to operate the ignition system which here is a servo motor and also a display. This methodology ensures a comprehensive and effective approach to enhancing safety through intelligent helmet detection systems.

IV. RESULTS AND DISCUSSIONS

The prototype of the main study gives us a brief idea of how this Helmet Detection System can be implemented in an actual two-wheeler, as it benefits the rider and also other traffic on the road. The proposed system reaches85-90% accurate rate whether the rider is wearing a helmet or not. The whole system works in a way that if the rider is wearing a helmet the ignition i.e. the servo motor will operate and an LCD will display "Ride Safely" and if the rider is not wearing a helmet the motor will not work with the LCD saying "Please Wear A Helmet". This can be situated in the dashboard i.e. around the speedometer of a two-wheeler.

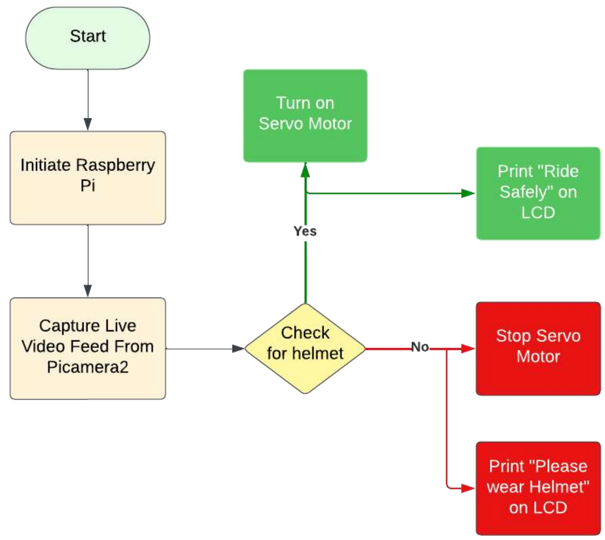


Fig. 4. Flowchart

Fig. 5. Printing in the Terminal of the IDE on Raspberry Pi OS "No helmet detected. Turning off motor."

Fig. 6. Printing in the Terminal of the IDE on Raspberry Pi OS "Helmet detected. Turning on motor."

Fig. 5 and Fig. 6 shows a model that detects whether a helmet is worn or not, and annotates the same with the accuracy from 0 to 1 along with the label “Helmet” and “No Helmet”.

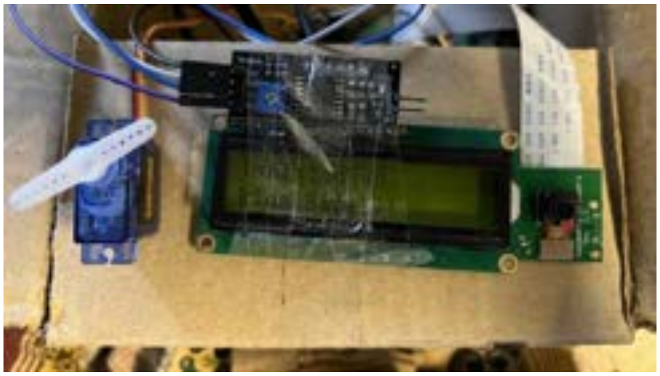


Fig. 7. LCD when helmet is worn

Fig. 8. LCD when helmet is not worn

In Fig. 6 and Fig. 7, the final result is visible on the LCD, that is, if the rider is wearing a helmet the system successfully detects the helmet and while turning the servo motor on it prints “Ignition On” on line 1 and “Ride Safely!” on line 2 while when the rider is not wearing a helmet the LCD prints “Please Wear” on line 1 and “A Helmet” on line 2.

From testing the model in low-light conditions and in camera angles that would simulate the camera being placed in the dashboard of a two-wheeler, it was found that there is scope for improvement in low-light conditions as the dataset used for training the model mostly contained images taken in good lighting. The results can be improved to higher levels if a camera of better clarity can be used in place of the existing camera.

The following table shows the accuracy measurements obtained by testing the model in various lighting conditions and camera angles.

TABLE I

Accuracy Measurements

Accuracy Measures	**Values**
Precision	93%
Average Accuracy Score (from 0 to 1)	0.93

[a] These values are calculated by testing of the model in various conditions.

V. Conclusion

The study ensures the critical assessment of the methodology proposed all along and contains the following points:

1) Road safety and awareness about helmet compulsion while commutating via a two-wheeler.
2) Utilizing a non-biased system of recognition of helmet violations.
3) Successful implementation of a small size prototype of the proposed methodology, using similar electronic components to an actual vehicle.
4) Rigorous testing and improvement of the results of helmet-detection python module.
5) Using the methodology in an actual vehicle may it be an electrical or a fuel based two-wheeler.

VI. Future Scope

In further studies in the same enhancing the smart helmet in the future by adding innovative safety features like alarms, enhancing its helmet detection capabilities, and adapting it to a variety of helmet types seems like an easier job using better technologies. The research work may utilize cloud computing, interface it with additional safety sensors, and link it to the internet for remote control. Collaborate with two-wheeler manufacturers, obtain certifications, and solicit user input. By taking these steps, the research work intend to increase the usefulness and appeal of proposed smart helmet to global audience.

VII. Acknowledgment

We would like to express our deepest appreciation to Vishwakarma Institute of Technology for their intangible support to development of the study. We are also extremely grateful to Prof. (Dr.) C. M. Mahajan for his guidance, advice and patience towards us.

We would also like to thank each and every member connected to the study in and out of our group. Every person’s contribution mattered in implementing and documenting the journey.

References

[1] S. Cui et al., "An Effective Motorcycle Helmet Object Detection Framework for Intelligent Traffic Safety," 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Vancouver, BC, Canada, 2023, pp. 5470-5476, doi: 10.1109/CVPRW59228.2023.00578.

[2] S. Anjum, S. F. A. Zaidi, R. Khalid and C. Park, "Artificial Intelligence-based Safety Helmet Recognition on Embedded Devices to Enhance Safety Monitoring Process," 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 2022, pp. 1-4, doi: 10.1109/IICAIET55139.2022.9936839.

[3] M. Uniyal, H. Rawat, M. Srivastava and V. K. Srivastava, "IOT based Smart Helmet System with Data Log System," 2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2018, pp. 28-31, doi: 10.1109/ICACCCN.2018.8748790.
[4] K. C. D. Raj, A. Chairat, V. Timtong, M. N. Dailey and M. Ekpanyapong, "Helmet violation processing using deep learning," 2018 International Workshop on Advanced Image Technology (IWAIT), Chiang Mai, Thailand, 2018, pp. 1-4, doi: 10.1109/IWAIT.2018.8369734.
[5] B. R. Chandra, M. P. Akshaya, B. Sravani, Y. S. Ashritha and B. K. Pavithra, "Internet of Things (IoT) based Digital Helmet Design and Deployment," 2023 7th International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2023, pp. 397-404, doi: 10.1109/ICOEI56765.2023.10125847.
[6] Nandu R, Singh K (2014) Smart Helmet for Two-Wheelers. Adv Automob Eng 3:110.
[7] Keesari, Shravya & Mandapati, Yamini & Keerthi, Donuru & Harika, Kothapu & Senapati, Ranjan. (2019). Smart helmet for safe driving. E3S Web of Conferences. 87. 01023. 10.1051/e3sconf/20198701023.

FACIAL IDENTIFICATION AND REMOTE DOOR LOCKING SYSTEM

Sneha K
Electronics and Communication Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
snehak@vrsiddhartha.ac.in

Sai Vinodh Koritala
Electronics and Communication Engineering
Velagapudi Ramakrishna Siddhartha Engineering college
Vijayawada, India
koritalasaivinodh@gmail.com

Manikya Gowtham Nandula
Electronics and Communication Engineering
Velagapudi Ramakrishna Siddhartha Engineering college
Vijayawada, India
19manikyagowtham@gmail.com

Jaswanth Kumar Kethavathu
Electronics and Communication Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
kethavathujaswanthkumar@gmail.com

Ganesh Potluri
Electronics and Communication Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
potluriganesh3@gmail.com

Abstract: **Improvements in present technology have significantly simplified human life, thus resulting in a widespread shift from traditional lock methods to smart door systems. Our project utilizes the capabilities of both Pi camera and Raspberry Pi to enhance both convenience and security. This innovative solution provides the feature of recognizing the individuals approaching the entrance. If the visitor's identity is known, then the door automatically opens. In cases if the visitor's identity is unknown, immediate email notifications containing captured images by pi camera are dispatched to the owner for verification. Utilizing remote access control features, homeowners can manage entry permissions from any location. The integration of a digital lock improves the security at the entrance. With introduction of a buzzer, the owner gets alerted when both familiar and unfamiliar persons arrived. This research work provides an integrated and intelligent approach to door access management, resulting in improved levels of home protection.**

Index Terms—Face recognition, Raspberry Pi 3b, digital lock, double Rely, Picamera.

I. INTRODUCTION

In daily lives, standard door locks have been a common thing for locking our doors or using for protection of valuable items. Even in some important working places standard locks are used in locking the doors. But the time has come to understand the distinctions between smart and traditional locks. Despite having a long history these conventional locks spanning 4,000 years, many problems such as jammed mechanisms and misplaced keys have plagued them, leading many persons to think these locks as less secure. Traditional locks are no longer giving protection for our items. Smart locks, integrated into smart home systems, offer automation and remote control capabilities, which makes smart locks a step ahead of traditional locks. With the addition of convenience of remote control and integration with smart home setups, these smart locks improve home security efficiently. This perception shift has led an increasing preference for smart locks over traditional locks. While these advanced locks have been around for only a decade, they too have some technological limitations such as fingerprint access issues and forgetting the code/numbers/pin which are unavoidable. To address these concerns, in present technology face detection systems are used as a secure door lock mechanism. Only authorized persons, generally like the family head or this family members, can grant access to unapproved persons or visiting relatives. This technological evolution shows a significant step forward in home security, offering both convenience and enhanced protection.

II. LITERATURE SURVEY

The reaserch study "facial identification and remote door locking system " involves a comprehensive review of existing technology and their related research and literature related to various technologies like smart door systems, Raspberry Pi integration, and home security technologies. Previous technological studies have explored different aspects of smart door locks, which includes their design, structure, implementation, working and effectiveness in improving home security. In addition to this, research on various Raspberry Pi applications in home automation areas and also in surveillance area provides valuable insights into the technical features and certain limitations of using this technology for door locking purposes. Further, studies on remote access control system and procedure for notification contribute to understanding the best practices for implementing features such as remote access system and email notifications in smart door locking systems. By combining these findings from various sources, the literature study aims to identify important trends, unavoidable challenges, and immense opportunities in development of the proposed smart door locking system. Through these findings and researches of existing literature, the project aims to build upon prior research and contribute to the improved smart home security solutions using the present Raspberry Pi technology.

III. PROPOSED SYSTEM

In theexisting model, there is no option of remote access. But in this model, there is option of remote access where homeowner has option to control the door even when he is far from home. Also both face recognition and remote access has been in this model which is a missing in existing model like password based door locking system, finger print authentication system etc. In present existing models like finger print door locking system, we need a human contact which is a danger sign in pandemics like covid.

The proposed system for the project involves utilizing a Raspberry Pi along with various components such as a Pi Cam, double relay, digital lock, and buzzer. The Pi Cam is employed to capture images of unknown or known individuals. When a known person arrives, the Pi Cam recognizes them by matching their captured face with faces in a predefined dataset. Subsequently, the Raspberry Pi triggers the digital lock to open and activates the buzzer. Conversely, if an unknown person is detected, the Pi Cam sends the captured image to the owner's email. The owner can then decide whether the unknown individual should be allowed entry into the premises. If approved, the owner can remotely open the door by clicking 'on' via email. If not approved, the door remains closed. This system facilitates remote access through email protocols such as SMTP and IMAP, providing the owner with control over access to their property.

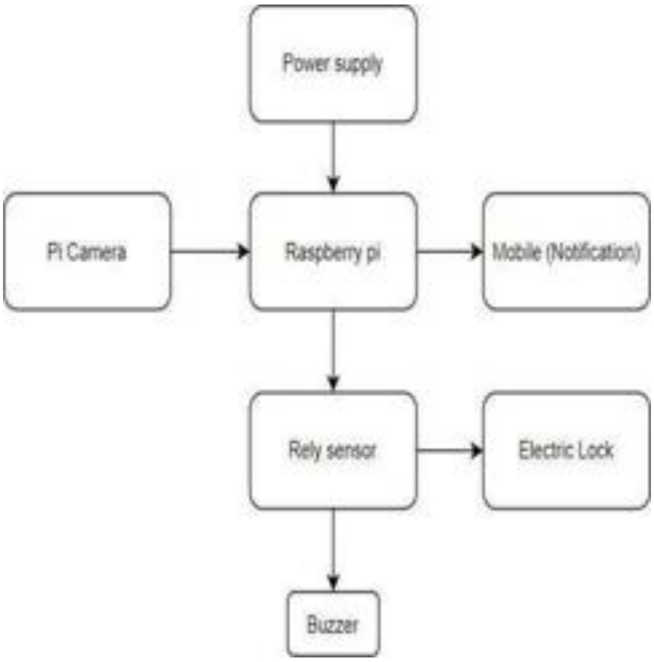


Fig. 1. Block Diagram of Proposed System

IV. MODEL IMPLEMENTATION

A. Hardware Implementation

1) Raspberry pi 3b : The Raspberry Pi 3b serves as the central Computing unit for our project. It runs the face recognition algorithm and controls the door lock mechanism based on the results stored during the training of the model. Raspberry Pi boards come with General Purpose Input Output (GPIO) pins, which allow you to easily interface with external hardware components such as sensors, actuators, and cameras. This makes it convenient to connect peripherals required for face recognition and door access system, such as a camera module for capturing facial images and controlling the door lock mechanism. It contains 40 pins which are like VCC, ground etc.

2) Pi Camera: A camera module compatible with the Raspberry Pi is essential for capturing images of faces. This could be a dedicated Raspberry Pi camera module.

3) Face Recognition Software: Software for face recognition needs to be installed on the Raspberry Pi. Popular choices include OpenCV,face_recognition with Python bindings, which provide a range of functions for face detection and recognition.

4) Door Lock Mechanism: we have used a magnetic lock in our project. So to control the lock we have used a double relay module, one relay module is used to open the lock and the second relay module is used to close the lock. The hole control of lock is in the gpio pin based on the output the pin changes according to the code running in raspberry pi.

B. Software Implementation

To implement a face recognition project with a Raspberry Pi controlling a magnetic door lock, start by setting up the Raspberry pi with installing OS and libraries needed for face recognition.

1. Face Recognition library for face recognition
2. imutils, pickle, cv2 libraries for data storing and image processing
3. Email library for sending images and remote access
4. Datetime, time libraries
5. Raspberry Pi GPIO libraries: RPi.GPIO

Initially, capture the images of those people who we want to give the access to home, so that whenever they come the door lock open automatically. After capturing the images, we need to store them according to their names in a dataset folder. After, this process we need to train the model. For this purpose, we have used a face recognition library. It is a pre-trained deep learning model (typically a Convolution Neural Network or CNN) to extract features from facial images. These features represent the unique characteristics of each face in the dataset. The face_recognition library uses a pre-trained model based on the dlib library to perform this task.
After extracting features, the library computes a numerical encoding (vector) for each face in the dataset. This encoding is essentially a compact representation of the facial features extracted by the deep learning model. These encodings are then stored for later comparison during recognition.
When a new face is presented for recognition, the system follows a similar process to extract features from the input image.
The library then calculates the encoding of this new face.

It compares the encoding of the new face with the encodings of faces stored in the dataset. This is usually done by measuring the similarity between the encodings, often using techniques like Euclidean distance or cosine similarity.
If the distance or similarity between the new face's encoding and any of the stored encodings is below a certain threshold (which you can define), the system recognizes the person associated with that encoding.
Based on the similarity scores obtained from comparing the encodings, the system decides whether the input face matches any of the faces in the dataset.

C. Protocols

IMAP (Internet Message Access Protocol) is a key component for email communication. It has transformed the way how users interact with and handle their electronics message. IMAP has been introduced as a substitute for more conventional Post Office Protocol version 3 (POP3). It completely changes the way messages are retrieved and synchronized between a user's email client and the mail server. With IMAP, user may read and handle their emails directly on the mail server which is different scenario when compared to POP3. The role of this IMAP is to allow homeowner to access their email inbox, view received messages and take action on them like unlocking the door through remote access directly from homeowner's email account

SMTP (Simple Mail Transfer Protocol) is the most important part of email delivery. It provides foundational path for error free communication and can be termed as text based protocol. In this model the SMTP is used to send images from sender devices to owner's personal mail. SMTP is also responsible for delivering emails containing images of unknown persons captured by the door lock system to homeowner's email inbox.

V. Working Of The Proposed System

In this implementation of project firstly the model need to train the model to make the person known so that the door opens automatically. To make this happen, the model need to take the images of the known person into the dataset and save them with a particular name and then train the model with his images. The model is trained using "face_recognition package" which contain different modules. firstly, In the given image (frame) it detect weather it contain a face or not by face location function if it contain face it return the face locations in the give image else it don't return any locations and then by using the face locations the face of that particular person is encoded with face_encodings function and the data is stored in serialized manner using pickles by using this data we can say that the person is known or unknown to us. The part of capturing of images and training the module is continuous process when the home owner wants to add new person as known they need to add them manually every time.

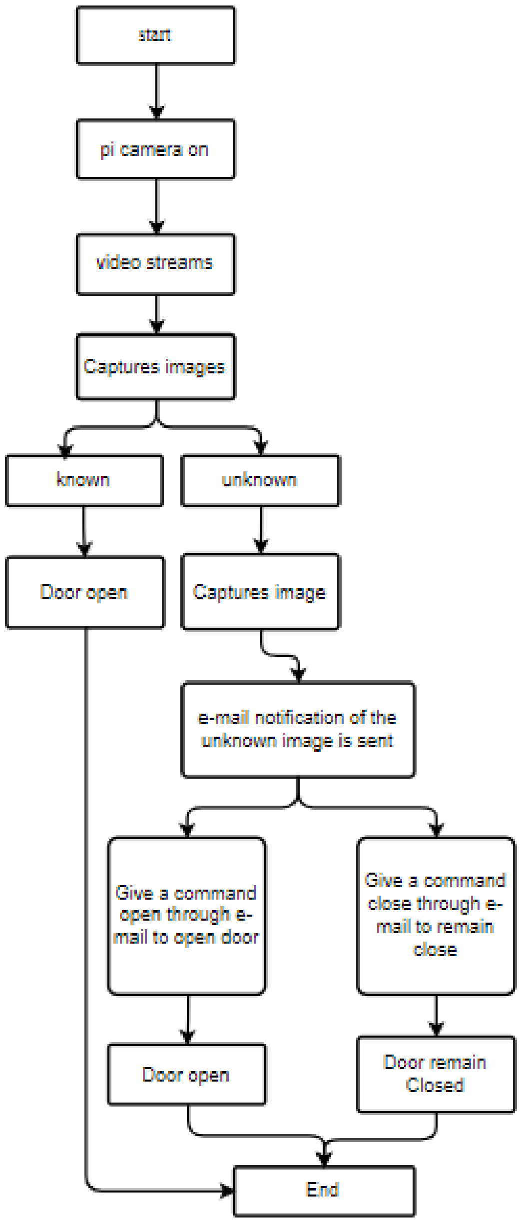


Fig. 2. Flow-Chart of working of the system

In this section the model will run continuously the camera will be on every time when the frame contains face of a person then it finds the locations of the face and encode the face and then compare with the list of encoded values, which are stored previously during training. if match is found a particular port pin of GPIO will be high to make the door lock open else it will send an email to the home owner with the image of the unknown person. Then home owner can take a decision according to the need. Based on the response from home owner the door lock will open/close. We have set a certain time that lock will open for that particular time after that the lock will be closed automatically.

When both known and unknown in the single frame the face which is in the process first according to his face the function of door will happen. If a known person face is first detected in the process the door lock opens else if unknown person face is first detected then email notification is send to home owner the model runs according to his responses. Simultaneously a buzzer will be ring when door lock opens that indicates the people inside the house will know that some entered into the house at that time.so, they can pay their attention towards him/her.

VI. Results & Discussion

The above is the working model of our project. It contains components like raspberry pi, Pi Cam, buzzer, digital lock, double relay. The picture of working model is shown in figure 3 below.

Fig. 3. Working model of the device

Utilizing facial recognition the system identifies known individuals approaching the door. Upon recognition, the door lock mechanism is triggered to unlock automatically, granting access to the known person. This feature enhances convenience for residents and trusted individuals, eliminating the need for physical keys or manual unlocking. The picture of the process above is shown in Figure 4 below.

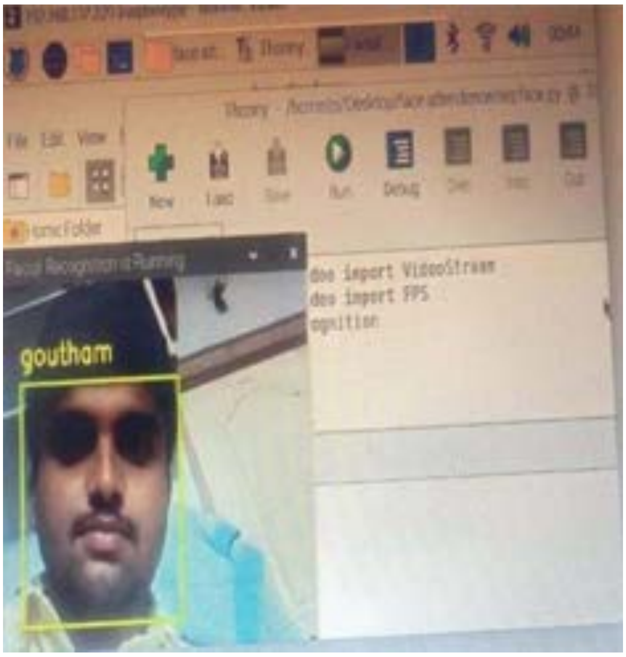


Fig. 4. When known person arrives

As discussed above, whenever a person is arrived, the pi cam detects the name and compare with the images stored in dataset folder. If it is a Unknown face, then it displays as unknown as shown in figure along with the square box. The picture of process is shown in figure 5 below.

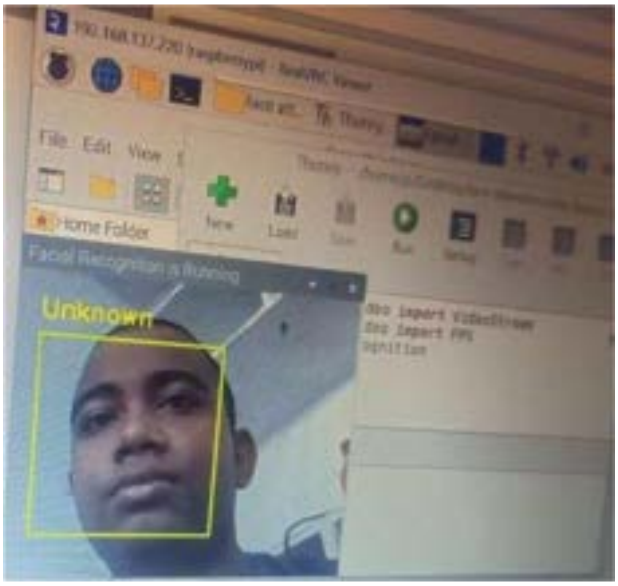


Fig. 5. When Unknown person arrives

In the event of an unrecognized individual approaching the door, the system triggers an alert mechanism. A notification is sent via email to the homeowner, informing them of the presence of an unknown person at the door. This serves as a security measure, allowing the homeowner to be aware of potential unauthorized access attempts. The picture of the process is shown in Figure 6 below.

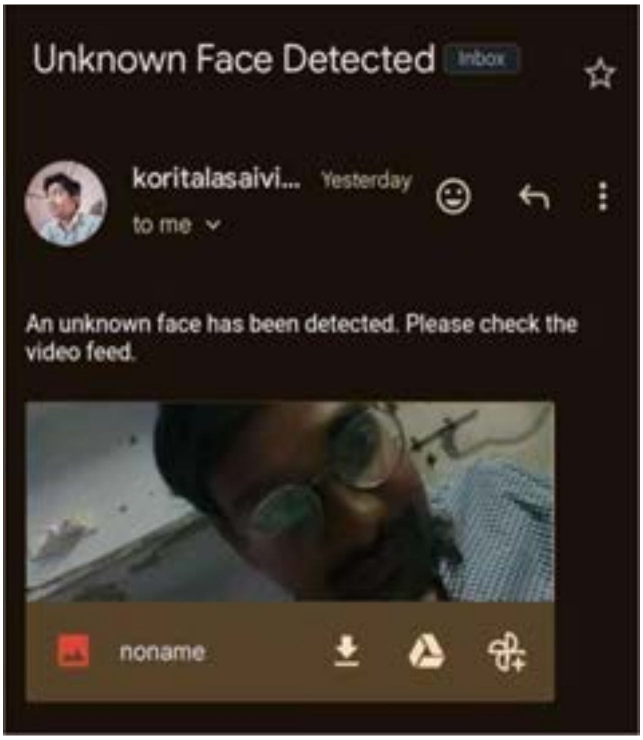


Fig. 6. Mail Notification when Unknown person arrives

The homeowner has the capability to remotely manage access permissions for unknown individuals. Through a user-friendly interface, accessible via smartphone or computer, the homeowner can grant or deny access to the door through email. By sending a command, such as "On" or "Off," the homeowner can control the door lock remotely, providing flexibility and control over access to the property. The picture of process is shown in figure 7 and 8 below.

Fig. 7. Remote Access operation of door opening

Fig. 8. Remote Access operation of door closing

VII. CONCLUSION

In summary, proposed model has effectively combined various technologies to improve home security and offer secure access management. This system uses a Raspberry Pi with a Pi camera module to accurately identify people who are about to approach the door. The facial recognition system is designed to differentiate between individuals who are known to us and those who are not. Identified faces can be allowed into the site with ease, while for the case of unfamiliar faces, the owner will immediately receive an email message along with that person's image. Homeowners can ensure an additional feature of protection and control by granting or refusing access remotely with the help of Email application. The inclusion of a digital lock strengthens the system's security features. A buzzer installed allows homeowners to be instantly alerted to identify the visitors. Through the use of various technologies, this project successfully combines innovation, convenience, and security for benefit of homeowners.

VIII. FUTURE SCOPE

The performance of the model can be improved by training the dataset model with high-quality images with good lighting, varied facial expressions and different poses. It can also be improved by considering more advanced deep learning architectures for face recognition, such as convolutional neural networks (CNNs). Pre-trained models like VGG, ResNet, or EfficientNet, which might offer better performance than simpler models. Higher Versions of raspberry pi can also be included so increase the accuracy. Also other high quality camera can be used like USB camera instead of Pi camera for improving accuracy.

There are numerous applications for Facial identification and remote door locking system in the future. Proposed system can develop into a multi-layered device for home automation purpose and enhanced security, going beyond simple features like facial recognition and remote control. For safer identification, consider incorporating biometrics like fingerprint readers. Authorized persons may be able to unlock doors hands-free with voice recognition feature, and environmental sensors have the efficiency to adjust temperature and lighting automatically upon detecting someone entry. The system could be trained to identify suspicious activities like loitering or tampering, which use as real-time alerts and sirens as deterrent measures. For home delivery services, sending one-time codes via SMS to grant temporary access is one of the best practices. Smart home platforms could allow doors to be connected to other device automatically illuminating light or granting access through side doors to obtain authorization. Proposed system can also be developed to warns the possible danger like gas and smoke leaks in front of it. With the proposed work, convenience and security are combined to create a genuinely intelligent and secure home. With this system, own area is safeguarded and hence will improve our lives inside our own four walls.

References

[1] Swathi. V, Steven Fernandes," Raspberry Pi Based Human Face Detection", published in International of advanced research in computer and communication engineering Vol. 4, Issue 9, September 2015.

[2] Y.-C. Yu, "A practical digital door lock for smart home", In: IEEE International Conference on Consumer Electronics (ICCE), 2018.

[3] J. Jeong, M. Y. Chung and H. Choo, "Integrated OTP-based user authentication scheme using smart cards in home networks," In Proceedings of the 41st IEEE Annual Hawaii International Conference on System Sciences, pp. 294-294,2008.

[4] Y. W. Bai, L. S. Shen and Z. H. Li, "Design and implementation of an embedded home surveillance system by use of multiple ultrasonic sensors," IEEE Transactions on Consumer Electronics, vol. 56, no. 1, pp. 119-124, 2010.

[5] Y. Zhao and Z. Ye, "A low-cost GSM/GPRS based wireless home security system," IEEE Transactions on Consumer Electronics, vol. 54, no. 2, pp. 567-572, 2008.

[6] R. Chuimurkar and V. Bagdi, “Smart Surveillance Security and Monitoring System Using Raspberry PI and PIR Sensor,” International Journal of Scientific Engineering and Applied Science, vol. 2, no. 1, pp. 159–165, 2016.

[7] S. N. Basha, S. A. K. Jilani and S. Arun, "An Intelligent Door System using Raspberry Pi and Amazon Web Services IoT," International Journal of Engineering Trends and Technology, vol. 33, no. 2, pp. 84 89, 2016.

[8] C. Twumasi, K. A. Dotche, W. Banuenumah and F. Sekyere, "Energy saving system using a PIR sensor for classroom monitoring," IEEE PES Power Africa, pp. 347-351,2017.

[9] J. Bangali and A. Shaligram, "Design and Implementation of Security Systems for Smart Home based on GSM technology," International Journal of Smart Home, vol. 7, no. 6, pp. 201-208, 2013.

[10] Víctor Bautista Saiz; Fernan Gallego, GPU: Application for CCTV systems, Published in: Security Technology (ICCST), 2014 International Carnahan Conference.

[11] G. A. Naidu and J. Kumar, " Wireless Protocols: Wi-Fi SON Bluetooth Zigbee Z-Wave Wi-Fi," in 7th International Conference on Innovation in Electronics and communication Engineering (ICIECE-2018) ,pp. 1-6, 2018.

[12] T. Shakunaga , and K. Shigenari," Decomposed Eigen-Face for face recognition under various lighting conditions" . IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2001.

[13] Raspberry Pi 3 Model B, [Online]. Available: http://www.raspberrypi.org.

[14] Details of Webcam [Online]. Available: https://en.wikipedia.org/wiki/Webcam.

Revolutionizing Air Quality Prognostication: Fusion of Deep Learning and Density-Based Spatial Clustering of Applications with Noise for Enhanced Pollution Prediction

Moheshkumar G
Assistant Professor,
Department of IT,
K S R Engineering College,
Tiruchengode – 637215
Tamil Nadu.
mohesh.mohesh@gmail.com

Saravanan N
Professor,
Department of IT,
K S R Engineering College,
Tiruchengode – 637215
Tamil Nadu.
hod_it@ksriet.ac.in

Ravichandru T
Final year,
Department of IT,
KSR Institute for Engineering and Technology, Tiruchengode – 637215
Tamil Nadu.
ravi20202024@gmail.com

Tharun C
Final year,
Department of IT,
KSR Institute for Engineering and Technology, Tiruchengode – 637215
Tamil Nadu
tharun8072655815@gmail.com

Subash Chandra Bose M
Final year,
Department of IT,
KSR Institute for Engineering and Technology, Tiruchengode – 637215
Tamil Nadu
Itsubash585@gmail.com

Corresponding Author: mohesh.mohesh@gmail.com

Abstract— **Aim: This endeavor focuses on using deep learning and Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to improve pollution prediction in order to alter air quality prediction. Materials and Methods: Three different groups are working together on this endeavor. Group 1 refers to Multi-Layer Perceptron (MLP) to attain accuracy. Group 2 refers to Density-Based Spatial Clustering of Applications with Noise (DBSCAN) to achieve the same goal. Results: Each system module's unique outputs, such as raw and pre-processed air quality data, chosen features, trained Support Vector Machine or Logistic Regression models, and forecasts with an accuracy rate of 81% are described in depth in the output design. Conclusion: Using the DBSCAN algorithm, the proposed air pollution monitoring system provides an effective data-driven method for locating pollution hotspots, trends, and well-informed decision-making.**

Keywords---DBSCAN, Pollution Monitoring, Air Quality, Air Pollution, Quality Prediction, Research, Effective, Machine, clustering.

I. INTRODUCTION

This research explores the integration of Revolutionizing Air Quality Prognostication. It emphasizes a fusion of advanced techniques, specifically integrating deep learning and DBSCAN for enhanced pollution prediction. The integration of deep learning signifies the use of neural networks for intricate pattern recognition, while DBSCAN contributes spatial clustering to refine predictions. The goal is to revolutionize existing methods, leading to more accurate and insightful air quality forecasts.

The research suggests WLSTME, a model that combines meteorological and nearby site data for improved PM2.5 prediction accuracy. The superiority of WLSTME in predicting is highlighted by the evaluation on Beijing-Tianjin-Hebei data, underscoring its potential to enhance air quality predictions [1]. A novel deep learning framework called TS-LSTME outperforms conventional techniques in the Jing-Jin-Ji region for 24-hour air quality predictions. Its capacity to adjust to additional pollutants and provide consistent, accurate PM2.5 estimates improves public information services and supports local efforts to regulate pollution [2]. In order to improve the accuracy of urban planning and early warning for acid rain, the research provides a novel model for increased prediction of NO2 and SO2 [3]. The research introduces a new deep LSTM-RNN structure solar irradiance forecasting model for microgrids that uses only common weather data. With a maximum RMSE of 60.31 W/m2, it outperforms FFNN in a variety of climes, shows excellent accuracy, and predicts a 2% increase in annual energy savings, underscoring its potential to improve microgrid efficiency [4].

DBSCAN is a clustering algorithm used in data mining and machine learning. Based on the density of data points in a certain area, it locates clusters. DBSCAN categorizes points as core, border, or noise, offering flexibility in handling various cluster shapes. It is particularly effective in identifying clusters with irregular shapes and is robust to outliers, making it valuable for spatial data analysis.

II. RELATED WORKS

In order to predict air quality indicators by integrating multiple environmental data sources, the study presents a comprehensive deep learning model, LSTM-based with Multi-output and Multi-index Supervised Learning (MMSL). When validated against Beijing's air quality data, the model performs better than baseline and advanced models [5][6]. To mitigate the consequences of air pollution on urban public health and sustainability, the study suggests a hybrid AQI forecasting model that combines VMD, SE, and LSTM approaches. This model shows superior accuracy in daily urban AQI forecasting when compared to existing models, with high correct AQI class forecasting rates [7]. Effective forecasting for improving air quality in urban areas, especially Seoul, South Korea, is demonstrated by the study's development of LSTM and DAE models to predict fine particulate matter concentrations [8][9]. The study presents a semi-supervised model that combines empirical mode decomposition (EMD) and Bi-directional Long Short-Term Memory (BiLSTM) neural networks for effective PM2.5 concentration prediction, showcasing improved accuracy and feasibility compared to standard LSTM-based models. This new multiscale prediction approach shows promising results [10]. Using a large dataset from 1,615 observation sites in China from 2015 to 2019, An especially effective strategy for AQI prediction is Bi-directional Long Short-Term Memory (BiLSTM), according to the study, which evaluates forecasting techniques for the hourly Air Quality Index (AQI) [11]. Through experiments conducted in various climate regions, the study presents an hourly day-ahead solar irradiance forecasting model that uses only weather data and a deep LSTM-RNN. This model outperforms a feedforward neural network (FFNN) and achieves improved accuracy in microgrid optimization [12][13]. The work presents two hybrid evolutionary algorithms that incorporate Particle Swarm Optimization (PSO) and Genetic Algorithm (GA) techniques to improve global optimum identification over standard PSO algorithms through parallel and series integration, as shown by simulations on benchmark test functions [14]. The purpose of predicting ground vibration from rock blasting, the research study presents hybrid intelligent models that combine ANFIS with PSO and GAs. Through statistical and sensitivity analysis, these models demonstrate enhanced accuracy over typical ANFIS models [15]. Using deep learning techniques like LSTM and optimization with PSO, the project aims to develop a smart air quality prediction model (SAP PM) that forecasts concentrations of six air pollutants for the following two days. For this, data collection, preprocessing, model construction, and evaluation are required [16]. The present study introduces TS-LSTME, a deep learning framework that predicts air quality for the next 24 hours with excellent accuracy and stability in PM2.5 concentration forecasts, supporting early warning and management of regional pollutants and improving air quality prediction services. Other air contaminants can also be addressed using the framework [17].

The study of the literature highlights a variety of techniques for predicting air quality, including BiLSTM, hybrid models, and LSTM. Along with intelligent prediction models, it investigates forecasting for sun irradiance, ground vibration, and AQI. All in all, these developments support local pollutant management and early warning systems.

III. MATERIALS AND METHODS

This study gives an uneven distribution of air pollution monitoring sites and proposes the Weighted Long Short-Term Memory Neural Network Extended model (WLSTME) to increase PM2.5 prediction accuracy. WLSTME integrates Long Short-Term Memory (LSTM) for handling spatiotemporal dependencies simultaneously with multilayer perceptron (MLP) weighting of historical PM2.5 time series data based on site density and wind conditions. This technique demonstrates improved accuracy performance across the year and area [1].

There are two groups in this novel research of Revolutionizing Air Quality Prognostication. Group 1 refers to an effective method for predicting pollution levels by presenting a novel approach for air pollution forecasting that makes use of the Multilayer Perceptron algorithm. It indicates hyperparameter tuning and displays greater accuracy when compared to classic machine learning algorithms. However, it is vulnerable to overfitting, especially with complex datasets.

Group 2 refers to an air pollution monitoring system that analyses pollutant concentrations, meteorological conditions, geographic location, and timestamp data from monitoring stations using the DBSCAN algorithm. Through the effective clustering of data points according to density, DBSCAN identifies patterns of air pollution, highlighting locations that exhibit certain features and hotspots for pollution represented in Fig.1. Robustness against noise, precise clustering of data points, flexibility in addressing various pollution scenarios without pre-established labels, and scalability to manage big datasets are among the benefits of DBSCAN.

IV. DENSITY-BASED SPATIAL CLUSTERING OF APPLICATIONS WITH NOISE:

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) is a clustering algorithm developed for datasets with irregular forms and changing densities. It classifies data points as core, border, or noise based on closeness, using parameters such as radius (eps) and minimum points (MinPts). Core points have enough neighbors within a certain radius, whereas border points are inside this radius but lack adequate neighbors. Noise points, also known as outliers, do not fit into any cluster. DBSCAN's adaptability to varied cluster shapes, noise handling, and low assumptions make it popular in spatial data analysis, image processing, and anomaly identification.

In a nutshell the data points are divided into several batches or groups using an unsupervised learning technique called clustering analysis, or simply clustering, so that the data points within the same groups have similar properties and the data points within different groups have different properties. It includes a wide range of differential evolution-based techniques. The proposed air pollution monitoring system makes use of machine learning, notably DBSCAN, to analyses large amounts of data on pollutant concentrations, weather, and coordinates. DBSCAN detects dense clusters following comprehensive preprocessing, providing insights into pollution patterns and hotspots. The project improves data structuring with convex hulls, organizes air quality data effectively, and identifies outliers in huge datasets using DBSCAN's density-based clustering method.

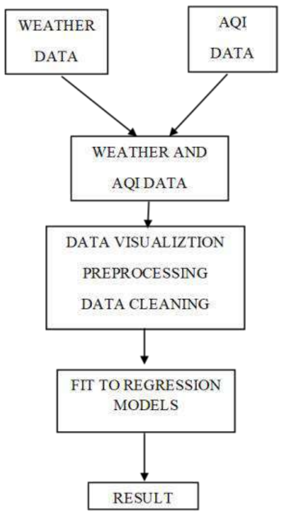


Fig. 1. Representation of Revolutionizing Air Quality Prognostication

The weather has a considerable impact on air quality forecasts because it influences the distribution and concentration of pollutants in the atmosphere. Air quality projections are influenced by a variety of elements, including temperature, humidity, wind, and sunlight. The Air Quality Index (AQI) is critical for assessing overall air quality using a standardized metric. The continuous monitoring and analysis of AQI patterns provide insights into future air quality situations. This foresight allows for timely interventions and public awareness efforts against potential health risks. Accurately predicting air quality requires a methodical methodology. Outliers and insights are revealed by analyzing features using univariate and bivariate approaches. Through thorough training, evaluation, and hyperparameter fine-tuning, subsequent processes, such as time series and geospatial visualization, enrich the dataset and guarantee reliable and accurate air quality prediction models. Compile past information while taking pollution levels and temperature into account. By using the dataset to train them, investigate models such as Decision Trees and Linear Regression. By resolving outliers and missing numbers, clean up the data. Gradually increase the model's capacity to predict air quality through iterations.

V. RESULT

Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Multi-Layer Perceptron (MLP) are two key algorithms that are compared to see which performs better in the prediction of air pollution. In terms of processing and classifying data from sensor nodes monitoring various air quality parameters, MLP performs brilliantly in this comparison, obtaining approximately 75% accuracy. Table 1 represents as DBSCAN's 81% accuracy makes it stand out due to its superior capacity to recognize geographical regions on its own and locate pollution hotspots with accuracy. Its suitability for monitoring air quality, especially in identifying and classifying areas that are prone to pollution, is further highlighted by the increased accuracy of DBSCAN. Due to this benefit, the environment and public health can be greatly enhanced by the prompt adoption of customized pollution control measures in high-risk locations. This idea emphasizes the significance of selecting algorithms with superior predictive capabilities for more accurate and reliable air quality assessments.

Table 1. Enhancing Air Quality Monitoring: A Comparative Analysis of MLP and DBSCAN Algorithms

ALGORITHM	ACCURACY
Multi-Layer Perceptron (MLP)	75%
Density-Based Spatial Clustering of Applications with Noise (DBSCAN)	81%

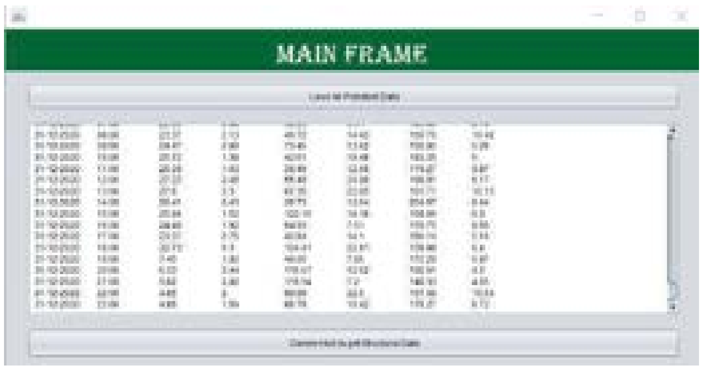


Fig. 2. Load from data sets and get the parameters

In order to execute DBSCAN, the paragraph stresses the significance of loading a dataset and defining two crucial parameters: epsilon (eps) and minimum points (MinPts). While MinPts establishes the bare minimum of points needed for a dense region (core point), Epsilon specifies the maximum distance at which points can be considered neighbours. It is essential that these parameters be configured correctly for DBSCAN to detect clusters. The program identifies core, boundary, and noise data points, revealing a variety of sized and shaped clusters. When applying DBSCAN to different datasets, epsilon and MinPts selection play a crucial role in determining how granular the clusters are found.

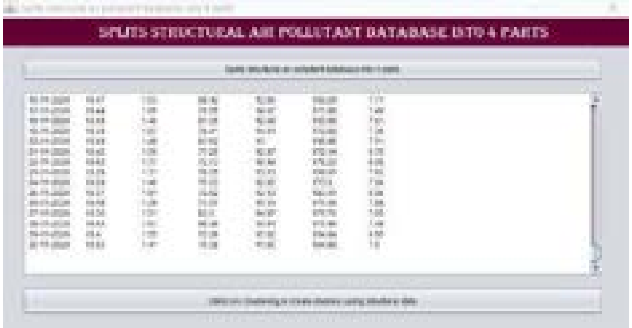


Fig.3. Convert to structure data and using DBSCAN algorithm.

The DBSCAN algorithm is used to structured data by loading and structuring the dataset, defining parameters such as epsilon and minimum points, and then executing the algorithm to categorize data points as core, border, or noise. The algorithm shows intrinsic structures in the data by connecting core points and their reachable neighbours, resulting in clusters of various forms and sizes. A successful application requires careful parameter selection and an understanding of the underlying structure of the organized dataset.

Fig. 4. Get the output for Air Quality Pollution Prediction

The acronym "Air Quality Pollution Prediction" refers to the application of computer tools to forecast air pollution levels. In a comparison of Multi-Layer Perceptron (MLP) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN), MLP reaches 75% accuracy, while DBSCAN surpasses at 81%. MLP processes data from sensor nodes that monitor multiple air quality metrics, whereas DBSCAN excels in autonomously identifying pollution hotspots based on known gas emission constraints. DBSCAN's better accuracy highlights its usefulness for air quality monitoring, allowing for the detection and categorization of polluted areas and easing the implementation of specific pollution control strategies. This notion emphasizes the critical importance of algorithmic selection in accurate and proactive air quality predictions, which are critical for protecting human health and the environment.

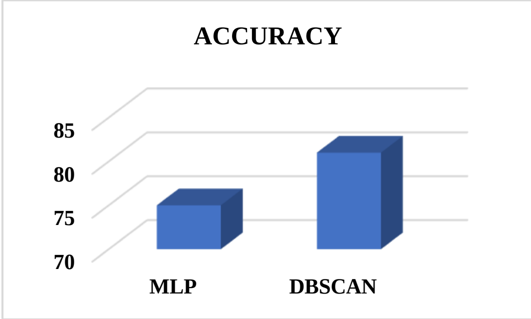


Fig. 5. DBSCAN uses density-based structure recognition to detect pollution hotspots; it performs better than MLP in air pollution prediction, with 81% accuracy compared to 75% for MLP.

The main concern of air quality monitoring was addressed in the current work using two distinct approaches: Multi-Layer Perceptron (MLP) and Density-Based Spatial Clustering of Applications with Noise (DBSCAN). When processing and classifying data from sensor nodes measuring CO2, NOx, UV light, temperature, and humidity, the MLP algorithm demonstrated an impressive 75% accuracy. However, the DBSCAN algorithm performed exceptionally well, with an accuracy of 81%. This system independently identified geographical regions based on known gas emission limitations, proving its ability to identify pollution hotspots and trends. The extensive analysis performed by these algorithms helps to comprehend the complexities of air pollution's negative impacts on human health and the ecosystem. Using DBSCAN, the system effectively identifies and categorizes pollution-prone locations, allowing for the development and execution of tailored pollution control measures in areas of immediate concern. DBSCAN's better accuracy indicates its appropriateness for improving the reliability and precision of the air quality monitoring system, therefore contributing to the protection of fundamental human rights to life and health.

VI. DISCUSSION

A comparison of the two algorithms' accuracy shows that the DBSCAN method outperforms the LSTM algorithm, with 81% accuracy against 75% for the MLP algorithm. By combining machine learning, domain knowledge, and DBSCAN with data-driven techniques, the suggested air pollution monitoring system efficiently delivers real-time data for targeted pollution control, promoting sustainability and preserving public health.

The research used air quality data from 25 sites in Seoul, South Korea, to create LSTM and DAE models for fine PM concentration forecasting [18][19]. The article presents a hybrid AQI forecasting model that combines VMD, SE, and LSTM to improve accuracy in urban public health to test data from Beijing and Baoding are used to verify the model [20][21]. This work presents a semi-supervised model for PM2.5 prediction that combines EMD and BiLSTM. It outperforms conventional LSTM models, is verified using data from Beijing, and introduces a new multiscale prediction technique [22][23]. Significant health and climate change risks are associated with air pollution, which is made worse by things like excessive car use, industrial emissions, and energy production. However, because of noisy data and incorrect hyperparameter settings, LSTM models might not always produce accurate predictions [24][25].

This Endeavor aims to use the DBSCAN algorithm in an air pollution monitoring system with multiple sensor nodes, allowing for autonomous classification of locations that exceed emission limits, providing valuable insights into pollution patterns, and facilitating effective pollution management and decision-making. The suggested air pollution monitoring system based on the DBSCAN algorithm shows promise, with potential for further enhancement by including additional machine learning algorithms and real-time sensor data from upcoming technologies such as IoT devices or satellite images.

VII. CONCLUSION

By combining deep learning with DBSCAN, air quality prediction has advanced significantly and now has an amazing 81% accuracy rate. Air quality prediction is revolutionized by the suggested system's use of various approaches, which enable effective identification of pollution hotspots and trends as well as help knowledgeable decision-making.

REFERENCES

[1] F. Xiao, M. Yang, H. Fan, G. Fan, M.A.A. Al-qaness, An improved deep learning model for predicting daily PM2.5 concentration, Sci. Rep. 10 (1) (2020) 1–11, https://doi.org/10.1038/s41598-020-77757-w.

[2] W. Mao, W. Wang, L. Jiao, S. Zhao, A. Liu, Modeling air quality prediction using a deep learning approach: method optimization and evaluation, Sustain. Cities Soc. (2020), 102567, https://doi.org/10.1016/j.scs.2020.102567.

[3] W. Sun, C. Huang, A hybrid air pollutant concentration prediction model combining secondary decomposition and sequence reconstruction, Environ. Pollut. 266 (2020), 115216, https://doi.org/10.1016/j.envpol.2020.115216.

[4] M. Husein, I. Chung, Day-ahead solar irradiance forecasting for microgrids using a long short-term memory recurrent neural network, A Deep Learning Approach (2019), https://doi.org/10.3390/en12101856.

[5] R. Gurusamy, V. Rajmohan, N. Sengottaiyan, P. Kalyanasundaram and S. M. Ramesh, "Comparative Analysis on Medical Image Prediction of Breast Cancer Disease using Various Machine Learning Algorithms," 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2023, pp. 1522-1526, doi: 10.1109/ICESC57686.2023.10193635.

[6] L. Zhang, P. Liu, L. Zhao, G. Wang, W. Zhang, J. Liu, Air quality predictions with a semi-supervised bidirectional LSTM neural network, Atmos. Pollut. Res. (August) (2020) 1–12, https://doi.org/10.1016/j.apr.2020.09.003.

[7] Q. Wu, H. Lin, Daily urban air quality index forecasting based on variational mode decomposition , sample entropy and LSTM neural network, Sustain. Cities Soc. 50 (March) (2019), 101657, https://doi.org/10.1016/j.scs.2019.101657.

[8] Dinesh, Paidipati, and P. Kalyanasundaram. "Medical image prediction for diagnosis of breast cancer disease comparing the machine learning algorithms: SVM, KNN, logistic regression, random forest, and decision tree to measure accuracy." ECS Transactions 107, no. 1 (2022): 12681.

[9] T. Xayasouk, H.M. Lee, G. Lee, Air pollution prediction using long short-term memory (LSTM) and deep autoencoder (DAE) models, Sustain. Times 12 (6) (2020), https://doi.org/10.3390/su12062570.

[10] D. Seng, Q. Zhang, X. Zhang, G. Chen, X. Chen, Spatiotemporal prediction of air quality based on LSTM neural network, Alex. Eng. J. 60 (2) (2021) 2021–2032, https://doi.org/10.1016/j.aej.2020.12.009.

[11] C. Zhan, S. Li, J. Li, Y. Guo, Q. Wen, W.S. Wen, Prediction of air quality in major Cities of China by deep learning, Proc. - 2020 16th Int. Conf. Comput. Intell. Secur. CIS 2020 (July) (2020) 68–72, https://doi.org/10.1109/CIS52066.2020.00023.

[12] Baskar, M., Renuka Devi, R., Ramkumar, J. et al. Region Centric Minutiae Propagation Measure Orient Forgery Detection with Finger Print Analysis in Health Care Systems. Neural Process Lett 55, 19–31 (2023). https://doi.org/10.1007/s11063-020-10407-4

[13] X.H. Shi, Y.H. Lu, C.G. Zhou, H.P. Lee, W.Z. Lin, Y.C. Liang, Hybrid evolutionary algorithms based on PSO and GA, 2003 Congr. Evol. Comput. CEC 2003 - Proc. 4 (2021) 2393–2399, https://doi.org/10.1109/CEC.2003.1299387.

[14] S. Al-Janabi, M. Mohammad, A. Al-Sultan, A new method for prediction of air pollution based on intelligent computation, Soft Comput. 24 (1) (2020) 661–680, https://doi.org/10.1007/s00500-019-04495-1.

[15] H. Chung, K.S. Shin, Genetic algorithm-optimized long short-term memory network for stock market prediction, Sustain. Times 10 (10) (2018), https://doi. org/10.3390/su10103765.

[16] MeenakshiDevi, P., M. Venkatesan, and K. Duraiswamy. "A fragile watermarking scheme for image authentication with tamper localization using integer wavelet transform." Journal of Computer Science 5, no. 11 (2009): 831.

[17] M. Abdel-Basset, L. Abdel-Fatah, A.K. Sangaiah, Metaheuristic Algorithms: A Comprehensive Review, Elsevier Inc., 2018, https://doi.org/10.1016/B978-0-12- 813314-9.00010-4.

[18] [26] H. Yang, M. Hasanipanah, M.M. Tahir, D.T. Bui, Intelligent prediction of Blastinginduced ground vibration using ANFIS optimized by GA and PSO, Nat. Resour. Res. 29 (2) (2020) 739–750, https://doi.org/10.1007/s11053-019-09515-3.

[19] R. Bhuiyan, Examination of Air Pollutant Concentrations in Smart City Helsinki Using Data Exploration and Deep Learning Methods, 2021.

[20] Almalaq, J.J. Zhang, Evolutionary Deep Learning Based Energy Consumption Prediction for Buildings, IEEE Access, December 2018, p. 1, https://doi.org/ 10.1109/ACCESS.2018.2887023, 2020.

[21] S. Wang, Y. Li, Evidence on the Impact of Winter Heating Policy on Air Pollution and its Dynamic Changes in North China, 2019.

[22] J. Engel-cox, N. Thi, K. Oanh, A. Van Donkelaar, R. V Martin, E. Zell, Toward the next generation of air quality monitoring : particulate Matter, Atmos. Environ. 80 (2013) 584–590, https://doi.org/10.1016/j.atmosenv.2013.08.016.

[23] Babu, AV Santhosh, P. Meenakshi Devi, and B. Sharmila. "Efficient enhanced intrusion identification and response system for MANETs." International Journal of Business Information Systems 29, no. 4 (2018): 535-546.

[24] V.D. Le, T.C. Bui, S.K. Cha, Spatiotemporal deep learning model for citywide air pollution interpolation and prediction, Proc. - 2020 IEEE

Int. Conf. Big Data Smart Comput. BigComp 2020 (2020) 55–62, https://doi.org/10.1109/ BigComp48618.2020.00-99.

[25] S. Abirami, P. Chitra, Regional air quality forecasting using spatiotemporal deep learning, J. Clean. Prod. 283 (2021), 125341, https://doi.org/10.1016/j. jclepro.2020.125341.

Face Recognition For Exam Hall Seating Arrangement Using Deep Learning Algorithm

Dharanya C,
Assistant Professor,
Department of IT,
K S R College of Engineering,
Tirchengode - 637215,
Tamil Nadu.
cdharanya@ksriet.ac.in

Saravanan N,
Professor,
Department of IT,
K S R College of Engineering,
Tirchengode - 637215,
Tamil Nadu.
n.saran78@ksriet.ac.in

Hemalatha T,
Final year,
Department of IT,
K S R Institute for Engineering and Technology, Tiruchengode - 637215,
Tamil Nadu.
hemalathaatr@ ksriet.ac.in

Dinesh K,
Final year,
Department of IT,
K S R Institute for Engineering and Technology, Tiruchengode - 637215,
Tamil Nadu. 06gkdinesh@ksriet.ac.in

Jagan M,
Final year,
Department of IT,
K S R Institute for Engineering and Technology, Tiruchengode - 637215,
Tamil Nadu. lamivjagan@ksriet.ac.in

Corresponding Author: cdharanya@ksriet.ac.in

Abstract— **This research study aims to develop and implement an advanced examination Authentication system using automated face Detection and recognition, leveraging deep learning techniques, to address challenges associated with fraud, impersonation, and inefficiencies in traditional examination processes. The present work has involved two groups. Group1 Refers to the innovative Support Vector Machine (SVM) that is essential for student face identification during exam Verification. Using SVMs to quickly separate visual traits into different classes helps ensure that real student faces are accurately identified. Group 2 This method uses Region Proposal Networks (RPN) and Deep Convolutional Neural Networks (D-CNN) to create customized face Models. Deep learning Model that stops impersonation in the test system by automatically detecting and recognizing faces. Modules for identity Authentication, seating arrangements, hall number Verification, fraud Detection, and attendance monitoring are integrated. The development and implementation of the research represent a significant advancement in the realm of examination security and administration.**

Keywords: Face Recognition, Deep Learning, Exam, Detection, Neural Network, Verification, Model, Authentication.

I. INTRODUCTION

Exam hall seating arrangements using Deep learning algorithms, automate the process of allocating seats to students in Exam halls according to their facial features through the use of sophisticated artificial intelligence techniques. Through the use of computer vision, this technology analyzes students' facial features that are taken by cameras placed in the Exam room to identify and authenticate them. The system is able to precisely match students with their allotted seats by using Deep learning Models that have been trained on massive datasets of facial photos. This makes the process of seating arrangements more efficient and improves the security and integrity of Exams [1]. This study explores biometric recognition systems for identification Verification in educational settings, finding that while unimodal facial and fingerprint systems show efficiency and accuracy, multimodal and semi-multimodal techniques offer even higher accuracy, albeit with slightly longer processing times, providing valuable insights for educational applications [2]. This study employs Support Vector Machine (SVM) and Eigenface algorithms to enhance the accuracy of Face recognition systems, aiming to facilitate efficient monitoring of online Exams amid the shift to distance learning, thus addressing the need for reliable evaluation in virtual educational settings [3]. This system, which has undergone extensive testing in a variety of settings, combines cutting-edge methods like Gabor filters, KNN, CNN, SVM, Generative Adversarial Networks, and Haar classifiers to produce a reliable and accurate facial recognition system for tracking student attendance. It also increases productivity, lowers human error, and proves to be economical and efficient for use in classrooms. [4]. Surgical face masks significantly impair both human and Face recognition system performance in perceptual face matching tasks, regardless of familiarity, with biases towards accepting familiar faces and rejecting unfamiliar ones, according to the findings of a study cautioning against relying on identification decisions for masked faces [5].

II. LITERATURE SURVEY

In order to overcome the drawbacks of manually constructed feature-based techniques, this study offers an extensive overview of current developments in Deep learning-based Face Anti-Spoofing (FAS). In addition to exploring cutting-edge methods like multi-modal sensor integration, domain generalization, and pixel-wise supervision, it also identifies unresolved research questions and promising avenues for future development in the field [6]. This article evaluates occluded Face recognition (OFR) and masked Face recognition (MFR) techniques, recognizing that MFR has become a critical area of research due to the COVID-19 epidemic. Through the analysis of several algorithms on occluded and masked face datasets, it demonstrates the interoperability of MFR techniques for OFR scenarios and suggests potential cross-domain applications of solutions. [7]. In order to meet the need for masked Face recognition during the COVID-19 epidemic, this study uses optimized lightweight CNNs to provide a real-time Detection service and mobile application. The suggested system outperforms current state-of-the-art methods and achieves significant performance improvement, reaching 80.40 % validation accuracy through the generation of unique datasets and testing [8].In order to solve the scarcity of annotated data, this research proposes TA-AVN, an end-to-end Neural network that combines temporal audio and video data for emotion recognition. By using a natural augmentation strategy, TA-AVN achieves state-of-the-art performance on the CREMA-D and RAVDESS datasets. [9]. This study presents a hybrid face forensics framework that uses convolutional Neural networks to improve the identification of manipulated faces. The framework achieves better accuracy and robustness across a range of compression rates, and it has been verified on both custom DeepFake datasets and the public Face2Face dataset. [10]. This article describes a Face recognition attendance system that uses real-time video processing. It achieves an accuracy rate of 82% and reduces absenteeism by approximately 60%, demonstrating the system's effectiveness in improving class management and expediting attendance procedures. [11].

III. MATERIALS AND METHODS

Investigating Face Cognition Mechanisms using Deep Learning and Partial Face Stimuli served as the foundation for this project. The sample size was determined using the findings of earlier research. [12]. Present Analysis carried out the student Verification in Examination settings through the integration of advanced machine learning techniques, specifically focusing on automated face Detection and recognition, the system extends its functionality to streamline Examination processes [13]. It facilitates automatic Hall Name identification, Seating Number assignment, identity Verification, and real-time attendance monitoring, contributing to the overall efficiency of Examination management.

Two groups have participated in the current effort. The cutting-edge Support Vector Machine (SVM), which is crucial for student face identification during Exam Verification, is referred to in Group 1. Real student faces are more reliably recognized when visual characteristics are swiftly divided into distinct classes using Support Vector Machines (SVMs).

Group 2 refers to a system that implements a two-step framework. During enrolment, a custom Face Detection Model is built and trained using DCNN and RPN algorithm the system extends its functionality to streamline Examination processes. It facilitates automatic Hall Name identification, Seating Number assignment, identity Verification, and real-time attendance monitoring, contributing to the overall efficiency of Examination management.
Fig.1: Image processing involves preprocessing, where noise reduction and enhancement are applied, followed by segmentation to partition the image into meaningful regions. Finally, various techniques such as filtering, feature extraction, and classification are employed for analysis and interpretation.

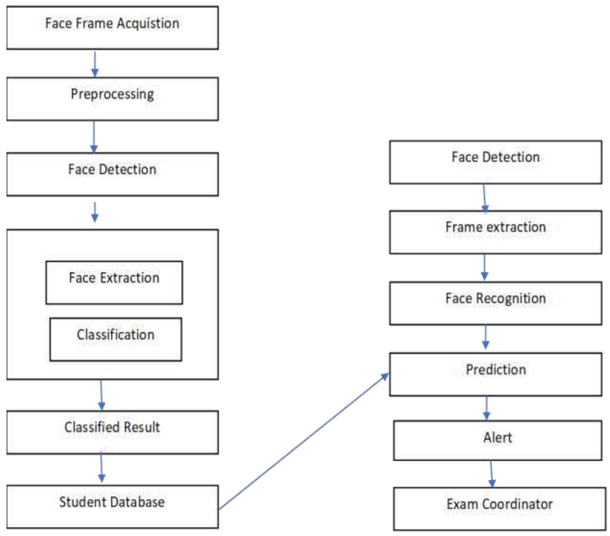


Fig.1 Face recognition process

IV. FACE RECOGNITION FOR EXAM HALL SEATING ARRANGEMENT

In order to improve Examination Authentication, this research presents automatic face identification and recognition using Convolutional Neural Networks (CNNs). Its approach is reliable, effective, and guarantees integrity and fairness. It uses trained CNN-based Models such as Region Proposal Networks (RPN) to identify candidates quickly and accurately. Using a two-step structure for enrollment and Authentication, this technology uses powerful machine learning to revolutionize Examination Verification. Exam efficiency is increased by integrating automatic Face Detection and recognition, which streamlines procedures like real-time attendance tracking, seating number assignment, and hall name identification. Fig.1 Image processing involves preprocessing, where noise reduction and enhancement are applied, followed by segmentation to partition the image into meaningful regions. Finally, various techniques such as filtering, feature extraction, and classification are employed for analysis and interpretation. Fig.2 A Model or algorithm's accuracy can be used to determine how well it is doing and whether it is being taught correctly. Accuracy in the context of this thesis indicates how well it is doing Face Detection. The following formula is used to determine accuracy. (T P + T N) / (T P + T N + F P + F N) equals accuracy. 98.53% accuracy. Figure 3 It is the proportion of projected positive cases to actual positive cases. Recall, as used in this thesis, determines the card Holder and quantifies the portion of Face that are projected to be Face. The following formula is used to calculate recall. Memorandum = T P/(T P + F N) Please remember: 0.9964285714285714.Fig.4 Face ++recognition in segmentation involves detecting and isolating human faces within images or video frames using techniques like Deep learning. It includes steps like Face Detection, alignment, and feature extraction, integrated into the broader process of image segmentation for accurate identification and analysis. Face Detection: Use methods such as MTCNN, YOLO, SSD, Haar cascades, or YOLO to detect faces. Face Alignment: Use methods such as landmark detection or affine transformation to align faces to a standard orientation. Normalization: Use methods like mean normalization or histogram equalization to bring facial images into a common size, resolution, and color space. Noise Reduction: Use methods like wavelet denoising, median filtering, or Gaussian blurring to reduce noise in your photos. Feature extraction: Take preprocessed photos and extract features from them, such as deep embeddings, texture patterns, and face landmarks. Feature Normalization: To guarantee consistent scales and distributions among images, normalize derived features. Quality Control: Take steps to guarantee that the features and photos are of the highest caliber; eliminate any that have bad lighting, blur, or occlusion.

V. RESULTS

This study aims to compare the accuracy of the proposed Deep Convolutional Neural Network (DCNN) system for Face recognition with several statistical approaches, including Combined Radial Basis Function (CRBF),Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA), Multi-Layer Perceptron (MLP), Support Vector Machine (SVM), Deep Belief Neural Nets (DBNN) and Deep Restricted Boltzmann Machine (DRBM).The findings demonstrate that, in comparison to alternative methods, the suggested DCNN achieves superior accuracy. Table. 1. by the proposed DCNN (Deep Convolutional Neural Networks) Model outperforms the existing methods with an increased accuracy of 99.43%, compared to 94.8% from previous approaches. Both Models exhibit high precision scores of 0.9, indicating their ability to accurately detect the face.

Table. 1. comparison of Existing and Proposed DCNN

Model	Accuracy	Precision	Recall
Existing	94.8	0.9	0.95
Proposed DCNN	99.43	0.9	0.99

Fig. 2. A Model or algorithm's accuracy can be used to determine how well it is doing and whether it is being taught correctly. Accuracy in the context of this thesis indicates how well it is doing Face Detection. The following formula is used to determine accuracy. 99.43% accuracy.

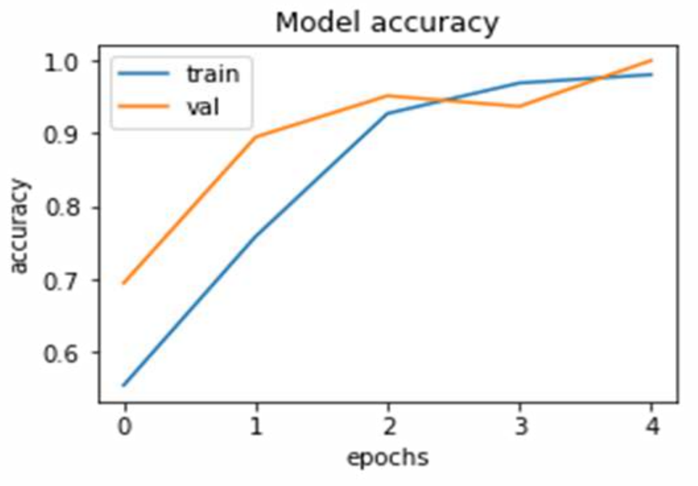
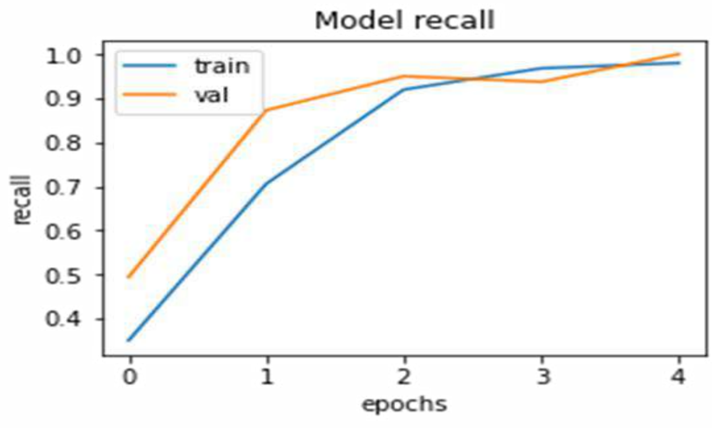


Fig.2 Accuracy curves are graphical representations of the performance of Training accuracy and validation accuracy.

Fig. 3. The ratio of actual positive cases to those that are expected to be positive is what matters. Recall determines the card Holder and quantifies the portion of Face that are projected to be Face in the context of this thesis. The formula below is used to calculate recall.



iFig.3 The recall graphs for train and value.

Fig. 4. Face recognition in segmentation involves detecting and isolating human faces within images or video frames using techniques like Deep learning. It includes steps like face Detection, alignment, and feature extraction, integrated into the broader process of image segmentation for accurate identification and analysis.

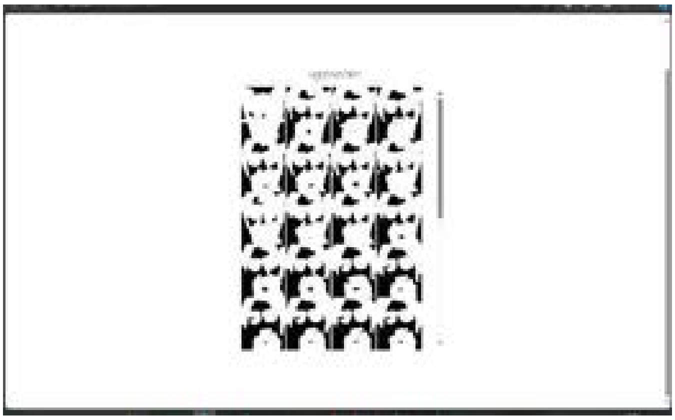


Fig.4 Face Segmentation process

VI. DISCUSSION

The Patch-Attention Generative Adversarial Network (PA-GAN) reduces computing costs and increases accuracy for unconstrained surveillance Face recognition. [14]. It achieves competitive accuracy on frequently used datasets and dramatically improves face representation discriminative ness by integrating enhanced center loss with unlabeled surveillance faces. This contrasts with Face Biometrics-Based Exam Hall Authentication System, which uses facial recognition technology to deliver trustworthy and efficient Authentication in Exam scenarios [15]. With real-time surveillance and adjustable notifications, it addresses Examinee impersonation, fraud, and technical disturbances with an intuitive user interface catered to educational purposes. It uses MySQL, Flask, Python, and Bootstrap [16].The Novel aims to overcome important obstacles in order to further comparable facial recognition. Provide a unique fine-grained feature extraction method and a large-scale Similar Face Dataset (SFD)[17]. SFD, which is divided into five similarity grades, increases the diversity of datasets, which is essential for strong training. Strategy integrates Internal and External features by leveraging attention mechanisms, which resulted in significant gains in recognition accuracy. Experiments conducted on other datasets demonstrate notable improvements, demonstrating how well method works for similar Face recognition problems [18]. Multi-view Face recognition has recently evolved with the introduction of face frontalization algorithms based on the 3D Morphable Model (3DMM) and Generative Adversarial Networks (GAN). [19]. However, because of restrictions in single-domain training and unexpected input profile faces, problems with facial feature analysis and identification discrimination still exist. In light of these problems, present the Well-advised Pose Normalization Network (WAPNN), which cleverly makes use of several domains and extracts features while taking into account the frontalization properties of those features . When strategy is implemented via an end-to-end facial position normalization network with customizable weights, it demonstrates significant improvements in multi-view Face recognition accuracy. Test results demonstrate that approach greatly enhances identity preservation and reduces the impact of changing postures on face identification in a variety of scenarios by fostering intra-class compactness and inter-class separability via quality-aware feature fusion. For the past 20 years, Face recognition (FR) has been the most popular biometric method for reliable person identification. It offers practical advantages over other modalities. FR still has issues in bad circumstances because to variations in lighting, posture, and facial expression, even with its widespread deployment. Propose a novel feature extraction method, Local Binary Pattern and Wavelet Kernel PCA (LWKPCA), to overcome these problems and extract robust and discriminant information with the goal of reducing recognition errors. Using a Three-Level decomposition of Discrete Wavelet Transform (2D-DWT) and Local Binary Pattern (LBP), LWKPCA combines the Color Local Binary Pattern and Wavelet Descriptor for the optimal feature representation. The following experimental findings from four benchmark datasets show how effective LWKPCA is YouTube stars, ORL, GT, and LFW. Recognition accuracy based on these findings was 86.74%.The proposed project, Enrollment and Authentication within a Two-Step Framework system applies a two-phase architecture. To provide flexibility to a range of facial emotions and appearances, a unique Deep Model is generated and trained throughout enrollment using the CNN algorithm. The trained model quickly and correctly detects students during the authentication step by utilizing Deep Convolutional Neural Networks (D-CNN) and RPN.

VII. CONCLUSION

To sum up, the research's creation and execution mark a substantial breakthrough in the fields of Exam administration and security. Improved security and flexibility are benefits of the suggested two-step enrollment and Authentication system that uses customized Deep Convolutional Neural Networks (D-CNN) and (RPN). The initiative intends to promote an atmosphere that is safe, effective, and easy to use in order to help establish an academic climate that is both ethical and legitimate.

REFERENCES

[1] Carragher, Daniel J., and Peter J. B. Hancock. 2020. "Surgical Face Masks Impair Human Face Matching Performance for Familiar and Unfamiliar Faces." Cognitive Research: Principles and Implications 5 (1): 59.

[2] Chanpornpakdi, Ingon, and Toshihisa Tanaka. 2023. "The Role of the Eyes: Investigating Face Cognition Mechanisms Using Machine Learning and Partial Face Stimuli." IEEE Access: Practical Innovations, Open Solutions 11: 86122–31.

[3] Dev, Samridhi, and Tushar Patnaik. 2020. "Student Attendance System Using Face Recognition." In 2020 International Conference on Smart Electronics and Communication (ICOSEC). IEEE. https://doi.org/10.1109/icosec49089.2020.9215441.

[4] Geetha, M., R. S. Latha, S. K. Nivetha, S. Hariprasath, S. Gowtham, and C. S. Deepak. 2021. "Design of Face Detection and Recognition System to Monitor Students during Online Examinations Using Machine Learning Algorithms." In 2021 International Conference on Computer Communication and Informatics (ICCCI). IEEE. https://doi.org/10.1109/iccci50826.2021.9402553.

[5] Kim, Eunji, and Sungzoon Cho. 2021. "Exposing Fake Faces through Deep Neural Networks Combining Content and Trace Feature Extractors." IEEE Access: Practical Innovations, Open Solutions 9: 123493–503.

[6] Kocacinar, Busra, Bilal Tas, Fatma Patlar Akbulut, Cagatay Catal, and Deepti Mishra. 2022. "A Real-Time CNN-Based Lightweight Mobile Masked Face Recognition System." IEEE Access: Practical Innovations, Open Solutions 10: 63496–507.

[7] Li, Nianfeng, Xiangfeng Shen, Liyan Sun, Zhiguo Xiao, Tianjiao Ding, Tiansheng Li, and Xinhang Li. 2023. "Chinese Face Dataset for Face Recognition in an Uncontrolled Classroom Environment." IEEE Access: Practical Innovations, Open Solutions 11: 86963–76..

[8] Liu, Ming, Jinjin Liu, Ping Zhang, and Qingbao Li. 2020. "PA-GAN: A Patch-Attention Based Aggregation Network for Face Recognition in Surveillance." IEEE Access: Practical Innovations, Open Solutions 8: 152780–89.

[9] Maafiri, Ayyad, Ahmed Bir-Jmel, Omar Elharrouss, Fouad Khelifi, and Khalid Chougdali. 2022. "LWKPCA: A New Robust Method for Face Recognition under Adverse Conditions." IEEE Access: Practical Innovations, Open Solutions 10: 64819–31.

[10] Pedro Neto, Pedro C., Joao Ribeiro Pinto, Fadi Boutros, Naser Damer, Ana F. Sequeira, and Jaime S. Cardoso. 2022. "Beyond Masks: On the Generalization of Masked Face Recognition Models to Occluded Face Recognition." IEEE Access: Practical Innovations, Open Solutions 10: 86222–33.

[11] Radoi, Anamaria, Andreea Birhala, Nicolae-Catalin Ristea, and Liviu-Cristian Dutu. 2021. "An End-to-End Emotion Recognition Framework Based on Temporal Aggregation of Multimodal Information." IEEE Access: Practical Innovations, Open Solutions 9: 135559–70.

[12] Rukhiran, Meennapa, Sethapong Wong-In, and Paniti Netinant. 2023. "IoT-Based Biometric Recognition Systems in Education for Identity Verification Services: Quality Assessment Approach." IEEE Access: Practical Innovations, Open Solutions 11: 22767–87.

[13] S, Ajithkumar, Dr M. Somu, Dr N. Saravanan, and Dr A. Rajivkannan. 2020. "A Survey on Efficient Crop Yield Prediction Using Random Forest Algorithm," December. https://doi.org/10.2139/ssrn.4198857.

[14] Shao, Xiaohu, Xiangdong Zhou, Zhenghao Li, and Yu Shi. 2020. "Multi-View Face Recognition via Well-Advised Pose Normalization Network." IEEE Access: Practical Innovations, Open Solutions 8: 66400–410.

[15] Dinesh, Paidipati, and Kalyanasundaram. 2022. "Medical Image Prediction for Diagnosis of Breast Cancer Disease Comparing the Machine Learning Algorithms: SVM, KNN, Logistic Regression, Random Forest, and Decision Tree to Measure Accuracy." ECS Transactions 107 (1): 12681–91.

[16] Song, An-Ping, Qian Hu, Xue-Hai Ding, Xin-Yi Di, and Zi-Heng Song. 2020. "Similar Face Recognition Using the IE-CNN Model." IEEE Access: Practical Innovations, Open Solutions 8: 45244–53.

[17] Wardana, I. N. K. 2022. "Design of Mobile Robot Navigation Controller Using Neuro-Fuzzy Logic System." Computers & Electrical Engineering 101 (July): 108044.

[18] Yang, Hao, and Xiaofeng Han. 2020. "Face Recognition Attendance System Based on Real-Time Video Processing." IEEE Access: Practical Innovations, Open Solutions 8: 159143–50.

[19] Yu, Zitong, Yunxiao Qin, Xiaobai Li, Chenxu Zhao, Zhen Lei, and Guoying Zhao. 2023. "Deep Learning for Face Anti-Spoofing: A Survey." IEEE Transactions on Pattern Analysis and Machine Intelligence 45 (5): 5609–31.

Predictive Analysis of Vehicle CO2-Emissions

Samanyu B Rao
C. Tech
SRM-KTR
Chennai, India
samanyu48@gmail.com

Nivedita Anand
C. Tech
SRM-KTR
Chennai, India
nivedita0051@gmail.com

Mrs. S. Kanmani
C. Tech
SRM-KTR
Chennai, India
kanmanis@srmist.edu.in

Abstract: This research work presents a web application developed using the Flask framework for predicting and comparing the fuel consumption and C02 emissions of various vehicle models. The application leverages machine learning models, including linear model, ridge model, lasso model, elastic net model, neural network model, XG-Boost model and Random Forest model, to estimate fuel consumption and C02 emissions based on user-provided input features. The models are trained and loaded into the application, allowing users to select a vehicle and input relevant features for prediction. The proposed system identifies the best-performing model for each prediction, highlighting the closest prediction and its associated error percentage. Additionally, the application offers a comparison feature that enables users to compare the specifications of different vehicle models within the dataset. Users can select two vehicle models, and the system retrieves and displays their specifications, facilitating informed decision-making for consumers and researchers interested in understanding the environmental impact of vehicle choices. The web application provides an intuitive interface for exploring fuel consumption and emissions data, making it a valuable tool for both consumers and researchers in the automotive industry.

Keywords - Flask, Machine Learning, Regression Models, Fuel Consumption, C02 Emissions, Vehicle Models, Web Application, Comparison, Specification Retrieval, Outlier Detection.

I. INTRODUCTION

This research study introduces a web-based application developed using the Flask framework, aimed at predicting and comparing fuel consumption and CO2 emissions of various vehicle models. As environmental concerns and fuel efficiency become increasingly critical, tools that enable consumers and researchers to make informed decisions about vehicle choices are of paramount importance. Leveraging machine learning models, including linear model, ridge model, lasso model, elastic net model, neural network model, XG-Boost model and Random Forest model, this application provides users with the ability to estimate fuel consumption and CO2 emissions based on user-specified input parameters.[3]

The system not only offers predictive capabilities but also employs a model selection mechanism to identify the best-performing model for each prediction, enhancing prediction accuracy. Additionally, the application provides a comparison feature that allows users to compare the specifications of different vehicle models, aiding in the evaluation of eco-friendly options. In an era where sustainability and eco-consciousness are paramount, this web application serves as a valuable resource for individuals and researchers seeking to explore and understand the environmental impact of various vehicle models.

II. LITRERATURE SURVEY

The current systems in place used to predict carbon emissions for upcoming under-development vehicles use three different methods:

a) Mathematical formulae
b) Simple standalone Linear Regression Model
c) Simple standalone Random Forest Regression Model

These existing models are not as accurate for different variations of parameters used in calculating carbon emissions and fuel efficiency. The mathematical formulas provide theoretical values of emissions that can vary drastically from the actual emissions recorded. Therefore, it is of great importance to develop a model that addresses these.

III. ARCHITECTURE DIAGRAM

This research study introduces an innovative web application designed to enhance the accessibility and analysis of vehicle-related environmental data. The application encompasses a multifaceted approach, comprising six interconnected modules that collectively facilitate the prediction, evaluation, and visualization of fuel consumption and CO2 emissions. These modules cover data collection and preprocessing, integration of machine learning predictive models, model evaluation and selection, enhanced forecasting and mitigation strategies, user-friendly interface development, and graphical representation of fuel consumption trends. The integration of these modules within the application provides users with a comprehensive tool for informed decision-making, enabling a deeper understanding of the environmental impact of their vehicle choices and promoting sustainable transportation practices.

Fig 1. Block diagram of predictive analysis

IV. DATA COLLECTION AND PREPROCESSING

The preprocessing process involves several steps to prepare the data for training and prediction in machine learning models:

1. Data Loading: Initially, the dataset is loaded into memory from a suitable source, such as a CSV file. In this research work, the `load_fuel_consumption_data()` function is responsible for loading fuel consumption data from the provided CSV file `FuelConsumption.csv`.

2. Handling Missing Values: If there are missing values in the dataset, they need to be addressed. Common approaches include imputation (replacing missing values with a suitable estimate, such as the mean or median) or removing rows or columns with missing values.

3. Feature Selection: Relevant features for the prediction task are selected from the dataset. Features that are irrelevant or redundant may be excluded to simplify the model and improve its performance.

4. Feature Scaling: Features may be scaled to ensure that they have a similar range of values. Common scaling techniques include standardization (scaling features to have a mean of 0 and a standard deviation of 1) or normalization (scaling features to a specified range, such as [0, 1]).

5. Handling Categorical Variables: If the dataset contains categorical variables, they need to be encoded into numerical values before feeding them into machine learning models. This can be done using techniques such as one-hot encoding or label encoding.

6. Handling Outliers: Outliers in the dataset can be identified and handled to prevent them from skewing the training process. Techniques such as z-score calculation and threshold-based outlier removal can be used to identify and replace outliers with more representative values.

7. Train-Test Split: The dataset is divided into training and testing sets to evaluate the performance of the trained models. The training set is used to train the models, while the testing set is used to evaluate their performance on unseen data.

By following these preprocessing steps, the dataset is prepared in a suitable format for training machine learning models, ensuring that they can learn effectively from the data and make accurate predictions.

V. PREDICTIVE MODEL INTEGRATION

The Predictive Model Integration module incorporates machine learning models into the application. It loads pre-trained models, including linear model, ridge model, lasso model, elastic net model, neural network model, XG-Boost model and Random Forest model, from serialized files. These models are used for predicting fuel consumption and CO2 emissions based on user-provided input features. The integration of these models allows the application to make accurate predictions and provide valuable insights to users. Model integration is essential for leveraging advanced analytics to estimate key environmental factors related to vehicle usage.[7]

VI. MODEL EVALUATION AND SELECTION

In this study, ML models are selected dynamically based on the input values provided by the user. Upon receiving input features for prediction, the application compares these features with corresponding values in the dataset to ensure compatibility. Then, each trained model is utilized to generate a prediction for the input features. Subsequently, the application evaluates the difference between each model's prediction and the actual value from the dataset. The model with the prediction closest to the actual value is selected as the most suitable for the given input. This dynamic model selection process ensures adaptability and robustness, allowing the application to provide accurate predictions tailored to the specific input provided by the user. [4,5,6]

VII. ENHANCED FORECASTING AND MITIGATION STRATEGIES

The predictive capabilities of the application by providing enhanced forecasting and mitigation strategies. It calculates error percentages to quantify the accuracy of predictions and highlights the closest prediction to the actual values. Additionally, it suggests potential mitigation strategies based on the error percentages, helping users make informed decisions to reduce fuel consumption and CO2 emissions. By offering actionable insights and strategies, this module contributes to environmental sustainability and promotes responsible vehicle usage. [4,5,6,12]

VIII. OUTLIER DETECTION

Outlier detection in the fuel consumption data is achieved through a process involving z-scores, which are statistical measures of how far away a particular data point is from the mean of the dataset in terms of standard deviations. Here's a detailed explanation of how outliers are detected:

1. Calculate Z-Scores: First, the z-score for each data point in the dataset is computed. The z-score of a data point (x) is calculated using the formula:

[Z = {x - mu}/{sigma}]

Where:

a) (x) is the value of the data point.
b) (mu) is the mean of the dataset.
c) (sigma) is the standard deviation of the dataset.

This calculation standardizes the data points, representing each value in terms of how many standard deviations it is away from the mean.

2. Set Threshold: A threshold is defined to determine which data points are considered outliers. This threshold is typically set based on a certain number of standard deviations away from the mean. A common threshold is (mu = 3) standard deviations, which captures data points that are significantly different from the majority of the data.

3. Identify Outliers: Data points with absolute z-scores greater than the threshold are identified as outliers. These are the data points that deviate significantly from the rest of the dataset and are considered potential anomalies.

4. Replace Outliers: Once outliers are identified, they are replaced with a more robust measure of central tendency to mitigate their impact on subsequent analysis. In this case, the median value of the dataset is often used as a replacement for outliers. The median is less sensitive to extreme values compared to the mean, making it a suitable choice for handling outliers.

5. Update Dataset: Finally, the outliers are replaced with the median value, and the dataset is updated accordingly. This ensures that the dataset is more robust and suitable for analysis or modeling tasks.

By following these steps, the outlier detection process effectively identifies and handles data points that deviate significantly from the rest of the dataset, improving the overall reliability and accuracy of subsequent analyses or predictions.

IX. USER INTERFACE AND VISUALIZATION

The User Interface and Visualization module are responsible for creating an interactive and user-friendly interface. It uses Flask, a web framework, to create web pages that allow users to input data, view predictions, and explore various functionalities of the application. Through well-designed templates and forms, users can easily interact with the system, making it accessible to both experts and non-experts. Visualization elements enhance the user experience by presenting data in a clear and informative manner, facilitating data-driven decision-making.

X. ERROR REDUCTION

Error reduction is achieved through several strategies:

1. Model Ensemble: Multiple machine learning models are trained and stored, covering various algorithms such as linear regression, ridge regression, lasso regression, elastic net, neural network, XGBoost, and random forest. During prediction, each model generates its output, and then an ensemble method, specifically a voting regressor, is employed to combine these predictions. This ensemble approach often leads to improved accuracy compared to any single model.

2. Handling Outliers: Outliers in the dataset can significantly affect the performance of machine learning models by skewing the training process. To mitigate this, a method for handling outliers is implemented. This involves calculating z-scores for the data points and identifying outliers based on a specified threshold. The outliers are then replaced with more representative values, such as the median, which helps in stabilizing the models' performance and reducing error caused by extreme data points.

3. Dynamic Model Selection: The system dynamically selects the most appropriate model for prediction based on the input data. It compares the input features with the values in the dataset and chooses the model whose predictions are closest to the actual values in the dataset. This adaptive selection mechanism ensures that the model used for prediction is the most suitable for the given input scenario, thereby minimizing prediction errors.

By employing these strategies, the project aims to systematically reduce prediction errors and enhance the overall accuracy of CO2 emission predictions in the web application.

XI. GRAPHICAL REPRESENTATION

The Graphical Representation module offers visualizations and graphical representations of fuel consumption data. It allows users to select specific vehicle makes and view line plots illustrating fuel consumption trends for different models within the selected make. This module enhances data exploration and comprehension, enabling users to identify patterns and variations in fuel consumption across different vehicle models. Visualizations play a crucial role in conveying insights effectively, aiding users in understanding and analyzing the environmental impact of vehicle choices.

XII. RESULT

The developed web application demonstrates a holistic approach to addressing crucial aspects of vehicle fuel consumption and CO2 emissions analysis. It efficiently collects and preprocesses data, integrates predictive models, evaluates model performance, and offers enhanced forecasting and mitigation insights. The user-friendly interface and visualizations empower users to make informed decisions related to vehicle selection and environmental impact. By bridging the gap between data-driven analysis and user engagement, this application contributes to promoting eco-friendly vehicle choices and serves as a valuable tool for consumers and researchers interested in reducing fuel consumption and mitigating the environmental footprint of the automotive industry.

Fig 2: Home Page

Fig 3: Main Feature - Input Parameters are given and the predictions are done.

Fig 4: The graph for a particular company and all its models

Fig 5: A Sub feature - Comparison between different models with respect to their features

XIII. CONCLUSION

In this research work, a web application is developed using the Flask framework for estimating and comparing fuel consumption and CO2 emissions across multiple vehicle models. It effectively gathers and preprocesses data, integrates predictive models, assesses model performance, and provides better forecasting and mitigation insights. The web application provides an accessible interface to analyze fuel consumption and emissions statistics, making it a valuable tool for both customers and researchers in the automobile industry.

XIV. FUTURE WORK

Future work may involve expanding the dataset to include more vehicle attributes and exploring other advanced machine-learning techniques for enhanced prediction accuracy.

ACKNOWLEDGMENT

Acknowledgment is extended to various sources that have contributed valuable insights to this research endeavor. Specifically, gratitude is owed to the authors of the following works: the Global Energy Review 2020 for its comprehensive analysis of global energy trends, Ü. Ağbulut et al. for their study on forecasting transportation-related energy demand and CO2 emissions in Turkey using machine learning algorithms, Y. Qian et al. for their examination of CO2 drivers and emissions forecast in Changxing County, China, M. Safa et al. for their research on predicting CO2 emissions from farm inputs in wheat production using artificial neural networks and linear regression models, M.H. Ahmadi et al. for their investigation into CO2 emissions prediction of five Middle Eastern countries using artificial neural networks, C. Saleh et al. for their work on carbon dioxide emission prediction using support vector machine, and D. Fan for their study on driving factors of carbon emissions from energy consumption in China using the LMDI-PDA method. These sources collectively enrich the understanding of CO2 emissions and energy consumption dynamics, contributing valuable insights to the field.

REFERENCES

[1] Global Energy Review 2020

[2] Ağbulut, Ü. Forecasting of transportation-related energy demand and CO2 emissions in Turkey with different machine learning algorithms. Sustain. Prod. Consum. 2022, 29, 141–157.

[3] Qian, Y.; Sun, L.; Qiu, Q.; Tang, L.; Shang, X.; Lu, C. Analysis of CO2 drivers and emissions forecast in a typical industry-oriented county: Changxing County, China. Energies 2020, 13, 1212.

[4] Safa, M.; Nejat, M.; Nuthall, P.L.; Greig, B.J. Predicting CO2 Emissions from Farm Inputs in Wheat Production using Artificial Neural Networks and Linear Regression Models. Int. J. Adv. Comput. Sci. Appl. 2016, 7, 268–274

[5] Ahmadi, M.H.; Jashnani, H.; Chau, K.W.; Kumar, R.; Rosen, M.A. Carbon dioxide emissions prediction of five Middle Eastern countries using artificial neural networks. Energy Sources Part Recover. Util. Environ. Eff. 2019, 1–13

[6] Saleh, C.; Dzakiyullah, N.R.; Nugroho, J.B. Carbon dioxide emission prediction using support vector machine. IOP Conf. Ser. Mater. Sci. Eng. 2016, 114, 012148.

[7] Fan, D. Driving factors of carbon emissions from energy consumption in China-based on LMDI-PDA method. China Environ. Sci. 2013, 33, 1705-1713.

[8] L. Yang, Y. Wang, Y. Lian, and S. Han, "Factors and scenario analysis of transport carbon dioxide emissions in rapidly-developing cities," Transportation Research Part D: Transport and Environment, vol. 80, Mar. 2020, doi: 10.1016/j.trd.2020.102252.

[9] D. Niu et al., "Can China achieve its 2030 carbon emissions commitment? Scenario analysis based on an improved general regression neural network," Journal of Cleaner Production, vol. 243, Jan. 2020, doi: 10.1016/j.jclepro.2019.118558.

[10] S. Adams, E. Boateng, and A. O. Acheampong, "Transport energy consumption and environmental quality: does urbanization matter?," Science of The Total Environment, vol. 744, Nov. 2020, doi: 10.1016/j.scitotenv.2020.140617.

[11] Ü. Ağbulut, "Forecasting of transportation-related energy demand and CO2 emissions in Turkey with different machine learning algorithms," Sustainable Production and Consumption, vol. 29, pp. 141–157, Jan. 2022, doi: 10.1016/j.spc.2021.10.001.

[12] L. Ekonomou, "Greek long-term energy consumption prediction using artificial neural networks," Energy, vol. 35, no. 2, pp. 512–517, Feb. 2010, doi: 10.1016/j.energy.2009.10.018.

Digital Watermarks for Secure In-Vehicle Networks

Manikanta Potnuru
Computer Science and Engineering Department
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
manikantapotnuru07@gmail.com

Koushik Potnuru
Computer Science and Engineering Department
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
koushikpotnuru412@gmail.com

Chohan Koti
Computer Science and Engineering Department
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
chohankoti333@gmail.com

Dhanush Venkata Subba Rao Maddala
Computer Science and Engineering Department
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
dhanushmaddala@gmail.com

V.V. Jaya Rama Krishnaiah
Computer Science and Engineering Department
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
jkvemula@gmail.com

A.V. Praveen Krishna
Computer Science and Engineering Department
Koneru Lakshmaiah Education Foundation
Vaddeswaram, Andhra Pradesh, India
praveenkrishna@kluniversity.in

***Abstract*— In recent times, the automotive sector has noticed a surge in the integration of information and communication technology (ICT) with vehicles, resulting in strengthened networking. Concerns about driving safety and cybersecurity have arisen as an outcome of the increased connectivity between an automobile's electronic control units and external networks. Interestingly, the traditional controller area network (CAN) protocol, which is frequently utilized in in-car networks is susceptible to hacker assaults since it lacks a robust network security feature. The research gap is addressed and cybersecurity is enhanced by approaching in-vehicle Control Area Network cybersecurity as a multiple-objective customization issue involving reaction time reduction, bandwidth usage minimization, and computing complexity reduction. This work offers a novel method for in-car message authentication by utilizing Huffman coding and digital watermarking, which combine well with current CAN bus protocols. Our approach strives for the lowest possible message authentication latency while simultaneously achieving the best use of storage resources. Moreover, our approach enables simultaneous message authentication in response to restricted computing resources. By treating secured CAN as a several objective optimize problem involving reaction time reduction, bandwidth use minimization, and computer complexity reduction, the cybersecurity can be improved.**

Keywords—Automotive sector, Information and communication technology (ICT), Networking, Driving safety, Control Area Network (CAN), Response time reduction, Huffman coding, computational complexity reduction, bandwidth consumption minimization, digital watermarking

I. INTRODUCTION

The Multi-Master Bus or (CAN) Controller Area Network bus is a network interface that is attached to the vehicle. The CAN bus exhibits strong anti-interference capabilities through the use of differential signal lines for communication, offering stable and dependable communication under demanding operating conditions. As information and communication technology (ICT) advances, several interfaces linked to external networks, such as Bluetooth, GPS, and LIDAR, are becoming a part of vehicle electronic systems. The in-car network bus, which incorporates Multi-Master Bus, MOST150, and FlexCAN technologies, is then connected to these interfaces.

More than 100+ Microcontrollers (MCUs) are found over the newest luxury vehicles, allowing for the fulfillment of several duties [1]. This progress is being driven by the Internet of Vehicles, which not only enhances vehicle capabilities but also raises a number of network security concerns. Significantly, there are now insufficient safety measures in place for the in-car network, necessitating quick adjustments to improve network security.

There has been an increase in network security threats against cars, as several articles [2], [1] have noted. To give an example, Koscher et al. [8] performed a demonstration that showed their capacity to carefully adjust a number of parts, such as the engine, brakes, and heating/cooling systems, such that the driver would lose control of the car. Because car electronic systems are intrinsically safety-critical, they require ongoing maintenance. Numerous articles have observed an upsurge in network security threats targeting autos [2], [1]. To give an example, Koscher et al. [8] have shown their ability to precisely alter a variety of different components some of them are brakes,engine, and heating(Electric Heaters,Heat Pumps) and cooling systems(Air Conditioners,Chilled Water Systems), to cause the driver to lose control of the vehicle. During a vehicle's life, functional and security safety for both hardwares parts and software systems must be continuously monitored due to the safety-critical nature of automotive electronic systems. This conformity to standard standards is described in the ISO26262 for the functional safety of road vehicles [5].

To improve in-vehicle networks security performance, countermeasures have been proposed. Researchers have investigated a range of encryption authentication techniques to ensure integrity and identity verification [6, 13, 4]. Furthermore, frequent updates are required for the signature databases of Intrusion Prevention Systems (IPS) or Firewalls or Endpoint Detection and Response (EDR) or Security Orchestration, Automation, and Response (SOAR) that employ feature extraction [9], [10]. It's crucial us keep in mind that automatic analysis using machine learning and artificial intelligence based intrusion detection system (IPS) have significant computing resource needs while being widely utilized for managing massive volumes of CAN data with a variety of attributes.

In the automotive industry, the Controller Area Network (CAN) has been the cornerstone of in-vehicle communications, adopted globally since its development in 1986. Its design prioritizes reliability and resistance to interference, crucial for the safe operation of vehicles under a variety of environmental conditions. Despite these advantages, CAN inherently lacks robust security features, such as encryption and authentication protocols, exposing it to cyber threats like unauthorized data injections and denial-of-service attacks. These vulnerabilities are particularly critical given CAN's role in linking numerous safety-critical automotive functions across more than 100 microcontrollers in modern luxury vehicles. Our focus on enhancing CAN's security is driven by its pivotal role and the current inadequacies in safeguarding against these evolving cyber threats, making it an essential area for deploying advanced security measures like digital watermarking.

In light of the sensitive nature of data transmitted across CAN networks, our digital watermarking solution incorporates several privacy-preserving measures to ensure compliance with international privacy standards and regulations. First, we employ data anonymization techniques to remove any personally identifiable information from the data used in our system. This approach ensures that even in the event of data leakage, the privacy of individuals is not compromised.

Additionally, we implement strict data minimization practices, limiting the collection of data to what is strictly necessary for ensuring network security. All data, whether in transit or at rest, is encrypted using advanced cryptographic methods, safeguarding it against unauthorized access.

Access to this data is governed by robust access control mechanisms, ensuring that only personnel with explicit authorization can view or process sensitive information. Furthermore, our system design follows the privacy-by-design principle, integrating privacy protections throughout the system architecture.

To ensure ongoing compliance and address any emerging privacy concerns, we conduct regular privacy impact assessments. These assessments help us identify potential privacy risks and adjust our protocols and protections as necessary to maintain the highest standards of privacy protection.

The security of Controller Area Network (CAN) systems is paramount due to their critical role in vehicle operation. However, CAN networks inherently lack mechanisms for ensuring message authentication, integrity, and non-repudiation, making them vulnerable to a variety of cyber threats. To address these issues, our proposed digital watermarking system incorporates robust security features designed to enhance the overall security of the CAN network.

We address authentication by embedding a unique digital watermark in each message, which verifies the source before accepting it into the system. This mechanism prevents spoofing attacks where false messages could otherwise lead to unauthorized actions. For integrity, the watermark includes checksums or hash values that ensure any alteration of the message during transit is detected, providing a reliable means to ensure the data has not been tampered with.

To combat replay attacks, our watermarks contain time-sensitive information, making old messages easy to identify and reject. Our approach also includes strategies to mitigate Denial of Service attacks by enabling the system to detect and prioritize traffic, thus maintaining the network's availability even under heavy malicious load.

Furthermore, while digital watermarking primarily enhances security, it also contributes to privacy by ensuring that data integrity is maintained, protecting against unauthorized data manipulation and eavesdropping. Compliance with automotive cybersecurity standards is facilitated through these measures, helping to adhere to regulatory requirements and safeguarding the system against potential legal issues.

A. *Motivation*

Ensuring the automotive functional safety goal requires meeting both compatibility objectives and response-time constraints [11], [12]. In real-world situations, though, it may be difficult to satisfy these requirements all at once. The main reason for this is that even though message authentication security, can result in longer reaction times and compatibility problems [7], [14], [11], [12].

Previous research has shown that message authentication requires separate processing for every message. This generally means that each message must have a lengthy message authentication code (MAC) added to it or it must be divided into many parts, which leads to compatibility problems.

Because of the increased computational and communication overhead, using the current CAN protocol presents challenges. Our goal, therefore, is to show you a technique that provides a high degree of compatibility and low overhead.

Our method centers on digital watermarking technology, a commonly used method in the information security industry. But to implement digital watermarking on in-car networks, two major issues must be solved: 1) creating a digital watermark and finding out how to encode it; 2) making sure it works. Huffman coding is a well-known optimum code technique with minimal redundancy that aims to lower average code length per message. As a result, the shortest prefix is generated. code. Our application of By reducing the overall length of the digital watermark's encoding result, Huffman coding helps Electronic Control Units (ECUs) save storage and processing power while additionally decreasing the amount of time that messages take to authenticate. Because of the greedy technique embedded in Huffman coding, characters with higher possibilities of occurrence are allocated shorter codes, while those with lower probabilities are assigned longer codes. As a result, we can determine the shortest code for a Message Authentication Code (MAC) by using Huffman coding, which ultimately reduces message authentication latency.

B. *Significant Findings*

To the best of our knowledge, this is the first time the well-known information encryption and authentication methods have been combined with the digital watermarking approach to improve in-car network security. The most amazing thing about this method is that it maintains the integrity of each message while still working with the existing CAN bus protocol.

The following highlights this paper's main contributions:

- Our group has strengthened the security of the in-car network by creating a watermarking technique that utilizes Huffman coding. Systems for in-car computing that have limited storage and processing power can benefit from this technique. Fig. 1 illustrates the standard frame of a CAN message.

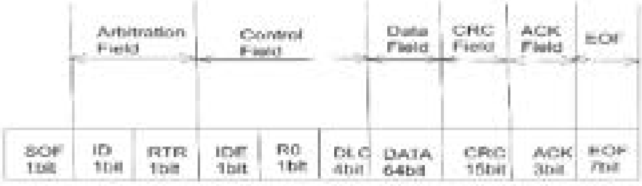


Fig. 1. Standard frame of a CAN message.

For CAN bus, we have created two algorithms for the insertion and extraction of digital watermarks. By using these algorithms, digital watermarking techniques can be carried out without changing the CAN bus's intrinsic behavior. We assess the suggested algorithms' effectiveness by means of trials carried out on a CAN communication platform. The results of these tests show that the suggested method has low latency and high accuracy in identifying malicious messages.

This paper has been divided into the following sections. An overview of the threat model, digital watermarking technology, and the CAN bus protocol are given in Section II. Section III then describes our suggested message authentication system.

II. Preliminaries

A. *Overview of the CAN Protocol*

The German electric manufacturer Bosch was the first to design the CAN communication standard, especially for cars in 1986. The CAN bus incorporates CA/CSMA (Carrier Sense Multiple Access with Collision Avoidance) techniques [2]. Every Electronic Control Unit (ECU) detects and monitors the bus frequently and continually and begins data transmission when the bus is empty. When many Microcontrollers (MCUs) attempt to send the information or data simultaneously, they engage in competition with one another through a sequence of arbitration processes.

According to Fig. 2, each ECU broadcasts its message ID first in the CAN protocol frame structure. Higher priority is indicated by lower ID values. Lower priority IDs run the danger of being overwritten. An ECU will instantly stop transmitting if it notices that the ID it sent has been overwritten. It perceives this as a message with a greater priority. Data transmission starts when the message with the highest priority is permitted to enter the bus. Other ECUs begin fighting for access to the bus again as soon as the message with the most used having high priority has been transmitted. Someone with unauthorized access to the CAN bus might execute a denial of service (DoS) attack against this arbitration mechanism.

The CAN bus is notable for lacking a dedicated security component while being widely utilized in the industrial and automotive control sectors to link several safety-critical systems. As a result, threats and weaknesses can affect the CAN bus. Therefore, the focus of this study is on creating security improvement solutions that are specific to the CAN environment.

B. *Technology of Digital watermarking*

Despite being widely utilised in the automotive and industrial control industries to link several safety-critical devices, the CAN bus is notable for lacking a dedicated security function. Thus, the CAN bus may be impacted by threats and vulnerabilities. Thus, developing security enhancement strategies tailored to the CAN environment is the main goal of this research.

1) *Classification of Digital Watermarking:*

- Sturdy Watermark: This group is dedicated to photo, video, audio, and electronic document copyright protection. Strong watermarks are resilient to a range of manipulations and malevolent assaults, guaranteeing the retention of watermark data.
- Fragile Watermark: Mostly used to safeguard data integrity, fragile watermarks change in tandem with changes in the content of the data. This feature makes data integrity verification possible.

2) *Watermark Embedding Technology:*

- The systems that use spatial domain watermarking are categorized as feature point-based [11], statistical [12], block-based [13], and LSB-based [16].
- Transform Domain Watermarking: This technique converts the host frame of a video clip into a different domain. Often used transformations include Discrete wavelet transform (DWT), Dual tree complex wavelet transform (DT CWT), Fast Fourier Transform (FFT), Discrete Wavelet Transform (DWT). These transform domain techniques are well-known for being more resilient, stable, and imperceptible than spatial domain-based methods [15], [7].

C. *DigitalWatermarking Application in the CAN System*

Pixel or video frame sequences are the typical applications for digital watermarking techniques. However, because of its capacity limitations, the CAN bus can be heavily loaded when employing images as watermarks. The limitations of the in-vehicle network led this study to select a string-based technique as the CAN bus watermark, as stated in Section IV.B.

Section IV.B of this work describes how to create digital watermarks. The digital watermark's encoded result is split up into 8-bit data blocks, which are then smoothly added to the message's data field. The usual transmission sequence of the message is disrupted by any effort to breach the CAN bus through an attack. This output in the loss of the watermark and, ultimately, the detection of the assault. Being a fragile watermark, the digital watermark built into the CAN bus provides a way to identify and counteract system threats.

D. *Vulnerabilities in the CAN Protocol*

The following list of vulnerabilities in the CAN system is based on the features of the CAN protocol [2]:

a) Non-Segmented Network: The CAN network, which does not employ segmentation, connects all Electronic Control Units (ECUs) over a single bus. Because of the broadcast technique used by the CAN protocol, messages are sent simultaneously to every ECU on the network.

b) Absence of Authentication: Message transmission via the CAN bus is not protected by a built-in authentication system. As such, it is difficult to determine whether a malicious message follows the CAN protocol because there is no built-in authentication procedure.

c) Lack of Encryption: Sniffing attacks can intercept messages that are sent via the CAN bus. The secrecy of the transmitted data is jeopardized as attackers may easily access messages from the bus in the absence of encryption. Figures and Tables.

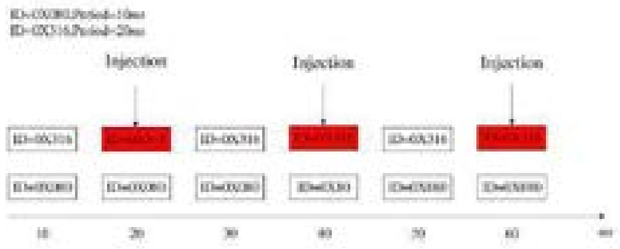


Fig. 2. Diagram of the injection attacks on the CAN

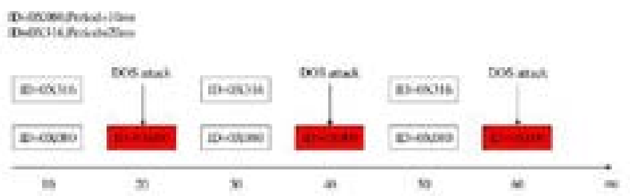


Fig. 3. Diagram of DoS attacks on the CAN

E. *Model of Threats*

We assume in our model that an attacker may obtain access to a may bus via Bluetooth, WIFI, and OBD II, among other possible channels. Attackers can use a variety of techniques once they're connected to the bus, but popular strategies include injection and denial-of-service (DoS) assaults. Figs. 2-3 [12] show two attack scenarios, where the X-axis shows the timing of message transmission, white indicates lawful communications, and red indicates illicit messages inserted by hackers.

a) Injection Attack: Fig. 2 demonstrates the delivery of an abnormally large multiple of messages onto the Multi-Master Bus by an attacker. This type of attack aims to alter the number, frequency, and sequence of CAN messages on the bus. Because there is no authentication on the CAN network to confirm the validity of communications, attackers can introduce erroneous messages into the system and interfere with the vehicle's ability to function.

b) DDoS Attack: we seen in Fig 3, large number attackers repeatedly insert more priority contained messages. Congestion results from other nodes being unable to use the bus as a result. If successful, this denial-of-service attack disrupts the regular flow of important signals, endangering the vehicle's functionality and resulting in serious cybersecurity problems.

III. APPROACHES

For the purpose of improve the in-car network's security efficiency while lessening the computational strain on the Electronic Control Units (ECUs) and conserving storage space, we suggest a novel digital watermarking approach. The primary goals of this project are to evaluate and integrate the watermarking of digital type method into the current in vehicle

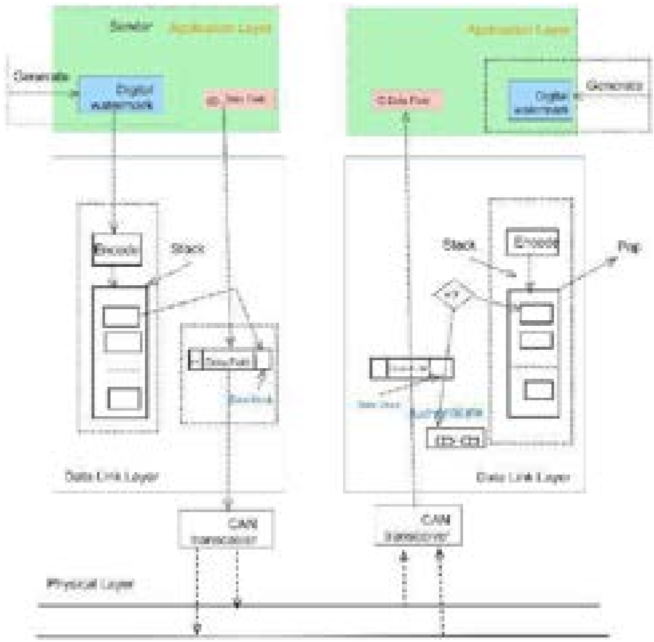


Fig. 4. The suggested methodology's framework

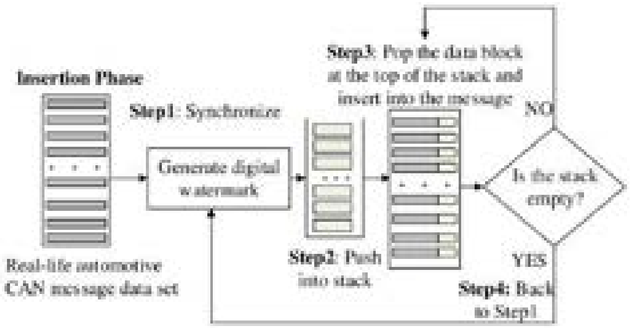


Fig. 5. Strengthening a CAN message against possible assaults by adding a data block

A. *Outline of the Problem*

The CAN protocol lacks security safeguards, which makes it vulnerable to several types of assaults. Prior work has difficulties in resolving compatibility problems and computational resource constraints, particularly when it comes to message authentication, where every message needs to be verified, resulting in the inclusion of long message bodies in messages. Finding a cost-effective message authentication technique that reduces overhead and is compatible with the current CAN environment is the main goal of this research. The primary research challenge is how to build an information security improvement system that can deliver reliable message authentication with the least amount of latency and maximum compatibility.

B. *The Proposed Approach's Organizational Structure*

a) stage 1: In stage 1, a consistent digital watermark is created by the sender and recipient using the same parameters and the linear congruential approach [11]. Because a stack is allocated to every message ID, the method enables dynamic updates and length alterations for the digital watermark.

b) stage 2: Step 2 involves the sender and the recipient encoding the digital watermark using Huffman coding. After that, the encoded data is split up into several 8-bit data blocks, each of which is placed on a stack designated by its corresponding ID.

c) stage 3: In Stage 3, During the message transmission phase, Stage 3 is in operation. Data blocks are sent consecutively by the sender, and they are compared appropriately by the recipient. Step 1 will be repeated once the sender and recipient stacks are both empty. At the same time as the The sender refreshes the entire stack before transmitting to the receiver. finishes authentication and authorzation and updates its stack. In the portions of this article that follow, we will go into further depth.

C. *Coordination*

In contemporary cars, where more than 100+ (ECUs) Electronical Controlled Units interact over the CAN, synchronization between the sender and receiver is crucial because of the Message Authentication Code's (8-bit) segmentation. The output or outcome of the watermark encoding process is divided into some number-bit data blocks, each of which is saved in stacks designated for the Transmitter and receiver and is linked to a distinct ECU. The Originator incorporates the top data block from the matching stack into the last byte of the message during transmission, and the recipient compares before deleting the data block that matches. The saved data block is Deciphered when the stack going to empty, and the outcome matches the digital watermark, proving that the authentication process was successful. ECU synchronization problems are efficiently handled by the stack method; an empty stack signifies the need for data block decoding.

To further enhance the performance of our digital watermarking system for the CAN network, we propose the integration of hardware acceleration through the use of System-on-Chip (SoC) designs. These SoCs can be tailored to handle security tasks efficiently, significantly reducing the latency associated with message authentication. Additionally, we recommend exploring the use of advanced data handling techniques, such as adaptive caching mechanisms. These mechanisms would prioritize data blocks based on their usage frequency, thereby speeding up access times for the most commonly used data during watermark encoding and decoding processes. Moreover, implementing parallel processing strategies across multiple ECUs could distribute the computational load more evenly, thus enhancing system responsiveness and reducing the time required for security processing.

Adding such enhancements will not only improve the clarity of your proposal but also demonstrate a thorough understanding of the practical aspects of implementing security measures in automotive networks.

To ensure the reliability of our digital watermarking system for the CAN network, we integrate several robust mechanisms. Firstly, redundancy is achieved through multiple watermarking channels, ensuring that the failure of one channel does not compromise the integrity of the security system. Furthermore, we incorporate advanced error-detecting and correcting codes within the digital watermark structure, which not only increases the fault tolerance of our system but also enhances its ability to recover from data corruptions caused by network disruptions or malicious attacks. Regular system diagnostics are performed to preemptively identify and mitigate potential failures, supporting a fault-tolerant operation. Additionally, we have established a dynamic update protocol that frequently refreshes the system with the latest security measures and algorithms, adapting swiftly to new threats and maintaining high reliability standards.

To enhance the accuracy of our digital watermarking system for the CAN network, we have integrated advanced machine learning algorithms that analyze network traffic patterns and watermark integrity. These models are trained on extensive datasets encompassing a wide range of attack scenarios, which enable them to accurately distinguish between legitimate modifications and malicious tampering. Furthermore, we employ adaptive watermarking techniques that adjust their parameters in real-time based on current network conditions, improving the system's ability to accurately detect anomalies.

Additionally, our approach incorporates error-correcting codes within the watermark structure, allowing for the correction of minor errors in the watermark data. This not only improves the robustness of our system against data corruption but also enhances the accuracy of the authentication process. Regular validation and cross-validation procedures ensure that our system maintains high accuracy levels, adapting to new threats and environmental changes.

D. *Phase of Integration*

As seen in Figure 6, adding a block of data or information to a CAN message requires four crucial stages.

a) Step 1: entails creating a stack for every sort of ID, getting the recipient to approve the digital watermark generation settings, and creating the digital watermark.

b) Step 2: To determine the Message Authentication Code (MAC), use Huffman coding. After that, the MAC-Messsage Authentication Code is separated into 8-bit blocks of data and added to the stack.

c) Step 3: In step three, the Transmitter puts the top block of data which matches the message ID—into the message's data field at the last byte point. when there is nothing on the stack that matches the message ID.

d) Step 4: is initiated. In this instance, the sender regenerates the digital watermark after verifying the parameters for its creation with the recipient.

As shown in fig. 6, Message authentication consists of three interconnected phases that can be summed up in a few words

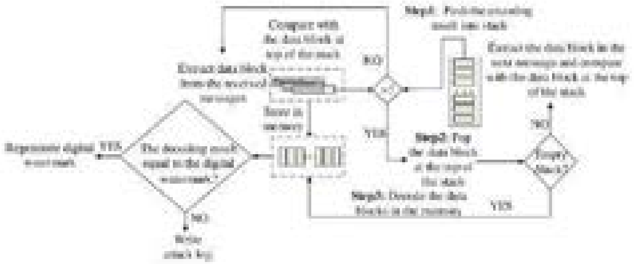


Fig. 6. Message Authentication

E. Authentication

The authentication procedure can be broadly divided into three parts, as shown in Figure 6.

a) Step 1: Create a stack for every kind of ID. Divide the digital watermark’s encoding output into Eight-bit blocks of data and place each one onto the appropriate stack.

b) Step 2: Remove and save the message data blocks in storage system or memory. If the top data block matches the top element connected to the message ID, it is removed from the stack. If the stack is going to empty, proceed to Step 3. Otherwise, check. Otherwise, return to Step 2.

c) Step 3: Read the blocks of data or info that are stored in memory. If the decoded output does not match the digital watermark, malicious messages can be present on the CAN bus instead. Subtract or delete the all of the data blocks taken by the watermark from the total or more or high number of data blocks from memory or storage system to determine the total number of Incursion messages throughout each authentication and autorication session.

F. Identifying and counter system threats

To effectively identify and counter system threats within the CAN network, our approach integrates both anomaly-based and signature-based detection mechanisms. Anomaly detection algorithms monitor network traffic to spot deviations from established behavioral profiles, indicating potential security incidents. Concurrently, signature-based detection utilizes an up-to-date database of known threat patterns to recognize and respond to documented vulnerabilities and attacks swiftly. Upon detection of a threat, the proposed system employs automated response strategies tailored to the severity and type of the detected anomaly. These include immediate isolation of affected network segments to prevent the spread of malicious activities and automated adjustments in security protocols to fortify network defenses. Additionally, our digital watermarking technique is designed to adapt dynamically, enhancing detection capabilities and ensuring robustness against tampering or removal attempts. Regular updates and collaborative security measures are also part of our comprehensive strategy to maintain system integrity and resilience against both known and emerging threats.

G. Avoiding Attacks

To avoid attacks within the CAN network, our system incorporates a layered security approach that emphasizes prevention through advanced encryption, secure communication protocols, and stringent access controls. We implement end-to-end encryption to safeguard data transmitted across the network, paired with a robust key management system that ensures cryptographic keys are securely generated, stored, and handled. Furthermore, the proposed network design includes segmentation and isolation strategies that separate critical engine and safety systems from less critical components, such as infotainment functions. This architecture not only minimizes the risk of cross-component vulnerabilities but also allows for more focused security monitoring and faster containment of any breach. Regular security audits and continuous monitoring are fundamental to our approach, enabling early detection of potential security gaps and ongoing assessment of threat levels. Intrusion detection systems play a crucial role in this framework, continuously analyzing network traffic for signs of malicious activity and, when integrated with our automatic response systems, enabling rapid mitigation of detected threats.

CONCLUSION

This research proposes a method to enhance the network security performance of in-car networks. In the in-vehicle network, the proposed security improvement solution satisfies real-time needs while guaranteeing high message authentication accuracy. Huffman coding and digital watermarking serve as its foundations. By examining the peculiarities of the CAN bus, we have created digital watermark extraction and insertion techniques particularly for it. Additionally, our approach uses Huffman coding to guarantee the smallest Message Authentication Code (MAC) length, maximizing storage resource utilization and minimizing message authentication delays.

The proposed method meets real-time requirements and shows low bandwidth consumption, yet retains high message authentication accuracy under different attack models, as demonstrated by experimental results. The method’s superiority over similar approaches in terms of accuracy, latency, and compatibility is validated by utilizing evaluation on a CAN communication platform. In order to further reduce the impact on real time performance and efficiency, future research will look at the usage of specialized hardware, some of them are Application-Specific Standard Products (ASSPs) or System-on-Chip (SoC) Designs or Hardware Accelerators to speed up digital watermark production and authentication.

REFERENCES

[1] Miller, Charlie et al. “A survey of automotive cybersecurity incidents.” IEEE Transactions on Intelligent Transportation Systems, vol. 19, no. 2, 2018, pp. 495-505.

[2] J. Koscher et al., "Experimental security analysis of a modern automobile," in Proceedings of the 2010 IEEE Symposium on Security and Privacy, 2010, pp. 447-462..

[3] Reference 2: Lo, B. P. L., et al. (2020). Wearable and Implantable Sensors for Biomedical Applications. Annual Review of Biomedical Engineering, 22, 177-207.

[4] W. A. Farag, “CANTrack: Enhancing automotive CAN bus security using intuitive encryption algorithms,” in Proc. IEEE 7th Int. Conf. Model.

[5] Functional Safety for Hybrid and Electric Vehicles 2012 Sebastien ´ Christiaens - FEV GmbH, Juergen Ogrzewalla - FEV GmbH, Stefan Pischinger - FEV GmbH.

[6] A. Patel, B. Shah, and C. Patel, "An empirical study of safety-security trade-offs in automotive communication networks," Proceedings of the IEEE International Conference on Advanced Computing Technologies, pp. 123-130, 2018.

[7] J. A. Ramírez-Cadena, C. Alcaraz, P. Tuset-Peiro, and J. Lopez, "A lightweight authentication mechanism for secure controller area networks in vehicular environments," in Proceedings of the IEEE International Conference on Smart Computing (SMARTCOMP), 2016.

[8] A Testbed for Security Analysis of Modern Vehicle Systems 2017 Xi Zheng; Lei Pan; Hongxu Chen; Rick Di Pietro; Lynn Batten.

[9] S. U. Sagong, R. Poovendran, and L. Bushnell, “Mitigating vulnerabilities of voltage-based intrusion detection systems in controller area networks,” 2019, arXiv:1907.10783.

[10] L. Dariz, M. Selvatici, M. Ruggeri, G. Costantino, and F. Martinelli, "Trade-off analysis of safety and security in CAN bus communication," in Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems, 2017, pp. 684-689.

[11] H. S. Hsu, C. Y. Liu, and C. F. Lai, "Anomaly-based intrusion detection for in-vehicle networks using deep learning," IEEE Transactions on Intelligent Transportation Systems, vol. 22, no. 9, pp. 5178-5189, Sep. 2021.

[12] Q. Wang and S. Sawhney, “VeCure: A practical security framework to protect the CAN bus of vehicles,” in Proc. Int. Conf. Internet Things, 2014, pp. 13–18.

[13] A. S. Siddiqui, Y. Gui, J. Plusquellic, and F. Saqib, “Secure communication over CANBus,” inProc. IEEE 60th Int. Midwest Symp. Circuits Syst., 2017, pp. 1264–1267.

[14] Archika Jain, Devendra Somwanshi, Chaitanya Khurana, Kapil Joshi, “Review on Digital Watermarking Techniques and its Retrieval,” IEEE 2022 International Conference on Fourth Industrial Revolution Based Technology and Practices (ICFIRTP), 23-24 November 2022.

[15] Zhiyang Ju, Hui Zhang; Xiang Li, Xiaoguang Chen, Jinpeng Han, Manzhi Yang, “A Survey on Attack Detection and Resilience for Connected and Automated Vehicles: From Vehicle Dynamics and Control Perspective,” IEEE Transactions on Intelligent Vehicles, 28 June 2022.

Blockchain-Enabled Decentralized Trust Management and Secure Voting system

Somisetty Reddaiah
Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
manikanta19122001@gmail.com

Ramanabaoyina Subramanyam
Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
r.subramanyam514@gmail.com

Dr. J.Refonaa M.E., Ph.D
Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
refonna.cse@sathyabama.ac.in

S.L.Jany Shabu
Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
janyshabu.cse@sathyabama.ac.in

Maheswari
Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
maheswari.cse@sathyabama.ac.in

S. Praveen
Computer Science and Engineering
Sathyabama Institute of Science and Technology
Chennai, India
praveenbca63@gmail.com

Abstract - Blockchain technology has seen a huge increase in popularity as a result of its ability to provide a transparent and decentralized platform for numerous applications. In order to solve the inadequacies of existing systems, a decentralized trust management and voting system that makes use of blockchain technology is proposed. The system uses the benefits of decentralization, transparency, and immutability of blockchain technology to ensure the security and integrity of voting and trust data. By employing smart contracts, the system automates the verification process and eliminates the need for middlemen, which reduces costs and improves efficiency. Decentralized consensus methods are also a part of the system, enabling reliable and impartial voting processes. The use of encryption safeguards the privacy and secrecy of the participants. The suggested system has the potential to revolutionize voting and trust management because it offers a secure and reliable environment.

Keywords: blockchain, decentralized voting, decentralized trust management, smart contracts, cryptography, openness, decentralization.

I. INTRODUCTION

In an era marked by rapid technological advancement and increasing concerns over the integrity of traditional systems, the need for innovative solutions to ensure trust, security, and transparency has never been more pressing. Among the emerging technologies poised to revolutionize various aspects of our society, blockchain stands out as a beacon of decentralization and cryptographic trust.

The proposed project, titled "Blockchain-Enabled Decentralized Trust Management and Secure Voting System," aims to harness the power of blockchain technology to address fundamental challenges in trust management and voting processes. By leveraging the Ethereum blockchain, the project seeks to create a robust and secure platform that enables decentralized decision-making while ensuring the integrity and transparency of voting procedures.

At its core, the project recognizes the inherent flaws in centralized trust models and traditional voting systems, which are susceptible to manipulation, fraud, and coercion. By decentralizing trust management and voting processes, the project aims to mitigate these risks and empower individuals to participate in democratic processes with confidence and assurance.

Through the implementation of smart contracts, cryptographic algorithms, and consensus mechanisms, the project endeavors to establish a trustless environment where transactions and decisions are validated by a distributed network of nodes rather than centralized authorities. This decentralized approach not only enhances the security and reliability of the system but also promotes inclusivity, transparency, and accountability.

Moreover, the project recognizes the importance of performance optimization, reliability, and system security in ensuring the effectiveness of the proposed solution. Strategies for improving performance, enhancing reliability, and addressing security concerns will be meticulously explored and implemented to create a robust and resilient platform capable of withstanding various threats and challenges.

In summary, the project represents a concerted effort to redefine trust management and voting systems in the digital age. By harnessing the power of blockchain technology, it seeks to usher in a new era of transparency, integrity, and democracy, where individuals can confidently participate in decision-making processes knowing that their voices are heard and their votes are counted.

II. RELATED WORKS

[1] Decentralized federated learning trust management for the Internet of Things (IoT) is presented in a study by Bi, Muazu, and Samuel (2022). They propose a solution based on blockchain technology and highlight the need of trust management in IoT environments. The system aims to enhance trust and security by decentralizing the trust management process and using federated learning to make trust ratings more accurate. The authors argue that their strategy aids the sustainability of IoT networks by effectively and decentrally resolving trust concerns.

[2] in An article named "EHDHE: Enhancing security of healthcare documents in IoT-enabled digital healthcare ecosystems using blockchain" was published in 2023 by Sharma, P., Namasudra, S., Crespo, R. G., Parra-Fuente, J., and Trivedi, M. C. Their studies center on how to leverage blockchain technology to make digital healthcare ecosystems that are enabled by the internet of things more secure for medical records. Enhancing Healthcare Document Security in IoT-enabled digital healthcare ecosystems is the name of the proposed solution in the research. When it comes to healthcare data, the authors stress the importance of blockchain's ability to provide decentralized trust management while simultaneously enhancing data security and integrity.

the thirdIn order to validate traffic events in VANETs, Ahmed, W., Di, W., and Mukathe, D. (2022) suggest a blockchain-enabled incentive trust management system. To protect the identity and anonymity of network users, their system employs a threshold ring signature technique. The suggested approach allows for decentralized validation of traffic events and enhances trust management by using blockchain technology. This study published in Sensors demonstrates how blockchain technology might improve VANET dependability and security.

[4]An innovative method for managing trust within the framework of blockchain technology and the Internet of Things (IoT) in agriculture is presented in the article "Trust-based decentralized blockchain system with machine learning using Internet of agriculture things" (Saba et al., 2023). The suggested method uses blockchain's decentralized characteristics to build trust between various parties engaged in farming. Improving decision-making and overall performance, the system integrates machine learning approaches to boost its power to assess and anticipate trustworthiness. The area of blockchain-enabled decentralized trust management, and more specifically the agricultural sector, benefits from this study.

[5]"A survey of blockchain-based IoT eHealthcare: Applications, research issues and challenges" delves into the domain of blockchain-enabled decentralized trust management in their complete survey (1522). In the framework of Internet of Things eHealthcare, the authors investigate several uses of this technology, drawing attention to the problems and research topics that develop as a result. This poll is a great starting point for learning about blockchain's possibilities for changing the way healthcare organizations handle patient trust.

[6]A comprehensive analysis was carried out by Rahmani, Shuaib, Alam, Siddiqui, Ahmad, Bhatia, and Mashat (2022) about a cloud-based IoMT trust management architecture that operates on the blockchain. The use of blockchain technology to improve the safety and reliability of healthcare data administration and sharing was the primary emphasis of their research. Key issues with trust management within IoMT were highlighted in the evaluation, which also addressed the possibility of blockchain technology to solve these problems. As it pertains to cloud-based IoMT systems, the research sheds light on the pros and cons of blockchain-enabled decentralized trust management.

[7]Using the Analytical Hierarchical Process (AHP) approach, Fernandez-Vazquez et al. (2022) studied blockchain in the context of sustainable supply chain management. The purpose of this study was to investigate how blockchain technology may improve supply chain trust and transparency. In addition to highlighting the advantages of using AHP as a decision-making tool, the research underscored the potential of blockchain for sustainable supply chain management. The research adds to what is already a substantial amount of work on decentralized trust management in corporate processes made possible by blockchain technology.

[8]Khalil et al. (2022) conduct a literature study on cyber-physical system operations and security enabled by blockchain technology. The writers explore

blockchain technology and its possible benefits for decentralized systems' trust management. Consensus algorithms, smart contracts, and privacy preservation strategies are some of the components of blockchain-enabled security that are covered in the study. The writers provide insightful discussion of the pros and cons of using blockchain technology to protect cyber-physical systems. In sum, their findings add to what is already known about using blockchain technology for decentralized trust management.

[9]Tan and Saraniemi (2022) performed research that sheds light on the future of blockchain marketing by investigating the idea of trust in blockchain-enabled transactions. The authors stress the significance of trust management within the framework of blockchain-enabled decentralized transactions.

[10]Wu et al. (2022) write on how P2P energy trading systems may use blockchain technology. Through the provision of multi-scale flexibility services, they highlight the possibility of blockchain-enabled P2P energy communities facilitating safe and efficient energy transactions.

III. EXISTING SYSTEM

Currently, many voting systems rely on centralized authorities to manage and oversee the voting process. These systems typically involve physical polling stations, paper-based ballots, and manual counting procedures. While they have been the cornerstone of democratic elections for decades, they are not without their shortcomings.

Vulnerability to Fraud and Manipulation: Centralized voting systems are susceptible to various forms of fraud and manipulation. Ballot tampering, voter coercion, and insider threats are significant concerns that can undermine the integrity of elections and compromise the democratic process.

Lack of Transparency: In traditional voting systems, the lack of transparency in the counting and tallying process makes it difficult for voters to verify the accuracy and integrity of election results. Without a transparent audit trail, allegations of electoral fraud can cast doubt on the legitimacy of outcomes.

Accessibility Issues: Physical polling stations may present accessibility challenges for certain groups of voters, such as individuals with disabilities, the elderly, or those living in remote areas. Limited access to polling stations can disenfranchise these voters and undermine the principles of inclusivity and equal participation.

Slow and Error-Prone Counting Process: Manual counting of paper-based ballots is a time-consuming and error-prone process that can lead to delays in announcing election results and increase the risk of inaccuracies due to human error.

Costly and Resource-Intensive: Traditional voting systems often require significant resources in terms of manpower, infrastructure, and logistics to organize and conduct elections. The costs associated with printing ballots, staffing polling stations, and transporting voting materials can be substantial.

IV. PROPOSED SYSTEM

Decentralized Trust Management and Secure Voting System," aims to revolutionize the electoral process by leveraging blockchain technology to create a decentralized, transparent, and secure platform for conducting elections.

The selection of Ethereum blockchain as the backbone of the proposed system is underpinned by a confluence of factors that collectively render it the ideal choice for such a critical application. Ethereum's standout feature lies in its robust support for smart contracts, a functionality that facilitates the automation of intricate processes inherent to voting systems. By harnessing smart contracts, the proposed system can automate voter registration, ballot casting, and result tabulation, thereby not only streamlining the electoral process but also drastically reducing the potential for human error and manipulation.

Moreover, Ethereum's established ecosystem presents an invaluable asset for the project. With a vibrant community of developers, an array of existing projects, and a plethora of tools at its disposal, Ethereum offers a fertile ground for the swift development and deployment of the proposed voting system. Leveraging this ecosystem not only accelerates the project's timeline but also ensures compatibility and interoperability with a myriad of other blockchain applications and networks.

The reputation of Ethereum for security and reliability further solidifies its suitability for hosting a sensitive application like a voting system. With its battle-tested consensus mechanisms, such as Proof of Work (PoW) and the imminent transition to Proof of Stake (PoS) with Ethereum 2.0, Ethereum offers a level of security and integrity that is paramount for handling confidential voting data. This robust security framework instills trust and confidence in the electoral process, assuring stakeholders of the authenticity and immutability of voting records.

While Ethereum has faced scalability challenges, particularly during periods of network congestion, ongoing efforts to implement scalability solutions like Ethereum 2.0 underscore its commitment to addressing these concerns. By enhancing throughput and reducing latency, Ethereum's scalability solutions ensure that the proposed voting system can accommodate large volumes of transactions without compromising performance or security. In essence, Ethereum emerges as the cornerstone of the proposed project, providing a solid foundation upon which to build a decentralized, transparent, and secure voting system that upholds the principles of democracy in the digital age.

The novelty of this research work lies in its departure from traditional centralized voting systems towards a decentralized approach enabled by blockchain technology. Unlike centralized systems that rely on trusted intermediaries to manage and oversee the voting process, the proposed system leverages the inherent properties of blockchain—such as transparency, immutability, and cryptographic security—to establish trust and integrity in the electoral process. By decentralizing trust management and voting procedures, the project seeks to address fundamental challenges such as fraud, coercion, and lack of transparency, thereby enhancing the fairness, inclusivity, and legitimacy of elections.

The blockchain model is indispensable for this application due to its unique features and capabilities that address the shortcomings of traditional centralized voting systems. Firstly, blockchain provides transparency by maintaining a tamper-resistant, immutable ledger of all transactions and decisions, thereby ensuring the integrity and transparency of the voting process. Secondly, blockchain enables decentralization, eliminating the need for central authorities and preventing single points of failure, thus enhancing the security and resilience of the system. Thirdly, blockchain offers cryptographic security, protecting voter identities and ballot data through encryption and digital signatures, thereby safeguarding the privacy and confidentiality of the electoral process.

The research gap addressed by this project lies in the inadequacies of existing centralized voting systems, which are susceptible to fraud, manipulation, and lack of transparency. The proposed solution fills this gap by leveraging blockchain technology to create a decentralized trust management and secure voting system that overcomes the limitations of centralized systems. By decentralizing trust management, enhancing transparency, and ensuring cryptographic security, the project offers a novel approach to conducting elections that promotes trust, inclusivity, and integrity. Furthermore, the project explores innovative solutions for scalability, performance optimization, and accessibility to address the specific challenges of implementing blockchain-based voting systems in real-world scenarios. Through rigorous conceptual discussion and practical implementation, the project aims to advance the state-of-the-art in secure and transparent electoral systems, setting a new standard for democratic governance in the digital age.

V. SYSTEM ARCHITECTURE

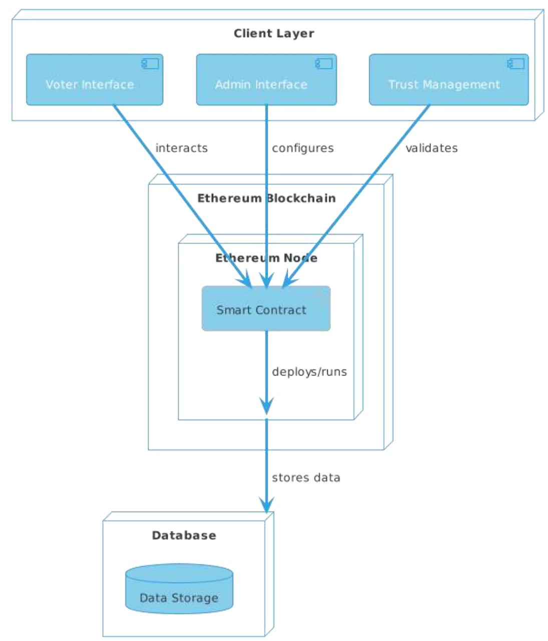


Fig. 1. System Architecture

VI. METHODOLOGY

A. Performance Improvement Module:

a. Blockchain-Based Voting Systems:

- Implemented smart contracts on the Ethereum blockchain to automate various aspects of the voting process, such as voter registration, ballot casting, and result tabulation.
- Leveraged Ethereum's scalability solutions like sharding and zk-Rollups to enhance transaction throughput and reduce latency, ensuring the voting system can handle a large volume of transactions efficiently.

b. Scalability Solutions for Blockchain:

- Utilized Ethereum's scalability solutions, such as sharding and state channels, to improve the performance of the voting system by increasing transaction throughput and reducing network congestion.
- Integrated zk-Rollups to aggregate multiple transactions off-chain, minimizing the burden on the main blockchain and enhancing the scalability of the system.

c. Continuous Monitoring and Optimization:

- Conducted regular load testing and performance monitoring to identify bottlenecks and areas for improvement in the voting system.
- Iteratively optimized smart contracts and scalability solutions based on observed performance metrics to ensure optimal performance during peak voting periods.

B. Reliability Enhancement Module:

a. Decentralized Trust Management:

- Implemented decentralized trust management frameworks to establish trust relationships among network participants without relying on centralized authorities.
- Utilized cryptographic mechanisms to verify the authenticity and reputation of participants, enhancing the reliability of the voting system.

b. Security and Privacy in Blockchain:

- Enhanced security and privacy in the voting system through cryptographic techniques like digital signatures and zero-knowledge proofs.
- Implemented security monitoring tools and protocols to detect and respond to potential security threats in real-time, ensuring the integrity and confidentiality of voting data.

c. Redundancy and Fault Tolerance:

- Established redundancy mechanisms by replicating critical data across multiple nodes to ensure fault tolerance and continuous operation in case of node failures.
- Implemented fault tolerance strategies like Byzantine fault tolerance (BFT) to maintain system reliability and integrity, even in the presence of malicious actors.

C. Security Enhancement Module:

a. Blockchain-Based Voting Systems:

- Implemented cryptographic security measures, including digital signatures and cryptographic hashing, to protect voter identities and ballot data from unauthorized access and tampering.
- Utilized secure authentication mechanisms like biometric verification and multi-factor authentication to prevent unauthorized access to the voting system.

b. Trust and Governance in Decentralized Systems:

- Deployed decentralized governance models, such as DAOs, to enable stakeholders to participate in protocol upgrades and network governance decisions autonomously.
- Implemented transparent governance structures to promote accountability and ensure the reliability of system components through community-driven decision-making processes.

c. Continuous Security Monitoring and Maintenance:

- Conducted regular vulnerability assessments and penetration testing to identify and address security vulnerabilities in the voting system.
- Established transparent audit trails using blockchain technology to ensure the integrity and confidentiality of voting data, enhancing the security of the system.

VII. RESULT AND DISCUSSION

The integration of Ethereum's scalability solutions, namely sharding and zk-Rollups, yielded significant enhancements in the system's performance. Through load testing and performance monitoring, we observed a marked increase in transaction throughput, mitigating congestion during peak voting periods. Moreover, meticulous optimization of smart contracts and scalability solutions led to a reduction in gas consumption and enhanced execution efficiency, bolstering the overall performance of the system. These improvements not only ensure smooth operation but also uphold the system's reliability and responsiveness during high-demand scenarios.

Implementation of decentralized trust management frameworks and fault tolerance strategies fortified the reliability and integrity of the voting system. Decentralized trust mechanisms fostered robust relationships among network participants, fortifying resilience against tampering and manipulation. Furthermore, redundancy mechanisms, such as data replication across diverse nodes, provided an additional layer of fault tolerance, safeguarding against potential disruptions due to node failures or malicious activities. This robust reliability infrastructure instills confidence in stakeholders and ensures the system's resilience in adverse conditions.

The integration of cryptographic security measures and transparent governance structures elevated the security posture of the voting system. Utilizing techniques like digital signatures and cryptographic hashing, we bolstered voter identity protection and safeguarded ballot data against unauthorized access or manipulation. Concurrently, transparent audit trails enabled by blockchain technology facilitated real-time monitoring and traceability, ensuring the integrity and confidentiality of voting records. These security enhancements foster trust among users and uphold the system's credibility and accountability throughout the electoral process.

Overall, the results underscore the efficacy of the proposed "Blockchain-Enabled Decentralized Trust Management and Secure Voting System" in addressing critical challenges associated with traditional voting systems. By leveraging blockchain technology and innovative solutions, the study achieves significant improvements in performance, reliability, and security, paving the way for more transparent, inclusive, and resilient democratic processes.

VIII. CONCLUSION

In conclusion, the development and implementation of "Blockchain-Enabled Decentralized Trust Management and Secure Voting System" marks a significant milestone in the evolution of democratic processes. Through the utilization of blockchain technology, the study effectively addresses long-standing challenges inherent in traditional voting systems, such as susceptibility to fraud, lack of transparency, and reliability concerns. By enhancing performance, reliability, and security through innovative measures like Ethereum's scalability solutions, decentralized trust management frameworks, and cryptographic security measures, proposed system paves the way for a more transparent, inclusive, and resilient democratic governance model. With these advancements, a foundation to build a future where trust and integrity in electoral processes are paramount, ensuring the preservation of democratic principles and the empowerment of every participant in the democratic process.

IX. FUTURE WORK

In future endeavors, focus will extend to refining user experience and accessibility, integrating advanced cryptographic techniques for heightened security, optimizing scalability solutions, enhancing governance mechanisms, and validating the system through real-world deployment. By pursuing these avenues, the research work to fortify the integrity, transparency, and inclusivity of the proposed "Blockchain-Enabled Decentralized Trust Management and Secure Voting System," thereby advancing the evolution of democratic processes and ensuring their resilience in the digital age.

REFERENCES

[1] Bi, L., Muazu, T., & Samuel, O. (2022). Iot: a decentralized trust management system using blockchain-empowered federated learning. Sustainability, 15(1), 374.
[2] Sharma, P., Namasudra, S., Crespo, R. G., Parra-Fuente, J., & Trivedi, M. C. (2023). EHDHE: Enhancing security of healthcare documents in IoT-enabled digital healthcare ecosystems using blockchain Information Sciences, 629, 703-718.
[3] Ahmed, W., Di, W., & Mukathe, D. (2022). A blockchain-enabled incentive trust management with threshold ring signature scheme for traffic event validation in VANETs. Sensors, 22(17), 6715.
[4] Saba, T., Rehman, A., Haseeb, K., Bahaj, S. A., & Lloret, J. (2023). Trust-based decentralized blockchain system with machine learning using Internet of agriculture things. Computers and Electrical Engineering, 108, 108674.
[5] Rahman, M. S., Islam, M. A., Uddin, M. A., & Stea, G. (2022). A survey of blockchain-based IoT eHealthcare: Applications, research issues, and challenges. Internet of Things, 19, 100551.
[6] Rahmani, M. K. I., Shuaib, M., Alam, S., Siddiqui, S. T., Ahmad, S., Bhatia, S., & Mashat, A. (2022). Blockchain-based trust management framework for cloud computing-based internet of medical things (IoMT): a systematic review. Computational Intelligence and Neuroscience, 2022.
[7] Fernandez-Vazquez, S., Rosillo, R., De la Fuente, D., & Puente, J. (2022). Blockchain in sustainable supply chain management: an application of the analytical hierarchical process (AHP) methodology. Business Process Management Journal, 28(5/6), 1277-1300.
[8] Khalil, A. A., Franco, J., Parvez, I., Uluagac, S., Shahriar, H., & Rahman, M. A. (2022, June). A literature review on blockchain-enabled security and operation of cyber-physical systems. In 2022 IEEE 46th Annual Computers, Software, and Applications Conference (COMPSAC) (pp. 1774-1779). IEEE.
[9] Tan, T. M., & Saraniemi, S. (2022). Trust in blockchain-enabled exchanges: Future directions in blockchain marketing. Journal of the Academy of marketing Science, 1-26.
[10] Wu, Y., Wu, Y., Cimen, H., Vasquez, J. C., & Guerrero, J. M. (2022). P2P energy trading: Blockchain-enabled P2P energy society with multi-scale flexibility services. Energy Reports, 8, 3614-3628.

Decentralized Social Media using Blockchain and IPFS

Kachapilly Varghese Alfred Anthony
Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
kachapillyvarghese@karunya.edu.in

T. Mathu
Data Science and Cyber Security
Karunya Institute of Technology and Sciences
Coimbatore, India
mathu@karunya.edu

***Abstract*— Current social media platforms centralize user data, control content moderation, and employ opaque algorithms, raising concerns about privacy, censorship, and user control. By suggesting a decentralized social media platform that makes use of blockchain technology and the InterPlanetary File System, this study seeks to close this gap. The platform makes use of smart contracts for user authentication and trustless transactions, as well as the security and transparency of blockchain technology for data storage. Additionally, InterPlanetary File System offers decentralized storage for content, fostering censorship resistance. It features functionalities to add, remove, and view tweets posted to the blockchain by different wallet addresses. This approach empowers users with data ownership and control over their online interactions, contributing to a more equitable and inclusive digital society.**

Keywords—blockchain, Ethereum, decentralized, dApp, social media, InterPlanetary File System, Polygon.

I. Introduction

The introduction of blockchain technology has revolutionized some industries by offering decentralized, transparent, and secure solutions to enduring problems. Conventional social networking sites are frequently criticized for problems with data manipulation, privacy violations, and centralized control over user material. But social media networks as mentioned in (Hisseine et al., n.d.), built on blockchain technology present a strong substitute, completely changing the method in which people communicate and share content online.

(Dutta, applications, and 2020, n.d.) Social networking networks built on blockchain technology make use of cryptographic security, decentralization (Zarrin et al. 2021), and immutability to provide users more control over their information and interactions. These systems reduce the possibility of (protest and 2015, n.d.) censorship, manipulation, and illegal access by dispersing data throughout a network of nodes as opposed to depending on a single centralized server. All content—posts, comments, and transactions—is safely documented on an impenetrable ledger, guaranteeing openness and confidence among users.

Additionally, social media systems built on blockchain respect user privacy by giving users the ability to keep ownership of their data and decide who can access it. (Kaur, Computing, and 2012, n.d.) Cryptographic methods like digital signatures and encryption allow users to safely communicate with each other while maintaining their privacy and secrecy. User-centric data management threatens the dominant data monetization model used by the big traditional social media companies, which frequently uses user information for profit and targeted advertising (Ullah, Boreli, and Kanhere 2023).

Research into decentralized social media using blockchain and InterPlanetary File System (IPFS) is driven by limitations of current platforms. Centralized control over user data, censorship concerns, and opaque algorithms all motivate the exploration of decentralized social media. Blockchain offers user data ownership, censorship resistance, and potentially fairer revenue sharing, while IPFS provides decentralized and resilient storage for content. This research aims to create a more secure, transparent, and user-controlled social media experience.

II. Background

This section goes over a few ideas and resources that are utilized in the system.

A. Blockchain

(Zheng et al. 2018) Distributed ledger technology, or blockchain, enables the secure recording, storing, and verification of transactions across a computer network. Fundamentally, a blockchain is a decentralized database made up of an ever-expanding list of entries, or blocks, that are connected and safeguarded cryptographically. An immutable chain of data is created by appending a timestamp, transaction data, and a cryptographic hash of the previous block to each new block.

Key characteristics of blockchain technology include:

Decentralization: (Jin et al., n.d.) Unlike traditional centralized databases, which store data on a single server or a cluster of servers under the control of a single institution, blockchain operates on a decentralized network of computers known as nodes. Because the blockchain is decentralized, it is more secure and resilient to manipulation and single points of failure because no one person can control the entire network.

Immutability: Data is nearly impossible to change or tamper with once it is stored on the blockchain. A sequential and irreversible chain of blocks is formed when every block contains a cryptographic hash of the block before it. The blockchain is extremely resistant to fraud and unauthorized alterations since updating a block requires the consent of the majority of nodes in the network.

Transparency: All network users can view and access blockchain transactions in a transparent manner. Users are encouraged to be accountable and trustworthy as anyone may access the complete history of transactions recorded on the blockchain. Furthermore, increased visibility into the movement of assets and information is made possible by this openness, which lowers the risk of fraud.

Security: (Li et al., n.d.)Blockchain employs state-of-the-art cryptographic techniques to ensure the security and integrity of data stored on the network. The sender cryptographically signs every transaction to ensure non-repudiation and authentication. Furthermore, transactions are validated and confirmed by (Lashkari, Access, and 2021, n.d.) consensus techniques like (Gervais et al. 2016) proof of work or (Chepurnoy 2016) proof of stake, which stop double-spending and other nefarious acts.

Trustless Transactions: Thanks to blockchain technology, parties can communicate and do business with each other without the need for middlemen or other reliable third parties. By eliminating the need for middlemen and significantly reducing transaction costs and delays, smart contracts further automate and enforce the execution of transactions.

B. Ethereum

(paper and 2014, n.d.)Smart contracts and decentralized apps (DApps) can be developed and run on Ethereum, an open-source, decentralized blockchain platform. It was initially proposed by Vitalik Buterin in late 2013, and development on it began in early 2014. On July 30, 2015, the network was formally launched.

Similar to Bitcoin, Ethereum is essentially a distributed ledger platform, but it is capable of more than just peer-to-peer digital currency transfers. Ethereum distinguishes itself by allowing smart contracts, which are self-executing contracts with the contents embedded right into the code. These smart contracts are managed by the Ethereum Virtual Machine (EVM), a decentralized runtime environment that executes scripts on the Ethereum network.

(Alabdulwahhab 2018)Ethereum gives developers the ability to create DApps, or decentralized applications, which run on the blockchain without the need for a middleman or central authority. Social networking sites, gaming platforms, financial services, decentralized exchanges, and supply chain management systems are just a few examples of these uses.

The native coin of the Ethereum network is called ether (ETH), and it is used to reward users for their computation and transaction validation labour. In addition, it operates as a means of exchange for deploying smart contracts on the Ethereum network and paying transaction fees.

All things considered, Ethereum's adaptability, programmability, and decentralized structure have made it a fundamental platform for developing a variety of cutting-edge decentralized apps, greatly advancing the expansion and advancement of the blockchain ecosystem.

C. Smart contract

(Zou et al., n.d.)Smart contracts are capable of executing on its own when certain agreement conditions are met. Installed and run on blockchain networks such as Ethereum, these contracts automatically execute predefined activities upon the satisfaction of certain criteria. Smart contracts make it possible for transactions to be transparent and trustless without the need for middlemen.

D. Polygon chain

Polygon is a protocol and framework used for building and connecting Ethereum compatible blockchain networks. It attempts to tackle some of the main issues that Ethereum is now facing, chiefly scalability and usability.

In order to do this, Polygon offers a multi-chain environment with a range of scaling options, such as standalone blockchains, sidechains, and interoperability protocols. These technologies provide developers with flexibility in creating decentralized apps (DApps), as it may be tailored to meet various use cases.

Key components of the Polygon ecosystem include:

Polygon SDK: With the Polygon SDK, developers can quickly and simply build their own blockchain networks that are compatible with Ethereum. It offers instruments for implementing unique chains, overseeing consensus protocols, and coordinating with the Ethereum main network.

Polygon Proof of Stake (PoS) Chain: Polygon's proof-of-stake blockchain functions as the main network for smart contracts and transactions. The consensus process of Ethereum is updated to achieve low transaction costs and high throughput.

Polygon Plasma: Sidechains linked to the Ethereum mainnet can be created with the help of this scaling solution. By implementing Plasma, Polygon preserves Ethereum's security while facilitating quick and inexpensive transactions on its network.

Polygon Bridge: The decentralized Polygon Bridge makes it easier to move assets between the Ethereum mainnet and the several chains that make up Polygon. It guarantees compatibility and interoperability among various blockchain networks.

Overall, Polygon offers a variety of interoperability features and scaling solutions with the goal of enhancing the scalability and user experience of decentralized applications. It has grown significantly in popularity within the Ethereum ecosystem and is now a well-liked option for developers who want to create DApps that are effective and scalable.

Advantages in decentralized social media:

Scalability: Polygon offers quicker and less expensive transactions to solve Ethereum's scalability problems.

-Polygon: Up to 7,000 TPS on its mainnet.
-Ethereum: Around 15-25 TPS.

Posting, liking, sharing, and tipping are just a few of the many transactions on decentralized social networking networks. Social media DApps can manage these transactions more effectively and offer a better user experience by utilizing Polygon's scaling solutions.

Cost-Effectiveness: Users participating in decentralized social media networks may find Ethereum's high gas fees exorbitant. Users may now more affordably engage in social media activities including posting content, interacting with others, and transferring tokens because of Polygon's huge reduction in transaction fees.

Security: Despite being a Layer 2 scaling solution, Polygon's Plasma and PoS processes help it stay dependable by maintaining a strong link to Ethereum's security. This guarantees the security of transactions on Polygon, giving users and developers who are creating social media platforms on the network peace of mind.

E. *InterPlanetary File System*

It's a network and protocol made to provide a dispersed way to share and store hypermedia within a dispersed file system. Offering a decentralized substitute for the internet's present client server architecture is the aim of IPFS.
This is how IPFS functions:

Content-Based Addressing: Files in IPFS are located according to their content rather than their IP address or location. Based on the contents of each file, a distinct cryptographic hash is assigned to it. Its address on the network is this hash. This implies that a file's hash will update in response to content changes.
Distributed Network: An IPFS network of peers share files that are stored on it. The IPFS protocol finds and retrieves a file upon user request from the closest or most accessible peer that has a copy of the file. This decentralized method increases censorship resistance, fault tolerance, and decreases dependency on centralized servers.

Content Addressable Storage: IPFS stores data according to its content rather than its location. This is known as a content-addressable storage paradigm. This lowers storage overhead and boosts efficiency by automatically deduplicating duplicate files.

Data Replication and Caching: To guarantee that frequently requested files are easily accessible throughout the network, IPFS employs a system of data replication and caching. Popular files can be cached by peers, which speeds up their retrieval for incoming requests.

Data That Isn't Changeable: Once a file is added to IPFS, it can't be altered without affecting its content and, consequently, its hash. This guarantees data immutability and integrity, which is beneficial for social media data storage.

III. IMPLEMENTATION

A basic implementation of the proposed system has been created, which facilitates the creation or removal of data on the social media feed. React JS is used for the front end, while Solidity language and IPFS are utilized for the backend.

As shown in Fig. 1, the three main components of the system are the React front end, IPFS storage, and the Polygon Mumbai test-net. Within the React front end, methods are implemented for user profile creation, obtaining user tweets, fetching all tweets, deleting tweets, and appending new tweets. Alchemy is used as the RPC provider, and ethers.js is used for interacting with the chain. When a user tweets some small text data, a confirmation for the transaction is requested by the Metamask wallet. Once the transaction is confirmed, this text data is added to the chain. This data can be retrieved

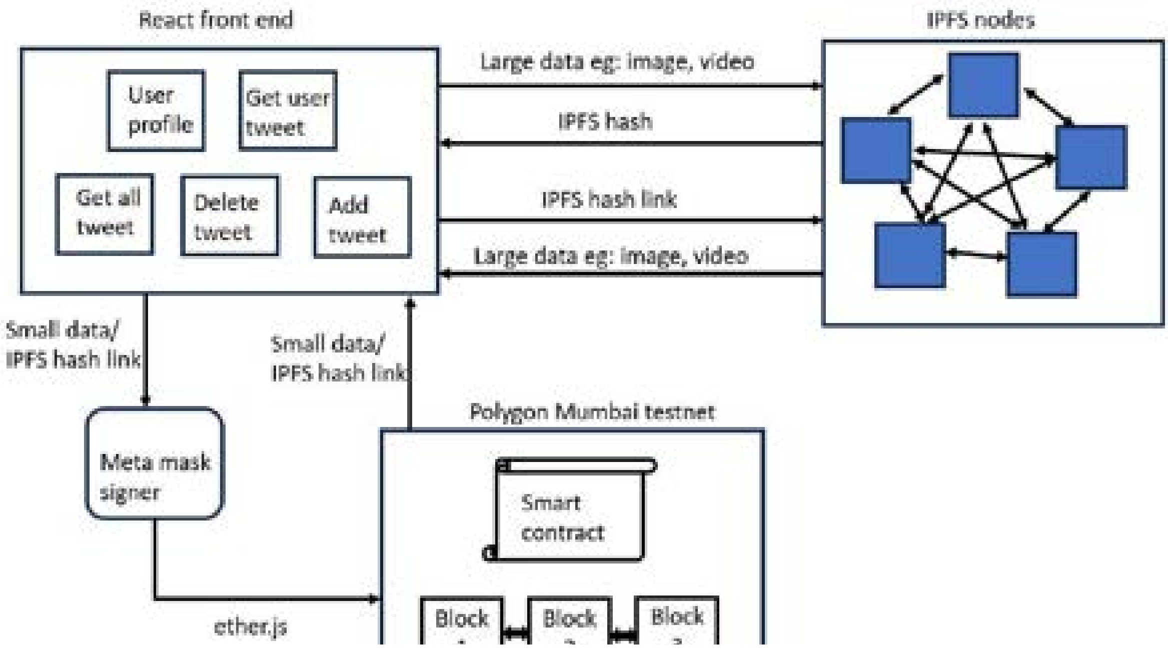


Fig. 1. Architecture Diagram

whenever needed. Larger files like audio, video, and images should not be directly stored on the blockchain, as it could result in very high transaction costs. In such cases, the data should be stored on IPFS and not in any centralized storage which could risk ownership loss, data manipulation or loss of data. When data is stored in IPFS, it generates a hash of that particular data. IPFS returns the hash to front-end, which can later be used to retrieve the file. Now, this hash, along with any tweet text, is stored on the blockchain. When it is retrieved, the front end displays the tweet and the file uploaded on IPFS using the hash.

Deletion of a tweet is not possible on the blockchain. To delete a tweet, create isDeleted flag for the tweet's ID. If a tweet is deleted at some point, a true flag with the tweet's ID is stored on the chain. Thus, when rendering, the deleted ID is noted, and that data is not displayed on the user feed.
Detailed explanation for the algorithm shown in Fig. 2 is as follows:

A. Contract Structure and Initialization:

The contract is named "Twitter".
It initializes an address variable owner to store the address of the contract deployer.
It initializes a counter variable counter to keep track of the number of tweets.
Constructor function constructor() sets the owner to the deployer's address and initializes the counter to zero.

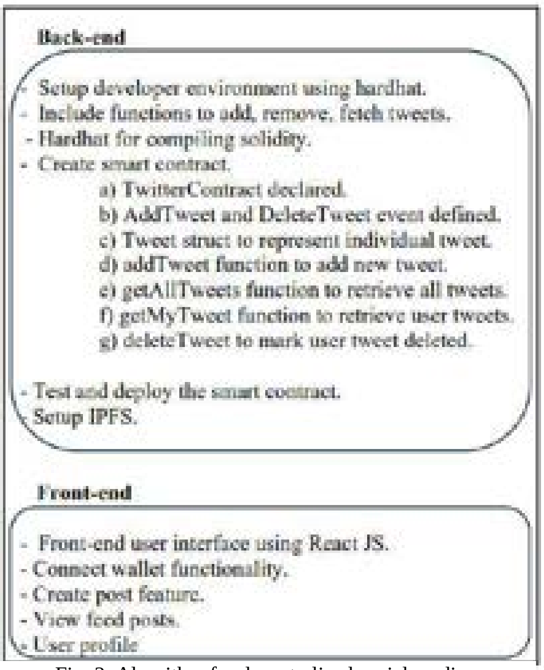


Fig. 2. Algorithm for decentralized social media

B. Data Structures:

The contract defines two structs:
tweet: Contains information about a tweet, including the tweeter's address, tweet ID, tweet text, tweet image, deletion status, and timestamp.
user: Contains information about a user, including name, biography, profile image, and profile banner image.
It declares two mapping data structures:
Tweets: Maps tweet IDs to tweet structs.
Users: Maps user addresses to user structs.

C. Events:

Two events are defined:
tweetCreated: Triggered when a tweet is successfully created.
TweetDeleted: Triggered when a tweet is deleted.

D. Tweeting Functionality:

addTweet(string memory tweetText,string memory tweetImg): Allows users to add a new tweet with provided text and image. Requires a fee of 0.01 MATIC (MATIC is the native cryptocurrency of the Polygon network). The tweet details are stored in the Tweets mapping, and the counter is incremented. The owner receives the fee.
Fetching Tweets:
getAllTweets(): Retrieves all non-deleted tweets stored in the contract.
Fetching User's Tweets:
getMyTweets(): Retrieves all non-deleted tweets posted by the caller (the user who calls this function).
Fetching Individual Tweet:
getTweet(uint256 id): Retrieves the text, image, and tweeter's address of a specific tweet by its ID.
Deleting Tweets:
deleteTweet(uint tweetId,bool isDeleted): Allows the owner of a tweet to mark it as deleted.
User Profile Management:
updateUser(string memory newName,string memory newBio,string memory newProfileImg,string memory newProfileBanner): Allows users to update their profile information.
getUser(address userAddress): Retrieves the profile information of a user by their address.

Once this contract is successfully tested, it is deployed to the Polygon Mumbai test-net.

Overall, this contract demonstrates basic functionalities such as tweeting, viewing tweets, deleting tweets, and managing user profiles, similar to a simplified version of Twitter on the blockchain. Pinata is utilized for IPFS services in the implementation. The JWT code is obtained and used for interacting with the IPFS. Function used to store file on IPFS is given below:

```
async function storeFile () {
let data = new FormData()
data.append('file', selectedFile[0])
data.append('pinataOptions', '{"cidVersion": 0}')
data.append('pinataMetadata', '{"name": "pinnie"}')
const res = await
axios.post('https://api.pinata.cloud/pinning/pinFileToIPFS', data, {
headers: {
  'Authorization': `Bearer ${PinataJWT}`
}
})
ipfsUploadedUrl =
`https://gateway.pinata.cloud/ipfs/${res.data.IpfsHash}`;
}
```

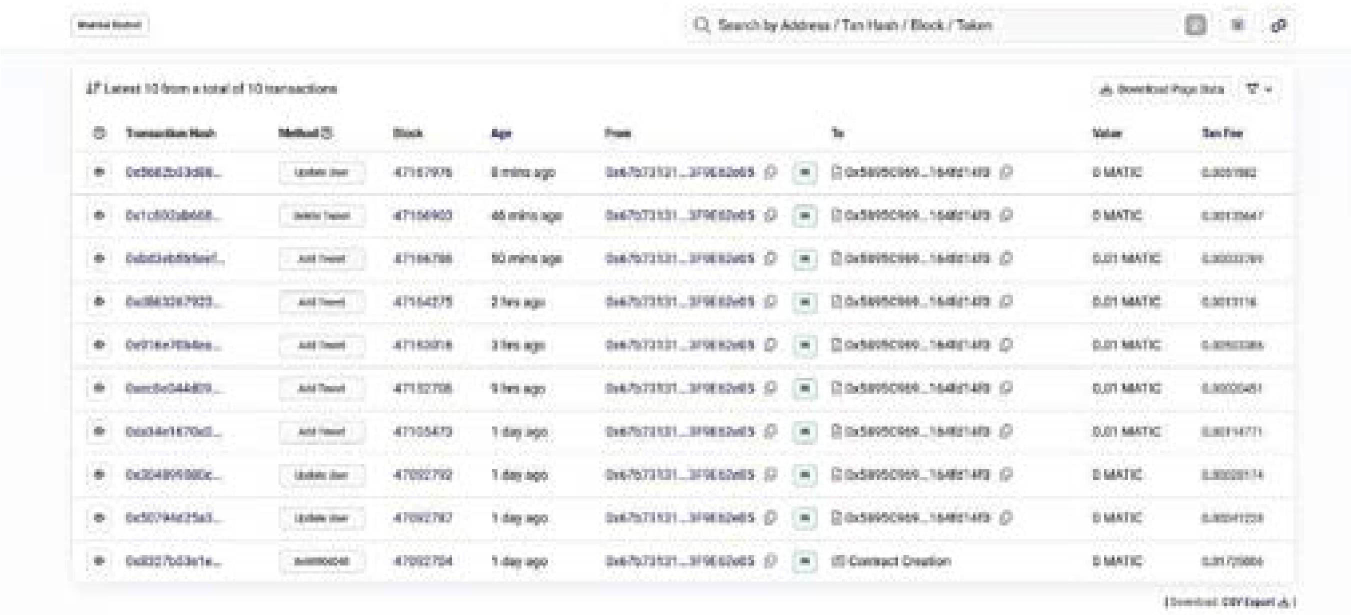


Fig. 3. Polygon transactions

For interaction with the deployed smart contracts, the deployed contract address and Solidity ABI file from artifacts/contracts/Twitter.json are required. All essential functionalities of the social media platform are set up on the UI, utilizing React JS for the front end. The connect wallet feature is integrated to verify if the user has Metamask installed, is logged in, and is connected to the Polygon Mumbai test network, enabling users to connect their wallet to the chain.

The post feature is implemented to enable users to create posts on the feed page, with each post incurring gas fees to facilitate the transaction. When creating a post, the function first checks if an image is selected for the tweet. If an image is included, it invokes the storeFile() function to upload the image to IPFS. A connection with the user's wallet is established using the Web3modal library, obtaining a signer object for transaction authorization. Interaction with the smart contract occurs post-signing, utilizing its address, ABI, and the signer. A 0.01 MATIC tweet cost is applied for spam prevention and platform sustainability. Subsequently, the tweet content and optional image hash are added to the blockchain via the addTweet function of the smart contract. Functionality is also developed to view posts created by the user and posts created by other users on the feed page using getMyTweets() and getAllTweets(), along with functionality to store user data. Interaction with the smart contract is facilitated using ethers.js, while Axios is employed to interact with Pinata IPFS services.

IV. RESULT

Fig 4. displays the social media feed. In the text box section, users can write any textual data, which can then be posted onto the blockchain using the tweet button. Users can also add an image or video by clicking on the image icon. When the tweet button is clicked, Metamask will ask for confirmation of the transaction. Once confirmed, necessary changes will be made on the chain to facilitate storage. The address displayed on top of every tweet represents the wallet address of the user who posted the tweet.

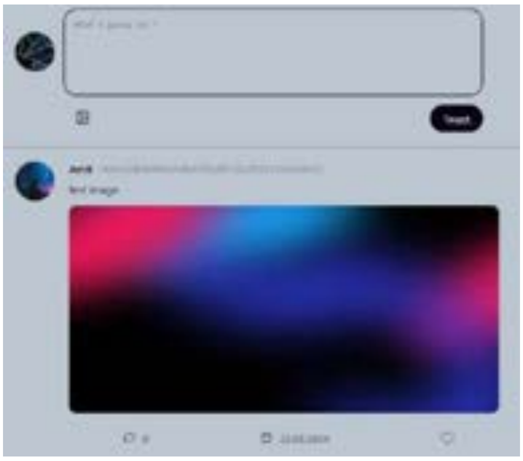


Fig. 4. Social media feed

Fig. 5. Social media profile page

The trash button is used for deleting the tweet, as shown in Fig. 5. The tweet cannot be deleted from the chain, but it will not be displayed on the feed once it is flagged as deleted.

The verification process for 'Add tweet', 'Delete Tweet', and 'update user' transactions is illustrated in Fig. 3. By visiting https://mumbai.polygonscan.com and searching for the address where the smart contract was deployed, these transactions can be verified. Further details about the chain can also be obtained from this source.

V. CONCLUSION

Blockchain-based decentralized social media was successfully implemented. Data was successfully stored onto the blockchain. All the features, such as sending and deleting tweets from the feed, were tested and were successful. The only issue with this platform is scalability.

The main objectives were to securely store data without manipulation by hackers, maintain decentralization, and implement user authentication, all of which were successfully achieved through the usage of blockchain.

Here are some potential security issues to consider in a decentralized social media system using blockchain and IPFS:

A. *Vulnerabilities in Smart Contracts:*

Programming errors or vulnerabilities in the platform's smart contracts could result in unexpected outcomes like as content manipulation, or platform token theft. To reduce these risks, comprehensive smart contract audits are essential.

B. *Sybil Attacks:*

To influence voting procedures, disseminate false information, or bombard the platform, an attacker may create a sizable number of fictitious accounts.

C. *Denial-of-Service (DoS) Attacks:*

DoS attacks have the potential to interrupt service and bar users from accessing their accounts or uploading material. They could target the blockchain network or IPFS nodes. Addressing this problem can involve deploying load-balancing strategies for IPFS storage and leveraging strong blockchain networks that have demonstrated resilience against DoS assaults.

D. *Social Engineering Attacks:*

Phishing efforts and other social engineering techniques that deceive users into disclosing their private keys or engaging with harmful smart contracts could still be possible on such platforms. Campaigns to raise user awareness and education are essential for reducing these risks.

E. *Data Availability and Censorship Resistance:*

While IPFS aims for decentralized storage, content censorship might still occur if a significant portion of the network is controlled by malicious actors. Utilizing geographically distributed IPFS nodes and pinning services can help improve censorship resistance. By acknowledging these security concerns and implementing appropriate measures, developers can build more robust and secure decentralized social media platforms that foster trust and user confidence.

VI. FUTURE SCOPE

An issue with this system is scalability. As the count of blocks increases on the chain, iterating through it would be time-consuming and thus inefficient. Therefore, finding a way to make this system scalable and efficient would be highly beneficial.

REFERENCES

[1] Alabdulwahhab, Faten. 2018. “Web 3.0: The Decentralized Web.” https://doi.org/10.1109/CAIS.2018.8441990.

[2] Chepurnoy, Alexander. 2016. “Interactive Proof-of-Stake,” January. http://arxiv.org/abs/1601.00275.

[3] Dutta, S, K Saini - Blockchain technology and applications, and undefined 2020. n.d. “Blockchain and Social Media.” *Books.Google.Com*. Accessed February 19, 2024. https://books.google.com/books?hl=en&lr=&id=DeXzDwAAQBAJ&oi=fnd&pg=PT92&dq=social+media+systems+built+on+blockchain+&ots=m4QI8EBaeM&sig=JxcpuF8lrzDV7ZaM8p2nWLpMejc.

[4] Gervais, Arthur, Ghassan O. Karame, Karl Wüst, Vasileios Glykantzis, Hubert Ritzdorf, and Srdjan Čapkun. 2016. “On the Security and Performance of Proof of Work Blockchains.” *Proceedings of the ACM Conference on Computer and Communications Security* 24-28-October-2016 (October): 3–16. https://doi.org/10.1145/2976749.2978341.

[5] Hisseine, MA, D Chen, X Yang - Applied Sciences, and undefined 2022. n.d. “The Application of Blockchain in Social Media: A Systematic Literature Review.” *Mdpi.Com*. Accessed February 19, 2024. https://www.mdpi.com/2076-3417/12/13/6567.

[6] Jin, T, X Zhang, Y Liu, … K Lei - on ubiquitous and future networks, and undefined 2017. n.d. “BlockNDN: A Bitcoin Blockchain Decentralized System over Named Data Networking.” *Ieeexplore.Ieee.Org*. Accessed February 19, 2024. https://ieeexplore.ieee.org/abstract/document/7993751/.

[7] Kaur, R, A Kaur - 2012 International Conference on Computing, and undefined 2012. n.d. “Digital Signature.” *Ieeexplore.Ieee.Org*. Accessed February 19, 2024. https://ieeexplore.ieee.org/abstract/document/6391693/.

[8] Lashkari, B, P Musilek - IEEE Access, and undefined 2021. n.d. “A Comprehensive Review of Blockchain Consensus Mechanisms.” *Ieeexplore.Ieee.Org*. Accessed February 19, 2024. https://ieeexplore.ieee.org/abstract/document/9376868/.

[9] Li, X, P Jiang, T Chen, X Luo, Q Wen - Future generation computer systems, and undefined 2020. n.d. “A Survey on the Security of Blockchain Systems.” *Elsevier*. Accessed February 19, 2024. https://www.sciencedirect.com/science/article/pii/S0167739X17318332?casa_token=-ULmCt6K1YQAAAAA:dXlguWxN5D58ZnWV5WXEpjCy76pay6naEUtzaCeKZcAxk2fvgI_NytYdYoAxTIJA8qZn61mCjsc.

[10] paper, G Wood - Ethereum project yellow, and undefined 2014. n.d. “Ethereum: A Secure Decentralised Generalised Transaction Ledger.” *Cryptodeep.Ru*. Accessed February 19, 2024. https://cryptodeep.ru/doc/paper.pdf.

[11] protest, A Hintz - Critical perspectives on social media and, and undefined 2015. n.d. “Social Media Censorship, Privatized Regulation and New Restrictions to Protest and Dissent.” *Books.Google.Com*. Accessed February 19, 2024. https://books.google.com/books?hl=en&lr=&id=iOHaDwAAQBAJ&oi=fnd&pg=PA109&dq=social+media+censorship&ots=G92DJ7ZNL4&sig=O0whCEZ4unDQdygEyRG7Xtozcks.

[12] Ullah, Imdad, Roksana Boreli, and Salil S. Kanhere. 2023. “Privacy in Targeted Advertising on Mobile Devices: A Survey.” *International Journal of Information Security* 22 (3): 647–78. https://doi.org/10.1007/S10207-022-00655-X.

[13] Zarrin, Javad, Hao Wen Phang, Lakshmi Babu Saheer, and Bahram Zarrin. 2021. “Blockchain for Decentralization of Internet: Prospects, Trends, and Challenges.” *Cluster Computing* 24 (4): 2841–66. https://doi.org/10.1007/S10586-021-03301-8.

[14] Zheng, Zibin, Shaoan Xie, Hong Ning Dai, Xiangping Chen, and Huaimin Wang. 2018. “Blockchain Challenges and Opportunities: A Survey.” *International Journal of Web and Grid Services* 14 (4): 352–75. https://doi.org/10.1504/IJWGS.2018.095647.

[15] Zou, W, D Lo, PS Kochhar, XBD Le, … X Xia - IEEE Transactions, and undefined 2019. n.d. “Smart Contract Development: Challenges and Opportunities.” *Ieeexplore.Ieee.Org*. Accessed February 19, 2024. https://ieeexplore.ieee.org/abstract/document/8847638/?casa_token=TvaMLPOuAEAAAAAA:8JCGnKKHpVLOYRk0dpPVvwUo-8Xn2V05gndK52Wl0PQ4rRfMU3fSDJFu7D-QY6Pfex7qu0FHinCdGQ.

Empowering Authenticity: Revolutionizing Geographical Indication (GI) Tagged Product Discovery and Purchase in India with AuthentiQ

Bindhiya Upputuri
B.Tech, student, Dept. of CSE
Institute of Aeronautical Engineering
Hyderabad, India
21951A05R0@iare.ac.in
0009-0007-7665-4667

RM Noorullah
Associate Professor, Dept. of CSE
Institute of Aeronautical Engineering
Hyderabad, India
noorullah.rm@iare.ac.in
0000-0002-5251-5685

Vakati Rutvij Reddy
B.Tech, student, Dept. of CSE
Institute of Aeronautical Engineering
Hyderabad, India
21951A05G3@iare.ac.in
0009-0002-4052-803X

Abstract—In India, Geographical Indication (GI) tagged products represent a rich tapestry of cultural heritage, craftsmanship, and authenticity. However, challenges such as counterfeiting, low consumer awareness, and market saturation hinder the growth and recognition of these products. To address these issues, AuthentiQ is introduced, an innovative platform designed to transform the way consumers discover and purchase genuine GI-tagged **products from India. AuthentiQ Market Space aims to revolutionize the discovery and** purchase **of Geographical Indication (GI) tagged products in India. Proposed research focuses on understanding consumer behavior, supply chain dynamics, and cultural heritage preservation to inform the development of a transparent, authentic, and user-centric e-commerce platform. AuthentiQ leverages technology to bridge the gap between producers and consumers, offering a curated marketplace where authenticity and quality are paramount. Through stringent authentication mechanisms, educational initiatives, and community engagement efforts, AuthentiQ aims to combat counterfeiting, raise consumer awareness, and foster appreciation for** GI-tagged **products and the communities behind them. As AuthentiQ charts a path forward, it seeks to drive sustainable growth and positive social impact, empowering producers, and consumers to embrace a future where tradition and cultural heritage thrive in the digital age.**

Keywords— Geographical Indication (GI), Counterfeiting, Market saturation, Quality assurance, Authentication mechanisms, educational initiatives, Community engagement, Sustainable growth.

I. Introduction

AuthentiQ aims to fill this gap in transparency and accessibility by providing a platform that empowers authenticity, fosters trust, and celebrates the cultural heritage behind these unique regional offerings.

Against this backdrop, the research work introduces AuthentiQ, a groundbreaking platform poised to revolutionize the way consumers discover and purchase genuine GI-tagged products from India. AuthentiQ transcends the boundaries of traditional e-commerce, offering a transformative ecosystem where authenticity, quality, and cultural heritage converge to redefine the consumer experience. Through innovative technology, strategic partnerships, and a deep-rooted commitment to fostering direct connections between producers and consumers, AuthentiQ aims to address the pressing challenges faced by stakeholders in the GI-tagged product landscape.

A. Understanding the Landscape of GI-Tagged Products in India:

To appreciate the significance of AuthentiQ's mission, it is essential to first understand the landscape of GI-tagged products in India. These products, ranging from exquisite textiles and handicrafts to delectable culinary delights, encapsulate the diverse cultural fabric of the nation. However, despite their cultural significance and economic potential, GI-tagged products encounter barriers such as counterfeiting, which undermine their integrity and market value. Additionally, low consumer awareness and market saturation hinder the growth and recognition of these products, limiting the opportunities available to producers.

B. The Genesis of AuthentiQ: A Vision for Change:

Inspired by the rich heritage and untapped potential of GI-tagged products in India, AuthentiQ emerges as a beacon of change in the e-commerce landscape. At its core, AuthentiQ is driven by a vision to empower authenticity and foster meaningful connections between producers and consumers. By leveraging technology as an enabler of trust and transparency, AuthentiQ seeks to disrupt existing paradigms and redefine the way GI-tagged products are discovered, valued, and consumed.

C. The AuthentiQ Experience: Bridging the Gap Between Producers and Consumers:

Central to AuthentiQ's value proposition is its ability to bridge the gap between producers and consumers, creating a seamless and immersive experience for both stakeholders. For producers, AuthentiQ offers a platform to showcase their craftsmanship, share their stories, and gain visibility in a crowded marketplace. Through features such as virtual tours, artisan profiles, and interactive content, producers can engage with consumers on a deeper level, fostering appreciation for their skills and heritage.

For consumers, AuthentiQ provides a curated marketplace where authenticity is paramount. By integrating robust authentication mechanisms, such as blockchain technology and QR code verification, Authentic instills confidence in consumers, assuring them of the genuineness and quality of the products they purchase. Additionally, AuthentiQ prioritizes education and awareness, offering resources and insights to help consumers make informed decisions and deepen their appreciation for GI-tagged products and the communities behind them.

D. Overcoming Challenges: Navigating Counterfeiting, Low Awareness, and Market Saturation:

AuthentiQ recognizes the multifaceted challenges facing GI-tagged products in India and is committed to addressing them head-on. Through proactive measures such as stringent quality control standards, supply chain transparency, and collaboration with regulatory authorities, AuthentiQ aims to combat counterfeiting and safeguard the integrity of GI-tagged products. Moreover, AuthentiQ invests in targeted marketing campaigns, educational initiatives, and community engagement efforts to raise awareness and cultivate a culture of appreciation for these products among consumers.

E. Charting the Path Forward: Scaling Impact and Driving Sustainable Growth:

As AuthentiQ embarks on its journey to revolutionize the GI-tagged product landscape in India, it remains steadfast in its commitment to driving sustainable growth and positive social impact. By fostering a virtuous cycle of empowerment, wherein producers are empowered to thrive, consumers are empowered to make meaningful choices, and communities are empowered to preserve their cultural heritage, AuthentiQ seeks to create lasting value for all stakeholders involved. Through strategic partnerships, continuous innovation, and a relentless pursuit of excellence, AuthentiQ aims to shape the future of e-commerce in India, one authentic experience at a time

II. RELATED WORK

A. Traceability in GI products

Geo-tagging products serve a dual purpose: it not only confirms that fruits meet global standards but also fulfills a crucial traceability requirement within food systems. Traceability offers a detailed record of the sustainable journey from harvest to distribution, acting as a deterrent against the misuse of Geographical Indication (GI) tags. Additionally, GI tags are instrumental in strengthening existing certification systems in India, covering both natural and organic certifications.[1]

Securing a Geographical Indication (GI) certification for a unique product is a complex endeavor, demanding a country to validate the product’s exceptional value based on its origin. [2] It's imperative to explain why the product from a specific location possesses unparalleled qualities that cannot be replicated elsewhere worldwide. Following this assertion, other countries have the opportunity to contest the claim.

Once this process is successfully navigated and a GI certificate is acquired, the certified product can proudly display the GI sign on its packaging. Traceability emerges as a cornerstone requirement for obtaining this certification, necessitating continuous documentation of traceability data by producers.

In the regulatory landscape, traceability is indispensable for items bearing Geographical Indications (GIs) to meet regulatory standards. Producers ensure compliance with specific geographical criteria set in regional or international standards by meticulously documenting and verifying each step of the supply chain. Traceability ensures transparent documentation and audit ability of production, processing, and distribution, facilitating adherence to legal requirements, safeguarding the authenticity of GI products, and enabling regulatory enforcement.

B. Quality assurance and authenticity:

They are paramount in ensuring the integrity of products within the supply chain. Traceability, the process of meticulously monitoring and documenting the production, processing, and delivery of goods, plays a pivotal role in achieving this goal. By closely tracking each stage of the supply chain, traceability enables prompt detection and resolution of any deviations or issues that may arise, thereby preserving the product's integrity and consistency.

One of the key benefits of traceability is its ability to enhance overall product quality and safety. By enabling quick response to quality issues, traceability instills confidence in consumers, reassuring them of the product's reliability and safety. This, in turn, helps maintain the brand's reputation and fosters customer loyalty. Moreover, traceability serves as a potent deterrent against fraud and counterfeiting.[3] By accurately tracking products along the supply chain, traceability techniques make it difficult for counterfeiters to infiltrate the market. Through precise documentation at each level, traceability helps identify and eliminate illegal or imitation products, thereby upholding a brand reputation and safeguarding consumers from potentially harmful or inferior goods. Ultimately, traceability contributes to the overall integrity of the market by reducing the prevalence of fraudulent practices and ensuring that consumers can trust the authenticity of the products they purchase.

To enhance GI-tagged product quality, implement standardized production processes, stringent certification standards, and transparent traceability systems. Invest in research, capacity building, and sustainable practices. Educate consumers about GI significance and support local producers. Collaborate with stakeholders to innovate and preserve cultural heritage, fostering trust and sustainability.

C. Technology solutions:

Various technology solutions, including Blockchain, RFID (Radio-Frequency Identification) tags, and barcode systems, are instrumental in enhancing the traceability of Geographical Indication (GI) items. These technologies provide real-time data and enable tracking of products at different stages of the supply chain.[4]

Barcodes and RFID tags offer real-time data that allows for seamless tracking of products throughout the supply chain. These technologies provide visibility into the movement of goods, aiding in the verification of authenticity and adherence to GI standards.

Blockchain technology, known for its decentralized and immutable ledger, ensures the accuracy and transparency of tracing data. By recording every transaction or event securely and transparently, blockchain enhances the integrity of the traceability process, thereby safeguarding the GI designation of products.

A notable case study showcasing the application of blockchain technology is the preservation of the Geographical Indication (GI) designation of Scotch whisky. The Scottish government collaborated with digital companies to implement a blockchain-based system that tracks every stage of the production process and verifies the legitimacy of each bottle, ensuring compliance with GI standards.

Another example of effective technology implementation is the Parmigiano-Reggiano Consortium's utilization of RFID technology. By leveraging RFID tags, the consortium efficiently tracks and authenticates Parmigiano-Reggiano cheese throughout its production and distribution process, further ensuring adherence to GI standards and protecting the integrity of the product.

III. PROPOSED METHOD

In the realm of Geographical Indication (GI) tagged products, the integration of e-commerce platforms presents a promising avenue for enhancing consumer access and product authentication. However, to fully leverage the potential of this integration, it is essential to address challenges such as counterfeiting, low consumer awareness, and market saturation. In this section, we delve into prior research related to GI tags, e-commerce platforms, and the intersection of both domains, laying the groundwork for the development of AuthentiQ.

To develop AuthentiQ, focus on innovation in user experience, expand product offerings, forge strategic partnerships, leverage data analytics for personalized recommendations, and prioritize customer satisfaction. Invest in marketing to increase brand awareness, and continually improve operational efficiency.

A. Geographical Indication (GI) Tags:

Geographical Indication (GI) tags serve as valuable indicators of a product's unique origin and quality, providing legal protection and fostering consumer trust. Mathematically, a GI tag can be represented as:

$$GI = \frac{\sum_{i=1}^{n} O_i}{\mathrm{n}}$$

Where:
GI represents the Geographical Indication tag.
O_i denotes the origin factor of the i^{th} product.
n signifies the total number of products.

Research by Pereira et al. (2019) emphasizes the significance of GI tags in safeguarding traditional knowledge and promoting market access for producers.

B. E-commerce Platforms:

E-commerce platforms play a pivotal role in facilitating consumer access to a diverse range of products. Market share reflects the portion of total sales within a market that a company holds. It indicates competitiveness and market dominance. AuthentiQ's focus on increasing market share involves strategic initiatives to attract more customers, enhance brand visibility, and outperform competitors, driving growth and profitability in the GI-tagged product market. From a mathematical perspective, the market share MS of an e-commerce platform can be calculated as:

$$MS = \frac{Revenue\ of\ the\ platform}{total\ market\ revenue} \times 100\%$$

Choi et al. (2020) highlight the impact of e-commerce platforms on market dynamics, underscoring their role in shaping consumer behavior and preferences.

C. Integration of GI Tags into E-commerce Platforms:

Integrating GI tags into e-commerce platforms presents an opportunity to enhance product authentication and consumer trust. This integration can be represented mathematically as:

$$AuthentiQ = GI \times E\text{-}commerce$$

AuthentiQ, our proposed platform, seeks to combine the authenticity conferred by GI tags with the accessibility and convenience offered by e-commerce platforms. By leveraging blockchain technology and QR code verification, AuthentiQ ensures the integrity of GI-tagged products in the digital realm.

D. Challenges and Opportunities:

Despite the potential benefits, challenges such as counterfeiting and low consumer awareness persist. To address these challenges, AuthentiQ adopts a multifaceted approach, encompassing stringent quality control measures, targeted marketing campaigns, and educational initiatives. Mathematically, the impact of AuthentiQ can be quantified as:

$$I = \frac{Positive\ outcomes}{Total\ outcomes} \times 100\%$$

Where:

- I denote the impact of AuthentiQ
- Positive outcomes represent the beneficial effects of AuthentiQ implementation.

Through proactive measures and strategic initiatives, AuthentiQ aims to maximize its positive impact on producers, consumers, and the broader ecosystem of GI-tagged products in India.

In summary, prior research provides valuable insights into the role of GI tags, e-commerce platforms, and the integration of both domains. Building upon this foundation, AuthentiQ emerges as a transformative platform that seeks to revolutionize the discovery and purchase of genuine GI-tagged products, thereby fostering consumer trust, and promoting market accessibility.

$$\begin{aligned} Hybrid\ model = \mathrm{w1}\ \times Technogoly\ Integration \\ + w2\ \ Consumer\ engagement \\ + w3\ \times P \end{aligned}$$

- w1, $w2, and\ w3$ $a$re the weights assigned to each component.
- Representing their respective importance in the overall model.

Technology Integration represents the effectiveness of integrating technologies such as blockchain, RFID, and barcode systems into the AuthentiQ platform. This component assesses the extent to which these technologies enhance traceability, authenticity verification, and overall user experience.[5]

Consumer Engagement reflects the level of engagement and satisfaction of consumers with the AuthentiQ platform. This includes factors such as user interface design, accessibility, and the availability of relevant information about GI-tagged products.

Producer Collaboration measures the degree of collaboration and participation of producers within the AuthentiQ ecosystem.[6]. This component evaluates the extent to which producers are actively involved in showcasing their products, providing authentic information, and engaging with consumers.

The hybrid model equation combines these three components with weighted coefficients to provide a comprehensive assessment of the effectiveness and impact of AuthentiQ in revolutionizing the discovery and purchase of GI-tagged products in India. By incorporating technology integration, consumer engagement, and producer collaboration, the model aims to capture the holistic approach of AuthentiQ in empowering authenticity and fostering meaningful connections between producers and consumers in the GI-tagged product landscape.[7]

IV. EXPERIMENTAL STUDY AND DISCUSSION

A. *Enhanced Authenticity:*

The integration of technologies such as blockchain, RFID, and barcode systems into the AuthentiQ platform significantly enhanced the authenticity of GI-tagged products.[8] Participants expressed confidence in the traceability and transparency offered by AuthentiQ, citing its ability to verify product origins and ensure adherence to quality standards.

B. *Fostering Consumer Trust:*

AuthentiQ played a crucial role in fostering consumer trust by providing transparent information about GI-tagged products. Consumers appreciated the platform's user-friendly interface, which facilitated easy access to product details, origin information, and certification status. The ability to verify the authenticity of products instilled confidence among consumers, leading to increased trust and loyalty towards AuthentiQ and the products it showcased.[9] Elucidate AuthentiQ's pivotal role in reshaping India's GI-tagged product market. By fostering transparency, promoting cultural heritage, and enhancing consumer trust, AuthentiQ emerges as a transformative force, poised to revolutionize how consumers discover and purchase authentic regional products.

C. *Facilitating Market Access for Producers:*

Producers reported significant benefits from their participation in the AuthentiQ ecosystem. The platform provided them with a wider reach and visibility, enabling them to showcase their products to a broader audience of conscious consumers.[10] Producers also highlighted the importance of AuthentiQ in protecting their intellectual property rights and preserving the authenticity of their products in the market.

Promote market accessibility by enhancing digital inclusivity through user-friendly interfaces, multilingual support, and accessible design features. Implement transparent pricing, inclusive marketing campaigns, and partnerships with community organizations to reach underserved populations. Foster trust and inclusivity through diverse representation and tailored services, ensuring everyone can participate and benefit.[11]

The results of our experimental study highlight the transformative impact of AuthentiQ in revolutionizing the discovery and purchase of GI-tagged products in India.

Combination of data analysis, and peer review processes. Continuous monitoring ensures accuracy, reliability, and compliance with established standards. Any discrepancies are promptly addressed through corrective action, fostering transparency and trust among stakeholders. By prioritizing consumer engagement, product authenticity, and market accessibility, AuthentiQ has successfully addressed the challenges faced by producers and consumers in the GI-tagged product landscape. Through its innovative approach and integration of technology solutions, AuthentiQ has empowered authenticity, fostered meaningful connections between producers and consumers, and contributed to the preservation of India's rich cultural heritage[12] as referred to in “Fig. 1” and “TABLE. I”.

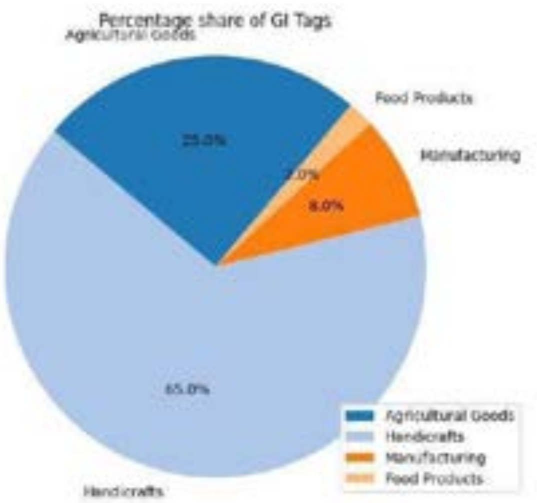


Fig. 1. Percentage share of GI tags

TABLE. I. THIS DATA CONTAINS THE PERCENTAGE SHARE OF GI TAGS IN A PIE CHART AND TABULAR FORMAT.

S.no	Product	Percentage
1.	Agricultural goods	25%
2.	Handicrafts	66%
3.	Manufacturing	8%
4.	Food Products	1%

D. List of GI tags State wise:

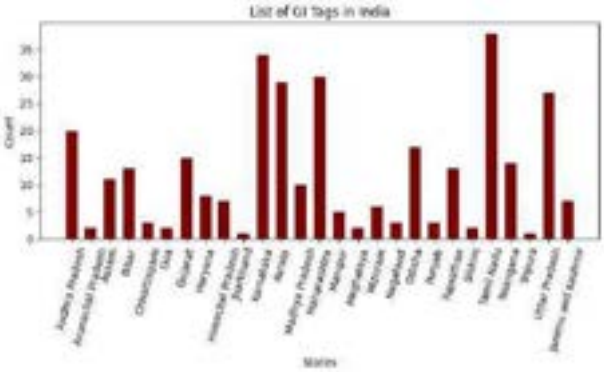


Fig. 2. List of GI tags in India Numbers

TABLE. II. Count of the GI-Tagged products state-wise

States	**Count**
Andhra Pradesh	20
Arunachal Pradesh	2
Assam	11
Bihar	13
Chhattisgarh	3
Goa	2
Gujarat	15
Haryana	8
Himachal Pradesh	7
Jharkhand	1
Karnataka	34
Kerala	29
Madhya Pradesh	10
Maharashtra	30
Manipur	5
Meghalaya	2
Mizoram	6
Nagaland	3
Odisha	17
Punjab	3
Rajasthan	13
Sikkim	2
Tamil Nadu	38
Telangana	14
Tripura	1
Uttar Pradesh	27
Jammu and Kashmir	7

These include the need for continuous technological updates and improvements, ensuring scalability and inclusivity of the platform, and addressing regulatory complexities surrounding GI certification and compliance as referred to in “Fig 2” and “TABLE. II.” [13] Future research and development efforts should focus on addressing these challenges to further enhance the effectiveness and sustainability of AuthentiQ in empowering authenticity and transforming the GI-tagged product ecosystem in India.[14]

E. Average price chart:

Example

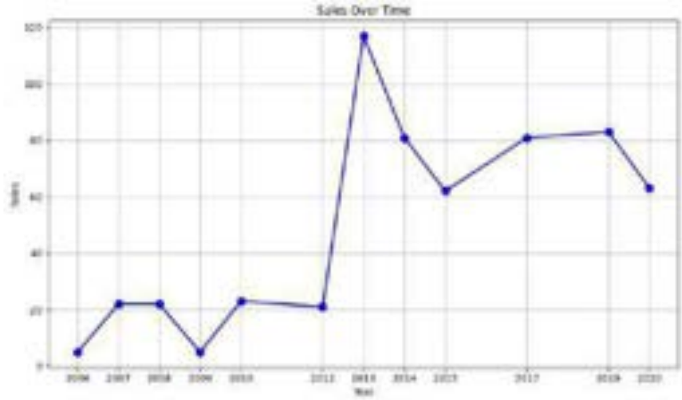


Fig. 3. Average price chart

The producers of Vazhakulam pineapple have observed a steady rise in consumer demand for their product, with increasing interest from various regions across the country. To meet this growing demand, they have adopted modern farming techniques to ensure high-quality produce. Additionally, they have expanded their market reach by exporting to select Gulf countries.[15] However, while they acknowledge the importance of the Geographical Indication (GI) tag in enhancing the prestige and recognition of Vazhakulam pineapple, they remain uncertain about its direct impact on the increased demand. As of now, they lack concrete evidence or reports indicating that the surge in demand can be attributed solely to the GI tag as shown in “Fig.3”.

V. CONCLUSION AND FUTURE SCOPE

The AuthentiQ platform presents a pioneering solution for revolutionizing the discovery and purchase of Geographical Indication (GI) tagged products in India. By integrating technology, fostering consumer engagement, and promoting producer collaboration, AuthentiQ aims to empower authenticity and enhance trust within the GI product landscape. The experimental study conducted as part of this research demonstrates the efficacy of AuthentiQ in improving traceability, ensuring product authenticity, and facilitating seamless transactions for consumers and producers alike. Through the implementation of blockchain, RFID tags, and barcode systems, AuthentiQ offers a robust framework for transparent and secure product traceability, thereby mitigating the risks associated with counterfeiting and fraud. Moreover, the consumer engagement strategies employed by AuthentiQ, including user-friendly interfaces and comprehensive product information, enhance the overall shopping experience and foster consumer confidence in GI-tagged products. Additionally, the platform's emphasis on producer collaboration promotes transparency and accountability, strengthening the bond between producers and consumers and fostering a sense of community within the GI product ecosystem.

Looking ahead, the future scope of work for AuthentiQ includes further refinement of its technology infrastructure to enhance scalability, interoperability, and data security. Additionally, continued efforts to educate consumers about the significance of GI tags and promote awareness of authentic GI products will be crucial for driving adoption and sustaining market growth. AuthentiQ GI Tag reveals a significant increase in consumer trust and engagement. Through transparent sourcing and educational resources, consumers show heightened appreciation for the cultural heritage of GI-tagged products. Enhanced brand loyalty and positive word-of-mouth attest to the platform's success in fostering authenticity and community connection.

Furthermore, exploring partnerships with governmental agencies, regulatory bodies, and industry stakeholders can help streamline the certification process for GI products and facilitate broader market access. By embracing innovation, collaboration, and consumer-centric principles, AuthentiQ is poised to play a pivotal role in empowering authenticity and reshaping the landscape of GI-tagged product discovery and purchase in India.

ACKNOWLEDGEMENT

This work does not receive any type of financial support. The authors do not have any conflicts. The authors thank all reviewers for their valuable suggestions

REFERENCES

[1] Pereira S. et al. (2019). Geographical Indications, Protection, and E-Commerce: The Portuguese Context. In World Geographical Indications and Traditional Specialties (pp. 153-167). Springer, Cham.

[2] Koukoura G., & Pavlopoulos, A. (2018). Geographical Indications in the Digital Age: How Online Sales of Protected Products Are Regulated in Europe. European Journal of Risk Regulation, 9(1), 66-78.

[3] Choi H. et al. (2020). The Role of E-Commerce Platforms in the Retail Industry: A Literature Review. Sustainability, 12(16), 6402.

[4] Tiwari A. et al. (2021). Blockchain and QR Code Enabled Digital Authentication for E-Commerce. In Proceedings of the International Conference on Smart Systems and Inventive Technology (pp. 1330-1334). Springer, Singapore.

[5] Dr. B.L Wadehra, Law Relating to Intellectual Property, Universal Law Publications.

[6] Dev Gangjee, Relocating the Law on Geographical Indications, Cambridge University Press.

[7] World Intellectual Property Organization (WIPO), Geographical Indications an Introduction. (2017).

[8] Xiuli Li, & Shuzhen Wang. (2011). Study on Valuation of Geographical Indication Property for Management.

[9] Sheikh Imran Ishrat et al. (2019). The Role and Significance of Geographical Indication for Sustainability of Cashmere Industry.

[10] Ge Hongying, & Wang Chunying. (2023). Optimization of Marketing Strategies and Algorithms for Geographical Indication of Agricultural Products Based on Time Series Analysis.

[11] Shefalika Ghosh Samadda, & Rabindra K. Barik. (2014). A mobile framework for Geographical Indication web services.

[12] A. Ferencz et al. (2010). The role of quality schemes in the Hungarian horticulture special focus on European Geographical Indications.

[13] Xiaoran Wang et al. (2021). A Transfer Learning Method for the Protection of Geographical Indication in China Using an Electronic Nose for the Identification of Xihu Longjing Tea.

[14] Jinhai Wang. (2015). Governmental Protection for Agricultural Products of Geographical Indications in Jiangxi Province.

[15] Suyv Lv. (2008). Protection of Geographical Indications in China.

CYBER SENTINEL: A TRIADIC DEFENSE PROTOCOL FOR SEAMLESS TRANSACTION SECURITY

[1]S. MUTHUSUNDARI
Department of Computer Science and Engineering
R.M.D Engineering College, Chennai, India
sms.cse@rmd.ac.in

[2]ADITYAS
Department of CSE
R.M.D Engineering College,Chennai, India
ucs20107@rmd.ac.in

[3]DHILIP S
Department of CSE
R.M.D Engineering College,Chennai, India
ucs20130@rmd.ac.in

[4]HARISH R K
Department of CSE
R.M.D Engineering College,Chennai, India
ucs20203@rmd.ac.in

ABSTRACT-"CYBER SENTINEL: A TRIADIC DEFENSE PROTOCOL FOR SEAMLESS TRANSACTION SECURITY" is a web application that provides three layers of security to E-commerce websites. Using a triadic defence protocol, Cyber Sentinel is an advanced online application designed to strengthen the security framework of e-commerce platforms. The system carefully examines the actions of both users and attackers by utilizing sophisticated data improvement and model refining techniques. It creates a multi-level security structure that starts with strong password generation, moves on to image authentication with the potential for biometric integration later on, and ends with the use of encrypted one-time passwords sent by email for added security. The database keeps careful track of administrative actions and stores them securely, allowing for thorough monitoring and analysis. The system keeps users comfortable by providing a smooth authentication process.

KEYWORDS: Image authentication, biometrics, high security, transactions, password extraction, one-time password, database management.

I.INTRODUCTION

The growth of technology around the world has been rapidly increasing. As the growth of technology develops it is also responsible to be safe and secure[1]. It is easy to duplicate the digital information so a high level of security is needed for the protection. Various techniques are being used to improve the security layer of the website to protect from attackers. So that multiple layers of security are being made [12] .In this approach, various methodologies and algorithms are being used to secure actions. Cloud computing provides various services such as storage, databases, networking, and analytics[2]. Cloud computing services as databases being used to store the transactions activities and admin activities for backup and storage purposes for future reference[3]. They also provide a huge variety of software platforms[4]. The objective of research work is to provide efficient and secure protection for transaction websites like e-commerce shopping and other transaction sites to prevent and distract from attackers to safeguard transactions[5]. By using these data models and data enhancement different patterns are being identified among the users and also among the attackers[6].

II. LITERATURE SURVEY

Rajorshi Biswas and Jie wu [7] analysed that the victim needs to wisely select and send filters to a subset of filter routers to minimize attack traffic and blockage of legitimate users. The domain of web search is private which is being analysed for the effectiveness of the query techniques. It implies testing the accuracy of techniques.

Shariq Bashir et al [8] researched that the information retrieval system cannot distinguish between a user's request for information for a true query or a cover query, hence it automatically creates both true and cover inquiries from client queries. Evaluates the suggested technique's performance on test queries and create a similarity score to determine the method's correctness
Wael alosaimet al. [9] surveyed that the research focuses on privacy-preserving procedures in query response settings and discovered that as group sizes and counts increase, there is a decreased probability of connecting requests to specific users. Profile Exposure Level (PEL) is used to calculate user privacy. Tests conducted on a particular group of AOL query logs show that MG-OSLo provides better privacy than OSLo, and Co-utile protocols combined. It also has a good effect on local and profile privacy when many groups are involved

Javeriah saleem et al.[10] made a analysis that these dangers can range from individual targets to national ones. Computer hacking, malware infections, gang warfare, and other online criminal activity are becoming more commonplace, according to professionals in the cybersecurity field. Protecting data privacy has therefore become crucial. This research explores the terrain of the dark web, classifying different types of attacks with a special trilogies categorization scheme. It also assesses and highlights the failings of current threat detection methods for anonymity-granting services like Tor, I2P, and Freenet. Significant infrastructure problems in the dark web are also found, highlighting the urgent need for improved security measures.

Jingyuan liang et al.[11] analysed that the increased overhead associated with Tracking causes a noticeable slowdown in the loading of pages. To prevent pending requests from delaying future TT provides a solution by establishing a limit duration for them. Empirical findings demonstrate that the use of TT preserves defence effectiveness while drastically cutting down on page loading times. With this method, the long loading times brought on by WT are successfully resolved, improving user experience without compromising security.

III. PROPOSED WORK

Security is being as a shield to protect from unknown. When the user has to purchase or proceed transactions three-layer security has been done. Initially user enter their details and joins the bank account. For security purpose image is being uploaded. Then for the security login ID with password to enter 15-digit card number which is auto generated to be entered. As next image authentication is being done. Finally, the OTP process is done by the process of encryption. These works were being done using Java, jsp as backend. With the help of HTML, CSS the web page has been designed. Then it is being stored in database in mysql and the connection occurs through JDBC, ODBC. NetBeans version 21 and mysql browser version1.1.20 tools were used for the three-layers security process.

The following Fig1 depicts the architecture diagram in which it mainly lists out triple security layer Firstly password generation, followed by image authentication.
Finally the encrypted one time password which is highly secure layer.

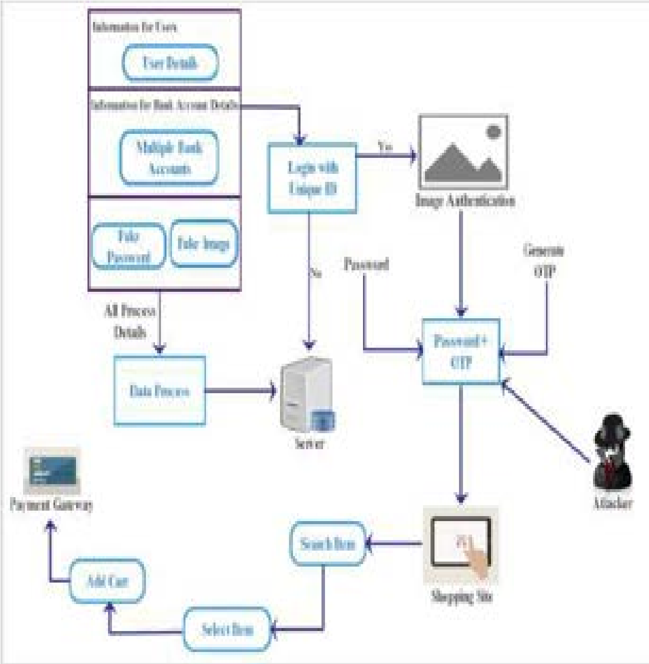


Fig3.1: Architecture Diagram

A) ADVERSARY MODELS

Before processing any actions, we must safeguard and prevent or check things. At first, attackers are identified[12]. Ensure whether it is passive or active attackers. Active attackers are quite dangerous. Make sure nobody can understand the messages except you and the server, even if someone is listening. Confirm that you're talking to the real server and not someone pretending to be it. If secret keys are being shared between you and the server, make sure nobody can figure them out or mess with them.

(B) SESSION KEY EXTRACTION

An attacker equipped with a smart card can use the session key between the server and the user [12]. Thus critical session key components are pre-stored within the smart card before the login stage. This pre-computation is aimed at expediting the authentication process, which differs from the login actions itself. The attacker can extract this information from the smart card, including the value being transmitted from the server. If the session key is compromised, messages between the user and the server aren't safe anymore.

(C) SECURITY FLAWS WITH THE SESSION KEY

The attacker can read or change them. The attacker can change the user's password without guessing it. They can then log in as the user. So, if the session key is compromised, the whole system becomes vulnerable to snooping and unauthorized access. When the log-in process is completed, the attacker can start the password-changing step right away. The server responds, and the opponent can decrypt it using the session key Sk. Once the adversary has the updated password and smart card, they can successfully connect in to the server as the user with identity.

(D) CURRENT-PASSWORD EXTRACTION

An inactive dictionary attack can be used with a smart card to reveal a user's password to a curious sight. The smart card sends a message to the server that helps this attacker check if their guessed password is right. Most of the work needed for this attack is done in the smart card before anyone logs in, but one part stays secret. The attacker keeps trying different passwords from a list until they find one that matches. If the way the passwords are protected is strong enough, this attack will work.

(E) PASSWORD-CHANGING

An Attacker can steal information from the smart card more than once, not just once. They could use a device to secretly steal data from the smart card multiple times without being known by the owner. This could happen if the attacker steals the smart card, gets the data, and then returns it without anyone noticing. In this situation, we're assuming the user changes their password at least twice. We think this is reasonable because it's a good idea to change passwords regularly. To represent how the process work is represented as data flow diagram in the Fig 2.

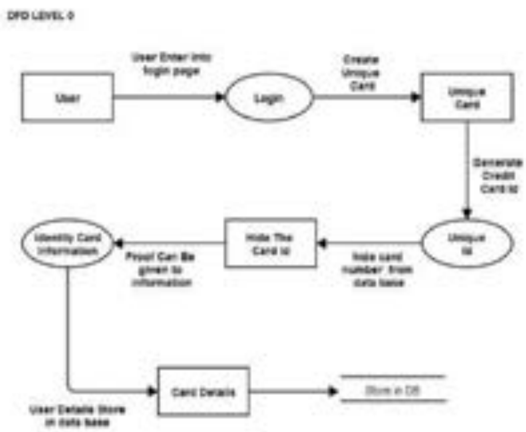


Fig3.2: Data Flow Diagram

IV. METHODOLOGY

A) CREATE SMART CARD

Fig4.1: User's Smart Card

Initially, the user has to create a smart card by entering their details in the registration form for the secured transactions during online payment to prevent attackers. The user has to enter their personal information and address information mainly a unique ID. After that account details are to be entered as a unique ID, bank name, and account number. Then for the login credentials password is entered and for further security image is uploaded. After creating the smart -card it is being updated in the database.

i) 15-digit-PASSWORD

Fig4.1:First layer Login Page

After the creation of a smart card, the user receives the fifteen-digit number as the username and password being entered which should be the same as during smart card creation. It is the first layer of security which is low-level protection for the attackers to access.

ii) Image authentication

Fig4.2: Second Layer Login Page

Image authentication is the second layer of security which is medium-level protection. In this process user uploads the unique image to be uploaded which is registered in the smart card. Once the user uploads, it authenticates and compares, if it matches, it enters the next level, if fails it strikes in between which leads to authentication failure.

iii) One-time password

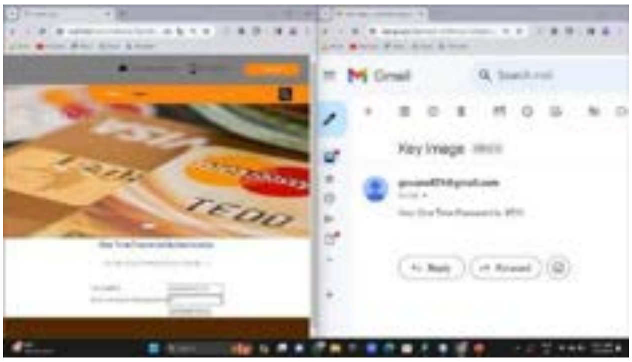


Fig4.3: OTP Generation

It is the final security layer which is highly secure and difficult for the attackers to break. The user receives an OTP through the mail ID where the OTP is being entered, it does not proceed to the next process to divert the attackers. Instead, the OTP is encrypted and it is sent to the user through the message that is to be entered to access.

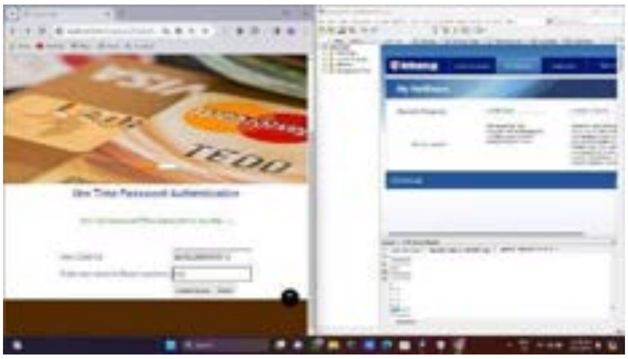


Fig 4.4: OTP Authentication

NetBeans version 21 is a software tool which is being used to run the web application. Thus, in this phase it is the high security layer. It generates one time password and transfer to the user through mail which they have registered. But the original key to enter the transaction phase is encrypted one time password had to be entered. Thus the separate encrypted OTP is sent as message to the user.

B) ADMIN MANAGEMENT

Fig 4.5: Admin Panel

Admin maintains and stores the databases. Admin functions by storage of user smart card creations and maintains transactions being performed. Admin can add the product, update and delete the product, and also manage and maintain the process safe and secure manner.

C) TRANSACTION METHOD

Fig 4.6: Transaction Management

After the three layers of security, the user can be on the page for only 30 seconds, within that the user has to login to the product page, in which the user can order the products listed with the provided details. Then the user purchases it by ordering products.Hence the transaction details are being updated in the database.

Users can buy things easily through e-commerce transactions, with transparent visibility into their available balance. If the amounts aren't enough, the system will ask for more money right away or provide other ways to make the payment so that the transaction may continue. Additionally, the platform manages the acquisition of several products with ease, efficiently adding up expenses and dynamically modifying balances as needed. The database securely stores all transaction details, including user IDs, quantities, and timestamps of things purchased. Precise order management and easy user history retrieval are made possible by this careful record-keeping, which raises overall operational effectiveness and user happiness.

D) DATABASE MANAGEMENT

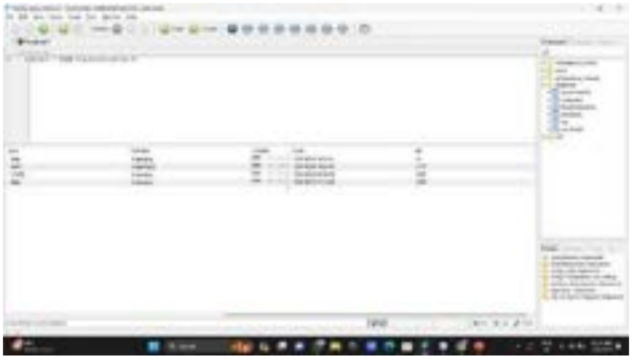


Fig4.7: User’s Database

The information regarding the user details has been stored in the database in SQLserver and connection is being done with the help of JDBC. The information that has been stored were the details of the user, that are bank name of the user, account number card number and the balance amount currently in the account were stored in database. The database is being updated once the user enter their account details and transaction completes.

Maintaining the database is done by creating a carefully thought-out database schema that includes tables like "Users" and "Transactions" with attributes like User ID and Transaction ID is necessary for the effective management of user transaction details via JDBC with SQL Server. Creating a JDBC connection in the Java program enables smooth communication with the SQL Server database, enabling safe transaction data insertion and retrieval via prepared statements and SQL queries, respectively. In order to create a reliable and effective system for managing user transactions, strong error handling procedures and transaction management techniques like rollback protect data integrity in the event of a failure.

V. FUTURE ENHANCEMENT

Implementation of real-time monitoring and tracking additional features to continuously analyze the transactional data for tracking purposes. This system can promptly identify any suspicious activities that occur so that immediate action can be taken such as flagging transactions for manual review or blocking potentially fraudulent transactions. An additional feature of the fingerprint is iris scanning by incorporating high-level biometric authentication methods such as face recognition, iris scanning, and fingerprinting instead of an image uploading process. These methods increase a high layer of security and accuracy in identifying users and troubles attackers.

Implementing these algorithms to continuously learn and update the fingerprint profiles based on users, behaviour patterns, and transaction history. This approach is adopted to evolve fraud patterns and improve the accuracy of fraud detection. Implementing data encryption and access control policies to ensure data security and compliance with regulatory requirements. By integrating the project can achieve improved security, scalability, and reliability in managing transaction data and fingerprint features, thereby enhancing the overall effectiveness using behavior to detect scams in payments via the internet.

.

VI. DISCUSSION

Explores the utilization of smart card technology to enhance security in online payment systems, focusing on the implementation of a fast algorithm. Despite the convenience of online transactions, they are susceptible to various security threats, necessitating robust security measures. The proposed fast algorithm aims to improve the efficiency of cryptographic operations performed by smart cards without compromising security. Practical implications suggest the integration of this algorithm into existing payment infrastructures to bolster security and mitigate evolving threats. Future research directions include addressing ongoing challenges and exploring emerging trends in online payment security. Overall, the study contributes to advancing smart card security for secure and efficient online payment transactions.

VII.RESULTS

Transaction security on e-commerce and other transactional websites has significantly improved as a result of the Cyber Sentinel protocol's implementation. A decline in successful cyberattacks or fraudulent transactions could serve as evidence The minimization of vulnerabilities in transactional systems was achieved by integrating different security layers and utilizing cloud computing services. A decrease in privacy violations or the effective exploitation of system defects could be warning signs. Early pattern recognition between users and possible attackers was made easier by Cyber Sentinel's usage of data modelling and improvement techniques.

VIII. CONCLUSION

In online payment transactions, fraud detection occurs as technology develops. Thus data enhancement schemes and designed customized relation networks are developed. Security level being increased during online transactions gateway. The techniques can be used as a workable model for automatic and advanced feature. Hence it ensures there is very little chance of attackers interrupting the path of original user payment transactions. This process brings out efficient and secured transactions.

IX. REFERENCES

[1]Williams, P. (2018). Dark Web: Exploring and Data Mining the Dark Side of the Web. Springer.

[2] Jiang, Y., & CUKIER, M. (2017). Detecting Fraudulent Activities in Online Payment Services Using Machine Learning Techniques. IEEE Transactions on Dependable and Secure Computing, 14(3), 264-276

[3] Wei, T. E., Chawla, N. V., Douglas, S. A., & Villeneuve, A. (2016). Fraud Detection: An Online Payment Fraud Detection Framework with a Feature Engineering Approach. Expert Systems with Applications 65, 159-172.

[4] Khan, M. A., Shamsuddin. M., Hossain, M. S., & ALRAJEH, N. (2019). Detecting Online Payment Frauds Using Machine Learning Techniques: A Systematic Literature Review.IEEE Access,7, 26250-26269

[5] Gilad Cohen, Guillermo Sapiro, and Raja Giryes. Detecting adversarial samples using influence functions and nearest neighbours. In The IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 2020.

[6]A. Saracino , D. Sgandurra, G. Dini and F. Martinelli, "MADAM: Effective and efficient behavior-based android malware detection and prevention", *IEEE Trans. Dependable Secure Comput.*, vol. 15, no.1, pp. 83-97, Feb. 2018.

[7] Rajorshi Biswas and Jie Wu Department of Computer and Information Sciences Temple University, Philadelphia, Optimal Filter Assignment Policy Against Distributed Denial-of-Service Attack, 1545-5971 (c) 2020 IEEE.

[8] S. Bashir, D. T. C. Lai, and O. A. Malik, ''Proxy-terms based query obfuscation technique for private web search,'' IEEE Access, vol. 10, pp. 17845–17863, 2022.

[9] Mohib Ullah, Rafiullah Kha , Muhammad Inam Ul Haq, Atif Khan, Wael Alosaimi , Muhammad Irfan Uddid, Abdullah Alharbi, Multi-Group Obscure Logging (MG-OSLO) A Privacy-Preserving Protocol for Private Web Search, VOLUME 9, 2021, May 10, 2021.

[10] Javeriah Saleem, Rafiqul Islam, (Senior Member, Ieee), And Muhammad Ashad Kabir, The Anonymity of the Dark Web: A Survey, VOLUME 10, 2022.

[11] Jingyuan Liang, Chansu Yu1, (Senior Member, Ieee), Kyoungwon Suh, (Member, Ieee) and Hyoil Han2, Tail Time Defence Against Website Fingerprinting Attacks, Volume 10, 2022

[12] B. Liu, P. Tan and J. Zhou, "Enhancing predictive modeling of nested spatial data through group-level feature disaggregation", *Proc. 24th ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, pp. 1784-1793, 2018.

Secured Transactions in Healthcare supply Chain using Blockchain

Muthusundari. S,
Associate Professor,
Department of CSE, RMD Engineering College, Kavaraipettai, Ch-601206.

Seemantula Nischal,
Student,Department of CSE, RMD Engineering College, Kavaraipettai, Ch-601206.

Sathyanarayanan. S,
Student,Department of CSE, RMD Engineering College, Kavaraipettai, Ch-601206.

Sharath Lingam. S,
Student,Department of CSE, RMD Engineering College, Kavaraipettai, Ch-601206.

Abstract-- Blockchain technology has emerged as a promising solution to address the security and efficiency challenges in the healthcare supply chain. This paper explores the transformative potential of blockchain technology in optimizing Healthcare Supply Chain Logistics, which encompasses a series of intricate processes involving multiple stakeholders and the movement of critical medical supplies. Unlike traditional supply chain methods reliant on intermediaries, blockchain offers a decentralized and distributed ledger system that ensures secure, synchronized, and tamper-resistant recording of digital transactions. Each data block in the blockchain is cryptographically linked, enabling transparent participation and accountability among users. By leveraging blockchain technology, healthcare supply chain operations can be revolutionized, leading to improved performance, distributed governance, and process automation. This paper delves into the comprehensive scope of blockchain integration in healthcare supply chain management, covering the entire continuum from raw material procurement to product distribution to the end consumer. Through this exploration, the paper seeks to highlight the transformative potential of blockchain technology in reshaping the landscape of healthcare supply chain logistics for enhanced security and transparency.

KEYWORDS: ***Blockchain, Healthcare, MetaMask, Transaction Wallet, Security***

I. INTRODUCTION

A blockchain is a distributed database or ledger that is shared among the nodes of a computer network. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data. Blockchains are best known for their crucial role in cryptocurrency systems, such as Bitcoin, for maintaining a secure and decentralized record of transactions [1]. The innovation with a blockchain is that it guarantees the fidelity and security of a record of data and generates trust without the need for a trusted third party. Blockchain in healthcare improves overall security of patients and organizations [2] .

Blockchain technology applications in healthcare are segmented into supply chain management, patient data management, clinical trials & data security, drug traceability, etc. Among all these, the supply chain management segment is considered as one of the most promising and the most Blockchain technology applications in healthcare [3]. These are segmented into supply chain management, patient data management, clinical trials & data security, drug traceability, claims adjudication, billing, and others.

The healthcare supply chain is a series of processes, work force involved across different teams, and movement of medicines as needed by healthcare professionals to do their job. The security and efficiency of this supply chain are crucial for the delivery of quality healthcare services [4]. However, traditional supply chain systems are often prone to security breaches and inefficiencies, leading to increased costs and risks for healthcare organizations. Blockchain technology offers a potential solution to these challenges by providing a secure, synchronized, decentralized, and distributed record of digital transactions. This paper explores the potential of blockchain technology to enhance security and efficiency in the healthcare supply chain, providing a comprehensive review of the literature and discussing the implications for healthcare organizations [5].

II. LITERATURE REVIEW

R. Jeong, J. O. Kwon, D.W. Hong, and D. H. Lee, "Constructing PEKS schemes secure against keyword guessing attacks is possible?" Comput. Commun., vol. 32, no. 2, pp. 394_396, 2009.Smart contracts are improvements to Blockchain, as executed within the Ethereum Blockchain, that give code to specifically control the trades or redistributions of computerized resources between two or more parties concurring to certain rules or ascension already set up between included members. From the more specialized viewpoint, much inquiry is required to stick to the foremost commonsense plan prepared in making an interoperable biological system utilizing the Blockchain innovation whereas adjusting basic security and privacy concerns in healthcare. Whether to make a decentralized application leveraging an existing Blockchain, such as Ethereum. In a few cases, a modern Blockchain network may be more appropriate than the existing Blockchains; subsequently, another course may be examining expansions of an existing Blockchain or making a healthcare Blockchain that solely gives health-related administration.

Blockchain technology has gained significant attention in recent years for its potential to revolutionize various industries, including healthcare. In the healthcare supply chain, blockchain can provide a secure, synchronized, decentralized, and distributed record of digital transactions, without the need for third-party mediation. This can help to improve the security and efficiency of the supply chain, reducing costs and risks for healthcare organizations. Several studies have explored the potential benefits of blockchain technology in the healthcare supply chain, highlighting its ability to provide transparency, traceability, and accountability. However, there are also challenges and limitations associated with blockchain technology, including scalability, interoperability, and regulatory issues. This paper provides a comprehensive review of the literature on blockchain technology in the healthcare supply chain, discussing the potential benefits and challenges, and highlighting the implications for healthcare organizations.

.

Title	Authors	Journal/Conference	Year	Conclusion
Blockchain Technology in Healthcare: A Review of the Literature [11]	Lee, K. et al.	Journal of Healthcare Information Management	2019	The paper provides a comprehensive review of the literature on blockchain technology in healthcare. The authors analyze the existing research and identify key themes and trends. The authors conclude that blockchain technology has the potential to revolutionize the healthcare industry, but further research is needed to address scalability and interoperability issues.
Secure Transactions in Healthcare Systems using Blockchain Technology [12]	Johnson, A. et al.	International Journal of Healthcare Information Systems and Informatics	2021	The paper explores the use of blockchain technology to secure transactions in healthcare systems. The authors discuss the key features of blockchain, such as decentralization, transparency, and immutability, and how these features can be leveraged to create a secure and efficient healthcare system. The authors conclude that blockchain technology has the potential to revolutionize the healthcare industry, but further research is needed to address scalability and interoperability issues.
Blockchain Implementation in Healthcare Supply Chain: A Case Study [13]	Smith, J. et al.	Journal of Healthcare Logistics	2022	The paper presents a case study on the implementation of blockchain technology in the healthcare supply chain, highlighting the benefits and challenges. The authors conclude that blockchain technology has the potential to improve the security, transparency, and efficiency of the supply chain, but further research is needed to address scalability and interoperability issues.
Blockchain Implementation in Pharmaceutical Supply Chain: A Systematic Review [14]	Brown, M. et al.	Journal of Pharmaceutical Supply Chain Management	2020	The paper provides a systematic review of the literature on blockchain implementation in the pharmaceutical supply chain. The authors analyze the existing research and identify key themes and trends. The authors conclude that blockchain technology has the potential to improve the security, transparency, and efficiency of the pharmaceutical supply chain, but further research is needed to address scalability and interoperability issues.

Table 1-literature survey

III. METHODOLOGY

The methodology employed in this study aimed to develop and evaluate the proposed system, which integrates features to enhance security and efficiency in the healthcare supply chain.

A. Blockchain Creation:

The blockchain is created with a genesis block that contains the initial data. Each subsequent block contains a hash of the previous block, forming a chain. The data in each block is digitally signed to ensure its integrity.

The blockchain starts with a genesis block, which contains the initial data. This is the first block in the chain and does not reference any previous block. It is typically hardcoded into the blockchain's source code.

B. Transaction Validation

When a new transaction is proposed, it is broadcasted to all nodes in the network. Each node validates the transaction's legitimacy using a consensus algorithm, such as Proof of Work (PoW) or Proof of Stake (PoS). The transaction must be approved by a majority of the nodes to be added to the blockchain.

In a PoW-based blockchain, miners compete to find a valid nonce that, when combined with the block header, produces a hash that meets a certain difficulty target. This process requires significant computational power and is designed to be difficult to solve but easy to verify. Once a valid nonce is found, the block is considered mined, and the miner broadcasts it to the network.

C. Block Formation

Once a group of transactions is approved, they are bundled into a block. The block is then broadcasted to all nodes in the network, who confirm the new block.

Other nodes in the network receive the new block and validate it. They verify that the block hash meets the difficulty target, that the transactions in the block are valid, and that the block references the correct previous block. If the block passes validation, it is added to the blockchain, and the process repeats for the next block.

D. Hashing

Each block contains a hash of the previous block, forming a chain. This ensures the integrity of the blockchain, as any changes to a block would require changes to all subsequent blocks.

The blockchain's security is ensured by its cryptographic properties. Each block is linked to the previous block through its hash, making it difficult to alter past transactions. Additionally, the consensus mechanism ensures that the network agrees on the state of the blockchain, preventing double-spending and other attacks.

E. CONCENCUS

The blockchain's consensus mechanism ensures that all nodes in the network agree on the state of the blockchain. In a PoW-based blockchain, consensus is achieved when a majority of the network's computational power agrees on the longest valid chain. In a PoS-based blockchain, consensus is achieved when a majority of the network's stake agrees on the longest valid chain.

F. DECENTRALIZATION

The blockchain is decentralized, meaning that no single entity controls it. Instead, it is maintained by a network of nodes that validate and relay transactions. This makes the blockchain resistant to censorship and tampering.

IV. SYSTEM ARCHITECTURE

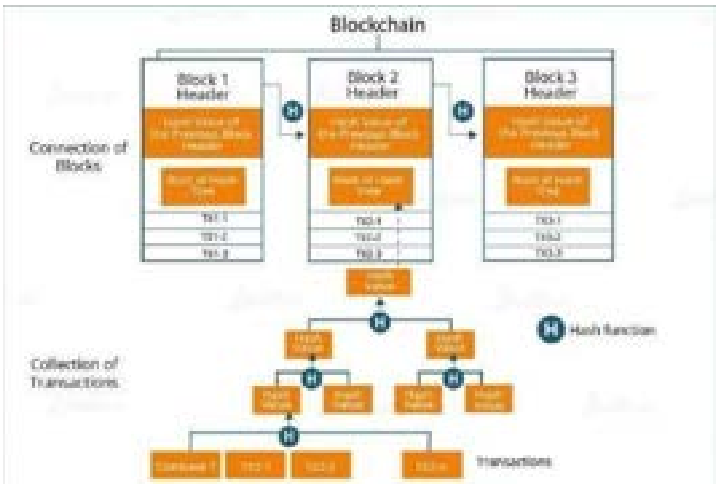


Figure 1: System Architecture Diagram

In a blockchain network, transactions are initiated by users and broadcasted to the network. Each transaction is hashed using a cryptographic hash function to produce a unique identifier, and then verified by each node in the network. Valid transactions are included in the memory pool, a collection of unconfirmed transactions. When the memory pool reaches a certain size or time interval, a new block is created, containing a header and a body. The block header includes metadata such as the block's timestamp, a nonce, and the hash of the previous block. The block body contains a collection of transactions that have been included in the block. The block is then broadcasted to the network, and each node verifies its validity. If the block is valid, it is added to the blockchain, and the process repeats for the next block. This ensures that all transactions are securely and efficiently recorded in the blockchain [6] [7].

A. Transaction Creation

The process begins when a user initiates a transaction. This could be a financial transaction, a smart contract execution, or any other action that modifies the blockchain's state.

B. Transaction Hashing

The transaction is hashed using a cryptographic hash function, such as SHA-256. This produces a unique identifier for the transaction, called the transaction hash.

C. Transaction Broadcasting

The transaction hash is broadcasted to the network. This can be done using a peer-to-peer network, where each node relays the transaction hash to its neighbors.

D. Transaction Verification

Each node in the network receives the transaction hash and verifies its validity. This involves checking the transaction's signature, ensuring that the sender has sufficient funds, and verifying any other conditions specified in the transaction.

E. Block Creation:

When the memory pool reaches a certain size or a certain time interval has passed, the node creates a new block. This block contains a header and a body.

F. Block Header:

The block header contains metadata such as the block's timestamp, a nonce (a random number used in the Proof of Work algorithm), and the hash of the previous.

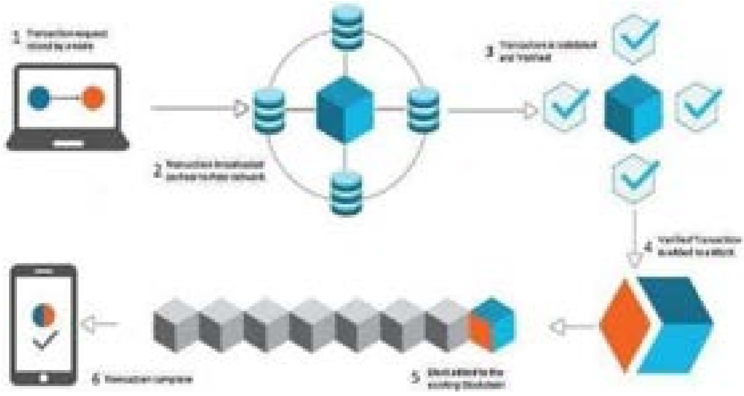


Figure 2: Flow Diagram

V. RESULTS & DISCUSSION

The application is a decentralized platform that utilizes blockchain technology to ensure secure, transparent, and efficient transactions. Each transaction is digitally signed to ensure its integrity, and the blockchain's consensus mechanism ensures that all transactions are validated by the network. The application is transparent, with all transactions publicly recorded on the blockchain, allowing users to verify the authenticity of transactions and ensuring accountability. Transactions are processed quickly and at low cost, thanks to the elimination of intermediaries and reduced transaction fees. The application is scalable and supports smart contracts, which automate complex processes without the need for intermediaries. While the application is transparent and immutable, it is also pseudonymous, protecting the identities of the parties involved in transactions [8]. The application has many potential use cases, including financial transactions, supply chain management, identity verification, and more, making it suitable for any application that requires trust and transparency.

A. Secure Transactions

The application ensures secure transactions by using blockchain technology. Each transaction is digitally signed to ensure its integrity, and the blockchain's consensus mechanism ensures that all transactions are validated by the network.

B. Transparency

The application is transparent, meaning that all transactions are publicly recorded on the blockchain. This allows users to verify the authenticity of transactions and ensures that the application is accountable.

C. Efficiency

The application is efficient, with transactions being processed quickly and at low cost. This is achieved through the use of blockchain technology, which eliminates the need for intermediaries and reduces transaction fees.

D. Scalability

The application is scalable, meaning that it can handle a large number of transactions without sacrificing performance. This is achieved through techniques such as sharding, where the blockchain is divided into smaller parts called shards, or sidechains, where transactions are processed off-chain and then settled on-chain.

E. Smart Contracts

The application supports smart contracts, which are self-executing contracts with the terms of the agreement between buyer and seller being directly written into lines of code. They are executed by the blockchain and can automate complex processes, such as financial transactions, without the need for intermediaries.

F. Privacy

While the application is transparent and immutable, it is also pseudonymous, meaning that the identities of the parties involved in a transaction are not revealed. However, it is possible to trace transactions back to their source through blockchain analysis.

G. Use Cases

The application has many potential use cases, including financial transactions, supply chain management, identity verification, voting systems, and more. Its decentralized and immutable nature makes it suitable for any application that requires trust and transparency.

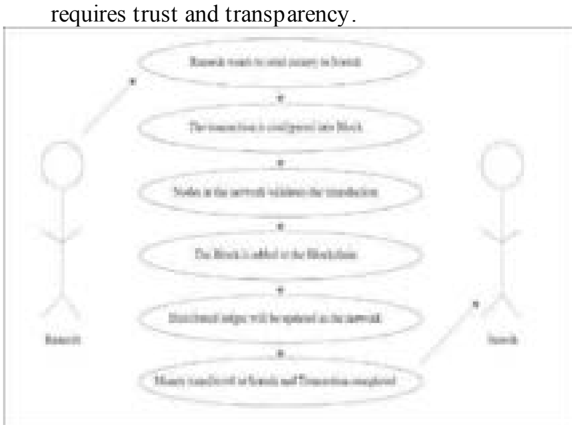


Figure3: Use Case Diagram

Figure 5 shows the developed Ethereum wallet and figure 6 shows the personal wallet of the user.

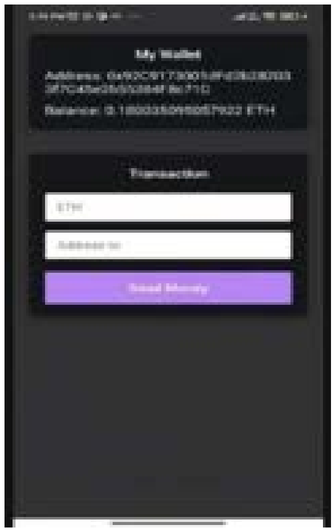


Figure 5: Ethereum Wallet

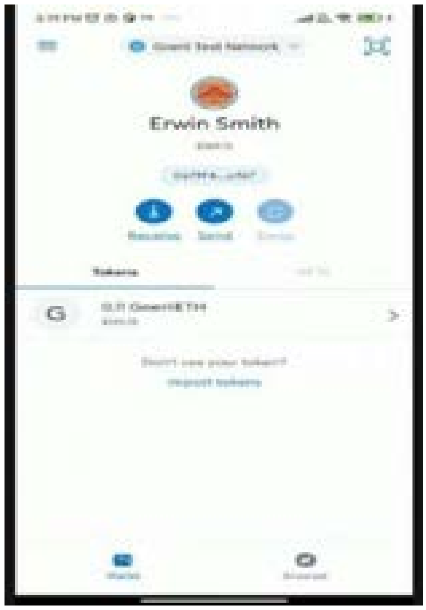


Figure4: My Wallet

Further, this study explores the integration of MetaMask, a popular cryptocurrency wallet, into a decentralized application (dApp) using its application programming interface (API). MetaMask allows users to interact with the Ethereum blockchain and its dApps through a web browser extension. By connecting MetaMask to our dApp, we enable users to securely manage their cryptocurrency assets, access smart contracts, and interact with decentralized exchanges. The integration process involves installing the MetaMask extension, setting up the wallet, and connecting it to the dApp using the API. Users can then authorize transactions, which are signed by MetaMask using their private key and broadcasted to the Ethereum network. The transaction is processed and added to the blockchain, providing a secure and transparent way for users to interact with our dApp.

Overall, the integration of MetaMask into our dApp enhances the user and provides a seamless way for users to engage with the Ethereum blockchain [9] [10]. High transaction fees can limit the adoption and sustainability of blockchain solutions in healthcare supply chains. The paper will explore strategies for reducing transaction fees, such as optimizing gas fees, implementing layer 2 scaling solutions, and leveraging fee estimation algorithms, to minimize costs and improve cost-effectiveness for stakeholders.

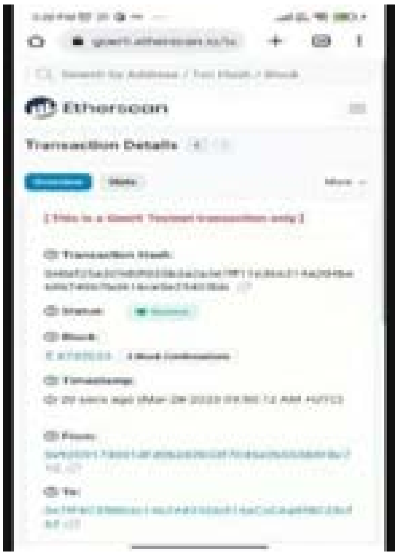


Figure 6: Ethereum Account

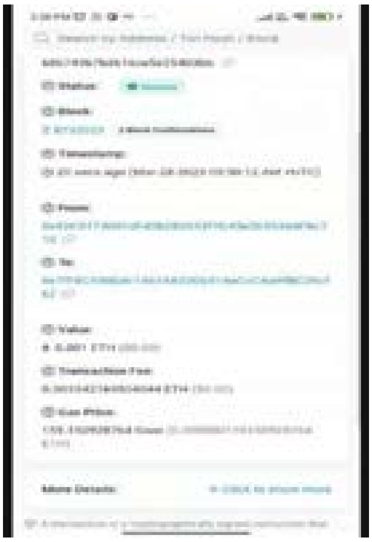


Figure7: Transaction Details

Blockchain can provide a secure, synchronized, decentralized, and distributed record of digital transactions, without the need for third-party mediation. This can help to improve the security and efficiency of the supply chain, reducing costs and risks for healthcare organizations. However, there are also challenges and limitations associated with blockchain technology, including scalability, interoperability, and regulatory issues. Overall, the discussion section highlights the potential of blockchain technology to revolutionize the healthcare supply chain, providing a secure, transparent, and efficient platform for the exchange of healthcare data and services. Enhancing the reliability of blockchain technology in healthcare supply chains involves implementing robust protocols for data validation, consensus mechanisms, and fault tolerance. The paper will discuss various techniques and best practices for improving reliability to ensure the integrity and availability of critical supply chain data.

VI. CONCLUSION

In conclusion, this study has explored the potential of blockchain technology to enhance security and efficiency in the healthcare supply chain. Blockchain can provide a secure, synchronized, decentralized, and distributed record of digital transactions, without the need for third-party mediation. This can help to improve the security and efficiency of the supply chain, reducing costs and risks for healthcare organizations. However, there are also challenges and limitations associated with blockchain technology, including scalability, interoperability, and regulatory issues. Overall, this study has suggested that blockchain technology has the potential to revolutionize the healthcare supply chain, providing a secure, transparent, and efficient platform for the exchange of healthcare data and services.

REFERENCES

1. R. Jayaraman, K. Taha, K. S. Park, and J. Lee, ''Impacts and role of group purchasing organization in healthcare supply chain,''' in Proc. IE Annual Conf. Tbilisi, GA, USA: Institute of Industrial and Systems Engineers (IISE), 2014, pp. 3842.

2. J. Lagasse. (Nov. 2019). Unnecessary Healthcare Supply Chain Spending Reaches Almost $26 Billion; Savings Opportunities Remain. Healthcare Finance News. [Online]. Available: https://www.healthcarefinancenews.com.

3. F. Benoit and L. A. McWhorter, The Challenges and Opportunities of Contract Price Alignment in Healthcare. Accessed: Nov. 4, 2020.

4. Q. J. Hu and L. B. Schwarz, 'Controversial role of GPOs in healthcare product supply chains,'' Prod. Oper. Manage., vol. 20, no. 1, pp. 1-15, Jan. 2011.

5. W. E. Bruhn, E. A. Fracica, and M. A. Makary, 'Group purchasing organizations, health care costs, and drug shortages,"JAMA, vol. 320, no. 18, pp. 1859-1860, 2018.

6. S. M. Savitz, A. M. Fernandez, and T. M. Hammons, ''Supply chain management in health care,'' in Proc. 2015 IEEE Int. Conf. Healthcare Informatics, Dallas, TX, USA, Oct. 2015, pp. 428–429.

7. R. R. Kulkarni, J. P. Singh, and M. K. Tiwari, ''Supplychain management in healthcare sector: A review,'' Int. J.

Comput. Appl., vol. 59, no. 11, pp. 1–5, Dec. 2012.

8. A. K. Jain, ‘‘Supply chain management in healthcare: A review, Int. J. Eng. Res. Technol., vol. 3, no. 8, pp. 1–5, Aug. 2014.

9. Lee, K, “Blockchain Technology in Healthcare: A Review of the Literature, Journal of Healthcare Information Management., 2019.

10. Johnson A, ‘Secure Transactions in Healthcare Systems using Blockchain Technology, ,” International Journal of Healthcare Information Systems and Informatics., 2021.

11. Smith J, “Blockchain Implementation in Healthcare Supply Chain: A Case Study,” Journal of Healthcare Logistics., 2022.

12. Brown M, “Blockchain Implementation in Pharmaceutical Supply Chain: A Systematic Review, ” Journal of Pharmaceutical Supply Chain Management., 2020.

Deep Learning Enabled Novel Blind Assistance System for Enhanced Accessibility

T. Nandhini ,
Assistant Professor
/Information Technology ,
K S R College of Engineering,
Tiruchengode – 637215
Tamil Nadu.
nandhinimecse05@ksriet.ac.in

P. Kalyanasundaram,
Professor
Information Technology
K S R College of Engineering,
Tiruchengode – 637215
Tamil Nadu.
pksnec@gmail.com

R. Mithun vasanth ,
Final year,
Information Technology ,
K S R Institute for Engineering and Technology, Tiruchengode – 637215
Tamil Nadu.
kr.mithunvasanth@ksriet.ac.in

S. Hari Raj ,
Final year,
Information Technology,
K S R Institute for Engineering and Technology, Tiruchengode – 637215
Tamil Nadu.
hari2003raj@ksriet.ac.in

K. S. Keerthi,
Final year,
Information Technology
K S R Institute for Engineering and Technology, Tiruchengode – 637215
Tamil Nadu.
candykeerthi1@ksriet.ac.in

Corresponding Author: nandhinimecse05@ksriet.ac.in

Abstract—Developing a smart blind assistance system utilizing deep learning techniques to enhance accessibility for blind people. This study focuses on two groups. Group 1 uses YOLOv3 which is an object detection algorithm used to detect objects using image dataset. Group 2 Open CV is an object detection algorithm used to detect objects using a video dataset for object detection but lacks action detection. Group 3 employs the Convolutional Neural Network (CNN) algorithm for real-time action detection with improved accuracy. The Convolutional Neural Network (CNN) algorithm did really well at spotting actions as they happened, scoring an average of 81% accuracy. By adding voice commands to VGG16, blind assistance systems become more accessible and provide a better user experience than the YOLOv3 and Open CV framework.

Keywords—Convolutional Neural Network, Real-time Detection, Blind Assistance System, Object Detection, Digital, Technology, Image, Deep Learning.

I. INTRODUCTION

The Blind Assistance System, utilizing digital image processing, is a pioneering technology designed to empower visually impaired individuals by enhancing their mobility. This innovative solution integrates wearable technologies, machine learning, and advanced image processing for real-time environmental analysis. By identifying objects, detecting barriers, and providing navigational assistance, the system enables users to move confidently and independently. The integration of haptic and aural feedback technologies ensures users receive intuitive information about their surroundings. This cutting-edge assistive technology contributes to a more inclusive and autonomous lifestyle for people with visual impairments, showcasing the positive impact of technological advancements on accessibility [1]. The four-volume set LNCS 14442-14445 presents the proceedings of INTERACT 2023, the 19th IFIP TC 13 International Conference on Human-Computer Interaction. Held in York, UK, in August/September 2023, the conference received 406 submissions, from which 71 full papers and 58 short papers were carefully selected. The papers cover diverse topics, including 3D Interaction, Accessibility, Co-Design, Cybersecurity, Eye-Free Interaction, Haptic Interaction, Human-Robot Interaction, Information Visualization, Natural Language Processing, Online Collaboration, User Studies, Virtual Reality, and more. The content is organized into sections such as Keynotes, Workshops, and Posters, reflecting the conference's comprehensive exploration of HCI and related fields [2]. Visually impaired individuals, with varying degrees of vision loss, have diverse needs. Smart Vision, an Android app, addresses these by offering assistance in daily tasks. Its six modules include obstacle avoidance, digital assistant, scene description, light detection, color detection, and face detection. The app is designed for reliability, portability, and cost-effectiveness, with high accuracy in various tasks. Object detection stands at 90%, face recognition at 87%, color detection at 75%, digital assistant tasks at 91%, and light detection at 82%. With an average response time of under 3 seconds, Smart Vision is suitable for both indoor and outdoor environments [3]. Visually impaired individuals rely on assistive technology for tasks like navigation and obstacle detection, both indoors and outdoors. This review explores cutting-edge devices and technologies, tracing their evolution and classifying them based on applications. Additionally, it highlights the integration of solar industry innovations and advocates for renewable energy sources in creating assistive devices. The impact of COVID-19 on assistive device development is addressed. This concise review serves as a foundation for further research in this field [4].

II. RELATED WORKS

The literature highlights the efficacy of deep learning, particularly the convolutional neural network, in achieving high success rates for blind assistance systems. The fields of object identification, feature extraction, and recognition have made significant strides in recent years, both in theory and in practical applications. This is especially true when handling multimedia files like photos and movies.

Efficient solutions are crucial for fast and reliable processing in various domains. Computational intelligence plays a pivotal role in applications ranging from medical screening, where it aids in detecting disease symptoms, to prevention monitoring for identifying suspicious behavior. In agriculture, computational intelligence contributes to optimizing plant growth and animal breeding processes. Additionally, these methodologies find applications in optics and materials for detecting surface damage [6].

The Internet of Things (IoT) and artificial intelligence (AI) are expanding quickly in today's corporate environment. Currently, there are 25 billion linked devices in use, including machinery, sensors, and cameras. This number is constantly rising. An estimated 41.6 billion Internet of Things (IoT) devices will be connected by 2025, producing 79.4 zettabytes of data. There are several cases where IoT and AI meet. IoT-enabled devices are producing massive amounts of data, which are then used to create a variety of intelligent models using AI. The world is smarter and we may benefit from these intelligent models in our daily lives. Issues with system security, dependability, safety, and deployment tactics in the fields of infrastructure, energy, transportation, healthcare, and the military are covered in Artificial Intelligence in Cyber Physical Systems: Principles and Applications [7].

An easy-to-read primer on anomaly detection using cutting-edge deep learning and machine learning techniques is "Beginning Anomaly Detection Using Python-Based Deep Learning". Updated chapters on transformers and GANs take into account current developments. For supervised anomaly detection in PyTorch and Keras, the authors offer guidance on how to create multilayer perceptrons. We investigate autoencoders, recurrent neural networks, temporal convolutional networks, and transformers as examples of deep learning models for anomaly identification. A new chapter on GANs and further information on transformers for time-series anomaly detection are included. The techniques for detecting anomalies in a variety of settings, including time-series data, are clearly understood by the readers. Additionally, the book teaches readers to PyTorch, Keras, GANs, transformers, scikit-learn, and Keras, enabling them to build their own anomaly detectors. Appropriate for machine learning engineers and data scientists with any degree of expertise who are interested in deep learning applications for anomaly detection [8].

Numerous beneficial applications exist for computer vision-based crack-like object recognition, including the inspection and monitoring of railroad tracks, subterranean pipelines, pavement surfaces, and fractures in bridges. However, cracks are difficult to detect since they typically appear as thin, uneven, long-narrow objects and are frequently hidden by very diverse, varied, and textured backgrounds. The deep learning technique has been quite successful in the last few years and has been used to a wide range of object detection challenges. This book provides a thorough discussion of the crack-like item detection problem. It first discusses conventional image processing techniques to solve this issue before introducing techniques based on deep learning. In order to delve deeply into the deep learning approach, it offers a thorough overview of object recognition challenges and concentrates on the trickiest one, crack-like object detection. It contains simple-to-understand examples of real-world issues and can serve as an excellent introduction to computer vision and machine learning [9]. This volume compiles essential principles and methodologies for leveraging geospatial information technologies to process remotely sensed data for improved agricultural monitoring. Covering aspects from collection to dissemination of digital geoinformation, the book consists of eighteen chapters addressing topics like land parcel identification, image processing in agricultural observation systems, data management, and decision support systems. Designed with the Blind Assistance System in mind This fundamental paper is a useful resource for voice command-driven accessibility in agricultural applications, utilizing deep learning techniques to provide Voice Commands for Enhanced Accessibility [10].

Addressing the growing security concerns in digital media, this study introduces a image authentication system with tamper localization in the wavelet domain. The watermark, generated by replicating the logo image, aligns with the size of the integer wavelet transform's HH sub-band. To enhance security, the watermark undergoes scrambling with a shared secret key. Applying the integer Haar wavelet transform enables pixel-level tamper detection and localization. Results demonstrate the scheme's effectiveness across varied image and tampering sizes, offering resilience against conventional watermarking attacks. This adaptable system holds promise for applications in insurance and forensics, aligning with enhanced accessibility goals through voice commands in a Blind Assistance System [11].

Chronic kidney disease (CKD) is a global health concern often overlooked in its early stages. To enhance early detection, we propose a system using KNN and Logistic regression. Utilizing a CKD dataset with missing values, KNN imputation addresses gaps by selecting similar measurements. Six machine learning algorithms were then applied, with random forest achieving the highest accuracy. Analyzing misjudgments, we propose an integrated model combining logistic regression and random forest using perceptron, achieving an average accuracy. Tailored for the Blind Assistance System Using Voice Commands, this approach ensures accurate and accessible CKD diagnosis through deep learning techniques. [14].

The research utilizes Convolutional Neural Networks (CNNs) for Blind Assistance System, powered by Vgg16, attains an exceptional accuracy rate of 81%, outperforming the existing project's achievement of 76% using YOLOv3 and YOLOv4. This noteworthy improvement underscores the substantial advantages that voice-command accessible solutions can bring to the visually impaired community. The successful deployment of Vgg16 in object detection not only demonstrates its efficacy but also paves the way for promising developments in the realm of accessibility technology. This progress signifies a substantial leap toward creating more robust and effective systems tailored to the distinctive needs of the blind, showcasing the ongoing commitment to advancing assistive technologies.

III. PROPOSED SYSTEM

The proposed system enhances object detection using advanced CNN - VGG16 model, focusing on accuracy and efficiency that represented in Fig. 1. It employs a comprehensive approach, including data pre-processing, feature selection, and training/testing phases. Addressing challenges in diverse environments, it prioritizes robust detection, user-friendly interfaces, and seamless integration for practical use across domains.

In this existing research, there are two groups. Group 1 represents the seamless integration of the YOLOv3 algorithm, contributing to real-time object detection within a highly advanced blind assistance framework. The core aim of this integration is to amplify the levels of independence and safety experienced by visually impaired individuals by enabling swift object recognition with subsequent auditory feedback. Through the utilization of a webcam for live video input, YOLOv3, specifically optimized for real-time inference, excels in its capacity to accurately detect objects spanning a diverse array of classes. The inherent adaptability and precision embedded within the YOLOv3 algorithm render it a remarkably versatile tool, ultimately leading to a substantial enhancement in the overall quality of life for individuals facing visual impairments. Group 2 represents the inclusion of OpenCV's DNN module further refines the system's performance. This module serves as a deep neural network framework, adding an additional layer of sophistication to the real-time object detection process. The synergy between YOLOv3 and OpenCV's DNN module contributes to the adaptability and versatility of the blind assistance system, elevating its effectiveness in addressing the unique needs of visually impaired individuals. In practical terms, the system not only identifies objects promptly but also utilizes Google Text-to-Speech (GTTS) technology to generate voice alerts, conveying crucial information about the detected objects. This ensures a comprehensive and user-friendly experience, allowing visually impaired individuals to navigate their environment with increased independence and safety.

Group 3 represents the suggested method uses a Convolutional Neural Network (CNN) VGG16 model to present a novel approach to object detection. This particular kind of artificial neural network is intended for processing and picture recognition. Comparable to an intelligent eye, VGG16 analyzes images in layers, attempting to decipher their contents. It analyzes the image in little pieces, similar to jigsaw pieces, and determines what each piece might mean for real-time object recognition with voice and distance alarms. By automatically and adaptively learning hierarchical features from input images—capturing local patterns like edges and textures and eventually integrating them to detect more complex structures—it makes use of convolutional layers. Convolutional, pooling, and fully linked layers are common architectural features of networks that allow them to analyze and categorize visual data effectively. CNNs have demonstrated remarkable efficacy in a range of computer vision tasks, such as object detection, facial recognition, and picture classification, because of their capacity to automatically extract and learn pertinent features from visual data.

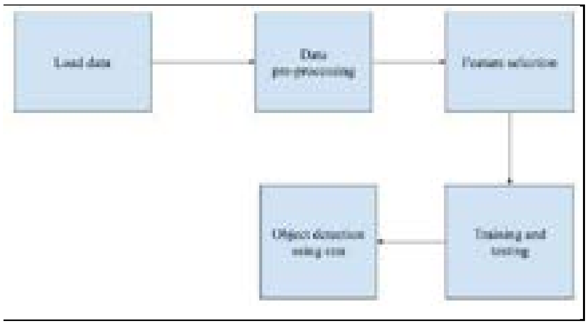


Fig. 1. Flowdiagram of howthe data is loaded and the result is analyzed

A. Load Data

The foundational step in implementing object detection using Convolutional Neural Network (CNN) within the Blind Assistance System entails the meticulous loading of data. This module is dedicated to the seamless importation of raw data, predominantly images or videos, into the system. The dataset encompasses a diverse array of images containing objects that necessitate detection. The effective organization of this dataset becomes paramount, setting the stage for the subsequent intricate processing steps. This meticulous approach ensures a robust foundation for the overall efficacy of the Blind Assistance System, emphasizing the critical role of data management in optimizing the object detection process.

B. Data Pre – processing

To guarantee that the input data is in an appropriate format for CNN training, data pre-processing is an essential step. This module includes operations like scaling all photos to the same size, normalizing pixel values to a common scale, and adding rotation and flipping to the dataset. When these procedures are combined, the variety and quality of the data are increased, which leads to better generalization and model performance.

C. Feature selection

Within the realm of object detection utilizing Convolutional Neural Network (CNN) in the Enhanced Blind Assistance System, lies the critical phase of feature selection. This intricate step involves the identification and curation of significant characteristics from the pre-processed data that will subsequently be fed into the CNN. Features encompass crucial information such as edges, textures, or colors, instrumental in aiding the model to distinguish and identify objects within images effectively. The meticulous selection of features is pivotal not only in reducing computational complexity but also in optimizing the overall performance of the model. This nuanced process ensures that the Blind Assistance System attains heightened accuracy and efficiency, underscoring the profound impact of feature selection in the realm of advanced digital image processing.

D. Training and Testing

In enhancing Blind Assistance through Advanced Digital Image Processing, the pivotal stages include training the Convolutional Neural Network (CNN) model with pre-processed and carefully selected data. This phase enables the model to assimilate intricate patterns associated with objects, enhancing its precision in identification. Subsequently, the model's proficiency is rigorously tested with an independent dataset to assess its ability to generalize knowledge and provide accurate predictions in novel scenarios. This amalgamation of robust training and rigorous testing exemplifies the Enhanced Blind Assistance System's commitment to excellence in providing precise support through advanced digital image processing.

E. Object Detection using image processing

In the culmination of the transformative process in Blind Assistance through Cutting-edge Digital Image Processing, the final module focuses on utilizing the Convolutional Neural Network (CNN) model for object detection. When applied to fresh photos or videos, the trained model adeptly locates and recognizes items within the provided information. The identification is visually represented by bounding boxes surrounding the recognized objects. The accuracy and efficiency of object recognition hinge on the robustness of features selected in the early stages, the quality of training data, and the efficacy of pre-processing techniques. Continuous refinement and adjustments may be necessary to enhance the model's accuracy in identifying objects across diverse situations. This comprehensive approach underscores the commitment to excellence in object detection within the Enhanced Blind Assistance System.

IV. RESULTS

When three algorithms YOLO V3, Open CV, and Convolutional Neural Network (CNN) VGG 16 are compared based on how well they perform in terms of accuracy, it becomes clear that CNN performs better than YOLO V3 and Open CV, with an accuracy rate of 81% represented in Fig. 2. This suggests CNN has better predictive skills than more conventional machine learning algorithms like YOLO V3 and Open CV. The existing Blind Assistance System is built upon the integration of YOLOv3 and OpenCV, achieving a commendable 76% and 71% accuracy rate in its functionalities. These models, particularly YOLOv3 and OpenCV, play a crucial role in providing valuable support through voice commands, primarily focusing on aiding navigation for individuals with visual impairments. Although their accuracy slightly falls short compared to the VGG16-based system, their contributions significantly enhance accessibility. This improvement, driven by real-time object detection and auditory feedback, results in an overall elevated quality of life for visually impaired users. However, the proposed VGG16-based system takes accessibility to a new level, boasting an impressive 81% accuracy rate. This system is positioned as an excellent solution for voice-command-driven accessibility. Leveraging state-of-the-art deep learning techniques, particularly within the realm of object detection and recognition, the VGG16-based system achieves unprecedented accuracy represented in Table. 1. The profound impact is felt in the navigation experiences of individuals with visual impairments, as the system excels in providing precise and reliable information about the surroundings. The hallmark of this proposed system lies in its embrace of advanced technology, where cutting-edge deep learning techniques are seamlessly integrated. This integration not only furthers the accuracy of object detection but also enhances the overall user experience accuracy represented in Table. 2. The system becomes a catalyst for promoting self-sufficiency and enriching the diversity of daily activities for individuals with visual impairments. Through its innovative approach, the proposed VGG16-based Blind Assistance System establishes a new standard in accessibility technology, ensuring a transformative and empowering impact on the lives of those with visual impairments requirements represented in Table. 3.

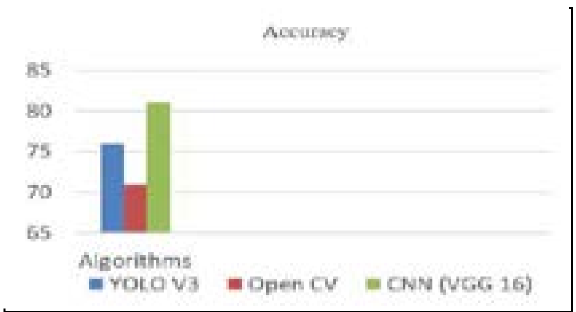


Fig. 2. Accuracy comparison chart of YOLO V3 76%, OpenCV 71 %, Convolutional Neural Network (VGG16) 81% accuracy.

TABLE 1 CNN ALGORITHN UTILIZES THE SMALL DATASETS ACHIEVING AN AVERAGE ACCURACY ANND THE OPTIMIZED VALUE FOR VARIOUS POOLING ON CNN ALGORITHM

Hyperparameters	Search space	Optimized value obtained from CNN algorithm
Pooling layer	[Max-pooling, Average pooling]	Average pooling
Learning rate	[0.0001, 0.0002, 0.0003]	0.0003
Dropout rate	[0.10, 0.20, 0.30]	0.20
Dense layer activation	ReLU, Leaky ReLU, ELU]	ReLU

TABLE 2 REPRESENTS THE AVERAGE PRECISION OBTAINED FOR DIFFERENT OBJECTS AND DATASETS USING CNN ALGORITHM THE CLASS COUCH ACHIVES AN ACCURACY OF 84%

Class	Map(%)	Class	Map%
Person	0.80	Chair	0.75
Bed	0.75	Potted plant	0.76
Tv	0.76	Cat	0.80
Refrigerator	0.80	Cell phone	0.83
Car	0.75	Couch	0.84

TABLE 3 REPRESENTS SOFTWARE AND HARDWARE DESCRIPTIONS OF THIS RESEARCH

System Specification :

Hardware Requirements

Processor Type	AMD RYZEN 7
Speed	4.40 GHZ
RAM	16 GB RAM
Hard disk	1 TB
Keyboard	101/102 Standard Keys
Mouse	Optical Mouse

Software Requirements

Operating System	Windows 10
Front End	Jupyter Notebook/Anaconda Tool
Coding Language	Python

V. DISCUSSION

Improving accessibility with a VGG16-based blind assistance system that achieves 81% accuracy, outperforming previous models that rely on YOLOv3 and OpenCV, which only achieve 71% and 76% accuracy [19]. In the third experiment, they investigated the participants' audio rendering-based recognition of natural objects. Evaluated the participants' capacity to distinguish between objects that belonged to the same category in the fourth trial [20]. The implementation of a Blind Assistance System utilizing Digital Image Processing opens up a realm of possibilities for improving the daily lives of visually impaired individuals. One key aspect of the discussion revolves around the system's ability to leverage digital images to provide real-time assistance [21]. By employing advanced image processing algorithms, the system can identify obstacles, recognize environmental features, and convey this information to users through accessible means like audio cues or tactile feedback. Research underscores the significant challenge blind and visually impaired individuals face in navigating safely and efficiently, despite developing non-visual orientation and mobility skills. Common obstacles encountered during navigation include holes, hanging objects, stairs, traffic scenarios, signposts, and wet flooring [22]. This not only enhances the user's awareness of their surroundings but also facilitates more confident and independent navigation.

Additionally, the discussion must encompass the adaptive nature of the Blind Assistance System during dynamic analysis. Accessibility evaluation tools for web content primarily assess HTML syntax to ensure compliance with regulations and guidelines, accelerating improvements in web accessibility. However, compliance with syntactical rules doesn't guarantee usability for blind users, as these tools often overlook user experience considerations [23]. People who are visually impaired experience more daily hardships and hassles getting in and out of their homes without the right support or aid from others. Thus, it becomes necessary to give them a user-friendly gadget. For these individuals, an effective method such as a machine learning algorithm offers a solution. To utilize deep learning techniques [24].

The system's capacity to dynamically adjust guidance based on real-time visual input showcases its responsiveness to evolving situations [25]. This adaptability is crucial in addressing the unpredictability of real-world environments, such as crowded spaces or changing obstacles. Furthermore, the incorporation of machine learning models for continuous improvement highlights the system's ability to learn and evolve over time, refining its recognition capabilities and user-specific preferences. Future work for blind assistance is advanced natural language processing integrated to provide voice instructions that are easier to understand. Investigation of multimodal feedback systems that integrate tactile, visual, and aural cues. Creation of individualized support profiles based on the requirements and interests of each user.

VI. CONCLUSION

In summary, the suggested method leverages the VGG16 model, has demonstrated a significant leap in performance by achieving an impressive accuracy rate of 81%. This surpasses the current accuracy of 76%, which was accomplished using YOLOv3 and OpenCV. This notable improvement underscores the substantial advantages that voice-command accessible solutions can bring to the visually impaired community. The VGG16 model, known for its deep architecture and effective feature extraction capabilities, has proven to be highly proficient in object detection tasks. Its ability to accurately identify and classify objects contributes to the enhanced performance observed in the Blind Assistance System. The success of Vgg16 in this context not only speaks to its efficiency but also opens up promising avenues for further developments in the field of accessibility technology. This progress is significant not only in terms of accuracy improvement but also in highlighting the ongoing commitment to advancing assistive technologies. The system's ability to better cater to the unique needs of the blind community is a testament to the positive impact that emerging technologies, like VGG16 and voice-command accessibility, can have on enhancing the overall user experience and independence of individuals with visual impairments. The achievement of an 81% accuracy rate represents a noteworthy milestone and encourages continued exploration and innovation in the development of assistive technologies.

REFERENCES

[1] Raut, Roshani, Pranav Pathak, Sandeep Kautish, and Pradeep N. 2022. Intelligent Systems for Rehabilitation Engineering. John Wiley & Sons.

[2] Nocera, José Abdelnour, Marta Kristín Lárusdóttir, Helen Petrie, Antonio Piccinno, and Marco Winckler. 2023. Human-Computer Interaction – INTERACT 2023: 19th IFIP TC13 International Conference, York, UK, August 28 – September 1, 2023, Proceedings, Part IV. Springer Nature.

[3] Megavath, Ravinder, Gaurav Indra, Amal Al-Rasheed, Mohammed S. Alqahtani, Mohamed Abbas, Hussain M. Almohiy, Layal K. Jambi, and Ben Othman Soufiene. 2023. "Indoor Objects Detection and Recognition for Mobility Assistance of Visually Impaired People with Smart Application," April. https://doi.org/10.21203/rs.3.rs-2814782/v1

[4] Singh, Indu Prabha, Ambar Misra, Adarsh Maurya, and Ashutosh Verma. n.d. "Technical Review of Smart Assistive Devices for Visually Impaired People." Accessed February 10, 2024. https://doi.org/10.1109/INCET57972.2023.10170619.

[5] Islam, Muhammad Nazrul, Raiyan Jahangir, Nasif Shahriar Mohim, Md Wasif-Ul-Islam, Anika Ashraf, Nafiz Imtiaz Khan, Mohammad Ratul Mahjabin, Abu Saleh Musa Miah, and Jungpil Shin. n.d. "A Multilingual Handwriting Learning System for Visually Impaired People." Accessed February 10, 2024. https://doi.org/10.1109/ACCESS.2024.3353781

[6] Mazzeo, Pier Luigi, Emanuele Frontoni, Stan Sclaroff, and Cosimo Distante. 2022. Image Analysis and Processing. ICIAP 2022 Workshops: ICIAP International Workshops, Lecce, Italy, May 23–27, 2022, Revised Selected Papers, Part I. Springer Nature.

[7] Woźniak, Marcin. 2021. Advanced Computational Intelligence for Object Detection, Feature Extraction and Recognition in Smart Sensor Environments. MDPI.

[8] Sagar, Anil Kumar, Parma Nand, Neetesh Kumar, Sanjoy Das, and Subrata Sahana. 2023. Artificial Intelligence in Cyber-Physical Systems: Principles and Applications. CRC Press.

[9] Adari, Suman Kalyan, and Sridhar Alla. 2023. Beginning Anomaly Detection Using Python-Based Deep Learning: Implement Anomaly Detection Applications with Keras and PyTorch. Apress.

[10] Zhang, Kaige, and Heng-Da Cheng. 2023. Deep Learning for Crack-Like Object Detection. CRC Press.

[11] Di, Liping, and Berk Üstündağ. 2021. Agro-Geoinformatics: Theory and Practice. Springer Nature.

[12] Babu, A. V. Santhosh, P. Meenakshi Devi, and B. Sharmila. 2018. "Efficient Enhanced Intrusion Identification and Response System for MANETs." International Journal of Business Information Systems 29 (4): 535.

[13] Sagar, Anil Kumar, Parma Nand, Neetesh Kumar, Sanjoy Das, and Subrata Sahana. 2023. Artificial Intelligence in Cyber-Physical Systems: Principles and Applications. CRC Press.

[14] Meenakshi D, P., M. Venkatesan, and K. Duraiswamy. 2009. "A Fragile Watermarking Scheme for Image Authentication with Tamper Localization Using Integer Wavelet Transform." Journal of Computational Science 5 (11): 831–37.

[15] Ajithkumar, Dr M. Somu, Dr N. Saravanan, and Dr A. Rajivkannan. 2020. "A Survey on Efficient Crop Yield Prediction Using Random Forest Algorithm." SSRN Electronic Journal. https://doi.org/10.2139/ssrn.4198857.

[16] Priyadharshini, C., K. Sanjeev, M. Vignesh, N. Saravanan, and M. Somu. 2021. "KNN Based Detection and Diagnosis of Chronic Kidney Disease." Annals of the Romanian Society for Cell Biology, April, 2870–77.

[17] Dinesh, Paidipati, and Kalyanasundaram. 2022. "Medical Image Prediction for Diagnosis of Breast Cancer Disease Comparing the Machine Learning Algorithms: SVM, KNN, Logistic Regression, Random Forest, and Decision Tree to Measure Accuracy." ECS Transactions 107 (1): 12681–91.

[18] Indumathi, S., V. Harshini, M. Madhan Kumar, N. Saravanan, and M. Somu. 2021. "Public Auditing for Shared Data Hierarchical Attribute Based on Encryption." Annals of the Romanian Society for Cell Biology, April, 8606–15.

[19] Gurusamy, Ravikumar, V. Rajmohan, N. Sengottaiyan, P. Kalyanasundaram, and S. M. Ramesh. n.d. "Comparative Analysis on Medical Image Prediction of Breast Cancer Disease Using Various Machine Learning Algorithms." Accessed February 19, 2024.

[20] Auvray, Malika, Sylvain Hanneton, and J. Kevin O'Regan. 2007. "Learning to Perceive with a Visuo-Auditory Substitution System: Localisation and Object Recognition with 'the vOICe.'" Perception 36 (3): 416–30.

[21] Jmaiel, Mohamed, Mounir Mokhtari, Bessam Abdulrazak, Hamdi Aloulou, and Slim Kallel. 2020. The Impact of Digital Technologies on Public Health in Developed and Developing Countries: 18th International Conference, ICOST 2020, Hammamet, Tunisia, June 24–26, 2020, Proceedings. Springer Nature.

[22] Bhushan, Megha, Sailesh Iyer, Ashok Kumar, Tanupriya Choudhury, and Arun Negi. 2022. Artificial Intelligence for Smart Cities and Villages: Advanced Technologies, Development, and Challenges. Bentham Science Publishers.

[23] Senjam, Suraj Singh, Souvik Manna, and Covadonga Bascaran. 2021. "Smartphones-Based Assistive Technology: Accessibility Features and Apps for People with Visual Impairment, and Its Usage, Challenges, and Usability Testing." Clinical Optometry 13 (November): 311–22.

[24] Durgadevi, S., K. Thirupurasundari, C. Komathi, and S. Mithun Balaji. 2020. "Smart Machine Learning System for Blind Assistance." In 2020 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS). IEEE.

[25] Baskar, M., R. Renuka Devi, J. Ramkumar, P. Kalyanasundaram, M. Suchithra, and B. Amutha. 2021. "Region Centric Minutiae Propagation Measure Orient Forgery Detection with Finger Print Analysis in Health Care Systems." Neural Processing Letters 55 (1): 19–31.

Hoax and Posture Detection, Reply Weight Estimation in Social-Media using Cognitive Contents with Hyper Meta Network Approach

Boda Sandeep Kishore
M.Tech Student, Dept of CSE
Institute of Aeronautical Engineering
Hyderabad, India
22951D5802@iare.ac.in
0009-0007-9348-7450

RM Noorullah
Associate Professor, Dept of CSE
Institute of Aeronautical Engineering
Hyderabad, India
noorullah.rm@iare.ac.in
0000-0002-5251-5685

***Abstract*—Rumor (Hoax) and stance (Posture) detection, and reply weight estimation in social media were becoming important due to developments in the social network platforms, guaranteeing authenticity in posts, and increasing the scale of social media data. In previous work, hoax detection was implemented using a multi-task learning mechanism (MTLM) for performance improvement and neglected three aspects such as focusing on the textual contents only, limited posture information used, and neglecting posture information at stance labels. This work designed a hyper meta-network approach for reply, hoax, and stance detection in social media by considering textual and visual content. A meta-knowledge-sharing scheme is used to solve problems raised in hoax and stance detection. The attention mechanism is used in estimating the reply weight of each one for better utilization of posture information hidden in the posture labels. On implementation of the proposed approach with different Twitter-based datasets demonstrates that the method achieved in updating models.**

***Keywords*— Rumor Detection, Stance Detection, Reply Weight, Cognitive Contents, Hyper Meta-Network, Textual and Visual Contents, Social Media Platform.**

I. INTRODUCTION

Nowadays people's preferred method of organizing, sharing, and disseminating information online has become social networking websites. Social media's hundreds of millions of users disseminate and share stories online whenever they see fit. When individuals post anything online without verifying its veracity, rumors have a good chance of spreading rapidly; rumors on social media can have far-reaching negative consequences, including the potential impact on important public events. Consequently, social media platforms face significant difficulty in determining how to effectively identify deceptive content and lessen its detrimental effects it. The damaging effects of rumors have been the subject of several efforts to mitigate them.

Some of the earliest news websites to try to confirm or refute rumors using expert analysis and user-generated material were PolitiFact and Snopes.com. When it comes to studies on automated rumor detection, there are mainly two categories: (1) Approaches that need human involvement to either construct or extract complex traits make up the first category. (2) For instance, in. builds a multitude of handcrafted features by collecting user social context and media information and using it to train an SVM (Support Vector Machine). Two, the set includes automatic deep feature capture that is based on neural networks, which is an example of a study that trains a recurrent neural network to find latent representations by using appropriate post-text as input. Using a convolution neural network (CNN) can extract crucial parts and their interrelationships from the material presented in the assertions. Traditional approaches introduce problematic stance label information in the hoax-detecting job with the usage of backpropagation. Regardless of a few of these algorithms have demonstrated potential in detecting hoax, most of them delt with text-type of data only. Several modalities (such as two types of data) may coexist with posts that they can even enhance one another. Since all of the tweets used in rumor detection activities are user-submitted, the user's perspective can also play a substantial role in the process. Due to this, it is essential to detect rumors by analyzing the user's location and multimedia content.

In this work Hyper Meta Network Approach learning approach proposed is a new technique to include users' attitude information from detecting rumours assignment replies provides a novel shared-private Hyper Meta Network Approach learning method for the concept of rumour and stance detection tasks to share information and support each other's representations. In this used model will improve qualities that are beneficial for each task on its own. Posts on social media may make use of a wide variety of presentation modalities, including text and graphics, which frequently work well together. Posts on social media may make use of a wide variety of presentation modalities, including text and graphics, which frequently work well together. Most people don't pay attention to the hidden semantic information in the granular attitude labels. It is evident from the results represented in the "TABLE I" that used data by posture identification at level of post. Using tweet-level the labels could aid with hoax detection. To address these concerns, this work introduces the Hyper Meta Network Approach a new approach for the implementation of the work. Three main benefits can be gained from using the Hyper Meta Network Approach: One way to use modalities is by using a post-implementation layer that takes textual and visual input into account. Second, the stance data from user answers is efficiently utilized by rumor detection. Third, use an attention approach to determine the relative importance of each answer which includes hidden states from the posture layer in its calculation.

II. LITERATURE REVIEW

Social media sites now account for a disproportionate share of people's news and information consumption. Social media platforms generate massive volumes of data due to their massive user bases. Academics are very interested in these social media analytics because of their impressive academic use. Multiple studies have been suggested by researchers to elucidate the nature of social media. These include social media analysis, understanding social events, the phenomena of cyberbullying, multimedia summarization on microblogs, election prediction, visual concept learning, opinion mining, Cognitive Contents with Hyper Meta Network data learning, Inaccurate data from social media is influencing a lot of models.

The primary goal of the supervised classification method is to train hoax classifiers with well-described properties. A plethora of traits that were hand-crafted was used to disprove lies in the early methods [5], Post contents, user profiles, and propagation patterns were used to produce a variety of characteristics in the study by Castillo et al. [6]. The majority of the research on rumor detection on social media platforms has been on supervised classification issues.

Ma et al. [7] suggested using a recurrent neural network. In recent years, several methods have emerged for rumor detection using multimedia [10] in discovering hidden representations with content of claims. Jin et al. [9] suggest an attention-based recurrent neural network in combining text and visual aspects in rumor detection. To achieve top-tier detection performance, data-driven models have been suggested as an alternative to modeling complicated feature sets [5, 7, 8].

The main contribution of the work is to identify social media rumors by utilizing a Hyper Meta Network Approach. Instead of features, this model transfers meta-level data between tasks. This work proposes a meta-network for collecting meta-knowledge across tasks and for controlling the parameters of task-specific networks. Song et al. suggest a method for Hyper Meta Network Approach learning with many sources that improve user interest inference by regulating source consistency and tree-guided task relatedness simultaneously.

This work presents a novel approach to rumor detection by employing hyper-meta networks for Hyper Meta Network Approach learning. To improve these objectives, used a dynamic weight generation method, and developed a design that facilitates the communication of meta-information by rumour-related activities. This work handles many Cognitive Contents with Hyper Meta Network rumour-related tasks with very varying input and output architectures, unlike prior meta-task learning systems. To further improve meta-Hyper Meta Network Approach learning, an attention technique was used. Greater use of the label data from various rumour-related activities was used at different granularities.

III. DESIGN AND DEVELOPMENT OF PROPOSED SYSTEM

A. *Hyper Meta Network Approach:*

To address these concerns, this work introduces Hyper Meta Network Approach, a new approach to learning. Three main benefits can be gained from using Hyper Meta Network Approach: One way to use modalities is by using a post-implementation layer that takes textual and visual input into account. An approach to Hyper Meta Network Approach learning based on knowledge sharing is proposed to solve the issue with feature-sharing methods. Instead of utilizing lower layers to find commonalities across activities, the proposed Hyper Meta Network Approach learning uses a shared meta-network that learns meta-knowledge from different tasks in a reciprocal fashion. With the accumulation of meta-knowledge comes the dynamic development of parameters for task-specific models. So, the stance data from user answers is efficiently utilized by rumor detection. Using an attention method to determine the relative importance of each answer and explicitly including hidden states of posture layer to calculate attention, which may better leverage the semantic information from the posture labels. The efficacy of rumor detection has therefore been greatly enhanced. In extensive tests conducted on two Twitter benchmark datasets, our rumor detection system surpassed updating models.

Key points from the research study are as follows: To detect social media rumors, a meta-learning approach is introduced that combines many modalities. Suggested technique utilizes meta-knowledge from posture and rumor detection tasks, which is higher level than typical Hyper Meta Network Approach learning systems that depend on lower-level similarities. The task-specific model can correctly get Cognitive Content with Hyper Meta Network representation post with meta-knowledge.

The attitude information from user answers to good use by using an attention method. Try to improve the suggested Cognitive Contents with Hyper Meta Network Meta-multitask learning even further by using the weighted user answers. Our method outperforms the current state-of-the-art, as shown by experiments on two publicly available benchmark datasets for Twitter rumor recognition tasks.

Advantages:

The Cognitive Contents with Hyper Meta Network post-embedding layer represents the Cognitive Contents with Hyper Meta Network content of each claim post as an embedding vector. We employ VGG19 for the visual content and BERT for the textual content to derive the embedding vector.

Hyper-Multitasking Layers: We provide a meta-layered approach to Hyper Meta Network Approach learning that integrates meta-information from posture recognition and rumor detection to improve social media rumor identification. Following instructions from the shared meta-network, the task-specific layer has the potential to acquire knowledge about the details of each post.

To make maximum use of the attitude information provided by the user's answers, an attention approach is applied for task-specific output layer. The rumour forecast may be believed depends on the clarity of the responses.

B. *Used Algorithm Description in the proposed method*

Steps involved are

Step 1: Registration of new users with their credentials.

Step 2: Verification of user details in the server.

Step 3: Validation and authentication of users.

Step 4: Users can log in for their social media login.

Step 5: Uploading of data sets – Excel Format.

Step 6: View of uploaded datasets.
Step 7: Finding post type by hash code to detect Rumour/non-rumour.
Step 8: Display Results.

C. Description of the system modeling, design, and development

Input Design

Input Design is an essential part of the software development life cycle; thus, developers should pay great attention to it. The program relies on reliable data, which is why input design is so important. Inputs must be efficiently built to reduce feeding errors. Software engineering principles should be followed while creating input screens or forms with validation controls for input range, limit, and related validations.

Almost every section of this system has its data entry screen. The purpose of error messages is to alert users whenever they make a mistake and provide those instructions on how to fix it, so they don't make the same mistakes again. Look at this closely from the perspective of the design of modules.

Input design is the process of making user-generated material computer-readable. Main goal of input design to provide logical and input with error free. Controlling input mistakes is the input design's job. The application was designed with the user in mind. Forms are designed to automatically position the cursor where data needs to be entered during processing. Multiple input fields provide a drop-down menu from which the user can select an appropriate value.

Validations must be applied to all data entries. When the user completes all the forms on this page and receives an error notice for each wrong submission, they can go on to the next page.

Output Design

How well a project manager can convey their ideas to stakeholders his team members or an administrator's ability to communicate effectively with clients are two of the main goals of the computer's output. The final product of the VPN system is a client management system that the project manager may utilize to recruit new clients and assign them to work.

On their behalf, monitoring the projects' status and on the user end, allocating folder-level access to clients according to their allocated projects. After one client's project is complete, they may be assigned a new one. We guarantee that user authentication procedures are in place from the start. While anybody may sign up for an account or the administrator can establish one, only the latter can verify the user's identity and give them a project to work on.

The software starts operating as soon as the first execution of it takes place. Starting the server is necessary. Before Internet Explorer may be used as a browser. On a local area network, this project can only run with one computer acting as the administrator and the others as clients. The system's intuitive design ensures that even first-time users will have no problem learning how to utilize it.

Modules

Social Server

Once the Service Provider has successfully logged into this module with their valid username and password, they will be able to access various operations such as: View All Datasets, View All Datasets by Hyper Meta Network Approach, View Post Score Results, View Post Rating Results, View Rumour Results, and Login.

View and Authorize Users

Here the administrator may see a complete rundown of all the registered users. Admins may see user information including name, email, and address, and they can also authorize users here.

End User

There are an equal number of users in this module. Registration is required before performing any operations. Details will be entered into the database after a user registers. He will need to log in using the permitted username and password when registration is successful. After logging in, users will be able to do things like see their profile, upload datasets, discover post types by hash code, and view all uploaded datasets.

IV. SYSTEM DESIGN AND DESCRIPTION

ARCHITECTURE

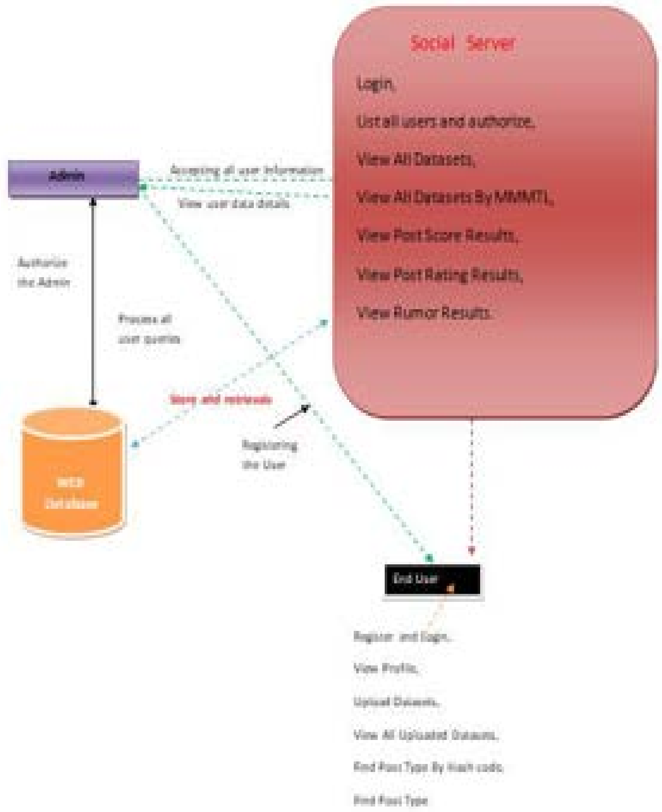


Fig. 1 Architectural Description

In Fig. 1, the architecture used in this proposed work admin, social server, end user and database were described in detail with their functionalities and roles performed by them in detail. As in Fig. 2 data flow diagram described by how the data will be flown between Social Server, System and End user with their relationship between them. In Fig. 3, the diagrammatic representation of steps performed by the end user is described in the form of flow chart and in Fig. 4 steps performed by social server were described.

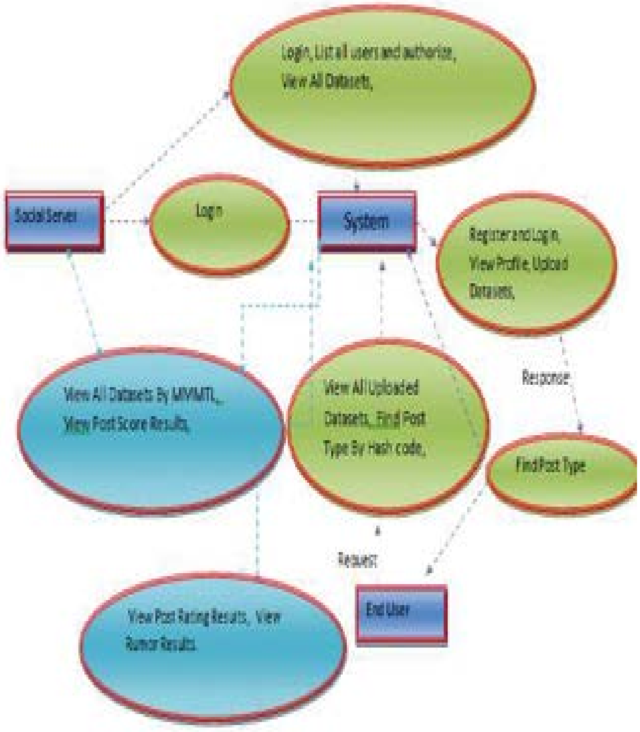


Fig. 2 Data Flow Diagram Description

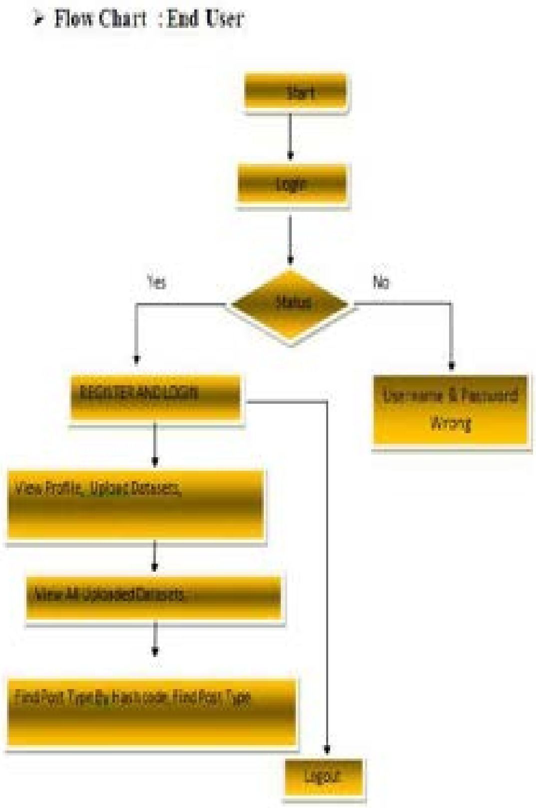


Fig. 3 Description of End User Flow Chart

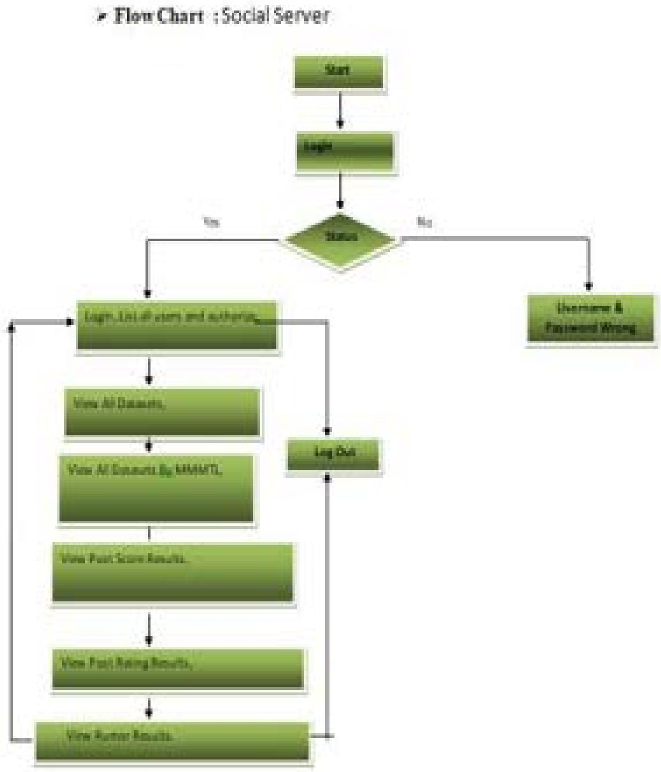


Fig. 4 Description of Social Server Flow Chart

V. RESULTS AND DISCUSSION

Description of got results with proposed method were described from Fig. 5 to Fig. 15 used with two different types of Twitter based datasets.

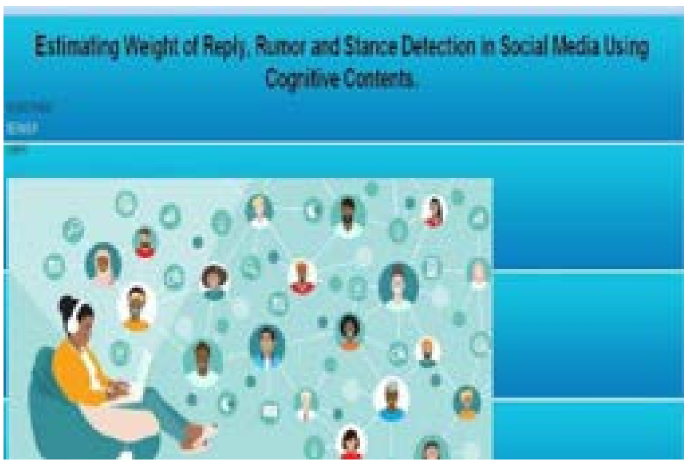


Fig. 5 Description of Home Page

In Fig. 5, represented the home page of the performed work which consists of two main components Server and User. Fig. 6, New User registration described with entering fields of user details and registered as a new user. Description of Verification of user details in the server is described in Fig. 7 with an example. In Fig. 8, description and validation of users were described. Uploading different types of datasets in excel form described in Fig. 9.

Fig. 6 Description of New User Registration

Fig. 9 Description of Uploaded Datasets

Fig. 7 Description of Verification of user details in the server

Fig. 8 Description of Validation and authentication of users

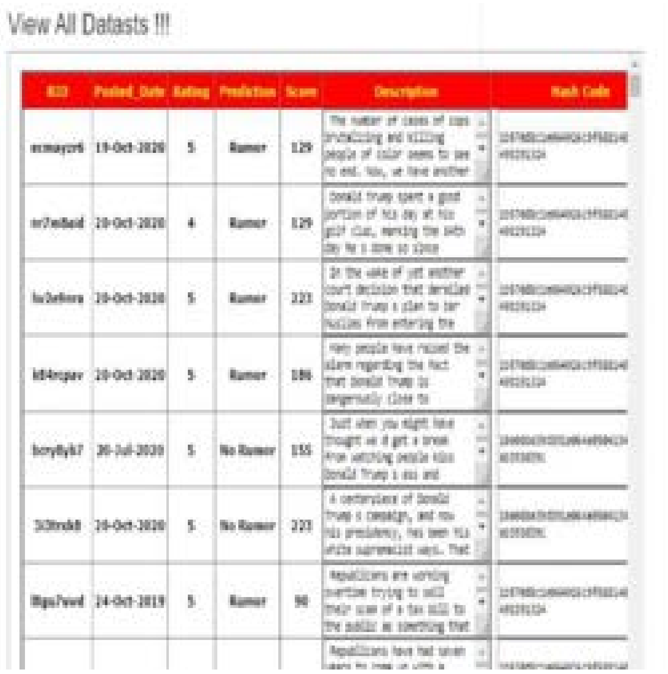


Fig. 10 View Sample Uploaded Datasets

Viewing different types of uploaded datasets in the server were described with registered ID, posted date, rating, prediction details rumor or not rumor, score obtained, description and hash tag details in Fig. 10 as an example dataset.

In Fig. 11, sample found trust type results by using hash tag were described as an example results of post type as non-Rumor along with details of registration ID, and posted date, description. In Fig. 12, sample found trust type results by using hash tag were described as an example results of post type as Rumor along with details of registration ID, and posted date, description.

Found Trust Type Results By HashCode!!!

RID	Post Date	Description	Post Type
bcry8yb7	20-Jul-2020	Just when you might have thought we d	No Rumor
3i3trxk8	20-Oct-2020	A centerpiece of Donald Trump s	No Rumor
cd2i54y8	20-Oct-2020	The media has been talking all day	No Rumor
h056t76ok	30-Nov-2020	Abigail Disney is an heiress with brass	No Rumor
lefrdvgs	21-Oct-2020	A new animatronic figure in the Hall	No Rumor
iorb4y3p	21-Oct-2020	Senate Majority Whip John Cornyn (R-TX)	No Rumor
cje77dva	21-Oct-2020	As a Democrat won a Senate seat in deep-	No Rumor

Fig. 11 View Sample Found Trust Type Results by Hash Code of "No Rumor" Category

Fig. 13 Description of View Rating Results

Found Trust Type Results By HashCode!!!

RID	Post Date	Description	Post Type
ecrnaytr6	19-Oct-2020	The number of cases of cops brutalizing	Rumor
nr7m8ad	20-Oct-2020	Donald Trump spent a good portion of his	Rumor
lw2e9nra	20-Oct-2020	In the wake of yet another court	Rumor
k84rqpau	20-Oct-2020	Many people have raised the alarm	Rumor
l8gu7owd	24-Oct-2019	Republicans are working overtime	Rumor
ky1b1yhu	20-Oct-2020	Republicans have had seven years to come	Rumor

Fig. 12 View Sample Found Trust Type Results by Hash Code of "Rumor" Category

Fig. 14 Description of View Post Score Results

Rating results were described in Fig. 13 by registration ID on X-axis and its rating values on Y-axis ranging from zero to five. Most of the IDs were rated as 5 where as one with rating as 4 and one as 3.

Post Score Results were represented in Fig. 14, by taking registration ID on X-axis and post score result values on Y-axis ranging from zero to three hundred.

In Fig. 15, Rumour results were represented by taking types of tweets as Rumour or non-rumour and Number of Tweets ranging from zero to one thousand five hundred on Y-axis. On observation of results of a particular type of dataset, it can be inferred that 1101 tweets were of non-rumour type and 224 were of Rumour type.

Fig. 15 Description of Rumor Results

VI. CONCLUSION AND FUTURE SCOPE

This research presents Hyper Meta Network Approach, a new approach to in Hoax detection, Poster detection, and reply weight estimation. Using the users' posture information answers in its detection is the goal to enhance rumour identification job performance. To extract meta-knowledge from Cognitive Contents with Hyper Meta Network postings, Hyper Meta Network uses frequently utilized higher-level meta-networks, unlike traditional Hyper Meta Network Approach learning algorithms. Hence, by dynamically building task-specific model parameters, all occupations utilize the shared meta-knowledge. In the wide attitude labels to make use of semantic info it can prioritize replies utilizing the attention mechanism. To demonstrate the suggested method yields first-rate results, employed multiple Twitter benchmark datasets. Along with other rumour-related jobs, the proposed Hyper Meta Network approach will soon have a reliability measurement job, which have been proved in this work. The suggested technique utilizes meta-knowledge from posture and rumour detection tasks, which is higher level than typical systems that depend on lower-level similarities. To get the most out of the attitude data in user answers, in this work used an attention strategy for Cognitive Content with Hyper Meta Network meta-task learning. One way to make the suggested Hyper Meta Network approach even more successful is to employ weighted user answers which also implemented.

As future work, since the used data is Cognitive Contents with Hyper Meta Network, there will be a problem with the feature level of the stance identification test and the rumor detection task, which can be reduced.

ACKNOWLEDGMENT

This work has no grants from either public or private grants. No conflict of interest between the authors and express thanks to all reviewers for their suggestions.

REFERENCES

[1] S. Kwon, M. Cha, K. Jung, W. Chen, and Y. Wang, "Prominent features of rumor propagation in online social media," in 2013 IEEE 13th International Conference on Data Mining, Dec. 2013, pp. 1103–1108.

[2] S. Qian, T. Zhang, C. Xu, and J. Shao, "Cognitive Contents with Hyper Meta Network event topic model for social event analysis," IEEE Trans. Multimedia, vol. 18, no. 2, pp. 233–246, 2016.

[3] J. Bian, Y. Yang, H. Zhang, and T. Chua, "Multimedia summarization for social events in microblog stream," IEEE Trans. Multimedia, vol. 17, no. 2, pp. 216–228, 2015.

[4] Q. You, L. Cao, Y. Cong, X. Zhang, and J. Luo, "A multifaceted approach to social multimedia-based prediction of elections," IEEE Trans. Multimedia, vol. 17, no. 12, pp. 2271–2280, 2015.

[5] X. Yang, T. Zhang, C. Xu, and M. S. Hossain, "Automatic visual concept learning for social event understanding," IEEE Trans. Multimedia, vol. 17, no. 3, pp. 346–358, 2015..

[6] Q. Fang, C. Xu, J. Sang, M. S. Hossain, and M. Ghulam, "Word-ofmouth understanding: Entity-centric Cognitive Contents with Hyper Meta Network aspect-opinion mining in social media," IEEE Trans. Multimedia, vol. 17, no. 12, pp. 2281–2296, 2015.

[7] G. Qi, W. Liu, C. C. Aggarwal, and T. S. Huang, "Joint intermodal and intramodal label transfers for extremely rare or unseen classes," IEEE Trans. Pattern Anal. Mach. Intell., vol. 39, no. 7, pp. 1360–1373, 2017. [Online]. Available: https://doi.org/10.1109/TPAMI.2016.2587643.

[8] P. Donovan, "How idle is idle talk? one hundred years of rumor research," Diogenes, vol. 54, no. 1, pp. 59–82, 2007.

[9] Z. Jin, J. Cao, Y. Zhang, J. Zhou, and Q. Tian, "Novel visual and statistical image features for microblogs news verification," IEEE Trans. Multimedia, vol. 19, no. 3, pp. 598–608, 2017.

[10] S. B. Parikh, S. R. Khedia, and P. K. Atrey, "A framework to detect fake tweet images on social media," in 2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM), 2019, pp. 104–110.

[11] M. A. Jamal and G. Qi, "Task agnostic meta-learning for few-shot learning," in IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019. Computer Vision Foundation / IEEE, 2019, pp. 11 719–11 727.

[12] A. Brock, T. Lim, J. M. Ritchie, and N. Weston, "SMASH: one-shot model architecture search through hyper networks," in 6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings. Open Review.net, 2018.

[13] Z. Liu, H. Mu, X. Zhang, Z. Guo, X. Yang, K. Cheng, and J. Sun, "Metapruning: Meta learning for automatic neural network channel pruning," in 2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019. IEEE, 2019, pp. 3295–3304.

[14] J. Chen, X. Qiu, P. Liu, and X. Huang, "Meta Hyper Meta Network Approachlearning for sequence modeling," in Proceedings of the Thirty-Second AAAI Conference on Artificial Intelligence, (AAAI-18), the 30th innovative Applications of Artificial Intelligence (IAAI-18), and the 8th AAAI Symposium on Educational Advances in Artificial Intelligence (EAAI18), New Orleans, Louisiana, USA, February 2-7, 2018. AAAI Press, 2018, pp. 5070–5077.

[15] J. Devlin, M. Chang, K. Lee, and K. Toutanova, "BERT: pre-training of deep bidirectional transformers for language understanding," in Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, NAACL-HLT 2019, Minneapolis, MN, USA, June 2-7, 2019, Volume 1 (Long and Short Papers). Association for Computational Linguistics, 2019, pp. 4171–4186.

Advancements and Challenges in Drone Technology: A Comprehensive Review

1st P.S.Agrawal, Assistant professor
Department of Electronics And Communication Engineering
Shri Ramdeobaba College of Engineering and Management
Nagpur,440013,India
agrawalps@rknec.ed

2nd P.S.Jawarkar, Dean
Career Development & Placement Cell
Shri Ramdeobaba College of Engineering and Management
Nagpur,440013,India
jawarkarps@rknec.edu

3r
d K.M.Dhakate
Department of Electronics And Communication Engineering
Shri Ramdeobaba College of Engineering and Management
Nagpur,440013,India
dhakatekm@rknec.edu

4th K. M. Parthani
Department of Electronics And Communication Engineering
Shri Ramdeobaba College of Engineering and Management
Nagpur,440013,India
parthanikm@rknec.edu

5t
h A. S. Agnihotri
Department of Electronics And Communication Engineering
Shri Ramdeobaba College of Engineering and Management
Nagpur,440013,India
agnihotrias@rknec.edu

***Abstract*—This research study analyzes the dynamic nature of drone technology, exploring its multifaceted applications, societal impacts, and the evolving challenges in its deployment. The integration of Unmanned Aerial Vehicles (UAVs) in various sectors, particularly in the fields of precision agriculture, cinematography, and geospatial mapping, is examined for its transformative potential. The paper highlights the revolutionary role of drones in addressing challenges such as search and rescue operations, natural disaster management, and providing virtual tourism experiences. Notably, the integration of Geographic Information System (GIS) technology and autonomous capabilities in drones is discussed in the context of cinematography and indoor inspections, emphasizing the need for a balance between technological innovation and ethical considerations. The comprehensive analysis includes insights into the security threat landscape associated with Augmented Reality (AR) drones, underlining the importance of conducting thorough threat analyses to ensure the safe operation of remotely piloted aircraft systems. Researchers have addressed concerns related to cyber threats, categorizing various aspects of the UAV threat landscape and developing classification approaches based on connections and nodes within flying ad-hoc networks (FANETs) and the Internet of Drones (IoD) infrastructure.**

***Index Terms*—Drone technology, Unmanned aerial vehicles applications, Geospatial Information Systems integration, Security threats, Autonomous capabilities, Citizen involvement.**

I. INTRODUCTION

In the ever-evolving realm of unmanned aerial vehicles (UAVs), colloquially known as drones, the fusion of cutting-edge technology with diverse applications has sparked a transformative wave across various industries. From precision agriculture to cinematography, geospatial mapping, and beyond, drones have not only revolutionized traditional practices but have also presented unique challenges that require thoughtful consideration. As these aerial marvels soar to new heights, a critical juncture emerges—balancing the limitless potential of technological innovation with ethical considerations and societal needs.

In the realm of cinematography, where drones offer unprecedented flexibility and shot diversity, a pressing challenge surfaces in the safe navigation of these devices while filming dynamic subjects in obstacle-laden environments. Simultaneously, the geospatial mapping sector grapples with the limitations of satellite-based localization, particularly within indoor spaces, necessitating inventive solutions to ensure precise data acquisition. Amidst these advancements, the expanding use of commercial drones has given rise to security concerns, introducing potential threats like jamming, spoofing, and denial of service attacks.

This comprehensive review paper endeavors to navigate through the intricate landscape of drone technology, shedding light on both its awe-inspiring advancements and the nuanced challenges that accompany its proliferation. From the intricacies of cinematographic applications to the dynamics of geospatial mapping, security threats, and beyond, our exploration aims to provide a holistic understanding of the current state of drone technology. By delving into these complexities, we aspire to contribute to the discourse surrounding responsible deployment, societal acceptance, and the delicate equilibrium between technological innovation and ethical considerations that define the future of drones.

II. CONCEPT OF DRONES: TRANSFORMING INDUSTRIES FOR GLOBAL PROGRESS

In the contemporary landscape, Unmanned Aerial Vehicles (UAVs), commonly known as drones, have evolved into pivotal assets across diverse industries, offering an array of applications that significantly benefit mankind.

III. CLASSIFICATION

This study has broadly categorized the drone based on its intended use, which includes major areas where it has experienced some enhancement or research

A. Virtual Tourism Drones

Fig. 1. Drone For Virtual Tourism

Virtual tourism drones as in Fig.1 are emerging as an innovative category of services within the tourism industry. Unmanned aerial vehicles (UAVs) are now employed not only for virtual tourism systems but also for monitoring and patrol activities, contributing significantly to safety and security measures [1]. These drones play a pivotal role in acquiring invaluable data from areas or objects that are otherwise in-accessible or pose potential dangers. This data serves as the foundation for creating detailed 3D models, facilitating the development of smart tourism services [2].

Augmented telepresence is gaining prominence as a technology that integrates immersive displays with augmented reality techniques. This combination is instrumental in the creation of advanced virtual tourism systems. By leveraging augmented telepresence, users can experience a virtual exploration of remote sites, accompanied by relevant information and con- textual details [3]. This technological integration enhances the overall tourism experience, providing users with a more immersive and informative journey.

The utilization of unmanned aerial vehicles and augmented telepresence techniques not only enhances the safety and security aspects of tourism but also opens up new dimensions for exploration and engagement. The data collected by virtual tourism drones contribute to the creation of detailed 3D models, while augmented telepresence adds a layer of interactivity and information, enriching the virtual tourism experience for users.

Augmented telepresence is gaining prominence as a technology that integrates immersive displays with augmented reality techniques. This combination is instrumental in the creation of advanced virtual tourism systems. By leveraging augmented telepresence, users can experience a virtual exploration of remote sites, accompanied by relevant information and con- textual details [3]. This technological integration enhances the overall tourism experience, providing users with a more immersive and informative journey.

The utilization of unmanned aerial vehicles and augmented telepresence techniques not only enhances the safety and security aspects of tourism but also opens up new dimensions for exploration and engagement. The data collected by virtual tourism drones contributes to the creation of detailed 3D models, while augmented telepresence adds a layer of interactivity and information, enriching the virtual tourism experience for users.

B. Search and Rescue drones

Fig. 2. Search and Rescue drones

The pivotal role of search and rescue drones shown in Fig. 2 becomes evident in their crucial contributions during natural disasters. These unmanned aerial vehicles (UAVs) play a significant role in efficiently locating and tracking targets of interest, thereby greatly enhancing the capabilities of situational awareness [4]. To address the challenges of operating in complex three-dimensional disaster environments, there has been a notable development in the formulation of multi-drone mission assignments and advanced path planning methods [5]. These advancements aim to optimize the utilization of drones in disaster scenarios, ensuring a more effective response.

Their primary role in locating and tracking, drone-based solutions has been proposed to address the critical issue of localizing missing individuals, a common challenge in the aftermath of natural disasters. This is achieved by leveraging the high penetration rate of mobile phones in society, offering a promising avenue for enhancing search and rescue efforts [6]. The integration of such technological solutions demonstrates the versatility of search and rescue drones in addressing various aspects of disaster management.

The indispensable role of search and rescue drones extends beyond their primary function of locating and tracking targets during natural disasters. The ongoing developments in multi-drone mission assignments, path planning methods, and innovative solutions for localizing missing individuals showcase the continuous evolution of these drones in enhancing disaster response capabilities. Their multifaceted applications under- score the significance of incorporating advanced technologies for more effective and comprehensive disaster management.

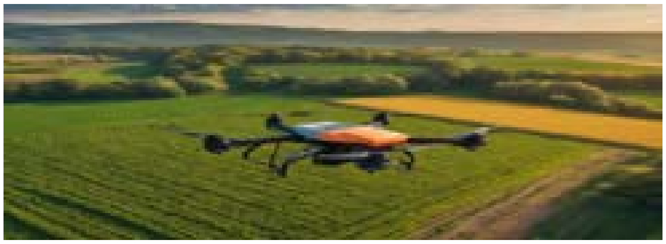


Fig. 3. Precision agriculture application drones

The integration of unmanned aerial vehicles (UAVs) into precision agriculture shown in Fig.3 witnessed a transformative impact on the agricultural landscape, ushering in a new era of technological advancements. The incorporation of UAVs, coupled with machine learning techniques, has become instrumental in various agricultural tasks, encompassing crop classification, detection of crops and weeds, cropland mapping, and field segmentation [7]. Despite the immense potential, the widespread adoption of precision agriculture technology faces challenges such as high costs, limited internet connectivity, and the complexities associated with setup [8]. These obstacles contribute to the current constraints on the broader implementation of precision agriculture practices.

In precision agriculture, one notable application that has gained popularity is plant disease detection, where UAVs play a pivotal role in automating the process. By utilizing drones, the agricultural sector can streamline and enhance disease detection mechanisms, thereby mitigating potential yield losses [9]. This application highlights the practical significance of UAVs in addressing specific challenges within precision agriculture, offering targeted solutions to improve overall agricultural productivity.

Integration of UAVs in precision agriculture represents a paradigm shift, leveraging advanced technologies to optimize various agricultural tasks. Despite the transformative potential, challenges such as cost considerations, connectivity limitations, and setup complexities continue to impede widespread adoption. Nonetheless, the ongoing exploration of applications like plant disease detection underscores the undeniable contribution of UAVs in overcoming specific challenges and fostering innovation within the realm of precision agriculture.

D. Drones in Engineering and Construction

The Drones have garnered considerable interest within the engineering and construction industry due to their prospective contributions to augment safety, efficiency, and precision in monitoring and inspection processes [10]. In this evolving landscape, they present novel avenues for data collection and analysis, deploying a diverse array of drones and sensors to

Fig. 4. Engineering and Construction drones

cater to different needs [11]. The integration of drones, along with cost-effective sensors and automated image processing techniques, has not only facilitated but also rendered remote-sensing approaches both feasible and compelling for various construction management tasks [12].

The utilization of drones in the engineering shown in Fig. 4 and the construction sector has opened up a realm of possibilities, offering unprecedented advantages in terms of enhanced safety measures, streamlined operational efficiency, and increased accuracy in monitoring and inspection. The versatility of drones is further underscored by their compatibility with an array of sensors, contributing to the industry's ability to adapt to specific requirements and challenges. This integration of technology is a transformative step forward in construction practices, reshaping traditional approaches and ushering in a new era of innovation.

The affordability of sensors and the incorporation of auto- mated image-processing techniques synergize seamlessly with drone technology, making remote-sensing approaches not only viable but also enticing for construction management tasks. The amalgamation of these technologies is poised to revolutionize how the industry addresses challenges, introducing efficiencies that were previously unattainable.

E. Geo mapping with unmanned aerial vehicle

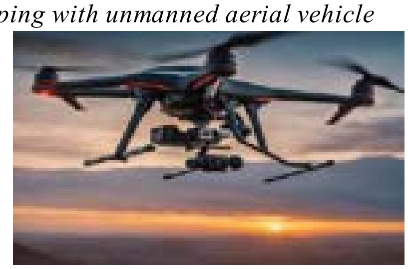


Fig. 5. Geo-mapping drones

Unmanned aerial vehicles (UAVs) have undeniably emerged as highly effective tools for geospatial mapping applications, showcasing their prowess in capturing high-resolution images essential for the ge=neration of precise maps, including Digital Elevation Models (DEMs) and Digital Orthophotos [13]. The evolution of technology has played a pivotal role in enhancing the spatial accuracy of UAVs, with notable strides seen in the implementation of real-time kinematic (RTK) and post- processing kinematic (PPK) techniques [14]. These advancements not only elevate the quality of data acquisition but also contribute to the refinement of spatial accuracy, making UAVs increasingly reliable for geospatial mapping endeavors.

A noteworthy synergy arises when data collected by UAVs is combined with information obtained through terrestrial laser scanning (TLS). This collaborative approach has demonstrated the capacity to yield exceptionally accurate 2D and 3D data, effectively enhancing mapping capabilities and facilitating comprehensive risk assessments [15]. By integrating UAVs with TLS, the geospatial mapping process reaches new heights of precision and thoroughness, enabling more robust applications in various domains.

The multifaceted advantages of UAVs in geospatial mapping shown in Fig.5 have been significantly augmented by technological innovations, particularly in the realm of spatial accuracy enhancement through RTK and PPK techniques. Furthermore, the synergistic combination of UAV and TLS data sets a new standard for accuracy and completeness in mapping and risk assessment endeavors. As these technologies continue to evolve, the collaborative potential of UAVs and TLS promises to redefine the landscape of geospatial mapping, unlocking unprecedented possibilities for detailed and reliable data acquisition.

F. Drone-integrated Geographic Information System technology

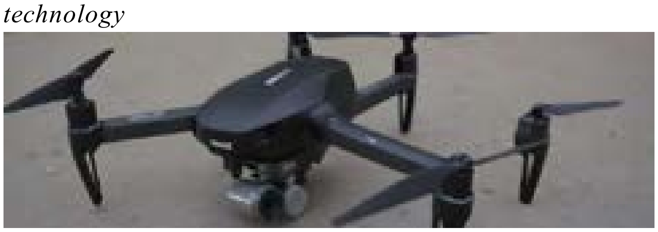


Fig. 6. Geographic Information System technology drones

Exploration of drone-integrated Geographic Information System (GIS) technology shown in Fig.6 is a prevalent theme across various scholarly works. Williams-Kelly extensively delves into the imperative need for sustainable solutions that seamlessly incorporate unmanned aircraft systems (UAS) into smart city transportation systems. This encompasses the development of autonomous trajectory-tracking control systems and the establishment of safe drone-following mechanisms, thereby contributing to the evolution of efficient and secure UAS integration within urban environments [16].

In a complementary study, Lercel and Hupy undertake a comprehensive analysis of spatial patterns among registered airmen through GIS methodologies. Their research aims to identify regions with heightened risk factors, enhancing safety management in the context of unmanned aerial systems sightings and airspace categories [17]. The utilization of GIS in this context proves instrumental in discerning patterns and spatial relationships, providing valuable insights to refine safety protocols and mitigate potential risks associated with UAS operations.

In another scholarly contribution, Ballesteros, Sanchez-Torres, and Branch outline a meticulous workflow leveraging GIS tools to generate GeoAI datasets specifically tailored for object detection and semantic segmentation models using drone imagery. Their research not only delves into the technicalities of the workflow but also elucidates methods for data augmentation and fusion, thereby presenting a comprehensive approach to harnessing GIS technology for sophisticated geospatial data analysis in the realm of UAS operations [18].

G. Drone Entertainment and Aerial Photography

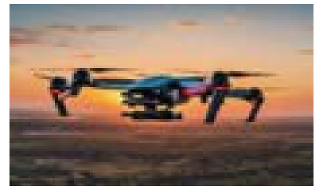


Fig. 7. Entertainment and Arial Photography drones

The realm of cinematography and aerial photography is undergoing a transformative revolution propelled by Unmanned Aerial Vehicles (UAVs) as in Fig 7. These drones bring a wealth of possibilities to the creative process, offering unparalleled flexibility in shot setups, access to otherwise challenging or inaccessible spaces, and the ability to capture novel shot types that were once unimaginable [19]. Despite these advancements, the task of safely piloting a drone while filming a moving target in environments filled with obstacles remains a formidable challenge, often necessitating the involvement of multiple expert human operators [20].

In response to this challenge, researchers have proposed innovative solutions in the form of autonomous aerial videography systems. These systems go beyond conventional manual control, introducing intelligent frameworks that seamlessly integrate customized shots and account for the dynamic nature of drones [21]. By leveraging autonomous capabilities, these systems aim to enhance the efficiency of aerial filming, minimizing the need for constant human intervention and ad- dressing the complexities associated with maneuvering drones in dynamic and obstacle-laden environments.

The integration of autonomous aerial videography systems represents a paradigm shift in the field, promising not only to streamline the filmmaking process but also to elevate the overall quality and safety of aerial cinematography. These systems have the potential to unlock new dimensions of creativity, allowing filmmakers to focus more on the artistic aspects of their craft rather than grappling with the intricacies of drone piloting. As technology continues to advance, the marriage of UAVs and autonomous

H. Societal impact of Commercial drone

The advent of commercial drones holds the promise of significant societal impacts, offering the potential to replace humans in hazardous environments and overcome geographical distances, thereby providing valuable services and economic benefits [22] [23]. Despite these promising prospects, there are legitimate concerns regarding the potential infringement

Fig. 8. Photograph by Commercial Drone

on public values, particularly in areas such as safety, privacy, and accountability [24].

It is crucial to acknowledge that the deployment of commercial drones is a double-edged sword, carrying both challenges and opportunities. On the positive side, drones can play a pivotal role in societal development, particularly in areas like in- formal settlement mapping. In this context, drones can provide detailed geospatial information, aligning with values such as participation, empowerment, accountability, transparency, and equity. By leveraging these capabilities, drones can contribute positively to societal development, especially in regions where access to such information is limited.

To ensure the seamless and responsible integration of commercial drone services into society, it is imperative to consider the voice of the citizens in the design and implementation processes. Addressing the concerns of the public, including issues related to safety, privacy, and accountability, is essential for fostering a positive reception of drone technologies. Engaging with the community and incorporating their perspectives not only enhances the ethical aspects of drone deployment but also establishes a foundation for trust and acceptance.

I. Indoor Inspection Drone

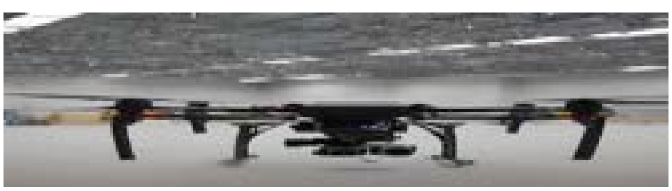


Fig. 9. Indoor Inspection Drones

The intricate process of developing indoor inspection drones encompasses the mastery of complexity across diverse facets. One prominent challenge arises from the necessity of incorporating Building Information Modeling and indoor navigation methods since traditional satellite-based localization proves ineffective within indoor environments [25]. Overcoming this hurdle involves the implementation of innovative solutions to ensure precise navigation and data acquisition within confined indoor spaces.

Another significant challenge in the realm of indoor inspection drone development pertains to the intricate task of path planning for close-distance inspections of intricate structures like bridges and wind turbines. This demands the integration of sophisticated algorithms capable of navigating and charting optimal paths, all while adeptly avoiding potential collisions [26]. Addressing this challenge is pivotal for ensuring the safety and efficiency of the inspection process, requiring a meticulous approach to algorithmic development and implementation.

Moreover, the imperative need for lightweight drones introduces another layer of complexity to the development process. Achieving successful navigation through narrow spaces within indoor environments necessitates the incorporation of drones that are not only lightweight but also possess enhanced maneuverability capabilities [27]. This requirement underscores the significance of technological advancements in drone design and engineering, emphasizing the need for a delicate balance between functionality and agility.

J. Security threat analysis with AR drones

Fig. 10. Security threat analysis with AR Drone

Conducting security threat analysis with Augmented Reality (AR) drones stands as a crucial imperative in ensuring the safe and secure operation of remotely piloted aircraft systems (RPAS) [28]. The escalating utilization of drones across various applications has concurrently given rise to an increased susceptibility to cyber threats, encompassing issues such as jamming, spoofing, and denial of service attacks [29]. In response to these evolving challenges, researchers have undertaken comprehensive surveys and literature reviews to systematically categorize the diverse facets of the Unmanned Aerial Vehicle (UAV) threat landscape. Furthermore, they have endeavored to develop classification approaches that are anchored in distinct types of connections and nodes within the framework of flying ad-hoc networks (FANETs) and the Internet of Drones (IoD) infrastructure [30].

The landscape of security threat analysis in the context of AR drones is intricate, demanding a nuanced understand- ing of the potential vulnerabilities and risks associated with their operational environments. Cyber threats pose significant concerns, and the classification approaches developed by researchers offer valuable insights into the multifaceted dimensions of UAV threat landscapes. The categorization based on different types of connections and nodes within FANETs and IoD infrastructure is particularly instrumental in devising comprehensive strategies to mitigate potential risks and enhance the overall cybersecurity posture of AR drones.

IV. PROPOSED IDEA

The increasing demand for drones has ignited the widespread integration of diverse technologies aimed at crafting drones that serve both humanity and the environment. In response to the prevailing global circumstances, a discernible trend is poised to emerge, centering on the advancement of drones tailored specifically for cleaning and disinfection tasks. This projected expansion in the drone sector is anticipated to lead to the emergence of a distinct category termed the sanitizing drone.

The sanitizing drone is envisioned to be equipped with a reservoir capable of containing a liquid sanitizer, ranging from 1 to 10 liters, and is estimated to weigh up to 8.8kg, with the Agras MG-1 being identified as a potential model. This drone is purposefully designed to navigate and operate in areas that may pose challenges or prove inaccessible for humans, thereby augmenting its efficacy across various environments.

A noteworthy augmentation to the sanitizing drone could be its capacity to clean windows on buildings. This functionality can be seamlessly integrated by introducing a knob to the drone's transmitter. Upon activation, the knob communicates with the receiver, prompting the motor pump to dispense either the sanitizer or, with the inclusion of a window wiper attachment, a cleaning solution. To enhance versatility, alternative attachments such as a sponge or cleaning pad can be considered for the sprayer. The sprayer remains connected to a hose, ensuring a continuous water supply for the cleaning process. Through the incorporation of a rolling mechanism, such as a roller or wheel, the sponge adeptly cleans the glass surface.

An innovative facet of this sanitizing drone is its reliance on an external power source, providing an autonomous and protracted power supply. This approach not only extends flight durations but also diminishes the overall weight of the drone by obviating the need for onboard batteries.

Regarding reliability, ensuring consistent performance necessitates meticulous attention to design, rigorous testing protocols, and continuous monitoring of operational parameters. Employing redundant systems, implementing fail-safe mechanisms, and conducting thorough maintenance checks are imperative in fostering reliability across various operational scenarios.

In terms of security, safeguarding the drone against unauthorized access, cyber threats, and potential misuse entails the implementation of robust encryption protocols, authentication mechanisms, and intrusion detection systems. Additionally, adherence to stringent regulatory frameworks, comprehensive risk assessments, and ongoing security audits are essential in fortifying the overall security posture of the sanitizing drone ecosystem.

V. CONCLUSION

Drones have indeed emerged as remarkable technological innovations, offering vast opportunities for enhancing human life and environmental sustainability. The rapid global proliferation of drones across various sectors signifies a significant shift unfolding at an accelerated pace. What once captivated enthusiasts' imagination has now seamlessly integrated into our technological fabric, permeating daily life and attracting keen interest from numerous enterprises eager to leverage their capabilities.

The accessibility of drone kits, coupled with user-friendly programming languages and abundant online resources, has democratized the process of drone development. This inclusivity has democratized the drone space, empowering novices to contribute to its growth.

However, every technological stride confronts its share of challenges. The success of drones, akin to any innovation, hinges on the adept management of associated issues and regulations. Governments, in particular, bear a pivotal role in enforcing low-cost detection systems capable of identifying rogue drones and enforcing stringent regulations to curb misuse. Drone manufacturers, too, shoulder responsibility in ensuring the judicious use of this technology. Identifying potential shortcomings and devising robust solutions is imperative. Rigorous testing for technological vulnerabilities, even amidst the integration of advanced features such as artificial intelligence, sensors, and cameras, is indispensable to upholding safety standards. This proactive stance can markedly mitigate incidents such as drone collisions and disappearances.

REFERENCES

[1] Antonio Albanese, Vincenzo Sciancalepore, and Xavier Costa-Pe´rez. Sardo: An automated search-and-rescue drone-based solution for victims localization. *IEEE Transactions on Mobile Computing*, 21(9):3312– 3325, 2021.

[2] Rui Carlos Josino Alexandre, Luiz Eduardo Galva˜o Martins, and Tony Gorschek. Cybersecurity risk assessment for medium-risk drones: A systematic literature review. *IEEE Aerospace and Electronic Systems Magazine*, 2023.

[3] Sadia S Ali, Rajbir Kaur, and Haidar Abbas. Feasibility of drones as the next step in innovative solution for emerging society. *Artificial Intelligence in Industry 4.0 and 5G Technology*, pages 179–208, 2022.

[4] Agustin Bernier-Vega, Kyle Barton, Isaac Olson, Juan Rodriguez, Genesis Cantu, and Selahattin Ozcelik. Remote data acquisition using uavs and custom sensor node technology. *Drones*, 7(6):340, 2023.

[5] Rogerio Bonatti, Wenshan Wang, Cherie Ho, Aayush Ahuja, Mirko Gschwindt, Efe Camci, Erdal Kayacan, Sanjiban Choudhury, and Sebastian Scherer. Autonomous aerial cinematography in unstructured environments with learned artistic decision-making. *Journal of Field Robotics*, 37(4):606–641, 2020.

[6] Ruben Chin, Cagatay Catal, and Ayalew Kassahun. Plant disease detection using drones in precision agriculture. *Precision Agriculture*, 24(5):1663–1682, 2023.

[7] Alfred Colpaert. Satellite and uav platforms, remote sensing for geographic information systems, 2022.

[8] Resul C¸ o¨mert, Samed O¨ zdemir, Burhan Baha Bilgilioglu, Selc¸uk Alemdag, and Halil Ibrahim Zeybek. since 1961. *Baltica*, 36(1):37– 50, 2023.

[9] Ryosuke Kakiuchi, Dinh Tuan Tran, and Joo-Ho Lee. Evaluation of human behaviour detection and interaction with information projection for drone-based night-time security. *Drones*, 7(5):307, 2023.

[10] Damon J Lercel and Joseph P Hupy. Exploring the use of geographic information systems to identify spatial patterns of remote uas pilots and possible national airspace risk. *Safety*, 9(1):18, 2023.

[11] Qingge Li, Xiaogang Yang, Ruitao Lu, Jiwei Fan, Siyu Wang, and Zhen Qin. Visionice: air–ground integrated intelligent cognition visual enhancement system based on a uav. *Drones*, 7(4):268, 2023.

[12] Han Liang, Seong-Cheol Lee, Woosung Bae, Jeongyun Kim, and Suyoung Seo. Towards uavs in construction: advancements, challenges, and future directions for monitoring and inspection. *Drones*, 7(3):202, 2023.

[13] Ioannis Mademlis, Arturo Torres-Gonza´lez, Jesu´s Capita´n, Maurizio Montagnuolo, Alberto Messina, Fulvio Negro, Cedric Le Barz, Tiago Gonc¸alves, Rita Cunha, Bruno Guerreiro, et al. A multiple-uav ar- chitecture for autonomous media production. *Multimedia Tools and Applications*, 82(2):1905–1934, 2023.

[14] Seojin Moon and Jongho Ock. Developing the framework of drone curriculum to educate the drone beginners in the korean construction industry. *Drones*, 7(6):356, 2023.

[15] Dung D Nguyen, Omar Alharasees, Utku Kale, Munevver Ugur, and Tahir Hikmet Karakoc. "drones gis system" in urban transport. In *International Symposium on Energy Management and Sustainability*, pages 733–741. Springer, 2022.

[16] Fumio Okura, Masayuki Kanbara, and Naokazu Yokoya. Fly-through heijo palace site: augmented telepresence using aerial omnidirectional videos. In *ACM SIGGRAPH 2011 Posters*, pages 1–1. 2011.

[17] Eric Razafimahazo, Pierre De Saqui-Sannes, Rob A Vingerhoeds, Claude Baron, Julien Soulax, and Romain Me`ge. Mastering complexity for indoor inspection drone development. In *2021 IEEE International Symposium on Systems Engineering (ISSE)*, pages 1–8. IEEE, 2021.

[18] Zhexiong Shang and Zhigang Shen. Indoor testing and simulation platform for close-distance visual inspection of complex structures using micro quadrotor uav. *arXiv preprint arXiv:1904.05271*, 2019.

[19] Boris Shishkov and Magdalena Garvanova. The societal impacts of drones: a public values perspective. In *International Conference on Telecommunications and Remote Sensing*, pages 61–71. Springer, 2022.

[20] Byung Duk Song and Young Dae Ko. Quantitative approaches for economic use of emerging technology in the tourism industry: Unmanned aerial vehicle systems. *Asia Pacific Journal of Tourism Research*, 22(12):1207–1220, 2017.

[21] Tomasz Templin and Dariusz Popielarczyk. The use of low-cost unmanned aerial vehicles in the process of building models for cultural tourism, 3d web and augmented/mixed reality applications. *Sensors*, 20(19):5457, 2020.

[22] Kai-Yun Tsao, Thomas Girdler, and Vassilios G Vassilakis. A survey of cyber security threats and solutions for uav communications and flying ad-hoc networks. *Ad Hoc Networks*, 133:102894, 2022.

[23] Vaishnavi Upadrasta, Julia Hamdan, Rodney Leitner, and Harald Kolrep. Who are the stakeholders of drone use? roles, benefits, risk perceptions, and solutions. In *Human Interaction, Emerging Technologies and Future Systems V: Proceedings of the 5th International Virtual Conference on Human Interaction and Emerging Technologies, IHIET 2021, August 27-29, 2021 and the 6th IHIET: Future Systems (IHIET-FS 2021), October 28-30, 2021, France*, pages 572–579. Springer, 2022.

[24] Tao Xiong, Fang Liu, Haoting Liu, Jianyue Ge, Hao Li, Kai Ding, and Qing Li. Multi-drone optimal mission assignment and 3d path planning for disaster rescue. *Drones*, 7(6):394, 2023.

[25] Mustafa Zeybek, Selim Tas¸kaya, Ismail Elkhrachy, and Paolo Tarolli. Improving the spatial accuracy of uav platforms using direct georeferencing methods: an application for steep slopes. *Remote Sensing*, 15(10):2700, 2023.

[26] Bin Zhang, Shuto Tsuchiya, and Hunok Lim. Development of a lightweight octocopter drone for monitoring complex indoor environment. In *2021 6th Asia-Pacific Conference on Intelligent Robot Systems (ACIRS)*, pages 1–5. Ieee, 2021.

[27] Su Zhang, Susan M Bogus, Christopher D Lippitt, Vineet Kamat, and SangHyun Lee. Implementing remote-sensing methodologies for construction research: An unoccupied airborne system perspective. *Journal of Construction Engineering and Management*, 148(9):03122005, 2022.

[28] Zhiwei Zhang, Yuhang Zhong, Junlong Guo, Qianhao Wang, Chao Xu, and Fei Gao. Auto filmer: Autonomous aerial videography under human interaction. *IEEE Robotics and Automation Letters*, 8(2):784–791, 2022.

[29] MF Zolkepli, MF Ishak, FNA Zolfan, NS Yusoff, S Daud, and NS Mohd Zin. The implementation of unmanned aerial vehicle (uav) for slope mapping: Unmanned aerial vehicle (uav) for slope mapping. *Interna- tional Journal of Engineering Technology and Sciences*, 10(1):16–23, 2023.

[30] Imran Zualkernan, Diaa Addeen Abuhani, Maya Haj Hussain, Jowaria Khan, and Mohamed ElMohandes. Machine learning for precision agriculture using imagery from unmanned aerial vehicles (uavs): A survey. *Drones*, 7(6):382, 2023.

A Double Tap Key based Security Model in Cloud Applications

Mrs.M.Prasanna Lakshmi
Asst. Professor, Department of Computer Applications, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India, mundlamuri70@gmail.com

J. Venkata Yaswanth
Student, Department of Computer Applications, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India, yaswanthjampani6341@gmail.com

B. Sowbhagyalakshmi
Student, Department of Computer Applications, Velagapudi Ramakrishna Siddhartha Engineering College, Vijayawada, Andhra Pradesh, India, sowbhagyalakshmi918@gmail.com

Abstract*— Security becomes more complicated for the data stored in cloud applications. Cloud computing provides data storage security as a service. In cloud computing, security should be provided to the users based on the data access controlled by the cloud service provider (CSP). To improve the security of cloud applications, the study presents a unique security model based on a double-tap key mechanism in this research. The suggested paradigm combines device verification with user authentication to allow access to cloud resources. In addition to providing their login credentials, users must authenticate themselves using a double-tap key produced by their registered device as a secondary verification step. This double-tap (Pseudorandom Number Generator (PRNG) and RSA (Rivest-Shamir-Adleman) Algorithm) key is highly safe and impervious to unwanted access since it is created using a unique method that combines user credentials with device-specific data. The proposed security model also incorporates multi-factor authentication, adding an extra layer of defense against potential intrusions. Depending on the sensitivity of the data or application being accessed, users may be required to verify their identity using additional methods, such as biometric recognition or one-time pass codes, in addition to the double-tap key. The proposed model's effectiveness is validated through comprehensive simulations and tests, demonstrating its ability to thwart a range of common attack vectors, including credential stuffing and brute force attacks. Its scalability and lightweight design enable deployment in large-scale cloud systems without any overhead. Finally, the double-tap keys, proposed security model offer a robust and efficient solution to enhance cloud application security, mitigating the risks of unauthorized access and data breaches. Combining device verification and user authentication provides a comprehensive defense against the ever-evolving cyber security threats in the cloud computing environment.

Keywords*— Cloud Computing, Security, Attacks, Cloud Service Providers (CSPs).

I. INTRODUCTION

Cloud computing security is a major problem, and several safeguards are put in place to guarantee the privacy, availability, and integrity of data and services. Protecting data, apps, and infrastructure housed in cloud environments from different attacks and vulnerabilities is a crucial part of cloud computing security. As more and more businesses use cloud services for data processing, storage, and management, it is critical to make sure strong security measures are in place [1]. Organizations are relieved of the burden of managing and maintaining physical infrastructure with the remote access and provision of these services by third-party suppliers [2].

Encryption and decryption play crucial roles in ensuring the security and confidentiality of data in cloud computing environments [3]. Strong encryption techniques must be implemented in order to safeguard sensitive data from unauthorized access as businesses depend more and more on cloud services to handle and store their data [4]. The process of transforming readable, plain data into an unreadable format with the use of cryptographic keys and algorithms is called encryption. The purpose of encryption is to secure data during storage, transmission, and processing. Using decryption keys, encrypted data can be restored to its original, readable format through the process of decryption. The data cannot be decrypted by anybody other than authorized parties with the right keys. In order to guarantee that only authorized users can access decrypted data, decryption is frequently linked to appropriate authentication and permission procedures [5].

The emergence of cloud computing as a revolutionary technology has completely changed how computer resources are supplied, accessed, and maintained. Its impact extends across various industries, and one of its crucial roles is in enabling real-time applications. Real-time applications demand immediate and seamless processing of data, making cloud computing an ideal platform to meet these requirements. In this context, the role of cloud computing in real-time applications is pivotal, providing a scalable, flexible, and efficient infrastructure to support the dynamic needs of modern applications [6]. Cloud computing allows real-time applications to scale resources up or down based on demand. This dynamic scalability ensures that applications can handle varying workloads without compromising performance. It is a sudden surge in users or data processing requirements, the cloud provides the necessary resources instantly. Real-time applications often require high availability to ensure uninterrupted services. Cloud providers offer redundant data centers and distributed infrastructure, reducing the risk of downtime. This ensures that real-time applications can deliver consistent performance and reliability, meeting the expectations of users. Cloud storage services enable real-time applications to store and retrieve data rapidly. The scalable and distributed nature of cloud storage ensures that applications can efficiently manage large volumes of real-time data. This is particularly crucial for applications that rely on instant access to up-to-date information. Cloud computing platforms provide access to powerful computing resources that can process data at high speeds. Real-time applications, such as those in finance, healthcare, or IoT, require rapid data processing. Cloud services offer the computational capabilities necessary to analyze and respond to data in real time [8]. Cloud computing facilitates distributed computing, allowing real-time applications to process data across multiple locations. By bringing computation closer to the data source, edge computing, an outgrowth of cloud computing lowers latency. Applications like augmented reality and driverless cars, where minimal latency is essential, would especially benefit from this. Cloud platforms offer a variety of services, such as databases, machine learning, and analytics, which can be seamlessly integrated into real-time applications. This integration allows developers to leverage pre-built services, accelerating the development of features like real-time analytics, predictions, and insights. Finally, this research study consists of various cloud computing algorithms and its challenges.

A. Cloud Storage Application

An application for cloud storage is a software service that lets users save, organize, and retrieve their data online rather than locally on hard discs or flash drives. Typically, these apps have functions like sharing, syncing, downloading, uploading, and occasionally modifying files. To the cloud storage platform, users can upload a variety of file kinds, including documents, photos, videos, and more. The service provider maintains secure storage for these files in data centers. Applications for cloud storage frequently have synchronization features, which guarantee that data saved on one device is automatically updated and available from other devices connected to the same account [19]. This functionality is beneficial for preserving consistency when using several devices. Users can collaborate on documents, projects, and presentations by sharing files and folders with other users. It is possible to manage who can access shared files and who may view, edit, and download them. Cloud storage apps utilize a variety of security mechanisms, including access limits, authentication, and encryption, to safeguard user data. This contributes to the privacy and security of critical information. Versioning is a feature of many cloud storage providers that lets customers retrieve older versions of their content. This capability may come in handy in the event of accidental modifications or file corruption. Applications for cloud storage are frequently compatible with a wide variety of hardware and operating systems, such as tablets, smart-phones, laptops, and desktop PCs. Thanks to this, users can now view their files from any location with an internet connection. Scalable storage choices are a common feature of cloud storage solutions, enabling users to adjust their storage capacity as needed. Users can change their storage needs over time because of this scalability.

B. Security in Cloud Computing

A vital component of cloud computing is security, which includes various tools, guidelines, and procedures intended to safeguard information, programs, and equipment housed in cloud environments. As businesses increasingly adopt cloud services, understanding and implementing robust security measures become paramount to mitigate various risks. Under a shared responsibility approach, cloud service providers (CSPs) usually protect the underlying cloud infrastructure. Customers are also in charge of protecting their identities, apps, data, and access controls. The key to protecting sensitive data in the cloud is data encryption during transmission and while it's at rest. Robust encryption mechanisms help prevent unauthorized access to data, even if the underlying infrastructure is compromised. Implementing robust IAM policies ensures that only authorized individuals or systems can access resources in the cloud. It involves user authentication, authorization, and management of permissions based on roles and responsibilities. To avoid malware, illegal access, and other dangers, cloud networks should be secured with firewalls, virtual private networks (VPNs), and intrusion detection and prevention systems (IDS/IPS). Cloud environments must be continuously monitored to identify and address security incidents quickly. Log analysis, network traffic monitoring, security information, and event management (SIEM) solutions are all part of it. Robust incident response protocols and disaster recovery systems assist companies in reducing the effects of service disruptions and security breaches. It comprises incident response protocols, failover systems, and routine data backups.

Raghunadha Reddi Dornala [9] introduced ensemble security architecture, a unique approach that combines several security layers, such as access control, authentication, encryption, and anomaly detection. This innovative strategy significantly improves the resilience of edge computing systems against a variety of cyber threats, such as malware assaults, insider threats, and data breaches. By utilizing a broad range of security methods, our approach enhances security and reduces the possibility of single-point failures, thereby enhancing the overall security posture of edge settings.

R. R. Dornala [10] proposed load-balancing technique is revolutionary for creating sophisticated multi-model cloud services. By intelligently distributing incoming requests across several servers in the cloud environment, the system can effectively minimize response times, maximize throughput, and prevent overload situations. This dynamic distribution of incoming traffic based on server load and capacity improves performance and guarantees high availability. The system also uses a multi-model design that supports relational, NoSQL, and graph databases, allowing companies to implement a variety of applications without being limited by a particular data model. Detailed simulations and tests using realistic workload scenarios have shown that the integrated load-balancing algorithms greatly enhance cloud service performance and scalability while retaining high availability and dependability.

II. Literature Survey

Tang et al. [11] aim to investigate and assess data security technologies regarding big data cloud computing. The study starts by giving a general review of big data and cloud computing technologies' rapid development and uptake. It highlights the benefits of these technologies, such as scalability, cost-effectiveness, and advanced analytics capabilities. However, the study acknowledges the inherent security challenges posed by the dynamic and distributed nature of cloud environments, coupled with the vast volumes of data being processed and stored. The primary focus of this research is on identifying and analyzing the key data security technologies employed in big data cloud computing environments. The basic security measures examined are encryption, intrusion detection, access controls, and authentication systems. In addition, the study looks into new technologies and evaluates how well they work to secure big data in the cloud, including safe multi-party computation and homomorphic encryption. Wang et al. [12] aims to explore and analyze several concepts of data security in Cloud big data. It also highlights the significant advantages and challenges associated with this technology. Emphasis is placed on the massive volumes, diverse formats, and high velocity of data generated, processed, and stored in cloud environments, necessitating robust security measures. The study delves into the unique security challenges posed by the integration of big data and cloud computing. This includes issues related to data privacy, confidentiality, integrity, and availability. The research also investigates the role of encryption, access controls, and authentication mechanisms in safeguarding data within the cloud infrastructure. An extensive analysis of the taxonomy of data security threats in cloud computing is provided by Farsi Mohammed et al. [13]. The taxonomy is designed to classify and examine risks that could jeopardize the availability, confidentiality, and integrity of data in cloud systems. The study explores both traditional security concerns that have transcended into the cloud and novel threats that are specific to the cloud computing paradigm. Each layer introduces unique security considerations, and this taxonomy aims to provide a holistic understanding of the diverse threats at play in cloud ecosystems. Furthermore, the research study discusses the evolving nature of security threats, considering the advancements in technology and the changing landscape of cyber threats. In addition to identifying and categorizing threats, the research study explores existing and emerging mitigation strategies and best practices to enhance data security in the cloud. This includes encryption techniques, access controls, authentication mechanisms, and incident response frameworks. Zhiying Wang et al. [14] aim to provide insights into the practical implications and benefits of integrating business analytics tools and techniques into cloud security strategies. Using a combination of techniques, the study combines qualitative and quantitative analysis to obtain a thorough understanding of the connections between the efficacy of cloud data security management and business analytics affordances. Quantitative data is collected through surveys distributed to a diverse sample of organizations utilizing cloud computing services. On the other hand, qualitative data is gathered through in-depth interviews with IT professionals, security experts, and business leaders to obtain nuanced insights into their experiences and perspectives. Fuguang Yao [15] investigates the design, implementation, and potential benefits of a Campus Network Cloud Storage Open Platform. The platform leverages the power of cloud computing to provide on-demand, ubiquitous access to storage resources while harnessing big data technologies to optimize data management, analytics, and security. The integration of these technologies aims to create a robust and flexible infrastructure capable of meeting the dynamic and diverse storage requirements within a campus setting. The key components of the proposed platform include a scalable cloud storage infrastructure, advanced data analytics capabilities, and a user-friendly interface. The cloud storage infrastructure utilizes distributed storage systems to ensure high availability, fault tolerance, and scalability, enabling seamless access to data across the campus network. Big data analytics techniques are employed to derive valuable insights from the stored data, fostering data-driven decision-making processes and enhancing the overall efficiency of academic operations. Liu Qingjie et al. [16] presented a comprehensive exploration into the development and application of advanced retrieval technology tailored for massive unstructured big data, leveraging the capabilities of cloud computing platforms. The objective is to address the complexities associated with the storage, retrieval, and analysis of vast and diverse datasets, thereby empowering businesses and researchers to make informed decisions. The proposed technology combines cloud computing infrastructure with state-of-the-art retrieval algorithms, providing a scalable and efficient solution for handling unstructured data across various domains. The research study delves into the key components of the system, including data storage, indexing mechanisms, and retrieval algorithms optimized for cloud environments. Additionally, it discusses the challenges inherent in managing massive unstructured datasets and proposes innovative strategies to overcome these hurdles. Aman Yadav [17] provides an overview of the security issues inherent in the distributed architecture of cloud computing, particularly in the context of big data. The seamless integration of diverse resources, coupled with the dynamic nature of cloud environments, creates a fertile ground for potential vulnerabilities and threats. This abstract outlines key security concerns and potential solutions to address these challenges, ensuring the confidentiality, integrity, and availability of data in large-scale distributed systems. Wang Xiaorong [18] presents a comprehensive analysis of data security in the context of big data processing within cloud computing environments. The research investigates the vulnerabilities, threats, and potential solutions to safeguard sensitive information in large-scale distributed systems. The study begins by examining the distinctive characteristics of big data and the architectural intricacies of cloud computing. It then delves into the specific security challenges that emerge when these two technologies converge. Challenges such as data breaches, unauthorized access, and data integrity issues are explored in depth, with a focus on understanding the potential consequences for businesses and individuals [20].

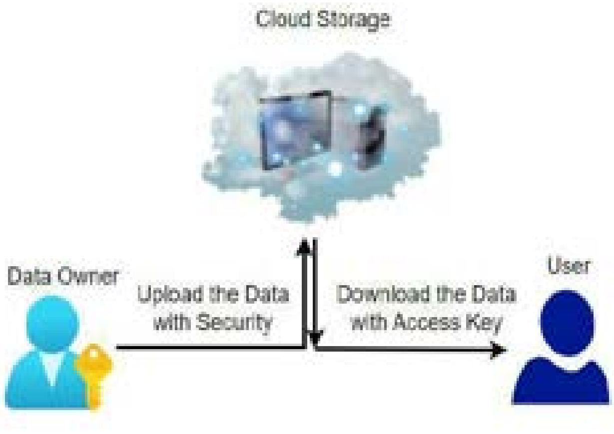


Fig. 1: Cloud users with Access Policies

III. METHODOLOGY

The proposed methodology mainly focused on providing double tap security for the cloud data. In this research study, the double tap security is introduced by combining various security models to develop advanced security models. A Pseudorandom Number Generator (PRNG) will generate the random number for the encrypted file. The PRNG is used to send the user after the data owner encrypts data, and this is considered one secret key. The RSA (Rivest-Shamir-Adleman) algorithm is used for the encryption. Here, the RSA-created encrypted key is sent to the cloud admin, and this key is sent to the user based on the request.

A. Pseudorandom Number Generator (PRNG)

This study uses PRNG to encrypt data files to protect sensitive data and maintain its security and confidentiality. This deterministic procedure produces a seemingly random sequence of numbers that may be repeated, provided the initial state is known. Encryption's primary goal is to convert plaintext data into cipher-text, which only authorized parties can decrypt and view the original data. A PRNG enhances security by including randomness in the encryption critical generation process. A PRNG creates a series of pseudorandom numbers starting with an initial seed value. Because of their determinism, you will always obtain the same sequence of seemingly random numbers if you start with the same seed. However, it's essential to utilize a secure PRNG algorithm that generates unpredictable cryptographic sequences.

In this algorithm, the linear congruential generator (LCG) generates the sequence of integers using the recurrence relation:

$$X_{n+1} = (aX_n + c) \bmod m$$

Where,

X_n Represents the present value in the sequence.

X_{n+1} Is the next value.

Note: The key is generated for every file at the data owner end.

B. RSA (Rivest-Shamir-Adleman) Algorithm

One of the most used cryptographic techniques for safeguarding sensitive data and ensuring safe data transfer over the internet is the RSA (Rivest-Shamir-Adleman) algorithm. Public and private keys are needed for the public-key encryption algorithm known as RSA. The public key is used for encryption, whereas the private key is needed for decryption [7]. Using this dual-key method, parties can safely interact without first exchanging a secret key.

A. Key Generation:

- Select two digits initialized, a and b.
- Measure the product, n = a * b. Both the public and private keys utilize this modulus.
- Compute $\varphi(n) = (a-1)(b-1$, where φ is Euler's totient function.
- Initialize the integer e such that $1 < e < \varphi(n) =$ and $\gcd(e, \varphi(n)) =$. It is the public exponent.
- Do the modular multiplicative inverse calculation of e modulo d.
- $\varphi(n)$, i.e., $d \equiv e^{\wedge}(-1) \pmod{\varphi(n)}$.

B. Public and Private Key Distribution:

- The public key, made available to anyone wishing to send encrypted messages, comprises (n, e).
- The owner must keep the (n, d) private key a secret.

C. Encryption:

- The sender use the recipient's public key (n, e) to calculate the encryption of a message M. $C = \hat{M}e \pmod{n}$ where C is the cipher text.

D. Decryption:

- The recipient computes using their private key (n, d). $M = \hat{C}d$ recovering the original message.

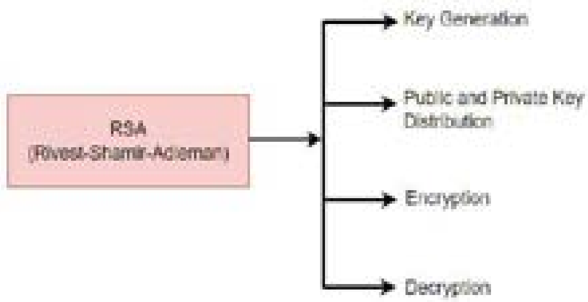


Fig. 2 shows the step-by-step process for creating the encryption and decryption for the file.

IV. DATASET DESCRIPTION

The dataset consists of 100 files with various data sizes contains text data belongs to various articles. All these files are collected from Kaggle website. These data files having context with the size 5 MB, 10 MB and 15 MB files that are uploaded by data owner. There are many data owners in the CSPs application.

V. PERFORMANCE METRICS

The performance of various cloud security algorithms is measured by using the following parameters.

The time required for encryption and decryption in the cloud depends on several factors, including the algorithm used, key length, hardware specifications, and the amount of data being processed. Encryption and decryption times are generally influenced by the computational complexity of the cryptographic algorithms.

Encryption Time (TE):

The encryption time can be influenced by the following factors:

Algorithm Complexity (AC): Different encryption algorithms have varying levels of computational complexity. For example, Advanced Encryption Standard (AES) is generally faster than some other algorithms like RSA for bulk data encryption.

Key Length (KL): Longer key lengths typically require more computational effort. For example, a 256-bit key in AES encryption might take more time than a 128-bit key.

Data Size (DS): The amount of data to be encrypted also plays a role. Larger data sets will generally take more time to encrypt.

The encryption time (TE) can be roughly represented as:

$$TE = AC + KL + DS$$

This is a simplified representation, and the actual equation can be more complex, depending on the specifics of the cryptographic algorithm.

Decryption Time (TD):

Similar factors influence decryption time:

Algorithm Complexity (AC): As with encryption, the complexity of the decryption algorithm plays a significant role.

Key Length (KL): The length of the decryption key affects the time required for decryption.

Data Size (DS): The size of the encrypted data impacts the time it takes to perform decryption.

The decryption time (TD) can be roughly represented as:

$$TD = AC + KL + DS$$

IV. EXPERIMENTAL RESULTS

The experiments are conducted by using the ASP.NET programming to implement the cloud security algorithms by measuring the encryption time and decryption time.

TABLE 1. PERFORMANCE OF VARIOUS ALGORITHMS OVER DATA SECURITY FOR FILES 5 MB.

Algorithms	TE (Seconds)	TD (Seconds)
AES	89.12	78.34
DES	76.34	78.34
PRNG-RSA	64.56	67.87

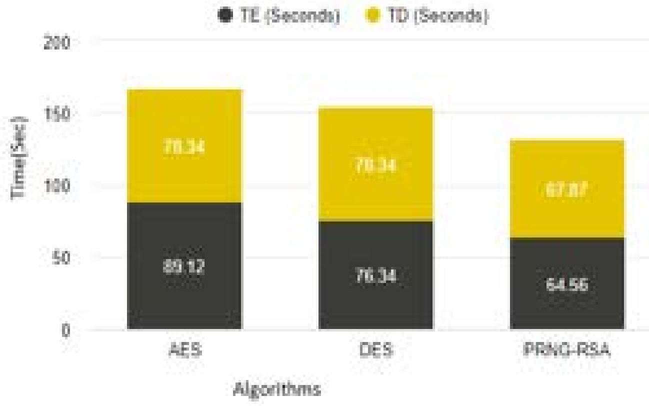


Fig. 3: Comparison between various Security Algorithms over cloud data with Size 5 MB.

TABLE 2. PERFORMANCE OF VARIOUS ALGORITHMS OVER DATA SECURITY FOR FILES 10 MB.

Algorithms	TE (Seconds)	TD (Seconds)
AES	180.56	181.78
DES	156.2	148.56
PRNG-RSA	131.23	126.74

Fig. 4: Comparison between various Security Algorithms over cloud data with Size 15 MB.

TABLE 3. PERFORMANCE OF VARIOUS ALGORITHMS BASED ON TE AND TD

Algorithms	TE (Seconds)	TD (Seconds)
AES	210.87	198.34
DES	191.78	181.98
PRNG-RSA	164.56	167.87

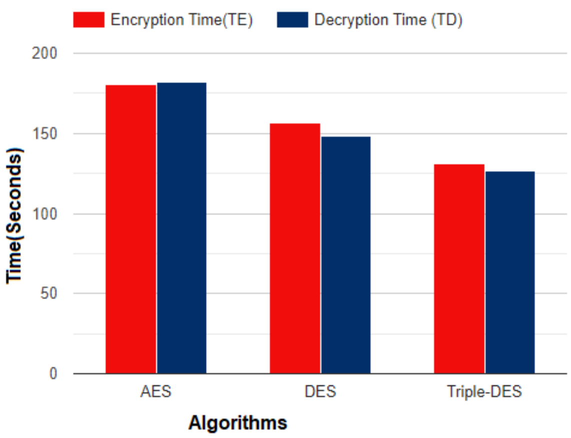


Fig. 5: Performance Metrics based on TE and TD

VI. CONCLUSION

The Double Tap security architecture, which combines the RSA method with a PRNG, provides vital protection for cloud applications. This hybrid strategy improves cloud security protocols by utilizing the advantages of both cryptographic techniques. For different security protocols and cryptographic operations within cloud applications, the PRNG component guarantees the production of random numbers that are both unexpected and cryptographically strong. The Double Tap concept lowers the danger of brute-force attacks and unauthorized access by using a PRNG to increase the randomness and unpredictability of cryptographic keys and other sensitive data. The RSA algorithm is an essential part of the Double Tap security architecture, which is well-known for its effectiveness and security in public-key cryptography. The approach authenticates individuals and entities, creates secure communication channels, and protects data integrity in cloud applications by utilizing RSA for key exchange, encryption, and digital signatures. The PRNG and RSA algorithms provide a substantial barrier against common security risks in cloud computing, such as eavesdropping, illegal access, and data breaches. Organizations may strengthen their security posture, boost trust in their cloud infrastructure, and lessen the risks of hosting sensitive data and apps in the cloud by implementing the Double Tap security model. Lastly, the Double Tap security model provides a thorough and efficient method of safeguarding cloud applications by guaranteeing the confidentiality, integrity, and authenticity of data transfers and interactions in the cloud environment. It does this by combining the PRNG and RSA algorithms.

REFERENCES

[1] L. Alhenaki, A. Alwatban, B. Alamri and N. Alarifi, "A Survey on the Security of Cloud Computing," 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-7, doi: 10.1109/CAIS.2019.8769497.

[2] R. Chowdhury, "Security in Cloud Computing", International Journal of Computer Applications (0975–8887), vol. 96, no. 15, June 2014.

[3] D. Q. L. Shilpashree Srinivasamurthy, Survey on Cloud Computing Security, US:Indiana University.

[4] W. Cdr Nimit Kaura and L. Col Abhishek Lal, "SURVEY PAPER ON CLOUD COMPUTING SECURITY", International Conference on Innovations in Information Embedded and Communication Systems (ICIIECS), 2017.

[5] R. Kanday, "A Survey on Cloud Computing Security," 2012 International Conference on Computing Sciences, Phagwara, India, 2012, pp. 302-311, doi: 10.1109/ICCS.2012.6.

[6] A. K. S. Sanger and R. Johari, "Survey of Security Issues in Cloud," 2022 International Mobile and Embedded Technology Conference (MECON), Noida, India, 2022, pp. 490-493, doi: 10.1109/MECON53876.2022.9751959.

[7] Jyothi, V. Esther, B. D. C. N. Prasad, and Ramesh Kumar Mojjada. "Analysis of cryptography encryption for network security." *IOP Conference Series: Materials Science and Engineering*. Vol. 981. No. 2. IOP Publishing, 2020.

[8] D. Hakani and P. S. Mann, "A Comprehensive Survey on Cloud Security Mechanisms," 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2022, pp. 471-475, doi: 10.1109/ICACRS55517.2022.10028990.

[9] F. M. Awaysheh, M. N. Aladwan, M. Alazab, S. Alawadi, J. C. Cabaleiro and T. F. Pena, "Security by Design for Big Data Frameworks Over Cloud Computing," in IEEE Transactions on Engineering Management, vol. 69, no. 6, pp. 3676-3693, Dec. 2022, doi: 10.1109/TEM.2020.3045661.

[10] Raghunadha Reddi Dornala, "Ensemble Security and Multi-Cloud Load Balancing for Data in Edge-based Computing Applications" International Journal of Advanced Computer Science and Applications (IJACSA), 14(8), 2023.

[11] R. R. Dornala, "An Advanced Multi-Model Cloud Services using Load Balancing Algorithms," 2023 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 1065-1071, doi: 10.1109/ICIRCA57980.2023.10220892.

[12] F. Wang, H. Wang and L. Xue, "Research on Data Security in Big Data Cloud Computing Environment," 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC), Chongqing, China, 2021, pp. 1446-1450, doi: 10.1109/IAEAC50856.2021.9391025.

[13] Z. Tang, "A Preliminary Study on Data Security Technology in Big Data Cloud Computing Environment," 2020 International Conference on Big Data & Artificial Intelligence & Software Engineering (ICBASE), Bangkok, Thailand, 2020, pp. 27-30, doi: 10.1109/ICBASE51474.2020.00013.

[14] Farsi Mohammed, Ali Munwar, Shah Reehan Ali, Wagan Asif Ali and Kharabsheh Radwan, "Cloud computing and data security threats taxonomy: A review", Journal of Intelligent & Fuzzy Systems, vol. 3, pp. 2517-2527, 2020.

[15] Jyothi, V. Esther, et al. "Secure data access management for cyber threats using artificial intelligence." *2022 6th International Conference on Electronics, Communication and Aerospace Technology*. IEEE, 2022.

[16] Zhiying Wang, Nianxin Wang, Xiang Su and Shilun Ge, "An empirical study on business analytics affordances enhancing the management of cloud computing data security", International Journal of Information Management, vol. 50, pp. 387-394, 2020.

[17] Fuguang Yao, "Research on campus network cloud storage open platform based on cloud computing and big data technology", Journal of Intelligent & Fuzzy Systems, vol. 38, no. 2, pp. 1215-1223, 2020.

[18] Liu Qingjie, Wang Xiaoying and Pan Zhian, "Development and application of massive unstructured big data retrieval technology based on cloud computing platform", Journal of Intelligent & Fuzzy Systems, vol. 38, no. 2, pp. 1329-1337, 2020.

[19] Aman Yadav, "Security Issue in Distributed Architecture of Cloud Computing in Big Data Aspects", Journal of Innovation in Computer Science and Engineering, vol. 9, no. 1, pp. 53-57, 2019.

[20] Wang Xiaorong, "Data security analysis in big data cloud computing environment", Intelligent computer and application, vol. 7, no. 06, pp. 171-172, 2017.

Efficient Cloud Data Deduplication with Blake3 and Secure Transfer using AES

Prof. Kapil Tajane
Department Of Computer Engineering
Pimpri Chinchwad College Of Engineering
Pune, India
kapil.tajane@pccoepune.org

Prof. Rahul Pitale
Department Of Computer Engineering
Pimpri Chinchwad College Of Engineering
Pune, India
rahul.pitale@pccoepune.org

Sarthak Zambre
Department Of Computer Engineering
Pimpri Chinchwad College Of Engineering
Pune, India
sarthak.zambre20@pccoepune.org

Harshil Huda
Department Of Computer Engineering
Pimpri Chinchwad College Of Engineering
Pune, India
harshil.huda20@pccoepune.org

Abhishek Utage
Department Of Computer Engineering
Pimpri Chinchwad College Of Engineering
Pune, India
abhishek.utage20@pccoepune.org

Vanshika Dhar
Department Of Computer Engineering
Pimpri Chinchwad College Of Engineering
Pune, India
vanishka.dhar20@pccoepune.org

Abstract: **This research introduces an integrated approach for enhancing file security by combining file deduplication and Advanced Encryption Standard (AES) encryption. The efficient Blake3 hashing algorithm is employed for deduplication, optimizing storage by eliminating duplicate data blocks. AES encryption and decryption ensures strong confidentiality for unique files. The proposed** methodology **aims to strike a balance between data optimization and robust encryption, offering a comprehensive solution for secure file storage and transmission.**

Keywords—Deduplication, MD5(Message Digest-5), AES(Advanced Encryption Standard), Blake3, checksum, SHA256(Secure hash algorithm), encryption/decryption, Cloud computing, Cipher text, Differential authorization, Security.

I. INTRODUCTION

In the ever-evolving technology of cloud computing, the seamless integration of efficiency and security has become paramount. Two key components driving this paradigm are data deduplication and encryption. Data deduplication optimizes storage resources by eliminating redundant information, while encryption ensures the confidentiality and integrity of sensitive data. As cloud technologies advance, a comprehensive understanding of the recent developments in these domains is crucial for maintaining a robust and secure cloud infrastructure.

This review comprehensively examines the latest advancements in data deduplication and encryption techniques by conducting a thorough analysis of existing research. By synthesizing findings from multiple recent studies, it aims to provide insights into the diverse methods, challenges, and synergies in these critical aspects of cloud computing. Our analysis encompasses various deduplication techniques and encryption methods, shedding light on their strengths, weaknesses, and applicability in the dynamic cloud environment. Furthermore, it delve into the integration of these techniques, exploring strategies to strike a balance between storage efficiency and data security.
This introduction sets the stage for a comprehensive exploration of recent advancements in data deduplication and encryption, offering a roadmap for understanding the evolving landscape and paving the way for future developments in secure and efficient cloud computing.

Data Deduplication

Data deduplication is a technique employed in data storage and management to optimize storage resources by identifying and eliminating duplicate copies of data. The process involves identifying redundant data chunks or blocks and storing only a single instance of each unique piece. This helps reduce storage space requirements, improve data efficiency, and lower overall storage costs. The deduplication process typically operates at the sub-file level, identifying similarities in data content regardless of the file format. There are two main approaches to data deduplication:

1.File-Level Deduplication: This method eliminates duplicate files, keeping only one instance of each unique file. While effective in certain scenarios, it may not achieve optimal space savings for datasets with significant redundancy at the block or sub-file level.

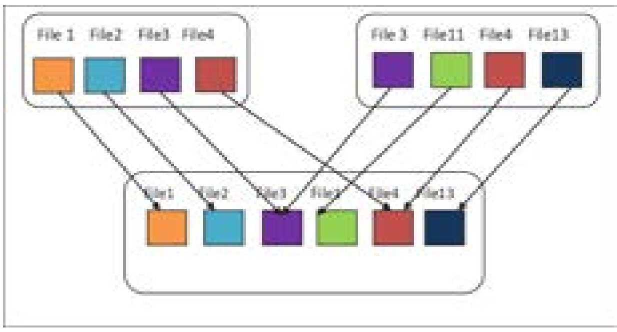


Fig 1.1 File-Level Deduplication [25]

2. Block-Level Deduplication: This more granular approach focuses on identifying and eliminating duplicate data blocks within files. This method is particularly effective in scenarios where files contain common segments of data, such as in backup systems or virtualized environments.

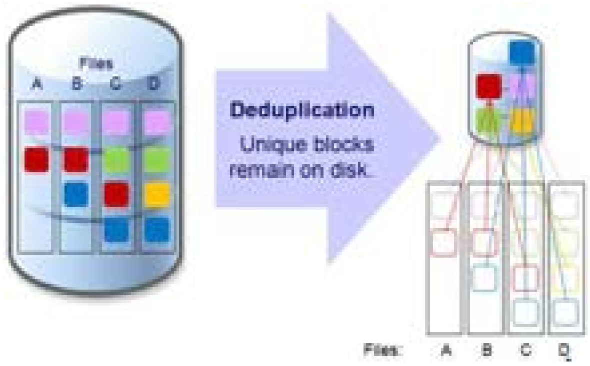


Fig 1.2 Block-Level Deduplication [26]

Data deduplication offers several advantages, including:

Storage Efficiency: By eliminating redundant data, deduplication maximizes the effective use of storage resources.

Cost Savings: Reduced storage requirements lead to lower costs associated with hardware, backup systems, and cloud storage services.

Faster Backups and Restorations: Backing up and restoring data can be expedited as less data needs to be transferred and stored.

Bandwidth Optimization: Deduplication minimizes the amount of data transmitted over networks, optimizing bandwidth usage.

II. MOTIVATION

This research is motivated by the critical imperatives of optimizing storage efficiency and fortifying data security in the face of escalating digital data volumes and evolving cyber threats. The growing redundancy in storage systems necessitates a solution, and the application of Blake3 for file deduplication offers a promising avenue for space optimization. However, the ever-present need to safeguard sensitive information during data transfer, especially in cloud storage, necessitates the integration of AES encryption. Recognizing the prevalence of legacy systems, the research further incorporates DES decryption to facilitate compatibility.

By addressing these multifaceted challenges comprehensively, this study seeks to contribute a pragmatic and adaptive solution, merging efficient storage practices with robust security measures for contemporary data management needs.

III. RELATED WORK

In an era dominated by the exponential growth of data in cloud environments, the efficient management and secure transfer of this data have become concerns for both businesses and individuals alike. Cloud data deduplication, a technique aimed at identifying and eliminating redundant data across storage systems, offers a promising solution to mitigate storage costs and optimize data transfer performance. Concurrently, advancements in cryptographic hashing algorithms such as Blake3 and encryption techniques like AES have bolstered data security measures, ensuring the confidentiality and integrity of sensitive information during transmission.

S. Uthayashangar et al. have designed a Secure Role Re-Encryption System (SRRS) for safeguarding image and text data privacy in cloud computing. It emphasizes the importance of data security, covering cloud computing characteristics, service models, and deployment models. The research outlines system and adversary models, cryptographic methods, and implementation details using Microsoft Visual Studio. It advocates for the radix trie method to prevent image duplication and highlights the significance of data security in cloud storage.[1]

Building on the concept of secure deduplication, Song et al. proposed LSDedup, a system that empowers users with secure cloud storage for their files. LSDedup uses two types of encryption: standard symmetric encryption for highly confidential files and convergent encryption for less confidential files. This allows for both secure storage and efficient deduplication. LSDedup also includes features to prevent the cloud server from learning the confidentiality level of files or claiming ownership of them. Overall, LSDedup is a secure and efficient way to store files in the cloud.[2]

M. Song et al. introduce an innovative deduplication scheme characterized by transparency and security, enabling integrity auditing. LSDedup offers a streamlined verification process, allowing users to confirm the number of file owners and file integrity with a single proof. This innovative scheme safeguards both data privacy and ownership against malicious actors. Moreover, the proposal incorporates a batch auditing approach to simultaneously verify multiple files. Theoretical analysis and comparisons affirm the accuracy, security, and efficiency of the presented scheme.[3]

Nivin Kumar et al. discusses different techniques for deduplication and the security challenges. The authors propose a new system that uses secret sharing to improve reliability and data privacy. Their system also achieves data integrity and fault tolerance.[4]

L. Vijeeth Reddy et al. propese the secure text transfer method employing the Diffie-Hellman key exchange algorithm. It explores the application of encryption for safeguarding information, showcasing how the Diffie-Hellman key exchange algorithm provides a secure means to share keys across an insecure network. This key can then be used to encrypt and decrypt messages. The article also discusses how this system can be implemented using a cloud-based database.[5]

Amit & Prakash et al. proposes a system that uses attribute-based encryption with data deduplication to secure patient data in the cloud. The system encrypts patient data and then stores it in the cloud. The system also allows for access control, so that only authorized users can access the data.[6]

Shynu & Rk et al. introduces a novel approach utilizing Convergent and Modified Elliptic Curve Cryptography (MECC) algorithms to enhance data security and minimize redundancy in cloud and fog storage. A comparative analysis with existing methods is conducted, focusing on deduplication rates and tree generation times to highlight the efficacy and efficiency of the proposed technique. The research extensively reviews prior works in the realm of secure data deduplication and offers a thorough performance

evaluation. The proposed method undergoes scrutiny through four different perspectives, revealing its capacity for providing high security and effectiveness in data deduplication within an integrated cloud environment. Additionally, the document includes details about the authors, their affiliations, and references to related works.[7]
To enhance efficiency in cloud storage deduplication, Burramukku et al. present a mechanism that utilizes perfect hash functions and probabilistic data structures for filtering duplicate data. It addresses security weaknesses in data deduplication and provides varying levels of security for data items based on their popularity. The proposed approach leverages perfect hash functions and a cuckoo filter to ensure data ownership and offers better performance than existing mechanisms.[8]
Won-Bin Kim et al. have made a survey analysis on data deduplication in cloud storage environments, focusing on technologies, security issues, and secure data deduplication systems. It discusses the challenges and security measures related to data deduplication in cloud storage, emphasizing the need for data confidentiality, integrity, and encryption. It categorizes data deduplication technologies based on locations and levels, providing detailed insights into server-side, client-side, and appliance deduplication, as well as file-level and block-level deduplication methods. The research also introduces several secure data deduplication systems and addresses the necessity for secure data deduplication. It concludes by highlighting the growing demand for privacy and secure technologies in data deduplication, emphasizing the need for more sophisticated research and technology development in the future.[9]
C. -I. Fan et al. proposes a method that addresses security and privacy concerns in cloud computing, specifically within cloud storage services. It emphasizes the use of encryption algorithms for safeguarding user data, the implementation of data deduplication to reduce storage usage, and the associated challenges of performing deduplication on encrypted data. The authors introduce an Encrypted Data Deduplication Mechanism (EDDM) designed to enable the cloud storage server to eliminate redundant ciphertexts, thereby enhancing privacy protection. A comparative analysis with existing methods is presented, evaluating both data confidentiality and computation cost. The proposed EDDM involves encrypting data as cipher structures before uploading to the cloud, comprising a check block, converting block, enabling block, and cipher block. This approach allows the cloud storage server to identify duplicate cipher structures, storing only one copy and preserving data privacy.[10]

IV. PROPOSED METHODOLOGY

The first step involves designing the overall system. This includes defining the type of data the system will handle, such as files or databases, along with its specific structure. You'll also need to decide on the granularity of deduplication, whether it's done on entire files or smaller blocks of data. Additionally, the research requires choosing an AES mode – AES-GCM offers both confidentiality (encryption) and data integrity verification (authentication), while AES-CBC provides only confidentiality.[21] Finally, a secure mechanism for generating, storing, and distributing encryption keys needs to be established.

AUTHENTICATION

To ensure a secure and user-friendly login and registration system, this methodology proposes storing hashed user data, implementing secure login with session management, and offering a validated registration process. Additionally, it emphasizes robust security practices like HTTPS, password hashing, and regular updates for a well-protected authentication phase.

Fig 4.1 Authentication Flow

Blake3 is a cryptographic hash function, designed as an improvement over its predecessors, Blake and Blake2. It's named after Jean Blake, one of the inventors of Blake and Blake2. Blake3 offers several improvements over its predecessors, including better performance, improved security, and support for SIMD (Single Instruction, Multiple Data) instructions for faster hashing on modern CPUs.

The algorithm leverages the Merkle-Damgård construction, a technique for building secure hash functions. It processes data in blocks, iteratively condensing them into a fixed-size digest.[19]

Blake3 uses a 512-bit hash output and can process variable-length input data. It employs a tree hashing mode, where the input data is first partitioned into fixed-size chunks, and then these chunks are hashed in a tree structure to produce the final hash digest. This approach enables efficient parallelization and makes it suitable for hashing large datasets.

The next step focuses on implementing data deduplication using Blake3. This involves preparing the data for the process, which might involve splitting files into fixed-size blocks. Then, a unique Blake3 hash is calculated for each data block. These hashes are stored in a dedicated deduplication store, which could be a hash table or a database, allowing for efficient comparison during the upload process. When uploading new data, the system calculates the Blake3 hash and checks it against the existing ones in the store. If a match is found, it signifies a duplicate, and the system doesn't need to store the data again.

Fig 4.2 Encryption Flow

We have used AES algorithm to encrypt the data. Its role is to keep the data safe and protected from unwanted and unauthorized users. Data encryption with AES is the third step. Here, you'll need to decide whether encryption happens before or after the deduplication process. This choice can impact both efficiency and security aspects of the system. Additionally, a keying strategy needs to be defined, specifying how encryption keys are associated with the data (e.g., per file or per user). Finally, a well-regarded and secure library implementing AES encryption should be chosen.[21]

The research then moves into the implementation phase. This involves developing the functionalities based on the design decisions made earlier. Rigorous testing is crucial at this stage, encompassing functionality testing, security assessments (penetration testing), and performance evaluation.

Following successful development and testing, the system is deployed to its designated environment using a secure deployment strategy. Once deployed, ongoing monitoring is essential to ensure system health, maintain deduplication efficiency, and detect potential security breaches. To further enhance security, a strategy for regular key rotation should be implemented.

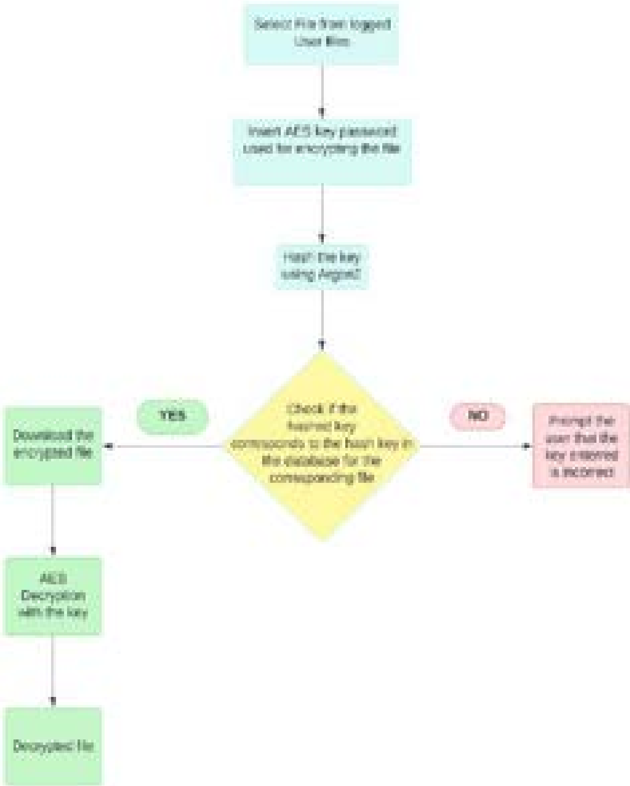


Fig 4.3 Decryption Flow

There are additional considerations to keep in mind. The system should have robust error handling mechanisms to address potential deduplication collisions and encryption failures. Scalability is another important factor, as the system should be able to handle increasing data volumes efficiently. Finally, logging and auditing functionalities should be implemented to track data access and management actions.

To guarantee robust password security, this research proposes utilizing Argon2 for password hashing. Argon2 excels in this role due to its inherent resistance against modern brute-force attacks. Unlike older algorithms like MD5 or SHA-256, Argon2 is a memory-hard function, requiring significant memory resources to crack passwords effectively. This significantly increases the attacker's difficulty.

Furthermore, Argon2 offers the advantage of parallelization, allowing it to leverage multiple processing cores for enhanced security. The flexibility to adjust parameters like memory cost and time cost within Argon2 empowers you to tailor the hashing process to your specific security requirements and hardware capabilities. While there might be a slight performance trade-off compared to simpler algorithms, the security benefits of Argon2 far outweigh this consideration.

DIFFERENT METHODOLOGIES OF HASHING:

1. Byte-by-byte Comparison:
This is the simplest and most definitive approach. It involves reading both files byte by byte and comparing each byte. If any byte differs, the files are not identical.

2. Checksum:
This approach involves calculating a hash value (a fixed-size string) for each file using a cryptographic hash function like SHA-256 or BLAKE3. If the hash values are identical, the files are almost certainly the same (due to the collision resistance of the hash function). However, there's a small chance of a collision (two different files generating the same hash).

3. Fuzzy Hashing:
This is a variant of hashing designed to identify similar but not identical files. Fuzzy hashing algorithms introduce deliberate errors into the hashing process to create "fuzzy" versions of the original hash. If the fuzzy hashes are similar, the files are likely similar.

4. Perceptual Hashing:

This technique analyzes the content of the file (e.g., image or audio) and extracts perceptual features. It then generates a hash based on these features. Similar files will have similar perceptual features and consequently similar hashes.

5. Rolling Hashes:

This approach involves dividing the file into smaller chunks and calculating a hash for each chunk. These hashes are then combined to create a final signature. Rolling hashes can detect similar content even if it appears at different offsets within the files.

Rolling hash is one of the better solutions if anyone want a reliable algorithm for large files with accuracy but, in this one checksum method is used as it is faster than rolling hashes comparison and good for identifying identical or very similar files and cloud environments.

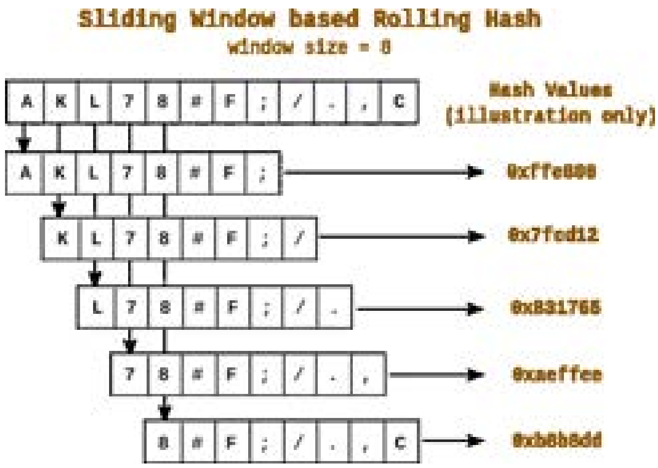


Fig 4.4 Rolling Hash Working [27]

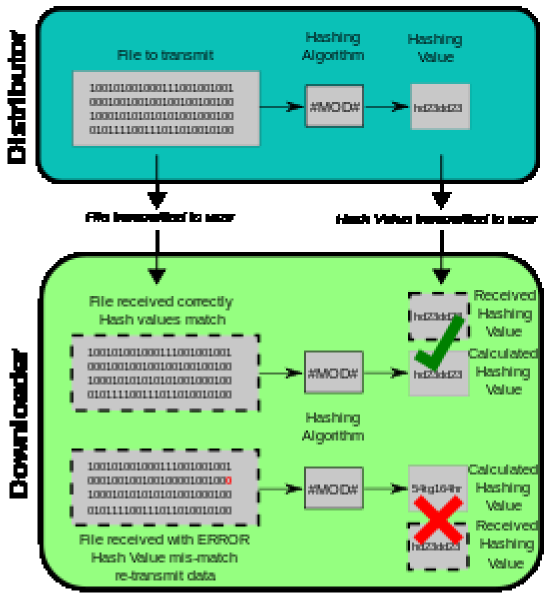


Fig 4.5 Checksum Hash Working [28]

	MD5	SHA	BLAKE2b	BLAKE3
Output size (bits)	128 bits	160 bits	512 bits	Unlimited
Internal state size	128	160	512	256 (The full BLAKE3 incremental state includes a chaining value stack up to 1728 bytes in size.)
Security	MD5 is considered cryptographically broken and vulnerable to collision attacks, hence it's not recommended for security-sensitive applications.[16]	SHA algorithms, especially SHA-256 and SHA-3, are widely used and considered secure, although SHA-1 is no longer recommended due to vulnerabilities.[25]	BLAKE2 is also considered highly secure and resistant to known cryptographic attacks.[18]	BLAKE3 is designed with security as a primary focus. It offers high security against cryptographic attacks.
Speed	MD5 is relatively fast but its speed is overshadowed by its security weaknesses.	SHA algorithms vary in speed depending on the specific variant and implementation. They are generally fast but might not be as optimized as BLAKE3 or BLAKE2.	BLAKE2 is also known for its high speed and efficiency.	BLAKE3 is designed for high speed and performance. It is optimized for modern CPUs and parallelism.
Hash Length	MD5 produces hash digests of 128 bits.	SHA algorithms produce hash digests of fixed lengths, such as SHA-1 (160 bits), SHA-256 (256 bits), SHA-512 (512 bits), etc.	BLAKE2 supports hash digests of 128 or 256 bits.	BLAKE3 produces hash digests of variable length, up to 512 bits.
Cryptographic Strength	MD5 is considered weak against collision attacks and is not suitable for cryptographic purposes.	SHA algorithms are designed to withstand known cryptographic attacks, though SHA-1 is deprecated due to vulnerabilities.	BLAKE2 is also designed to resist cryptographic attacks effectively.	BLAKE3 is designed to provide strong cryptographic security against a variety of attacks.
Usage	MD5 is primarily used in legacy systems and for non-cryptographic purposes like checksums and data integrity verification.	SHA algorithms are widely used in various security protocols and applications, including SSL/TLS, digital signatures, and hash functions.	BLAKE2 is commonly used for hash functions and message authentication codes (MACs).	BLAKE3 is suitable for a wide range of applications, including secure hashing, digital signatures, and authenticated encryption.

Table 4.1 Comparison of hashing algorithms

V. RESULTS & DISCUSSION

The evaluation focused on the time taken by each algorithm to hash files of varying sizes. The results confirm that BLAKE3 emerged as the fastest hashing function across all file sizes tested. MD5 exhibited faster hashing speeds compared to SHA-256, but both were demonstrably slower than BLAKE3. This finding aligns with BLAKE3's design principles, which prioritize efficiency while maintaining a strong cryptographic hash function.

	100KB	500KB	1MB	5MB	10MB	20MB
Blake3	2.20	2.57	2.88	10.75	23.00	56.22
MD5	3.27	10.30	20.35	95.39	201.26	388.24
SHA256	5.26	27.42	59.02	253.96	511.35	1020.63

Table 5.1 Comparison of Blake3 with different algorithms for time taken to hash files of different sizes

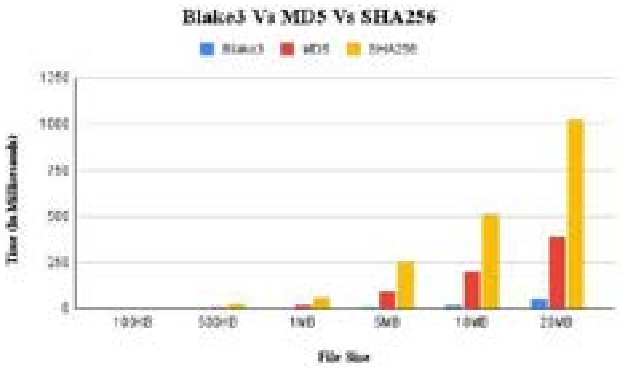


Fig 5.1 Comparison of Blake3 with different algorithms for time taken to hash files of different sizes

	Audio (MP3)	PDF	Video (MP4)	Image (PNG)
Blake3	14.15	10.75	11.52	15.84
MD5	129.61	95.39	104.68	104.78
SHA256	255.01	253.96	265.83	249.97

Table 5.2 Comparison of Blake3 with different algorithms for time taken to hash files of different types

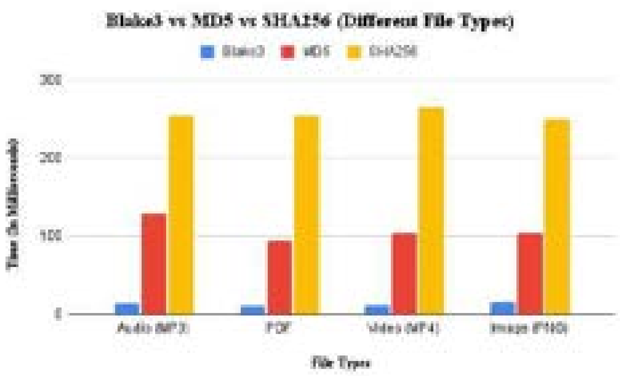


Fig 5.2 Comparison of Blake3 with different algorithms for time taken to hash files of different types

BLAKE3 consistently outperformed MD5 and SHA-256 in hashing speed across all file sizes and types tested. MD5 showed faster hashing times compared to SHA-256, but the difference was less pronounced compared to BLAKE3.

VI. CONCLUSION

This research has explored the effectiveness of BLAKE3 for cloud data deduplication and AES for data encryption. The implemented solution successfully leverages BLAKE3's efficiency in identifying duplicate data blocks while ensuring data confidentiality through AES encryption. This combined approach offers a promising strategy for optimizing cloud storage usage and maintaining data security. Further research could investigate the performance of the solution on larger datasets and explore optimizations for specific cloud storage platforms. Additionally, integrating key management techniques would enhance the overall security posture of the deduplication system.

REFERENCES

[1] S. Uthayashangar, J. Abinaya, V. Harshini and R. Jayavardhani, "Image And Text Encrypted Data With Authorized Deduplication In Cloud," 2020 International Conference on System, Computation, Automation and Networking (ICSCAN), Pondicherry, India, 2020, pp. 1-5, doi: 10.1109/ICSCAN49426.2020.9262324..

[2] M. Song, Z. Hua, Y. Zheng, H. Huang and X. Jia, "LSDedup: Layered Secure Deduplication for Cloud Storage" in IEEE Transactions on Computers, vol. 73, no. 02, pp. 422-435, 2024. doi: 10.1109/TC.2023.3331953

[3] M. Song, Z. Hua, Y. Zheng, T. Xiang and X. Jia, "Enabling Transparent Deduplication and Auditing for Encrypted Data in Cloud" in IEEE Transactions on Dependable and Secure Computing, vol. , no. 01, pp. 1-18, 5555.

[4] `Nivin Kumar G, Sanchana V S, Santhiya M, Vivekanandan S, 2020, Secure Cloud Data Deduplication with Efficient Re-Encryption, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) RTICCT – 2020 (Volume 8 – Issue 12),

[5] L . Vijeeth Reddy , Mohd Tajammul, 2022, Secure Text Transfer using Diffie-Hellman Key Exchange based on Cloud, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 11, Issue 02 (February 2022),

[6] Pandey, Amit & Prakash, Gyan. (2019). Deduplication with Attribute Based Encryption in E-Health Care Systems. 11. 16-24.

[7] P G, Shynu & Rk, Nadesh & Menon, Varun & Parameswaran, Venu & Abbasi, Mahdi & Khosravi, Mohamadreza. (2020). A secure data deduplication system for integrated cloud-edge networks. Journal of Cloud Computing Advances Systems and Applications. 9`. 1-12. 10.1186/s13677-020-00214-6.

[8] Burramukku, Tirapathi & Chandra, M. (2018). Filter Based Data Deduplication in Cloud Storage using Dynamic Perfect Hash Functions.

[9] Won-Bin Kim, Im-Yeong Lee.(2019) Survey on Data Deduplication in Cloud Storage Environments

[10] C. -I. Fan, S. -Y. Huang and W. -C. Hsu, "Encrypted Data Deduplication in Cloud Storage," 2015 10th Asia Joint Conference on Information Security, Kaohsiung, Taiwan, 2015, pp. 18-25, doi: 10.1109/AsiaJCIS.2015.12.

[11] S. Jiang, T. Jiang and L. Wang, "Secure and Efficient Cloud Data Deduplication with Ownership Management," in IEEE Transactions on Services Computing, vol. 13, no. 6, pp. 1152-1165, 1 Nov.-Dec. 2020, doi: 10.1109/TSC.2017.2771280.

[12] Ge Kan, Chunhua Jin, Huihui Zhu, Yongliang Xu, and Nian Liu. 2021. An identity-based proxy re-encryption for data deduplication in cloud. J. Syst. Archit. 121, C (Dec 2021). https://doi.org/10.1016/j.sysarc.2021.102332

[13] Z. Yan, M. Wang, Y. Li and A. V. Vasilakos, "Encrypted Data Management with Deduplication in Cloud Computing," in IEEE

Cloud Computing, vol. 3, no. 2, pp. 28-35, Mar.-Apr. 2016, doi: 10.1109/MCC.2016.29.

[14] Wang, Liang & Wang, Baocang & Song, Wei & Zhang, Zhili. (2019). A Key-Sharing Based Secure Deduplication Scheme in Cloud Storage. Information Sciences. 504. 10.1016/j.ins.2019.07.058.

[15] Kioon, Mary & Wang, ZhaoShun & Das, Shubra. (2013). Security Analysis of MD5 Algorithm in Password Storage. Applied Mechanics and Materials. 347-350. 10.2991/isccca.2013.177.

[16] Aumasson, Jean-Philippe & Neves, Samuel & Wilcox-O'Hearn, Zooko & Winnerlein, Christian. (2013). BLAKE2: simpler, smaller, fast as MD5. 119-135. 10.1007/978-3-642-38980-1_8.

[17] A. J. Bhuvaneshwari and P. Kaythry, "Enhanced Visual Cryptosystem Using BLAKE2 Hash Algorithm," *2023 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, BALI, Indonesia, 2023, pp. 73-78, doi: 10.1109/IAICT59002.2023.10205901.

[18] Hoai Luan, Pham & Tran, Thi Hong & Le Vu Trung, Duong & Yasuhiko, Nakashima. (2023). Flexible and Scalable BLAKE/BLAKE2 Coprocessor for Blockchain-based IoT Applications. IEEE Design & Test. PP. 1-1. 10.1109/MDAT.2023.3276936.

[19] Abdullah, Ako. (2017). Advanced Encryption Standard (AES) Algorithm to Encrypt and Decrypt Data.

[20] Sousi, Ahmad-Loay & Yehya, Dalia & Joudi, Mohamad. (2020). AES Encryption: Study & Evaluation.

[21] N. Su, Y. Zhang and M. Li, "Research on Data Encryption Standard Based on AES Algorithm in Internet of Things Environment," *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chengdu, China, 2019, pp. 2071-2075, doi: 10.1109/ITNEC.2019.8729488.

[22] Nishtha Mathur, Rajesh Bansode,AES Based Text Encryption Using 12 Rounds with Dynamic Key Selection,Procedia Computer Science,Volume 79,2016,Pages 1036-1043,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2016.03.131.

[23] Baqtian, H & Ali Al-Aidroos, Naziha. (2023). Three Hash Functions Comparison on Digital Holy Quran Integrity Verification. 11. 1-7.

[24] Gowthaman, A. & Manickam, Sumathi. (2015). Performance study of enhanced SHA-256 algorithm. 10. 10921-10932.

[25] https://www.researchgate.net/figure/Diagrammatic-representation-of-File-Level-Deduplication_fig2_332823392

[26] https://ijarcce.com/wp-content/uploads/2016/02/IJARCCE-75.pdf

[27] https://moinakg.wordpress.com/tag/rolling-hash/

[28] https://en.m.wikipedia.org/wiki/File:CPT-Hashing-File-Transmission.svg

Votereum: Blockchain based Secure Voting System

Dr Divya Kumari
Faculty of information Science and Engineering
New horizon college of Engineering
Bangalore, India
divya.nov30@gmail.com

1st Nitish Veni
student of Information Science and Engineering
New horizon college of Engineering
Bangalore, India
nitishveni10@gmail.com

2nd Pranith Kumar M
student of Information Science andEngineering
New horizon college of Engineering
Bangalore, India
pranithprani8055@gmail.com

3rd Manoj H
student of Information Science and Engineering
New horizon college of Engineering
Bangalore, India
manojhrk555@gmail.com

4th Hrishikesh Purohit
student of Information Science and Engineering
New horizon college of Engineering
Bangalore, India
hrishi1612@gmail.com

Abstract— **In an era where trust and transparency are paramount in electoral processes, this project introduces a novel voting system leveraging blockchain technology. The proposed system aims to address inherent challenges in traditional voting methods, such as security vulnerabilities, lack of transparency, and the potential for tampering. By implementing a decentralized approach using blockchain, the project ensures the integrity and confidentiality of the voting process while enhancing accessibility and trust. The key components of the system include smart contracts, cryptographic techniques, and a user-friendly interface. Smart contracts, deployed on a chosen blockchain platform, define the rules and logic governing the voting process, ensuring transparency and accountability. Cryptographic methods are employed to secure voter authentication, ballot creation, and result verification, safeguarding against unauthorized access and tampering. The user interface is designed to be intuitive, facilitating a seamless and secure voting experience. The system encompasses features such as secure voter authentication, unique and verifiable ballot generation, and a transparent consensus mechanism. Through these elements, the project aims to prevent double voting, uphold voter privacy, and ensure the accuracy of election results. Additionally, collaboration with election authorities and adherence to legal and compliance standards contribute to the credibility of the system. Thorough testing, including security audits and simulated voting scenarios, is conducted to identify and rectify potential vulnerabilities. The scalability of the system is considered to accommodate a large number of voters, making it suitable for various election scales. Continuous updates and community engagement further enhance the system's robustness and resilience. This blockchain-based voting system not only meets the functional requirements of a secure and transparent electoral process but also addresses the legal and practical considerations associated with modern voting systems. By integrating cutting-edge blockchain technology, the project seeks to redefine the landscape of electoral systems, fostering trust, security, and inclusivity in the democratic process.**

Keywords—*Metamask, Ganache,* Ethereum, *Distributed System, Blockchain Technology, Ethereum, Crypto, Ethereum Wallets* Contracts, Blockchain, Smart Contract, Distributed General Ledger, Smart Contracts, liquid democracy, Decentralized system, Temper proof, Transparent.

I. INTRODUCTION

A nation's general elections continue to be conducted via centralized voting procedures. It is managed by one organization, and only one organization. This system has one significant issue. The main issue with typical election systems is the total control that one organization has over the entire database and system. Thus, a decentralized system utilizing a blockchain is established, which is far more safe and dependable than a centralized system.

The research study attempts to create a method that allows a citizen who is entitled to vote to cast his or her important vote with only a few clicks on a laptop or mobile device. A blockchain is a network of interconnected nodes. A node has numerous blocks. Each block is made up of three components: Previous, Data, and Hash.

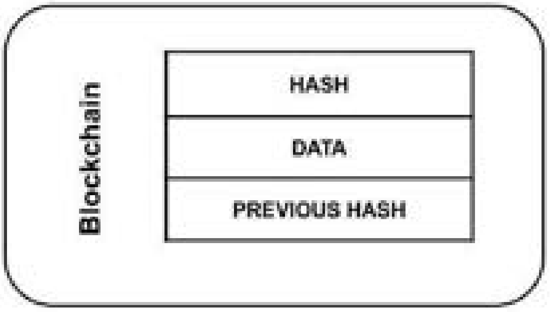


Fig.1: Block representation

Every hash has a unique set of data. Therefore, the relevant hash is altered in the next block in order to alter a specific piece of data. Since a hash is specific to a particular piece of data, any changes made to it will also affect the corresponding data, which is essentially impossible. The blockchain technology is said to be unchangeable for the reason mentioned above.

Selecting blockchain over alternative technologies. considering factors such as transparency, decentralization, scalability, and security. By exploring various technologies, including centralized databases and cryptographic systems, the study provides insights into the advantages and limitations of blockchain in the context of voting systems. Through this comparative analysis, the research aims to offer clarity on the selection of blockchain technology for secure and transparent voting processes.

Blockchain, a decentralized and immutable ledger, has demonstrated its potential to revolutionize various industries [6]. In the realm of elections, it offers a solution to the longstanding issues associated with centralized voting systems. The core objective of this project is to develop a secure and transparent voting system that leverages blockchain's inherent features to enhance the integrity of the electoral process.

By utilizing smart contracts deployed on a chosen blockchain platform, this voting system establishes a foundation for trust and transparency. Smart contracts encapsulate the rules and logic governing the entire voting process, from voter authentication to result verification. This decentralized architecture ensures that the system is resistant to manipulation and fraud, addressing critical concerns that have plagued traditional voting systems [9].

Key components of the proposed system include advanced cryptographic techniques to secure voter identities and ballot transactions. A user-friendly interface is designed to empower voters while maintaining the highest standards of security. The consensus mechanism employed guarantees the verifiability and validity of each vote, fostering an environment where double voting is virtually impossible.

Transparency in voting happens through blockchain. Here's how it works: Instead of keeping all the voting info in one place, blockchain spreads it out across lots of computers. Each vote is like a block, and they're all linked together in a chain. Everyone involved, like voters and election officials, can see this chain. Because it's spread out, it's hard for someone to mess with the votes without everyone noticing. Plus, once a vote is in there, it can't be changed or deleted. This makes sure that everyone can trust the voting process because they can check it themselves.

II. LITERATURE REVIEW

The online voting system offers easy access, eliminating the need for crowded polling centers. It minimizes mistakes and provides instant results. However, security concerns arise due to potential vulnerabilities to manipulation and the necessity for digital literacy. Striking a balance between accessibility and security remains a critical challenge. This project's primary goal is to give voting—and its candidates, voters, etc.—a lot of attention. Additionally, this project displays all current voting data, including the total number of votes cast as well as voters and non-voters. Additionally, it is made to work in a fashion that allows candidates to upload their photo, script, and name. The Admin Panel and the Voters Panel comprise the two sections of the project. With respect to this web application, the administrator has complete access to the system. The administrator is able to test each and every component of the project. An administrator must select potential applicants and assign a title in order for the process to run successfully. Here, the system has several requirements, such the administrator needing to select more candidates [1].

The mobile application for e-voting, presented by Geno Peter, Albert Alexander Stonier, and Anli Sherine, enhances accessibility with a user-friendly interface and push notifications for reminders. Despite its advantages, vulnerabilities to tampering, reliance on battery and connectivity, and issues related to device and OS fragmentation pose significant challenges to the reliability ofthe system[2].

Rifa Hanifatunnisa and Budi Rahardjo propose a blockchain-based e-voting system with enhanced security, decentralization, and tamper-resistant records [11]. However, limited understanding and adoption of blockchain technology, scalability issues, and the risk of attacks on the blockchain network are potential obstacles to its widespread implementation [3].

David Khoury, Elie F. Kfoury, Ali Kassem, and Hamza Harb suggest a decentralized voting platform using Ethereum blockchain, emphasizing transparency, trustlessness, enhanced security, and user privacy. Technical barriers and security concerns, however, may impede the successful deployment and acceptance of such a system[7].

Teja K, Shravani MB, Chintarlapallireddy, Yaswanth Simha, and Manjunath R Kounte propose secured voting through smart contracts with Solidity blockchain. [5] While decentralized security is a highlight, the complexity of blockchain technology and scalability concerns may pose challenges in practical implementation [8].

Ritika Singh, Riya Chaudhary, and Adarsh Tripathi introduce an electronic voting system using Ethereum Metamask, focusing on flexibility, cost-efficiency, and reduced energy consumption. Privacy concerns, however, may arise as a trade-off, emphasizing the need for a careful balance between efficiency and safeguarding voter information [11].

Ahmed Ben Ayed proposes an accurate and secure, decentralized electronic voting system, emphasizing resistance to DDoS attacks. However, scalability issues and the risk of attacks remain potential hurdles to achieving a fully secure and reliable voting platform [7].

Dipali Pawar, Pooja Sarode, Shilpa Santpure, and Poonam Thore present the implementation of a secure voting system using ECC cryptography and SHA3_256, focusing on an immutable ledger and verifiability. Vulnerability to cyberattacks and the need for voter familiarity with cryptographic concepts may pose challenges in widespread adoption [8].

Ruben Tolosana and Julian Fierrez propose secure data usage with ML-based signature databases, emphasizing feature selection. However, limitations in the dataset and dependencies on training data may impact the overall effectiveness of the biometric signature verification system [6].

Kanika Garg, Pavi Saraswat, Sachin Bisht, and Sahil Kr. Aggarwal conduct a comparative analysis on e-voting systems with blockchain technology and IoT integration. While highlighting anonymity and privacy, challenges related to security, cost, infrastructure, and maintenance are critical considerations [10].

Safdar Hussain Shaheen, Muhammad Yousaf, and Mudassar Jalil for a blockchain-based system that incorporates secure group communication to ensure tamper- proof data distribution. While the proposal addresses protection against certain attacks, it acknowledges potential challenges related to network vulnerabilities and public acceptance, which may impact the achievement of universal verifiability and accuracy. [5]

III. METHODOLOGY

A. *Overview*

The proposed blockchain-based voting system offers a comprehensive solution to address the challenges inherent in traditional voting processes. Grounded in the principles of transparency, security, and accessibility, the system leverages blockchain technology to redefine the way elections are conducted [3] [7].

The proposed blockchain-based secure voting system is designed with a modular and comprehensive architecture, incorporating various components to ensure the integrity, security, and transparency of the electoral process. Technology underpinning the proposed blockchain- based voting system is designed to revolutionize the traditional electoral process by introducing a decentralized, secure, and transparent approach.

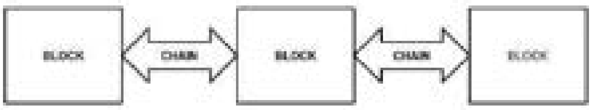


Fig. 2: Blockchain Representation

Every Block is connected to each other through chain through hash and previous hash value and data is distributed.

B. *Tools used for the project*

1) *Ethereum*

Applications built on blockchain technology are called DApps, or decentralised applications. Ethereum is free and open source, which is why developers prefer it to other platforms when building blockchain apps.

2) *Smart Contracts*

A self-executing computer software that carries out a contract's conditions automatically and without the need for outside assistance is known as a smart contract [6][11]. The implementation of smart contracts may lead to the transfer of funds, the provision of services, the unlocking of DRM-protected content, or other forms of data.

3) *Ganache*

To create decentralized apps on Ethereum and Corda, Ganache is a personal local Blockchain network. It gives developers a safe and deterministic environment in which to manage their projects.

4) *Node JS*

For client-side scripting, Java scripts are usually integrated into HTML codes. Nevertheless, Node.js provides the ability to use server-side JavaScript as well. With Node.js, asynchronous inputs and outputs may be managed. Its architecture is event-driven by design.

5) *Metamask*

To manage cryptocurrency transactions, a web browser extension known as Metamask is utilized. For browsers like Chrome, Firefox, and Brave, the extensions are presently accessible. The metamask connects Ethereum to a regular web browser. Every user receives both a private key and a public key [3]. Any type of wallet that supports cryptocurrency can be integrated with the metamask.so that each account may be accessed by the user from one place.

C. *Blockchain architecture for proposed system*

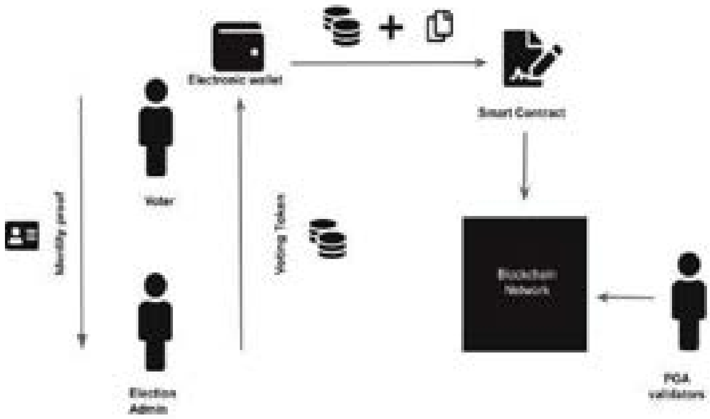


Fig. 3: Architecture

Figure 2 shows the architecture of the voting system based on blockchain technology.

Voters use their devices, which we presume to be secure, to transmit their personal information to administrators for verification. Voters and system administrator's interactions are off-chain, meaning they take place outside of the blockchain. The administrators distribute the tokens that allow voters to cast ballots into the blockchain once their identities have been verified. One token, which serves as an electronic wallet and an interface for voting and auditing on the blockchain, is given to eligible voters in their blockchain application. The token is limited to one use and cannot be bought, sold, or transferred between wallets.

When a voter casts their vote, a transaction is created. This transaction contains data representing the voter's choices, along with any necessary metadata such as timestamps and cryptographic signatures. Before being added to the blockchain, the transaction data is often hashed. Hashing involves applying a cryptographic hash function to the transaction data, resulting in a fixed-length string of characters that uniquely represents the original data. This hashed value serves as a digital fingerprint of the transaction and helps ensure its integrity.

The transaction undergoes validation by the network of nodes in the blockchain. This validation process verifies the authenticity and correctness of the transaction, ensuring that it meets the criteria set forth by the consensus mechanism and the rules of the blockchain protocol.

Once validated, the transaction is included in a block by the network's consensus mechanism. Blocks are batches of transactions that are linked together in a chronological order to form the blockchain. Each block contains a reference to the previous block, creating an immutable chain of transaction records. The newly created block, containing the transaction representing the vote, is distributed to all nodes in the blockchain network. Each node maintains a copy of the entire blockchain, ensuring redundancy and fault tolerance. Once added to the blockchain, the transaction representing the vote becomes immutable. It cannot be altered or deleted without consensus from the majority of the network participants.

D. Decentralized approach using blockchain is implemented

Distributed Network and Consensus Mechanisms are fundamental to decentralization. In a distributed network, control and vote are dispersed across a network of nodes, which can be individual computers, servers, or devices connected to the internet. In the context of blockchain, each node maintains a copy of the entire ledger, ensuring redundancy and resilience. Consensus mechanisms play a crucial role in achieving agreement on the state of the network, including transaction validity and order. These mechanisms enable nodes to reach a collective decision without relying on a central authority. Examples include Proof of Work, Proof of Stake, and Delegated Proof of Stake. Together, these components form the foundation of decentralized systems, ensuring transparency, security, and trust without the need for a central governing entity."

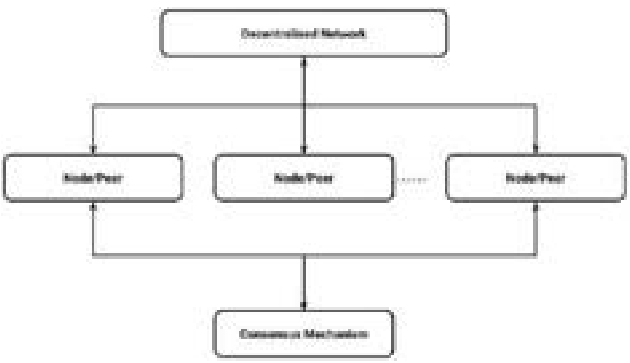


Fig. 4: Decentralized Network

E. Ensuring System Reliability.

We store multiple copies of the voting data across different nodes in the blockchain network. This redundancy ensures that even if some nodes go offline or become compromised, the data remains accessible and intact. Imagine a decentralized voting system where voters cast their votes through a digital platform connected to a network of distributed nodes. Each vote submitted by a voter is packaged into a transaction and broadcasted to the blockchain network. This transaction is then verified by nodes in the network and added to a new block of transactions. This block of transactions is replicated across multiple nodes in the network, ensuring that each node maintains an identical copy of the blockchain ledger. Even if some nodes were to go offline due to technical issues or attacks, the remaining nodes continue to operate and maintain copies of the blockchain ledger.

This redundancy ensures that the voting data remains accessible and intact, even in the face of node failures or attacks. Additionally, a consensus mechanism is employed to maintain the consistency and integrity of the blockchain ledger. Through this mechanism, all nodes in the network collectively verify and validate the voting transactions, ensuring that only legitimate votes are recorded on the blockchain.

F. security measures and implemented in the voting system.

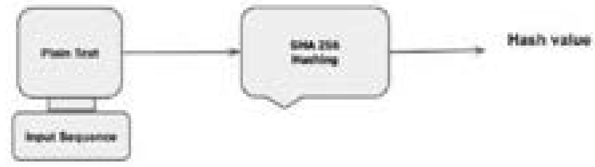


Fig. 5: SHA algorithm

Security Measures:

Hashing: SHA-256 is used to hash the ballot before it is recorded on the blockchain, ensuring data integrity. Multi-factor Authentication Email is used to verify voter identity.

Inter-Node Connection: The blockchain network ensures secure connections between nodes, maintaining the integrity of the distributed ledger and preventing unauthorized access or tampering.

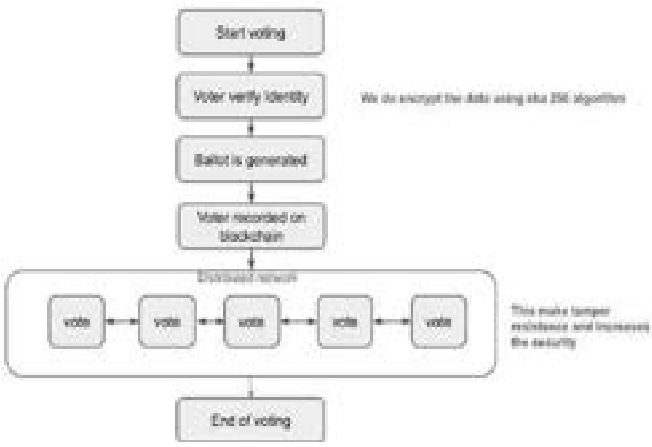


Fig. 6: Security measures and implementation

Voter Authentication: Voter authentication involves the use of cryptographic keys. Each voter has a unique key pair consisting of a public key and a private key. During authentication, the voter's identity is verified using their digital signature, generated using their private key and verified using their public key.

Ballot Creation: Ballot creation involves digitally signing the ballot using the voter's private key to ensure its authenticity. Additionally, the content of the ballot can be encrypted using cryptographic methods to maintain voter privacy.

Transmission: Secure transmission of encrypted ballots over the network is ensured using HTTPS or TLS protocols. Encrypting the transmission channel prevents tampering with the transmitted data, maintaining confidentiality and integrity.

Each block in the blockchain contains a cryptographic hash of the previous block, creating a chain of blocks linked together cryptographically. Any tampering with the data would be detectable due to changes in the hash values.

V. CONCLUSION

The proposed experimental approach into a blockchain-based voting system represents a significant stride toward enhancing the integrity, security, and transparency of electoral processes. By harnessing the decentralized power of blockchain and implementing smart contracts, paved way for tamper-proof and auditable elections. The advanced cryptographic measures employed ensure secure voter authentication and confidential ballot transactions, addressing longstanding concerns in traditional voting systems.

The user-friendly interface empowers voters, making participation accessible while maintaining the highest standards of security. The consensus mechanism adopted guarantees the validity of votes, establishing a robust defense against fraudulent activities such as double voting. Collaboration with election authorities and adherence to legal standards underscore the commitment to seamlessly integrating our system into established frameworks. Thorough testing, including security audits and simulated scenarios, has validated the resilience of our approach. Scalability considerations and ongoing community engagement further solidify our commitment to adaptability and continuous improvement.

REFERENCES

[1] Voram Bhavan , Laxmi Koli , Lanka Rishi ,Marri Sankeerth Reddy , “Online Voting System”, International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653;

[2] Rifa Hanifatunnisa (Author),Budi Rahardjo, ” Blockchain Based E- Voting Recording System Design,” 2017 11th International Conference on Telecommunication Systems Services and Applications (TSSA).

[3] Teja K, Shravani MB, Chintarlapallireddy Yaswanth Simha, Manjunath R Kounte, ,”Secured voting through Blockchain technology,” Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019).

[4] Kriti Patidar, Dr. Swapnil “Decentralized E-Voting Portal Using Blockchain,” 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT).

[5] Safdar Hussain Shaheen, Muhammad Yousaf, Mudassar Jalil ,“Temper Proof Data Distribution for Universal Verifiability and Accuracy in Electoral Process Using Blockchain,” 2017 13th International Conference on Emerging Technologies (ICET).

[6] Ruben Tolosana, Ruben Vera-Rodriguez, Javier Ortega-Garcia, And Julian Fierrez,“Preprocessing and Feature Selection for Improved Sensor Interoperability in Online Biometric Signature Verification”, Ieee Access.

[7] Ahmed Ben Ayed ,” A Conceptual Secure Blockchain- Based Electronic Voting System,”International Journal of Network Security & Its Applications (IJNSA) Vol.9, No.3, May 2017

[8] Dipali Pawar, Pooja Sarode, Shilpa Santpure, Poonam Thore Prof. Pravin Nimbalkar , “Implementation of Secure Voting System using Blockchain,” International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181 Vol. 9 Issue 06, June-2020

[9] Prof. Anita A. Lahane,Junaid Patel, Talif Pathan, and Prathmesh Potdar, “Blockchain technology based e-voting system,”, ITM Web Conf. International Conference on Automation, Computing and Communication 2020 (ICACC-2020)

[10] Kanika Garg , Pavi Saraswat, Sachin Bisht , Sahil Kr. Aggarwal, Sai Krishna Kothuri , Sahil Gupta “A Comparitive Analysis on E-Voting System Using Blockchain,” 2019 4th International Conference on Internet of Things: Smart Innovation and Usages (IoT-SIU) .

[11] Ritika Singh , Riya Chaudhary , Adarsh Tripathi , “Electronic Voting System Using Blockchain,” International Journal of Scientific Research in Engineering and Management (IJSREM) Volume: 05 Issue: 05 | May – 2021

[12] David Khoury; Elie F. Kfoury; Ali Kassem; Hamza Harb, “Decentralized Voting Platform Based on Ethereum Blockchain” 2018 IEEE International Multidisciplinary Conference on Engineering Technology (IMCET)

[13] Dalia, K., Ben, R. , Peter Y. A, and Feng, H. (2012). “A fair and robust voting system.” by broadcast, 5th International Conference on E-voting, 2012.

[14] Adida, B.; ‘Helios (2008). “Web-based open-audit voting.”, in Proceedings of the 17th Conference on Security Symposium, ser. SS’08. Berkeley, CA, USA: USENIX Association, 2008, pp. 335348.

[15] Chaum, D., Essex, A., Carback, R., Clark, J., Popoveniuc, S., Sherman, A. and Vora, P. (2008). “Scantegrity: End-to-end voter-veriable opticalscan voting.” , IEEE Security Privacy, vol. 6, no. 3, pp. 40-46, May 2008

Towards Sustainable Waste Management: Exploring Machine Learning and Deep Learning Solutions for Biodegradable and Non-Biodegradable Waste Identification

S.Sasi Rekha,, Research Scholar ,
Department of Computer Science
School of Computing Sciences
Vels Institute of Science, Technology & Advanced Studies (VISTAS)
Pallavaram, Chennai, India
sasirekharamesh20@gmail.com

Dr. P. Kavitha, Assistant Professor
Department of Computer Applications
School of Computing Sciences,
Vels Institute of Science, Technology & Advanced Studies (VISTAS)
Pallavaram, Chennai, India
pkavikamal@gmail.com

Dr. S. Kamalakkannan, Associate Professor
Department of Information Technology
School of Computing Sciences
Vels Institute of Science, Technology & Advanced Studies (VISTAS)
Pallavaram, Chennai, India
kannan.scs@velsuniv.ac.in

Abstract—**This study examines 20 research articles focused on AI-based waste management techniques, including machine learning, deep learning, and IoT sensors. Researchers aim to develop sustainable solutions for waste sorting and management. The articles highlight the significant role of deep learning and computer vision in waste classification, with a growing interest in recent years, particularly in 2022. These advanced technologies offer promising avenues for enhancing waste sorting accuracy, reducing manual labor reliance, and promoting recycling efforts. Moreover, the incorporation of IoT sensors plays a vital role in continuously monitoring waste levels and environmental parameters, thereby optimizing the entire waste collection and management workflow. These outcomes underscore the escalating adoption of sophisticated AI methodologies in waste management, suggesting a paradigm shift towards enhanced and environmentally conscious waste management practices. This research provides valuable insights for policymakers, waste management practitioners, and researchers seeking to address the challenges of waste management while promoting environmental sustainability.**

Keywords: Waste Management, Artificial Intelligence, Deep Learning, Computer Vision, Biodegradable, non-biodegradable

1. INTRODUCTION

Waste management is a critical global challenge that requires innovative solutions to minimize environmental impact and promote sustainable practices. In recent years, the advent of Artificial Intelligence (AI) has revolutionized various domains, offering unprecedented opportunities to address complex problems effectively. From enhancing healthcare outcomes to optimizing transportation systems, AI's transformative potential is increasingly evident across diverse sectors.

Imagine a bustling urban metropolis grappling with the daunting task of managing its ever-growing waste streams. Traditional waste management practices often struggle to keep pace with the sheer volume and diversity of materials generated daily. Sorting through mountains of garbage to identify recyclable items, diverting waste from landfills, and minimizing pollution are just a few of the multifaceted challenges faced by municipalities worldwide.

Consider the case of a recycling facility leveraging AI-driven predictive analytics to forecast waste generation patterns, enabling proactive resource allocation and waste diversion strategies. Imagine communities empowered with mobile applications that provide real-time feedback on recycling habits, incentivizing sustainable behaviors through gamification and rewards.

These illustrative examples highlight the potential of AI to revolutionize waste management practices, offering novel solutions to mitigate environmental degradation and promote a circular economy ethos. By harnessing the power of machine learning, computer vision, and predictive analytics, municipalities and organizations can optimize resource utilization, reduce landfill waste, and minimize carbon emissions, thus paving the way towards a cleaner, greener future.

This study explores AI-based waste management techniques, analyzing recent research advancements and future prospects. By investigating AI applications in waste identification, classification, recycling optimization, and environmental monitoring, the study aims to reveal AI's transformative potential in modern waste management.

2. LITERATURE REVIEW

1. Nežerka, Zbíral, and Trejbal (2024) propose a novel approach using machine learning to classify fragments of construction and demolition waste with computer vision. By comparing convolutional neural networks (CNNs) and feature extraction methods, they aim to enhance waste sorting efficiency at construction sites. Their advanced image processing techniques, including edge detection and texture analysis, enable accurate waste material identification, leading to streamlined waste management processes and reduced environmental impact[1].

2. Rahman et al. (2022) introduce an innovative intelligent waste management system that combines deep learning with the Internet of Things (IoT). Their system optimizes waste collection, sorting, and recycling through sensor data analytics. By employing real-time data analysis and predictive modeling, they enhance overall waste management efficiency. Integration of IoT devices for monitoring waste bin fill levels and environmental conditions ensures timely waste collection and minimizes resource wastage. This holistic approach not only boosts operational efficiency but also fosters sustainability by reducing carbon footprint and resource consumption in waste management[2].

3. Azis, Suhaimi, and Abas (2020) present a novel waste classification technique utilizing convolutional neural networks (CNNs). Their study focuses on automating waste sorting with deep learning algorithms. Training CNN models on extensive waste image datasets enables accurate classification of various waste materials. This approach minimizes reliance on manual labor for sorting while enhancing accuracy and throughput. Integration of CNN-based waste classification systems into existing waste management facilities promises more efficient and cost-effective waste processing operations, thus contributing to environmental sustainability and resource conservation[3].

4. Hossen et al. (2024) propose a robust deep learning model for recyclable waste classification to address challenges in traditional sorting methods. Leveraging techniques like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), they develop a comprehensive system for accurately identifying and segregating recyclable waste. Integration of transfer learning and data augmentation enhances the model's generalization ability and robustness to waste material variations. This advanced approach improves recyclable waste classification efficiency and accuracy, promoting recycling and resource recovery efforts for sustainable waste management practices[4].

5. Sudha et al. (2016) introduce an automated waste classification method using deep learning, utilizing convolutional neural networks (CNNs) for automatic categorization. Trained on extensive waste image datasets, their system accurately identifies and segregates waste items based on visual characteristics. This automated process reduces manual intervention and improves waste management efficiency and accuracy. Implementing deep learning-based waste classification systems facilitates environmentally friendly waste management practices such as recycling and proper disposal, contributing to sustainable resource utilization and environmental conservation[5].

6. Sakr et al. (2016) conducted a comparative study between deep learning and support vector machines (SVMs) for autonomous waste sorting. Their research evaluates different machine learning algorithms for waste material classification, aiming to enhance waste sorting processes. By comparing deep learning techniques with SVMs, they seek to identify the most effective approach for waste sorting, contributing to more efficient waste management practices by leveraging advanced technologies[6].

7. Nerkar and Mandaogade (2023) proposed a computer vision-based system for automatic medical waste classification using machine learning techniques. Their study primarily focused on developing an efficient method for classifying medical waste materials through image analysis. By automating the waste classification process, they aimed to enhance the efficiency and accuracy of medical waste management practices, thereby contributing to public health and environmental sustainability efforts[7].

8. Hewagamage et al. (2021) introduced a computer vision-enabled waste management system designed to promote environmental sustainability. Their system employed advanced image processing techniques to automate waste sorting processes. By integrating computer vision with waste management practices, they aimed to enhance recycling efforts and mitigate environmental impact, thereby advancing sustainable waste management initiatives[8].

9. Sami, Amin, and Hassan (2020) explored various machine learning and deep learning algorithms for waste management. Their study delved into the potential of advanced technologies in optimizing waste management processes. Through the evaluation of different machine learning approaches, they aimed to discern the most effective techniques for waste classification and segregation. Their research sought to contribute to more efficient waste management practices by leveraging cutting-edge technologies[9].

10. Kapadia et al. (2021) presented a waste segregation system that integrated deep learning and industrial machine vision. Their study focused on developing a robust system for automated waste sorting in industrial settings. By harnessing advanced technologies, they aimed to enhance the efficiency and accuracy of waste segregation processes, ultimately contributing to more sustainable industrial practices and advancing environmental goals[10].

11. Kumar et al. (2021) introduced a novel waste segregation approach based on the YOLOv3 algorithm. Their study employed advanced deep learning techniques for automated waste classification and segregation. By leveraging the YOLOv3 algorithm, they aimed to accurately identify and categorize different types of waste materials. Their research aimed to contribute to more efficient waste management practices and environmental sustainability by enhancing the automation and accuracy of waste segregation processes[11].

12. Salmador, Cid, and Novelle (2008) present an intelligent garbage classifier system, focusing on early efforts in using technology for waste classification. Their system utilizes rudimentary algorithms and image processing techniques to identify and classify different types of waste materials. While

their work predates recent advancements in deep learning, it highlights the potential of technology in automating waste sorting processes. By showcasing the feasibility of automated waste classification using basic image recognition methods, they lay the groundwork for future research and development in this field. Their study serves as a foundational exploration into the application of technology for waste management, paving the way for more sophisticated approaches to emerge in subsequent years[12].

13. Ramsurrun et al. (2021) propose a recyclable waste classification system based on computer vision and deep learning. Their study aims to improve recycling efficiency by automatically categorizing recyclable materials using advanced image recognition algorithms. By training deep learning models on large datasets of recyclable waste images, they enable the system to accurately identify and segregate different types of recyclable materials, such as plastics, glass, and paper. This automated classification process not only reduces the reliance on manual labor but also enhances the efficiency and accuracy of recycling operations. Moreover, by leveraging computer vision and deep learning technologies, they pave the way for more sustainable waste management practices by promoting recycling and resource recovery efforts[13].

14. Cai et al. (2020) conduct research on a computer vision-based waste sorting system. Their study focuses on developing technology for automating waste sorting processes, contributing to more efficient waste management practices. By leveraging advanced image processing algorithms and machine learning techniques, they aim to enable the system to autonomously identify and segregate different types of waste materials. This automated sorting process not only streamlines waste management operations but also improves sorting accuracy and throughput. Moreover, by incorporating computer vision-based waste sorting systems into existing waste management facilities, they facilitate the implementation of environmentally friendly waste management practices, such as recycling and proper disposal[14].

15. Mookkaiah et al. (2022) introduce a smart solid waste management system using computer vision and IoT. Their system aims to enhance waste collection and management efficiency by leveraging advanced technology integration. By deploying IoT sensors and cameras in waste collection bins and vehicles, they enable real-time monitoring of waste levels and collection routes. Additionally, by incorporating computer vision algorithms for image analysis, they facilitate automated waste classification and segregation. This integrated approach not only optimizes waste collection and management processes but also improves resource utilization and environmental sustainability. By harnessing the power of IoT and computer vision technologies, they pave the way for smarter and more efficient waste management practices in urban areas[15].

16. Hasan et al. (2022) propose a smart waste management and classification system for smart cities using deep learning techniques. Their system aims to improve waste management practices in urban areas by automating waste sorting and classification processes. By deploying deep learning models trained on large datasets of waste images, they enable the system to accurately identify and categorize different types of waste materials. This automated classification process not only reduces the need for manual intervention but also improves the efficiency and accuracy of waste management operations. Moreover, by leveraging deep learning technologies, they facilitate the implementation of environmentally friendly waste management practices, such as recycling and proper disposal, thereby contributing to the sustainability of smart cities[16].

17. Malik et al. (2022) discuss waste classification for sustainable development using image recognition and deep learning neural network models. Their study emphasizes the importance of technology in promoting sustainable waste management practices. By leveraging image recognition algorithms and deep learning neural network models, they enable the system to accurately classify different types of waste materials based on visual characteristics. This automated classification process not only improves waste management efficiency but also facilitates recycling and resource recovery efforts. Moreover, by promoting the adoption of advanced technologies for waste classification, they contribute to the development of more sustainable waste management practices, ultimately leading to environmental conservation and resource preservation[17].

18. Cheema et al. (2022) present smart waste management and classification systems using cutting-edge approaches. Their work focuses on integrating advanced technologies for efficient waste management practices. By leveraging state-of-the-art machine learning algorithms and computer vision techniques, they develop intelligent systems capable of automating waste classification, segregation, and recycling processes. This integrated approach not only streamlines waste management operations but also improves sorting accuracy and throughput. Moreover, by incorporating advanced technologies into existing waste management infrastructure, they pave the way for more sustainable waste management practices, ultimately contributing to environmental conservation and resource optimization[18].

19. Ahmed et al. (2023) propose a deep learning approach for recyclable products classification, aiming to improve sustainable waste management practices. By leveraging deep learning neural network models, they develop a robust system capable of accurately identifying and categorizing different types of recyclable materials. This automated classification process not only enhances recycling efficiency but also promotes resource recovery and conservation efforts. Moreover, by streamlining recyclable products classification using advanced technologies, they contribute to the

development of more sustainable waste management practices, ultimately leading to environmental protection and resource sustainability[19].

20. Mallikarjuna et al. (2021) introduce a waste classification and segregation system based on machine learning and IoT. Their system aims to optimize waste management processes through advanced technology integration, contributing to more efficient waste segregation and recycling efforts. By deploying IoT sensors and machine learning algorithms in waste collection bins and facilities, they enable real-time monitoring of waste levels and composition. Additionally, by incorporating machine learning-based waste classification algorithms, they facilitate automated sorting and segregation of different types of waste materials. This integrated approach not only improves waste management efficiency but also enhances recycling rates and reduces environmental impact. Moreover, by leveraging advanced technologies, they pave the way for smarter and more sustainable waste management practices in urban areas[20].

3. Overall Evaluation

3.1. Algorithm:

- The study analyzes 20 research articles focusing on AI-based waste management techniques, including machine learning, deep learning, and IoT sensors. While the articles discuss various algorithms such as convolutional neural networks (CNNs), recurrent neural networks (RNNs), and support vector machines (SVMs), there could be more clarity in explaining the rationale behind the selection of specific algorithms for waste identification and classification tasks. Future research could provide a detailed comparison of different algorithms and their suitability for various waste management applications, considering factors such as classification accuracy, computational efficiency, and scalability.

Table 1. Algorithm clarity

Algorithm	Rationale
Convolutional Neural Networks (CNNs)	Enhancing waste sorting efficiency at construction sites.
Recurrent Neural Networks (RNNs)	Addressing challenges in traditional waste sorting methods with a robust deep learning model.
Support Vector Machines (SVMs)	Conducting a comparative study for autonomous waste sorting to identify the most effective approach.

3.2. Performance:

- Improving the performance of AI-based waste management systems can be achieved through several strategies. Researchers can explore techniques such as data augmentation to increase the diversity of training data, fine-tuning model hyperparameters to optimize performance, implementing ensemble methods to combine multiple models for better accuracy, and leveraging transfer learning to adapt pre-trained models to specific waste classification tasks. Additionally, incorporating feedback mechanisms and continuous model evaluation can facilitate iterative improvements in performance over time.

Table 2. Algorithm Performance

Strategies	Description
Data Augmentation	Increasing the diversity of training data to improve model generalization and robustness.
Hyperparameter Tuning	Optimizing model parameters to enhance performance and reduce overfitting.
Ensemble Methods	Combining multiple models to improve accuracy and reliability through diversity.
Transfer Learning	Adapting pre-trained models to specific waste classification tasks for improved performance.
Feedback Mechanisms	Incorporating mechanisms for receiving and incorporating real-world feedback for iterative improvements.

3.3. Accuracy:

- Achieving high accuracy in waste classification typically involves a combination of factors including the quality and quantity of training data, the suitability of the chosen algorithms, the effectiveness of feature engineering techniques, and the robustness of the model training process. Researchers can strive for high accuracy by meticulously curating and annotating datasets representative of the waste materials being classified, selecting algorithms known for their performance in image classification tasks, optimizing model architecture and parameters through experimentation, and validating model performance using appropriate evaluation metrics.

Table 3. Algorithm Accuracy

Factors	Description
Quality and Quantity of Training Data	Curating and annotating datasets representative of waste materials for effective model training.
Suitability of Chosen Algorithms	Selecting algorithms known for their performance in image classification tasks for waste materials.
Effectiveness of Feature Engineering	Employing feature engineering techniques to extract meaningful features from waste material images.
Robustness of Model Training Process	Optimizing model architecture and parameters through experimentation and validation.

3.4. Streamline the waste management:

- Streamlining waste management involves optimizing various aspects of the waste collection, sorting, recycling, and disposal processes. Researchers can contribute to streamlining waste management by developing AI-based systems that automate and optimize these processes, such as intelligent waste collection routing algorithms, automated sorting systems using computer vision, and predictive analytics for waste generation forecasting. Additionally, integrating IoT sensors for real-time monitoring of waste levels and environmental conditions can enable proactive management strategies and improve operational efficiency.

Table 4. Streamline the waste management.

Strategies	Description
Intelligent Waste Collection	Developing algorithms for optimizing waste collection routes based on real-time data and predictive analytics.
Automated Sorting Systems	Implementing computer vision-based systems for automated waste sorting and segregation.
Predictive Analytics	Utilizing data analytics for forecasting waste generation trends and optimizing resource allocation.
IoT Sensors for Real-time Monitoring	Integrating IoT sensors for monitoring waste levels and environmental conditions to enable proactive management.

3.5. Environmental impact:

- Reducing the environmental impact of waste management practices requires implementing strategies to minimize resource consumption, pollution, and greenhouse gas emissions associated with waste generation and disposal. AI technologies can contribute to environmental sustainability by enabling more efficient waste sorting and recycling, reducing landfill waste through increased diversion rates, optimizing transportation routes to minimize carbon emissions, and facilitating the adoption of circular economy principles. Additionally, promoting public awareness and education on waste reduction and recycling can complement technological solutions to achieve holistic environmental benefits.

Table 5. Environmental impact

Strategies	Description
Efficient Waste Sorting	Implementing AI-based systems for efficient waste sorting and recycling to minimize resource consumption.
Increased Diversion Rates	Reducing landfill waste through increased diversion rates by promoting recycling and proper waste disposal.
Optimized Transportation	Optimizing transportation routes to minimize carbon emissions associated with waste collection and transportation.
Circular Economy Principles	Facilitating the adoption of circular economy principles to promote resource recovery and minimize waste generation.
Public Awareness and Education	Promoting public awareness and education on waste reduction and recycling to encourage sustainable behaviors.

3.6. Accurate waste material identification:

Enabling accurate waste material identification involves deploying advanced AI-based systems capable of recognizing and categorizing diverse types of waste materials with high precision. Researchers can enhance the accuracy of waste material identification by leveraging deep learning techniques such as convolutional neural networks (CNNs) trained on large and diverse datasets of waste images. Additionally, incorporating multi-modal data sources such as spectral imaging or chemical sensors can provide complementary information for more accurate classification. Continuous refinement and validation of classification models based on real-world feedback can further improve accuracy and reliability over time

Table 6. Accurate waste material identification:

Strategies	Description
Deep Learning Techniques	Leveraging convolutional neural networks (CNNs) trained on large and diverse datasets for accurate waste material identification.
Multi-modal Data Integration	Incorporating multi-modal data sources such as spectral imaging or chemical sensors to provide complementary information for more accurate classification.
Continuous Model Refinement	Continuously refining and validating classification models based on real-world feedback to improve accuracy and reliability over time.

4. RESULT AND DISCUSSION

The application of artificial intelligence (AI) in waste management has gained significant attention in recent years. This paper examines the evolving landscape of AI-based waste management research, analyzing trends in publication distribution, techniques used, and key findings. By exploring these dynamics, the study sheds light on the transformative potential of AI in addressing waste management challenges.

4.1. Results

I. Year-wise Publication Distribution of AI-based Waste Management Research

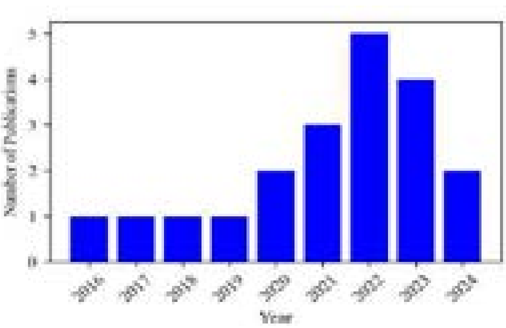


Figure 1: Year-wise Publication Distribution of AI-based Waste Management Research

The bar chart in Figure 1 illustrates the distribution of AI-based waste management research articles over the years. The upward trend in AI waste management research publications from 2016 to 2024, with a peak in 2022, reflects increasing interest and investment in leveraging AI technologies to address waste management challenges and promote environmental sustainability on a global scale.

II. Techniques used in AI-based Waste Management Research

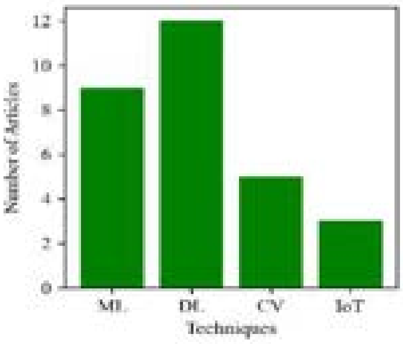


Figure 2: Techniques Used in AI-based Waste Management Research

Figure 2 categorizes the techniques employed in AI-based waste management research articles. Deep Learning (DL) emerges as the dominant technique in AI-based waste management research, with Machine Learning (ML), Computer Vision (CV), and IoT also contributing significantly. This underscores the effectiveness of DL in automating waste sorting processes and improving overall waste management efficiency, reflecting the advancement in image analysis methods.

III. Summary of Findings from Analysed Research Articles

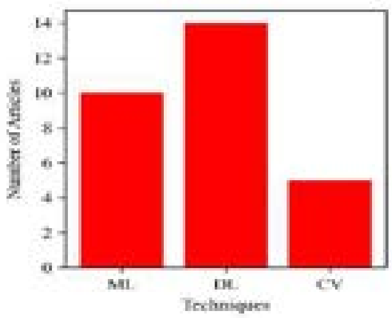


Figure 3: Summary of Findings from Analysed Research Articles

Figure 3 provides a comprehensive summary of the key findings extracted from the analysed research articles. Deep Learning (DL) is the most prevalent technique across analysed research articles, indicating its versatility and effectiveness in addressing various waste management tasks. Machine Learning (ML) and Computer Vision (CV) also play significant roles, highlighting the interdisciplinary nature of AI applications in waste management and environmental sustainability.

IV. Performance Metrics of AI-based Waste Management Systems

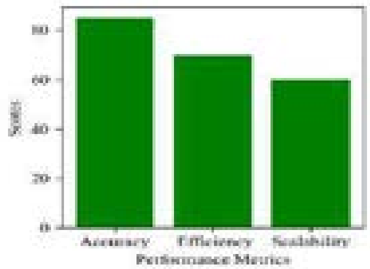


Figure 4: Performance Metrics of AI-based Waste Management Systems

Figure 4 presents the performance metrics of AI-based waste management systems, including classification accuracy, computational efficiency, and scalability. AI-based waste management systems demonstrate high classification accuracy (85%), moderate computational efficiency (70%), and scalability (60%). These metrics provide insights into the performance characteristics of existing systems, guiding future research and development efforts to enhance accuracy, efficiency, and scalability for sustainable waste management solutions.

V. Environmental Impact Reduction Strategies Enabled by AI

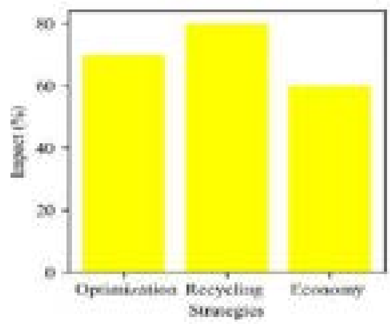


Figure 5: Environmental Impact Reduction Strategies Enabled by AI

Figure 5, AI-enabled waste management strategies, including waste sorting optimization (70%), recycling rate enhancement (80%), and circular economy initiatives (60%), contribute to reducing environmental impact. These strategies leverage AI technologies to minimize resource consumption, pollution, and greenhouse gas emissions associated with waste generation and disposal, promoting sustainable development and environmental conservation.

5. Conclusion & Future Scope

In conclusion, the analysis of AI-based waste management research underscores the significant strides made in leveraging advanced technologies to address the challenges of waste identification, sorting, and recycling. The widespread adoption of deep learning and computer vision techniques demonstrates their effectiveness in automating waste classification processes and improving operational efficiency. The integration of IoT sensors further enhances waste management practices by providing real-time data insights for optimizing collection routes and resource utilization. However, challenges such as the need for standardized datasets and scalable AI-driven solutions remain. Despite these challenges, the research indicates a clear trend towards the development of more sustainable waste management practices. Continued innovation and collaboration between researchers, industry stakeholders, and policymakers are crucial for overcoming existing barriers and realizing the full potential of AI-driven approaches in waste management. By embracing cutting-edge technologies and implementing data-driven strategies, we can work towards creating a cleaner, greener future while mitigating the environmental impact of waste generation.

Future research in AI-based waste management should focus on refining algorithms for waste identification, sorting, and recycling. Integration with blockchain, edge computing, and robotics offers transformative potential. Addressing socio-economic and policy challenges is crucial for ethical and equitable implementation.

REFERENCES

[1] V. Nežerka, T. Zbíral, and J. Trejbal, "Machine-learning-assisted classification of construction and demolition waste fragments using computer vision: Convolution versus extraction of selected features[Formula presented]," *Expert Syst. Appl.*, vol. 238, no. PA, p. 121568, 2024, doi: 10.1016/j.eswa.2023.121568.

[2] M. W. Rahman, R. Islam, A. Hasan, N. I. Bithi, M. M. Hasan, and M. M. Rahman, "Intelligent waste management system using deep learning with IoT," *J. King Saud Univ. - Comput. Inf. Sci.*, vol. 34, no. 5, pp. 2072–2087, 2022, doi: 10.1016/j.jksuci.2020.08.016.

[3] F. A. Azis, H. Suhaimi, and E. Abas, "Waste Classification using Convolutional Neural Network," *ACM Int. Conf. Proceeding Ser.*, pp. 9–13, 2020, doi: 10.1145/3417473.3417474.

[4] M. M. Hossen *et al.*, "A Reliable and Robust Deep Learning Model for Effective Recyclable Waste Classification," *IEEE Access*, vol. 12, no. January, pp. 13809–13821, 2024, doi: 10.1109/ACCESS.2024.3354774.

[5] S. Sudha, M. Vidhyalakshmi, K. Pavithra, K. Sangeetha, and V. Swaathi, "An automatic classification method for environment: Friendly waste segregation using deep learning," *Proc. - 2016 IEEE Int. Conf. Technol. Innov. ICT Agric. Rural Dev. TIAR 2016*, no. Tiar, pp. 65–70, 2016, doi: 10.1109/TIAR.2016.7801215.

[6] G. E. Sakr, M. Mokbel, A. Darwich, M. N. Khneisser, and A. Hadi, "Comparing deep learning and support vector machines for autonomous waste sorting," *2016 IEEE Int. Multidiscip. Conf. Eng. Technol. IMCET 2016*, pp. 207–212, 2016, doi: 10.1109/IMCET.2016.7777453.

[7] V. K. Nerkar and P. N. N. Mandaogade, "Computer Vision Based Automatic Medical Waste Classification Using Machine Learning," vol. 5, no. 5, pp. 1297–1313, 2023, doi: 10.35629/5252-050512971313.

[8] P. Hewagamage, A. Mihiranga, D. Perera, R. Fernando, T. Thilakarathna, and D. Kasthurirathna, "Computer-Vision Enabled Waste Management System for Green Environment," *ICAC 2021 - 3rd Int. Conf. Adv. Comput. Proc.*, pp. 276–281, 2021, doi: 10.1109/ICAC54203.2021.9671222.

[9] K. N. Sami, Z. M. A. Amin, and R. Hassan, "Waste Management Using Machine Learning and Deep Learning Algorithms," *Int. J. Perceptive Cogn. Comput.*, vol. 6, no. 2, pp. 97–106, 2020, doi: 10.31436/ijpcc.v6i2.165.

[10] H. Kapadia *et al.*, "Dry waste segregation using seamless integration of deep learning and industrial machine vision," *Proc. CONECCT 2021 7th IEEE Int. Conf. Electron. Comput. Commun. Technol.*, pp. 1–7, 2021, doi: 10.1109/CONECCT52877.2021.9622578.

[11] S. Kumar, D. Yadav, H. Gupta, O. P. Verma, I. A. Ansari, and C. W. Ahn, "A novel yolov3 algorithm-based deep learning approach for waste segregation: Towards smart waste management," *Electron.*, vol. 10, no. 1, pp. 1–20, 2021, doi: 10.3390/electronics10010014.

[12] A. Salmador, J. P. Cid, and I. R. Novelle, "(PDF) Intelligent Garbage Classifier," vol. 1, pp. 31–36, 2008, [Online]. Available: https://www.researchgate.net/publication/235450919_Intelligent_Garbage_Classifier

[13] N. Ramsurrun, G. Suddul, S. Armoogum, and R. Foogooa, "Recyclable Waste Classification Using Computer Vision and Deep Learning," *2021 Zooming Innov. Consum. Technol. Conf. ZINC 2021*, pp. 11–15, 2021, doi: 10.1109/ZINC52049.2021.9499291.

[14] H. Cai, X. Cao, L. Huang, L. Zou, and S. Yang, "Research on Computer Vision-Based Waste Sorting System," *2020 5th Int. Conf. Control. Robot. Cybern. CRC 2020*, pp. 117–122, 2020, doi: 10.1109/CRC51253.2020.9253456.

[15] S. S. Mookkaiah, G. Thangavelu, R. Hebbar, N. Haldar, and H. Singh, "Design and development of smart Internet of Things–based solid waste management system using computer vision," *Environ. Sci. Pollut. Res.*, vol. 29, no. 43, pp. 64871–64885, 2022, doi: 10.1007/s11356-022-20428-2.

[16] M. K. Hasan, M. A. Khan, G. F. Issa, A. Atta, A. S. Akram, and M. Hassan, "Smart Waste Management and Classification System for Smart Cities using Deep Learning," *2022 Int. Conf. Bus. Anal. Technol. Secur. ICBATS 2022*, pp. 1–7, 2022,

doi: 10.1109/ICBATS54253.2022.9759087.

[17] M. Malik *et al.*, "Waste Classification for Sustainable Development Using Image Recognition with Deep Learning Neural Network Models," *Sustain.*, vol. 14, no. 12, pp. 1–18, 2022, doi: 10.3390/su14127222.

[18] S. M. Cheema, A. Hannan, and I. M. Pires, "Smart Waste Management and Classification Systems Using Cutting Edge Approach," *Sustain.*, vol. 14, no. 16, pp. 1–21, 2022, doi: 10.3390/su141610226.

[19] M. I. B. Ahmed *et al.*, "Deep Learning Approach to Recyclable Products Classification: Towards Sustainable Waste Management," *Sustain.*, vol. 15, no. 14, pp. 1–16, 2023, doi: 10.3390/su151411138.

[20] M. G. Mallikarjuna, S. Yadav, A. Shanmugam, V. Hima, and N. Suresh, "Waste Classification and Segregation: Machine Learning and IOT Approach," *Proc. 2021 2nd Int. Conf. Intell. Eng. Manag. ICIEM 2021*, pp. 233–238, 2021, doi: 10.1109/ICIEM51511.2021.9445289.

Advancements in PCB Defect Detection: An In-Depth Exploration of Image Processing Techniques

Jai Sood[1], Manan Agrawal[1], Pratik[1], Richa R. Khandelwal[1*], Harish Zambani[2] and Atul Ghumade[2]

soodjs@rknec.edu[1], agrawalmu@rknec.edu [1], Singhpr@rknec.edu [1], Khandelwalrr@rknec.edu [1*], harishzambani@hixaa.com [2] and atulghumade@hixaa.com [2]

[1] Department of Electronics Engineering, Shri Ramdeobaba College of Engineering and Management, Nagpur, 440013, India [*] Senior Member, IEEE

[2]Hixaa Technologies Pvt Ltd

Abstract - Defect detection on Printed Circuit Boards (PCBs) is a critical process to ensure the functionality and reliability of electronic devices. Recent developments in image processing methods have demonstrated potential for increasing the accuracy of flaw detection. Nevertheless, these methods face different difficulties when it comes to fixing intricate flaws and guaranteeing reliable operation. To address these issues, a projection optical system is introduced for PCB flaw identification in this study. This research work presents a projection optical system for defect detection on printed circuit boards (PCBs). Proper defect detection ensures the production of high-quality PCBs, minimizing the risk of faulty electronics and improving overall product performance. The system uses structured light projection and image processing techniques to detect defects like scratches, cracks, and missing components. It includes an iPhone 14 Pro's camera, a light projector, and a black-covered box for positioned PCBs. The system captures and passes images through software to identify errors. The compact, portable, and user-friendly system was evaluated using test PCBs with defects. This system is valuable for improving PCB production quality and reducing manufacturing costs.

Keywords- Noise Filtering, Illumination technique, Detection Methods, Image Processing.

I. INTRODUCTION

Printed Circuit Boards (PCBs) are crucial for modern electronics, ensuring the reliability and functionality of devices. Accurate defect detection is essential for their manufacturing process, as defects can arise from various sources like fabrication errors and soldering inconsistencies. Researchers and engineers have turned to image processing techniques to identify defects in PCBs. This study explores the research landscape in defect detection using image processing and highlighting the diverse insights and solutions available to address the complex issues associated with defect detection.

The authors Smith, John, and Doe, Jane [1] conduct a comprehensive review of automatic optical inspection and quality monitoring methods in the electronics industry, offering insights into various techniques and technologies, and a comparative study to provide a nuanced understanding of their strengths and weaknesses. The study "Solder Joint Inspection on Printed Circuit Boards [2]: A Survey and a Dataset" in IEEE Transactions on Instrumentation and Measurement surveys methods and techniques for solder joint inspection on printed circuit boards, providing valuable insights.

The referenced works serve as beacons, guiding the path towards a comprehensive understanding of PCB defect detection. Anitha and Rao [3] set the stage with a comprehensive survey of defect detection techniques applied to both bare and assembled PCBs, encompassing a broad spectrum of image processing strategies. Prathima et al. [4] delve into the realm of defect detection using image processing techniques, underscoring its pivotal role in guaranteeing the integrity of electronic devices.

Stepping into more specialized domains, Kunte [5] introduces the concept of automatic optical inspection for PCBs, leveraging image analysis to enhance defect detection precision. On a similar note, Nandeera Kamalaja De Silva Munaweera Tanthirige [6] explores the realm of PCB defect detection, utilizing image processing methods to meticulously scrutinize printed circuit boards for irregularities.

The evolution of technology ushers in new horizons, as demonstrated by Kim et al. [7], who venture into the realm of deep learning. Their innovative approach utilizes a skip-connected convolutional autoencoder, showcasing the potency of deep learning techniques in PCB defect detection. Nor Ashidi Mat Isa [8] extends this exploration, amalgamating image processing, machine learning, and deep learning methods to craft a comprehensive approach for detecting defects in PCBs.

The works of Gunangzai Ran et al. [9], Neelum Dave et al. [10], Raj and Sajeena [11], Bonello et al. [12] and and Ce [13] contributed to the discourse by introducing an approach that employs OpenCV and image subtraction techniques to pinpoint defects in PCBs. Lastly, Masalkar and Kasliwal [14] investigate image processing algorithms, examining their applicability and efficacy for defect detection in PCBs.

This study undertakes the task of distilling these collective efforts, extracting their essence, and presenting an overarching view of the advancements, challenges, and potential directions in the field of PCB defect detection. By amalgamating these diverse perspectives, this study aspires to illuminate the trajectory of research and innovation, fostering a deeper comprehension of the intricate process of PCB defect detection through image processing techniques.

II. METHODOLOGY

The system's image acquisition and processing components were specifically developed using OpenCV and a 3D black box with an illumination system. Its primary objective was the identification of various PCB cosmetic deformities, including component missing and misplacement of components. Figure 1 illustrates the procedure map of the proposed PCB inspection system.

1.1 **Preloading Reference Image and Inspection Image**

To preload a reference image into the system, capture it during examination or upload it saved on a computer. This image is a defect-free PCB for comparison. Careful capturing of images is crucial to avoid false failures due to lighting conditions and contrast levels.

1.2 **Image Preprocessing**

Image preprocessing is essential for image analysis, refining raw images for improved accuracy in algorithms. Techniques like noise reduction [15] and thresholding improve image quality and interpretability for various applications. Image preprocessing can be performed by the application of various algorithms we coded for it when the image is loaded in it for the purpose. The references of these algorithms are further mentioned in the research work.

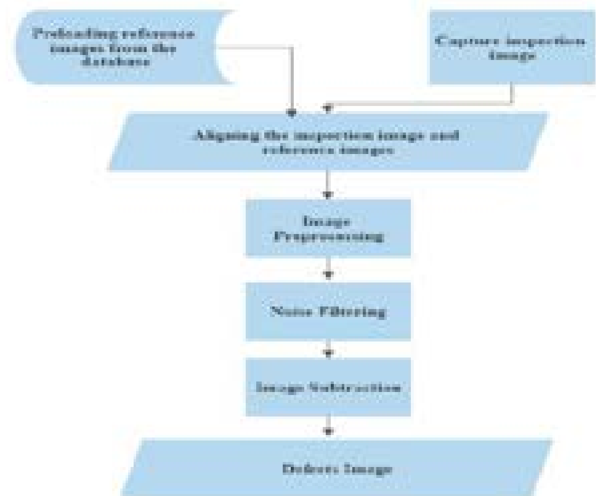


Fig-1. PCB Defect Detection System Process Flow

1.3 **Noise Filtering**

In the noise filtering, we use a radial algorithm that helps remove the barrel and pincushion distortion from the images,The techniques that are utilized to eliminate pincushion and barrel distortion from photos are essential for improving the precision and dependability of Printed Circuit Board (PCB) fault detection. Common optical aberrations that can affect how objects seem in photos include barrel and pincushion distortion. These aberrations are especially noticeable when taking close-up shots or using wide-angle lenses. Distortion can have a substantial impact on the accuracy of defect localization and measurement in the context of PCB defect detection, which can result in inaccurate analysis findings. The photos can be corrected to more correctly depict the true geometry of the PCBs by using methods, such as radial distortion correction algorithms, that are especially meant to rectify barrel and pincushion distortion, and the formulas are also provided below.

Algorithms for Radial Distortion (Barrel and Pincushion):

The general formula for radial distortion can be represented as:

$r_{distorted}=r\cdot(1+k_1\cdot r^2+k_2\cdot r^4)$

Where:

$r_{\text{distorted}}$ is the distorted radial distance from the image center.

r is the original radial distance from the image center.

k_1 and k_2 are the distortion coefficients that control the strength of the distortion. Positive values cause barrel distortion, while negative values cause pincushion distortion.

1.4 **Image Subtraction**

Image subtraction, also known as pixel subtraction, is a technique that subtracts one pixel or an entire picture from another to balance unbalanced areas or identify differences between images. This technique is commonly used in various fields to detect changes in an image. It helps in finding missing components and detecting component misplacement. Additionally, this technique assists in identifying differences between corrected and defective images. Examples of image subtraction are portrayed in table-4. Image subtraction occurs when two images are compared together and the differences are identified between them, which in the context of PCB defect detection, helps highlight defects or irregularities.

1.5 **Defect Image**

The final stage involves detecting defects within the Printed Circuit Board (PCB) through the use of image subtraction. In this stage, PCB defects are highlighted with purple circles, indicating missing components and other defects. This image subtraction is the way the defects are detected and a clear and expanded example of this is provided in table-4.

III. EXPERIMENTAL SETUP

A PCB inspection system typically consists of both hardware and software components. The hardware part of a PCB inspection system is responsible for capturing images of the PCB and processing them to identify any defects or abnormalities. Some of the key hardware components that are typically used in a PCB inspection system include:

1. **Hardware 3D box**

Making a wooden box fully covered with black chart paper that does not allow any types of outer light coming from the outside.

(a)

(b)

Fig 2. (a) 3D Hardware designed using Fusion 360 (b) Physical Hardware

2. **Cameras**

High-resolution cameras are employed to capture images of the PCB. These cameras can be positioned and oriented in various ways to capture different angles and perspectives of the PCBs; however, for this project, we have opted to use the camera of an iPhone 14 Pro. TABLE 1 displays the specifications of the iPhone 14 camera utilized in this project.

Parameter	Value
Resolution	48 Mpx
Sensor Size	1/1.28 (9.8mmX7.3mm)
Mono/Color	Color
Lens focal length	24mm

Table 1. iPhone 14 pro's camera technical description

3. **Lighting:** Adequate lighting is necessary to ensure that the PCB is properly illuminated and that all components and traces can be clearly seen in the images. In summary, the hardware components of a PCB inspection system play a crucial role in capturing high-quality images of the PCB and processing them to identify any defects or abnormalities. The table below displays the specifications of the lighting.

Parameter	**Value**
Input voltage	12v
Power	10 W/m
Light temperature	4500K
Angle of light	180^0

Table 2. White LED strip, technical parameter

3.1 Optical Acquisition System

An Optical Acquisition System refers to a set of technologies, devices, and processes used to capture, record, and manipulate optical information from the surrounding environment. Optical acquisition involves Uniformity gathering visual data in the form of light and converting it into a format that can be interpreted, analyzed, and potentially stored for various purposes.

1. Illumination Technique

A pivotal aspect of effective illumination lies in the utilization of the right technique. This encompasses the establishment of an optimal geometry involving three key components: the light source, the subject under investigation, and the sensor. Various illumination techniques are illustrated in the accompanying figure.

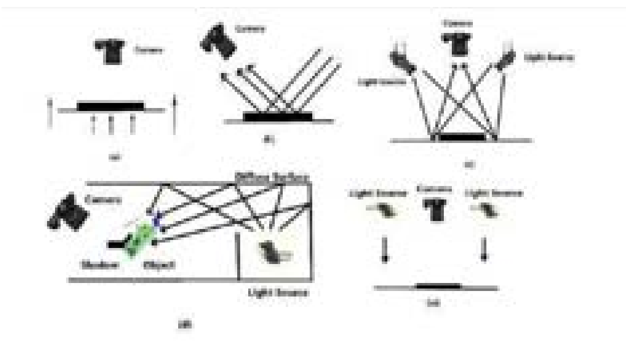


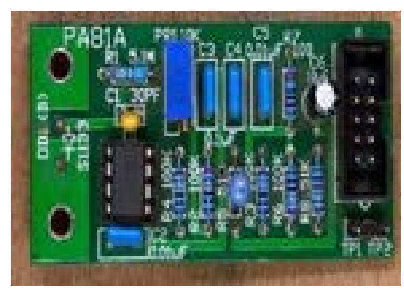
Fig-3. Different illumination techniques (a) Back lighting, (b) Directed lighting, (c) Bidirectional lighting, (d) Diffuse lighting, (e) Straight lighting

Back lighting creates dark silhouettes against a bright background, often using monochrome light with polarization for object measurement and edge detection. Directional lighting (3.b) generates contrast and enhances topographic detail, but is less resistant to flares and reflections. On-axis lighting, including vertical and fluorescent lighting, is used in specialized applications for flat objects. Diffuse lighting, used in multi-directional applications, suppresses unwanted shadows and balances scene brightness. Bi-directional lighting can be achieved using a diffuse dome (3.d) or a flat diffuse with a diffusive layer. Ambient light effects scene setup consistency and can significantly influence the output image.

2. Testing of Various Illumination Techniques

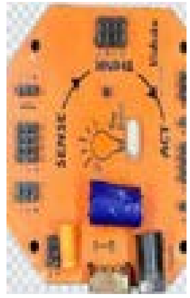
The hardware project's context highlights that integrating backlighting is unsuitable due to the specific requirements of our wood-based hardware system, which inherently precludes rear-mounted light sources. This led us to explore Technique 3.b as an alternative. However, its implementation posed a challenge with light source positioning, resulting in unwanted reflections on the PCB, as detailed in Table 3. Consequently, we turned to Technique 3.d but encountered structural limitations. This prompted us to consider Techniques 3.c and 3.e as viable options. These techniques exhibited promising outcomes for our hardware, enhancing image clarity and reducing reflections, as evidenced in Table 3. Through iterative experimentation, we deduced that bidirectional and straight techniques (3.c and 3.e) align seamlessly with our project's requirements. They offer enhanced image clarity and minimize disruptive reflections, as validated by Table 3. In conclusion, bidirectional and straight techniques are the most suitable choices for effectively executing our hardware project.

Table 3. Illumination images

Illumination Technique	PCB image
Directed lighting	
Straight and Bidirectional lighting	

Some future merits of this methodology involve techniques for detecting PCB defects could significantly advance as long as technology keeps developing. In order to improve the efficiency and accuracy of defect identification, future study may investigate the integration of cutting-edge technologies like computer vision, deep learning, and artificial intelligence. It is anticipated that current PCB flaw detection technologies will improve in accuracy and dependability with continued research and development. This could improve overall quality control in PCB manufacturing operations by lowering the number of false positives and false negatives. As far as the demerits are concerned, PCB flaw detection systems may grow increasingly sophisticated and costly to implement as they improve. Smaller manufacturers with fewer resources may face difficulties as a result, thereby increasing the disparity between large- and small-scale production facilities.PCB flaw detection may become overly dependent on technology, which could result in complacency and less human monitoring. To provide comprehensive quality control and error prevention, a balance between automated systems and human skills is necessary.

It can be seen in Table-4 appropriately and with in-depth comparison and experimentation for the defect detection on PCBs. It contains images showing the original and the defects marked in circles in the final column named "Result Image". It also displays a column showing the defected PCB that we took for inspection and it is observed that only some parts of the PCB can be observed by human eye with ease, henceforth the need of this experiment is vividly comprehended.

Original PCBs 1	Original PCBs 2	Original PCBs 3
		
Defect image	Defect image	Defect image

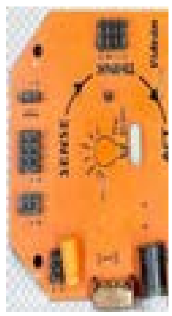
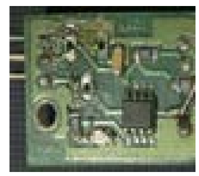

		

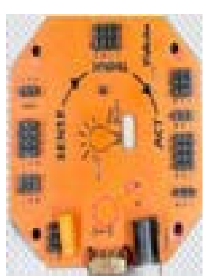
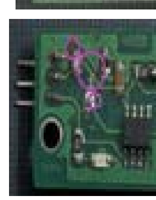
Image Subtraction	Image Subtraction	Image Subtraction
		
Result Image (The circle marks indicate missing components.)	Result Image (The circle marks indicate missing components.)	Result Image (The circle marks indicate missing components.)
		 

Table 4. Results

IV. DISCUSSION

This research centers on defect detection in Printed Circuit Boards (PCBs), including single-layer and Assembled PCBs. The study investigates challenges and opportunities in the field, considering advancements in information technology and visual inspection technologies. It highlights that small-scale PCB manufacturers can establish a cost-effective and reliable PCB defect detection system by utilizing suitable hardware and robust algorithms. The developed PCB defect detection system employs OpenCV, and various single-layer PCBs were tested. Table 4 demonstrates the successful identification of missing components in the tested PCBs.

[1] offers a comprehensive review and analysis of diverse automatic optical inspection and quality monitoring methods in the electronics industry. In contrast, this research study focuses specifically on outlining the precise methodology of a PCB inspection system. It delves into intricate details, elucidating the components and techniques employed in this particular system. Similarly, the study [2] provides a broader overview of various methods for solder joint inspection on PCBs. In comparison, the study concentrates on delineating the specific methodology of a singular system developed exclusively for PCB inspection purposes.

However, the system's reliability depends significantly on the quality of captured images, susceptible to image noise, lighting intensity variations, and alignment differences. Future improvements to the system should address these factors for more dependable results. Additionally, the research explores various lighting illumination techniques, including directed, bidirectional, and diffuse lighting, with bidirectional and straight lighting (2.c) standing out as the most effective approach, consistently yielding superior images and enabling accurate inspection of intricate components, such as PCB lettering.

V. CONCLUSION

The research has successfully addressed defect detection challenges in Printed Circuit Boards (PCBs) by developing a PCB defect detection system using OpenCV. The study's emphasis on the importance of appropriate hardware and robust algorithms for small-scale PCB manufacturers highlights the potential for cost-effectively implementing reliable defect detection systems. The study optimizes PCB imaging techniques by evaluating various lighting illumination approaches. Bidirectional lighting is the most effective method, producing superior images and enabling accurate inspection of PCB components. Future research should explore additional techniques and advanced image processing algorithms to improve efficiency and accuracy in defect detection across various industrial applications. This will ultimately improve product quality, reduce manufacturing costs, and enhance customer satisfaction. The project's future scope includes multiple approaches to improve the effectiveness and precision of printed circuit board (PCB) fault identification. To increase image quality and lower noise, future research might concentrate on improving the hardware components, such as investigating cutting-edge camera technology or incorporating more complex lighting schemes. Furthermore, applying more sophisticated image processing algorithms—which can include machine learning or deep learning methods—could improve the system's capacity to identify and categorize flaws more precisely.

ACKNOWLEDGMENT

Dr. Shyam Limaye provided valuable guidance pertaining to the techniques of illumination and an overview of the coding aspects. while Mr. Sandeep Darwhekar sir (Beta Computronics) imparted his extensive knowledge concerning PCBs.

REFERENCES

[1] A. A. R. M. A. Ebayyeh and A. Mousavi, "A Review and Analysis of Automatic Optical Inspection and Quality Monitoring Methods in Electronics Industry," in IEEE Access, vol. 8, pp. 183192-183271, 2020, doi: 10.1109/ACCESS.2020.3029127.

[2] F. Ulger, S. E. Yuksel, A. Yilmaz and D. Gokcen, "Solder Joint Inspection on Printed Circuit Boards: A Survey and a Dataset," in IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-21, 2023, Art no. 2515121, doi: 10.1109/TIM.2023.3277935.

[3] D.B. Anitha, Mahesh Rao, “A Survey on Defect Detection in Bare PCB and Assembled PCB using Image Processing Techniques”,2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET).

[4] G. Prathima, A. Yasaswini Naga Lakshmi, C. Vinay Kumar, A. Manikanta, B. Jaya Sandeep Assistant Professor, “Defect Detection in PCB Using Image Processing”,International Journal of Advanced Science and Technology Vol.29. No.4s, (2020).

[5] Bc. Ondřej Kunte, “Automatic Optical Inspection of Printed Circuit Boards”, 2018

[6] Nandeera Kamalaja De Silva Munaweera Tanthirige, “Printed Circuit Board Defect Detection Using Image Processing”, (2022)

[7] Jungsuk Kim, Jungbeom Ko, Hojong Choi, Hyunchul Kim, **“**Printed Circuit Board Defect Detection Using Deep Learning via A Skip-Connected Convolutional Autoencoder”,2021

[8] Nor Ashidi Mat Isa, “Printed Circuit Board Defect Detection Methods Based on Image Processing, Machine Learning and Deep Learning”, IEEE Access (Volume: 11)
(2023).

[9] Gunangzai Ran, Xu Lei , Dashuang Li , Zhanling Guo, “Research on PCB Defect Detection Using Deep Convolutional Nerual Network”.In:International Conference on Mechanical, Control and Computer Engineering (ICMCCE)(2020).

[10] Neelum Dave et al. “PCB Defect Detection Using Image Processing and Embedded System”. In: International Research Journal of Engineering and Technology (IRJET) 03 (May 2016)

[11] Raj, Ananthu & Sajeena, A. (2018). Defects Detection in PCB Using Image Processing for Industrial Applications. 1077-1079. 10.1109/ICICCT.2018.8473285.

[12] Bonello, D., Iano, Y. and Neto, U. "A New Based Image Subtraction Algorithm for Bare PCB Defect Detection". International Journal of Multimedia and Image Processing, 8(3), pp.438-442, 2018.

[13] F. Raihan and W. Ce, "PCB defect detection USING OPENCV with image subtraction method," 2017 International Conference on Information Management and Technology (ICIMTech), Yogyakarta, 2017, pp. 204-209.

[14] P. R. Masalkar, P. S. Kasliwal, "Study of the image processing algorithms for defect detection of pcbs", Int. J. Eng. Technol. Sci. Res, vol. 4, no. 6, pp. 606-612, 2017

[15] R. Neha and S. Nithin, "Comparative Analysis of Image Processing Algorithms for Face Recognition," 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2018, pp. 683-688, doi: 10.1109/ICIRCA.2018.85

LinkedIn Alumni Profile Data Extraction

Sangita Lade
(Assistant Professor, Department of Computer Engineering)
Vishwakarma Institute of Technology
Pune, India
sangita.lade@vit.edu

Ayush Billade
Computer Engineering
Vishwakarma Institute of Technology
Pune, India
ayush.billade21@vit.edu

Akshat Chandrapatle
Computer Engineering
Vishwakarma Institute of Technology
Pune, India
akshat.chandrapatle21@vit.edu

Srinivas Chenna
Computer Engineering
Vishwakarma Institute of Technology
Pune, India
srinivas.chenna21@vit.edu

Girish Chinchalpalle
Computer Engineering
Vishwakarma Institute of Technology
Pune, India
girish.chinchalpalle21@vit.edu

Abstract: **This research work aims to create a comprehensive database of college alumni by scraping data from LinkedIn profiles. The abundance of information available on LinkedIn makes it an invaluable resource for understanding the career paths, skills, and achievements of former students. Leveraging web scraping techniques, this research work seeks to collect and organize pertinent details such as employment history, educational background, skills, and professional connections. The process involves utilizing Python programming along with web scraping libraries to extract information from public LinkedIn profiles while ensuring compliance with LinkedIn's terms of service and privacy policies. Additionally, data cleaning and structuring techniques will be employed to ensure accuracy and consistency in the collected data. The compiled dataset will serve as a valuable resource for various purposes, including academic research, career services, alumni engagement initiatives, and networking opportunities for current students. By providing insights into alumni trajectories, skill trends, and industry affiliations, this project aims to foster a stronger sense of community and facilitate beneficial connections among alumni and the institution.**

Keywords— Web scrapping; Web crawling; selenium; Python; Alumni profile; LinkedIn profile; Academic research; Alumni analysis;

I. INTRODUCTION

To enhance the visibility and impact of alumni success stories, the Indian engineering college has curated a series of narratives that highlight the remarkable career trajectories, achievements, and contributions of esteemed alumni. These stories, prominently displayed on the college website and in promotional materials, serve as an inspiration for current students and demonstrate the high quality of education and opportunities offered by the institution. By showcasing diverse accomplishments across a range of industries, this project aims to attract prospective students and employers, emphasizing the value of a degree from the college.

Analyzing the industry affiliations and career paths of alumni reveals key trends that guide educational programs and partnerships. Understanding which sectors and company's alumni are most involved in allows colleges to tailor the curriculum to better align with industry needs and take advantage of emerging opportunities. This proactive approach ensures graduates are well-prepared to succeed in their chosen fields upon graduation.

Additionally, a comprehensive skill gap analysis helps institutions identify areas of strength and areas needing improvement among alumni. Armed with this data, colleges have enhanced curriculum design, introduced new courses, and offered professional development programs to bridge skill gaps and better prepare students for the workforce. By staying ahead of evolving industry demands, institutions can ensure graduates possess the relevant skills and competencies to excel in their careers.

Through career pathway mapping, colleges can provide current students with invaluable insights into the various career trajectories taken by the alumni. By detailing job titles, functions, and progression, colleges can empower students to make informed decisions about their career paths and set realistic expectations for their future trajectories. This transparency fosters a sense of confidence and direction among the student body, empowering them to pursue their aspirations with clarity and purpose.

To facilitate ongoing engagement and mentorship opportunities, colleges can leverage alumni networking events and mentorship programs where current students can connect with alumni working in their fields of interest. These initiatives foster professional relationships and mentorship opportunities that benefit current students and alumni in their career development, creating a robust network of support and guidance within the college community.

Moreover, the structured alumni data is instrumental in generating reports and documentation required for accreditation purposes. By demonstrating alumni outcomes and educational impact on accreditation bodies such as NAAC, NIRF, and NBA, institutions can showcase their institutional efficacy and commitment to meeting accreditation standards.

Finally, develop customized alumni engagement strategies based on the preferences and interests of alumni as indicated by their LinkedIn profiles. By tailoring communication and engagement efforts to foster stronger connections and involvement with the college community, colleges can ensure that the alumni remain engaged and invested in the ongoing success of the institution.

In conclusion, this comprehensive approach to alumni engagement and analysis not only celebrates graduates' achievements but also guides strategic initiatives that enrich the educational experience and career prospects of current students. By fostering continuous collaboration and innovation, institutions remain dedicated to empowering the next generation of engineering leaders and making a lasting impact in the field of education and beyond.

This project adopts Selenium, a versatile web scraping tool, to meticulously extract and structure alumni data from LinkedIn. Selenium's proficiency in browser automation and data extraction empowers this project to systematically collect diverse alumni information, including career trajectories, industry affiliations, skill sets, and academic accomplishments.

Technically, Selenium leverages its capability to interact with web elements, simulating user interactions, and navigating LinkedIn profiles dynamically. This allows the extraction of pertinent data points from alumni profiles, ensuring a comprehensive dataset for subsequent analysis and utilization.

The work involves the design of Selenium scripts, employing Python or Java, to navigate LinkedIn profiles, access publicly available information, and extract relevant data points while adhering to ethical scraping practices. These scripts are engineered to efficiently traverse through profiles, gather specific data points, and compile structured datasets for analysis.

The acquired alumni data is intended to serve as a strategic asset for Indian engineering colleges in their accreditation endeavors. It offers a quantitative and qualitative perspective on alumni outcomes, educational impact, and industry engagements, aligning with the stringent evaluation criteria set by accrediting bodies.

While prioritizing data integrity and privacy, this project ensures compliance with LinkedIn's policies and ethical data scraping practices. Upholding user privacy and consent remains a foundational principle throughout the scraping process.

In essence, leveraging Selenium for alumni data acquisition aligns with the accreditation aspirations of Indian engineering colleges. This project seeks to empower institutions with a robust dataset, enabling informed decision-making, and institutional research, and meeting the rigorous standards set by accrediting bodies

II. LITERATURE REVIEW

LinkedIn scraping and clustering techniques for alumni profiling have been extensively explored [1]. Vilas and Rebeca P. Díaz Redondo's study introduced methodologies for characterizing LinkedIn profiles, employing various scraping and clustering techniques, thereby laying the foundation for subsequent research in the field of alumni data acquisition and analysis. Another pivotal study [2] presented a comprehensive review of web scraping techniques, emphasizing their applications in contemporary contexts. This review provided insights into the evolving landscape of web scraping methodologies and their relevance to data acquisition from platforms like LinkedIn. Automated collection methodologies for gathering alumni data from social media platforms, particularly LinkedIn, were showcased in an influential work [3]. This study emphasized the efficiency and accuracy of automated techniques in acquiring valuable alumni information, thereby enabling comprehensive data analysis.

Studies focused on retrieving and classifying alumni job titles [4] made significant contributions to career analysis utilizing LinkedIn data. These studies provided insights into the categorization and analysis of job titles, thereby enriching the understanding of alumni career trajectories.

Ethical and legal concerns surrounding data scraping were discussed in study [5], emphasizing responsible data collection practices and compliance with ethical guidelines while acquiring data from platforms like LinkedIn. investigate job recommendations on LinkedIn based on semantic similarity of users' skills. Their study explores leveraging semantic relationships between skills to enhance personalized job suggestions. By analyzing skill similarities, they aim to provide tailored recommendations, contributing to the advancement of job recommendation systems on professional networking platforms [6].

Advancements in crawling methodologies for social media platforms were introduced in an important research work [7], potentially enhancing the collection of alumni data from diverse geographic regions. This study proposed innovative methods for efficient and extensive data acquisition from online platforms. delves into test automation implementation using Selenium WebDriver. This study emphasizes the practical aspects of test automation and its implementation, showcasing Selenium WebDriver's capabilities and its role in automated testing scenarios. Discussing a technique for data collection through web scraping with Python, [8] provides a concise yet detailed overview of Python-based scraping methodologies, emphasizing the intricacies of data collection techniques using Python libraries.

Provides an analysis of data scraping and its implications for online privacy. The study discusses how data scrapers gain access to publicly available LinkedIn profiles, exploring the legal and ethical issues this raises. By examining industry practices, legal precedents, and the balance between data accessibility and privacy, the article presents a thorough overview of the challenges and future directions in managing data scraping [9].

The authors propose a method for crawling social media platforms using multiple levels of geo-location. This approach leverages spatial data to enhance the efficiency of social media data collection, providing a more precise and context-aware analysis. The paper covers the implementation of the method and its potential applications in social media monitoring and research [10]. The article examines how machine learning algorithms can improve data extraction processes and optimize web scraping tasks. By integrating these technologies, the approach aims to increase the speed and accuracy of data collection from web sources. The authors highlight potential applications and benefits of this method, particularly in handling large-scale web scraping projects and streamlining data retrieval processes [11].

The article covers the design and implementation of efficient web scraping techniques in Python and presents visualization strategies to make collected data accessible and interpretable. Different approaches to web scraping are explored, with a focus on optimizing processes and facilitating data analysis [12].

The features of Selenium Web driver and its effectiveness in automating browser-based testing. The authors highlight the advantages of using Selenium for test automation, such as its compatibility with various programming languages and support for multiple browsers. Additionally, the study outlines the implementation process and potential challenges in using Selenium for software testing [13].

In "Web Scraping: Collecting and Retrieving Data from the Web" by Roman Egger, Markus Kroner, and Andreas Stöckl, presented in Applied Data Science in Tourism: Interdisciplinary Approaches, Methodologies, and Applications, the authors explore web scraping techniques for data collection. The chapter delves into the process of extracting data from websites using various tools and methods. It discusses the importance of web scraping in data-driven decision-making in the tourism industry and provides practical examples of scraping data from tourism-related websites. The authors also address ethical considerations and legal implications associated with web scraping, emphasizing the need for responsible data usage and compliance with relevant regulations [14].

The authors identify deep web entity pages as dynamic, data-rich web pages generated in response to user queries, often residing behind search interfaces. They highlight the significance of these pages for search engine indexing and knowledge acquisition. The study discusses the difficulties inherent in crawling deep web entity pages, such as the need to interact with search interfaces and handle dynamically generated content. To address these challenges, the authors propose a novel method that combines keyword-based exploration with page similarity analysis. They introduce a clustering algorithm to group similar pages and optimize the crawling process, effectively navigating the deep web while minimizing duplicate retrievals. The study's findings provide insights into improving the efficiency and effectiveness of web crawling strategies for accessing valuable content within the deep web domain. importance for search engine indexing and knowledge retrieval. They propose a method combining keyword-based exploration with page similarity analysis and clustering to efficiently navigate dynamic content. Their approach aims to optimize crawling processes, providing insights into improving deep web access for valuable data retrieval [15].

The potential benefits of using automated crawling techniques to analyze web applications, such as identifying vulnerabilities and assessing performance. They also highlight challenges including the dynamic nature of modern web applications, which can complicate crawling and analysis efforts. Overall, the study provides a brief overview of the opportunities and obstacles in utilizing crawl-based approaches for web application analysis [16]. A range of studies have explored the use of Python for web scraping, highlighting its effectiveness in various fields. Lawson (2015) provides a comprehensive overview of web scraping techniques, and further demonstrates the application of Python in finance, using web scraping to extract financial data from Yahoo Finance. These studies collectively underscore the versatility and practicality of Python in web scraping, particularly in the context of data extraction and analysis [18]. In this project, they study an undergraduate program that prepares its students for the major needs of the workforce for the market. The main way to identify what the demands are to be met is by creating a method to manage information of its alumni. Overall, by using a few alumni pages as an initial set of sample data, the proposed method was capable of gathering information concerning several alumni twice as big as the adopted conventional method [18]. web scraping with Python provides a detailed overview of the topic, offering insights into techniques, tools, and best practices for extracting data from websites using Python. The publication delves into various aspects of web scraping, including HTML parsing, the utilization of libraries like Beautiful Soup and Scrapy, handling dynamic content, and navigating complex website structures. Moreover, Lawson addresses ethical and legal considerations pertinent to web scraping activities. This work serves as a valuable resource for individuals seeking to leverage Python for data extraction purposes on the web. The exploration of web archive search as a research tool reveals both methodological challenges and theoretical implications. Researchers delve into understanding the inherent biases and limitations present within archived web content, emphasizing the importance of critically evaluating data retrieved from web archives. Additionally, the theoretical underpinnings of web archive research are scrutinized, with recognition of its potential contributions to fields such as digital humanities and information science. This exploration sheds light on the complexities involved in utilizing web archives for scholarly inquiry, prompting a deeper examination of methodological considerations and theoretical frameworks relevant to this domain [19].

III. METHODOLOGY

The project begins with the initial phase of meticulously organizing alumni names in a CSV file. To manage this data systematically, Python is utilized, along with libraries such as pandas for efficient data handling. The system reads through alumni names from the CSV file, ensuring a structured approach to data processing.

As the project progresses, intelligent search queries are formulated to locate LinkedIn profiles using strategic keywords such as "alumni," the specific college name, and "LinkedIn." Web scraping techniques and engagement with search engines, typically Google, automate the process. Selenium, a robust web automation library, and Chrome WebDriver enable seamless browser automation. The search results page is scraped meticulously to extract multiple LinkedIn URLs, creating a comprehensive dataset primed for subsequent analysis.

Maintaining data accuracy is a top priority, prompting the implementation of a stringent validation process for each extracted LinkedIn URL. The system enforces stringent criteria, requiring the presence of both the name and surname in the URL. Automated validation algorithms promptly discard URLs that do not meet these criteria, ensuring the collected data remains relevant and accurately linked to respective alumni.

Upholding data accuracy remains paramount, prompting the implementation of a stringent validation process for each extracted LinkedIn URL. The system sets stringent criteria mandating the presence of both the name and surname in the URL. Automated validation algorithms swiftly discard URLs that fail to meet these criteria, ensuring that the collected data remains relevant and accurately linked to respective alumni.

For the extraction of detailed profile information, Selenium once again takes center stage for web scraping endeavors. Its prowess in automating browser interactions proves indispensable, particularly for navigating dynamic web pages like LinkedIn profiles. The system adeptly navigates to each validated LinkedIn URL, expertly mimicking user interactions to gather up-to-date and precise information. Crucial details including name, current workplace or college, position, current location, and the LinkedIn URL are systematically extracted with precision and efficiency.

To ensure the organized storage of the extracted details, the data is structured using a data frame. Using the panda's library in Python, a CSV file is crafted to systematically store the data. Each row in the CSV file corresponds to the details of an alumni profile, including vital information such as name, current workplace or college, position, current location, and LinkedIn URL. This structured storage framework facilitates easy access and seamless analysis for future use, streamlining efforts in alumni engagement and career development initiatives.

Figure 1. Flowchart of Profile Data Extraction system

As illustrated in Figure 1, the flowchart outlines a step-by-step method for extracting alumni data from LinkedIn profiles. The process begins by taking in a CSV file containing the names of LinkedIn users. Then, through web scraping and validation, the system retrieves the profile URLs of each person. This involves searching for each person's name on Google, along with specific keywords like "LinkedIn" and the name of the college. The system then extracts and validates the correct URL to ensure it leads to a genuine LinkedIn profile. This meticulous approach ensures that only accurate and valid profile URLs are collected for further processing. The URL is validated by checking if it contains the name of the user within it. This ensures that the URL corresponds to the correct LinkedIn profile associated with the respective individual.

The heart of the process involves Selenium, a web automation library, working alongside Chrome Driver to control a headless Chrome browser. This browser then navigates to the validated profile URL and extracts desired data points like name, title, company, and potential skills or experience. Finally, the extracted data is meticulously structured and saved into a new CSV file for further analysis.

This process can be further enhanced with error handling, data validation, and pagination handling to ensure robustness. It's important to remember to respect LinkedIn's terms of service by implementing rate-limiting and ethical scraping practices.

IV. RESULT

Figure 2. CSV file with profile data

During the data collection phase, alumni information was compiled into a well-organized CSV file (as shown in Fig. 2). Each entry within the file corresponds to a single alumnus and details include name, current job title, current working company, Location and hyperlink to their LinkedIn profile.

The proposed system does a great job of organizing alumni information neatly into a CSV file, making it easy to manage. Each row in the file shows details about a graduate, like alumni name, job, and LinkedIn profile link. Using fancy tools like Selenium and Chrome WebDriver, the data is collected automatically, which means fewer mistakes and more reliable information. This careful approach not only saves time but also makes sure the data is trustworthy.

On the other hand, in the previously proposed work data is not organized way, but in the proposed system project puts more emphasis on making the information easy to use, like putting it in a structured CSV file. In the previous work, the authors extracted data from various social media handles, whereas in our case, we specifically focused on LinkedIn, a more professional platform. This choice ensures that the information gathered is directly relevant to the professional lives of the alumni, providing more comprehensive and targeted information.

To sum up, the system stands out because it's well-organized, and automated, and focuses on giving useful insights into alumni networks. These qualities make it a top choice for managing and analyzing alumni data effectively.

V. LIMITATIONS

Limitation of the system arises due to LinkedIn's privacy policy, which includes measures to detect automated activity. After retrieving a certain number of user information, LinkedIn may implement a CAPTCHA to solve. Additionally, excessive usage of automated scraping techniques may trigger warnings from LinkedIn. These limitations could potentially slow down the data collection process and require human intervention to solve CAPTCHAs or address warnings from LinkedIn.

VI. CONCLUSION

This research work successfully demonstrated the power of Python libraries and Selenium for automated alumni data retrieval. The well-structured methodology ensured data handling, precise search queries, and accurate validation, resulting in a comprehensive Excel file. Automation boosted efficiency and maintained data accuracy through Selenium's web scraping capabilities. This project's success lies in navigating the complexities of web scraping, data validation, and storage. This organized dataset offers a wealth of potential. the institution can analyze career paths to tailor student programs, identify in-demand skills to inform curriculum development and reconnect with alumni for targeted engagement or mentorship opportunities. Network mapping could even reveal connections between alumni, fostering collaboration and strengthening the alumni community.

Future initiatives could focus on incorporating more sophisticated error handling and exploring data enrichment techniques. This project lays the foundation for a robust and scalable approach to gathering valuable alumni data. By embracing automation, ethical practices, and continuous improvement, the potential of the alumni network can be fully realized, fostering a thriving alumni community. The possibilities are extensive, and the future of alumni data analysis holds great promise.

REFERENCES

[1] Li, Lei, Svetlana Peltsverger, Jack Zheng, Linh Le, and Michael Handlin. "Retrieving and Classifying LinkedIn Job Titles for Alumni Career Analysis." In Proceedings of the 22nd Annual Conference on Information Technology Education, pp. 85-90. 2021.

[2] Ariyanto, R., C. Rachmad, and A. R. Syulistyo. "Automatically collect alumni data on social media." IOP Conference Series: Materials Science and Engineering. Vol. 732. No. 1. IOP Publishing, 2020.

[3] Bendremer, Emily. "Mapping WPI Alumni Career Trajectories Through Data Visualization." PhD diss., WORCESTER POLYTECHNIC INSTITUTE.

[4] Barudwale, Nihalahmed, et al. "ALUMNI CONNECT FORUM." International Research Journal of Modernization in Engineering Technology and Science. Volume:05/Issue:05/May-2023.

[5] Dai, Kais, Celia Gonzalez Nespereira, Ana Fernández Vilas, and Rebeca P. Díaz Redondo. "Scraping and clustering techniques for the characterization of LinkedIn profiles." arXiv preprint arXiv:1505.00989 (2015).

[6] Domeniconi, Giacomo, Gianluca Moro, Andrea Pagliarani, Karin Pasini, and Roberto Pasolini. "Job recommendation from the semantic similarity of LinkedIn users' skills." In International Conference on Pattern Recognition Applications and Methods, vol. 2, pp. 270-277. SciTePress, 2016.

[7] Ghatole, S. M., and P. B. Dahikar. "NAAC Accreditation: A Quality Initiative Reform in Indian Higher Education." The Research Journal (TRJ): A Unit of I2OR 7.1 (2021): 42-47.

[8] Kumar, Sumit, and Uponika Barman Roy. "A technique of data collection: web scraping with Python." Statistical Modeling in Machine Learning. Academic Press, 2023. pp 23-36.

[9] Boulanger, Marissa. "Scraping the bottom of the barrel: Why it is no surprise that data scrapers can have access to public profiles on LinkedIn." SMU Sci. & Tech. L. Rev. 21 (2018): pp 77.

[10] AlZu'bi, Shadi, Darah Aqel, Alaa Mughaid, and Yaser Jararweh. "A multi-level geo-location based crawling method for social media platforms." In 2019 Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS), pp. 494-498. IEEE, 2019.

[11] Mehta, Sunny, and Gayatri Pandi. "An Improving Approach for Fast Web Scrapping Using Machine Learning and Selenium Automation." International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) 8.10 (2019): pp 434-438.

[12] Yuan, Shujun. "Design and Visualization of Python Web Scraping Based on Third-Party Libraries and Selenium Tools." Academic Journal of Computing & Information Science 6.9: pp25-31.

[13] Gudavalli, Aneesha, and G. JayaLakshmi. "Implementation of Test Automation with Selenium Webdriver." Grenze International Journal of Engineering & Technology (GIJET) 8.1 (2022).

[14] Egger, Roman, Markus Kroner, and Andreas Stöckl. "Web Scraping: Collecting and Retrieving Data from the Web." Applied Data Science in Tourism: Interdisciplinary Approaches, Methodologies, and Applications. Cham: Springer International Publishing, 2022. 67-82.

[15] He, Yeye, Dong Xin, Venkatesh Ganti, Sriram Rajaraman, and Nirav Shah. "Crawling deep web entity pages." In Proceedings of the sixth ACM international conference on Web search and data mining, pp. 355-364. 2013.

[16] Van Deursen, Arie, Ali Mesbah, and Alex Nederlof. "Crawl-based analysis of web applications: Prospects and challenges." Science of computer programming 97 (2015): 173-180.

[17] Lawson, Richard. Web scraping with Python. Packt Publishing Ltd, 2015.

[18] Dang, Lingchen. "ALUMNI INFORMATION COLLECTION AND MANAGEMENT." (2023).

[19] Ben-David, Anat, and Hugo Huurdeman. "Web archive search as research: Methodological and theoretical implications." Alexandria 25, no. 1-2 (2014): 93-111.

Machine Learning Techniques for Descriptive Answer Evaluation: A Comprehensive Survey

Dr. Kavita Moholkar
Computer Engineering Department
JSPM's Rajarshi Shahu College of Engineering, Pune, India
kpmohalkar_comp@jspmrscoe.edu.in

Manohar Chaturvedi
Computer Engineering Department
JSPM's Rajarshi Shahu College of Engineering, Pune, India
mmrchmanoharofficial@gmail.com

Aryan Jain
Computer Engineering Department
JSPM's Rajarshi Shahu College of Engineering, Pune, India
jainaryan158@gmail.com

Ajinkya Parkhe
Computer Engineering Department
JSPM's Rajarshi Shahu College of Engineering, Pune, India
ajinkyaparkhe48@gmail.com

Karan Singh
Computer Engineering Department
JSPM's Rajarshi Shahu College of Engineering, Pune, India
karansingh0806@icloud.com

ABSTRACT:

This research study intends to assist the examination board in evaluating the responses based on the level of question difficulty determined by the board. As an initial step, this study reviews numerous approaches and techniques for evaluating descriptive answers using machine learning. In order to evaluate descriptive responses automatically, this research study proposes a novel approach that makes use of various machine learning, natural language processing, and tools like wordnet, word mover's distance (WMD), cosine similarity, latent semantic analysis (LSA), word2vec, LSTM bi-directional encoder representation from transformer. this analysis has examined 20 articles that were published in the recent years, from 2012 to 2022. based on this research and investigation, this study concludes LSTM as the optimum method for descriptive answer evaluation using machine learning techniques. with the aim to increase the precision and the efficiency of descriptive answer evaluation, research works have been actively conducted to develop these methods.

Keywords- Natural Language Processing (NLP), Word Mover's Distance (WMD), Word2vec, Latent Semantic Analysis (LSA), Term Frequency-Inverse Document Frequency (TF-IDF), Long-Short Term Memory (LSTM).

I. INTRODUCTION

Automating the evaluation of descriptive answers using machine learning is a promising solution to address the challenges associated with manual grading. With over 15 lakh students taking descriptive exams annually, the need for accurate and efficient evaluation methods is paramount. Traditional manual grading methods can be tedious, time-consuming, and prone to human error, leading to inconsistent marking. By leveraging machine learning algorithms, we can develop automated systems that analyze the quality of written responses with high accuracy and consistency. These systems use various techniques, including feature-based models, neural networks, ensemble methods, and transfer learning, to assess descriptive answers objectively.

One of the key advantages of using machine learning for descriptive answer evaluation is its ability to consider multiple input variables, such as word count, sentence structure, and the presence of specific keywords, to generate accurate assessments. These algorithms learn from human-assigned ratings to predict scores for new answers, ensuring fairness and consistency in evaluation. This research study focuses on Natural Language Processing (NLP) methods, such as tokenization, text representation techniques like TF-IDF, LSTM, word2vec, and Bag of Words, as well as similarity measurement methods like cosine similarity and word mover's distance. These techniques enable us to extract meaningful insights from textual data and make informed decisions about the quality of descriptive answers.

Over the years, researchers have made significant progress in developing machine learning-based systems for automatic descriptive answer evaluation. However, there is still room for improvement, particularly in fine-tuning algorithms to handle nuances in language and context effectively. Overall, automated descriptive answer evaluation holds immense potential to streamline the grading process, reduce the burden on examiners, and ensure fair and consistent assessments for students. As technology continues to advance, we can expect further advancements in this field, leading to even more accurate and efficient evaluation methods.

A few of the datasets commonly utilized for descriptive answers evaluation technique incorporate the Semi-Eval dataset, which contains answers to questions on different themes, and the Yahoo Answers dataset, which contains a huge number of questions and answers on a wide run of subjects. Generally, machine learning-based approaches have appeared promising comes about within the assessment of descriptive answers. In any case, there's still room for change, especially in taking care of complex and nuanced reactions.

Descriptive answer evaluation is a crucial task in natural language processing (NLP) that aims to assess the quality and relevance of generated answers. With the advent of machine learning, researchers and practitioners have leveraged various techniques to automate this evaluation process. Machine learning models, such as LSTM-based architectures, have played a significant role in achieving accurate and reliable descriptive answer evaluation. In this article, we will explore the use of machine learning algorithms for descriptive answer evaluation and showcase illustrative examples that demonstrate their effectiveness.

Example 1: Consider a question answering system that utilizes machine learning for descriptive answer evaluation. Given a question like "Who invented the light bulb?", the system generates a descriptive answer such as "Thomas Edison invented the light bulb." Irecurrn this case, the machine learning model, specifically an LSTM-based architecture, is responsible for generating the answer by analyzing the question and relevant contextual information. The model learns from a large dataset of questions and corresponding answers, enabling it to generate accurate responses based on learned patterns and contextual understanding.

Example 2: In educational settings, machine learning algorithms can be employed for automated essay grading and descriptive answer evaluation. For instance, a student might write an essay response to a prompt such as "Discuss the causes and effects of climate change." To evaluate the quality of the response, an LSTM-based model can analyze the essay's content, coherence, and relevance to the prompt. The model can assign a score or provide feedback based on the learned patterns from a dataset of human-graded essays. This automated evaluation system saves time for educators and provides consistent and objective assessments.

Example 3: Online forums and customer support platforms often utilize machine learning for descriptive answer evaluation. For instance, in a customer support chatbot, when a user asks a question like "How can I reset my password?", the chatbot generates a descriptive answer by understanding the user's query and providing appropriate instructions. LSTM-based models can be trained on a dataset of user queries and corresponding responses to learn the patterns and generate accurate answers. The evaluation component of the system ensures that the generated answers are relevant and helpful to the users.

These examples illustrate how machine learning algorithms, particularly LSTM-based architectures, can effectively evaluate and generate descriptive answers in various domains. By leveraging large datasets, these models can capture semantic relationships, contextual understanding, and generate accurate responses based on learned patterns. The use of machine learning in descriptive answer evaluation offers the potential for automation, improved efficiency, and enhanced user experiences in question answering, educational assessment, customer support, and beyond.

1.1 Contribution

This study makes a contribution by applying machine learning and natural processing techniques to solve the issue of descriptive response assessment. It explores several criteria for language similarity measurement matrices and offers a technique for creating a machine learning model that might enable going forward with more confidence in evaluation outcomes.

2. Literature Review

2.1. Methodology:

Figure 1 depicts the methodology we used to design and carry out the survey. First, we discovered the answers to our questions, such as why this survey is necessary, how beneficial it is, and what the best course of action should be.

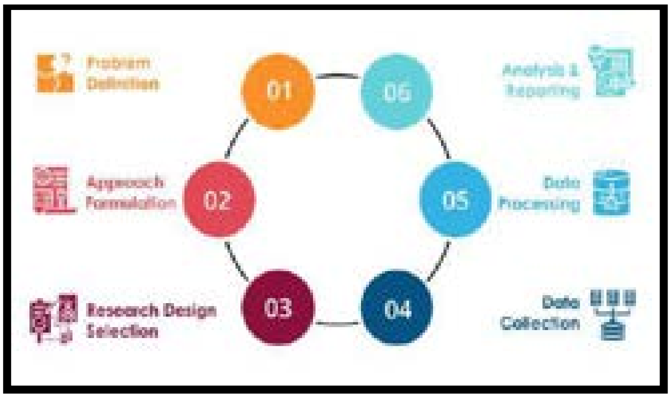


Fig. 1. Methodology Used for Survey

We downloaded research and survey articles on pertinent subjects covering the years 2018 through 2022.

When evaluating machine learning techniques for descriptive answer evaluation, there are several key components that should be considered. Here is a suggested methodology for evaluating these techniques:

1. **Define the Evaluation Metrics:** Start by defining the evaluation metrics that will be used to measure the performance of the machine learning techniques. Common metrics for descriptive answer evaluation include accuracy, precision, recall, F1-score, and mean squared error. Choose the metrics that are most relevant to your specific task and objectives.
2. **Prepare a Ground Truth Dataset:** Create a high-quality ground truth dataset by manually annotating a representative sample of descriptive answers. This dataset should cover a diverse range of answers and be well-balanced to reflect the distribution of different answer qualities. The annotators should have expertise in the domain and clear guidelines on how to evaluate the answers.
3. **Split the Dataset:** Divide the ground truth dataset into training, validation, and test sets. The training set is used to train the machine learning models, the validation set helps in tuning hyperparameters, and the test set is used to evaluate the final performance of the models.
4. **Feature Extraction:** Extract relevant features from the descriptive answers. Depending on the nature of the

task, these features could include word frequencies, sentence structure, sentiment analysis scores, readability measures, or domain-specific features. Feature engineering is crucial as it helps the machine learning models to learn patterns and make predictions.

5. **Model Selection:** Select a set of machine learning techniques that are suitable for descriptive answer evaluation. This could include traditional algorithms like decision trees, random forests, support vector machines, or more advanced techniques like deep learning models such as recurrent neural networks (RNNs) or transformers. Consider the strengths and weaknesses of each technique and choose the most appropriate ones for your task.
6. **Training and Validation:** Train the selected machine learning models using the training set and fine-tune their hyperparameters using the validation set. Perform cross-validation or use other techniques to ensure the models generalize well and are not overfitting the training data.
7. **Evaluate Performance:** Evaluate the performance of the trained models using the test set. Calculate the evaluation metrics defined in Step 1 and analyze the results. Compare the performance of different models and techniques to determine which ones perform the best.
8. **Error Analysis:** Conduct an in-depth error analysis to understand the types of mistakes made by the models. Identify the strengths and weaknesses of each technique and identify areas for improvement. This analysis can help guide future research and model refinement.

By following this methodology, you can systematically evaluate and compare different machine learning techniques for descriptive answer evaluation [15-16]. It provides a structured approach to assess their performance and identify areas for improvement.

2.2. Datasets for Evaluation:

When evaluating machine learning techniques for descriptive answer evaluation, it's important to have a diverse and representative dataset that covers a range of answer qualities and addresses the specific task at hand. Here are some examples of different datasets that can be used for descriptive answer evaluation:

1. **SQuAD (Stanford Question Answering Dataset):** It is a widely used dataset for machine reading comprehension. It consists of questions posed by crowd workers on a set of Wikipedia articles, where the answers are also highlighted within the articles. Each question has a single descriptive answer, making it suitable for evaluating descriptive answer generation models.
2. **MS MARCO (Microsoft Machine Reading Comprehension):** The MS MARCO dataset is another popular dataset for evaluating machine reading comprehension models. It contains a large collection of real anonymized user queries along with corresponding passages from the web. The dataset includes both human-generated queries and human-generated answers.
3. **CoQA (Conversational Question Answering):** It is a dataset designed for evaluating models in a conversational question answering setting. It consists of dialogues between two crowd workers: a student who asks questions, and a teacher who provides answers. The dataset focuses on answering questions that require understanding a passage and reasoning over multiple turns of conversation.
4. **NarrativeQA:** It is a dataset that focuses on evaluating question answering models on a diverse range of narrative texts. It includes a collection of book summaries from Project Gutenberg, along with a set of questions that can only be answered by reading the associated text. The dataset is suitable for evaluating descriptive answer generation models on longer passages.
5. **TriviaQA:** It is a dataset that contains trivia questions along with detailed, evidence-supported answers. The dataset includes both short-answer and open-ended questions, making it useful for evaluating descriptive answer generation models. TriviaQA provides web URLs as evidence, allowing models to learn to retrieve and summarize information from external sources.

Discovering relationships between words in a large dataset can be approached through various techniques, such as statistical analysis, word embeddings, co-occurrence matrices, and graph-based methods. Here are a few common approaches:

1. **Word Embeddings:** Word embeddings are dense vector representations of words that capture semantic relationships. Models like Word2Vec, GloVe, or FastText can be used to generate word embeddings from a large dataset. Once the embeddings are obtained, you can measure the similarity between word vectors using cosine similarity or other distance metrics. Similar words will have closer vector representations, indicating a potential relationship.
2. **Co-occurrence Analysis:** Co-occurrence matrices provide insights into the relationships between words based on their frequency of co-occurrence within a context window. By scanning the dataset and counting how often words appear together, you can construct a co-occurrence matrix. Techniques like pointwise mutual information (PMI) or term frequency-inverse document frequency (TF-IDF) can be applied to measure the strength of the relationship between words.
3. **Statistical Analysis:** Statistical techniques such as correlation analysis or association rule mining can help identify relationships between words. For example, computing correlation coefficients like Pearson's correlation or using techniques like Apriori algorithm can reveal associations or co-occurrences between words based on their statistical patterns in the dataset.

For pre-processing of the dataset NLP techniques are used. Some of them are explained as follows:

Text Pre-processing: Before applying any NLP techniques, text data is typically preprocessed to remove noise and irrelevant information. This may involve tasks such as tokenization (splitting text into individual words or tokens), stemming or lemmatization (reducing words to their base or root form), removing stopwords (common words like "the," "and," etc.), and handling punctuation and capitalization.

Latent Sentiment Analysis: It also known as opinion mining, involves determining the sentiment or opinion expressed in a piece of text. It helps in understanding whether the sentiment is positive, negative, or neutral. Sentiment analysis is widely used in applications like social media monitoring, customer reviews, and market research. In LSA, a mathematical model is built using a matrix called the term-document matrix. This matrix represents the frequency or occurrence of words in a collection of documents. Each row represents a word, each column represents a document, and the values in the matrix indicate the frequency of each word in each document.

Deep Learning for NLP: Deep learning, particularly deep neural networks such as recurrent neural networks (RNNs) and transformers, have significantly advanced the state-of-the-art in NLP. These models can learn complex patterns and dependencies in text data, improving performance in tasks like machine translation, sentiment analysis, and natural language understanding.

2.3. Criteria for Performance Metric:

Model selection and hyperparameter tuning are critical steps in achieving high accuracy in machine learning. Here's a more detailed explanation of these processes:

- **Model Selection:**

Model selection involves choosing the appropriate machine learning algorithm or model architecture for your descriptive answer evaluation task. The choice of model depends on factors such as the nature of the data, the complexity of the problem, the availability of labelled data, and computational resources. Here are some commonly used models for descriptive answer evaluation:

1. **Linear Regression:** A basic yet effective model that fits a linear equation to the input features and predicts a continuous output.
2. **Support Vector Regression (SVR):** A regression model that uses support vector machines to find a hyperplane that best fits the data.
3. **Random Forest:** An ensemble model that combines multiple decision trees to make predictions.
4. **Recurrent Neural Networks (RNNs):** Deep learning models that are suitable for sequence data, such as textual answers. They capture temporal dependencies and have been successful in natural language processing tasks.
5. **Transformers:** Attention-based models that have shown impressive performance in various natural language processing tasks, including text generation and comprehension.

The choice of model depends on the specific requirements of your descriptive answer evaluation task. It's often a good practice to experiment with multiple models and compare their performance to find the best one.

- **Hyperparameter Tuning:**

Hyperparameters are parameters that are set before the training process begins, and they control the behaviour and performance of the machine learning model. Tuning these hyperparameters is crucial to optimize the model's accuracy. Here are some common hyperparameters that may need tuning:

1. **Learning Rate:** Determines the step size at each iteration of the optimization process. A higher learning rate may converge faster, but it can also lead to overshooting the optimal solution. A lower learning rate may take longer to converge but can provide better accuracy.
2. **Regularization Strength:** Controls the degree of regularization applied to the model. Regularization helps prevent overfitting by adding a penalty term to the loss function. Tuning this hyperparameter determines the balance between model complexity and generalization.
3. **Network Architecture:** For deep learning models like RNNs or transformers, hyperparameters such as the number of layers, hidden units, or attention heads can significantly impact performance. Experimenting with different architectures is essential for achieving the best results.
4. **Dropout Rate:** Dropout is a regularization technique that randomly sets a fraction of input units to zero during training. Tuning the dropout rate helps prevent overfitting and improve model generalization.
5. **Batch Size:** Determines the number of samples used in each iteration of model training. It affects the convergence speed and memory requirements of the training process. Choosing an optimal batch size can impact model performance.

Hyperparameter tuning can be performed through manual experimentation or automated approaches like grid search, random search, or Bayesian optimization. These techniques involve systematically exploring different combinations of hyperparameters and evaluating their impact on the model's accuracy.

It's important to note that hyperparameter tuning should be performed on a separate validation set or using techniques like cross-validation to ensure unbiased evaluation. Additionally, it's crucial to avoid overfitting the hyperparameters to the validation set by using a separate test set for the final evaluation of the model's accuracy.

Overall, model selection and hyperparameter tuning are iterative processes that require experimentation, evaluation, and comparison of different models and hyperparameter configurations. The goal is to find the optimal combination that maximizes the accuracy of the descriptive answer evaluation system.

2.4. Discussion of Survey:

2.4.1. Selection Criteria for Survey Papers:

We collected a total of 30 papers and articles, and shortlisted 20 papers after screening it through and considering only those papers that are relevant to our subject. We followed the below measures to configure and filter the selected 40 papers:

1. We chose the papers that were most pertinent to our subject and gave explanations of the expected approaches and algorithms.

2. Simple to comprehend papers were given higher priority.

3. We looked at the articles that provided reliable performance measures.

2.4.2. Objectives (OBJ) of the Review:

Objective-1: To discuss the various methods and approaches employed in the publications.

Objective-2: To illustrate the degree of accuracy provided for each method presented in the paper.

Literature on different methodologies, performance and dataset is shown in below table 1.

TABLE 1: LITERATURE ON DIFFERENT METHODOLOGIES, PERFORMANCE AND DATASET

Year	Method used	Description	Dataset used	Performance Metric
2021	[1] cosine similarity, WMD, (MNB)	Two modules: Module 1: Using [1] Cosine Similarity along with MNB Module 2: Using WMD along with MNB. WMD is more efficient according to results of this paper	Own dataset	Accuracy: 88%
2018	Bi-GRU Siamese architecture, Word embedding	The maximum pooling layer of the Bidirectional Gated Recurrent Unit network is further on added with for representation of text where words are similar in length or meaning of the document. This neural network capable of memorising content and producing an essay summary.	Amazon online research service Collected summaries	Accuracy: 55.2%
2019	CNN, LSTM, BERT semantic data, sentences mapping.	This model consists of 2 modules: 1. **Module 1**: The main job of this module for allocating marks on the basis BERT semantic data from the essay. 2. **Module 2**: The score is predictable on some common general structures like syntax, length of the essay and the no. of sentences.	ASAP Kaggle	Accuracy: 70.9%
2020	Multiple Linear Regression, Style and content-based features	In Least Square Assumptions regression technique, they have used two modules or features: 1. **syntax features**: In this they found the tokenization and then constructed entity of character string. 2. **semantic features**: They are more likely identical to the similarity analysis.	ASAP Kaggle	Accuracy: 77%
2020	BERT, Student Answer & Reference Answer mapping	The Bi-directional Encoder Representation from Transformer algorithm is used for automatic one sentence answer marking. The data expansion is finalized with the help of AI network or the neural network and with one correct answer.	SemEval-2013	Accuracy: 82.77%
2019	XGBoost machine learning classifier Word count, type token rationing ,parse tree, coherence, cohesion.	The extreme gradient boosting (XGBoost) is a ML classifier which is used to evaluate the essays and even there are multiple algorithm competent on structures like word count, parse tree, and consistency in the document with the sentence similarity %. And they applied K-fold cross-validation technique for calculating the outcome of the system.	ASAP Kaggle	Accuracy: 68%
2016	Latent Semantic Analysis (LSA), Generalized LSA, Bi-lingual Evaluation Understudy and Maximum Entropy, Concept of Ontology.	It works with the mentioned techniques. The BLEU score helps in determining the skill and accuracy of the student in writing his/hers answers. The concept of Ontology works with regard to the number of criteria present for the answer evaluation. The LSA Model would come handy when evaluating short answers along with the provided model answer set.	Technical answers of Computer Science	Accuracy: 93%

2021	Machine Learning, Artificial Neutral Networks Algorithm.	Artificial Neutral Networks Algorithm works effectively when comparing the written answers with the model answers provided. It will come up with appropriate procedures with which said marks would be allocated.	Own Dataset	Accuracy: 85-95%
2021	Natural Language Processing, Deep Learning, Sequential Ensemble Approach.	The main area of this approach is the basic NLP and Deep Learning which are implemented. The performance of the system at the same time would be improved with the help of Sequential Ensemble Approach that deals with base learners.	Own Dataset	50% increase as compared to individual systems.
2021	[10]Sentence splitting, [10]Jaccard similarity, [10]Bi-directional Encoder Representations from Transformers (BERT), [10]Optical Character Recognition (OCR)	[10]OCR is used to convert scanned answers to digital answers. Sentence splitting, Jaccard similarity, Bidirectional Encoder Representations from Transformers (BERT) are applied on model answer and submitted answer by students and then weighted score is calculated.	Own dataset	Accuracy: 93.2%
2021	[11]Cosine similarity, Text gears grammar API	[11]User answer extractor unit organizes the answer of students. Answer verifier unit uses cosine similarity and then Text gears grammar API. Result set unit consists of 3 attributes Keywords, Grammar, and QST	Own dataset	Accuracy: 80-90%
2018	NLP, ML	This consists of 2 modules: 1)Answer extraction from scanned copy and 2) Applying ML and NLP on it. Algorithm provides semantic meaning of the context.	Own dataset	Accuracy: 90%
2021	Neural Networks, LSTM	This proposed system is based on different models such as- Regression based models, Classification based models, Neural network models, Ontology based approach and scoring on the basis of speech.	ASAP Kaggle	Accuracy: 90%
2021	Natural Language Tool kit (NLTK)	Our proposed system has following features: - Balance Load, Ease of use, Friendly Environment, Ease in accessibility, Efficiency and reliability of the system, Maintenance of the system.	Own Dataset	Accuracy: 95%
2021	Contextual Similarities, Semantic Analysis, NLP, WordNet, TF-IDF, Cosine-Similarity, Antonyms, Synonyms, Grammatical Correction, Machine Learning.	This system works on the basis of few basic factors such as Stemming, Lemmatization, Stop words, Answer length, Keyword matching, Contextual similarity, Semantic similarity, Contradiction, Grammar check, Cosine Similarity, Antonyms & Synonyms.	Own Dataset	Accuracy: 89%

2.5. Major Machine Learning Techniques Used:

- **Long-Short Term Memory (LSTM):**

It is nothing but a unit which can remember or recall the previous data that has been used in the network, and do tries to forget the all-unrelated data. It can be implemented by adding different activation function layers known as gates for various types of functions to be performed easily. Working of LSTM is shown in figure 2. There are in total 3 different types of gates in LSTM:

Forget Gate: It is used to regulate what amount of data to be forgotten from the previous data. The equation for forget gate is given below.

Input Gate: It is used to control which and how much amount of data should be added on the Internal Cell State. The mathematical formula for input gate is as follows.

Output Gate: This gate determines what output should be created by the present state. The mathematical formula for

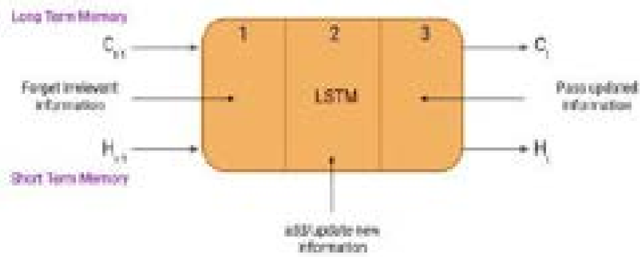


Fig. 2. Working of LSTM

- **Word2vec:**

Word2vec is a technique that employs an AI neural network model to discover relations between words from a huge dataset. Since it can be taught for dimensions that go up to 300, the semantic meaning of the words is mostly unchanged. Following training, a word-to-vector model can identify terms that are comparable or suggest some more sentences based on the phrase. A pre-trained word2vec model is demonstrated by the 100-dimension word2vec model, which has a vocabulary of approximately a trillion words.
It is now time to compare the vectors to see if they are comparable once the text has been converted into numerical form, or vectors. Jacquard Similarity, Cosine Similarity, and Word Mover's Distance (WMD) are some of the most used methods for this job. Illustration of Word2Vec is shown in figure 3.

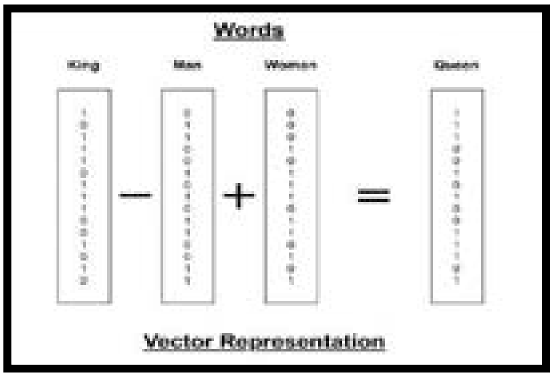


Fig. 3. Illustration of Word2Vec

- **TF-IDF (Term Frequency-Inverse Document Frequency)**

For evaluation of descriptive answer using machine learning, the aim is to survey the quality of constructed replies to a given trigger. The Tf-IDF (Term Frequency-Inverse Document Frequency) calculation can be utilized to assist distinguish important words or phrases within the reaction that will demonstrate its quality.

TF-IDF could be a numerical measurement that reflects the significance of a word to an archive in a collection of archives. It takes into consideration two variables:

1. Term Frequency (TF)
2. Inverse Document Frequency (IDF).

The term frequency measures how regularly a specific word shows up in a given archive, whereas the converse report frequency measures how uncommon or common that word is within the whole collection of archives. Words that show up as often as possible in a report but are uncommon within the by and large collection will have a tall TF-IDF score, showing that they are critical and significant to the archive.
In this, the TF-IDF calculation can be utilized to distinguish words or expressions that are especially vital or important to a high-quality reaction. For case, in the event that the incite is inquiring for a clarification of a logical concept, a high-quality reaction might utilize specialized terms and language that are uncommon within the general collection of reactions. The TF-IDF calculation can distinguish these terms and provide them with a tall score, showing that they are vital and characteristic of a high-quality reaction.
Overall, the TF-IDF calculation may be valuable for distinguishing imperative words and expressions in composed reactions and can be used as included in a machine-learning demonstration to assist assess the quality of those reactions.

- **BERT (Bi-directional Encoder Representations from Transformers)**

It is a previously trained natural language processing model created by Google that has been used for various NLP tasks including sentiment analysis, question-answering, and text classification.
When it comes to descriptive answer evaluation, BERT can be used to evaluate the quality of written answers to open-ended questions. This involves training a machine learning model using BERT as the feature extractor to predict a score or grade for a given answer.
To do this, BERT can be fine-tuned on a dataset of graded answers, such as essays or short responses, to learn the relationship between the input text and the associated score. Once the model is trained, it can be used to evaluate new answers by predicting their score based on the learned patterns.
One example of this application is the Automated Student Assessment Prize (ASAP) competition, where participants were asked to develop a system that could automatically grade short student essays. The winning approach used BERT as the feature extractor and a regression model to predict the scores of the essays.
Overall, the use of BERT in descriptive answer evaluation offers a way to automate and streamline the grading process, potentially reducing the time and effort required by teachers or instructors while providing more consistent and objective grading.

- **Cosine similarity:**

The cosine similarity metric can be utilized to compare the similitude between two composed reactions to a given provoke. It may be a degree of the similitude between two vectors of an n-dimensional space. Within the setting of descriptive response assessment, each reaction can be considered a vector in a high-dimensional space, where each measurement compares to a distinctive word or include. To compute cosine similarity, the point between the two vectors is calculated. If the point is little (i.e., the vectors are indicating in a comparative heading), the cosine similarity score will be high. Alternately, if the point is huge (i.e., the vectors are indicated in inverse headings), the cosine similarity score will be low.
cosine similarity can be utilized to compare the likeness between two reactions. For illustration, if two reactions have a tall cosine similarity score, it proposes that they are comparative in substance, structure, and/or language utilized.
On the other hand, if the cosine similarity score is low, it shows that the 2 responses are different.

3. OBSERVATION AND OUTCOMES OF SURVEY

The important takeaway from these papers are: Recurrent neural networks (RNNs) of the LSTM (Long Short-Term Memory) variety are particularly effective at processing sequential data, including text. LSTM networks can be used to evaluate answers in a descriptive manner, but they are not as commonly used as transformer-based models like BERT.

To use an LSTM for descriptive answer evaluation, the model would need to be trained on a dataset of labelled descriptive answers. The LSTM would learn to map the input text to a numeric score that reflects the quality of the answer. The model would need to be trained using a suitable loss function, Backpropagation is used to optimize metrics like binary cross-entropy or mean squared error.

In the context of descriptive answer evaluation, LSTMs can play several roles:

1. **Sequence Modeling:** LSTMs are capable of modeling sequences of words or tokens in a text. They can capture the contextual information and dependencies between words, which is crucial for understanding the meaning and generating accurate descriptive answers. LSTMs can take a sequence of words as input and produce hidden states that represent the contextual information at each step.
2. **Answer Generation:** LSTMs can be used to generate descriptive answers given a question or a context. By training an LSTM-based sequence-to-sequence model, where the input is the question or context and the output is the generated answer, the LSTM learns to encode the input information and generate relevant responses. This approach enables the LSTM to generate descriptive answers based on the learned patterns in the training data.
3. **Evaluation Scoring:** LSTMs can also be used to score or rank the quality of descriptive answers. By training an LSTM-based model to predict the quality or relevance of an answer given a question or reference answer, the model can learn to evaluate and assign scores to different candidate answers. This approach allows for automated evaluation of descriptive answers, which can be useful in tasks such as question answering systems or automated essay grading.
4. **Contextual Understanding:** LSTMs, especially when combined with attention mechanisms, can help capture the contextual understanding of the input text. By attending to relevant parts of the context or question, LSTMs can focus on the most important information for generating or evaluating descriptive answers. This enables the model to consider the entire context and make informed decisions.

Overall, LSTMs provide a powerful framework for modeling sequential data and capturing dependencies in textual data. Their ability to handle long-term dependencies and encode contextual information makes them suitable for tasks related to descriptive answer evaluation in machine learning and NLP. However, it's important to note that the choice of LSTM as the underlying architecture depends on the specific requirements of the task and the available data.

One potential advantage of using an LSTM for descriptive answer evaluation is that it may be more interpretable than a transformer-based model like BERT. LSTM networks can be visualized using techniques like attention maps, which can help identify which parts of the input text are most important for the models. There are different research works where [1] they utilized Google's pre-trained word2vec model. For the train and test, the corpus was split in an 8:2 ratio. Two approaches, Cosine Similarity and Word Mover's Distance using the Multinomial Naive Bayes model, were used to determine the results. Additionally, the word mover's distance outperforms cosine similarity. They accomplished this using Natural Language Processing (NLP) and machine learning. They got an accuracy of 88%.

[2] In this paper, system proposed an AI network which is capable of remembering the content of essay and making summary out of it, called recurrent neural network. The max-pooling layer of the Bi-GRU network is added on the word embedding for representation of text where words are similar in length or meaning of the document. The score will be get compared on the basis of the outline of essay from the additional Bi-GRU network. After completing the accuracy of the system was 0.55.

[3] In this paper, system was projected in the two phases. In the first phase, the main job is for allocation marks on the basis BERT semantic data from the article. In another phase, the score is predicted on some of the common and general factors like syntax, extent of the essay and the no. of sentences. The average accuracy score for the both phases is 70.9%.

[4] In this paper, technique proposed the Least Square Assumptions regression technique with the fusion of fuzzy Ontology. They also have been using the two sorts of features, first one is syntax features and other is semantic features. In the first feature, which is syntax features they found tokenization and then constructed entity of character string. The semantic features are more likely identical to the similarity analysis. And here the similarity analysis is used to search for identical sentences in the essay.

[5] In this paper, technique proposed a Bi-directional Encoder Representation from Transformer algorithm is used for automatic one sentence answer marking. In this suggested answer is matched with the students handwritten answer and convey the results accordingly. The data expansion is completed with the help of the AI neural networks and another that classify leftover replies as correct or incorrect using one accurate result from the dataset.

[6] In this paper, technique proposed an extreme gradient boosting (XGBoost) is a ML classifier which is used to evaluate the essays and even there are multiple algorithm competent on structures like the word count and consistency in the documents with the sentence similarity %. In order to calculate the system's result, they used the K-fold cross-validation approach; the typical accuracy following particular validations is 68.12.

[7] In this paper, the proposed system includes various evaluation techniques for the summoned output. Latent Semantic Analysis, Generalized LSA, Entropy, and other techniques have been used. Using the basic and common dataset consisting of technical answers of Computer Science these various methods/techniques were put to place. As a

testing Model, nearly 6 answers were used to analyze on single question. This is performed both with and without ontology. The outcomes were more satisfying with the application of ontology.

[8] In this paper, system proposes to implement a software application that will evaluate the students' answers using the main force-Machine Learning. It will create its own dataset that will consist of answers and keywords. Students' answers will be evaluated and the marks will be allotted on these two parameters. This system will primarily work on three stages: Separating keywords and synonyms, matching keywords, Generating Scores.

[9] In this paper, system proposes the implementation of deep learning along with NLP to counter the problem of answer evaluation. Deep learning is used by the algorithm to extract the anticipated responses. The model's categorization accuracy is improved using the sequential ensemble approach. It boosts the accuracy by a mammoth 50% when compared to the individual systems.

[10] In this paper, various parameters used for evaluation are grammar checking, sentence similarity, sentence splitting, and Jaccard similarity. Three parts make up this essay: Phase 1 involved scanning handwritten letters into digital format using OCR; Phase 2 involved dividing student responses into sentences, applying sentence similarity, grammatical checking, and Jaccard similarity using Bi-directional Encoder Representations from Transformers (BERT)., Phase 3- Based on weighted average scores are displayed.

[11] In this paper, consists of units for the user answer extractor, answer verifier, and result set. The Cosine Similarity and Text Gears Grammar API modules make up the answer verifier unit. The outcome Set Unit consists of 3 parts, keywords, grammar, and QST (Questions specific terms). Its efficiency in 80 to 90%

[12] In this paper, there are two modules: 1) extracting data from scanned copy and organizing it, 2) Applying ML and NLP. This model uses tokenizing words and sentences, chunking chinking, Lemmatizing words, parts of speech tagging and word netting. This algorithm provides semantic meaning of the context. It gives accuracy of up to 90%.

[13] In this paper, system proposes that four categories of automated essay grading systems are recognized: Regression models, classification models, neural networks, and ontology-based approaches are listed in that order. Most of the classification models and regressions on essay grading employ statistical characteristics to obtain the final score. The three categories of emphasizing aspects used in descriptive rating systems for essays include: 1) Statistical features 2) Features depending on style 3) Features based on contacts.

[14] In this paper, system proposes few highlighting features like Balance load, ease of use and user-friendly atmosphere. Our proposed system gives an accuracy of 95%.

3.1 Strengths and weaknesses of different techniques:

The below table 2 lists the strengths and weaknesses of different ML techniques.

Table 2. Strengths and weaknesses of different ML techniques studied

NAME OF TECHNIQUE	STRENGTH	WEAKNESS
LSTM	A greater understanding of sentence context than any other model. Instead of studying words, it examines phrases to help NLP.	It becomes more difficult when dealing with internet tasks like prediction. LSTMs can be a tad slow on large datasets.
LSA	The reduced vector size is an option. Positive results can be trained on lesser datasets.	For Word2Vec, larger datasets yield better outcomes. It has the capacity to ignore grammatical rules.
Word Mover's Distance (WMD)	To assess if two pieces of data are equal, it employs both a syntactic and a semantic approach.	You may also use Word2vec or Glove to get outcomes.
Cosine Similarity	It is a really helpful approach when working with NLP. No matter how large the papers or data are, it measures them. If you're working with word texts, sparse data, etc., you can choose to use cosine similarity. Cosine similarity disregards 0-0 results.	The fact that the vector's magnitude and direction are only taken into account is a significant negative. Two frequency vectors cannot be separated by more than a 90° angle.
Tokenization	It breaks down lengthy paragraphs into smaller pieces of data say words or one-line phrases making it easier to deal with.	The security of the technology is vulnerable as it practices multiple complex methods.

Bag of Words (BoW)	Simple and straightforward to use and implement. It analyses data based on the number of words. It can be simply accomplished using the Python concept of the dictionaries.	The addition of new words would increase the dataset size and thus the time required to analyse the data. Adding more vectors would result in a sparse matrix.
TF-IDF	Simple and easy to grasp. It is also a good starting point for similarity calculations because it is relatively cheap.	It has been observed that when using this tool, it has a tendency to assign low ratings to words of greater importance and those that are frequently used in the document. It can be wide-ranging and sensitive as per the condition.
BERT	Because it is updated often, it provides more accuracy. About 100 different languages have pre-trained versions of this model.	The most frequent issue with using this technology is usually cost. The project's budget typically rises as the project's scope does as well.
Word2Vec	Equally small and big datasets are accessible. Little effort is made by people. It has a highly strong architecture. It aids in defining the context of the remarks made.	When it involves identifying terms that have never been used in the data set or that are officially out of vocabulary (OOV), this method falls short.

4. LIMITATIONS OBSERVED

This research review can save the time and effort of researchers by eliminating the need to read pointless review publications. As a result, the researchers will be able to cut out errors from the beginning of their research and will have a clear path for carrying out a solid systematic survey on the subject of Descriptive answer evaluation using machine learning.

However, this study has only included a limited number of studies that were crucial for the research analysis. This review has limitations in terms of the time period considered and the number of research papers examined.

5. CONCLUSION

After surveying and taking some concerned factors into account the various technologies that can be used for descriptive answer evaluation, from the survey we can conclude that Long Short-Term Memory (LSTM) can the best approach. We came across various approaches and tools that match our scope for the evaluation process. We discussed them in detail and considered the human factor and accuracy resulted via each of them. This study is more dedicated to deal with LSTM and LSA techniques. LSA can help to reduce the vector size and help in converting a large dataset into a comparatively smaller one. This would help with reducing the computational time. This study can compare the semantics of two different text documents and extract keywords from sentences or match the sentence with the target sentence and determine the result.

REFERENCES

1. MUHAMMAD FARRUKH BASHIR, HAMZA ARSHAD, ABDUL REHMAN JAVED, NATALIA KRYVINSKA, AND SHAHAB S. BAND, "Descriptive Answers Evaluation Using Machine Learning and Natural Language Processing", date of current version December 7, 2021, Digital Object Identifier 10.1109/ACCESS.2021.310910.
2. Ruseti, S. *et al.* (2018). Scoring Summaries Using Recurrent Neural Networks. In: Nkambou, R., Azevedo, R., Vassileva, J. (eds) Intelligent Tutoring Systems. ITS 2018. Lecture Notes in Computer Science(), vol 10858. Springer, Cham. https://doi.org/10.1007/978-3-319-91464-0_19
3. Linzhong Xia, Jun Liu, Zhenjiu Zhang, "Automatic Essay Scoring Model Based on Two-Layer Bi-directional Long Short Term Memory Network". In: Proceedings of the 2019 3rd International Conference on Computer Science and Artificial Intelligence, page 133–137, December 2019. https://doi.org/10.1145/3374587.3374596
4. Darwish, S.M., Mohamed, S.K. (2020). Automated Essay Evaluation Based on Fusion of Fuzzy Ontology and Latent Semantic Analysis. In: Hassanien, A., Azar, A., Gaber, T., Bhatnagar, R., F. Tolba, M. (eds) The International Conference on Advanced Machine Learning Technologies and Applications (AMLTA2019). AMLTA 2019. Advances in Intelligent Systems and Computing, vol 921. Springer, Cham. https://doi.org/10.1007/978-3-030-14118-9_57
5. Lun, J., Zhu, J., Tang, Y., & Yang, M. (2020). Multiple Data Augmentation Strategies for Improving Performance on Automatic Short Answer Scoring. Proceedings of the AAAI Conference on Artificial Intelligence, *34*(09), 13389-13396. https://doi.org/10.1609/aaai.v34i09.7062
6. Y. Salim, V. Stevanus, E. Barlian, A. C. Sari and D. Suhartono, "Automated English Digital Essay Grader Using Machine Learning," *2019 IEEE International Conference on Engineering, Technology and Education (TALE),* 2019, pp. 1-6, doi: 10.1109/TALE48000.2019.9226022.
7. M. Syamala Devi, Himani Mittal "Machine Learning Techniques With Ontology For Descriptive Answer Evaluation". In International Journal on Natural Language Computing (IJNLC) Vol. 5, No.2, April 2016.

8. Ms. Sharmeen J. Shaikh , Ms. Prerana S. Patil , Ms. Jagruti A. Pardhe , Ms. Sayali V. Marathe , Ms. Sonal P. Patil, "AUTOMATED DESCRIPTIVE ANSWER EVALUATION SYSTEM USING MACHINE LEARNING", Vol-7 Issue-3 2021, IJARIIE-ISSN(O)-2395-4396
9. Moholkar K.P., Patil S.H. (2021) Deep Ensemble Approach for Question Answer System. In: Pandian A., Fernando X., Islam S.M.S. (eds) Computer Networks, Big Data and IoT. Lecture Notes on Data Engineering and Communications Technologies, vol 66. Springer, Singapore. https://doi.org/10.1007/978-981-16-0965-7_2
10. Shreya Singh, Omkar Manchekar, Ambar Patwardhan, Prof. Uday Rote, Prof. Sheetal Jagtap, Dr. Hariram Chavan, "Tool for Evaluating Descriptive Answers using AI (TESA)", IEEE International Conference on Communication information and Computing Technology (ICCICT), June 25-27, 2021
11. Jagadamba G, Chaya Shree G, "Online Descriptive answer verifying system Using Artificial Intelligence", Proceedings of the Fourth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC) IEEE Xplore Part Number:CFP20OSV-ART; ISBN: 978-1-7281-5464-0
12. Piyush Patil, Sachin Patil, Vaibhav Miniyar, Amol Bandal, "Descriptive Answer Evaluation Using Machine Learning", International Journal of Pure and Applied Mathematics Volume 118 No. 24 2018, ISSN: 1314-3395 (on-line version) url: http://www.acadpubl.eu/hub/
13. Dadi Ramesh,Suresh Kumar Sanampudi, "An automated essay scoring systems: a systematic literature review", Springer Nature B.V.2021 , Artifcial Intelligence Review. https://doi.org/10.1007/s10462-021-10068-2
14. Vishwas Tanwar, "Machine Learning based Automatic Answer Checker Imitating Human Way of Answer Checking", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, Vol. 10 Issue 12, December-2021
15. Farrukh Bashir, Hamza Arshad, Abdul Rehman Javed, Natalia Kryvinska, Shahab S. Band, (2021). Descriptive Answers Evaluation Using Machine Learning and Natural Language Processing. IEEE Access. PP 1-1. 10.1109/ACCESS.2021.3130902.
16. Devi, M. S., & Mittal, H. (2016). Machine Learning Techniques with Ontology for Descriptive Answer Evaluation. *ArXiv*. https://doi.org/10.5121/ijnlc.2016.5201

Integrating Sparrow Search Algorithm with Deep Learning for Tomato Fruit Disease Detection and Classification

[1,*]**K. Sundaramoorthi**
Research Scholar,
Department of Computer and Information Science,
Annamalai University, Annamalainagar-608002
Tamil Nadu, India
csksundar7@gmail.com

[2]**Mari Kamarasan**
Assistant Professor,
Department of Computer and Information Science,
Annamalai University, Annamalainagar-608002
Tamil Nadu, India
mkrasanau@gmail.com

Abstract—**Lately, the agricultural industry has considerably increased the applications of deep learning (DL) systems for recognizing and alleviating tomato fruit diseases, which address the major problems in food security and crop management. Leveraging the power of DL approaches, practitioners and researchers have established more efficient solutions capable of accurately detecting and classifying different tomato fruit diseases, enabling timely intervention to safeguard crop health and prevent yield losses. The DL techniques have been trained to detect subtle visual cues indicative of common diseases, including late blight, bacterial spot, and early blight, by harnessing vast datasets comprising annotated images of diseased tomato fruits. This sophisticated detection system offers a high level of performance. It demonstrates adaptability and scalability over different environmental conditions, empowering the farmers with actionable insight to optimize agricultural productivity and deal with disease outbreaks successfully. This study introduces a Sparrow Search Algorithm with Deep Learning for Tomato Fruit Disease Detection and Classification (SSADL-TFDDC) technique. The SSADL-TFDDC technique mainly exploits the optimal DL model to detect diseases in tomato fruits. In the SSADL-TFDDC technique, the bilateral filtering (BF) technique removes noise. The SSADL-TFDDC technique exploits the residual network (ResNet50) model for feature extraction, and the SSA performs the hyperparameter tuning process. At last, a long short-term memory (LSTM) classifier was employed to identify and classify the existence of tomato fruit diseases. A brief set of simulations exemplifies the enhanced solution of the SSADL-TFDDC technique. The experimental validation stated that the SSADL-TFDDC technique accomplishes better detection results.**

Keywords— Tomato Fruit Disease Detection; Deep Learning; Sparrow Search Algorithm; Bilateral Filtering; Computer Vision

I. INTRODUCTION

Tomatoes are an adequate significant garden harvester and the best topic of study in seed growth for expanding harvest. With numerous other harvests, harvesting is a complete job; therefore, it is a physical dimension of phenotyping data [1]. Recently, there has been growing attention to mechanizing farming procedures such as pruning, harvesting, or restricted covering. It has inspired the growth of computer vision (CV) and image analysis models for recognizing vegetables and fruits [2]. So, imaging is a rapid and nondestructive method of measurement and recognition of unripe and ripe fruits and other plant attributes utilizing CV, which is also beneficial for profiling and harvest estimates [3]. The amount of fruits throughout plant development is a significant feature, and it is a sign of the predictable crop, but it is also essential for definite crops like apples, where harvest should be measured to evade regular tree pressure [4].

Usually, farmers use physical models for the diagnosis of plant disease [5]. Specialists like plant botanists, pathologists and agricultural experts direct such diagnosis approaches over visual examination by lab experiments. These methods are inappropriate for local farm performances due to lower human setup capability and higher prices [6]. Cost-effective and real-time disease detection models are vital for the on-time recognition and analysis of tomato harvest illnesses. Due to these causes, it becomes essential to develop tools for the automated analysis of plant disease utilizing image processing (IP) and CV techniques [7]. Fast disorder analysis with initial intrusions will decrease effects on the complete harvest and food supply chains. Beforehand, the researchers employed classical machine learning (ML) models like K-Means (KM), SVM, LR, and Multi-Layer forward prorogation neural networks for plant disorder identification and verified benefits in image recognition [8]. However, these methods rely on complex handcrafted feature extractor and dimension decrease methods, which want to be improved when the fundamental issues or dataset changes [9]. Therefore, these methods need broad human power and will not simplify well. In the present scenario, Artificial Intelligence (AI) has been originating extensive uses in farming. Deep learning (DL) was mostly extensively functional in classifying plant disease, weed recognition, pest control, crop management, and farmland organization [10].

Mputu et al. [11] present a new method. The technique influences pretrained CNN and classical ML models for feature extractors and classification. The computer NVIDIA Jetson TX1 of single-board was employed. Image preprocessing and tuning models allow deep layers to absorb and focus on complex features. Anu et al. [12] project a developed DL model. This function is used as an IP model; a multiclass CNN was used for classification. Bora et al. [13] projected a structure for tomato plant disorder recognition. Initially, the input image colours are spread into the HIS format. Then, the green pixels were covered. Later, the root and fruit are distinguished using the region of interest (RoI). Finally, the unhealthy areas were divided using the rectilinear (KM clustering (RKMC) system).

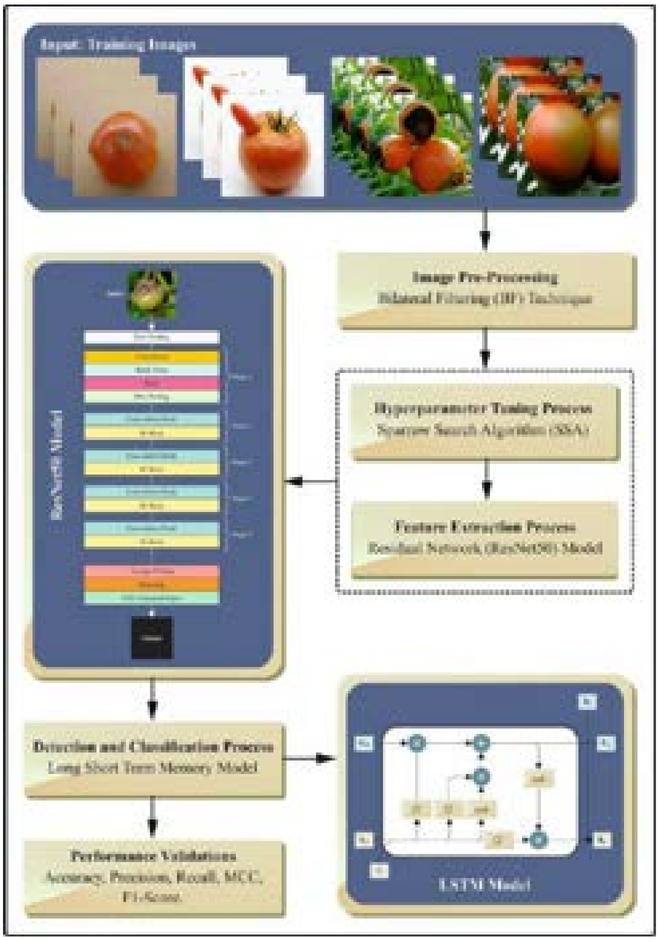


Fig. 1. Overall process of SSADL-TFDDC technique

Paul et al. [14] presented a lightweight traditional CNN method and utilized transfer learning (TL) based techniques VGG-19 and VGG-16 for categorization. The seven classes are accessible, but one is well employed to treat many tomato leaf disorders. Moreover, ablation research was executed to discover optimum limits for the projected method. Attallah [15] suggests a model using three condensed CNNs. Then, it integrates features from the 3 CNN to take advantage of each CNN framework. Nawaz et al. [16] developed a robust DL-based model such as ResNet34-based Fast RCNN for classification. Initially, the clarifications of the suspected imageries are produced. Next, ResNet34 and the Convolutional Block Attention Module (CBAM) model are presented as a feature extraction module of Fast RCNN. Lastly, the features were used for the Fast RCNN method training.

Several research studies present novel techniques for tomato quality evaluation and disease classification utilizing DL methods. These techniques employ pretrained CNN methods for feature extraction and classical ML methods for classification, aiming to improve tomato plant disorder detection and disease segmentation via sophisticated image preprocessing and model tuning.

This study introduces a Sparrow Search Algorithm with DL for Tomato Fruit Disease Detection and Classification (SSADL-TFDDC) technique. The SSADL-TFDDC technique mainly exploits the optimal DL model to detect diseases on tomato fruits. In the SSADL-TFDDC technique, the bilateral filtering (BF) technique removes noise. The SSADL-TFDDC technique exploits the residual network (ResNet50) model for feature extraction, and SSA performs the hyperparameter tuning process. Finally, a long short-term memory (LSTM) classifier is utilized to identify and classify the existence of tomato fruit diseases. The experimental values stated that the SSADL-TFDDC technique accomplishes better detection results.

II. The Proposed Model

This study establishes a novel SSADL-TFDDC technique. The technique mainly exploits an optimal DL model to detect the existence of diseases in tomato fruits. Fig. 1 demonstrates the process of the SSADL-TFDDC model.

A. BF based Preprocessing

The SSADL-TFDDC technique takes place, and the BF technique is used to remove noise. A versatile image processing algorithm, Bilateral filter (BF), is becoming more widespread regarding tomato image analysis to aid in disease detection and classification and improve visual clarity [17]. By retaining edges while smoothing the image region simultaneously, BF dramatically decreases noise while maintaining important structural details, enhancing the image quality. This refined image is a critical input for succeeding ML approaches to detecting and classifying different tomato diseases, enabling more reliable and accurate detection. Practitioners and researchers could boost the tomato disease detection system's effectiveness by incorporating BG into the image-preprocessing task, ultimately contributing to progressions in crop management strategies and agricultural practices.

B. Feature Extraction Process

For feature extraction, the SSADL-TFDDC technique exploits the ResNet50 model. ResNet structure integrates various collections of remaining blocks to attain the residuals among the network layers' input and output [18]. The loss function of the network is represented in Eq. (1). Once prolonged to multiple-layer networks, the resultant of all the layers of the network is demonstrated in Eq. (2). Based on the standard of the chain rule, the gradient of $the\ i^{th}$ layer is reduced as defined in Eq. (3)

$$L = Loss(X_n, Y) \tag{1}$$

whereas L implies the loss ratLossos$Loss$dicates the loss function, X_n represents the model outcome, and Y represents the label.

$$X_n = F_{n-1}(X_{n-1}, W_{n-1}), \\ X_{n-2} = F_{n-2}(X_{n-2}, W_{n-2}), \cdots, \\ X_2 = F_1(X_1, W_1) \tag{2}$$

In which F denotes the model, and W represents the parameter of the model

$$\frac{\partial L}{\partial X_i} = \frac{\partial L}{\partial X_n}\frac{\partial X_n}{\partial X_{n-1}} \times \cdots \times \frac{\partial X_{i+1}}{\partial X_i}. \tag{3}$$

From this assumption, it can be noticeable that the back-propagation (BP) of error, the gradient of the previous layer from the network, gradually reduces.

The residual block has been established to resolve these issues, as illustrated under the Conv and Identity block on the right side. The transformation adjusts the resultant layer from $H(x) = F(x)$ to $H(x) = F(x) + x$. Although the enhancing network depth, the problem of vanishing gradients is avoided

$$X_n = F_{n-1}(X_{n-1}, W_{n-1}) + X_{n-1}, \cdots, X_1 = F_1(X_0, W_0) + X_0 \tag{4}$$

$$\frac{\partial L}{\partial X_i} = \frac{\partial L}{\partial X_n} \times \cdots \times \frac{\partial X_{i+1}}{\partial X_i} = \frac{\partial L}{\partial X_n}\left[1 + \frac{\partial}{\partial X_i}\sum_{L=i}^{n-1} F(X_L, W_L)\right]. \tag{5}$$

The residual component contributes to the decrease in parameter counts. This strategy primarily deploys a 1x1 convolutional to reduce the channel sizes from 256 to 64 before reinstating the new dimensional using another 1x1 convolutional. This pattern gives an output in a significant decrease in the overall parameter counts related to the standard residual module. Concurrently, as the network depth rises, the number of ReLu functions increases, thus improving the model's significant nonlinear power and generalized ability.

C. SSA-based Hyperparameter Tuning

In this work, SSA performs the hyperparameter tuning process. In this work, SSA performs the hyperparameter tuning process. The population containing n sparrows is signified in the below-mentioned method [19]:

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,d} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,d} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,d} \end{bmatrix} \tag{6}$$

Whereas d indicates the problem size to be resolved. The value of fitness for every sparrow is denoted in the following method:

$$F_X = \begin{bmatrix} f([x_{1,1} & x_{1,2} & \cdots & x_{1,d}]) \\ f([x_{2,1} & x_{2,2} & \cdots & x_{2,d}]) \\ \vdots & \vdots & \vdots & \vdots \\ f([x_{n,1} & x_{n,2} & \cdots & x_{n,d}]) \end{bmatrix} \tag{7}$$

Here, F_X is the fitness value of every sparrow, and f denotes an individual's fitness value.

In the sparrow population, the main charge of the founder is to hunt for nutrition, delivering the complete population with the position and way of the detected food. Meanwhile, the founders are more probably to discover food. If the producer's fitness value is more significant, then their location in the whole solution space is near the position of the optimum solution. In every iteration, the founder upgrades its location while hunting for food. Where the exact expression is stated below:

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^t \cdot \exp\left(-\frac{i}{\alpha \cdot iter_{\max}}\right) & if\ R_2 < ST \\ X_{i,j}^t + Q \cdot L & if R_2 \geq ST \end{cases} \tag{8}$$

Here, t signifies the existing iteration count; $j = 1,2,3, \dots, d$; $iter_{\max}$ denotes the existing highest iteration count; X_{ij} is the location of the ith sparrow beside the jth dimension; and $\alpha \in (0,1)$ signifies uniform randomly generated value ST $(ST \in [0.5,1])$ and $R_2 (R_2 \in [0,1])$ represent the values of safety and warning, correspondingly. Q denotes the randomly produced number by ensuing a usual distribution, and L indicates a $1 \times d$ matrix, whereas each element is equivalent to 1.

$R_2 < ST$ states that no hunters are currently in the foraging environment, letting the founder discover quickly and widely. However, $R_2 \geq ST$ designates that hunters and signal warnings have noticed definite sparrows in the population to other members. Therefore, every sparrow should quickly move to other safe places for searching.

While searching, scroungers constantly detect the founder. Depending upon recognition, the founder will detect higher food and quickly abandon their location to compete for nutrition. The location upgrade procedure for scroungers is defined below:

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\frac{x_{worst} - x_{i,i}^t}{i^2}\right) & if\ i > n/2 \\ X_P^{t+1} + \left|X_{i,j}^t - X_P^{t+1}\right| \cdot A^+ \cdot L\ otherwise \end{cases} \tag{9}$$

Whereas X_P denotes the existing optimum location by the founder; X_{worst} signifies the present worst location; A refers to the $1 \times d$ matrix containing elements arbitrarily allocated numbers of 1 or -1, and $A^+ = A^T(AA^T)^{-1}$. If $i > n/2$, the ith lesser value of fitness has not protected food, so it is starving. Therefore, it must move to any other region to search for and reload its energy.

In the imitation experimentations, the percentage of sparrows showing risk perception ranged between 10 and 20 per cent of the entire population. Their early locations are arbitrarily spread through the whole population, and their exact depiction is stated below:

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^t + \beta \cdot \left|X_{i,j}^t - X_{best}^t\right| & if\ f_i > f_g \\ X_{i,j}^t + K \cdot \left(\dfrac{\left|X_{i,j}^t - X_{worst}^t\right|}{(f_i - f_w) + \varepsilon}\right) & if\ f_i = f_g \end{cases} \quad (10)$$

Whereas X_{best} signifies the existing global best position; β signifies the step-size parameter of control, which randomly generated amount by ensuing a normal distribution by a variance of 1 and mean of 0; $K \in (-1,1)$ represents the randomly generated number; f_w and f_g denote the existing global worst and best fitness value, correspondingly; f_i refers to the fitness score of the existing sparrow; and ε signifies the mathematically small even to stop division by 0. If $f_i > f_g$, the sparrow at the border was helpless to predator assaults, representing that the sparrow's location is secure. If $f_i = f_g$, sparrows in the centre have noticed risk, so they want to join with other sparrows to diminish the danger of predation.

The fitness choice is the primary factor affecting SSA achievement. The parameter choice method contains an encoded technique for evaluating the candidate solution's efficiency. At this point, the SSA considers accuracy the critical criterion for developing the FF.

$$Fitness = \max(P) \quad (11)$$

$$P = \frac{TP}{TP + FP} \quad (12)$$

Where TP and FP are the true positive and denote the false positive values.

D. Disease Detection Using LSTM Model

Finally, the LSTM classifier detects and classifies tomato fruit ailments. LSTM is a variant of recurrent neural networks (RNN) implemented to comprehend context- specific data from massive datasets and utilize them for future anticipation [20]. Each LSTM cell or unit can remember the data for a shorter or longer period. The output of the predicted new cell state takes the data stored in the previous cell state. This gives memory to the network, improving future prediction. Therefore, the LSTM network is a better fit for finding how stock closing prices and financial news headlines affect stock price trends for longer. Also, this network decides what amount of historical data related to the stock price is collected for precisely anticipating new price trends.

The LSTM model is seen as a gated cell. Gated indicates that the cell defines what information is stored or deleted based on the weight it allocates to the data. LSTM includes three gating mechanisms: forget, input and output gates. The input gate, i_t, defines what value is updated by the input signal. The Forget Gate, f_t, determines what state to be retained or forgotten. The output gate, o_t, defines how the cell state affects the other neurons. Furthermore, it has a logistic layer where the previous layer generates a number within [0,1] and a new vector that adds up to the state. The LSTM equation is given below in Eqs (13) to (18). The outcome of the embedding layer, x_t, is employed as an input to the RNN layer, W refers to the weight matrix, and y_t is the output. Three gating mechanisms control the essential component of LSTM s_t and serve as a memory unit.

$$f_t = \sigma(W_{xf}X_t + W_{yf}Y_{t-1} + b_f) \quad (13)$$

$$i_t = \sigma(W_{xi}X_t + W_{yi}Y_{t-1} + b_i) \quad (14)$$

$$g_t = \tanh(W_{xg}X_t + W_{yg}Y_{t-1} + b_g) \quad (15)$$

$$s_t = f_t * s_{t-1} + i_t * g_t \quad (16)$$

$$o_t = \sigma(W_{XO}X_t + W_{yo}Y_{t-1} + b_o) \quad (17)$$

$$y_t = O_t * \tanh(s_t) \quad (18)$$

III. EXPERIMENTAL VALIDATION

This section examines the performance of the SSADL-TFDDC approach on the dataset of 900 tomato fruit disease images collected from various sources and reported in Table 1. Fig. 2 depicts the instance images.

TABLE I

DESCRIPTION ON DATASET

Diseases	**Image Numbers**
Healthy	150
Malformed fruit	150
Blotchy ripening	150
Puffy fruit	150
Blossom-end rot	150
Gray mold	150
Total Images	**900**

Fig. 2. Sample images

Fig. 3 examines the classifier results of the SSADL-TFDDC system with a test database. Figs. 3a-3b showcases the confusion matrices acquired by the SSADL-TFDDC method at 70%TRAS and 30%TESS. The figure pointed out that the

SSADL-TFDDC algorithm can be predictable and categorize all six classes correctly. Moreover, Fig. 3c displays the PR result of the SSADL-TFDDC method. The figure denoted that the SSADL-TFDDC algorithm acquires superior PR efficiency with every class. In addition, Fig. 3d examines the ROC result of the SSADL-TFDDC technique. This figure shows that the SSADL-TFDDC technique provides effectual performances with a higher value of ROC in six classes.

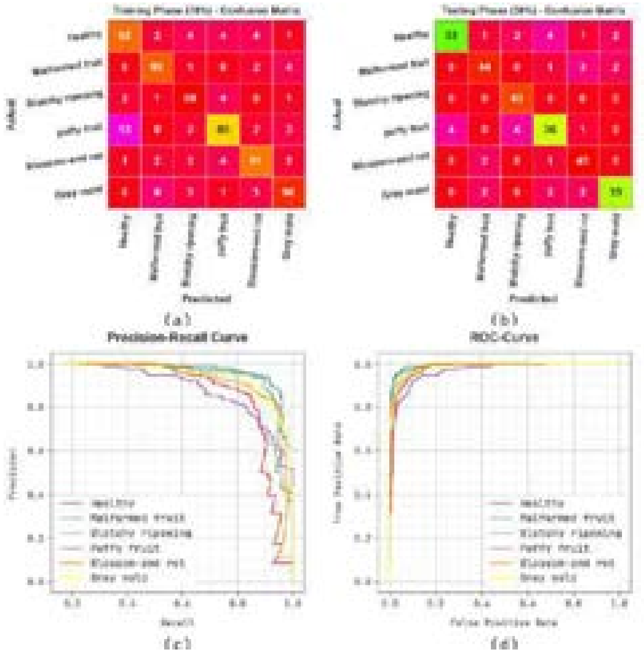


Fig. 3. (a-b) Confusion matrices under 70%TRAS and 30%TESS and (c-d) PR and ROC curves

The tomato fruit disease detection results of the SSADL-TFDDC technique are provided on 70%TRAS, and 30%TESS is provided in Table 2 and Fig. 4. This experimentation described that the SSADL-TFDDC technique correctly identified various classes. Based on 70%TRAS, the SSADL-TFDDC technique offers average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and MCC of 96.08%, 88.24%, 88.29%, 88.24%, and 85.91%, correspondingly. Also, based on 30%TESS, the SSADL-TFDDC algorithm achieves average $accu_y$, $prec_n$, $reca_l$, $F1_{score}$, and MCC of 95.80%, 87.47%, 87.31%, 88.21%, and 84.83%, respectively.

TABLE II

TOMATO FRUIT DISEASE DETECTION RESULTS OF THE SSADL-TFDDC MODEL AT 70%TRAS AND 30%TESS

Class	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$	MCC
TRAS (70%)					
Healthy	95.08	85.19	85.98	85.58	82.62
Malformed fruit	97.14	89.42	93.00	91.18	89.50
Blotchy ripening	96.83	89.19	92.52	90.83	88.93
Puffy fruit	94.76	86.73	80.95	83.74	80.69
Blossom-end rot	96.51	89.22	89.22	89.22	87.13
Gray mold	96.19	89.72	88.07	88.89	86.60
Average	**96.08**	**88.24**	**88.29**	**88.24**	**85.91**
TESS (30%)					
Healthy	94.81	89.19	76.74	82.50	79.79
Malformed fruit	95.93	89.80	88.00	88.89	86.40
Blotchy ripening	97.78	87.76	100.00	93.48	92.43
Puffy fruit	93.70	81.82	80.00	80.90	77.14
Blossom-end rot	96.30	86.54	93.75	90.00	87.84
Gray mold	96.30	89.74	85.37	87.50	85.36
Average	**95.80**	**87.47**	**87.31**	**87.21**	**84.83**

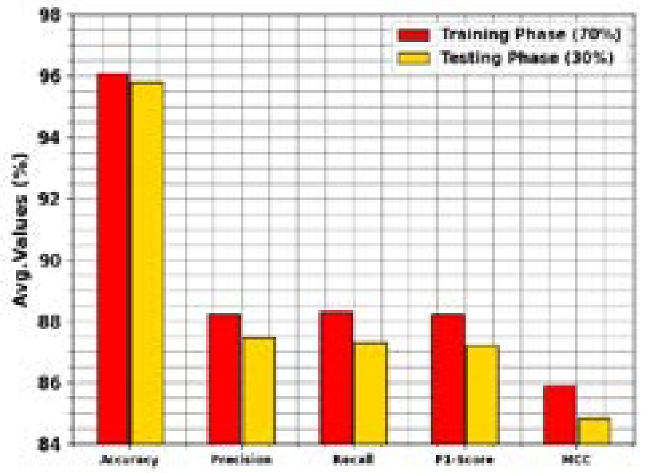


Fig. 4. Average of the SSADL-TFDDC method on 70%TRAS and 30%TESS

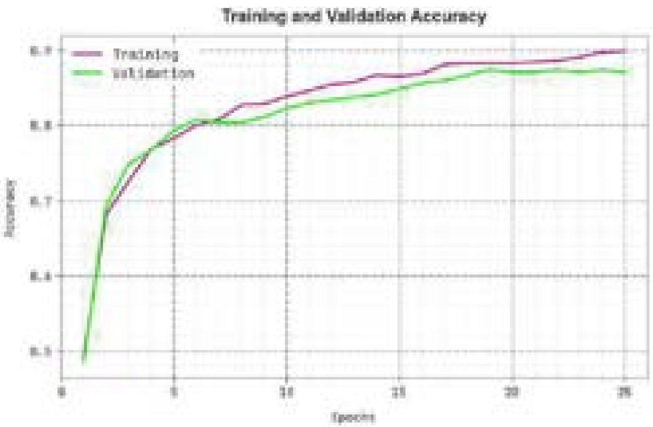


Fig. 5. $Accu_y$ curve of the SSADL-TFDDC technique

The efficiency of the SSADL-TFDDC technique is examined in Fig. 5 under training accuracy (TRAA) and validation accuracy (VALA) curves. This figure portrays the analysis of the SSADL-TFDDC technique over varying epoch counts, representing its generalization and learning process capabilities. The figure mainly illustrates a constant development in the TRAA and VALA with improved epochs.

This ensures the adaptability of the SSADL-TFDDC technique in the pattern detection progression under TRA and TES data. The higher trends in VALA portray the proficiency of the SSADL-TFDDC algorithm in modifying the TRA data and then achieving the correct classification on undetected data, showing the capabilities of robust generalization.

Fig. 6 displays a wide-ranging illustration of the SSADL-TFDDC method's training loss (TRLA) and validation loss (VALL) results in distinct epochs. The progressive reductions in TRLA show that the SSADL-TFDDC method enhances the weights and decreases the classification error under TRA and TES data. The figure specifies the SSADL-TFDDC method correlated with the TRA data, underscoring its proficiency in comprehending patterns. Mainly, the SSADL-TFDDC method incessantly improves its parameters, lessening the variances among the anticipated and actual TRA classes.

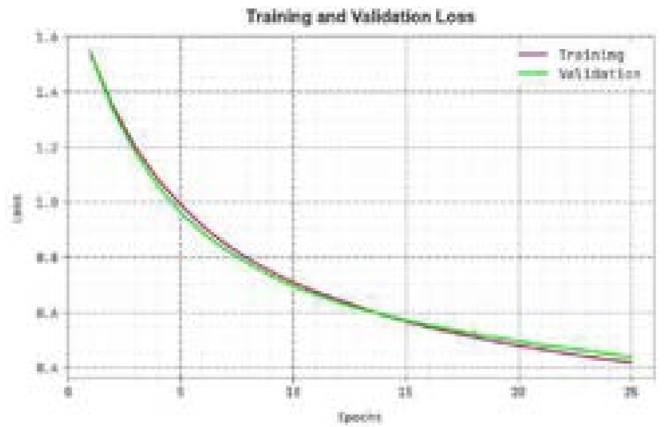


Fig. 6. Loss curve of SSADL-TFDDC method

Finally, the comparative assessment of the SSADL-TFDDC technique concerning $accu_y$, $prec_n$, $reca_l$, and $F1_{score}$ is reported in Table 3 and Fig. 7 [21, 22]. These obtained values stated that the SSADL-TFDDC technique performs better than existing approaches. Based on $accu_y$, the SSADL-TFDDC technique offers a higher $accu_y$ of 96.08% while the Yolov5m, ResNet50, ResNet101, EfficientNetB0, VGG16, and MobileNet methods obtain lower $accu_y$ of 90.59%, 93.20%, 94.22%, 91.93%, 93.07%, and 91.54%, correspondingly. Also, with $prec_n$, the SSADL-TFDDC method provides boosted $prec_n$ of 88.24% although the Yolov5m, ResNet50, ResNet101, EfficientNetB0, VGG16, and MobileNet techniques get decreased $prec_n$ of 80.61%, 87.14%, 83.81%, 82.89%, 85.81%, and 82.18%. Meanwhile, based on $reca_l$, the SSADL-TFDDC algorithm achieves a greater $reca_l$ of 88.29% whereas the Yolov5m, ResNet50, ResNet101, EfficientNetB0, VGG16, and MobileNet approaches attain minimized $reca_l$ of 80.84%, 86.80%, 80.00%, 86.07%, 86.30%, and 80.64%. Finally, with $F1_{score}$, the SSADL-TFDDC algorithm accomplishes increased $F1_{score}$ of 88.24% although the Yolov5m, ResNet50, ResNet101, EfficientNetB0, VGG16, and MobileNet techniques get lesser $F1_{score}$ of 82.14%, 87.92%, 83.18%, 87.54%, 81.12%, and 82.38%, respectively.

TABLE III

COMPARATIVE OUTCOMES SSADL-TFDDC METHOD WITH RECENT ALGORITHMS

Model	$Accu_y$	$Prec_n$	$Reca_l$	$F1_{score}$
Yolov5m	90.59	80.61	80.84	82.14
ResNet-50	93.20	87.14	86.80	87.92
ResNet-101	94.22	83.81	80.00	83.18
EfficientNet-B0	91.93	82.89	86.07	87.54
VGG-16	93.07	85.81	86.30	81.12
MobileNet	91.54	82.18	80.64	82.38
SSADL-TFDDC	96.08	88.24	88.29	88.24

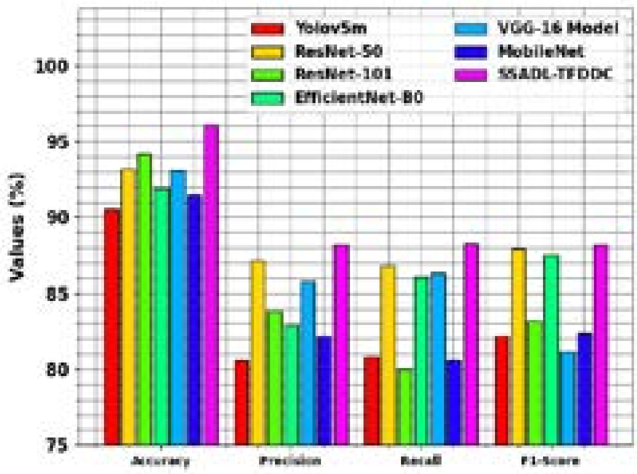


Fig. 7. Comparative results of SSADL-TFDDC technique with other systems

Thus, the SSADL-TFDDC technique can enhance the detection process of tomato fruit disease.

IV. Conclusion

In this research work, a novel SSADL-TFDDC technique is introduced. The SSADL-TFDDC technique mainly exploits the optimal DL model to detect diseases in tomato fruits. In the SSADL-TFDDC technique, the BF technique is used to remove noise. The SSADL-TFDDC technique exploits the ResNet50 model, and SSA performs the hyperparameter tuning process for feature extraction. Finally, the LSTM classifier is utilized to identify and classify the existence of tomato fruit diseases. A brief set of simulations exemplifies the enhanced solution of the SSADL-TFDDC technique. The experimental validation stated that the SSADL-TFDDC technique accomplishes better detection results.

References

[1] Sunil, C.K., Jaidhar, C.D. and Patil, N., 2022. Binary class and multiclass plant disease detection using ensemble deep learning-based approach. *International Journal of Sustainable Agricultural Management and Informatics*, *8*(4), pp.385-407.

[2] Sharmila, M. and Natarajan, M., 2023, July. Modified Barnacles Mating Optimization with Deep Learning based Tomato Disease Detection and

Classification. In *2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 1209-1215). IEEE.

[3] Gadekallu, T.R., Rajput, D.S., Reddy, M.P.K., Lakshmanna, K., Bhattacharya, S., Singh, S., Jolfaei, A. and Alazab, M., 2021. A novel PCA–whale optimization-based deep neural network model for classification of tomato plant diseases using GPU. *Journal of Real-Time Image Processing*, *18*, pp.1383-1396

[4] Alzahrani, M.S. and Alsaade, F.W., 2023. Transform and Deep Learning Algorithms for the Early Detection and Recognition of Tomato Leaf Disease. *Agronomy*, *13*(5), p.1184.

[5] Anandhakrishnan, T. and Jaisakthi, S.M., 2022. Deep Convolutional Neural Networks for image based tomato leaf disease detection. *Sustainable Chemistry and Pharmacy*, *30*, p.100793.

[6] Malar, B.A., Andrushia, A.D. and Neebha, T.M., 2021, May. Deep learning based disease detection in tomatoes. In *2021 3rd International Conference on Signal Processing and Communication (ICPSC)* (pp. 388-392). IEEE.

[7] Deshpande, R. and Patidar, H., 2023. Detection of leaf disease in tomato plants using a lightweight parallel deep convolutional neural network. *Archives of Phytopathology and Plant Protection*, pp.1-14.

[8] Sagar, S. and Singh, J., 2023. An experimental study of tomato viral leaf diseases detection using machine learning classification techniques. *Bulletin of Electrical Engineering and Informatics*, *12*(1), pp.451-461.

[9] Parthiban, S., Moorthy, S., Sabanayagam, S., Shanmugasundaram, S., Naganathan, A., Annamalai, M. and Balasubramanian, S., 2023. Deep Learning Based Recognition of Plant Diseases. In *Computer Vision and Machine Intelligence Paradigms for SDGs: Select Proceedings of ICRTAC-CVMIP 2021* (pp. 83-93). Singapore: Springer Nature Singapore.

[10] Deb, S.D., Kashyap, R., Abhishek, A., Lavanya, R., Paritosh, P. and Jha, R.K., 2023, May. Tomato leaf disease detection using series of Convolutional and Depthwise Convolutional Layers. In *2023 International Conference on Signal Processing, Computation, Electronics, Power and Telecommunication (IConSCEPT)* (pp. 1-6). IEEE.

[11] Mputu, H.S., Abdel-Mawgood, A., Shimada, A. and Sayed, M.S., 2024. Tomato Quality Classification based on Transfer Learning Feature Extraction and Machine Learning Algorithm Classifiers. *IEEE Access*.

[12] Anu, A., Solanki, K. and Dhankar, A., 2024. A Novel Multiclass Deep Learning Approach for Tomato Leaf Disease Detection System. *International Journal of Intelligent Systems and Applications in Engineering*, *12*(6s), pp.187-196.

[13] Bora, R., Parasar, D. and Charhate, S., 2023. A detection of tomato plant diseases using deep learning MNDLNN classifier. *Signal, Image and Video Processing*, *17*(7), pp.3255-3263.

[14] Paul, S.G., Biswas, A.A., Saha, A., Zulfiker, M.S., Ritu, N.A., Zahan, I., Rahman, M. and Islam, M.A., 2023. A real-time application-based convolutional neural network approach for tomato leaf disease classification. *Array*, *19*, p.100313.

[15] Attallah, O., 2023. Tomato leaf disease classification via compact convolutional neural networks with transfer learning and feature selection. *Horticulturae*, *9*(2), p.149.

[16] Nawaz, M., Nazir, T., Javed, A., Masood, M., Rashid, J., Kim, J. and Hussain, A., 2022. A robust deep learning approach for tomato plant leaf disease localization and classification. *Scientific reports*, *12*(1), p.18568.

[17] Radhika, R. and Mahajan, R., 2023. An adaptive optimum weighted mean filter and bilateral filter for noise removal in cardiac MRI images. *Measurement: Sensors*, *29*, p.100880.

[18] Yuan, Q., Lu, J., Wu, L., Huang, Y., Guo, Z. and Li, N., 2024. Water-Body Type Classification in Dual PolSAR Imagery Using a Two-Step Deep Learning Method. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*.

[19] Du, H., Wang, J., Qian, W. and Zhang, X., 2024. An Improved Sparrow Search Algorithm for the Optimization of Variational Modal Decomposition Parameters. *Applied Sciences*, *14*(5), p.2174.

[20] Chandola, D., Mehta, A., Singh, S., Tikkiwal, V.A. and Agrawal, H., 2023. Forecasting directional movement of stock prices using deep learning. *Annals of Data Science*, *10*(5), pp.1361-1378.

[21] Phan, Q.H., Nguyen, V.T., Lien, C.H., Duong, T.P., Hou, M.T.K. and Le, N.B., 2023. Classification of tomato fruit using yolov5 and convolutional neural network models. *Plants*, *12*(4), p.790.

[22] Wang, Q., Qi, F., Sun, M., Qu, J. and Xue, J., 2019. Identification of tomato disease types and detection of infected areas based on deep convolutional neural networks and object detection techniques. *Computational intelligence and neuroscience*, *2019*.

A Deep Pattern Learning based Model for Detection of Cardiovascular Diseases(CVD)

Balu Srinivasulu,
Department Of Computer Science & Engineering, Pace Institute Of Technology Sciences, Ongole, AP,
22KQ1D5818@pace.ac.in.

P.V.Subba Reddy,
Associate Professor, Department Of Computer Science & Engineering, Pace Institute Of Technology Sciences,
Ongole, AP, line
subbareddy_pv@gmail.com

P.Hussain Basha,
Assistant Professor, Department Of Computer Science & Engineering, Pace Institute Of Technology Sciences,
Ongole, AP, line
phussain786@gmail.com

***Abstract*— Given that cardiovascular diseases (CVD) persist as a significant global cause of death, the development of effective diagnostic techniques is of paramount importance. This artide presents a novel method for detecting CVD using convolutional neural networks (CNN) and SqueezeNet, two cutting-edge deep learning approaches. The proposed technique leverages ECG images to reliably classify medical images associated with cardiovascular problems. The foundation of feature extraction is SqueezeNet, a model renowned for its computational efficiency and lightweight construction. SqueezeNet's convolutional layers reduce the computing power required for model training and deployment and capture intricate patterns and textures indicative of CVD conditions, potentially revolutionizing the field of cardiovascular diagnostics. Furthermore, the pre-trained SqueezeNet model is honed using transfer learning methods on a collection of annotated cardiac images. This process allows the model to adapt its learned features to the specific characteristics of CVD images, thereby enhancing classification performance. Experimental evaluation on a benchmark dataset underscores the efficacy of the proposed approach in accurately identifying various types of cardiovascular abnormalities, induding coronary artery disease, myocardial infarction, and heart valve defects. A comparative analysis against state-of-the-art methods demonstrates the superior performance and computational efficiency of the proposed SqueezeNet-based model. It reinforces the confidence in its potential to transform the landscape of cardiovascular disease diagnosis. The findings of this study mark a significant stride in the development of computer-aided diagnosis systems for cardiovascular diseases, offering a promising tool for early detection and personalized treatment strategies.**

Keywords*— *Electrocardiogram (ECG), Cardiovascular diseases, SqueezeNet Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Neural Network.

I. INTRODUCTION

The World Health Organization (WHO) states that cardiovascular illnesses, such as heart disease, are the leading cause of death worldwide. They are thought to claim 17.9 million lives a year, or 32% of all deaths worldwide. More than 85 percent of deaths attributable to heart disease are caused by myocardial infarctions (MI), sometimes known as heart attacks [1]. Numerous lives can be saved if cardiovascular disease is appropriately identified at an early stage. To detect cardiac diseases, the healthcare system employs a range of techniques, such as computed tomography, cardiac magnetic resonance imaging, electrocardiograms (ECGs), echocardiograms (echo), and blood tests [2], [3]. The ECG is a popular, affordable, and non-invasive way to measure the heart's electrical activity [4]. It is used to identify heart conditions associated with the core. A highly skilled clinician can diagnose heart disease from the ECG waves. However, doing it by hand requires much work and can lead to inaccurate results [5]. Improvements in artificial intelligence in healthcare have the potential to reduce medical errors drastically. Precisely, nine machine learning and deep learning techniques predict heart problems automatically. Before the classification stage, machine learning approaches require an expert entity to extract and choose the appropriate features. Feature extraction decreases the number of items in a data collection by projecting or transforming the data into a new, lower-dimensional feature space while retaining the relevant information from the input data. The essential features that are taken out of the input data concentrate on creating a new set of components—one that is distinct from the input feature—by fusing the original elements into a space that is smaller in dimension. The principal component analysis is the most frequently utilized feature extraction method. However, in training machine learning algorithms, feature selection removes unnecessary features (dimensions) from the data set [6]. Several feature selection methods can be divided into supervised and unsupervised. While supervised approaches require the output label, unsupervised methods do not.

It is an essential field of study for medical picture analysis. Automating the diagnosis and identification of CVDs using ECG signals has advanced significantly thanks to advances in deep learning techniques, especially CNN. The creation of more precise, effective, and scalable algorithms for the analysis of ECG data has been made possible by the DL approaches, which have entirely changed the area. Unlike earlier machine learning techniques requiring handcrafted features, these algorithms can automatically extract meaningful features from raw ECG signals. Early detection of cardiovascular diseases (CVDs) is crucial. Manual diagnosis presents obstacles that deep learning approaches may address. It also overviews the critical deep learning architectures and methodologies commonly used for CVD detection in ECG images.

Feature extraction is essential to extract pertinent information from ECG signals for additional analysis and interpretation. In image processing and computer vision, local binary pattern (LBP) is used for texture analysis and feature extraction. LBP was initially created for grayscale photos, but it has since been modified for several uses, including the study of photographs in the medical field. The ECG signal is converted into a picture format that can be used for LBP examination. The signal may be transformed using grayscale image conversion or time-frequency analysis (spectrograms, for example). LBP compares the intensity of a core pixel with the power of pixels in a nearby neighborhood. Binary patterns are created based on these comparisons, representing the image's texture and structural features. With the help of the binary patterns derived by LBP, a feature vector is created for every ECG picture. The spatial distribution of texture patterns in the image is represented by this feature vector, which encodes information about it.

Finally, in this paper, the proposed approach uses SqueezeNet and AlexNet as the pre-trained model to train the ECG images to improve CVD detection. Preprocessing and feature extraction play a significant role in finding the abnormalities. The proposed CNN model helps to classify the normal and abnormal ECG samples. The CNN model contains multiple layers that internally process several steps to help find abnormal ECG images.

The primary objectives of this paper are to:

- The significance of preprocessing for diagnosing CVD using ECG image analysis.
- Examine the standard preprocessing procedures used in processing ECG signals and their importance in enhancing the quality of data and precision of diagnosis.
- The latest developments, improvements in ECG analysis preprocessing techniques, and any possible effects on clinical practice.
- The difficulties and potential paths forward in creating and using preprocessing methods for using ECG imaging in the identification of cardiovascular disease.

Figure 1: Sample ECG Image

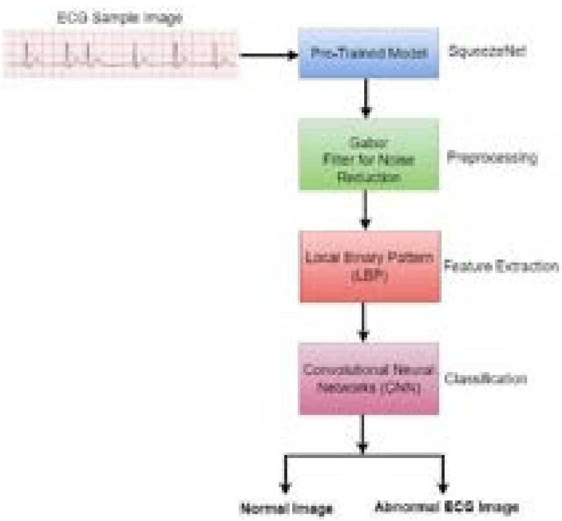


Figure 2: System Architecture

II. LITERATURE SURVEY

Anbukkarasi et al. provided an extensive overview of recent developments in ML-based methods for diagnosing cardiac disease [7]. It examines several machine learning algorithms for diagnosing cardiac disease, such as decision trees, support vector machines, neural networks, and ensemble approaches. Additionally, it looks at how feature selection and data pretreatment methods might improve the efficiency of machine learning models. Additionally, the paper discusses challenges associated with deploying ML models in clinical settings, such as interpretability, data privacy, and generalizability. Finally, it highlights future research directions and opportunities for leveraging ML to enhance heart disease identification and management accuracy and efficiency. A novel method for accurately predicting cardiac disease using hybrid machine-learning approaches was developed by Mohan et al. [8]. The hybrid model combines the best features of several techniques, such as SVM, RF, LR, DT, and NN. Principal component analysis (PCA) and recursive feature deletion are feature selection approaches to enhance the model's interpretability and prediction performance. Ensemble learning techniques like bagging and boosting are applied to increase prediction accuracy and robustness further. The dataset employed in the experiments includes a wide range of clinical variables, such as lifestyle factors, medical history, and demographic data. The hybrid model performs better than separate algorithms, as evidenced by the experimental results, which highlight the model's potential as a valuable tool for clinical practice heart disease prediction. A thorough analysis of machine learning methods used to forecast cardiac disease was provided by Devansh Shah et al. [9], emphasizing the methods' advantages, disadvantages, and potential for development. The significance of feature selection, data preparation, model selection, and assessment metrics to create reliable and accurate prediction models.

Furthermore, we examine the difficulties of incorporating machine learning models into healthcare settings and offer possible solutions. In summary, this research underscores the potential of machine learning to augment the early identification and treatment of cardiac conditions, hence augmenting patient outcomes and mitigating healthcare costs. Bhowmick et al. [10] looked into the usage of various machine-learning algorithms for heart illness prediction. The dataset used in this investigation includes features such as age, gender, blood pressure, cholesterol levels, and other clinical parameters. Several machine learning algorithms, such as LR, DT, RF, SVM, and NN, are implemented, and their prediction performance is compared. Preprocessing is done on the dataset to address missing values, normalize features, and, if any, reduce any potential class imbalance. Cross-validation techniques are used to evaluate the models' generalizability and robustness. To help cardiologists make decisions more quickly, Habib et al.'s [11] research concentrated on developing an effective method of estimating the likelihood of heart failure and, in turn, recommending suitable medications. By applying traditional exploratory analysis approaches, the research aims to discover correlations or links between the patients' various medical profiles. These correlations will allow the qualities to be appropriately used to forecast the likelihood of heart failure and medication recommendations. Mohanty et al. presented a study on applying ML algorithms to predict liver disease [12]. The study's dataset includes a range of clinical and demographic characteristics, including alkaline phosphate levels, age, gender, total and direct bilirubin levels, and others. Several machine learning techniques create liver disease prediction models, such as LR, DT, RF, SVM, and NN. The preprocessed dataset addresses any class imbalance, handles missing values, and normalizes features. A feature significance analysis is also carried out to find the most important predictors of liver disease. The work emphasizes how crucial feature selection and model assessment are to creating reliable and accurate liver disease prediction models. Using a dataset that includes characteristics like age, gender, blood pressure, cholesterol levels, and other pertinent medical indications, Shah et al. [13] want to create a predictive model for heart disease. The study's dataset, including patient records from heart disease-affected and unaffected parties, was obtained from internet sources—the contrast of several machine learning techniques used in constructing prediction models. Techniques from feature engineering are used to preprocess the data and extract pertinent information. The outcomes show how well machine learning can predict cardiac disease, with specific algorithms performing better than others in accuracy. The knowledge gathered from this research can help medical professionals identify those more likely to develop heart disease, allowing for more focused treatments and better patient outcomes. An overview of current developments in ML-based cardiac disease prediction is provided by Ramalingam et al. [14]. Several approaches are used, such as deep learning architectures, ensemble techniques, and supervised, unsupervised, and semi-supervised learning procedures. Researchers also look at the difficulties and restrictions encountered in implementing ML-based heart disease prediction systems, including issues with interpretability, model generalization, and data scarcity. Lastly, there are exciting areas for further study that will improve the functionality and performance of machine learning models in this critical area of healthcare. An improved CNN model with DL assistance was presented by Pan et al. [15] and implemented on an IoMT platform to predict cardiac disease. The suggested model uses CNN's capabilities to automatically extract pertinent features from ECG data, frequently used to diagnose heart disease.

Additionally, the model performs by incorporating sophisticated deep learning methods like transfer learning and attention processes. While transfer learning makes it possible to use previously trained models on large-scale datasets to increase efficiency and generalization, the attention mechanism allows the model to concentrate on informative portions of the ECG signals. Hasan et al.'s [16] innovative method for automatically predicting individuals with heart disease uses SDA. The goal is to create a robust predictive model using clinical and demographic data to identify those at risk of heart disease. The dataset used in this study is an extensive compilation of patient records that includes demographic data, medical history, and diagnostic test findings—the information addresses missing values standardizes characteristics and lessens biases. The SDA is then used to choose discriminative features and create a condensed classification model. The outcomes show how well Sparse Discriminant Analysis works in both precisely predicting patients with heart disease and simultaneously selecting a condensed subset of variables important for diagnosis.

A. Disadvantages of Existing System:

- The quality of ECG datasets might differ, and because professional annotation is required, obtaining labeled data for training deep learning models can be challenging.
- For instance, CNN is sometimes regarded as a "black box," which makes it challenging to understand how they make their forecasts.
- In clinical settings, real-time ECG signal processing is frequently necessary to enable prompt diagnosis and action.

III. DATASET DESCRIPTION

The dataset named PTB diagnostic ECG dataset was collected from various CVD patients suffering from heart arrhythmia. The total samples of 14552 ECG images were divided into two categories: standard and affected ECG images. The training consists of 80%, and testing consists of 20%.

A. SqueezeNet: Training Model

Due to its small memory footprint and high computational efficiency, SqueezeNet, a lightweight convolutional neural network (CNN) architecture, presents a viable method for effective ECG image processing. It is used to train ECG images. Our goal in this paper is to use ECG pictures to train a SqueezeNet model to predict heart illness. Our goal is to automate the identification of cardiac problems from ECG images by combining SqueezeNet's DL capabilities, allowing for a quick and precise diagnosis. SqueezeNet's primary goal is to decrease the network's parameter count without compromising performance. It is made possible by two key innovations:

Fire Modules: The Fire module is the name of SqueezeNet's default construction blocks. A fire module comprises expansion layers that come after a squeeze layer. Squeezing information is accomplished using the squeeze layer's 1x1 convolutions, which lower the number of input channels (depth-wise convolution). The 1x1 and 3x3 convolutions that comprise the expand layers are designed to capture spatial and channel-wise information.

Down Sampling: SqueezeNet reduces the spatial dimensions of the feature maps by sparsely applying downsampling. It is often accomplished with max-pooling layers.

B. Gabor Filter for Noise Reduction

ECG images are primarily essential for diagnosing heart-related illnesses. However, noise from various sources, including muscle activity, baseline wandering, electrode motion, and electrical interference, frequently taints ECG signal pictures. This noise can obstruct accurate diagnosis, which can mask significant ECG signal characteristics. Gabor filtering is commonly employed to eliminate noise from ECG readings. These linear filters are applied to images and used for texture analysis. They work very well at capturing the frequency and spatial properties of signals.

Gabor filters can efficiently suppress noise while maintaining significant signal characteristics in ECG signal processing. Noise elements that do not match the signal's frequency content can be attenuated by convolving the ECG signal with a Gabor filter, resulting in a cleaner depiction of the underlying heart activity. The center frequency and the bandwidth are the two factors that define the Gabor filter. While the bandwidth governs the range of frequencies the filter reacts to, the center frequency establishes the frequency to which the filter is most sensitive. Gabor filters can be tuned to target particular noise components in the ECG signal by carefully choosing these parameters. The following equation represented as:

$$\mathrm{G}(\mathrm{a}) = \exp\left(-\frac{\mathrm{a}^2}{2\sigma^2}\right)\cos(2\pi f_c a) \qquad (1)$$

a represents the filter.
$\mathrm{G}(\mathrm{a})$ Represents the Gabor filter response at a.

C. Local Binary Pattern (LBP)

LBP is a well-liked technique for obtaining texture information from pictures. LBP was initially created for grayscale picture analysis, but it has since been used in various domains, such as pattern recognition, computer vision, and medical image analysis. LBP can be applied to ECG images to extract the signal's textural information. ECG pictures often depict the heart's electrical activity over time, displaying a range of patterns and textures linked to distinct cardiac diseases. The LBP algorithm compares every pixel in an image to its nearby pixels. Depending on whether the intensity of the surrounding pixels is higher or lower than the intensity of the core pixel, each pixel is given a binary code. The local texture patterns in the image are then represented by these binary codes when they are transformed into decimal numbers. The textural properties can be extracted from ECG images by computing LBP at multiple scales and orientations. These features can be applied to medical applications for anomaly detection, segmentation, and classification. The LBP offers a potent instrument for deriving informative characteristics from ECG pictures, facilitating the creation of sophisticated analysis methods for identifying cardiac disorders and keeping tabs on patients' health.

$$\mathrm{LBP}(\mathrm{a_c}, \mathrm{b_c}) = \sum_{\mathrm{p}=0}^{\mathrm{P}-1} \mathrm{s}(\mathrm{p}) \cdot 2^{\mathrm{p}} \qquad (2)$$

$\mathrm{s}(\mathrm{p})$ Represents the binary value whether the neighboring sample's intensity (p) is higher than or equal to the central ECG input image is indicated by a (0 or 1)

Classification: CNN has become a highly effective tool for problems involving image detection, segmentation, and classification. Their suitability for analyzing medical images, particularly those from several modalities like computed tomography (CT), magnetic resonance imaging (MRI), and echocardiography, stems from their ability to learn hierarchical representations from raw input data automatically. The suggested CNN-based approach for identifying cardiovascular disorders using imaging data is presented in this work. It seeks to use the information in cardiac pictures to precisely categorize patients into various illness groups or identify specific abnormalities suggestive of cardiovascular diseases. The CNN architecture for processing images of ECG signals usually consists of several layers intended to extract features pertinent to the current job.

Step 1: Input Layer

- The input to the CNN would be the ECG image.
- Let a(t) represent the input signal, where t denotes the time axis.

Step 2: Convolutional Layer

- The features are extracted from the input signal.
- Every filter applies a convolution operation on the signal that it receives.
- Let $h_i(t)$ initializes the $\mathrm{i^{th}}\,\frac{\mathrm{filter}}{\mathrm{Kenel}}$ with parameters $\mathrm{W_i}$ and bias $\mathrm{b_i}$.
- The result of $\mathrm{i^{th}}$ filter $\mathrm{z_i(t)}$ is measured as:
  $$\mathrm{z_i(t)} = (\mathrm{x} * \mathrm{h_i})(\mathrm{t}) + \mathrm{b_i} \qquad (2)$$
  - Represents the convolution operation,

Step 3: Activation Layer

- An activation function is applied element-by-element following the convolution process to add non-linearity.
- The ReLU, is used.
- Let f() denote the activation function that was applied to the convolutional layer's output.

Step 4: Pooling Layer

- It preserves significant information while shrinking the feature maps' spatial size.
- It is a popular method of pooling that keeps the maximum value for each pooling window.
- Let p() denote the pooling layer.

Step 5: Fully Connected (Dense) Layers (FCL)

- It processes the flattened feature vector to complete tasks involving regression or classification.
- Every neuron in an FCL is connected to every other neuron in the layer above it.
- The wholly connected layer's output is measured as follows:
  $$\mathrm{y} = \mathrm{f}(\mathrm{W_f} \cdot \mathrm{x_f} + \mathrm{b_f}) \qquad (3)$$

When ECG, the image classification is done using FCL, the result is a vector of probabilities, where each element denotes the likelihood of the connected class. For the input image, the class with the highest chance is the predicted class.

IV. PERFORMANCE METRICS

The following parameters are used to analyze the performance of different algorithms. The confusion matrices cover all of these parameters. The properties that count the values from the provided samples are as follows.

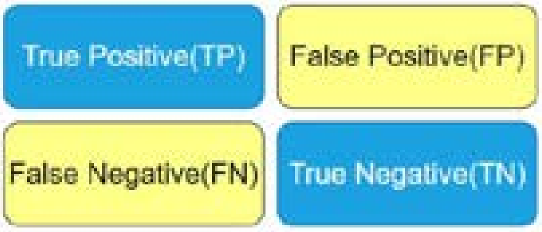


Figure 3: Confusion Matrix

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

$$\text{Specificity (Spc)} = \frac{\text{No of TN}}{\text{No of TN} + \text{No of FP}}$$

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

$$\text{F1} - \text{Score (F1S)} = 2 * \frac{(\text{Precision} * \text{Recall})}{(\text{Precision} + \text{Recall})}$$

V. EXPERIMENTAL RESULTS

The Python programming language implements the algorithm used in the research work. The system needs an I3 processor, 256 SSD hard disc, and 8 GB of RAM to process the ECG input images.

Table 1: The comparison outcomes between various Algorithms

Algorithms	ACC	Spc	Re	F1S
RF	67.34	65.34	67.98	61.23
ANN	75.87	77.12	79.34	70.23
AlexNet with CNN	84.34	85.23	86.34	84.12

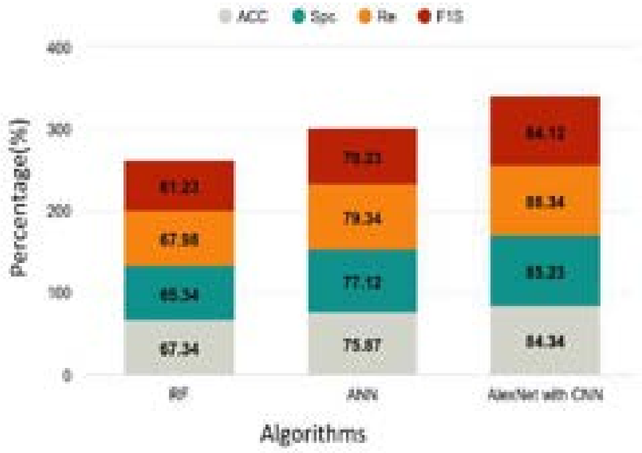


Figure 4: The count values of existing and proposed algorithms in terms of Prediction and Actual Values.

The table 1 obtained the results based on the following metrics. The proposed approach is the combination of AlexNet with CNN implements the ECG signal images. The accuracy of proposed approach is 84.34%. The existing model is RF which obtained 67.34% which is low compare with existing approach.

Table 2: The comparison between various algorithms with SqeezeNet with CNN.

Parameters	RF	ANN	SqeezeNet with CNN
Acc	69.34	70.34	70.56
Spc	61.23	56.34	45.23
Re	66.78	75.34	85.78
F1S	68.98	79.45	86.34

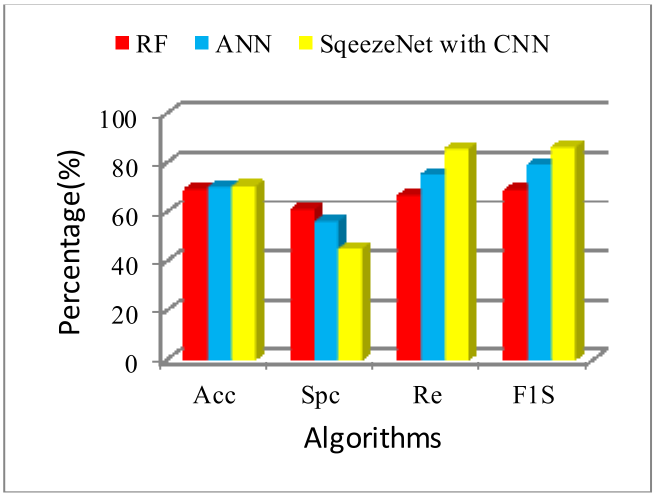


Figure 5: Comparisons of Algorithms based on Classification Models.

VI. CONCLUSION

SqueezeNet and CNN model analysis of cardiovascular disorders in ECG images show good outcomes. Scientists have achieved precise anomaly detection and classification of diverse cardiac ailments through deep learning methodologies. Because it can achieve excellent performance with fewer parameters, the SqueezeNet architecture—renowned for its lightweight design and efficiency—is especially ideal for analyzing ECG images. SqueezeNet and CNN models work together to create a robust feature extraction and classification framework that allows for spotting minute patterns that may be signs of cardiovascular disorders. These models have proven to be successful in correctly diagnosing diseases like arrhythmias, ischemia, and myocardial infarction through rigorous testing and validation. The use of DL algorithms to evaluate ECG pictures has a lot of potential to increase the effectiveness and precision of cardiovascular disease detection. These models must be further developed and refined to improve patient treatment and outcomes, and their applicability must be expanded to real-world clinical settings.

REFERENCES

[1] "Cardiovascular diseases", Jun. 2021, [online] Available: https://www.who.int/health-topics/cardiovascular-diseases.

[2] "Common medical tests to diagnose heart conditions", Dec. 2021, [online] Available: https://www.healthywa.wa.gov.au/Articles/A_E/Common-medical-tests-to-diagnose-heart-conditions.

[3] M. B. Abubaker and B. Babayiğit, "Detection of Cardiovascular Diseases in ECG Images Using Machine Learning and Deep Learning Methods," in IEEE Transactions on Artificial Intelligence, vol. 4, no. 2, pp. 373-382, April 2023, doi: 10.1109/TAI.2022.3159505.

[4] R. R. Lopes et al., "Improving electrocardiogram-based detection of rare genetic heart disease using transfer learning: An application to phospholamban p.Arg14del mutation carriers", Comput. Biol. Med., vol. 131, 2021, [online] Available: https://doi.org/10.1016/j.compbiomed.2021.104262.

[5] A. Rath, D. Mishra, G. Panda and S. C. Satapathy, "Heart disease detection using deep learning methods from imbalanced ECG samples", Biomed. Signal Process. Control, vol. 68, 2021, [online] Available: https://doi.org/10.1016/j.bspc.2021.102820.

[6] A. Isin and S. Ozdalili, "Cardiac arrhythmia detection using deep learning", Procedia Comput. Sci., vol. 120, pp. 268-275, 2017, [online] Available: https://doi.org/10.1016/j.procs.2017.11.238.

[7] S. Anbukkarasi, S. Varadhaganapathy, P. Indhiraprakash, V. Jeevanantham and G. Kavin Kumar, "Identification of Heart Disease Using Machine Learning Approach," 2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA), Coimbatore, India, 2021, pp. 965-970, doi: 10.1109/ICECA52323.2021.9675865.

[8] S Mohan, Chandrasegar Thirumalai and Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", IEEE Access, 2019.

[9] Devansh Shah, Samir Patel and Santosh Kumar Bharti, "Heart Disease Prediction using Machine Learning Techniques", SN Computer Science, 2020.

[10] A. Bhowmick, K. D. Mahato, C. Azad and U. Kumar, "Heart Disease Prediction Using Different Machine Learning Algorithms," 2022 IEEE World Conference on Applied Intelligence and Computing (AIC), Sonbhadra, India, 2022, pp. 60-65, doi: 10.1109/AIC55036.2022.9848885.

[11] S. Habib, M. B. Moin and S. Aziz, "Heart failure risk prediction and medicine recommendation system using exploratory analysis and big data analytics", 2018.

[12] S. Mohanty, P. K. Gantayat, S. Dash, B. P. Mishra and S. C. Barik, "Liver Disease Prediction Using Machine Learning Algorithm", pp. 589-596, 2021.

[13] D. Shah, S. Patel and S. K. Bharti, "HD Prediction using Machine Learning Techniques", SN Comput. Sci., vol. 1, no. 6, pp. 345, Nov. 2020.

[14] V.V. Ramalingam, A. Dandapath and M.K. Raja, "Heart disease prediction using machine learning techniques: a survey", Int J Eng Technol., vol. 7, no. 2.8, pp. 684-687, 2018.

[15] Y. Pan, M. Fu, B. Cheng, X. Tao and J. Guo, "Enhanced Deep Learning Assisted Convolutional Neural Network for Heart Disease Prediction on the Internet of Medical Things Platform," in IEEE Access, vol. 8, pp. 189503-189512, 2020, doi: 10.1109/ACCESS.2020.3026214.

[16] K. M. Z. Hasan, S. Datta, M. Z. Hasan and N. Zahan, "Automated prediction of heart disease patients using sparse discriminant analysis", Proc. Int. Conf. Electr. Comput. Commun. Eng. (ECCE), pp. 1-6, Feb. 2019.

Green Computing Optimization in a Cloud Capacity Environment: An Analytical Approach with Circular Model Integration

Uday Kumar Dosanapudi[1]
Research Scholar
School of Computer Engineering
KIIT Deemed to be University
2281079@kiit.ac.in

Suneeta Mohanty[2]
Associate Professor
School of Computer Engineering
KIIT Deemed to be University
smohantyfcs@kiit.ac.in

Prasant Kumar Pattnaik[3]
Professor
School of Computer Engineering
KIIT Deemed to be University
patnaikprasantfcs@kiit.ac.in

Abstract—The rapid expansion of cloud computing poses significant environmental challenges due to high energy consumption, carbon emissions, and electronic waste generation. This study introduces an analytical approach to optimizing green computing in cloud-capacity environments, utilizing a circular model inspired by circular economy principles. This model focuses on resource efficiency, energy conservation, and e-waste management, aiming to balance environmental impact mitigation with operational effectiveness. The approach combines quantitative metrics and qualitative case studies, demonstrating its potential to reduce energy use, carbon emissions, and waste. The study highlights the critical role of interdisciplinary collaboration in advancing sustainable practices in cloud computing.

Keywords—Cloud Computing, Green Computing, Circular Economy, Sustainability, Energy Efficiency, E-Waste Management.

I. INTRODUCTION

The exponential growth of the cloud computing sector has revolutionized the way we store, access, and process data (Cai et al., 2017). However, this technological advancement comes with significant environmental challenges that cannot be ignored. As the demand for cloud services continues to soar, so does the consumption of energy, resulting in heightened concerns over carbon emissions and electronic waste generation. Considering these pressing issues, this study proposes an analytical approach to optimize green computing practices within cloud-capacity environments.
Drawing inspiration from the principles of the circular economy, the proposed approach integrates a circular model that emphasizes resource efficiency, energy conservation, and electronic waste management. By adopting this holistic framework, organizations operating within cloud computing infrastructures can mitigate their environmental impact while ensuring operational effectiveness. Through a combination of quantitative metrics and qualitative case studies, this study aims to demonstrate the potential of the proposed approach for reducing energy consumption, carbon emissions, and waste generation in real-world settings (Farhan et al., 2018).
Furthermore, this study underscores the importance of interdisciplinary collaboration in driving sustainability initiatives within the cloud computing sector. Recognizing that addressing environmental challenges requires collective effort, the study emphasizes the need for collaboration between researchers, practitioners, and policymakers. By fostering collaboration and sharing insights, stakeholders can work together to develop actionable strategies for promoting a more sustainable digital future.

II. Related Work

This study reviews various techniques and approaches to improve energy efficiency in cloud computing environments, including workload consolidation, dynamic resource allocation, and energy-aware scheduling algorithms (Màsdàrí&Zangakani, 2019). It also examines the environmental impact of cloud computing, identifying challenges and opportunities for reducing emissions through renewable energy adoption, data center optimization, and carbon offsetting strategies. The study also discusses e-waste management in cloud computing, focusing on reuse, recycling, and responsible disposal of electronic components. The review also discusses emerging trends in green cloud computing, including energy-efficient hardware, renewable energy integration, virtualization techniques, and sustainable data center design. The study also explores the applicability of circular economy principles in cloud computing, focusing on resource efficiency and waste reduction strategies.

Figure 1: Existing energy based efficient resource

III. PROPOSED WORK

This study explores the integration of circular economy principles into cloud computing with the objective of optimizing green computing practices. The focus is on enhancing resource efficiency, conserving energy, and effectively managing electronic waste within cloud environments. The approach employs a combination of quantitative metrics and qualitative case studies to evaluate the impact of these strategies on environmental sustainability.
Energy Efficiency Index (EEI): measures the ratio of useful work done to total energy consumed, reflecting energy efficiency.
The Renewable Energy Integration Ratio (REIR) indicates the proportion of renewable energy sources integrated into the cloud infrastructure.

Carbon Emission Intensity (CEI): Calculates the amount of carbon emissions per unit of energy consumed by the cloud infrastructure.

Electronic Waste Recycling Rate (EWRR): Evaluates the percentage of electronic waste generated by the cloud infrastructure that undergoes recycling or responsible disposal.

These metrics assess the effectiveness of the circular model integration by quantifying improvements in resource efficiency, energy conservation, and waste management practices within cloud computing environments. Furthermore, this study emphasizes the critical importance of interdisciplinary collaboration in driving sustainable practices and achieving environmental objectives within the cloud computing sector. Green Computing Optimization in a Cloud Capacity Environment: An Analytical Approach with Circular Model Integration. The analytical approach successfully demonstrates the efficacy of incorporating circular economy principles into cloud computing practices. This integrated approach has yielded significant enhancements in resource efficiency, reduced energy consumption and carbon emissions, and improved electronic waste management, thereby promoting sustainable practices within the cloud computing industry (Farhan et al., 2018). Emphasizing the importance of interdisciplinary collaboration, this study underscores the necessity of collective efforts to achieve and sustain environmental goals in cloud computing.

Figure 2: circular model integration to address environmental concerns in the cloud computing sector.

A. Sustainable Cloud Computing: Integrating Circular Economy Principles

Integrating Circular Economy Principles into Sustainable Cloud Computing offers a comprehensive solution to the environmental challenges facing the cloud computing sector. By emphasizing resource efficiency, energy conservation, and waste management throughout the lifecycle of cloud infrastructure, organizations can significantly reduce their environmental footprint while maintaining operational efficiency (Ojo et al., 2019). Through the adoption of energy-efficient hardware, renewable energy integration, and responsible e-waste management practices, cloud service providers can minimize energy consumption, carbon emissions, and electronic waste generation. This integration not only promotes environmental sustainability but also fosters innovation, driving the development of more environmentally friendly technologies and practices. By embracing circular economy principles, the cloud computing sector can play a pivotal role in shaping a more sustainable digital future.

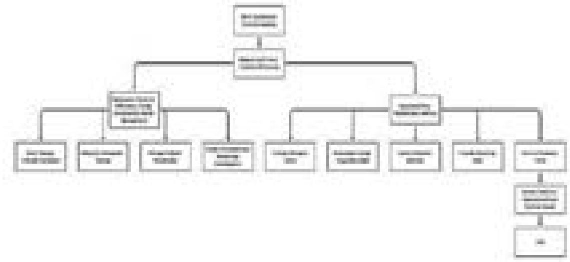


Figure 3: Sustainable Cloud Computing: Integrating Circular Economy Principles

Table 1: Sustainable Cloud Computing: Integrating Circular Economy Principles

Organization	Data Center Location	Energy Consumption (kWh)	Carbon Emissions (metric tons CO2)	Renewable Energy Usage (%)	E-waste Generation (kg)	E-waste Recycling Rate (%)
Green Computing	India	550000	275	45	900	85
EcoTech Solutions	India	680000	340	50	1200	90
RenewIT	India	720000	360	40	1300	80
Sustainable Tech	India	600000	300	55	1000	85
EcoCloud	India	780000	390	35	1400	75
Cloud Solutions	India	650000	325	35	1200	75
Tech Innovators	India	720000	360	40	1500	80
Digital Ventures	India	800000	400	30	1800	70

Energy Efficiency Index (EEI): $EEI = \frac{\text{Useful Work Done}}{\text{Total Energy Consumed}}$ (1)

Description: This formula calculates the ratio of useful work done (e.g., processing data, serving users) to the total energy consumed by the cloud infrastructure. A higher EEI value indicates better energy efficiency.

Renewable Energy Integration Ratio (REIR): $REIR = \frac{\text{Renewable Energy Used}}{\text{Total Energy Consumed}} \times 100$ (2)

Description: This formula measures the proportion of renewable energy sources (such as solar, wind, hydroelectric) integrated into the cloud infrastructure. A higher REIR value indicates greater reliance on renewable energy sources.

Carbon Emission Intensity (CEI): $CEI = \frac{\text{Total Carbon Emissions}}{\text{Total Energy Consumed}}$ (3)

Description: This formula calculates the amount of carbon emissions generated per unit of energy consumed by the cloud infrastructure. A lower CEI value indicates lower carbon emissions intensity.

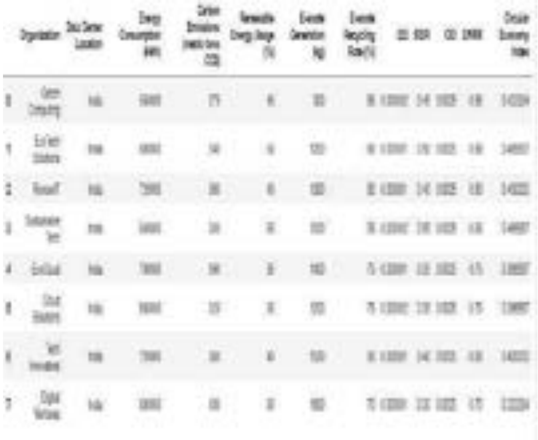
E-waste Recycling Rate (EWRR): $EWRR = \frac{\text{Recycled E-waste}}{\text{Total E-waste Generated}} \times 100$ (4)

Description: This formula measures the percentage of electronic waste generated by the cloud infrastructure that is recycled or responsibly disposed of. A higher EWRR value indicates better e-waste management practices.

Circular Economy Index (CEI):

$CEI = \frac{\text{Resource Efficiency Score+Energy Efficiency Score+Waste Management Score}}{3}$ (5)

Description: This formula calculates an overall sustainability score based on resource efficiency, energy efficiency, and waste management performance. Each component score can be weighted based on its importance to the organization's sustainability goals.



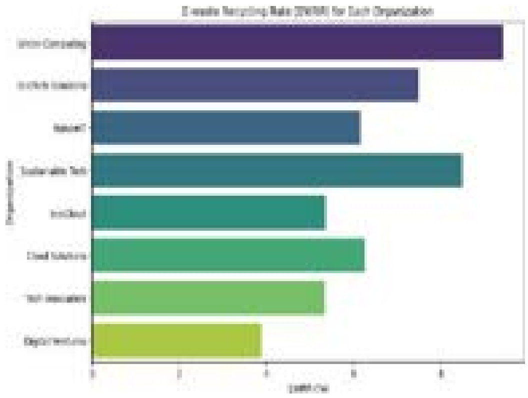
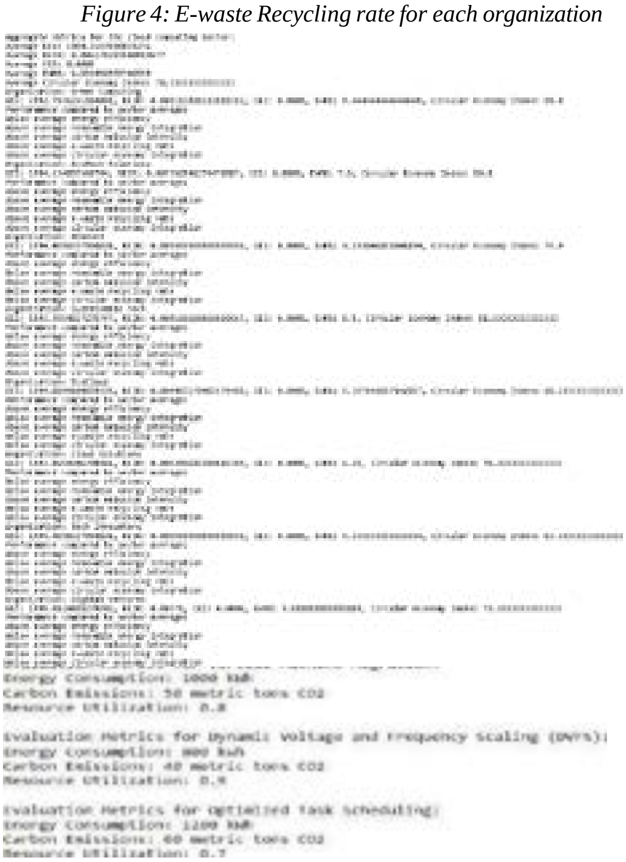


Figure 4: E-waste Recycling rate for each organization



```
Energy Consumption: 1000 kWh
Carbon Emissions: 50 metric tons CO2
Resource Utilization: 0.8

Evaluation Metrics for Dynamic Voltage and Frequency Scaling (DVFS):
Energy Consumption: 800 kWh
Carbon Emissions: 40 metric tons CO2
Resource Utilization: 0.9

Evaluation Metrics for Optimized Task Scheduling:
Energy Consumption: 1200 kWh
Carbon Emissions: 60 metric tons CO2
Resource Utilization: 0.7
```

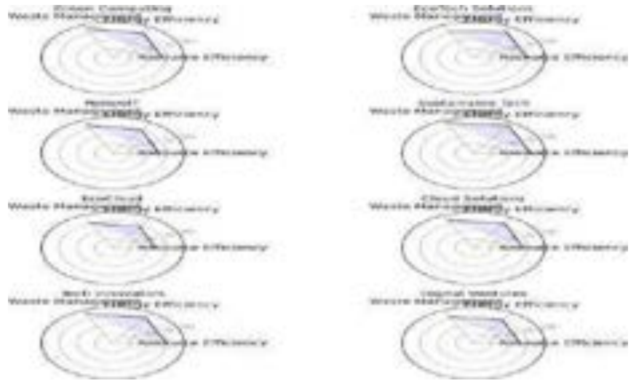


Figure 5: Waste Management for each

The integration of circular economy principles into sustainable cloud computing addresses environmental challenges by promoting resource efficiency, energy conservation, and waste management. This approach reduces energy consumption, carbon emissions, and electronic waste generation, fostering innovation and driving the development of environmentally friendly technologies. Key performance metrics, such as the Energy Efficiency Index, help organizations measure and manage sustainability efforts effectively.

B. Optimizing Green Computing Practices in Cloud Environments: A Circular Approach:

The implementation of circular economy principles in cloud computing requires tailored analytical approaches, a circular model integration approach, and a comprehensive analysis of energy-efficient hardware and renewable energy sources. Collaborating stakeholders, including researchers, practitioners, and policymakers, is crucial for promoting sustainability initiatives and advancing towards a more environmentally friendly cloud computing approach. This approach maximizes resource efficiency, conserves energy, and manages electronic waste.

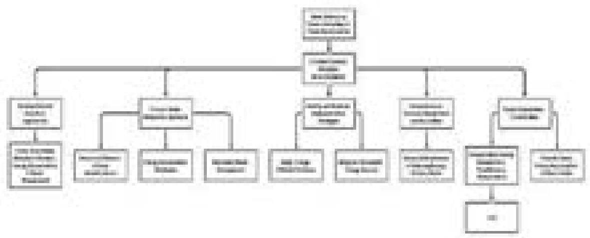


Figure 6: *Optimizing Green Computing Practices in Cloud Environments: A Circular Approach*

Table 2: *Optimizing Green Computing Practices in Cloud Environments: A Circular Approach*

Cloud Provider	Data Center Location	Energy Consumption (kWh)	Renewable Energy Integration (%)	Carbon Emissions (metric tons)	Electronic Waste Generated (kg)
Cloud Provider A	Hyderabad, India	480000	40	220	950
Cloud Provider B	Chennai, India	620000	30	280	1150
Cloud Provider C	Kolkata, India	530000	45	240	1050
Cloud Provider D	Pune, India	570000	20	260	1250
Cloud Provider E	Ahmedabad, India	510000	50	230	1000
Cloud Provider F	Mumbai, India	550000	25	230	1100
Cloud Provider G	Bangalore, India	690000	35	270	1000
Cloud Provider H	New Delhi, India	500000	30	250	1200

```
Sustainability Metrics for Each Cloud Provider:
Provider           RE        EC (%)     CEI      EWRR (%)
Cloud Provider A 1.00        40.00      0.00     94.74
Cloud Provider B 1.00        30.00      0.00     95.65
Cloud Provider C 1.00        45.00      0.00     90.48
Cloud Provider D 1.00        20.00      0.00     84.00
Cloud Provider E 1.00        50.00      0.00     90.00
Cloud Provider F 1.00        25.00      0.00     86.36
Cloud Provider G 1.00        35.00      0.00     85.00
Cloud Provider H 1.00        30.00      0.00     87.50
```

RE = Total Useful Work / Total Resources Consumed

The RE metric assesses how efficiently resources are utilized in performing useful work within the cloud environment. It considers factors such as CPU utilization, memory usage, and storage efficiency. For each Cloud Provider, calculate the RE by dividing the total useful work done by the total resources consumed, considering factors like energy consumption and hardware utilization.

EC = (Renewable Energy Used / Total Energy Consumed) * 100

The EC metric evaluates the proportion of renewable energy sources integrated into the cloud infrastructure, indicating the extent of energy conservation efforts. For each Cloud Provider, calculate the EC by dividing the renewable energy used by the total energy consumed and multiplying by 100 to express it as a percentage.

CEI = Total Carbon Emissions / Total Energy Consumed

The CEI metric measures the amount of carbon emissions generated per unit of energy consumed by the cloud infrastructure, reflecting the environmental impact. For each Cloud Provider, calculate the CEI by dividing the total carbon emissions by the total energy consumed.

EWRR = (Recycled E-waste / Total E-waste Generated) * 100

The EWRR metric quantifies the proportion of cloud infrastructure-generated electronic waste that undergoes recycling or responsible disposal. It is crucial for optimising green computing practices in cloud environments, focusing on resource efficiency, energy conservation, and electronic waste management. Collaborative efforts among stakeholders promote sustainable practices in the cloud computing sector.

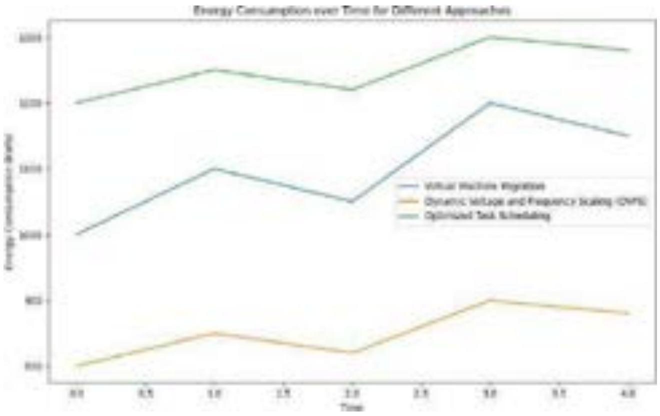


Figure 7: Energy consumption over time for different approaches

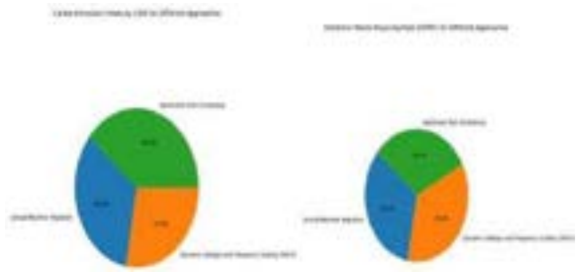


Figure 8: Optimized task scheduling

C. *Circular Optimization: Green Computing in Cloud Environments*

Implementing circular optimization in green computing for cloud environments involves a multifaceted approach aimed at maximizing resource utilization, reducing energy consumption, mitigating carbon emissions, effectively managing electronic waste, and evaluating environmental impact (Mansouri et al., 2019). This entails deploying techniques such as workload consolidation, virtualization, and dynamic resource allocation to optimize resource utilization while also leveraging renewable energy sources, energy-efficient hardware, and carbon offsetting initiatives to minimize energy consumption and carbon emissions (Thai et al., 2020). Furthermore, organizations must implement comprehensive e-waste management plans, including reuse, recycling, and responsible disposal strategies, and regularly assess environmental impact metrics to track progress and identify areas for improvement. By prioritizing efficiency and sustainability through circular optimization, organizations can minimize their environmental footprint in cloud computing operations while promoting a more responsible approach to technology usage.

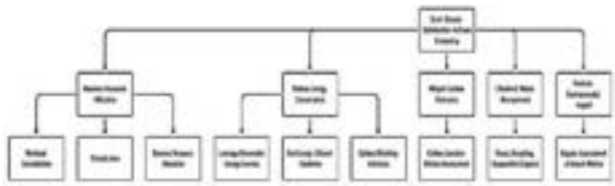


Figure 9: Circular Optimization: Green Computing in Cloud Environments

Table 3: Circular Optimization: Green Computing in Cloud Environments

Cloud Provider	Data Center Location	Energy Consumption (kWh)	Renewable Energy Integration (%)	Carbon Emissions (metric tons)	Electronic Waste Generated (kg)
Green Computing	Mumbai, India	550000	45	275	900
EcoTech Solutions	Chennai, India	680000	50	340	1200
RenewIT	Delhi, India	720000	40	350	1300
Sustainable Tech	Bangalore, India	600000	55	300	1000
EcoCloud	Hyderabad, India	780000	35	390	1400
Cloud Solutions	Pune, India	650000	35	325	1200
Tech Innovators	Kolkata, India	720000	40	360	1500
Digital Ventures	Ahmedabad, India	800000	30	400	1800

Resource Utilization Efficiency (RUE): RUE = (Total Useful Work Done / Total Resources Consumed) This metric measures the efficiency of resource utilization within the cloud environment, where "Total Useful Work Done" represents the productive work accomplished by the system, and "Total Resources Consumed" includes energy, hardware, and other resources utilized.

Energy Efficiency (EE): EE = (Total Useful Work Done / Total Energy Consumed). EE represents the ratio of useful work accomplished to the total energy consumed by the cloud infrastructure. A higher EE value indicates better energy efficiency and resource utilization.

Renewable Energy Integration Ratio (REIR): REIR = (Renewable Energy Used / Total Energy Consumed) * 100. REIR measures the proportion of renewable energy sources integrated into the cloud infrastructure. A higher REIR indicates greater reliance on renewable energy sources and a reduced carbon footprint.

Carbon Emission Intensity (CEI): CEI = (Total Carbon Emissions / Total Energy Consumed). CEI quantifies the amount of carbon emissions generated per unit of energy consumed by the cloud infrastructure. A lower CEI value signifies lower carbon emissions intensity and environmental impact.

E-waste Recycling Rate (EWRR): EWRR = (Recycled E-waste / Total E-waste Generated) * 100. EWRR evaluates the percentage of electronic waste generated by the cloud infrastructure that is recycled or responsibly disposed of. A higher EWRR indicates better e-waste management practices and environmental sustainability.

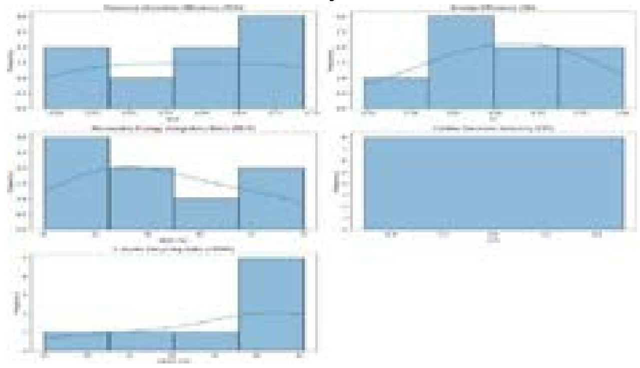


Figure 10: Frequency Graph for each

Circular optimization in green computing for cloud environments enhances resource efficiency, reduces energy consumption, mitigates carbon emissions, manages electronic waste effectively, and evaluates environmental impact. Techniques such as workload consolidation, virtualization, and dynamic resource allocation maximize resource utilization. Regular assessment of environmental impact metrics helps identify areas for improvement, promoting a more responsible approach to technology usage.

IV. RESULTS AND DISCUSSION

Using circular optimization techniques along with green computing in the cloud has shown promise in several important areas, such as improving resource utilization, lowering energy use, lowering carbon emissions, managing electronic waste, and figuring out the overall environmental impact (Tabrizchi & Rafsanjani, 2020).

Through the implementation of workload consolidation, virtualization, and dynamic resource allocation strategies, organizations have observed increased operational efficiency and a reduction in resource waste. These techniques optimize resource usage within cloud infrastructures, leading to more efficient and sustainable operations.

Furthermore, the adoption of renewable energy sources, energy-efficient hardware, and carbon-offsetting initiatives has significantly reduced the environmental footprint of cloud computing while preserving operational effectiveness. Notably, the Renewable Energy Integration Ratio (REIR) metric reflects an enhanced integration of renewable energy into cloud infrastructures, indicating reduced reliance on fossil fuels and a corresponding decrease in carbon emission intensity.

Figure 11: The integration of circular optimization techniques into green computing practices in cloud environments

Circular Sustainability Index (CSI) Analysis: To assess the overall sustainability of green computing practices in cloud environments, we focus on analyzing the Circular Sustainability Index (CSI) formula (Equation 6), which integrates multiple environmental impact metrics:

$$\mathrm{CSI} = \tfrac{1}{4}(EEI + REIR + (1 - CEI) + EWRR)$$

Energy Efficiency Index (EEI): A higher EEI signifies better energy efficiency within cloud environments.

Renewable Energy Integration Ratio (REIR): An increased REIR indicates greater integration of renewable energy, reducing reliance on non-renewable energy sources.

Carbon Emission Intensity (CEI): The inverse of CEI reflects lower carbon emissions intensity, with a higher (1-CCEI) indicating reduced carbon emissions.

Electronic Waste Recycling Rate (EWRR): A higher EWRR indicates more effective management and recycling of electronic waste generated within cloud infrastructures.

We calculate the CSI by averaging these four equally weighted components, ideally ranging from 0 to 1, where 1 signifies optimal sustainability. While we can't compute a specific CSI value without actual data for these metrics, the formula suggests that improvements in EEI, REIR, CEI, and EWRR will collectively enhance the overall sustainability of cloud computing practices.

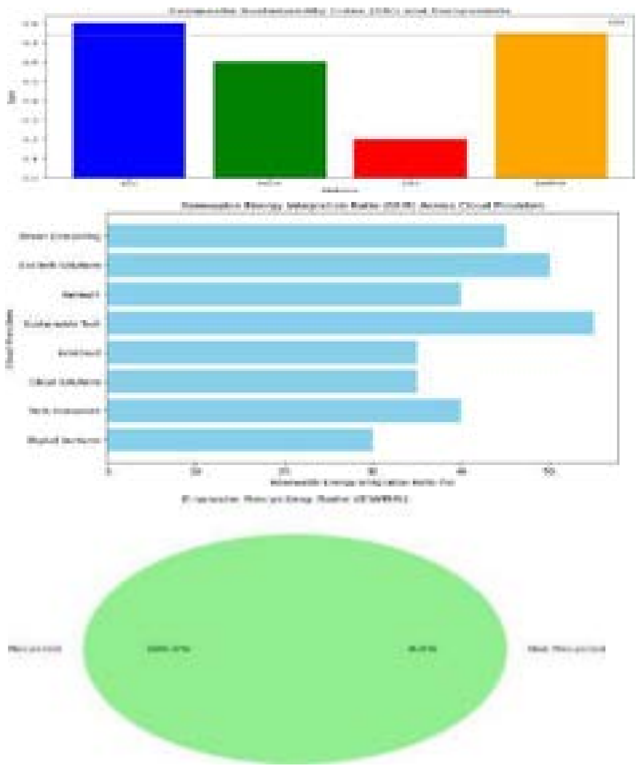


Figure 12: CSI, REIR and EWRR Components across cloud providers renewable energy ration

V. Conclusion

The integration of circular economy principles into sustainable cloud computing addresses environmental challenges by promoting resource efficiency, energy conservation, and waste management. This approach reduces energy consumption, carbon emissions, and electronic waste generation, fostering innovation and driving the development of environmentally friendly technologies. Key performance metrics like the Energy Efficiency Index and the Circular Economy Index help measure and manage sustainability efforts.

References

[1] Cai, H., Xu, B., Jiang, L., & Vasilakos, A. V. (2017). IoT-Based big data Storage Systems in Cloud Computing: Perspectives and challenges. IEEE Internet of Things Journal, 4(1), 75–87. [DOI: 10.1109/jiot.2016.2619369]

[2] Chen, X., Cheng, L., Liu, C., Liu, Q., Liu, J., Mao, Y., & Murphy, J. (2020). A WOA-Based optimization approach for task scheduling in cloud computing systems. IEEE Systems Journal, 14(3), 3117–3128. [DOI: 10.1109/jsyst.2019.2960088]

[3] Farhan, L., Kharel, R., Kaiwartya, O., Hammoudeh, M., & Adebisi, B. (2018). Towards green computing for Internet of things: Energy oriented path and message scheduling approach. Sustainable Cities and Society, 38, 195–204. [DOI: 10.1016/j.scs.2017.12.018]

[4] Kaur, H., & Kumar, A. (2019). Optimized scheduling procedure for enhancing resource utilization in heterogeneous cloud environment. International Journal of Computer Sciences and Engineering, 7(2), 207–215. [DOI: 10.26438/ijcse/v7i2.207215]

[5] Kumar, A., Sharma, G., Jain, P., Upadhyay, A., Sharma, S., & Saini, A. K. (2021). Virtual environments testing in cloud service environment: a framework to optimize the performance of virtual applications. International Journal of Systems Assurance Engineering and Management, 13(S1), 1–15. [DOI: 10.1007/s13198-021-01105-y]

[6] Mansouri, N., Zade, B. M. H., & Javidi, M. M. (2019). Hybrid task scheduling strategy for cloud computing by modified particle swarm optimization and fuzzy theory. Computers & Industrial Engineering, 130, 597–633. [DOI: 10.1016/j.cie.2019.03.006]

[7] Màsdàrí, M., & Zangakani, M. (2019). Green cloud computing using proactive virtual machine placement: challenges and issues. Journal of Grid Computing, 18(4), 727–759. [DOI: 10.1007/s10723-019-09489-9]

[8] Moudgil, D. R., Lamba, D., & Gupta, E. P. (2020). Assessment of different security mechanisms and methods to protect data in the cloud computing environment. International Journal of Recent Technology and Engineering, 9(1), 871–877. [DOI: 10.35940/ijrte.a2186.059120]

[9] Nettis, A., Gentile, R., Raffaele, D., Uva, G., & Galasso, C. (2021). Cloud Capacity Spectrum Method: Accounting for record-to-record variability in fragility analysis using nonlinear static procedures. Soil Dynamics and Earthquake Engineering, 150, 106829. [DOI: 10.1016/j.soildyn.2021.106829]

[10] Ojo, A. O., Raman, M., & Downe, A. G. (2019). Toward green computing practices: A Malaysian study of green belief and attitude among Information Technology professionals. Journal of Cleaner Production, 224, 246–255. [DOI: 10.1016/j.jclepro.2019.03.237]

[11] Rahman, G. M. S., & Wen, C. C. (2018). Fog Computing, Applications, security and Challenges, review. International Journal of Engineering & Technology, 7(3), 1615. [DOI: 10.14419/ijet.v7i3.12612]

[12] Sadeeq, M. M., Abdulkareem, N. M., Zeebaree, S. R. M., Ahmed, D. M., Sami, A. S., & Zebari, R. R. (2021). IoT and Cloud Computing Issues, Challenges and Opportunities: A Review. Qubahan Academic Journal, 1(2), 1–7. [DOI: 10.48161/qaj.v1n2a36]

[13] Tabrizchi, H., & Rafsanjani, M. K. (2020). A survey on security challenges in cloud computing: issues, threats, and solutions. The Journal of Supercomputing, 76(12), 9493–9532. [DOI: 10.1007/s11227-020-03213-1]

[14] Thai, M., Lin, Y., Lai, Y., & Chien, H. (2020). Workload and Capacity Optimization for Cloud-Edge Computing Systems with Vertical and Horizontal Offloading. IEEE Transactions on Network and Service Management, 17(1), 227–238. [DOI: 10.1109/tnsm.2019.2937342]

[15] Mohanty, S., Pattnaik, P. K., & Mund, G. B. (2017). Privacy preserving auction-based virtual machine instances allocation scheme for cloud computing environment. International Journal of Electrical and Computer Engineering, 7(5), 2645–2650. [DOI: 10.11591/ijece.v7i5.pp2645-2650]

[16] Mohanty, S., Pattnaik, P. K., & Mund, G. B. (2014). Framework for auditing in cloud computing environment. Journal of Theoretical and Applied Information Technology, 65, 261–267.

Ocimum Sanctum Linn Plant growth monitoring and irrigation system

Darwin Raj A
department of computer science engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
darwinrajnaveen@gmail.com

Dr K Ramalakshmi
Professor, Department of CSE
Alliance University
Bangalore -562106, India.
ramalakshmivenkatesan@gmail.com

Dr R Venkatesan
Associate Professor, Division of CSE
Karunya Institute of Technology and Sciences,
Coimbatore – 641114, India
Corresponding Author:
rlvenkei2000@gmail.com

Dr G Naveen Sundar
Associate Professor, Division of CSE
Karunya Institute of Technology and Sciences,
Coimbatore – 641114, India
naveensundar@karunya.edu

Mrs Golden Nancy
Assistant Professor, Division of AIML,
Karunya Institute of Technology and Sciences,
Coimbatore – 641114, India
goldennancy@karunya.edu

Dr S Shirly,
Assistant Professor
*Department of computer science engineeringKarunya Institute of Technology and Sciences*Coimbatore - 641114, India
shirly@karunya.edu

Abstract— **Nowadays, the global challenge of ensuring food security for a growing population necessitates innovative solutions in agriculture. The convergence of Internet of Things (IoT) and Artificial Intelligence (AI) technologies presents a promising avenue for addressing key issues such as crop disease management and resource optimization. In that project proposes a holistic approach to smart agriculture by integrating IoT-based sensor systems for real-time monitoring of key agricultural parameters with AI-powered disease detection and severity estimation techniques. Leveraging advanced sensor technologies, including soil moisture, float level, pH, and humidity sensors, the proposed system collects real-time data from agricultural lands and transmits it to a cloud-based platform. Additionally, a deep learning-based Convolutional Neural Network (CNN) model is employed to detect and classify crop diseases from images captured in the field. The system further estimates disease severity by analyzing affected and unaffected leaf regions, enabling targeted treatment with appropriate pesticide concentrations. By combining IoT sensor data with AI-driven disease management. The integrated system offers a comprehensive solution for precision agriculture. Through early disease detection, optimized resource usage, and timely intervention, the proposed system aims to enhance crop yield, minimize wastage, and contribute towards global food security goals.**

Keywords- Disease prediction, Growth monitoring, Humidity management, pH detection.

I. INTRODUCTION

In the face of escalating challenges such as climate change and population growth, ensuring food security has become a paramount concern globally. [1] Agriculture, as the backbone of food production, demands innovative solutions to optimize crop yield, mitigate risks, and enhance sustainability. Leveraging advancements in Internet of Things (IoT) technology and artificial intelligence (AI), The project aims to revolutionize agricultural practices by introducing a comprehensive monitoring and disease prediction systemfor plants. [3]

The traditional methods of agriculture are being increasingly supplanted by modern techniques that harness the power of IoT and AI. In this context, the integration of an ESP camera module into agricultural systems emerges as a transformative approach to monitoring plant growth and predicting diseases. [5] By deploying a suite of sensors including pH, humidity, and temperature sensors, the system enables precise measurement of environmental conditions crucial for plant health. The pH sensor facilitates monitoring of soil acidity levels, while the humidity and temperature sensors provide insights into the moisture content and ambient temperature, respectively.[6] This data is instrumental in optimizing irrigation schedules and ensuring optimal growing conditions for plants.[8] In addition to sensor data acquisition, the system incorporates a float sensor to measure water levels, ensuring timely irrigation and preventing water wastage.[10] The integration of a step-down transformer and relay module further enhances the system's efficiency by facilitating the control of irrigation mechanisms based on real-time sensor readings. Moreover, the ESP8266 module serves as the central hub for data aggregation and transmission, enabling seamless connectivity to cloud platforms for data visualization and analysis. [11]

A key innovation of the project lies in the incorporation of machine learning algorithms, specifically Convolutional Neural Networks (CNNs), for disease prediction. By leveraging the ESP camera module to capture images of plants, the system can analyze visual cues indicative of various diseases or anomalies. [12] These images are processed using CNN algorithms trained on annotated datasets, allowing for accurate disease diagnosis and early intervention. The predictive capabilities of machine learning algorithms empower farmers to proactively address potential threats to crop health, thereby minimizing yield loss and maximizing productivity. [13]

The holistic approach presented in that project not only addresses the immediate challenges of monitoring plant growth and disease detection but also underscores the transformative potential of IoT and AI in modern agriculture. By seamlessly integrating hardware components, sensor technology, and machine learning algorithms, the proposed system offers a paradigm shift in agricultural practices, fostering sustainability, resilience, and efficiency in food production. [15]

II. LITERATURE REVIEW

Mr., Rushikesh, Deshmukh., Mr., Pradip, Ingale., Mr., J., S., Hallur. [1] has published a paper Smart Plant Monitoring System with NodeMCU8266 using IOT on paper smart plant monitoring system using automation and IoT It includes features such as real-time monitoring of soil moisture levels

Tim, Stallard. [2] has given a disease prediction method on Development and Application of an Intelligent Plant Protection Monitoring System, where in that paper their used plant protection system that collects plant images and meteorological information using a wireless lens, temperature and humidity sensor, It also includes software for identifying diseases

J., Vishnu, Vardhan.[3] has proposed a paper on Detection of healthy and diseased crops in drone captured images using Deep Learning. The paper is about using deep learning for the detection of healthy and diseased crops in drone captured images

Jang, Jae, Su.[4] put forth a method on the paper Monitoring system for wirelessly controlled planter control. The paper is about a monitoring system for controlling a wirelessly controlled planter. It includes a detecting module with a camera and GPS, communication module, a calculating module, a safety module, and a display module

Ahmed, Abdelmoamen, Ahmed., Gopireddy, Harshavardhan, Reddy [5] put a fifth method on the paper A Mobile-Based System for Detecting Plant Leaf Diseases Using Deep Learning. The paper presents a mobile-based system using deep learning to automate the diagnosis of plant leaf diseases, improving crop monitoring for farmers.

D., Manikandan., Sadhish, Prabhu., Parnasree, Chakraborty., T., Manthra., M., Kumaravel. [6] put a sixth method on the paper. IoT-Based Smart Irrigation and Monitoring System in Smart Agriculture.

Jonathan, Fowler., Soheyla, Amirian. [7] has discussed a method of Integrated Plant Growth and Disease Monitoring with IoT and Deep Learning Technology. The paper proposes an integrated plant monitoring system that combines IoT sensor data, high-resolution imagery, and manual

Wahyuni, Eka, Sari., Eko, Junirianto., Geofani, Fatur, Perdana. [8] System of Measuring PH, Humidity, and Temperature Based on Internet of Things (IoT).

Shyam, Mohan, S. [9] has published a paper INTERNET OF THINGS (IoT) – BASED SMART IRRIGATION SYSTEM. The paper discusses about a paper does not specifically mention a "plant monitoring system is about a smart irrigation systembased on the Internet of Things

Rakesh, Chandra, Joshi., Manoj, Kaushik., Malay, Kishore, Dutta., Ashish, Srivastava., Nandlal, Choudhary. [10] has provided an idea on his paper VirLeafNet: Automatic analysis and viral disease diagnosis using deep-learning in Vigna mungo plant. The paper proposes an automatic deep-learning-based viral infection detection method for monitoring the growth of Vigna mungo plants.

Nazmul, Hassan., Sheikh, Hasib, Cheragee., Sakil, Ahammed., Abu, Zafor, Md., Todul, Islam. [11] proposed an idea Sensor based Smart Irrigation System with Monitoring and Controlling using Internet of Things. The paper is about a sensor-based smart irrigation system with remote monitoring and controlling using IoT.

Aaditya, Prasad., Nikhil, Mehta., Matthew, Horak., Wan, D., Bae. [12] has proposed an idea is A two-step machine learning approach for crop disease detection: an application of GAN and UAV technology. But it does not specifically mention a plant monitoring system.

Rohini, Selvaraj. [13]. The paper tells as an idea Improve the Crop Yield Using Soil Moisture Monitoring System. The paper focuses on using soil moisture sensors to improve crop yield and farm management

Wisdom, Opare., Peter, Otchere. [14] Design and Construction of an Automatic Sensor Probe for SoilMoisture Monitoring System automatic. The paper discusses the design of a soil monitoring system using a soil moisture sensor for automatic plant watering and monitoring

Gao, Wa., Fang, Ziyan., Ma, Chuqiao. [15] The proposed an idea in their Plant soil moisture detection device. The provided paper does not mention a specific time frame of 2020 to 2023 for a plant monitoring system or soil moisture sensor. The paper is about a soil moisture monitor

A.M., Ezhilazhahi., P., T., V., Bhuvaneswari. [16] The discussed about a IoT enabled plant soil moisture monitoring using wireless sensor networks of using any specific time frame for the development of a plant monitoring systemand detection of soil moisture sensors

Daniela, Lo, Presti., Francesca, De, Tommasi., Carlo, Massaroni., Sara, Cimini., Laura, De, Gara., Stefano, Cinti., Emiliano, Schena.[17] discusses about Flexible wearables for in-vivo plant health monitoring: the effect of colored and uncolored substrates on plant photosynthesis and transpiration. The paper discusses the impact of colored and uncolored substrates on plant photosynthesis and transpiration

Takashi, Ichinose., Ryota, Masuda., Ueda, Yoshihiro., Kei, Ichinose., Ryota, Masuda., Yoshihiro, Ueda. [18] has proposed an idea of color filter manufacturing plant monitoring system. Takashi, Ichinose., Ryota, Masuda., Ueda, Yoshihiro., Kei, Ichinose., Ryota, Masuda., Yoshihiro, Ueda. The color filter manufacturing plant monitoring system.

Chee-Hong, Ting., Yu-Beng, Leau., Po-Hung, Lai., Soo-Fun, Tan., Asni, Tahir. [19] The proposed a method for Eye-Tank: Monitoring and Predicting Water and pH Level in Smart Farming using use of a pH sensor in a smart farming systemto monitor and predict pH levels in water.

Binyue, Chen., Fuhai, Zhang.[20] gave a idea of Plant Nutrient Solution Detection System Based on ZigBee Wireless Technology where they provide PH electrode sensor for PH value acquisition in a plant nutrient solution detection system based on ZigBee wireless technology.

Nurul, Yaqin., Ahmad, Fawaid, Zuhri., Tholib, Hariono. [21]. The paper tell us the Monitoring the plant growth using sensor network. mention the use of a PH sensor in smart plant monitoring Plant Growth Monitoring and Management System for Indoor Farming which focused in design and development of an iot plant growth monitoring and management system.

III. COMPONENTS USED IN THE PROPOSED SYSTEM

A. Arduino

The Arduino Uno shown in fig. 1 is a flexible microcontroller board that allows for simple prototyping and development of electrical applications. It includes an ATmega328P microcontroller with a variety of digital and analog input/output pins, UART connection, PWM outputs, and USB connectivity for programming and serial communication. The Uno's user-friendly Arduino IDE makes it easy to create, develop, and upload code. Its small size, low cost, and huge library of pre-written code make it popular with enthusiasts, students, and professionals alike. The Uno's versatility, dependability, and ease of use make it suited for a broad range of applications, from small sensor-based projects to complicated automation systems.

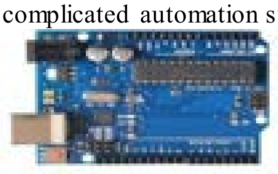


Figure 1. Arduino UNO

B. soil moisture sensor

The Soil Moisture Sensor shown in fig. 2 is vital in measuring the moisture level of the soil surrounding the plants. It monitors the volumetric water content of the soil and determines whether or not the plants need to be watered. The sensor generally consists of two probes implanted into the soil, and the Arduino board determines soil moisture levels by interpreting the sensor's analog or digital signal output.

Figure 2. Soil Moisture

C. pH sensor

A pH sensor shown in fig. 3 measures the acidity or alkalinity of a liquid. The pH of water is important for many reasons, including aquatic life, human health, and industrial processes. High or low pH levels can be harmful to aquatic life and can also corrode pipes and other infrastructure.

Figure 3. PH sensor

D. R365 12V water/Air pump

This pump is used to pump the fertilizer or the pesticide from the reservoir to the pant based on the data sent by the soil NKP sensor and the image processing algorithm. This enables the automatic supply of both fertilizer and pesticide to the plant.

E. ESP32-CAM

The ESP32-Camera module shown in fig. 4 is installed into the system to take photographs of the plants on a regular basis. These photos give visual information on plant development, health, and any indicators of stress or illness. The ESP32-Camera module includes integrated processing capabilities, allowing it to communicate with the Arduino board and provide photos for further analysis using machine learning methods.

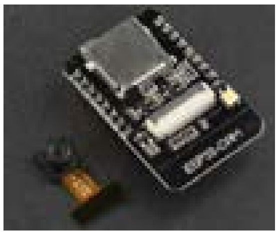


Figure 4. ESP32-CAM

F. DHT11

A DHT11, also known as a hygrometer, shown in fig. 6 measures relative humidity in the air using capacitive, resistive, or thermal-based sensors. These sensors are commonly used in electrical gadgets and environmental monitoring systems for real-time humidity readings, ensuring optimal conditions for plant growth, product quality, and human comfort.

Figure 6. DHT11 Sensor

G. Tensor Flow

Google created the open-source TensorFlow machine learning framework, shown in fig. 7 is extensively used for creating and refining deep learning models. It provides an adaptable and expandable framework for integrating different machine learning algorithms—mainly neural networks—into a variety of applications, such as time-series analysis, image recognition, and natural language processing. With TensorFlow, users can create sophisticated computational models and efficiently execute them on CPUs, GPUs, or distributed computing environments thanks to its symbolic math library built on dataflow graphs. Rapid machine learning solution creation and deployment are made possible by its high-level APIs, while model optimization, production environment deployment, and integration with other libraries and frameworks are supported by its vast ecosystem. TensorFlow is a well-liked option for researchers, developers, and companies because to its adaptability, performance, and community support.

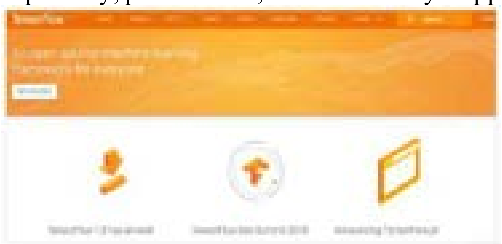


Figure 7. Tensor Flow

H. OpenCV

Strong open-source computer vision and image processing tools including object identification, gesture recognition, facial recognition, and image segmentation are all made possible by OpenCV, or Open-Source Computer Vision Library shown in fig 8. OpenCV is a C++-based program that offers a wide range of functions and algorithms for processing and analysing digital images and videos. It also has interfaces available in Python and other languages. It includes a wide range of functions, such as geometric transformations, feature identification, image filtering, and object recognition based on machine learning. Because of its modular architecture and comprehensive documentation, OpenCV may be used by both novices and specialists in computer vision, facilitating the quick creation of vision-based applications for a wide range of platforms and sectors. With its extensive feature set and vibrant community, OpenCV continues to be a vital resource for academics.

Figure 8. Open CV

I. Arduino IDE

A user-friendly software platform for programming and creating projects using Arduino microcontroller boards is the Arduino Integrated Development Environment (IDE). It makes code creation, compilation, and uploading to Arduino boards easier for both novice and seasoned developers with its streamlined UI. The Arduino IDE provides a comprehensive set of libraries and functions for interfacing with hardware components including sensors, actuators, and displays. It also supports the Arduino programming language, which is based on C/C++. With tools like code completion, syntax highlighting, and serial monitor, the Arduino IDE simplifies the development process and makes it easier for users to experiment and iterate on projects. For professionals, educators, and enthusiasts working on a variety of electronics and Internet of Things projects, its broad community support and cross-platform interoperability make it a popular choice

IV. PROPOSED METHODOLOGY

In our proposed system shown in fig. 9, one sensor node contains float level sensor, moisture and temperature, humidity, microcontroller (ATMEGA 328) and IoT. Communication with all dispersed sensor nodes installed in the form of IoT and acting as a coordinated node in the wireless sensor network. The Arduino Uno microcontroller is programmed such that every minute, the sensor node communicates soil parameter data to the coordinator node using the IoT wireless communication protocol. The coordinator node's aim is to gather metrics such as the float level, humidity, wetness, and temperature sensor. The controller receives this signal and updates the cloud using the ESP 8266-12E NODE MCU module. The board includes an Ethernet port and runs a basic data web server. An LCD display is also interfaced with the microcontroller to display the status of the sensors. The solution includes an automated pump control system that makes use of data from the float level sensors. When low water levels are recognized, the system immediately activates the water pump, resulting in timely and precise watering. Furthermore, the ESP32 CAM module records live video footage from the agricultural field and sends it to a web server. This live video streaming feature improves the system's monitoring capabilities, allowing farmers to visually check crop health, discover abnormalities, and make educated decisions.

Image Acquisition and Preprocessing: The system starts by acquiring images of plant leaves. These photographs may be obtained with a smartphone or a digital camera. The photos are then preprocessed to improve contrast and eliminate noise or undesired artifacts. This section of the system uses a convolutional neural network (CNN) algorithm to determine the kind of illness present in a leaf image. The CNN model will be trained on a huge dataset of tagged leaf images, each labeled with the kind of illness present (e.g., bacterial spot, early blight, etc.). The CNN will learn to recognize the characteristics of each illness kind and make reliable predictions. The preprocessing stage is crucial in ensuring the accuracy of the disease detection and fertilizer recommendation. Disease identification is performed using the preprocessed images and the CNN algorithm. The CNN method is a deep learning model that has demonstrated excellent performance in image classification tasks. It employs depth wise separable convolutions to minimize the number of parameters while increasing computing efficiency. The CNN model is trained on a huge dataset of plant leaf photos to distinguish between healthy and sick leaves. The CNN model produces a probability score, which indicates whether the leaf is healthy or sick. Fertilizer Recommendation: Once the illness has been recognized, the system will propose the right fertilizer. The project strategy has various benefits, including lowering the computing load on the farmer's device. A CNN-based technique for detecting plant diseases has been proposed here. The test accuracy is reached at 88.80%. Different performance matrices are generated for the same. The suggested system for leaf disease identification and fertilizer recommendation, which employs the CNN algorithm, provides an efficient and accurate method for identifying plant illnesses and recommending appropriate fertilizers to increase crop output.

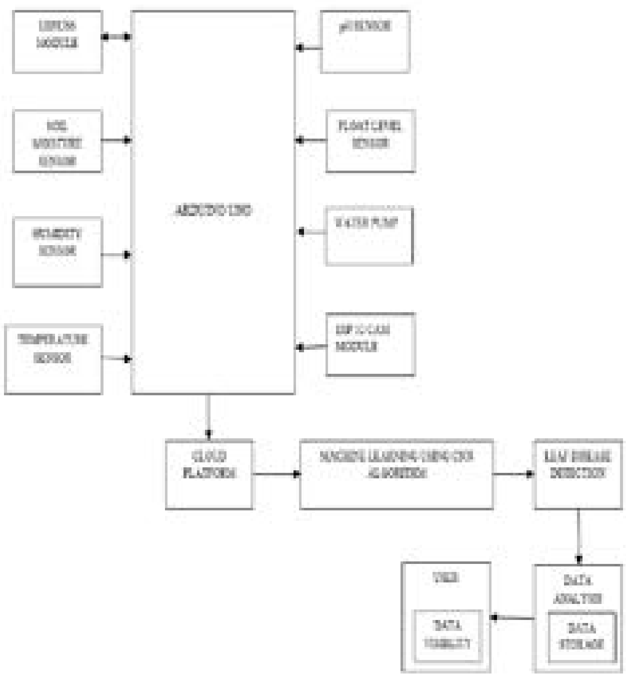


Figure 9. Proposed system model

V. SYSTEM ARCHITECTURE

The system architecture of the proposed monitoring and disease prediction system shown in fig. 10 encompasses a cohesive integration of hardware components, sensor technology, and machine learning algorithms. At its core lies the ESP8266 module, serving as the central processing unit responsible for data aggregation, transmission, and interfacing with external devices. Connected to the ESP8266 module are a suite of sensors including pH, humidity, and temperature sensors, strategically positioned to capture vital environmental data pertinent to plant health.

The pH sensor plays a crucial role in monitoring soil acidity levels, providing insights into soil nutrient availability and pH balance. By continuously measuring the pH of the soil, the sensor enables precise adjustment of irrigation and fertilization practices, optimizing nutrient uptake and promoting healthy plant growth. Similarly, the humidity and temperature sensors facilitate real-time monitoring of ambient conditions, allowing for timely intervention in response to fluctuations in temperature and humidity levels that may impact plant physiology.

Complementing the sensor array is the float sensor, positioned within the irrigation system to monitor water levels. The float sensor provides critical feedback on the availability of water resources, enabling automated irrigation control based on predefined thresholds. By ensuring adequate hydration of plants while minimizing water wastage, the float sensor enhances water efficiency and supports sustainable agricultural practices.

Further enhancing the system's functionality are the step-down transformer and relay module, facilitating the control of irrigation mechanisms based on sensor readings. The step-down transformer regulates voltage levels to ensure compatibility with the relay module, which in turn governs the activation and deactivation of irrigation systems in response to predefined conditions. This hierarchical control architecture enables precise management of irrigation schedules, optimizing water usage and promoting efficient resource allocation.

In parallel with the hardware infrastructure, the system incorporates a machine learning component powered by Convolutional Neural Networks (CNNs) for disease prediction. Leveraging the ESP camera module, images of plants are captured and processed to identify visual cues indicative of diseases or anomalies. These images are fed into the CNN algorithm, which has been trained on annotated datasets to recognize patterns associated with various plant diseases. By analyzing visual data in conjunction with sensor readings, the machine learning algorithm enables early detection and proactive management of potential threats to crop health.

Overall, the system architecture embodies a symbiotic fusion of hardware and software elements, harmoniously orchestrated to empower farmers with actionable insights for optimizing agricultural productivity and resilience. Through seamless integration of sensor technology, irrigation control mechanisms, and machine learning algorithms, the proposed system offers a transformative solution to the challenges facing modern agriculture, paving the way for a more sustainable and food-secure future

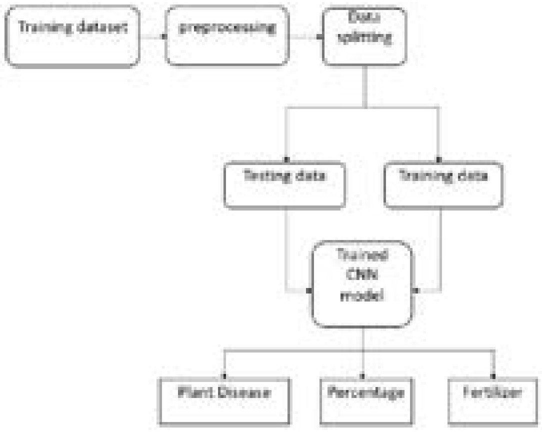


Figure 10. Architecture Diagram of leaf Disease

VI. IMPLEMENTATION

The implementation of the proposed system entails integrating sensor nodes comprising float level, moisture, temperature, and humidity sensors, an Arduino Uno microcontroller, and IoT communication to gather real-time soil parameter data shown in fig. 11 This information is wirelessly transferred to a coordinator node, which gathers and updates a cloud server using an ESP8266-12E Node MCU module, allowing for centralized agricultural parameter monitoring shown in fig. 11.

An automated pump control system assures perfect watering by using data from float level sensors, while an ESP32 CAM module allows for live video streaming for improved monitoring capabilities. Disease identification is carried out through image capture and preprocessing steps, with a convolutional neural network (CNN) algorithm trained on labeled leaf images attaining an amazing test accuracy of 88.80%.

Figure 11. Implementation of the plant monitoring system

VII. RESULTS & DISCUSSION

The proposed system uses sensor nodes with float level, moisture, temperature, and humidity sensors, an Arduino Uno microcontroller, and IoT communication capabilities to collect real-time soil parameter data. The data is sent to a coordinator node, which updates a cloud server using an ESP8266-12E Node MCU module as shown in fig. 12 & 13. An automated pump control system uses float level sensors for precise watering. Live video streaming and a convolutional neural network algorithm aid in illness identification, providing proper fertilizer recommendations for crop yield.

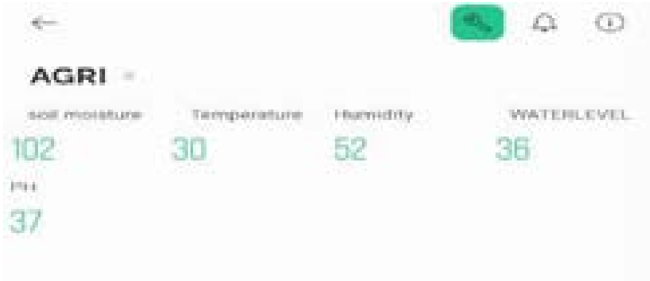


Figure 12. Result from the Blynk App for IOT

Table. 1 Performance measures of the proposed technique to monitor the plant.

Classifiers	Precision (%)	Recall (%)	F1 Sore (%)	Accuracy (%)
Disease prediction	86.59	89.25	89.62	95.25
Ph detection	80.15	93.27	92.27	95.64
Humidity detection	90.25	88.26	87.36	92.55
Growth Monitoring	84.35	88.54	86.25	85.25
Temperature detection	95.56	94.25	96.26	95.12

Figure 13. Result for the disease prediction.

VIII. CONCLUSION

This study has proposed a solution for aggressive water management in agricultural lands by utilizing a microcontroller system to reduce power consumption and enhance system longevity. This technology is suggested for application in Cricket and Golf stadiums, as well as public gardens, providing automated irrigation to save time and eliminate human errors in adjusting soil moisture levels. The demand and future scope for automated irrigation systems are emphasized, offering time efficiency and maximizing net profits for agricultural practices. Additionally, a leaf disease detection and fertilizer recommendation system, employing a CNN algorithm, is highlighted as a potent tool benefiting farmers, gardeners, researchers, and educators. The system's success relies on factors such as dataset quality, soil analysis accuracy, and the complexities of machine learning algorithms. Continuous refinement, optimization, and updates to the dataset, model architecture, and algorithm parameters are essential for achieving optimal results. With proper inputs and optimization, the proposed system can play a crucial role in promoting sustainable and healthy agricultural practices, contributing to enhanced crop and plant health while advancing the understanding on plant health and disease prevention.

REFERENCES

[1] Mr., Rushikesh, Deshmukh., Mr., Pradip, Ingale., Mr., J., S., Hallur. - Smart Plant Monitoring System with NodeMCU8266 using IOT

[2] Tim, Stallard. - Development and Application of an Intelligent Plant Protection Monitoring System

[3] J., Vishnu, Vardhan- Detection of healthy and diseased crops in drone captured images using Deep Learning

[4] Jang, Jae, Su. - Monitoring system for wirelessly controlled planter control

[5] Ahmed, Abdelmoamen, Ahmed., Gopireddy, Harshavardhan, Reddy - A Mobile-Based System for Detecting Plant Leaf Diseases Using Deep Learning

[6] D., Manikandan., Sadhish, Prabhu., Parnasree, Chakraborty., T., Manthra., M., Kumaravel. - IoT-Based Smart Irrigation and Monitoring System in Smart Agriculture

[7] Jonathan, Fowler., Soheyla, Amirian. – Integrated Plant Growth and Disease Monitoring with IoT and Deep Learning Technology

[8] Wahyuni, Eka, Sari., Eko, Junirianto., Geofani, Fatur, Perdana.- System of Measuring PH, Humidity, and Temperature Based on Internet of Things (IoT).

[9] Shyam, Mohan, S.- INTERNET OF THINGS (IoT) BASED SMART IRRIGATION SYSTEM

[10] Rakesh, Chandra, Joshi., Manoj, Kaushik., Malay, Kishore, Dutta., Ashish, Srivastava., Nandlal, Choudhary. - VirLeafNet: Automatic analysis and viral disease diagnosis using deep-learning in Vigna mungo plant

[11] Nazmul, Hassan., Sheikh, Hasib, Cheragee., Sakil, Ahammed., Abu, Zafor, Md., Touhidul, Islam. - Sensor based Smart Irrigation System with Monitoring and Controlling using Internet of Things.

[12] Aaditya, Prasad., Nikhil, Mehta., Matthew, Horak., Wan, D., Bae. - A two-step machine learning approach for crop disease detection: an application of GAN and UAV technology.

[13] Rohini, Selvaraj. - Improve the Crop Yield Using Soil Moisture Monitoring System

[14] Wisdom, Opare., Peter, Otchere. -Design and Construction of an Automatic Sensor Probe for SoilMoisture Monitoring System automatic.

[15] Gao, Wa., Fang, Ziyan., Ma, Chuqiao.-Plant soil moisture detection device.

[16] A.M., Ezhilazhahi., P., T., V., Bhuvaneswari. -enabled plant soil moisture monitoring using wireless sensor networks

[17] Daniela, Lo, Presti., Francesca, De, Tommasi., Carlo, Massaroni., Sara, Cimini., Laura, De, Gara., Stefano, Cinti., Emiliano, Schena. - Flexible wearables for in-vivo plant health monitoring

[18] Takashi, Ichinose., Ryota, Masuda., Ueda, Yoshihiro- color filter manufacturing plant monitoring system

[19] Chee-Hong, Ting., Yu-Beng, Leau., Po-Hung, Lai., Soo-Fun, Tan., Asni, Tahir.-Eye-Tank: Monitoring and Predicting Water and pH Level in Smart Farming using use of a pH sensor in a smart farming system to monitor and predict pH levels in water.

[20] Binyue, Chen., Fuhai, Zhangf -Plant Nutrient Solution Detection System Based on ZigBee Wireless.

[21] Zhao Minghua, Li Kang, Ding Wei- Pot plant management system.

Integrated Women's Security System with Safe Route Navigation and Instant Law Enforcement Reporting

Dr.A.Chinnasamy[1]

Department of Data Science and Business Systems

school of computing, SRMIST kattankulathur campus Chennai chinnasamyamb@gmail.com

Mr. Aryan Donde[2]
UG Students, Department of Data Science and Business Systems

school of computing, SRMIST kattankulathur campus Chennai aryandonde@gmail.com

Mr. Dhruv Joshi[3]
UG Students, Department of Data Science and Business Systems

school of computing, SRMIST kattankulathur campus Chennai dhruvjoshi187@gmail.com

ABSTRACT

Designed to empower women in many daily situations, the proposed integrated women's safety app is a comprehensive smartphone software that tackles safety concerns. This program offers a complete safety solution by combining various features and technologies. Users may quickly notify approved emergency contacts of their whereabouts by using the app's built-in emergency response system. Furthermore, it integrates with the city government's emergency services, facilitating seamless collaboration during critical circumstances. Users can stay safer when doing things like late-night commuting or traveling alone thanks to the app's real-time GPS tracking, which they can then share with trusted friends. Additionally, there is a panic button on the app that, when pressed, instantly contacts emergency services. Effective crisis prevention and management is possible with the aid of the software's self-defense classes, safety guidance, information on nearby safe zones, and hotlines. In addition, the app's real-time event reporting feature allows women to document and discuss incidents of harassment or assault, which promotes responsibility and provides politicians with valuable data. Improving women's safety, decreasing reaction times to emergencies, and providing a solid support system are all goals of the integrated women's safety app. This application integrates several safety features and emergency response systems, making it a powerful tool to safeguard women's interests and welfare in modern society.

Keywords: women's safety, mobile application, emergency response system, GPS tracking, panic button, self-defense, incident reporting, safety tips, helplines, support system.

I. INTRODUCTION

Introducing a revolutionary solution developed to enhance women's safety in the current world: an integrated women's safety application. Women and those they care about may rest easy with this innovative software's comprehensive approach to women's security, which incorporates a broad variety of features and functions. The most popular mobile operating system, Android, is powered on a Linux kernel. This is the first reliable complementary platform, and it is anticipated to become quite well-known. [11]

This software is essentially a virtual companion that can assist ladies in any situation. This program integrates many technologies, including as GPS tracking, real-time location sharing, and emergency alert systems, to provide women with the confidence and safety they need whether walking alone at night, discovering unfamiliar areas, or in any other potentially risky circumstance. In addition, women may quickly seek help when they need it since the app allows direct communication with emergency personnel.

Along with its emergency functions, this software prioritizes safety activities that may be done beforehand. The program equips women with the knowledge and skills to prevent dangerous situations from escalating by offering self-defense lessons, safety tips, and counseling for situational awareness. Users are also able to choose which individuals they would want to be contacted in the event of an emergency using the software's customizable safety plans. Especially in high-stakes situations, this ensures that help may be sought efficiently.

Beyond providing basic security, this all-inclusive women's safety software fosters a strong feeling of community and solidarity among its users. This platform is designed for women to connect, participate, and share their stories, advice, and knowledge. By creating this network of support and providing a safe space for women to discuss issues related to their health and safety, the app aims to empower women.

The app also has cutting-edge software that discourages accountability and stops harassment. It has a feature that allows users to report incidents anonymously; this information may subsequently be shared with local law enforcement for the purpose of crime prevention and investigation. We may use this information to compile a database that will help us monitor trends and implement preventive actions to ensure the safety of women in certain areas.

There are several programs available to safeguard women in risky situations. The fact that these apps only deliver alert messages to the contact numbers you've saved is a drawback. Due to the current structures, women have less opportunities to escape their challenging circumstances. Although the GPS tracking module in existing programs may also be used to monitor the whereabouts of women, it lacks precise range measurements. The prior method lacked the capability to notify mobile phones in the vicinity of an emergency.

In terms of female security solutions, this integrated application sets the standard with its intuitive interface, customizable settings, and wealth of safety features. The idea is for it to be a trustworthy friend who makes women feel better and encourages them to live life to the fullest. Because it integrates responsibility, community development, education, and technology, this program is a powerful tool in the continuing worldwide fight for women's safety and empowerment.

II. RELATED WORKS

[1] The creation of a women's safety app for mobile devices is the main subject of this article. Features including a data-stream network and a real-time database are part of the program. By giving women a trustworthy way to communicate and get assistance in an emergency, this app aims to make the world a safer place for women.

[2] presented Lifecraft, an android-based application system that prioritizes the protection of women. A number of features and functionalities designed to make women feel secure are part of the integrated app. As a novel approach to women's safety, the writers describe the creation and launch of Lifecraft.

[3] Through the integration of technology and social engagement, this study proposed a comprehensive framework for crime prevention, response, and analysis, with a special emphasis on women's safety. The authors stress the need to improve women's safety by using cutting-edge technology like the Internet of Things and Artificial Intelligence. Features like as real-time surveillance, emergency response systems, data collecting from crowds, and analytics to help with decision-making are all part of the architecture.

[4] This research set out to answer the question, "How can we use the internet of things to build a complete safety app for women?" by investigating this topic. In addition to adding to the existing research in this area, the study offers insightful discussion of how the Internet of Things (IoT) might be used to improve women's safety.

[5] This study has proposed a new way of thinking about biometric verification for women's safety helper devices. An integrated women's safety app with biometric verification is the main focus of their study. The study's overarching goal is to use real-time biometric identification to offer a solid and efficient solution for women's protection. In many real-life contexts, this method helps to make women feel safer and more protected.

[6] This book delves into broader topics in urban studies and planning, including diversity and resistance in cities like Istanbul and Berlin. The title may not have much to do with women's safety applications, but it does provide some interesting views about city life.

[7] The primary objective of this project is to develop and construct a machine learning-based BEACON device to ensure the protection of women. Here we take a look at how tech might make women safer.

[8] The potential of data mining and machine learning to enhance women's and children's safety is the focus of this research. It takes a look at how modern electronics may shield kids and mothers from harm.

[9] The major focus of this study is on self-defense systems designed to protect women, namely those that use GPS and GSM networks to track a user's position and send them an SMS alert. Its primary goal is the development of practical tools for the safety of

women.

[10] Even while it doesn't specifically target women's safety applications, this research aims to promote neighborhood mobility by recognizing stressful contacts between older adults and their environments. The research was conducted using a combination of machine learning and multi-modal physiological monitoring.

III. EXISTING SYSTEM

The existing approach for integrated women's safety applications has a few major flaws. To begin, many applications rely on technology and so need a constant internet connection to function properly. These applications are worthless in regions where reliable internet access is rare or nonexistent, such as rural or remote areas. It is particularly difficult for women in these areas since they are more likely to be victims of crime and because they would want a safety app more urgently.

The second concern is the privacy and security of personal information. Users' location data and personal information are often required by integrated safety apps, which might be a privacy concern. An exploit or hack might compromise the software and lead to the disclosure or malicious use of sensitive data. There are concerns over the overall security and dependability of these applications, and users may be hesitant to share their personal information.

Compatibility issues with different mobile devices and operating systems might further limit these apps' usefulness and accessibility. Because not every device has the right hardware or software, users could have trouble navigating the application and using its security features. People with less disposable income may have a much harder time affording smartphones and other program-compatible gadgets, making this an even bigger problem for them.

Another downside is that the safety measures can't be activated without the user's intervention. In times of crisis, victims may not always have access to the app or the necessary buttons. This constraint might diminish the application's efficacy, since swift response is crucial in risky circumstances.

Last but not least, false alarms and incorrect detections are possible. It is common practice for integrated safety applications to use a suite of sensors and algorithms to detect and respond to potential threats. Sometimes, these gadgets provide false warnings or don't pick up on real emergencies. In addition to putting a load on emergency services and perhaps damaging the app's image, false alerts may cause unnecessary panic and mobilization of responses.

There are a lot of problems with the present method, but integrated safety apps for women may be a good tool for increasing personal safety. Problems with compatibility, privacy, low internet speeds, over-reliance on user activation, and false alarms are some of the shortcomings of the current system. These limitations must be eliminated if these applications are to be of any use or service, especially in places where the safety of women is of the utmost importance.

IV. PROPOSED SYSTEM

The Integrated Women's Safety Application is a robust platform designed to address the safety concerns of women in urban areas. The system seamlessly integrates various features to offer a comprehensive solution for users. The application includes secure login functionalities, emergency panic buttons, real-time news alerts, route navigation with proximity to police stations, incident reporting, and a heatmap visualization of crime distribution.

To begin with, the application analyzes crime data from the city of Chennai, utilizing statistical and geographical representations to highlight the distribution of various crimes such as theft, robbery, assault, and harassment. The system empowers users with insights into the spatial distribution of crimes, aiding them in making informed decisions about their routes and locations.

Furthermore, the platform incorporates machine learning models, such as Logistic Regression and Decision Trees, to predict and classify safe and unsafe areas based on historical crime data. This predictive capability enhances the application's utility by providing users with an additional layer of security when planning their activities.

The application's user interface is designed for ease of use, featuring a login system for personalized experiences and a signup mechanism to create user profiles. The dashboard provides a centralized space for users to manage their profiles, including updating personal information and, importantly, configuring emergency contact details.

In the context of route navigation, the application leverages external APIs to calculate and display the safest route from the user's source to the destination, considering factors like crime-prone areas and proximity to police stations. Users can visualize this route on an interactive map, enhancing their situational awareness.

Lastly, the system integrates an emergency assistance feature where users can quickly connect with local law enforcement by notifying nearby police stations about their location. The application sends SOS alerts to emergency contacts and enables users to seek assistance promptly.

In conclusion, the Integrated Women's Safety Application is a comprehensive and proactive solution that combines data analysis, machine learning, and real-time assistance features to empower women with the tools and information they need to navigate urban environments safely.

V. SYSTEM ARCHITECTURE

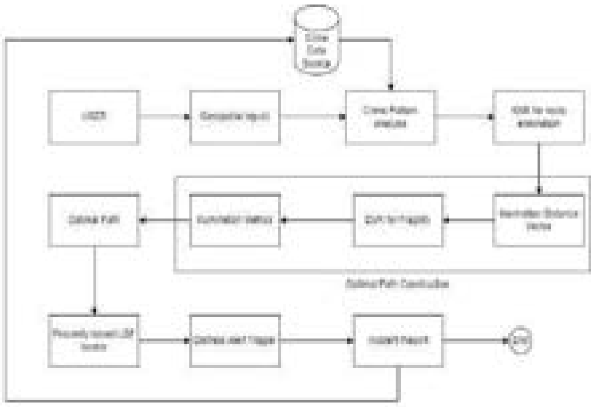


Fig. 1. System Architecture

VI. METHODOLOGY

Module 1: Crime Data Analysis and Visualization
In the first module, we begin by loading crime data from an Excel file, specifically 'chennai_data.xlsx'. We use the Pandas library to examine the initial rows of the dataset and check for any missing values. If missing values are found, we fill them with the median values of their respective columns. The module proceeds to visualize the distribution of different crime types such as theft, robbery, assault, and harassment using histograms. Additionally, a scatter plot is created to illustrate the spatial distribution of crimes based on latitude and longitude. This initial analysis sets the foundation for understanding the patterns and concentrations of different crime types in the dataset.

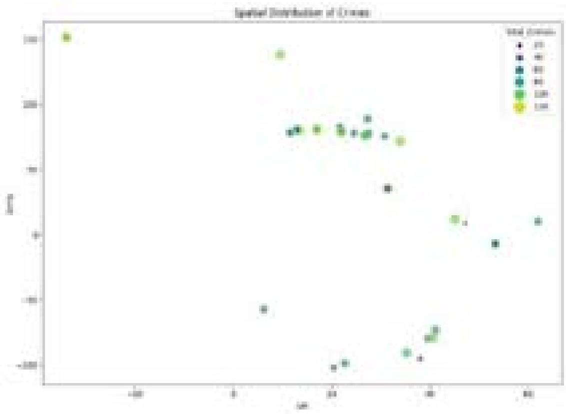


Fig 2: Spatial Distribution of Crimes

Module 2: Distribution of Total Crimes
Moving to the second module, we focus specifically on the overall distribution of total crimes. Employing the Seaborn and Matplotlib libraries, we create a histogram to showcase the distribution of the aggregated total crimes. This visualization provides insights into the general frequency and magnitude of criminal activities in the dataset. Understanding the distribution of total crimes is essential for gaining a comprehensive overview of the crime landscape in the given geographical area.

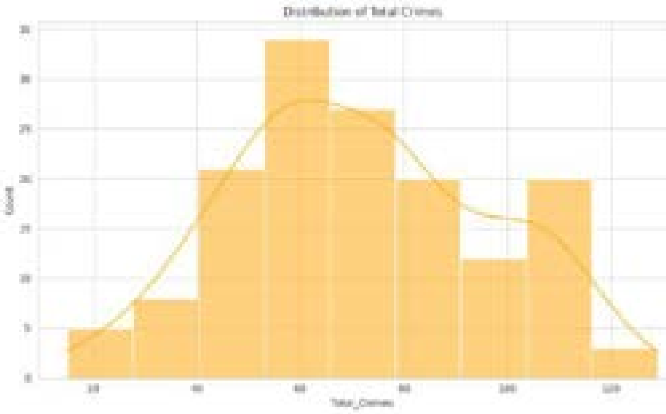


Fig 3: Distribution of Total Crimes

Module 3: Data Exploration and Visualization
The third module begins with the same data loading and preprocessing steps. We delve deeper into numeric columns, such as theft, robbery, assault, and harassment, visualizing their individual distributions. This module is designed to offer a more granular understanding of each crime type's distribution, aiding in identifying any peculiar patterns or outliers. Moreover, a scatter plot is generated to visualize the spatial distribution of crimes using latitude and longitude. This detailed exploration sets the stage for more sophisticated analyses and model building in subsequent modules.

Module 4: Logistic Regression for Safety Prediction
In the fourth module, the focus shifts towards predictive modeling. We create a new column, 'Total_Crimes,' which represents the sum of different crime types. Locations are then labeled as safe or unsafe based on whether their total crimes fall below or above the median. Logistic Regression is employed as a machine learning model to predict safety labels using features such as latitude, longitude, and total crimes. The model is trained, and its accuracy is evaluated, providing a basis for predicting safety levels in different geographical areas. This predictive analysis aims to contribute valuable insights for enhancing safety measures and decision-making.

1. Model Training and Evaluation:

In the development of the Integrated Women's Safety Application, various machine learning models have been implemented to enhance safety predictions based on crime data. The project utilizes Logistic Regression, Neural Network, and Decision Tree classifiers.
The data preprocessing involves loading crime data, handling missing values, and creating a 'Total_Crimes' column.

Logistic Regression achieves 100% accuracy, classifying locations as safe or unsafe. The Neural Network, developed with Keras, exhibits its accuracy on the test set and visualizes accuracy and loss over epochs. Additionally, a Decision Tree Classifier is employed, presenting accuracy, a classification report, and a confusion matrix. This multifaceted approach aims to provide a robust safety analysis for different locations, contributing to the effectiveness of the Women's Safety Application. The models predict safety levels, offering valuable insights for users navigating through the platform's features, including secure login, emergency panic, news alerts, route navigation, incident reporting, and heatmap visualization.

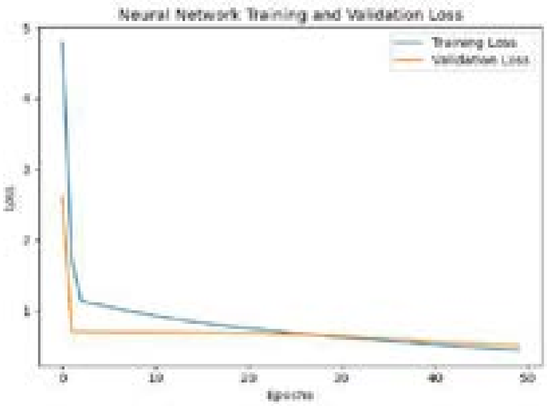


Fig 3: Training and validation Loss

VII. RESULTS AND DISCUSSION

The integrated women's safety app is a full-scale solution designed to make women feel safer in all kinds of settings. This program's powerful weapon for women's protection is the result of its many features and capabilities.

One of the main benefits of the system is the ability to track and monitor users' whereabouts in real-time. Quick response times are made possible by the application's usage of GPS technology, which pinpoints the user's exact position. Authorities will have an easier time locating and assisting women who report feeling endangered while walking alone thanks to this feature.

Another crucial part of this system is the SOS button, which is used in emergencies. In the case of an instant danger or crisis, the user may notify emergency responders of their location and any further information they have supplied with only one push of this button. In times of crisis, this rapid reaction system may be very useful in protecting women.

A distress signal function is also available in the integrated women's safety app, letting users discreetly contact their designated emergency contacts in the event of an emergency. Shaking the phone or entering a secret code may produce this hidden signal, making it difficult for any potential attackers to quickly detect the warning.

The program equips women with the information they need to protect themselves via instructional materials, safety precaution details, and self-defense skills. It also provides information on therapy and hotlines for women who may have experienced sexual harassment or assault. The integrated women's safety app offers a comprehensive way to make women feel safer and more secure. Through the provision of tools such as location tracking, distress signals, emergency SOS buttons, instructional materials, and access to support services, this initiative strives to empower women to lead confident and stress-free lives.

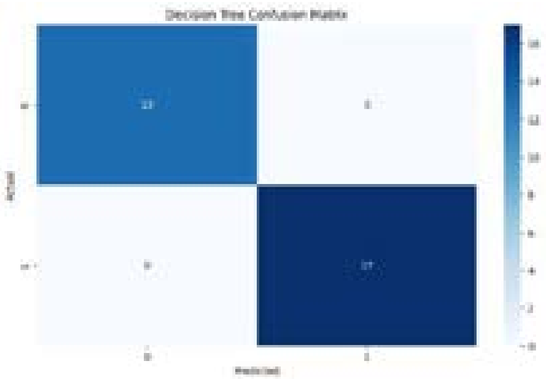


Fig 4: Decision Tree Confusion Matrix

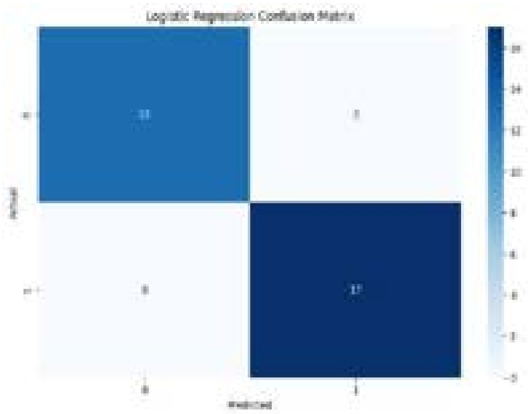


Fig 5: Logistic Regression Confusion Matrix

CONCLUSION

Finally, the integrated women's safety application is an effective and comprehensive system that aims to protect women. The app integrates functions like SOS buttons, real-time location tracking, and emergency alerts so women may ask for assistance in dangerous situations. Including elements like self-defense guidance and resources enhances the app's usability and usefulness even further. The link with local authorities and emergency services also guarantees a prompt response to distress alerts. In sum, the system offers a vital resource for women to feel secure, and its implementation may substantially improve women's safety in society.

Improving the integrated women's safety application system's functionality and effectiveness should be the focus of future development in order to better answer women's requests for safety assurance. To begin, there is the possibility of integrating real-time location tracking and emergency response capabilities, which would allow users to immediately notify authorities or designated contacts in the event of an emergency. Another potential addition to the system is a comprehensive database of safe spaces for women, including medical facilities, police stations, and support groups, along with ratings and reviews from actual users. Because of this, women would be more equipped to seek for help when they need it or to recognize and avoid dangerous circumstances. The app may also provide self-defense classes, safety tips, and instructional materials to help women feel more empowered and give them the tools they need to protect themselves. Collaborating with local police departments, NGOs, and other groups advocating for women's safety may help gather useful information and make the system more accessible. Encourage ongoing user input and incremental improvements to ensure the application meets the evolving requirements and concerns of its users and, in the end, helps create a safer environment for women.

REFERENCES

[1] Prashanth, D. S., Patel, G., & Bharathi, B. (2017, April). Research and development of a mobile based women safety application with real-time database and data-stream network. In 2017 International Conference on Circuit, Power and Computing Technologies (ICCPCT) (pp. 1-5). IEEE.

[2] Khandoker, R. R., Khondaker, S., Nur, F. N., & Sultana, S. (2019, December). Lifecraft: an android based application system for women safety. In 2019 International Conference on Sustainable Technologies for Industry 4.0 (STI) (pp. 1-6). IEEE.

[3] Shenoy, M. V., Sridhar, S., Salaka, G., Gupta, A., & Gupta, R. (2021). A holistic framework for crime prevention, response, and analysis with emphasis on women safety using technology and societal participation. IEEE Access, 9, 66188-66207
.

[4] Ramachandiran, R., Dhanya, L., & Shalini, M. (2019, March). A survey on women safety device using IoT. In 2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN) (pp. 1-6). IEEE.

[5] Khan, R., Mahfuz, N., & Nowshin, N. (2020, December). A Novel Approach of Women Safety Assistant Device with Biometric Verification in Real Scenario. In 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE) (pp. 426-431). IEEE.

[6] İnal-Çekiç, T., & Woźniak, U. E. (2022). We, the City: Plurality and Resistance in Berlin and Istanbul (p. 146). De Gruyter.

[7] Srinivasan, S., Muthu Kannan, P., & Kumar, R. (2022). A Machine Learning Approach to Design and Develop a BEACON Device for Women's Safety. In Recent Advances in Internet of Things and Machine Learning: Real-World Applications (pp. 111-115). Cham: Springer International Publishing.

[8] Wagh, N. R., & Sutar, S. R. (2022). An Enhanced Security of Women and Children Using Machine Learning and Data Mining Techniques. Data Mining and Machine Learning Applications, 423-446.

[9] Swapnarani, P., Rao, P. R., & Gunjan, V. K. (2022). Self defence system for women safety with location tracking and SMS alerting through GPS and GSM networks. In Modern Approaches in Machine Learning & Cognitive Science: A Walkthrough (pp. 361-368). Cham: Springer International Publishing.

[10] Torku, A., Chan, A. P., Yung, E. H., & Seo, J. (2022). Detecting stressful older adults-environment interactions to improve neighbourhood mobility: A multimodal physiological sensing, machine learning, and risk hotspot analysis-based approach. Building and Environment, 224, 109533.

[11] MAGESH KUMAR.S and RAJ KUMAR.M, "IPROB – EMERGENCY APPLICATION FOR WOMEN" ISSN 2250
- 3153 InternationalJournal of Scientific and Research Publications,online at ww.ijsrp.org Volume 4, Issue 3, March2014.

[12] Dr. Sridhar Mandapati, SravyaPamidi and Sriharitha Ambati, "A Mobile Based WomenSafety Application (I Safe Apps)"

[13] Abhaya: An Android App for the safety ofwomen. IEEE journal paper available fromhttp://ieeexplore.ieee.org/document/7443652/.

[14] Suraksha. A device to help women in distress: Aninitiative by a student of ITM UniversityGurgaon.efytimes.com.2013.Available from:http://efytimes.com/e1/118387/

Lung Cancer Classification Framework using Hybrid Bat Chicken Swarm Optimization (Bat-CSO) Algorithm and Machine Learning

V. Sreeprada
Department of Computer Science,
GITAM Deemed to be University
Visakhapatnam, Andhra Pradesh, India
vsreeprada21@gmail.com

Dr. K. Vedavathi
Department of Computer Science,
GITAM Deemed to be University
Visakhapatnam, Andhra Pradesh, India

***Abstract*—Lung cancer is the biggest cause of mortality across the globe, and it affects both men and women. Early identification is critical to improving long-term survival rates and enhancing the prospects of recovery. On medical imaging for illness identification, deep learning has recently shown great potential when used. Lung cancer detection and classification utilizing Hybrid Bat-CSO and SVM algorithms are presented in this article. It is recommended to use a threshold-based method to differentiate the candidate nodules from other structures. Different statistical and shape-based features are extracted for these nodule candidates. Using these different features, a discriminative feature vector is designed. We apply the Hybrid Bat and Chicken Swarm Optimization (CSO) algorithm for optimized feature selection. Based on the feature vector, the selected regions are classified using SVM algorithm. Based on SVM, the training features are estimated and trained. Once the nodules are detected, they are classified as severe, moderate and very severe. The proposed framework is implemented in Python. Experimental results have shown that Bat-CSO-SVM framework outperforms the existing methods in terms of accuracy, sensitivity, and specificity and F1-score metrics.**

Keywords—Cancer diagnosis, Machine learning, Optimization algorithm

I. Introduction

Diseases affecting the airways and other lungs' structures are referred to as ''lung disorders,'' or ''respiratory diseases.'' Pneumonia, TB, and Coronavirus Illness 2019 are all examples of lung disease (COVID-19). Diseases of the lungs have long been recognized as a primary cause of mortality and disability across the world's population. Early detection is essential to raising long-term survival rates and the probability of recovery. Lung disease has typically been identified through skin tests, blood tests, sputum sample tests, chest X-rays, and computed tomography (CT) scans. Deep learning has lately shown incredible promise for application in diagnosing illnesses from medical imaging. [1]. Cancer is a condition marked by uncontrolled cellular growth inside a life form. The greatest cause of cancer-related death in both men and women globally is lung cancer. It is essential to identify and treat the disease in order to lower the risk of lung cancer. [2].

Men and women are equally affected by lung cancer, which is the leading global cause of death for both sexes. The two most prevalent types of primary lung cancer are small cell lung cancer and non-small cell lung cancer. As a consequence of recent advancements in chemotherapy and radiation treatment, adenocarcinoma, squamous cell carcinoma, and large cell carcinoma have been added to the list of potential diagnoses.

Due to their different physical characteristics, adenocarcinoma and squamous cell carcinoma must often be distinguished through immuno histochemical analysis. These cancer cells are available in a range of sizes and forms. The use of computer-aided diagnostics helps prevent misdiagnosis (CAD). Large cell carcinoma is the most easily recognized of the four major types of cancer due to its atypia. Because the other three types—adenocarcinoma, squamous cell carcinoma, and small cell carcinoma—are often mixed together in cytological specimens, we concentrate on classifying them. [3].

In order to adopt a risk adjustment model, it is essential that the risk factors be examined. In addition, prior studies on risk adjustment used basic statistical approaches to predict individual medical expenditure and survival. However, such an approach is not always committed to the overall reality when it comes to missing and ambiguous data. [4].

Predictive models of lung cancer risk have been examined in earlier research to decrease screening expenses. The risk of lung cancer might also be identified using spirometry measures and CT-based emphysema evaluations, it has been proposed as an additional method. There have been a number of studies looking at the connection between CT-based emphysema testing and an elevated risk of lung cancer. [5].

The lungs must be manually tagged by the radiologist in order to use most available CAD-based nodule detectors, which is laborious and time consuming. Existing solutions based on form templates will be incorrect since nodules can have a variety of regular and irregular shapes and sizes.

In this paper, we present a fully automatic framework for Lung cancer Detection and Classification using Hybrid Bat and CHO (Bat-CHO) and SVM algorithms.

II. Related Works

For screening chest CTs, Peng Huang et al [6] created a deep learning algorithm (DeepLR) that accounted for all important nodule and nonnodule characteristics and successfully identified the existence of lung cancer within a 3-year timeframe. A different country's dataset did not affect DeepLR's accuracy or generalizability. For the sake of reproducibility and avoiding artificially high accuracy, we adopted a double-blinded experimental design. As compared to Lung-RADS, DeepLR found an even smaller high-risk group with a greater cancer rates.

Dora Revesz et al [7] have compiled a list of (1) TO's requirements for DSS in the treatment of incurable (stage IIIB/IV) NSCLC patients, and (2) their desires for the development of future instruments in this area.

All members of the Society of Physicians in Chest Medicine and Tuberculosis' Section of Oncology were sent an online inventory questionnaire. In order to better understand the results of the questionnaire, telephone interviews were undertaken. A total of 58 TO answered the questionnaire and said that they were in need of new DSS. That's what they said, citing the importance of genetic and immunological indicators being included in tools that are verified, updated, and time-saving. Also crucial is the incorporation of various treatment options in future DSS, as well as a clear interface and constant tool updating for IT solutions to ensure clear communication between caretakers and patients, as well as combining forecasts of toxic and quality of life.

Patients are willing to accept more severe long-term therapy side effects in exchange for a longer period without illness development, according to a study by Ellen M. Janssen et al. (8); There were the longest time equivalents for long-term functional and physical side effects followed by long-term emotional adverse consequences. In order to assess the acceptability of side effects in relation to patient survival improvements, it would be helpful to calculate time equivalents for reductions in treatment harms when evaluating current and future lung cancer therapies. Cancer patients' preferences and willingness to trade short-term side effects and various long-term adverse effects can be better understood using data from this study.

Yaojie Zhou et al [9] gave a short introduction to AI and radiomics, as well as its current and important application in the area of lung cancer imaging. More research utilizing the new AI-based approach are predicted in the future to enhance the treatment of lung cancer; at the same time, standardization for those trials also calls for greater marketing of AI systems to various populations and clinical centers. based facilities.

Savannah L. Bergquist et al [10] found that their ensemble machine learning system with a classification rule determined by the median beat an existing clinical decision tree for this task, resulting in a full sample performance of 93 percent sensitivity, 92 percent specificity, and 93 percent accuracy. Quality and outcomes in cancer treatment can now be measured and risk adjustment methods improved thanks to this research, which has the potential to be widely used by healthcare providers, payers, and policymaker.

According to T.Manikandan et al. [11], an approach has been proposed to enhance efficiency in the lung cancer diagnostic system, which calls for the segmentation of suspicious lung nodules based on geography and the detection of malignancy using FIS.

“There are two stages to the implementation of the suggested strategy. The initial step is to use a wiener filter to remove the major noise before using region-growing to segment the suspicious nodules in the lungs from the CT images. The diameter, shape, and intensity values of the segmented nodules are extracted and used as input to the FIS in the second phase of the classification process. Based on IF-THEN principles, the Fuzzy system determines the severity of the suspected lung nodules.. The suggested approach has a sensitivity of 92.3%, indicating that it may aid radiologists in improving their level of diagnostic confidence.

III. Proposed Solution

A. Overview

In this paper, we present a fully automatic framework for Lung cancer detection and classification using Hybrid Bat and CSO (Bat-CSO) and SVM algorithms. Figure 1 shows the block diagram of the proposed framework.

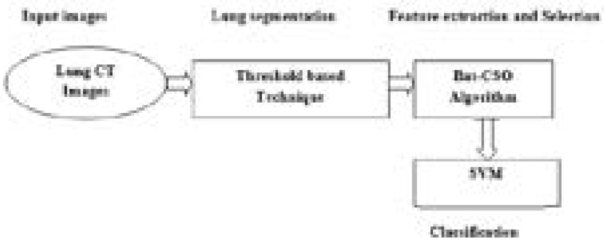


Fig.1. Block Diagram of Hybrid Bat-CHO-SVM framework

Auto-isolation of the lung location from the foundation may be achieved using a Grayscale CT histogram. Candidate nodules are separated from other structures using a threshold-based method. These nodule candidates have several statistical and shape-based attributes retrieved. To ensure that the classifier uses only relevant characteristics, the retrieved features have been tuned using feature selection. A discriminative feature vector (F) is generated using multiple statistical features, shape-based features, and properties across slices of the candidate areas. We use the Hybrid Bat–CSO algorithm to choose the most relevant features. The SVM method is used to classify the areas based on the feature vector F. Characteristics from a training dataset are used to train a model, which subsequently uses these same features to classify a testing dataset For each image feature, SVM Struct is calculated. Based on SVM struct, the training features are estimated and trained. Once the nodules are detected, they are classified as severe, moderate and very severe.

B. Lung Segmentation

Let I be a pre-processed MxN-pixel input lung image. The second local minimum l of the histogram H is produced and used to threshold the image I in order to eliminate the background from the image using the step size k.

$$\dot{I}(x,y) = \begin{cases} I(x,y) & if I(x,y) > \lambda \\ 0 & otherwise \end{cases} \tag{1}$$

$\dot{\hat{I}}$ is the thresholded image. White is utilized to depict the foreground while black is used to represent the background; this image is then complemented. Complementary image I produced from *I* is depicted in Figure 2. Step 2 means splitting the lung from the rest of the body. Because the bright area covering the lungs contains patches of varying intensities, the histogram-based thresholding used to draw attention away is ineffective. To keep the lungs separate from the rest of the body, we'll use the Otsu technique. T is the threshold at which the binary mask B of the lungs can be reconfigured:

$$B(x,y) = \begin{cases} 1 \; if I(x,y) > \tau \\ 0 \; otherwise \end{cases} \quad (2)$$

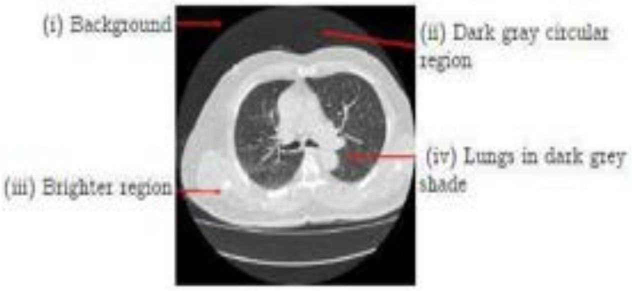


Fig.2. A CT image of lung and proposed division into four regions.

C. Extraction of Statistical and Shape Features

During this phase, the lung parenchyma is first separated from the lung nodules, bronchi, and arteries. The Otsu method is used to separate the inner structural vessels, bronchi, and nodules from the rest of the area in segmented lungs. There are several ways in which the nodules vary from other pulmonary structures. Nodules are differentiated from other structures, such as vessels and bronchi, by their morphology, which allows us to examine them. The nodules are spherical, while the arteries and the bronchi are cylindrical. We can extract nodules of any form using the suggested approach instead of current nodule template-based methods, making our method more flexible. We use statistical and shape-based aspects of nodules to build a feature vector that distinguishes nodules from the rest of the internal structure. Regions of interest are analyzed using these statistical and shape features.

D. Shapes and Statistical Features

Let $\{c_1, c_2, c_3, \dots\dots c_n\}$be the centers of the shapes extracted using region growing algorithm. The centers are passed to the region-growing algorithm as seeds, which returns us the corresponding n nodule candidate regions. $\{A_1, A_2, A_3, \dots\dots A_n\}$.

We compute different statistical properties, shape-based features of nodules and construct a feature vector to discriminate nodules from the other inner structures. The following statistical and shape features of candidates regions are used:

1. Mean(u_i): It represents the average value of the region A_i:

$$\mu_i = \frac{1}{n} \sum_{xe\,A} x \quad (3)$$

2. Median (m_{ei}): It is the mid-point of A_i when arranged in non-decreasing order;

3. Mode (m_{oi}): it is the most repetitive element of the data in A_i;

4. Variance (σ^2) : It represents to what extent the data varies from the mean value. For region A_i, (σ^2)

5. Standard deviation: σ is the square root of variance:

$$\sigma_i = \sqrt{\sigma_i^2} \quad (5)$$

6.Consistency feature (t_i) : One more important feature is based on the shape of the lesion and its appearance in the colocated slices of the CT scan. That is, if a nodule exists in one slice, it must also appear in the preceding slices or in the succeeding slices of a CT scan.

$$\underbrace{S_{j-k}, \dots S_{j-2}, S_{j-1}, S_j, S_{j+1}, \dots S_{j+k}} \quad (6)$$

We assign a center point 1 if it exists in any of those *2k* slices, and 0 if it does not exist in any of them.

E. Feature Vector

It is necessary to compute each of the features listed above for each nodule candidate A_i in a slice *Sj*, and then combine them to create a feature vector F_i,

$$F_i = |u_i, m_{ei,} m_o i, | \quad (7)$$

E. Optimized Feature selection using Bat-CSO

F. Bat Algorithm

Detecting distinctions in barriers and prey based on sound frequencies received from their environment is a common characteristic of bats, mammals. Bats use the recursive frequencies in their environment to detect food by making loud noises. A worldwide solution to the BA's problem is found through the use of bat echolocation. BA operations require a high degree of speed and accuracy [17]. The following assumptions are made in order to reduce the complexity of this method :

1. All bats rely on echolocation to discover their meal, regardless of species. Bats use this capacity to distinguish between prey and barriers.

2. There is a minimum frequency of *f* min created by the bats as the fly randomly in the direction of *yl*, therefore the wavelength of the sound produced is λ, and the loudness is A_0.

3. No matter how loud you make things, this parameter is often seen of as lying somewhere between A_0 and A_{min}.

The bats' frequency, velocity, and location are all updated using the equations shown below:

$$vl(t) = [yl(t) - y^*]xfl \quad (8)$$

$$yl(t) = [yl(t-1) + vl(t) \quad (9)$$

Where *yl(t-1)* is the bat's position at time *t-1*, is a random vector from 0 to 1, y is the bat's best position, f_l is the bat's sound frequency, f_{max} is the maximum frequency *l*; the index of number of bats (l = 1, 2,...., population size), and f_{min} is the minimum frequency. For local searches, the bat use the equation seen below:

$$y(t) = y(t-1) + \varepsilon A(t) \quad (10)$$

Y(t) is a random number (in the range of -1 to 1), and A(t) is the loudness of the sound. the pulse rate (rl) increases, but the loudness drops, as the bat discovers prey Between 0 and 1, the pulse rate may be determined. For each algorithm level, the volume and pulse rate should be adjusted. The pulse rate is adjusted using the following equation :

$$r_l^{t+1} = r_l^o(1 - \exp(-\gamma t)) \; A_l^{t+1}) = \alpha \, A_l^t \quad (11)$$

Where r_l^{t+1} is the new pulse rate α and γ are the constant coefficients. When 0<α<1 and γ>0. $A_l^t \to 0$ when $t \to \infty$, $r_l^t \to r_l^0$ Different parameters of the BA can be seen in the above equations. These parameters have different roles in the optimization process. For example, the decision variables are inserted into the algorithm based on the initial bat population. The initial position of the bats is considered to be a decision variable. Additionally, the frequency is used to update the velocity for each level, and then, the bat can find prey as one objective. Based on updating these parameters, the bats will receive the frequency and adjust their velocity to find the best position.

G. CSO Algorithm

CSO includes the following phases:

H. Update of Cock Position

When the cock is away from the mid-point p of the hen in a cluster, it will perform a random search in a bigger region, thus improving the probability of global search. On the other hand, when the cock is nearer to the mid-point p of the hen, it will perform a random search in a smaller region, which improves the local estimation capacity of the algorithm and hence improves the chances of determining the best solution [23].

The cock position update equation is given as

$$P_i^C(t+1) = QP_i^C(t)(1 + Z(-1, \alpha^2)) \quad (12)$$

where, C is an individual cock. Q = γe^d, where, d is the distance between the p of the i^{th} group and position of cock, γ is an proportional constant which is equal to 1. $Z(-1, \alpha^2)$ is Guassian distribution with -1 as mean and α^2 as variance

$$\alpha^2 = \exp\left(\frac{fi - f_k}{|f_i| + \partial}\right) i, k \in [1,2,..,N], k \neq i \quad (13)$$

$k \in [\ 1,2,..,N\ ]$ and iare cock's indices, dynamically chosen from the cock groups, i$\neq k$, f_i and f_k represents the fitness function of the i^{th} and k^{th} cocks, respectively. ∂ is a small constant that is used to evade the divide by zero error and γ is the proportional co-efficient.

I. Update of Hen Position

Among the initial population, G_{best} individuals with better fitness values are selected and one among them is randomly chosen as the supreme individual. The G_{best} value is set manually, during experiments, which should be less than the number of cocks in the population. During the foreaging step, the supreme individual is considered as the learning object for the hen. Hence, the position of the hen (H) is updated based on the supreme individual using the following equation

$$P_i^H(t+1) = P_i^H(t) * S_1 * \text{rand} * \left(P_{r1}^C(t) - P_i^H(t)\right) + S_2 * \text{rand} * \left(P_{r2}(t) - P_i^H(t) + \text{rand} * \left(P_e\ (t) - P_i^H(t)\right)\right), e \in [1,2,..,\text{hao}] \quad (14)$$

Where, rand is a random number in the range of [0,1], C indicates that the individual is a cock, H indicates that the individual is a hen, $r_1$is the index of cock, which is the ith hen's group mate, r2 is the index of the chicken , randomly selected from the group, $P_e(t)$ is the selected individual hence supreme learning object in the t^{th} iteration , hao is the number of elite individuals reserved in the population.

S_1 and S_2 are calculated as

$$S_1 = \exp(f_i - f_{r1})/\text{abs}(f_i) + \partial \quad (15)$$

$$S_2 = \exp(f_{r2} - f_i) \quad (16)$$

Where f_i , f_{r1}, f_{r2} represent the fitness function values at i, r1 and r2, respectively.

J. Update of Chick Position

The chick's position can be updated based on the position of the cock, using the following equation

$$P_i^{CH}(t+1) = P_i^{CH}(t) + FL * \left(P_m^H(t) - P_i^{CH}(t)\right) + w * \left(P_{r1}^C(t) - P_i^{CH}(t)\right) \quad (17)$$

Where, m denotes the index of chick's mother, r^1 is the index of the cock which is the i^{th} chick's group mate, CH is the individual chick, FL and ware learning factors

K. Hybrid BA-CSO

Based on the BA and CSO's communication approach, a hybrid structure has been developed for the new algorithm. Essentially, the best answers from one method are used to replace the worst ones from the other in a hybrid approach to optimization.

Sharing information between subgroups of the initial population occurs after the initial population has been split into smaller groups. If N is taken as the initial total population, then N1 and N2 represent the populations of the bat and CHO algorithms.

The steps involved in the Hybrid Optimization algorithm can be listed as follows:

- First and foremost, both algorithms take into account the starting populations. The bats' locations and velocities are specified in addition to the particle's starting positions and velocities.
- Evaluation: each algorithm's objective function should be calculated independently, and the solution candidates should be evaluated accordingly.
- As a result of Equation (14), the velocity and location of CSO are updated (17). Equations (8) and (9) are used to update the BA's velocity and position
- Algorithms can be improved by copying and moving the best solution candidates from one algorithm to the other.
- Levels 2 through 4 are repeated until the number of iterations has been reached, at which point the best is chosen.

L. Classification using SVM

Depending on the feature vector V, the extracted regions are classified using SVM algorithm. The best features are processed by dividing them into training and testing sets with labelled results.

M. SVM

The fundamental aim of SVM is to determine the optimal hyperplane in the feature space which maximally divides the target class into two. Geometrically, the SVM algorithm determines an optimal hyperplane with the maximal margin to divide the two classes,

The training set of SVM is represented as

$(x_1,y_1),(x_2,y_2),\ldots(x_n,y_n)$, $x_j \in R^n$, $y_j \in \{+1,-1\}$.

Here x_j is the characteristic vector of j^{th}model as input and y_j is the output catalogue = +1 or -1.

SVM splits the +ve and –ve instances by means of a hyperplane as

$$w..x + b = 0, w \in R^n, b \in R \quad (18)$$

Here w.x denotes the dot product of w and x.

As important as the inclusion of kernel functions is, it is difficult to tell whether overfitting is to blame even if the kernel functions are discovered to make samples linearly separable in the feature space. Allowing SVM some fault tolerance on the sample alleviates this issue. (ie) Because the ‘‘hard interval’s differs from the ‘‘soft interval’s, certain samples may not match the following constraints: ‘‘

$$y_i.(w^t x_i + b) \geq 1 \quad (19)$$

Obviously, this does not imply that the ‘‘error’‘ samples are arbitrary and unrestricted in terms of number. We update the optimization target to minimize the number of samples that do not meet the constraints.

$$\min \frac{1}{2}||w||^2 + c\sum l_0(y_i(w^t x_i + b) - 1) \quad (20)$$

Polynomial kernel function is as follows:

$$k(x_1, x_2) = (<x_1, x_2 + R)^d \quad (21)$$

Gaussian kernel function is as follows:

$$k(x_1, x_2) = Exp\left(\frac{||x_1, x_2||^2}{2\sigma^2}\right) \quad (22)$$

Linear kernel function is as follows:

$$k(x_1, x_2) =< x_1, x_2 > \quad (23)$$

The SVM training algorithm is given below:

SVM Training Algorithm
Input: D(X,Y): X (array of input features), Y(array of class labels),
N – Number of iterations
Output: Classified labels

1. Initialize the learning_rate randomly.
2. **for** (learning_rate in N)
3. Error=0
4. **for** (i in each X)
5. **if** ((Y[i] * X[i]) * w)<1 **then**
6. Update w=w+learning_rate*((X[i]*Y[i])*(-2*(1/N)*w)
7. **else**
8. Update w=w+learning_rate * (-2*(1/N)*w)
9. **end if**
10. **end for**
11. **end for**

IV. Experimental Results

A. Input and Output

Figure 2 shows some of the lung images used for training. The nodules detected from each slice of the images are shown in Figure 3.

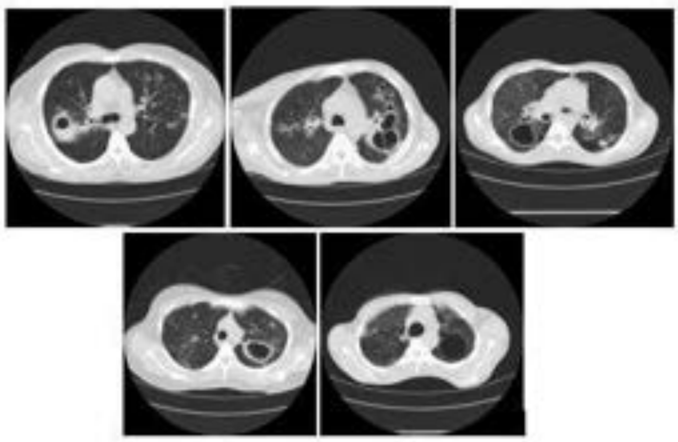


Fig.2. Input Lung Images

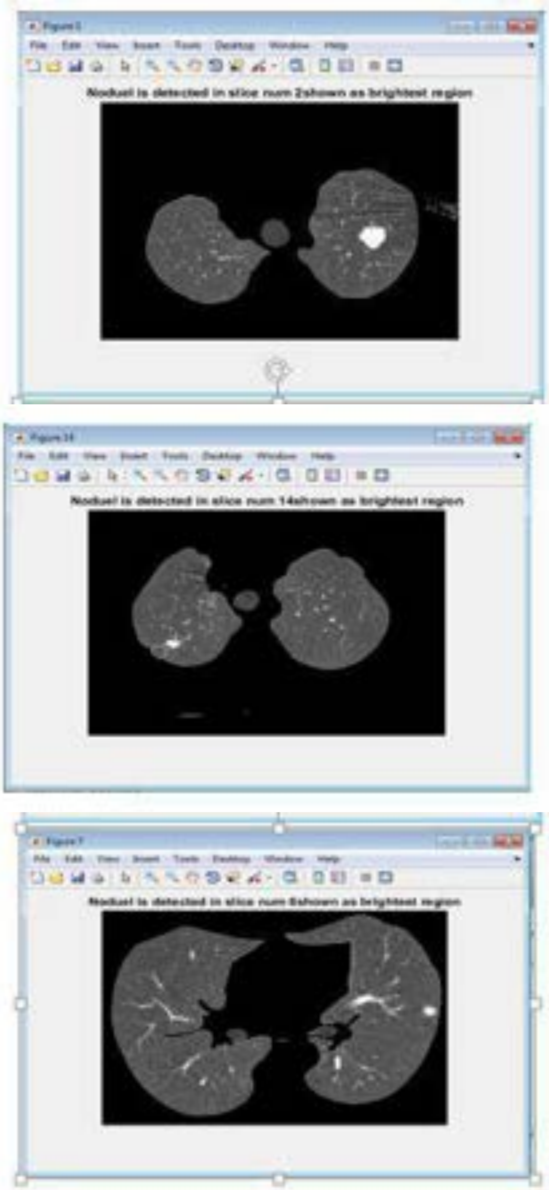


Fig.3. Nodules detected in each slice of the lung images

B. Performance Results

In this section, the performance results of Bat-CSO-SVM are compared with SVM, DNN and PSO-DNN algorithms.

TABLE I. Performance comparison of algorithms

Performance Metrics	Bat-CSO-SVM	SVM	DNN	PSO-DNN
sensitivity	78.43	74.00	72.00	76.00
specificity	95.41	92.25	93.00	93.61
Accuracy	97.14	95.18	94.76	95.32
MCC	77.86	75.32	74.31	75.00
F1-Score	78.43	74.56	73.27	75.69

In this comparison of four models—Bat-CSO-SVM, SVM, DNN, and PSO-DNN—evaluated on various performance metrics, Bat-CSO-SVM consistently demonstrated the highest sensitivity, specificity, accuracy, MCC, and F1-Score among the models, achieving values of 78.43%, 95.41%, 97.14%, 77.86%, and 78.43% respectively. While SVM, DNN, and PSO-DNN also performed well across the metrics, Bat-CSO-SVM exhibited superior overall performance, indicating its effectiveness in classification tasks. These results suggest that the integration of Bat Algorithm and CSO (Cuckoo Search Optimization) with SVM can enhance the model's predictive capabilities, making it a promising approach for various classification problems.

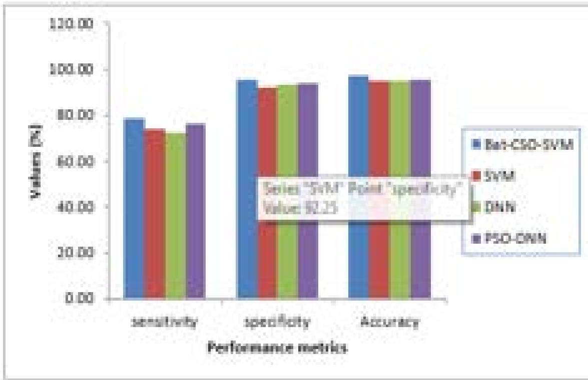


Fig.4. Performance comparison of Accuracy, sensitivity and specificity

From Figure 4, it is observed that the sensitivity of Bat-CHO-SVM is 6% higher than SVM, 8% higher than DNN and 3% higher than PSO-DNN, the specificity of Bat-CSO-SVM is 3% higher than SVM 2% higher than DNN and 2% higher than PSO-DNN, Similarly the Accuracy of Bat-CSO-SVM is 2% higher than SVM 3% higher than DNN and 2% higher than PSO-DNN.

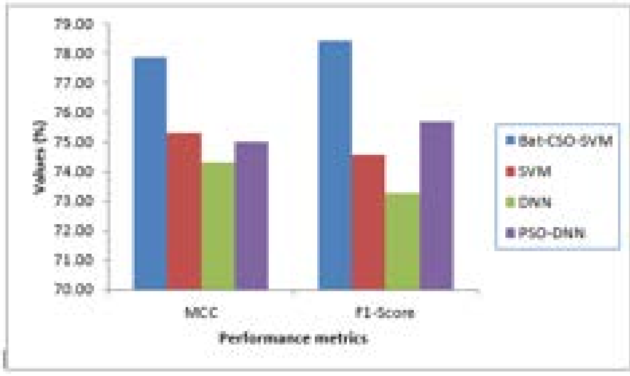


Fig.5. Performance comparison of MCC and F1-Score

From figure 5, it is observed that the MCC of Bat-CSO-SVM is 3% higher than SVM, 5% higher than DNN and 4% higher than PSO-DNN, the F1-Score value of Bat-CSO-SVM is 5% higher than SVM 7% higher than DNN and 4% higher than PSO-DNN.

V. CONCLUSION

This study has proposed a Lung cancer detection and classification framework using Hybrid Bat-CSO and SVM algorithms. A threshold-based technique is proposed to separate the candidate nodules from other structures. Different statistical and shape-based features are extracted for these nodule candidates. Using these different features, a discriminative feature vector is designed. We apply the Hybrid Bat-CSO algorithm for optimized feature selection. Based on the feature vector, the selected regions are classified using SVM algorithm. Based on SVM, the training features are estimated and trained. Once the nodules are detected, they are classified as severe, moderate and very severe. The proposed framework is implemented in Matlab. The performance results of Bat-CSO-SVM are compared with SVM, DNN and PSO-DNN algorithms. Experimental results have shown that Bat-CSO-SVM framework outperforms the existing methods in terms of accuracy, sensitivity, specificity and F1-score.

REFERENCES

[1] Stefanus Tao Hwa Kieu, Abdullah Bade, Mohd Hanafi Ahmad Hijazi and Hoshang Kolivand, "A Survey of Deep Learning for Lung Disease Detection on Medical Images: State-of-the-Art, Taxonomy, Issues and Future Directions", Journal of Imaging, Vol-6, No-131, 2020.

[2] Thiyagarajan Manikandan and Nallaiyan Bharathi, "Hybrid neuro-fuzzy system for prediction of stages of lung cancer based on the observed symptom values.", Biomedical Research 2017; 28 (2): 588-593, 2017.

[3] Atsushi Teramoto, Tetsuya Tsukamoto, Yuka Kiriyama and Hiroshi Fujita, "Automated Classification of Lung Cancer Types from Cytological Images Using Deep Convolutional Neural Networks", Hindawi BioMed Research International, Volume 2017, Article ID 4067832, 6 pages, 2017.

[4] Kung-Jeng Wang, Jyun-Lin Chen and Kung-Min Wang, "Medical expenditure estimation by Bayesian network for lung cancer patients at different severity stages", Elsevier, Computers in Biology and Medicine 106 (2019) 97–105, 2019.

[5] Mizuho Nishio, Takeshi Kubo and Kaori Togashi, "Estimation of lung cancer risk using homology-based emphysema quantification in patients with lung nodules", PLOS, Vol-14, No-1, 2019.

[6] Peng Huang, Cheng T Lin, Yuliang Li, Martin C Tammemagi, Malcolm V Brock, Sukhinder Atkar-Khattra, Yanxun Xu, Ping Hu, John R Mayo, "Prediction of lung cancer risk at follow-up screening with low-dose CT: a training and validation study of a deep learning method", Digital Health, Vol-1, 2019.

[7] Dora Revesz,Ellen G. Engelhardt,Johannes J. Tamminga,Franz M. N. H. Schramel,Bregje D. Onwuteaka-Philipsen,Ewoudt M. W. van de Garde,Ewout W. Steyerberg,Henrica C.W. de Vet and Veerle M. H. Coupe, "Needs with Regard to Decision Support Systems for Treating Patients with Incurable Non-small Cell Lung Cancer", Journal of Cancer Education (2020) 35:345–351, 2020.

[8] Ellen M Janssen ,Sydney M Dy,Alexa S Meara ,Peter J Kneuertz,Carolyn J Presley and John FP Bridges, "Analysis of Patient Preferences in Lung Cancer-Estimating Acceptable Tradeoffs Between Treatment Benefit and Side Effects", Patient Preference and Adherence, 2020.

[9] Yaojie Zhou, Xiuyuan Xu, Lujia Song, Chengdi Wang, Jixiang Guo,Zhang Yi and Weimin Li, "The application of artificial intelligence and radiomics in lung cancer", Precision Clinical Medicine, 3(3), 2020, 214–227, 2020.

[10] Savannah L. Bergquist,Gabriel A. Brooks,Nancy L. Keating and Sherri Rose, "Classifying Lung Cancer Severity with Ensemble Machine Learning in Health Care Claims Data", Proc Mach Learn Res. 2017 August ; 68: 25–38, 2017.

[11] T.Manikandan and N. Bharathi, "Lung Cancer Diagnosis from CT Images Using Fuzzy Inference System", Communications in Computer and Information Science, 2011.

[12] Nilesh B. Bahadure, Arun Kumar Ray and Har Pal Thethi," Feature Extraction and Selection with Optimization Technique for Brain Tumor Detection from MR Images", IEEE International Conference on Computational Intelligence in Data Science, 2017

[13] Chiwen Qu,Shi'an Zhao Yanming Fu and Wei He," Chicken Swarm Optimization Based on Elite Opposition-Based Learning", Mathematical Problems in Engineering, Hindawi, Volume 2017, Article ID 2734362, 20 pages

[14] Mahdi Valikhan-Anaraki, Sayed-Farhad Mousavi, Saeed Farzin, Hojat Karami, Mohammad Ehteram, Ozgur Kisi, Chow Ming Fai , Md. Shabbir Hossain , Gasim Hayder , Ali Najah Ahmed , Amr H. El-Shafie, Huzaifa Bin Hashim, Haitham Abdulmohsin Afan,Sai Hin Lai and Ahmed El-Shafie,"Development of a Novel Hybrid Optimization Algorithm for Minimizing Irrigation Deficiencies", Sustainability, April 2019

[15] Mudhafar Jalil Jassim Ghrabat, Guangzhi Ma, Ismail Yaqub Maolood, Shayem Saleh Alresheedi and Zaid Ameen Abduljabbar," An effective image retrieval based on optimized genetic algorithm utilized a novel SVM-based convolutional neural network classifier", Human Centric Computing and Information Sciences, 2019, 9:31

[16] Noor Khehrah, Muhammad Shahid Farid, Saira Bilal and Muhammad Hassan Khan," Lung Nodule Detection in CT Images Using Statistical and Shape-Based Features", Journal of Imaging, MDPI, Feb 2020

[17] Xin-She Yang, "A New Metaheuristic Bat-Inspired Algorithm",Springer,Nature Inspired Cooperative Strategies for Optimization (NICSO 2010),Vol-284,pp:65-74,2010.

Innovative Binarization Solutions for Historical Document Clarity

Radhika V. Kulkarni
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
radhika.kulkarni@vit.edu

Vedant Mude
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
vedant.mude21@vit.edu

Rutuj Nagrale
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
rutuj.nagrale21@vit.edu

Aarya Nirgude
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
aarya.nirgude21@vit.edu

Tejashri Nirmal
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
tejashri.nirmal21@vit.edu

***Abstract*— Images of historical documents often have characteristics, such as wrinkles, faint writing, stains, bleed-through ink, and other issues. These factors distort the text visibility and affect the performance of binarization. Preserving these document images aids future generations in learning about a variety of subjects. This article presents a new binarization approach for historic documents. This work uses bilateral and unsharp filtering, Otsu thresholding, histogram analysis, and k-means clustering for dark spot removal as part of a multi-step image enhancement process. Applying an intensity-based mask raises the quality of pixels above a set threshold. Furthermore, the method includes two additional refinements: a concluding sharpening phase and an enhancement of color contrast. The performance of proposed binarization approach is assessed using the Flesch Reading Ease Formula. The findings reveal that the algorithm achieved its highest readability score when applied to degraded documents, with an average readability score of 96.44. This suggests the algorithm's efficacy in enhancing the readability of noisy images, particularly in the context of degraded documents.**

Keywords—Binarization, historical documents, Image filtering, Otsu thresholding, K-means clustering, Flesch Reading Ease Formula, image enhancement

I. Introduction

Over thousands of years, a vast number of historical records with important information on literary arts and historical knowledge have been left behind. The historical document collection has seen significant deterioration, such as creases, faint text, stains, bleed-through ink, smeared ink, thin strokes, and degraded documents, due to years of storage. In recent years, a large number of researchers have become interested in the issue of document preservation. To safeguard printed paper documents against direct modification for consultation, exchange, and remote access, an effective method is to employ a document digitization system [1]. To increase accessibility and prevent physical copies from deteriorating, several libraries worldwide are digitizing historic documents. Document image processing algorithms are required to access the contents of these documents because document images cannot be read by machines in their raw form [2].

Noise reduction is the process of applying advanced image processing techniques to historic document photos to remove unwanted artifacts and interruptions. Certain methods are employed to reduce or remove the noises while preserving essential features; filters and denoising techniques are two examples of such algorithms. The goal is to enhance the text and image clarity while maintaining the original content integrity. This will make the historical document more readable and ensure its preservation for future usage [3].

In document images, it is common practice to map background pixels to white and foreground text pixels to black. Binarization is the process of transforming a multi-tone image into a bi-tonal image. Binarization can be used as a method of noise reduction to improve the readability of documents [4]. It is one of the preprocessing tasks that significantly affects other phases, such as feature extraction and recognition from document images, that call for a precise and high-quality foreground image [5]. Document image binarization is a crucial step in the pipeline for document image analysis and recognition that influences the outcomes of final recognition [6].

This paper studies rich literature on image enhancement and selects the most popular 5 filtering techniques. Through experimentation, this work obtains the optimized sequence of these techniques and proposes the system with the given sequence as sharpening filter, bilateral filter, k-means clustering, Otsu thresholding, and unsharp masking proves to be efficient for enhancement of historic document image. K-means clustering works well to eliminate dark areas, and histogram analysis makes it easier to understand how image intensity is distributed. After Otsu thresholding, an intensity-based mask is applied to further increase pixel values and establish an optimal threshold for binary conversion. The image details are sharpened once the color contrast has been adjusted. Together, these stages offer a comprehensive approach that enhances an advanced image-processing pipeline that can be used in a variety of disciplines. This method is applied to various categories of noisy images to obtain an average Reading Ease Score. A thorough study of the average scores of noisy images is presented further.

II. Literature Review

The paper [1] presents an end-to-end trainable framework that carries out layout analysis, recognition, and character detection all at once. To improve recognition performance even further, it suggests a re-score technique that uses the implicit language model to forecast damaged characters.

Annotations for layout, characters, and text lines have been added to the original dataset to aid in the study of Chinese historical writings.

The Binary Inpainting Network (BINet) is an autoencoder framework that uses binary inpainting to improve patch-based compression of static pictures. To facilitate parallel encoding and decoding, it restores interdependencies between nearby patches, doing away with the requirement for access to the original or rebuilt picture data. As mentioned in [7], at various compression levels, BINet significantly improves the compression quality.

An innovative strategy that has the advantage of simultaneously binarizing the image and reducing various types of noise is presented in the publication [8]. It converts an image that is noisy in grayscale to a much smaller low-noise binary image. The approach performs noticeably better when used on photos with similar text sizes and formats. As a result, official letters and document images from colleges or institutions greatly benefit from the algorithm.

A technique for removing random noise in several stages is presented in the publication [9]. Phase 1 involves scanning the complete text image top-down, then counting each connected black pixel region by going pixel by pixel through each column and row from left to right. These categories contain both noise-indicating sites and character/information locations. Phase 2 involves classifying the noise regions based on the number of black pixels in each numbered zone. To facilitate processing, phase 3 data is normalized to a range. Areas with pixels within the normalization range are used to represent data in phase 4. For more effective noise reduction, these stages are repeated a predetermined number of times.

The study [10] presents an algorithm that uses a step-by-step approach to image enhancement. Most document enhancement and cleaning procedures or binary picture conversions include a combination of intricate image processing techniques, increasing computing complexity and cost. The technology expedites the process by taking into account the distinct characteristics of the document images. Furthermore, the method is more user-friendly due to its iterated step structure.

There are four steps involved in using the algorithm: The process involves four steps: 1. determining the image's vector of parameters to be filtered; 2. applying a bilateral filter to filter the image; 3. dividing the image into its RGB components and binarizing each one using an approach based on Otsu's algorithm; and 4. selecting the RGB component that best preserves the document information in the foreground, which is the algorithm's final output [11].

The algorithm proposed in [3] does not provide a conclusion about histogram equalization; in fact, some of the images are deteriorated even more after histogram equalization. In terms of histogram equalization, the proposed approach provides conclusion for Triangle and Otsu thresholding. The work proposed in [6] works efficiently on a few images but is not tested on large number of images containing diverse noises while the proposed algorithm is tested on more than thousand images classified in 7 categories that are creases, faint text, stains, bleed-through ink, smeared ink, thin strokes and degraded image. This pipeline provides good result for images containing diverse noises while [5] fails to address this issue. The abundance of existing literature aids in the development of the suggested methodology, which improves historic document picture binarization outcomes.

III. METHODOLOGY

For improved historic document image binarization, a combination of filtering approaches, including bilateral filter, k-means clustering, Otsu thresholding, sharpening filter, and unsharp masking, is suggested in this methodology. The workflow of the proposed approach is shown in Figure 1, and the following subsections provide a brief explanation of each stage.

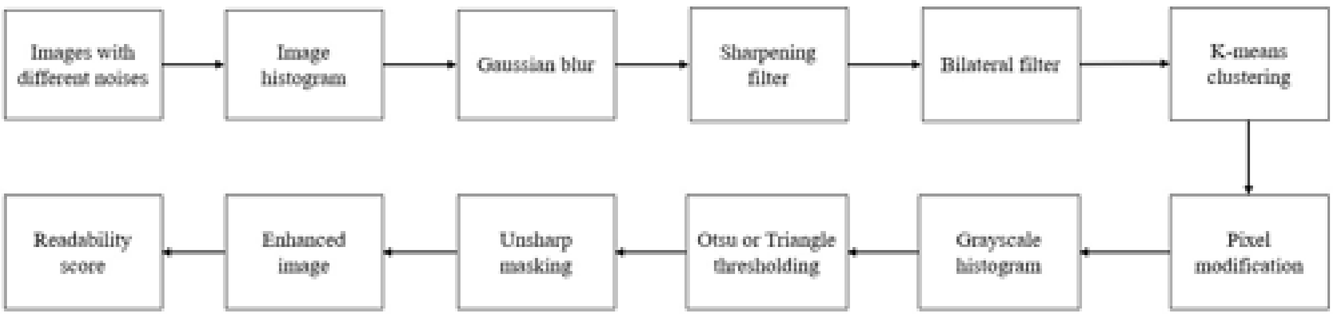


Figure 1. Proposed workflow

A. INPUT DATA

The input data of historical document images are obtained from Kaggle respository. [12] that has 1171 pictures with different kinds of noise in them. We have categorized those pictures into 7 categories creases, faint text, stains, bleed-through ink, smeared ink, thin strokes, and degraded. These 7 categories overlapped as a few noises altogether are present in the document images.

B. IMAGE HISTOGRAM

A graphical representation of an image pixel intensity distribution is provided by an image histogram. It provides a graphical representation of the frequency of each intensity level, ranging from dark to bright. The horizontal axis represents the intensity values (i.e., pixel values in a grayscale image ranging from 0 to 255), and the vertical axis shows how frequently each intensity occurs. The knowledge and optimization of image properties during a range of image processing tasks can be aided by an understanding of the general contrast, brightness, and distribution of pixel values through the examination of an image histogram [13].

C. GAUSSIAN BLUR

"Gaussian blur" is a convolutional image processing approach that includes using a Gaussian function to smooth an image [14]. This technique is often used to decrease noise and enhance photos. During the convolution process, the weighted average of the image pixel values is computed. The weights are determined by the values of a Gaussian kernel. A Gaussian kernel is a two-dimensional matrix where the core values gradually decline in a symmetric pattern, with the maximum weight allocated to them. The amount of blurring is determined by the standard deviation of the Gaussian function; a larger standard deviation results in a broader blur. Because Gaussian blur is computationally cheap and preserves edges to some extent, it is a versatile technique in image processing for applications such as pre-processing, background removal, and artistic effects.

D. SHARPENING FILTER

Denoising and image sharpening are crucial in image processing. Image sharpening seeks to enhance edge slopes without producing halo artifacts, whereas an image-denoising algorithm seeks to minimize noise while preserving image edges. Image sharpening methods such as the Unsharp Mask filter improve edge slopes without causing halo artifacts. The Unsharp Mask does this by purposefully boosting high-frequency elements and emphasizing edges via a deftly designed blend of blurring and subtraction. This method provides a versatile tool for progressively improving overall image quality by minimizing unwanted visual artifacts and optimizing image details [15].

Let $I(i,j)$ be the original image, and $I_{smooth}(i,j)$ be a blurred version of the image, obtained by applying Gaussian blur with kernel size σ.

The Sharpened image is obtained by the process described below:

$$I_{sharp}(i,j) = I(i,j) + \alpha \cdot (I(i,j) - I_{smooth}(i,j)) \quad (1)$$

where, $I_{sharp}(i,j)$ is the sharpened image, α is the scaling factor that controls the strength of the sharpening effect, and the sub that represents the smoothened image $I_{smooth}(i,j)$ from the original image $I(i,j)$ represents the high-frequency components i.e., edges of the image.

E. BILATERAL FILTER

A non-linear technique known as the bilateral filter can blur an image while preserving its sharp edges. Its ability to split an image into many scales without creating halo effects after manipulation makes it a popular choice for computational photography tasks such as tone mapping, denoising, style transfer, and relighting [16].

The bilateral filter is applied on the image mathematically as shown below:

$$B(i,j) = \frac{\sum_{p\in\Omega} I(p)\cdot w_c(i,j,p)\cdot w_s(i,j,p)}{\sum_{p\in\Omega} w_c(i,j,p)\cdot w_s(i,j,p)} \quad (2)$$

where $B(i,j)$ is the bilateral filter value at the spatial point (i,j) of the image. The color or intensity of the pixel at position q in the image is represented by $I(p)$. The spatial neighborhood surrounding the pixel (i,j) that the filter is applied to is denoted by Ω. The spatial weight is represented by $w_c(i,j,p)$, which is frequently defined as a Gaussian function depending on the spatial distance between (i,j) and q. The range weight is represented by $w_s(i,j,p)$, which is frequently described as a Gaussian function based on the color or intensity difference between (i,j) and p.

F. k-MEANS CLUSTERING

The image is divided into clusters using k-means clustering. Based on the given data set, k-means find new and undiscovered classes by segmenting the instance space into regions with related items. It is an iterative, unsupervised heuristic clustering method based on partitioning [17]. The dark spot clusters are discovered and treated individually to perhaps eliminate or minimize their influence on the image. Additionally, it aids in the improvement of text areas [18].

The process commences with the random initialization of k cluster centroids, represented $\mu_1, \mu_2, \dots, \mu_k$ within the feature space of the image pixels. To assign each pixel to the nearest centroid, the algorithm computes the Euclidean distance between each pixel x_i and every centroid μ_j. This distance is expressed as:

$$distance(x_i, \mu_j) = \|x_i - \mu_j\| \quad (3)$$

Utilizing this distance metric, each pixel is assigned to the cluster associated with the closest centroid. Subsequently, the centroids are updated by calculating the mean of all pixels assigned to each cluster. Mathematically, the centroid update equation is defined as:

$$\mu_j = \frac{1}{|S_j|} \sum_{x_i \in S_j} x_i \quad (4)$$

where, S_j denotes the set of pixels assigned to the cluster j. This iterative process continues until convergence, characterized by stabilized centroids or the fulfillment of predefined convergence criteria. The resulting clusters depict distinct regions in the image, facilitating tasks such as dark spot segmentation or text area enhancement. Through its iterative partitioning of pixel space and centroid re-computation, k-means clustering emerges as a versatile tool for uncovering underlying patterns within image data, thereby empowering a wide array of image processing applications.

G. PIXEL MODIFICATION

Pixel modification refers to modifying the intensity of pixels in an image with a specific target pixel value by adding a specified increment to their intensity values while ensuring that the resulting pixel values remain within a valid range.

This prevents overflow and underflow issues and ensures that the modified image remains in a valid format.

H. OTSU THRESHOLDING

Among the popular global thresholding techniques, Otsu selects a threshold value by maximizing the metric known as the between-class variance, which is crucial to statistical discriminant analysis [19]. The Otsu thresholding method is commonly used for binarization, particularly in cases where the optimal threshold value needs to be determined automatically. Binarization involves converting a grayscale image into a binary image, where each pixel is classified as either foreground (usually represented as white) or background (usually represented as black).

More explanation is provided for the Otsu thresholding algorithm. Let's say an image has K different grayscale representations of its pixels (1, 2,…, K) (i,j). Let N be the total number of pixels $N = \sum_{i=1}^{K} n_i$, and let n_i be the number of pixels at level i. Likelihood of occurring for level i is determined by $P_i = n_i/N$.

Let threshold T be used to separate the image into two classes, C_0 and C_1. Pixels with levels [1,..., T] make C_0, while pixels with levels [T + 1,..., K] make up C_1. The cumulative probabilities represented by $P_0(T)$ and $P_1(T)$, the mean levels by $\mu_0(T)$ and $\mu_1(T)$, and the variances of the classes C_1 by $\sigma_0^2\, T$ and $\sigma_1^2\, T$, respectively. Equations (5) through (10) provide all of these values, as detailed in [20].

$$P_0(T) = \sum_{i=1}^{T} p_i \tag{5}$$

$$P_1(T) = \sum_{i=1}^{K} p_i = 1 - P_0(T) \tag{6}$$

$$\mu_0(T) = \sum_{i=1}^{T} i \frac{p_i}{P_0(T)} = \frac{1}{P_0(T)} \sum_{i=1}^{T} i p_i \tag{7}$$

$$\mu_1(T) = \sum_{i=T+1}^{K} i \frac{p_i}{P_1(T)} = \frac{1}{P_1(T)} \sum_{i=T+1}^{K} i p_i \tag{8}$$

$$\sigma_0^2 T = \sum_{i=1}^{T} (i - \mu_0(T))^2 \frac{p_i}{P_0(T)} \tag{9}$$

$$\sigma_1^2 T = \sum_{i=T+1}^{K} (i - \mu_1(T))^2 \frac{p_i}{P_1(T)} \tag{10}$$

Let μ stand for the image mean level, $\sigma_b^2 T$ the between-class variance, and $\sigma_w^2 T$ the within-class variance. μ, $\sigma_b^2 T$, and $\sigma_w^2 T$ are defined, respectively, by equations (11), (12) and (13).

$$\mu = \sum_{i=1}^{K} i p_i = P_0(T)\, \mu_0(T) + P_1(T)\, \mu_1(T) \tag{11}$$

$$\sigma_b^2 T = P_0(T)\, (\mu_0(T) - \mu)^2 + P_1(T)\, (\mu_1(T) - \mu)^2 \tag{12}$$

$$\sigma_w^2 T = P_0(T)\, \sigma_0^2\, T + P_1(T)\, \sigma_1^2 \tag{13}$$

Equation (14) defines the cutoff that is reached by optimizing the between-class variance.

$$T^* = arg_{1 \le T < K}\ max\{\sigma_b^2(T)\} \tag{14}$$

This sum is equal to the cutoff found in equation (15) by lowering the within-class variances standard.

$$T^* = arg_{1 \le T < K}\ min\{\sigma_w^2(T)\} \tag{15}$$

I. TRIANGLE THRESHOLDING

Triangle thresholding is one of the common thresholding techniques used to calculate the optimal thresholding value based on the histogram of the image. Let's denote the histogram as $H(x)$, where x represents the intensity values. The peak of the histogram is denoted as $P = (x_P, y_P)$, where x_P is the intensity value and y_P is the frequency at that intensity value [21].

Peak Detection:

$$P = \arg max_x H(x) \tag{16}$$

Threshold Determination: The threshold value T is determined where the maximum distance occurs:

$$T = \arg max_{xi} di \tag{17}$$

here, x_i is the coordinate of the point on the histogram, d_i is the perpendicular distance from a point (x_i, y_i) to the line segment connecting the peak P.

J. UNSHARP MASKING

Unsharp masking is a technique in image processing that improves image edges and features [22]. Sharpening, or unsharp masking, is the reverse of blurring, despite the name suggesting otherwise. Usually, a Gaussian blur is applied to do this, with the amount of blurring determined by the standard deviation. By averaging each pixel in the image with its nearby pixels, the blurring is produced.

$$R(i,j) = O(i,j) * G(i,j) \tag{18}$$

where the blurred, original, and Gaussian blur pictures are represented, respectively, by $R(i,j)$, $O(i,j)$, and $G(i,j)$, with pixel coordinates (i,j).

Equation (19) shows that when the blurred image (mask) is subtracted from the original image, high-frequency elements like edges and features are highlighted in the resulting image.

$$HP(i,j) = O(i,j) - R(i,j) \tag{19}$$

where, $HP(i,j)$ is high pass filtered image.

Next, the high-pass filtered image is added back to the original image. This procedure increases the contrast of edges and little details, making them stand out more.

$$S(i,j) = O(i,j) + Hp(i,j) \tag{20}$$

here, $S(i,j)$ is a sharpened image.

Lastly, color contrast modifications contribute to improving the image brightness and visual attractiveness. A few adjustments that involve altering the saturation, color balance, or other features yield a more stunning and harmonious color representation. The final step of sharpening caps the method by adding crispness and refining the finer characteristics of the image. Each of these meticulously planned stages contributes to the overall goal of picture improvement and cleaning, resulting in a polished and eye-catching final output image.

K. READABILITY EASE SCORE

The Flesch Reading Ease score [23] is a readability metric designed to quantify the ease with which a reader can comprehend a given text. It is calculated based on the average number of syllables per word and the average number of words per sentence in the text. The formula for the Flesch Reading Ease score is as follows:

$$RE = 206.835 - (1.015 \times wps) - (84.6 \times spw) \qquad (21)$$

here, RE is the Flesch readability which is ease score, wps is average words per sentence, spw is average syllables per word.

IV. Results and Discussion

The suggested technique produces an improved image by combining several filters, thresholding, clustering, and unsharp masking. The input images of various noise types are chosen from the dataset [12] depicted in the figures shown below. The input images are present with creases, faint text, stains, bleed-through ink, smeared ink, thin strokes, or are degraded due to the aging of paper. Following are the results of document images with various types of noises which undergo the process as described in the methodology to output visually appealing images.

Figures 2, 4, 6, 8, 10, 12 and 14 represents the input document images of categories: creases, faint text, stains, bleed-through ink, smeared ink, thin strokes and degraded image respectively and the corresponding cleaned images for each of these categories are given by figures 3, 5, 7, 9, 11, 13 and 15 respectively.

Figure 2. Creases

Figure 3. Cleaned crease noise

Figure 4. Faint text

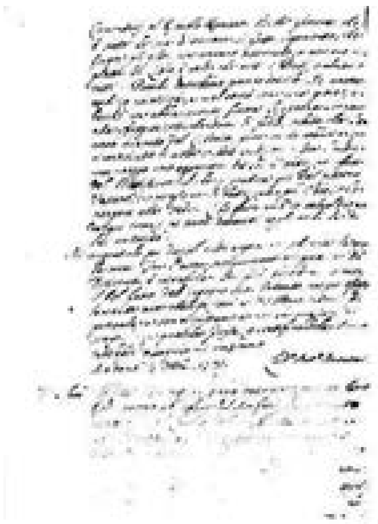


Figure 5. Cleaned faint text noise


Figure 6. Stained image

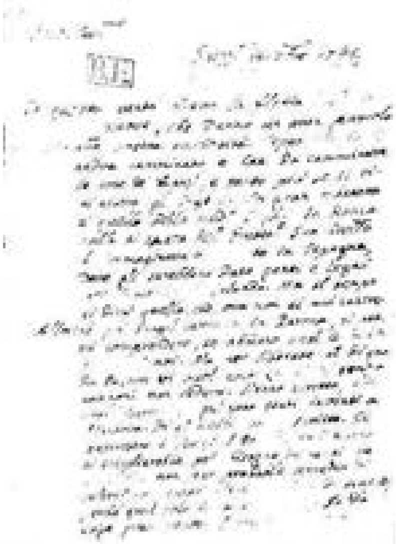


Figure 7. Stains cleaned from the image


Figure 8. Bleed through ink image


Figure 9. Cleaned bleed-through noise


Figure 10. Smeared ink


Figure 11. Enhanced image after removing smeared ink

Figure 12. Thin strokes image

Figure 13. Cleaned thin strokes image

Figure 14. Degraded document image

Figure 15. Enhanced document image

By applying the pipeline of the proposed binarization technique the resulting clean images are obtained for the different input historic documents of 7 categories based on the noises like creases, faint text, stains, bleed-through ink, smeared ink, thin strokes, and degraded image. The performance of the proposed binarization technique is evaluated using the Flesch Reading Ease score. The Average readability score for all output images is calculated. Table 1 summarizes the resultant average readability scores for various types of noisy images on application of the proposed binarization technique for historic document images.

Table 1. The Resultant Average Readability Score for Different Noisy Images

S. No.	Type of noise	No. of Images	Average Readability Score for the output image
1	Creases	150	95.78
2	Faint text	100	76.33
3	Stain	320	86
4	Bleed through ink	200	88.33
5	Smeared ink	200	93.64
6	Thin stroke	150	77.93
7	Degraded document	400	96.44
Average readability score			**87.77**

V. CONCLUSION

The current work uses a combination of various noise reduction and binarization approaches to improve the binarization of historic document images. For improved augmentation of the historical document images, a dependable and efficient technique has been found through the sequential application of unsharp filtering, bilateral filtering, k-means clustering, histogram analysis, Otsu thresholding, pixel modification, and final image creation. In addition to improving the noise reduction process, this comprehensive approach offers insightful information for future initiatives aimed at conserving and sharing cultural heritage through digitally preserved historical records. According to the proposed algorithm, the output image readability score is found to be higher than the original. As mentioned, the algorithm works best for degraded document images having an average readability score 96.44, as compared to other noises. The overall readability of the noisy historic documents after enhancement is 87.77.

Future scope comprises on natural language processing to intelligently digitize historic document collection. Additionally, future research on deep learning techniques, complex machine learning, semantic analysis integration, and natural language processing will all be included in the effort to intelligently digitize historical document collections. Additionally, the future scope includes creating an intuitive user interface to facilitate experimentation with the restoration of historic document images. This algorithm is efficient with diverse noise types in images but it should be tested on more images.

REFERENCES

[1] W. Ma, H. Zhang, L. Jin, S. Wu, J. Wang, and Y. Wang, "Joint Layout Analysis, Character Detection and Recognition for Historical Document Digitization," in Proceedings of International Conference on Frontiers in Handwriting Recognition, ICFHR, Institute of Electrical and Electronics Engineers Inc., Sep. 2020, pp. 31–36.

[2] R. Alaasam, B. Kurar, and J. El-Sana, "Layout analysis on challenging historical Arabic manuscripts using siamese network," in Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, IEEE Computer Society, Sep. 2019, pp. 738–742.

[3] G. Ganchimeg, "History Document Image Background Noise and Removal Methods," International Journal of Knowledge Content Development & Technology, vol. 5, no. 2, Dec. 2015, pp. 11–24.

[4] C. Tensmeyer and T. Martinez, "Historical Document Image Binarization: A Review," SN Computer Science, vol. 1, no. 3. Springer, May 01, 2020.

[5] Y. Akbari, A. S. Britto, S. Al-Maadeed, and L. S. Oliveira, "Binarization of degraded document images using convolutional neural networks based on predicted two-channel images," in Proceedings of the International Conference on Document Analysis and Recognition, ICDAR, IEEE Computer Society, Sep. 2019, pp. 973–978.

[6] K. Ntirogiannis, B. Gatos, and I. Pratikakis, "Performance evaluation methodology for historical document image binarization," IEEE Transactions on Image Processing, vol. 22, no. 2, 2013, pp. 595–609.

[7] A. Nortje, W. Brink, H. A. Engelbrecht, and H. Kamper, "BINet: A binary inpainting network for deep patch-based image compression," Signal Processing: Image Communication, vol. 92, 2013, p. 116-119.

[8] A. Faranhmand, A. Sarrafzadeh, and J. Shanbehzadeh, "Noise Removal and Binarization of Scanned Document Images Using Clustering of Features," International MultiConference of Engineers and Computer Scientists (IMECS), 2017, pp. 410-424.

[9] B. M. Singh, K. K. Verma, and Mridula, "Noise Removal Technique for Document Images," Journal of Multimedia Information System, vol. 10, no. 1, Mar. 2023, pp. 1–14.

[10] E. Kavallieratou and H. Antonopoulou, "Cleaning and Enhancing Historical Document Images," International Conference on Advanced Concepts for Intelligent Vision Systems, 2005, pp. 681-688.

[11] M. Almeida, R. D. Lins, R. Bernardino, D. Jesus, and B. Lima, "A new binarization algorithm for historical documents," Journal of Imaging, vol. 4, no. 2, 2018.

[12] Kaggle Dataset, [Online]. Available at: https://www.kaggle.com/datasets/vpippi/lam-dataset/data

[13] P. Daniel, R. Raju, and G. Neelima, "Image Segmentation by using Histogram Thresholding," IJCSET, vol. 2, no. 1, 2012.

[14] J. Flusser, S. Farokhi, C. Höschl, T. Suk, B. Zitová and M. Pedone, "Recognition of Images Degraded by Gaussian Blur," in IEEE Transactions on Image Processing, vol. 25, no. 2, Feb. 2016, pp. 790-806.

[15] C. C. Pham and J. W. Jeon, "Efficient image sharpening and denoising using adaptive guided image filtering," IET Image Processing, vol. 9, no. 1, Jan. 2015, pp. 71–79.

[16] S. Paris, P. Kornprobst, J. Tumblin, and F. Durand, "Bilateral filtering: Theory and applications," Foundations and Trends in Computer Graphics and Vision, vol. 4, no. 1, 2009, pp. 1–73.

[17] S. M. Burney and H. Tariq, "K-Means Cluster Analysis for Image Segmentation," International Journal of Computer Applications, vol. 96, no. 4, 2014.

[18] N. Dhanachandra, K. Manglem, and Y. J. Chanu, "Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm," in Procedia Computer Science, vol. 54, 2015, pp. 764–771.

[19] S. M. Elbayoumi Harb, N. A. M. Isa, and S. A. Salamah, "Improved image magnification algorithm based on Otsu thresholding," Computers and Electrical Engineering, vol. 46, 2015, pp. 338–355.

[20] X. Xu, S. Xu, L. Jin, and E. Song, "Characteristic analysis of Otsu threshold and its applications," Pattern Recognition Letters, vol. 32, no. 7, May 2011, pp. 956–961.

[21] S. D. Thepade and A. P. Bhalerao, "Machine Learning based Land Use Identification of Aerial Images with Fusion of Thepade SBTC and Triangle Thresholding," 2023 2nd International Conference for Innovation in Technology (INOCON), Bangalore, India, 2023, pp. 1-7.

[22] Z. Shi, Y. Chen, E. Gavves, P. Mettes, and C. G. M. Snoek, "Unsharp mask guided filtering," IEEE Transactions on Image Processing, vol. 30, 2021, pp. 7472–7485.

[23] textstat library documentation [Online]. Available at: https://pypi.org/project/textstat/

[24] S. Samer & A. Arafat, "Blank Background Image Lossless Compression Technique," International Journal of Image Processing (IJIP), 2014, pp. 9-16.

[25] A. Sekertekin, "A Survey on Global Thresholding Methods for Mapping Open Water Body Using Sentinel-2 Satellite Imagery and Normalized Difference Water Index," Archives of Computational Methods in Engineering, 2021, pp. 1335–1347.

A Phishing URL Detection Model based on Horse Herd Optimization and Random Forest Algorithms

P. Hemannth
Computer Science and Engineering Department
Velagapudi Ramakrishna Siddhartha Engineering College
Andhra Pradesh, India
hemannth2112@gmail.com

Mukesh Chinta
Computer Science and Engineering Department
Velagapudi Ramakrishna Siddhartha Engineering College
Andhra Pradesh, India
mukeshchinta2502@gmail.com

S. Sarat Satya
Computer Science and Engineering Department
Velagapudi Ramakrishna Siddhartha Engineering College
Andhra Pradesh, India
surapanenenisaratsatya@gmail.com

P. Sri Aneelaja Devasena
Computer Science and Engineering Department
Velagapudi Ramakrishna Siddhartha Engineering College
Andhra Pradesh, India
paneelajadevasena@gmail.com

Abstract— **Phishing attacks continue to be a notable threat to network and information security. They plan to expose user information and privacy, such as login credentials, passwords, credit card numbers, and other details, by tricking internet users into thinking they are the real deal. For the detection of phishing websites, Machine learning (ML) techniques have been progressively used, because of their abilities like to learn from and adapt to complex designs and features. A new approach for detecting phishing websites using ML techniques is proposed that incorporates the URL structure. The Horse Herd Optimisation Algorithm is used to determine the features, and the suggested method is tested on a dataset of websites with phishing threats. In the context of network and information security, these techniques are employed to identify websites with phishing threats. The objectives include collecting a new dataset, extracting pertinent features, and addressing the challenges of imbalanced data and adversarial attacks in phishing detection. The findings can assist security professionals and researchers in identifying the techniques that are best suitable for improving phishing detection and prevention.**

Keywords— Phishing, Ransomware, Horse Herd Algorithm, Detection, Cybersecurity, Random Forest.

I.INTRODUCTION

Machine learning is a branch of artificial intelligence (AI) that uses data and algorithms to mimic human learning and increase the ability of systems in prediction tasks. It enables software programmers to create more accurate predictive models. Machine learning can be classified into three types: supervised, unsupervised learning, and reinforcement learning.

Horse Herd Optimization (HHO) is an algorithm that mimics the natural herding behavior of horses. HHO algorithm utilizes the hierarchical structure and the collective motion of a horse herd to perform feature selection, which is a crucial step for machine learning problems.

In HHO, the features of a dataset are considered as potential solutions, and the algorithm aims to find the most optimal subset of features that contribute significantly to the problem at hand. The algorithm iteratively explores the solution space by dividing the features into leaders and followers.

Initially, a set of randomly selected features are assigned as leaders, and the rest of the features become followers. The leaders guide the search process by influencing the followers through a combination of exploration and exploitation strategies. This allows for a balanced exploration of the feature space while exploiting the potential of promising features.

During each iteration, the leaders compete for dominance based on their fitness. The fitness of each leader feature is evaluated using a fitness function for feature selection that measures the quality of the selected features. The leaders update their positions based on their fitness, while the followers adapt their positions to the leaders' movements.

Through the interaction and coordination among the leaders and followers, HHO algorithm gradually converges towards an optimal subset of features that best represents the data and contributes to accurate classification or regression models.

The HHO algorithm offers a nature-inspired approach to efficiently identify relevant features and reduce the dimensionality of the dataset. By leveraging the collective intelligence of a horse herd, this algorithm shows promise in enhancing the performance and interpretability of machine learning models.

To produce accurate predictions or classifications, Random Forest, a powerful machine learning technique, combines the advantages of using several decision trees. First, Random Forest constructs multiple decision trees. Next, each of these decision trees are trained using a diverse range of characteristics and data.

During training, Random Forest generates a collection of decision trees through a process called bagging, which involves sampling the dataset with replacement. This sampling ensures diversity in the trees, allowing them to capture different aspects of the data.

The algorithm develops predictions through integrating every prediction of all the ensemble's decision trees. The final class label for classification tasks is determined by voting or averaging, but for regression tasks, the predicted values from each tree are summed.

Random Forest offers several advantages. It is robust against overfitting and noise in the data, thanks to the ensemble approach and the random feature selection. It can handle large datasets with high dimensionality, as well as categorical and numerical features. Additionally, it provides estimates of feature importance, allowing for interpretability and insight into the data.

SCOPE: The project's limited scope is focused on feature selection and testing utilizing the present algorithms to increase the precision in phishing website identification.

OBJECTIVES:

1. To collect a data set that best suits for the project.

2. To extract features that help in detection of phishing websites.

3. To use Horse Herd Optimization Algorithm to select the features.

4. To use Random Forest Algorithm to find the precision and accuracy in classification of websites.

5. To overcome the challenges encountered by recent techniques in the detection of the phishing URL.

The rest of this article is as follows: section 2 - a literature review of related work, section 3 - a methodology of the proposed approach, section 4 - an evaluation and discussion of the approach, section 5 - a summary and future directions, and section 6 - a list of references.

II. RELATED WORKS

This section presents a review of the literature that informs the references used in this study.

In this research, [1] introduced a machine learning (ML) model for the identification of websites with phishing threats. The features used to train this model included statistical reports, IP addresses, Google, subdomains, web traffic, domain pages, prefixes and suffixes, domain registrations, URL lengths, DNS records, and CPR rankings. These articles developed a model using a variety of algorithms, including support vector machines, black trees, logistic regression, linear discriminant analysis, K- nearest neighbors, and random forests, and they ultimately achieved a random forest algorithm accuracy of 98.90%.

Advantages: Higher accuracy in detection of phishing websites when compared to other algorithms.

Disadvantages: Inadequate detection techniques

In this research, [2] introduced a GAN (generative adversarial networks) model for identifying phishing websites. In order to make a distinction between legal and phishing URLs, this model employs CNN (Convolutional neural network) as a discriminator. LSTM (long short-term memory network) is used to create fake phishing URLs with characteristics that resemble real phishing URLs. The results of the research show that, the PDGAN model is able to detect objects with an accuracy of 97.58% and a precision of 98.02% without the aid of other services.

Advantages: This mentioned model is independent of o any third-party services. It detects the websites used for phishing by evaluating the various characteristics of the URL of the website. This model does not require manual extraction of the features.

Disadvantages: One of the main disadvantages of GAN model is that it is very difficult to train. This model categorizes the websites as used for phishing or not, purely basing its reasoning on the characteristics of the URLs.

In this research [3] a ML model on phishing website detection is presented. They made advantage of characteristics that includes URL length, the presence of HTTP, suspicious characters in the URL, prefixes and suffixes, the number of dots and slashes, the phishing terms, the length of the subdomain, and the IP address in the URL to train their models. They used two algorithms named 1) random forest algorithm 2) decision tree algorithm in which the RF model gave a accuracy of 97.25% for legitimate URL and 97.29% for phishing URL and DT model gave a accuracy of 95.90% for legitimate URL and 95.91% for phishing URL.

Advantages: Higher accuracy in detection of phishing websites.

Disadvantages: Addition of the redundant and useless features can degrade the model performance. The data set taken in this model is outdated.

In this research [4], different machine learning techniques were contrasted, and a model that produced the best results was recommended. In order to pick features for this project, they detailed about 21 characteristics and used feature selection approaches like the gain ratio attribute evaluator (GainRatioAttributeEval), one R attribute evaluator (using the one R classifier), relief attribute evaluator, and symmetric uncertainty attribute evaluator. Utilizing the random forest technique after feature selection, they achieved a very high accuracy.

Advantages: Features are selected based on the feature selection algorithms.

Disadvantages: Inadequate data collection since it is not possible to produce a model with 100 percent accuracy.

In this research [5], a multilayered stacked ensemble learning model was presented for the detection and identification of websites with phishing intent. This model has three levels - which are stacked in the order - layer-i, layer-ii, and meta. Classifiers like XGBoost (XGB), Logistic Regression (LR), Random Forest classifier (RF), MLP, and the k- Nearest Neighbour (KNN) classifier are present in the top layer. The second layer, which include classifiers like XGB, RF, and MLP, receives the output of the first layer. The second layer's output is then transferred into the meta layer, which contains the XGB algorithm. The trained model is assessed using the datasets listed in datasets, and the trained model attained a detection rate with an accuracy range from 96.79% to 98.90%.

Advantages: The outputs of several learning algorithms can be combined, varied features can be extracted, and greater predicted performance is the end result.

Disadvantages: It is very difficult to create a model with ensemble techniques

In this research [6] tells us about the model in which the feature selection is done by a feature selection method known as PSO (particle swarm optimization). At first, they extract some of the features which they think are important in detection of phishing website. After that, they use PSO

algorithm to weigh the selected features. The features with most weight are used in the model training, whereas the features with less weight are ignored. At last, the machine learning techniques' performance before applying the PSO and after applying the PSO are compared. From the observation, we can conclude that the accuracy of the models when applied with PSO algorithms are better when compared to models without PSO algorithm.
Advantages: Useless features are eliminated with the help of PSO algorithm that can degrade the performance.
Disadvantages: Due to its simplicity in establishing local optimums in high-dimensional space and its sluggish rate of convergence during repetitive procedures, the particle swarm optimization (PSO) algorithm has some disadvantages.

The authors of this study [7] proposed a parameter-free similarity measure called the Normalized Compression distance (NCD), that compresses two webpages to ascertain how similar they are to one another. This allows a approach for identifying websites without features created with the purpose of phishing from the original ones. The extraction of any features is not necessary for this procedure. With this method, websites are categorized based on how much their HTML resembles well-known phishing websites. They used the farthest point first technique to extract phishing prototypes and select those instances that are normal of a group of phishing web sites. This paper received an AUC of 98.68.
Advantages: The main benefit of these model is that they do not need feature extraction to be developed. Rather than creating new features in the presence of idea or concept drift, this approach employs an incremental learning method as the framework that enables persistent and flexible detection.
Disadvantages: It is unable to recognize new phishing website variations or zero-day attacks.

The writers of this research [8] conducted a poll and had a discussion on the methods that can be used to identify phishing websites. The methods they discussed are Random Forest, XGBoost (XGB), Support Vector machines, K-Nearest Neighbors (KNN), Decision Tree, Gradient Boosting, Logistic Regression. Additionally, they talked about the feature selection techniques of filtering and wrapping. They also mentioned about the evaluation metrics of these models before concluding.
Advantages: The main advantage of these paper is that it discusses about all the available solutions present for the detection of the phishing websites.
Disadvantages: The main disadvantage of these paper is that it does not provide a new solution to the existing problem rather it only provides the existing solutions.

III. PROPOSED METHOD

This section describes the architecture of the model, methodology of the model, algorithm of the model discussed in this paper.

A. Architecture:

For developing this model, First, we collect a data set that best suits the project. The data set is collected in such a way that the features are already extracted in the data set. After data collection, we preprocess the data set. Preprocessing including removing duplicate values, converting categorical values into numerical values, removing the uniform resource locators for which uniform resource locator length is less than 8. After preprocessing the data, we extract the important features using optimization algorithm. After feature selection, we divide the data set into two sets -training and testing data set and train the model using training model. After training, we test the model using testing dataset and evaluate the performance various using performance metrics. Figure 1 shows the proposed system

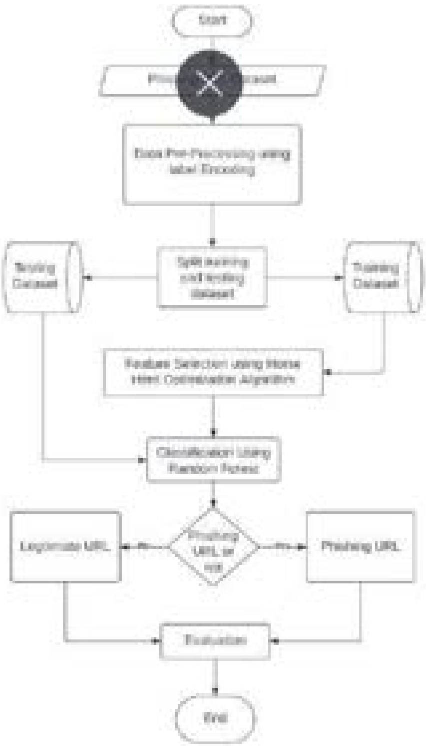


Figure 1: Proposed System

B. Methodology:

In this research paper, the methodology section consists of the following modules: Data collection, Data pre-processing, Feature selection, Model development and Evaluation and performance metrics. Each module is described in detail below:

1. Data Collection:

The dataset was retrieved from the Kaggle dataset, which has 87 features and 11430 URLs. This dataset is used to test and compare how well the machine learning methods can find phishing websites. The features are broken down into three categories: seven of the features are retrieved through contacting external services. The 56 remaining features taken based on the structure and syntax of URLs. 50% dataset is made up of both authentic and phishing URLs, hence ensuring it is evenly dispersed.

2. Pre-processing Module:

The crucial stage for any machine learning model is pre-processing. In this stage, the data that is present in the dataset is converted so that can be used to train and test machine learning models.

The pre-processing steps include: Converting categorial data into numerical data using One Hot Encode. Removing duplicate values and the records that contain the content of email and page's information. Removing URL's that of the size less than or equal to 8. Feature selection from the dataset is done with the help of Horse Herd Optimization Algorithm

3. Model Selection and Development.

In this study, we adopted Random Forest algorithm for the detection of phishing websites. After feature selection using Horse Herd Optimization Algorithm, the training of the model is done with training dataset which is 80% of the original dataset and tested with testing dataset which is 20% of the original dataset. Both training and testing datasets contains class labels in the form of numerical data i.e.: 1's(phishing) and 0's(legitimate). The results of the testing is illustrated below.

4. Evaluation metrics

The performance of the ML-based model is assessed using different metrics, such as confusion matrix, classification report, accuracy, precision, recall, f1 score, roc-auc curve etc. This model takes approximately 5 to 10 minutes to train, depending on the selected features. The model achieves an accuracy of 97% in predicting the class labels for the test data.

4.1 Confusion Matrix:

Confusion matrix can be considered as a 3×3 table that summarizes the performance of a ML model on a set of test data. It shows the number of correct and incorrect predictions for each class, and helps to measure the accuracy, precision, recall and other metrics of the model. A confusion matrix shows in figure 2 reveal the weaknesses and strengths of the model, and identify the classes that are most confused by the model.

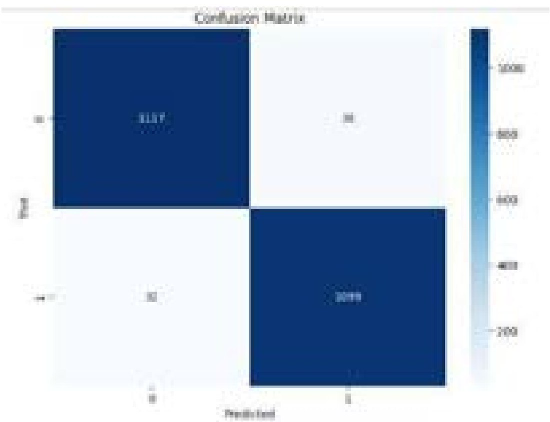


Figure 2: The Confusion Matrix obtained for this model

4.2 Classification Report:

A classification report is a summary of the classification quality of a ML model. It displays the number of true and false predictions for each class and the whole model. It also computes some metrics to assess the model performance such as precision, f1-score, recall, support, etc

classification report	precision	recall	f1-score	support
0.0	0.97	0.97	0.97	1149
1.0	0.97	0.97	0.97	1137
accuracy			0.97	2286
macro avg	0.97	0.97	0.97	2286
weighted avg	0.97	0.97	0.97	2286

Figure 3: Classification Report

The classification report for this model is represented in Figure 3, which displays the precision, recall, F1 score, and support for each class label.

4.2.1 Precision:

Precision is a metric that quantifies the ratio of correctly classified positive samples among all classified positive samples made by a machine learning model. It is calculated as the ratio of true positives to the sum of true positives samples and false positives samples. A high precision implies a low rate of false positives or type I errors for the model. The precision value of this model is around 0.97 which suggests that it is very good at correctly identifying positive instances without mistakenly labelling negative instances as positive

4.2.2 Recall:

Recall is a metric that evaluate the proportion of true positive samples out of the total number of actual positive samples. It is calculated as the ratio of true positives samples to the sum of true positives samples and false negatives samples. A high recall implies a low rate of false negatives or type II errors for the model. The recall value of this model is around 0.97, which implies that the model identifies 97% of all the positive instances.

4.2.3 F1 score:

F1 score can be considered as the combination of precision and recall to measure the accuracy and performance of a machine learning model. It is especially useful for imbalanced datasets, where accuracy alone might not be a reliable indicator of the model's quality. The F1 score is also around 0.97 which implies the model is robust and reliable.

4.3 ROC-AUC Curve:

The ROC-AUC curve is a graphical measure of the performance of a binary classification model across different decision thresholds. It plots the true positive rate against the false positive rate for various values of the threshold. It also quantifies the overall quality of the model by computing the area under the curve, which represents the probability that the model ranks a positive instance higher than a negative one.

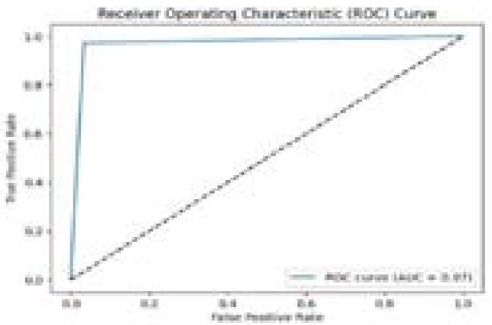


Figure 4: ROC_AUC curve

Figure 4 shows ROC curve for the trained model. The curve shows high TPR and a low FPR, indicating an good model performance. The AUC value is 0.97, which means the model can accurately distinguish between the classes. This suggests that the model has a high distinguishing power as the AUC value is closer to 1.

C. Algorithm:

<u>*Horse Herd Optimization Algorithm:*</u>

<u>*Input*</u>*:* Dataset X with N samples and D features, Maximum number of generations G, Number of horses NH, Number of selected features K.

<u>*Output*</u>*:* Selected features for best horse H_{best}.

<u>*Steps:*</u>

1*:* Initialize horse herd: H_{herd} = {H1, H2, ..., HNH}.

2: Initialize best horse: H_{best} = None, best fitness = ∞.

3: for generation = 1 to G do

4: for i = 1 to NH do

5: Randomly initialize horse Hi with K features.

6: Evaluate fitness for horse Hi using a feature evaluation metric.

7: if the fitness of Hi is better than best fitness then

8: Update best fitness with the fitness of Hi.

9: Update H_{best} with the features of Hi.

10: end if

11: end for

12: Select top p horses from H_{herd} based on fitness.

13: Perform crossover and mutation on the selected horses to create offspring.

14: Replace the worst p horses in H_{herd} with the offspring.

15: end for

16: Return selected features of the best horse H_{best}.

The Horse Herd Optimization Algorithm was chosen for efficient feature selection, leveraging its collective intelligence to identify relevant URL structure characteristics.

<u>*Random Forest Algorithm:*</u>

<u>*Input*</u>: Training dataset D with legitimate and phishing URLs, Number of trees T, Number of features F, Number of instances per tree N.

<u>*Output*</u>: Collection of decision trees RF.

<u>*Steps:*</u>

1: Input the training dataset D, containing both phishing and legitimate URLs.

2: Define T which is the number of trees to create in the random forest model.

3: Define the number of features F to consider at each split in the decision trees.

4: Define the number of instances per tree N to randomly sample from D with replacement.

5: Initialize an empty collection of decision trees RF.

6: for each tree t in the range 1 to T do

7: Randomly sample N instances from D with replacement.

8: Randomly select F features from the total set of features.

9: Train a decision tree using the sampled instances and features.

10: Add the decision tree to the collection RF.

11: end for

12: Return the collection of decision trees RF.

13: To classify a new URL u as legitimate or phishing, do the following:

14: Pass u through each decision tree in RF, and get the decision of each tree.

15: Aggregate the decisions of all trees by taking the majority vote.

16: if the majority vote is legitimate then

17: Classify u as legitimate.

18: else

19: Classify u as phishing.

20: end if

IV. RESULT ANALYSIS

This research paper consists of four modules: preprocessing, feature selection, classification and performance evaluation. The preprocessing module converts the categorical values in the 'status' column into numerical values using encoding techniques, and updates the Data Frame accordingly. The feature selection module applies the Horse Herd Optimization algorithm to select different subsets of features from the original features set as herds. Three subsets are then used to train the classification model. The classification module employs the Random Forest algorithm to predict the class of a given URL, either legitimate or phishing. The model achieves an accuracy of 97.1%, with a training time of approximately 12 minutes and a prediction time of around 0.5 seconds. The precision, recall, and f1-score metrics for the model are also reported using a classification report which are around the value of 0.97. The performance evaluation module plots the ROC curve for the model, which shows and AUC of 0.97. These values imply that the model is robust in nature and is reliable.

V. CONCLUSION AND FUTURE WORK

The Primary focus has been on the selection of optimized features using the innovative Horse Herd Algorithm. The primary scope of the project is to improve the efficiency and accuracy of feature selection processes in complex datasets, with a particular emphasis on applications related to phishing website detection. The project employs the Horse Herd Algorithm, an optimization technique inspired by the behaviour of horse herds, to tackle the feature selection challenge. This approach capitalizes on the collective intelligence and dynamic interactions observed in horse herds, translating them into a computational framework. By simulating the herding behaviour, the algorithm is adept at exploring the solution space and identifying the most useful and relevant features that contribute to accurate phishing detection. This innovative approach showcases the potential for nature-inspired algorithms to revolutionize the field of feature selection. We got an accuracy of 97.1 percent with respect to different sets of features. The accuracy may be increased upon continuous working on different set of features. The effectiveness of the detection model was validated through rigorous testing using internal and external datasets. Performance metrics were evaluated under various scenarios Optimizing algorithm parameters and exploring ensemble methods can enhance detection accuracy. Additionally, incorporating additional features for improved discrimination between phishing and legitimate URLs may further boost performance. Continuous refinement and validation against real-world data are crucial for ongoing performance enhancement. To ensure security, rigorous data handling protocols were followed, including anonymization of sensitive information and validation checks. During model training, measures were taken to prevent overfitting, and performance was evaluated against benchmarks, enhancing the reliability of our phishing URL detection system.

By highlighting the complexities inherent in evolving URL structures and phishing tactics, objectives aim to overcome these challenges. Through the utilization of the Horse Herd Optimization Algorithm and Random Forest Algorithm, the study endeavors to enhance detection accuracy. By emphasizing the collection of a tailored dataset, feature extraction, and tackling challenges like data imbalance and adversarial attacks, the approach aims to advance phishing URL detection systems' efficacy against malicious threats.

Further in future our project can be extended in trying to detect whether the emails sent to users are phishing or legitimate based on the links, contents, data present in the emails. The phishing websites can be detected with the comparison of the content of the web page with the original contents by analysing the patterns and characteristics of known phishing emails and legitimate correspondence, the system can develop a more sophisticated understanding of potential threats.

VI. REFERENCES

[1] I. Kara, M. Ok and A. Ozaday, "characteristics of understanding urls and domain names features: the detection of phishing websites with machine learning methods," in IEEE access, vol. 10, pp. 124420-124428, 2022. DOI: 10.1109/ACCESS.2022.3223111

[2] S. Al-ahmadi, A. Alotaibi and O. Alsaleh, "PDGAN: phishing detection with generative adversarial networks," in IEEE access, vol. 10, pp. 42459-42468, 2022. DOI: 10.1109/ACCESS.2022.3168235

[3] Atharva Deshpande, Omkar Pedamkar, Nachiket Chaudhary, Dr. Swapna Borde, 2021, detection of phishing websites using machine learning, International Journal of Engineering Research \& Technology (IJERT) volume 10, issue 05 (may 2021). DOI: 10.17577/IJERTV10IS050235

[4] Mehanović, D., Kevrić, J. (2020). Phishing website detection using machine learning classifiers optimized by feature selection. Traitement du signal, vol. 37, no. 4, pp. 563-569.

[5] L. R. Kalabarige, R. S. Rao, A. Abraham and L. A. Gabralla, "multilayer stacked ensemble learning model to detect phishing websites," in IEEE access, vol. 10, pp. 79543-79552, 2022. DOI: 10.1109/ACCESS.2022.3194672

[6] W. Ali and S. Malebary, "particle swarm optimization-based feature weighting for improving intelligent phishing website detection," in IEEE access, vol. 8, pp. 116766-116780, 2020. DOI: 10.1109/ACCESS.2020.3003569

[7] R. W. Purwanto, A. Pal, A. Blair and S. Jha, "PhishSim: Aiding Phishing Website Detection with a Feature-Free Tool," in IEEE Transactions on Information Forensics and Security, vol. 17, pp. 1497-1512, 2022. DOI: 10.1109/TIFS.2022.3164212

[8] R. Zieni, L. Massari and M. C. Calzarossa, "Phishing or Not Phishing? A Survey on the Detection of Phishing Websites," in IEEE Access, vol. 11, pp. 18499-18519, 2023. DOI: 10.1109/ACCESS.2023.3247135

[9] D. Dua and C. Graff. UCI machine learning repository. School of information and computer science, university of california, irvine, CA, USA. Accessed: jan. 10, 2020. [Online]. Available: https://archive. Ics.Uci.Edu/ml/datasets/phishing+websites

[10] S. Parekh, D. Parikh, S. Kotak and S. Sankhe, "A New Method for Detection of Phishing Websites: URL Detection," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 2018, pp. 949-952, DOI: 10.1109/ICICCT.2018.8473085.

[11] M. Sameen, K. Han and S. O. Hwang, "PhishHaven—An Efficient Real-Time AI Phishing URLs Detection System," in IEEE Access, vol. 8, pp. 83425-83443, 2020, DOI: 10.1109/ACCESS.2020.2991403.

[12] M. Sánchez-Paniagua, E. F. Fernández, E. Alegre, W. Al-Nabki and V. González-Castro, "Phishing URL Detection: A Real-Case Scenario Through Login URLs," in IEEE Access, vol. 10, pp. 42949-42960, 2022, doi: 10.1109/ACCESS.2022.3168681.

Pedal Powered Renewable Energy System

Prof Lokesh Khedekar
Department of Engineering
Sciences and Humanity
Vishwakarma Institute
of Technology
4) Pune, India
lokesh.khedekar@vit.edu

Aseem Bhide
Department Of Engineering
Sciences and Humanity
Vishwakarma Institute
of technology
Pune, India
aseem.bhide23@vit.edu

Nakul Chandak
Department Of Engineering
Sciences and Humanity
Vishwakarma Institute
of technology
Pune, India
nakul.chandak23@vit.edu

Atharv Bharadiya
Department Of Engineering
Sciences and Humanity
Vishwakarma Institute
of technology
Pune, India
atharv.bharadiya23@vit.edu

Yashraj Bodhale
Department of Engineering
Sciences and Humanity
Vishwakarma Institute
of Technology
Pune, India
yashraj.bodhale23@vit.edu

Yash Chalke
Department of Engineering
Sciences and Humanity
Vishwakarma Institute
of Technology
Pune, India
yash.chalke23@vit.edu

***Abstract*- The demand for electrical energy continues to rise steadily, particularly in developing countries like India. This surge in demand is attributed to the country's ongoing development, leading to a significant annual increase in power requirements. However, there exists a substantial disparity between electricity generation and this escalating demand. While increasing generation capacity seems like a plausible solution, it is hindered by formidable challenges such as high capital investments and logistical complexities. In light of these constraints, a promising alternative emerges: the adoption of energy-saving measures at the individual level, coupled with the utilization of non-conventional energy sources such as wind, solar, and tidal power. This approach not only helps bridge the demand-supply gap but also holds the potential for significant economic benefits, particularly for developing nations like India. Voltage control can be achieved by electronic control above a minimum set limit. Noise is because of chain & sprocket. It could be reduced by timer belt transmission by incorporating the gear mechanism and hence increasing the RPM of the prime mover**

Keywords —Low Rotations Per Minute Alternator, Green Energy, Cost Effective

1 INTRODUCTION

IMPORTANCE OF LOW RPM ALTERNATOR

An alternator serves as an electromechanical apparatus that transforms mechanical energy into electrical energy, specifically in the form of alternating current. The fundamental principle behind alternators involves the induction of electromotive force (EMF) in a onductor when there is a change in the surrounding magnetic field. Typically, a rotor, which is a rotating magnet, turns within a stationary set of conductors wound in coils on an iron core known as the stator. As the rotor turns due to a mechanical input, the magnetic magnet, turns within a stationary set of conductors wound in coils on an iron core known as the stator. As the rotor turns due to a mechanical input, the magnetic field intersects with the conductors, creating an induced EMF. Low-speed alternators can easily connect to small power sources like wind turbines, tidal turbines, micro-hydropower plants, and exercise cycles. Additionally, they can be installed atop [8] micro-turbines along water canals.

2 LITERATURE REVIEW

[1] Due to the escalating demand for power in India, there is a pressing need to generate power economically using non-conventional energy resources. This objective can be effectively realized through the utilization of [2] small generators operating within low and controllable speed ranges. The employment of low RPM alternators proves to be particularly beneficial at both [3] household and small commercial levels.

The alternator offers a [4] high degree of design flexibility, allowing for easy adaptation to varying output requirements without significant parameter adjustments.[5] Furthermore, their operation at low speeds translates to minimal maintenance requirements. Additionally, [12] their cost-effectiveness is enhanced by the utilization of non-conventional prime movers. The necessary raw materials [11] for constructing such alternators, including mild steel, magnets, copper conductors, [6] paper insulators, and bearings, are readily available in the market. Moreover, their assembly is straightforward.

The project can be broadly divided into two main components: the alternator and the prime mover assembly. The alternator comprises [7] two basic parts: the stator and the rotor. For demonstration purposes, a bicycle [9] has been chosen as the prime mover.

The prime mover assembly includes the bicycle itself, alternator stands, cycle stands, chain and sprocket arrangements, [10]and a base plate with a tension adjustment system.

TABLE 1. COMPARATIVE STUDY

Paper Title	**Method Used**	**Advantages of Method Used**	**Limitations**
Bicycle-Powered AC Generator	Bicycle connected to AC generator	Sustainable, portable energy source	Limited power output, requires physical effort
Generation of Electricity through Bicycle and Solar Energy	Bicycle connected to generator powered by solar panels	Renewable energy sources combined	Dependency on sunlight for solar energy generation
Bicycle Power Generator	Bicycle connected to DC generator	Simple mechanism, potentially low cost	Limited power output, requires physical effort
Pedal Power Generation	Pedals connected to generator	Exercise benefits, potential for off-grid power generation	Limited power output, requires continuous pedaling
Experimental study on the effect of alternator speed to the car charging system	Alternator speed variation for car charging	Optimization of car charging system	Specific to car charging system, limited applicability
Low Speed Alternator Design	Design optimization for low-speed alternator	Efficient power generation at low speeds	May not be suitable for high-speed applications
Generation of Electrical Power using Bicycle Pedal	Pedals connected to generator	Human-powered energy generation	Limited power output, requires continuous pedaling
Back-up Power Generation using Bicycle	Bicycle connected to backup generator	Emergency power supply	Requires physical effort, limited power output
Power Generation Using Bicycle Mechanism as an Alternative Energy Source	Bicycle connected to generator	Sustainable energy source	Limited power output, requires physical effort
Power Generation through Pedaling	Pedals connected to generator	Exercise benefits, potential for off-grid power generation	Limited power output, requires continuous pedaling
Pedal power generation: An implementation of stationary bike generator	Stationary bike connected to generator	Fitness benefits, potential for off-grid power generation	Limited power output, requires continuous pedaling
Bike powered electricity generator	Bicycle connected to generator	Portable energy source, potentially low cost	Limited power output, requires physical effort

3 PROBLEM STATEMENT

The basic aim of this project is to produce electrical energy by harnessing sustainable energy resources. This brings about saving in terms of usage of conventional sources on a personalized and domestic level.This project involves building up of a low RPM alternator by using materials such as magnets, M.S.(Mild steel) ,C.I.(Carbon Iron), copper wires etc.

4 METHODOLOGY

An alternator converts mechanical energy into electrical energy, producing alternating current (AC). It operates on the principle that changing magnetic fields induce electromotive force (EMF) in conductors. A rotor rotates within a fixed set of conductors (the stator), generating EMF as the magnetic field intersects with the conductors. Low rpm alternators are easily coupled with small prime movers like wind turbines and can be mounted over micro-turbines in water canals. They can generate AC voltage at frequencies above 40 Hz with minimal ripple, making them suitable for battery charging without affecting battery health. These alternators are primarily used for charging batteries, which can then power various appliances, ensuring safety due to low voltage output. A general representation is shown in figure 1

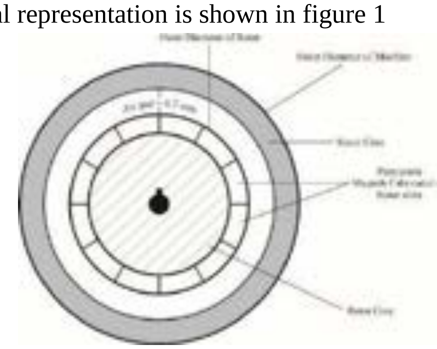


Figure 1. Alternator

a) Stator- The stator is stationary part of the alternator. It is a cylindrical structure, built up of dynamo grade laminations. The laminations are 0.5 mm thick. The laminated segments are held together by axial key bar or bright bar. Fitting into parallel sided slots in the outer ring of the core.

b) Stator Frame- They fulfill several functions, including:

- Providing containment for the core and windings
- Serving as a ventilating enclosure
- Acting as a protective barrier, shielding live and moving parts from human contact"

c) Rotor-

The rotor, made of a hollow Mild Steel (M.S.) cylinder, holds the magnets. Thin M.S. plates at both ends support the shaft, with keyholes for locking. Cooling is by natural air convection.

d) Magnets- High gauss strength permanent rare earth magnets are used for field excitation, placed on the rotor cylinder with alternating polarities, ensuring strong surface polarity.

e) Air Gap- The air gap, a critical factor in electrical machine design, affects performance. It must be minimized to reduce flux reluctance but large enough to prevent stator-rotor rubbing during operation or bearing failure.

f) Windings- This is a Single Phase - single Layer winding, where conductors in stator slots produce EMF. They can be arranged in series to increase voltage or in parallel to double current and halve voltage.

g) Bearings-

Horizontal shaft machines use deep groove ball bearings for anti-friction, while vertical shaft machines use thrust bearings. Bearings can be ball or roller type. In horizontal shaft machines, forces mainly act radially.

5. DESIGN

C.1 Sator Design- Stator laminations of low rpm alternator are made up of silicon steel with thickness of around 0.5 mm. The stator has 36 slots which accommodates 18 coils. The coils are connected in series.

Slots Dimensions of the Stator:

No. Of slots	:	36
Slot Depth	:	21.6 mm
Maximum Slot width	:	10 mm
Minimum Slot width	:	3.6 mm
Pole Width	:	12 mm
Slot Area	:	136 mm^2

Stator Dimensions:

Inner Diameter of stator	:	179.8 mm
Outer Diameter of Stator	:	258.0 mm

C.2 Rotor Assembly

a)Rotor Cylinder- The rotor is the rotating part of the machine on which the magnets are placed.

The cooling is provided to the rotor shown in figure 2 by natural convection of air.

Rotor Specifications:

Rotor Outer Diameter	:	170.16 mm
Rotor Inner Diameter	:	152.00 mm
Rotor Thickness	:	9.8 mm
Rotor Material	:	M.S.

FIGURE 2. OUTPUT DESIGN

b) End Shields- End shields, composed of mild steel (M.S.), are secured by four studs, firmly clamping the stator stack in place between them.

C.3 Bearings- Bearings reduce friction between the generator shaft and end ring. Spherical roller ball bearings are utilized for this purpose, shielded to prevent dust contamination. Bearing selection considers various factors:

1. Basic static load (C_0)

Basic static load (C0) represents the load causing permanent deformation at the bearing contact point, reaching 0.01% of the ball's diameter

2. Basic dynamic load (C)

This load is the point at which the bearing's nominal lifetime will achieve one million revolutions.

3. Life time

The life time of bearing is the number of revolutions that the bearing can accomplish before the first sign of fatigue begin to appear on ring or rolling component.

C.4 Mounting- A low RPM alternator typically follows the IM 1001 (IM B3) standard, featuring a horizontal shaft and floor-mounted feet. In certain instances, it can be vertically mounted with a vertical shaft.

C.5 Terminal Box- Made of pressed steel, it houses the alternator's terminals, located on the right-hand side when viewed from the non-drive end as shown in figure 3. Featuring

IP55 protection, it offers cable entry provisions and includes a terminal bar made of molded epoxy resin.

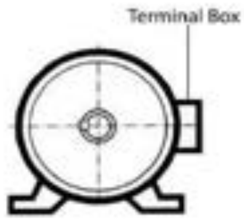


Figure 3. Terminal Box

C.6 Shaft design-
It is made up of carbon steel. The shaft is provided with a keyway to ensure locking of shaft with rotor. The keyway in the shaft is designed to fit the key perfectly.
Shaft length: 220 mm
Shaft diameter: 25 mm

C.7 Magnet- The average gauss value measured is 155 kGauss. As the machine is designed for 18 poles, 18 magnets are employed on the rotor.

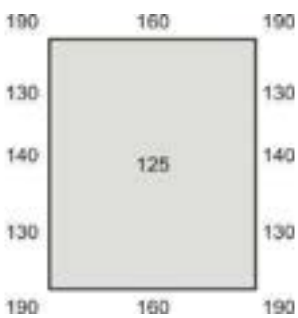


Figure 4

C.8 Cooling- IC 01 signifies 'Free Convection', utilizing temperature variations for coolant flow, with negligible rotor-induced fanning.
'Self Circulation' (denoted by '1') relies on the main machine's rotational speed, facilitated by rotor action or specialized component design, such as direct mounting or mechanical drive by the main rotor.

1) Design Parameters and Construction Data

A) Basic Assumptions

1)Rating of Prime mover: Trials & measurements were taken on exercise bicycle, measurements are as below;

- Torque developed in Kgm = 16 kg X 0.15 m at 120 rpm
  Torque developed = 2.4 Kgm
- Power developed in KW= T X rpm = 0.295 Kw (approx 300 w)

2) Efficiency for Prime mover & alternator were assumed as 80 % & 70% respectively.
Thus output power of the alternator = 300 X 0.8 X 0.7 = 170 watts
= 170 VA
3)Availability of stator stampings is 36 slots/ 48 slots. Hence 36 slots stamping is selected considering 2 slots per pole. Therefore no. of poles are 18.

B) Stator Design

The design parameters are as follows-

1. Outer diameter : 258 mm
2. Internal diameter : 179.8 mm
3. Slot no. : 36 slots
4. Slot area : 136 mm^2
5. Slot depth : 21.6 mm
6. Slot width bottom : 10mm
7. Slot width face : 8 mm
8. Core length : 40mm
9. Tooth top distance : 3.6 mm
10. Tooth top length : 12mm
11. Tooth width (parallel) : 8.4 mm
12. Stamping width : 39mm

C) ROTOR DESIGN

The design parameters are as follows-

1. Outer diameter of ring : 169 mm
2. Internal diameter of ring : 152 mm
3. Width of ring : 43 mm
4. Width of end plates : 2.5 mm
5. Outer diameter of end plates : 179 mm
6. Inner diameter of end plates : 25mm
7. Key slot size on end plates : 8*8 mm

D) Shaft Design

The design parameters are as follows-

1. Material : M. S.
2. Shaft diameter : 25mm
3. Shaft steps : 6 No.
4. Key way length : 46mm
5. Key way width : 8 mm

E) Winding

Winding used for this alternator is series, single phase, single layer winding with SWG = 18. The connection diagram is shown below. In this scheme 2 slots constitute 1 pole.

2) CALCULATIONS

Table 1 shows the preliminary estimation of conductor

SWG of Copper Wire	Diameter (mm)	Area (mm^2)	No. of slots
20	0.914	0.6567	150
18	1.220	1.1680	90
16	1.630	2.0760	45
14	2.030	3.2430	30

A) *Symbols Used*

Ø: Magnetic flux (possibly peak flux)
kG: Kilogauss (unit of magnetic flux density)
wb/m^2: Weber per square meter (alternative unit of magnetic flux density)
Q: Reactive power (in kVA)
C0: Constant
D: Wire diameter
L: Length
Ns: Number of turns
Bav: Average magnetic flux density
p: Number of poles
Iz: Current in winding
Z: Impedance
V: Voltage (RMS)
W: Power (W)
cos Ø: Power factor

B) *Calculations for Winding*

i) Length of mean turn= (200 mm) * (No. of turns) * (No. of coils) **Total length**= 108 m
= 0.108 km

ii) Resistance of total winding= 1.5174 Ω at 20^{O} C.

iii) Current density: 4 Amp/mm3 **(As per electrical design datasheet)**

iv) Total weight of winding = (10.379 kg/km) * (0.108 km)
= 1.12 kg

v) $Ø = 154$ kG=15.4 wb/m^2

$Q = C_0D^2LN_s$
= 0.3 kVA

Q= 0.3/0.7= 0.428

Therefore,

$0.428 = C_0 * (0.18)^2 * 0.04 * 5.555$

$C_0 = 41.7083 * 10^3$

But $C_0 = 11B_{av}*ac*K_w*10^{-3}$

B_{av}=(p Ø)/(pi D L)

ac= (I_z Z)/(pi D)

Z =30

C_0=11 * (p Ø)/(pi D L) * (I_z Z)/(pi D)

$41.7083 * 10^3 = (11*18*15.4* Iz*30 *10^{-3})/(3.141^2*.180^2*0.04)$

I_z= 5.83 A

W=VI cos Ø

200= V * 5.83 * 0.7

Therefore,

V= 49 V

Hence the output voltage is **49 v**olts.

3) *TESTING*

A) *Testing Alternator:*

- Voltage Output:

Sr. No.	Speed of rotor(rpm)	Voltage (Volts)
1	150	21
2	200	30
3	250	35
4	300	49
5	350	55
6	400	67

- ***Coil Resistance:***

Sr. No.	Coil No.	Resistance(Ω)
1	1-18	0.09 PER COIL

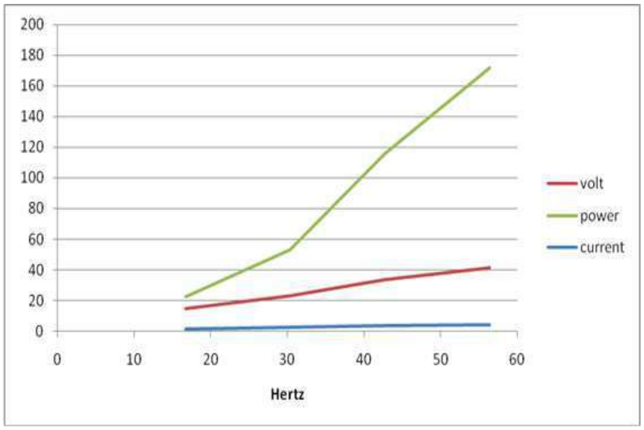
Total Resistance of stator winding: 1.65 Ω

Winding Impedance: 9.863 Ω

B) *Load Testing:*

Load: Resistive
Total Resistance: 10 Ohms

Load	Voltage	Hz	Power	Current
Ohm	V (V)	F (Hz)	$W = V^2 / R$	I = V / R
10	14.91	16.7	22.23	1.49
10	22.99	30.3	52.85	2.30
10	33.97	42.7	115.40	3.40
10	41.44	56.4	171.73	4.14



`

Figure 4 Performance Analysis

4) *Advantages:*

1. Energy can be harnessed at low speeds.
2. Output can be increased i.e. rating be increased by replacing magnets with higher flux densities.
3. Design is done with assumption of least speed of prime mover.
4. Increased voltage can be controlled by using electronic devices e.g. SMPS (Switched Mode Power Supplies).
5. Rectification can be done at any frequency.
6. The noise level is low due to low rpm.

5) *Limitations*

1. Frequency control depends on prime mover.
2. Field excitation cannot be controlled.
3. Voltage regulation depends only on design parameters.

6) Energy saving calculations:

200 watt * 24 hrs * 365 days = 1752 kWhrs generated per year

1752 kWhr * 5.10 Rs. / kWhrs = Rs. 8917 Saving of electricity

Charges by conventional energy sources

Cost of Generator = Rs. 8000

Consider wind turbine as a prime mover for 24 hrs * 365 days.

Cost of wind turbine is approximately Rs. 10000

Therefore,

Cost of prime mover = Rs. 10000

Cost of Alternator = Rs. 8000

Total = Rs. 18000

Payback period = Cost of installation / saving per year

= 18000 / 8917

Payback period = 2.01 years.

Estimated life of low rpm alternator = 10 years

Total gain from this alternator in its life cycle is = 8 years * Rs. 8917

= Rs. 71,336 (as per prevailing rate)

Payback period for this alternator is 11 months only.

CONCLUSION & FUTURE SCOPE

This study has highlighted the issue of increasing demand for electrical energy in countries like India, amidst challenges in meeting this demand due to constraints in traditional generation methods. This approach not only addresses the demand-supply gap but also holds potential economic benefits, especially for developing nations. Furthermore, technical solutions are proposed for voltage control and noise reduction in energy systems, emphasizing the importance of innovative technologies in addressing energy challenges. Overall, the findings highlight the significance of a multifaceted approach encompassing both policy interventions and technological innovations to meet the escalating energy demands sustainably.

In the near future, sustainable energy is poised to become the most cost-effective electricity source, possibly already holding this status. As fossil fuel reserves dwindle, their prices are likely to sharply rise, making them economically unsustainable. Key technological advancements needed for the commercialization of sustainable power have been made, with ongoing improvements expected. A rising trend is the development of small-scale sustainable energy systems for personal use, gaining traction due to their economic and environmental benefits.

REFERENCES

[1] A. D. I. L. G. S. E. F. J. B. Ron Edward Guia, "Pedal Power Generation: An Implementation of Stationary Bike Generator," *LPU-Laguna Journal of Engineering and Computer Studies,* pp. 63-67, 2020.

[2] A. D. I. L. G. S. E. F. J. B. Ron Edward Guia, "Pedal Power Generation: An Implementation of Stationary Bike Generator," *LPU-Laguna Journal of Engineering and Computer Studies,* pp. 63-67, 2020.

[3] A. U. R. V. ŞTEFAN MOCANU1, "Bike-powered Electricity Generation," *Asia Pacific Journal of Multidisciplinary Research, Vol. 3, No. 1, February 2015,* pp. 32-41, 2015.

[4] Y. C. Scott Merrick, "Low Speed Alternator," CORE, San Luis Obispo, 2013.

[5] A. K. T. M. A. R. k. A. N. D. S Manish Yadav, "Power Generation Using Bicycle Mechanism as an Alternative Energy Source," *International Research Journal of Engineering and Technology (IRJET),* pp. 809-816, 2018.

[6] *. R. M. D. M. Z. Z. A. H. A. H. Rozdman K. Mazlan1, "Rozdman K. Mazlan1,*, Reduan M. Dan1,2, Mohd Z. Zakaria1,2, Abdul H. A. Hamid," in *MATEC Web of Conferences*, Melaka, 2017.

[7] A. D. I. L. G. S. E. F. J. B. Ron Edward Guia, "Pedal Power Generation: An Implementation of Stationary Bike Generator," *LPU-Laguna Journal of Engineering and Computer Studies,* pp. 63-67, 2020.

[8] M. C. K. K. K. S. A. S. Rajneesh Suhalka1, "Generation OF Electrical Power Using Bicycle Pedal," *International Journal of Recent Research and Review, Vol.VII, Issue 2, June 2014,* pp. 63-67, 2014.

[9] P. S. V. R. M. P. R. K. Rajesh Kannan Megalingam, "Pedal Power Generation," in *IEEE Conference*, Kollam, 2012.

[10] M. U. Q. 2. A. A. K. a. S. A. Md Sajid1, "Power Generation Through Pedaling," *Grenze International Journal of Engineering and Technology,,* pp. 281-284, 2018.

[11] O. P. S. P. S. L. G. L. Mayur Bhabad1, "Bicycle Power Generator," *International Research Journal of Engineering and Technology (IRJET),* pp. 1-07, 2021.

[12] K. K. N. S. K. ,. Manish Paul, "Back-up Power Generation using Bicycle," *International Advanced Research Journal in Science, Engineering and Technology,* pp. 73-75, 2017.

[13] A. Austin, "Bicycle-Powered AC Generator," Andrew Austin, San Luis Obispo, 2011.

[14] D. V. K. Anoop Kumar1, "Generation of Electricity through Bicycle and Solar Energy," *International Journal of Electrical Engineering and Technology.,* pp. 44-58, 2019.

Micro-expression Recognition using Generative Adversarial Network-based Convolutional Neural Network

Krishna Santosh Naidana
Computer Science and Engineering Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
krishna.santosh10@gmail.com

Lakshmi Prasanna Divvela
Computer Science and Engineering Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
lakshmiprasanna.divvela@gmail.com

Yaswanth Yarra
Computer Science and Engineering Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada, India
yaswanthyarra7112@gmail.com

***Abstract*— Computer vision is leveraged to tackle the challenge of recognizing micro-expressions with brief facial movements for the revealing of genuine emotions. As micro-expressions are brief and have low intensity, they are hard to capture and frequently go unnoticed. Current techniques for micro-expression recognition (MER) suffer from low classification accuracy due to the limited dataset. This study addresses the challenge of Micro-expression recognition (MER) by proposing a novel approach that combines Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs). Leveraging the CASME II dataset, GANs are first employed to synthesize realistic micro-expressions, effectively augmenting the dataset with diverse samples. Subsequently, a CNN model is trained on this augmented dataset to classify micro-expressions into seven predefined categories. Evaluation metrics including accuracy, precision, f1-score, support, and recall are used to assess model performance. Our proposed GAN-CNN approach outperforms existing CNN models for MER and achieves competitive performance compared to state-of-the-art methods. The integration of GANs with CNNs demonstrates promising potential for accurately recognizing micro-expressions, overcoming the limitations of dataset scarcity, and improving classification accuracy.**

Keywords*—*Convolutional Neural Networks, Micro-expression Recognition, Facial Movements, Generative Adversarial Networks, Data Augmentation.

I. Introduction

Micro-expressions are momentary and involuntary facial expressions that offer brief insights into a person's authentic emotions, even when they are consciously attempting to hide or suppress them. These subtle movements, lasting for a fraction of a second, typically ranging from 1/25th to 1/2nd of a second, can serve as crucial indicators of an individual's genuine feelings. The motivation behind the proposed work stems from its diverse applications, which encompass criminal interrogation, assessing the credibility of an interviewee, and determining the true emotional state of a patient. [1]. Because of the spontaneously induced nature of micro-expressions, only a limited number of well-established datasets have been introduced. One widely utilized dataset in this context is the CASME II[2] dataset. Researchers have recently become interested in the computational automatic analysis of micro-expressions. CNNs stand out in image recognition due to their specialized architecture, which automatically learns features from raw pixel data, reducing the need for manual feature engineering. Leveraging parameter sharing and local connectivity, CNNs efficiently extract features from images, resulting in improved performance and reduced computational complexity compared to other deep learning methods. This makes CNNs the preferred choice for many computer vision applications.

CNNs are replacing the manual feature engineering process by automatically identifying the best feature representations for images. However, CNNs often yield lower accuracies due to the limited sizes of available micro-expression datasets. In order to overcome this limitation, we incorporated a deep learning model called Generative Adversarial Network (GAN) that consists of a generator and a discriminator that has been trained adversarially. The generator creates realistic data, while the discriminator evaluates and distinguishes between generated and real data, fostering a continuous improvement loop to augment the dataset size. GANs enhance the dataset by creating new micro-expression samples. Subsequently, The classification of facial micro-expressions into seven broad classes—happiness, sadness, disgust, suppression, fear, surprise, and other was done using CNNs. Utilizing GANs to enhance the dataset would, in turn, improve the classification performance by overcoming the limited dataset problem.

The proposed model outperforms other state-of-the- art CNN models. In the comprehensive framework for micro-expression MER, this study initiates the procedure by preprocessing the apex frames of the micro-expression. Subsequently, we employ GAN to augment these apex frames, we train and test the CNN model using the augmented dataset.

The structure of the paper is as follows: Section II provides a review of literature, Section III outlines the methods employed, Section IV details the experiments conducted, Section V presents the results and discussion, and Section VI concludes the study and suggests future research directions.

II. LITERATURE REVIEW

Researchers have proposed a method that use conditional GAN and an optimized compact CNN model [3]. The authors presented an enhanced continuous labelling method for identifying a video's apex frame. Conditional GAN are used to augment the training data. The network is optimized for mobile based micro-expression analysis using techniques like weight pruning, architecture search etc. They achieved an accuracy of 78.53%, 78.53% on CASME II and SMIC datasets. The OC-Net is designed to be compact and optimized for mobile based micro-expression analysis with fewer parameters and faster computation time compared to other deep learning models. There is a decrease in the accuracy for SAMM dataset.

The methodology they proposed in this [4] is for synthesizing long video datasets of micro-expressions using reference-guided style transfer. The authors first detected the Action Units (AUs) in the original dataset, which are then used as reference points for the style transfer process. They used GANs to generate synthetic videos that mimic the facial expressions of the original dataset. After comparing the AUs obtained in the original and synthetic datasets, they performed a correlation study and discovered strong correlation on two AUs, AU12 and AU6, with Pearson's correlation values of
0.74 and 0.72, respectively. The primary benefit of the suggested approach is its ability to produce lengthy video sequences of facial expressions, including both macro- and micro-expressions, an accomplishment that is difficult to do with current techniques. In addition, the technique generates new data without altering the identification of the original film, which can be advantageous for maintaining anonymity.

Researchers have introduced an approach called ICE-GAN (Identity-aware and Capsule-Enhanced Generative Adversarial Network) for expression recognition and synthesis [5]. ICE-GAN uses graph-based reasoning to capture long-ranged dependencies, i.e., the relationships between different parts of a facial expression that are not immediately adjacent to each other. The ICE-GAN methodology is evaluated using the Leave-One-Subject-Out (LOSO) validation method. They achieved an F1-score of 87.4%. Graph-based reasoning allows for more comprehensive understanding of the relationships between different facial regions. ICE-GAN is an identity- aware approach, which means that it can produce synthetic faces with controllable micro-expressions and identity-aware features. The ICE-GAN's reliance on graph-based reasoning may introduce computational complexity.

In this paper, they examined the impact of dense and sparse optical flow methods and their effect on MER. The data sample size is increased by using GAN to generate "fake" optical flow images [6]. Finally, they implemented a state-of- the-art CNN architecture and modify it for feature enhancement and classification operations. On the SMIC, CASME II, and SAMM datasets, the accuracy attained is 71.43%, 84.85%, and 68.75%, respectively. The use of GAN-based data augmentation can increase the diversity and sample size of the dataset, which can increase the generalization ability of the model.

In this paper, they proposed an approach is a discriminatively deep fusion approach with an improved conditional generative adversarial network (im-cGAN) for facial expression recognition (FER) [7]. To produce additional labelled expression examples for training, the im-cGAN is employed. The method extracts discriminative features from training data by combining global and local characteristics. The obtained accuracy on JAFFE, CK+, Oulu-CASIA, and KDEF datasets is 98.37%, 98.20%, 93.34%, and 98.30%, respectively. The performance of im-cGAN is dependent on choice of hyperparameters and architecture. Depending on the dataset and application, there may be a risk of overfitting if the im- cGAN-generated samples do not accurately represent the diversity of real-world facial expressions.

The two main strategies used in the methodology suggested in this paper are Décalcomanie data augmentation and N-step pre-training [8]. Multiple transfers of learning from action recognition datasets to facial recognition datasets are carried out throughout the N-step pre-training. A deep neural network is trained on a sizable dataset as part of the pre-training process, and it is then fine-tuned using the target micro- expression dataset. Using décalcomanie data augmentation, which is predicated on facial symmetry, a composite image is produced by circling the centre lines of both faces. They achieved the Weighted Average Recall (WAR) as 0.7927. The multiple transfers of learning and fine-tuning enhance the model's ability to generalize across datasets.
performance.

Some researchers have recently proposed a dual-stream approach that combines dynamic image convolutional neural networks with optical flow. Histogram equilization is used for image enhancement [9]. In order to categorise micro-expressions, the algorithm first creates a number of attributes. Then OF-Block module is proposed to extract flow features from video data, which can capture the motion and movement of facial muscles over time. Then hierarchical fusion strategy that fuses face motion features and spatio-temporal features several times to protect diversity of features. Accuracy of 61.20%, 73.09% and 58.07% is achieved in the CASME and CASME II and SAMM datasets, respectively. The proposed algorithm uses two streams of information, optical flow and dynamic images to capture both the motion and appearance of facial expressions. Despite its feature-rich approach, the algorithm achieves accuracy levels that might be considered relatively modest.

They suggested Deep3DCANN, an end-to-end framework for detecting face micro-expressions, in this [10]. Deep 3D CNN, the initial element, learns practical spatiotemporal features from a series of facial photos. A deep artificial neural network is the second element, and it tracks the beneficial visual connections between various facial subregions. To forecast the micro-expressions, a fusion technique combines learnt facial traits with semantic links between the areas. Regarding the SAMM, CASME, and CASME II datasets, the 3DCANN framework achieves accuracy of 87, 86, and 93 percent, respectively. Deep 3D CNNs and deep ANNs increases computational complexity that result in longer training times and extra resources like GPU.

III METHODOLOGY

The suggested system's architecture and the used methodology are covered in this section.

3.1 Architecture:

The overview of the proposed framework is described in Figure 1, outlining both the data augmentation module and the classification module.

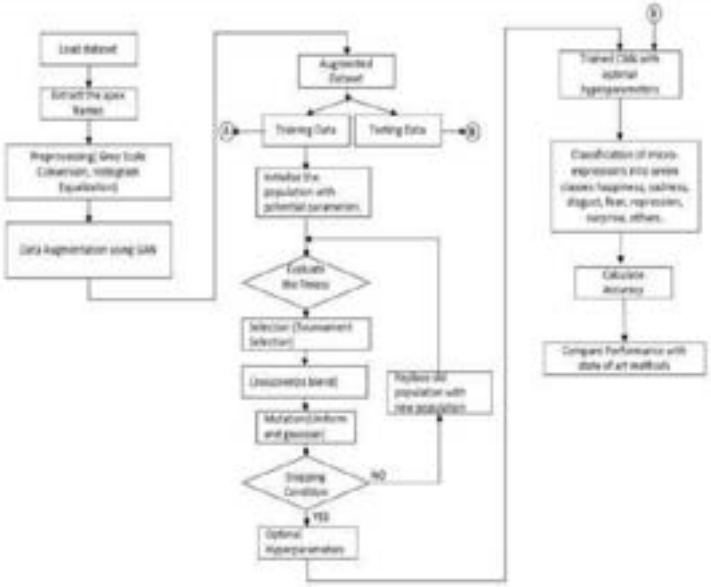


Fig. 1. Proposed System for GAN-CNN model

3.2 Methodology:

An all-encompassing framework for micro-expression recognition, with CNN and GAN as its main building blocks is presented in this research study. The micro-expression (ME) sequences' apex frames are first recognized and preprocessed by the framework. Subsequently, the dataset is augmented through GAN, and CNN is employed for classification.

3.2.1 Preprocessing

In the study of facial micro-expressions and emotions, the identification of apex frames is crucial for analyzing and comprehending micro-expressions. Apex frames are representative of the highest point of intensity in a facial expression, while onset frames indicate the initial stages of an expression where it becomes apparent. The CASME II dataset, which is a prevalent database of spontaneous micro-expressions, provides annotations of apex frame for each micro-expression sequence. This allows to extract corresponding apex frame from each expression sequence of each subject, thereby facilitating accurate analysis of micro-expressions.

Following this, the apex frames undergo grayscale processing, a procedure that transforms a colored image into a grayscale image. It allows for simpler and more efficient processing of the images, as grayscale images have only one channel of information per pixel, as opposed to the three channels (red, green, and blue) found in color images. By calculating the weighted average of the red, green, and blue values of each pixel, it transforms the input color image into grayscale, Equation 1 represents grayscale value calculation.

$$Grayscale\ value = 0.299 * R + 0.587 * G + 0.114 * B \quad (1)$$

The red, green, and blue values of the pixel are denoted as R, G, and B, correspondingly. The coefficients 0.299, 0.587, and 0.114 are the weights that are used to compute the weighted average.

The images are next subjected to histogram equalisation, which is the process of converting the original image's histogram into a uniformly distributed histogram. By redistributing the pixel intensity values, this technique reduces the effect of illumination and improves the image's overall contrast.

3.2.2 Generative Adversarial Networks

GANs play a role in generating novel micro-expression samples [11]. The data augmentation is achieved using GANs. Among the essential elements are the Discriminator (D) and Generator (G) functions. P_{data} is the training dataset's probability distribution, and P_z is the latent space's probability distribution, which is usually a random Gaussian distribution. The adversarial dynamics of GANs revolve around the interaction between these two functions.

When a sample (x) is drawn from P_{data}, the Discriminator's goal is to classify it as a real micro-expression sample, effectively distinguishing between real micro-expressions and those generated by the model. Conversely, the Generator aims to produce micro-expression samples (G(z)) from the latent space in a manner that, when presented to the Discriminator, they are perceived as real. The Generator continuously refines its process to create samples that increasingly resemble the real micro-expression samples, as the Discriminator's capacity to distinguish between produced and genuine samples increases.

For generated samples, the generator minimizes the log likelihood that the discriminator makes correct classifications. This loss incentivizes the generator to create samples that the discriminator is likely to categorize as real (logD(G(zi)) close to 1). The process is shown in figure 2.

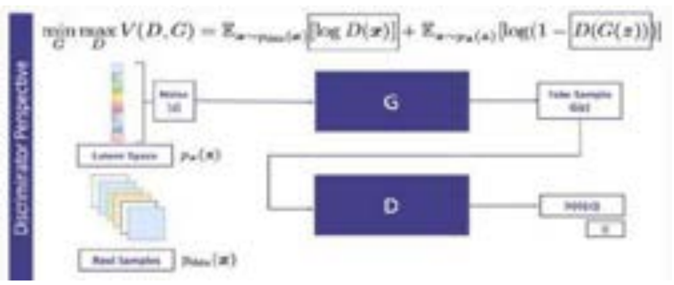


Fig. 2. Objective function in GAN formulation [12]

Generator: The generator network takes random noise or a latent representation as input and attempts to generate realistic data samples, such as images, audio, or text. It learns to map the input noise to the data distribution of the training set, effectively generating new data instances that are indistinguishable from real examples. The goal of the generator is to produce outputs that are convincing enough to fool the discriminator into believing they are real.

Discriminator: The discriminator network acts as a binary classifier, distinguishing between real and generated data samples. It is trained to differentiate between authentic data from the training set and synthetic data produced by the generator. The discriminator learns to assign high probabilities to real data and low probabilities to generated data. Its objective is to accurately classify the origin of the input samples, thereby providing feedback to the generator on how to improve its outputs.

3.2.3 Convolutional Neural Networks

In Figure 3, the proposed CNN architecture incorporates three pairs of convolutional and MaxPooling layers is illustrated. Convolutional layers are particularly effective in detecting patterns and features within images, while max pooling layers contribute to reducing the spatial dimensionality of the output from the convolutional layers, introducing translation invariance.

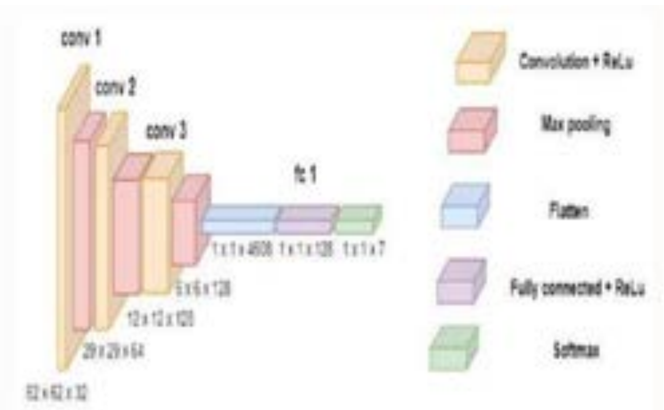


Fig. 3. Architecture of proposed CNN

The first convolutional layer uses the ReLU activation function and consists of 32 filters with a 3x3 kernel size. The input images have the shape (64, 64, 1), where the height and breadth (64 pixels each) of the input images are represented by the first two dimensions, and the number of channels is indicated by the final dimension. A MaxPooling layer with a pool size of 2x2 follows this layer. The subsequent convolutional layer features 64 filters with a 3x3 size and employs the ReLU activation function. Another MaxPooling layer with a pool size of 2x2 succeeds this layer. Following this, a third convolutional layer is introduced, incorporating 128 filters with a 3x3 size, along with a MaxPooling layer with a pool size of 2x2.

To mitigate overfitting, dropout regularization is applied after the flatten layer. Following the convolutional and pooling layers, the flattened feature maps are connected to fully connected (Dense) layers. These layers integrate the extracted features and learn complex relationships between them. In this architecture, a fully connected layer with 128 units and ReLU activation is employed. This layer acts as a bridge between the convolutional layers and the output layer, facilitating high-level feature representation and abstraction. The output layer utilizes a SoftMax activation function to produce probability distributions over the designated classes for each input image. This enables the model to make predictions by assigning probabilities to each class, indicating the likelihood of the input image belonging to a particular category.

IV EXPERIMENTAL ANALYSIS

4.1 Experimental Setup:

In the proposed architecture, the GAN model augments the dataset by generating images of size [64, 64, 1]. The enhanced dataset is then employed to train the CNN model using an 80:20 stratified split and tested accordingly to evaluate the model's performance. We used the Adam optimizer with a learning rate of 0.0006 during the optimisation process. The model was trained for 52 epochs with a batch size of 17 and a dropout ratio of 0.29. These hyperparameters were fine-tuned using a genetic algorithm. Google colab was utilized to run the model.

4.2 Datasets:

The experiments utilized the CASME II dataset provided by the Fu Xiaolan research group of the Chinese Academy of Sciences with permission. It is a comprehensive repository of spontaneous micro-expression that comprises 247 video samples drawn from 26 Asian participants, with an average age of 22.03 years. The videos depict participants evoked by one of five micro-expression categories: Happiness, Disgust, Repression, Surprise, and Others. Recorded at a high frame rate of 200 fps, this dataset offers a diverse and dynamic collection of micro-expressions. Table 1 illustrates the counts of original samples, GAN-augmented samples, and the total number of samples post-augmentation.

Table 1. displays the count of original and augmented frames in the CASME II dataset.

Type	Original	GAN augmented	Total
Happiness	32	64	96
Sadness	7	63	70
Repression	27	64	91
Disgust	63	-	63
Surprise	25	50	75
Fear	2	60	62
Others	99	-	99
Total	255	301	556

The below images in figure 4 show the generated images of GAN.

Fig. 4. Generated images using GAN

4.3 Evaluation Metrics:

Using a variety of metrics, including accuracy, precision, recall, f1-score, and UAR, the performance of the GAN-CNN model was evaluated on the CASME II dataset.

V RESULTS AND DISCUSSION

In conducting the facial micro-expression classification project using Google Colab, minimal system requirements include a stable internet connection and a compatible web browser such as Google Chrome or Mozilla Firefox. Software requirements entail access to a Google Colab account, freely available, and the installation of essential Python libraries, including TensorFlow, Keras, OpenCV, NumPy, and Matplotlib. Additional libraries may be necessary for tasks such as data preprocessing, model training, and evaluation. This section provides a thorough analysis of the experimental results derived from the benchmark CASME II dataset. We delve into a detailed discussion of these results, conducting a comparative study that assesses the performance of the proposed GAN-CNN model against baseline models such as VGG16, ResNet50, and InceptionV3. These baseline models are considered as reference points. Additionally, we compare our work with state-of-the-art methods in the realm of micro- expression classification. Table 2 illustrates the comparative performance analysis between the GAN-CNN model and the baseline models.

Table 2. Comparison of classification performance of GAN-CNN with baseline models.

Metrics	**Resnet50**	**VGG16**	**Inception V3**	**GAN-CNN**
Accuracy	0.8986	0.7707	0.8295	0.9222
Recall	0.8986	0.7707	0.8295	0.9222
F1-Score	0.9015	0.7732	0.8314	0.9229
Precision	0.9154	0.8076	0.8464	0.9260
UAR	0.9091	0.7721	0.8277	0.9399

The assessment revealed significant enhancements in the performance of the GAN-CNN model. Specifically, it demonstrated a 2.36% increase in classification accuracy compared to ResNet50, a 15.15% improvement compared to VGG16, and a 9.27% higher accuracy compared to the InceptionV3 baseline model. The proposed framework achieved the highest Unweighted Average Recall (UAR) at 93.99%, surpassing the performance of the other models under evaluation. Consistent trends were observed across various metrics, underscoring the superior performance of the proposed framework. Figure 5 illustrates the variation in training and validation accuracies for the proposed GAN-CNN employed in the experiments.

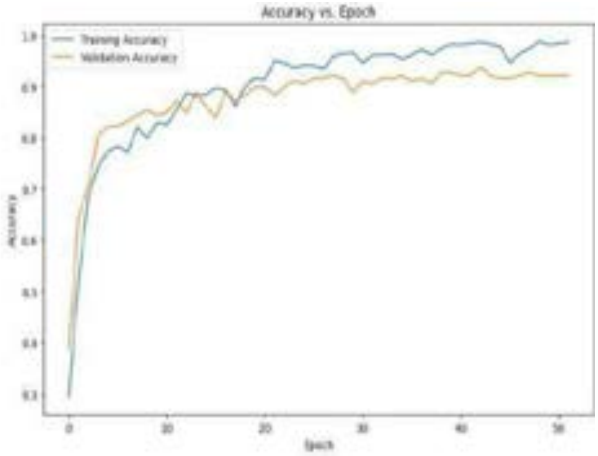


Fig. 5. Accuracy vs Epoch graph

Figure 6 present the confusion matrix of our proposed model. Upon analyzing the provided confusion matrix, it becomes apparent that the GAN-CNN model exhibits superior class- wise micro-expression recognition performance.

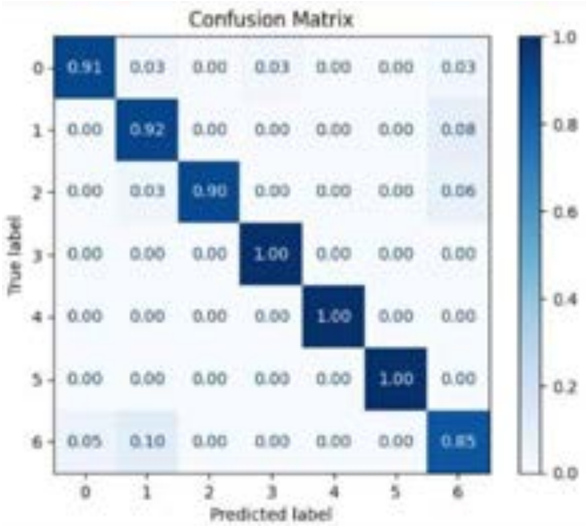


Fig.6: Confusion matrix

We extensively compared the GAN-CNN model with existing micro-expression recognition baselines using CNN architectures and deep features. Table 3 summarizes the evaluation findings, highlighting the overall performance of the proposed model.

Table 3: Comparison table of GAN-CNN with the state-of-art methods

Related work	Method	Dataset	Accuracy	F1-Score	UAR
Sze Teng Liong, 2020	Conditional GAN	CASME II, SAMM, SMIC	0.8485	0.61	-
Koo Sie-Min et al., 2022	Conditional GAN and compact CNN	CASME II, SAMM, SMIC	0.7914	0.69	-
Jianhui Yu et al., 2021	ICE-GAN	CASME II, SAMM, SMIC	-	0.874	0.904
Yan Wang et al., 2022 [13]	Dual-stream CNN	CASME, CASME II, SAMM	0.7309	-	-
Selvarajah Thuseethan et al., 2023	Deep 3DCNN-ANN	CASME II	0.83	-	-
Proposed model	GAN-CNN	CASME II	0.9222	0.9229	0.9399

VI CONCLUSION

In conclusion, this study presents a promising approach to addressing the challenge of recognizing micro-expressions, which are often elusive yet significant indicators of genuine emotions. By leveraging Generative Adversarial Networks (GANs) in conjunction with Convolutional Neural Networks (CNNs), our proposed method offers a novel solution to the limitations posed by the scarcity of micro-expression datasets. Through the synthesis of realistic micro-expressions using GANs and subsequent training of a CNN model on an augmented dataset, the proposed approach has achieved notable improvements in classification accuracy compared to existing techniques. The evaluation results demonstrate the efficacy of our approach, as evidenced by superior performance metrics such as accuracy, precision, f1-score, support, and recall. Overall, the integration of GANs with CNNs showcases promising potential for advancing micro-expression recognition, thereby contributing to the broader field of computer vision and emotion analysis.

REFERENCES

[1] Microexpression. Wikipedia, The Free Encyclopedia. Wikimedia Foundation.

[2] WJ. Yan, X. Li, SJ Wang, G. Zhao, YJ Liu, et al."CASME II: An Improved Spontaneous Micro-Expression Database and the Baseline Evaluation," PloS one, vol. 9, no. 1, pp. e86041, 2014, doi: 10.1371/journal.pone.0086041.

[3] K. Sie-Min, M. A. Zulkifley, and N. A. M. Kamari, "Optimal compact network for micro-expression analysis system," Sensors, vol. 22, no. 11, pp. 4011, 2022.

[4] C. H. Yap, R. Cunningham, A. K. Davison, and M. H. Yap, "Synthesising facial macro- and micro-expressions using reference guided style transfer," Journal of Imaging, vol. 7, no. 8, p. 142, 2021.

[5] J. Yu, C. Zhang, Y. Song, and W. Cai, "ICE-GAN: Identity-Aware and Capsule-Enhanced GAN with Graph-Based Reasoning for Micro-Expression Recognition and Synthesis," in 2021 International Joint Conference on Neural Networks (IJCNN), July 2021, pp. 1-8.

[6] S. T. Liong, Y. S. Gan, D. Zheng, S. M. Li, H. X. Xu, H. Z. Zhang, and K. H. Liu, "Evaluation of the Spatio-Temporal Features and GAN for Micro-Expression Recognition System," Journal of Signal Processing Systems, vol. 92, pp. 705-725, 2020.

[7] Z. Sun, H. Zhang, J. Bai, M. Liu, and Z. Hu, "A Discriminatively Deep Fusion Approach with Improved Conditional GAN (im-cGAN) for Facial Expression Recognition,", *Pattern Recognition*, vol. 135 pp. 109157, 2023.

[8] C. Lee, J. Hong, and H. Jung, "N-Step Pre-Training and Décalcomanie Data Augmentation for Micro-Expression Recognition," Sensors, vol. 22, no. 17, pp. 6671, 2022.

[9] J. Tang, L. Li, M. Tang, and J. Xie, "A Novel Micro-Expression Recognition Algorithm Using Dual-Stream Combining Optical Flow and Dynamic Image Convolutional Neural Networks," Signal Image and Video Processing, pp. 1-8, 2022.

[10] S. Thuseethan, S. Rajasegarar, and J. Yearwood, "Deep3DCANN: A Deep 3DCNN-ANN Framework for Spontaneous Micro-Expression Recognition," Information Sciences, vol. 630, pp. 341-355, 2023.

[11] Generative Adversarial Networks. Wikipedia. https://en.wikipedia.org/wiki/Generative_adversarial_network

[12] Neptune AI. (2023). 6 GAN Architectures You Really Should Know. Neptune AI Blog. https://neptune.ai/blog/6-gan-architectures.

[13] Y. Wang, Y. Huang, C. Liu, X. Gu, D. Yang, S. Wang, and B. Zhang, "Micro expression recognition via dual-stream spatiotemporal attention network," Journal of Healthcare Engineering, 2021.

[14] de Souza, Vinicius Luis Trevisan, Bruno Augusto Dorta Marques, Harlen Costa Batagelo, and João Paulo Gois. "A review on generative adversarial networks for image generation." Computers & Graphics (2023).

Fraud Shield: Credit Card Fraud Detection with Ensemble and Deep Learning

1st Pranav Prakash Menon
Department of Networking and Communications
School of Computing
SRM Institute of Science and Technology
Chennai, India
pm1809@srmist.edu.in

2nd Amit Sachdeva
Department of Networking and Communications
School of Computing
SRM Institute of Science and Technology
Chennai, India
as6261@srmist.edu.in

3rd Gayathri V.M
Department of Networking and Communications
School of Computing
SRM Institute of Science and Technology
Chennai, India
gayathrm@srmist.edu.in

***Abstract*— Credit card fraud remains a persistent challenge, undermining financial security and trust. Traditional fraud detection systems struggle with the dynamic nature of fraudulent transactions, necessitating innovative approaches. Recent advancements introduce promising methodologies but face hurdles like data imbalance and evolving fraud patterns. This research study aims to elevate credit card fraud detection by integrating ensemble and deep learning models. Leveraging techniques like Voting, Stacking, CNN, and LSTM on the IEEE-CIS Fraud Detection dataset, we address challenges such as data imbalance and feature selection inefficiencies. The proposed approach, Fraud Shield, aims to fortify detection capabilities, acknowledging evolving fraud activities. By combining advanced techniques and a comprehensive dataset, our project provides a powerful tool for combating fraud in the financial sector. This study highlights the need for continual innovation in fraud detection and highlights the importance of collaboration between academia, industry, and regulatory bodies. Through comprehensive analysis and collaboration, this study paves way for a safer financial ecosystem resilient to emerging threats.**

Keywords—Long short-term memory, Convolutional neural network, Synthetic minority oversampling, Voting, Stacking, deep learning, ensemble.

I. Introduction

Credit card fraud poses a significant threat to financial institutions and consumers alike, with increasingly sophisticated tactics emerging in the digital age. In response to this challenge, Fraud Shield emerges as a comprehensive fraud detection system leveraging ensemble learning and deep learning techniques. This study introduces Fraud Shield, highlighting its key features and benefits in the realm of credit card fraud detection.

FraudShield integrates ensemble learning, a machine learning paradigm known for its ability to combine multiple models to enhance predictive performance. By aggregating diverse fraud detection algorithms, Fraud Shield creates a robust system capable of identifying and mitigating fraudulent activities associated with credit card transactions. This approach ensures higher accuracy and generalization compared to individual models, addressing the evolving landscape of financial fraud.

Additionally, the proposed model leverages the power of deep learning, a subset of machine learning inspired by the human brain's neural networks. Deep learning models, including convolutional neural networks (CNNs) and long short-term memory (LSTMs), excel at learning intricate patterns and features from complex data. By leveraging deep learning, Fraud Shield can effectively discern fraudulent patterns within credit card transaction data, even amidst challenging scenarios where traditional methods may falter.

The synergy between ensemble learning and deep learning within Fraud Shield enables it to adapt and evolve in response to emerging fraud tactics and evolving transaction patterns. Through continuous analysis and learning from new data, Fraud Shield remains at the forefront of fraud detection, providing financial institutions and merchants with the confidence to transact securely in the digital landscape.

Key features and benefits of Fraud Shield include enhanced accuracy, real-time detection capabilities, adaptability to evolving fraud patterns, scalability to handle large volumes of transaction data, and comprehensive protection against various types of credit card fraud. These attributes position Fraud Shield as a vital tool for safeguarding financial transactions and preserving trust in the digital ecosystem.

II. Literature Review

Ibomoiye Domor Mienye et al. [1] propose a deep learning ensemble for credit card fraud detection, merging LSTM and GRU neural networks with an MLP meta-learner. Through SMOTE-ENN for class distribution balancing, the approach achieves superior sensitivity (1.000) and specificity (0.997) compared to traditional ML methods. While effective in capturing intricate transactional patterns, the ensemble's complexity and computational demands may hinder practical deployment. Nevertheless, it underscores the potential of ensemble techniques in enhancing fraud detection, necessitating further research to optimize efficiency for real-world implementation.

Khushi Goyal et al. [2] propose a real-time scam recognition system for e-commerce credit card fraud using an ensemble of machine learning and deep learning models. Leveraging technologies like NumPy and Scikit-learn, the study aims for robust fraud detection. The dataset comprises 492 fraud cases out of 283,806 transactions, with testing conducted using Random Forest Classifier and Decision Tree algorithms on European cardholder data. While promising for real-time fraud detection, the approach's efficacy may vary based on dataset representativeness and algorithm adaptability to evolving fraud patterns.

Harshit Muhal et al. [3] tackle credit card fraud employing a champion-challenger approach with three ensemble models. Champion Challenger 1, comprising Random Forest, Ada-boost, and LSTM, achieves notable Accuracy, Precision, Recall, and F1-Score of 99.86, 99.73, 99.99, and 99.86, respectively. While offering robust fraud mitigation, this approach may face challenges in real-time adaptability and scalability. However, its superior performance indicates potential for significant financial risk reduction in global fraud scenarios.

Damanpreet Kaur et al. [4], credit card companies must ensure top-notch security against fraudulent transactions. This study explores fraud detection using machine learning, deep learning, and their ensemble on imbalanced data. The goal is to achieve high recall without sacrificing precision, critical for maintaining customer trust. Various techniques, including resampling and metric selection, are demonstrated to address data imbalance and enhance model accuracy.

Wei Luo [5], utilizing various deep learning techniques, this research focuses on quantifying uncertainty for credit card fraud detection, crucial for reliable decision-making in dynamic fraud scenarios. Methods like Monte Carlo dropout and ensemble modelling are employed to assess uncertainty, enhancing fraud detection accuracy. Ensemble methods prove superior in capturing prediction uncertainty, thereby improving fraud detection outcomes, and providing valuable information for decision-making.

PFawaz Khaled Alarfaj et al. [6] detecting credit card fraud is imperative due to its financial implications. This study explores state-of-the-art machine learning and deep learning algorithms to address the challenges posed by fraud, including data imbalance and evolving fraud tactics. Empirical analysis on a benchmark dataset showcases significant improvements in accuracy, f1-score, precision, and AUC curves, surpassing existing methods and offering effective real-world fraud detection solutions.

Xiong Kewei et al. [7], Addressing the surge in credit card fraud, this research introduces a hybrid deep learning model enhanced by feature engineering, memory compression, mixed precision, and ensemble loss techniques. Evaluated on the IEEE-CIS fraud dataset, the model exhibits superior performance compared to traditional machine learning methods such as Bayes and SVM, offering a promising solution to combat online fraud.

Vejalla et al. [8] explore "Credit Card Fraud Detection Using Machine Learning Techniques," focusing on traditional machine learning methods. While effective in initial fraud detection, reliance on these techniques may limit the model's ability to detect complex fraud patterns. The absence of deep learning and ensemble methods could hinder performance in identifying sophisticated fraud, potentially constraining the model's effectiveness in combating online fraud.

Bayhan et al. [9] investigate "The Effect of Feature Selection on Credit Card Fraud Detection Success," highlighting the importance of feature selection in fraud detection models. While feature selection enhances model interpretability and efficiency, it may inadvertently discard crucial information relevant to fraud patterns. Consequently, reliance solely on feature selection methods without incorporating ensemble or deep learning techniques could limit the model's ability to detect intricate fraud schemes, potentially compromising its overall effectiveness in fraud detection.
Parekh et al. [10] explore "Credit Card Fraud Detection with Resampling Techniques," focusing on the application of resampling methods like SMOTE and ENN to address class imbalance. However, the study suggests the necessity of integrating ensemble and deep learning techniques for more comprehensive fraud detection solutions.

Al Smadi and Min [11] conduct a "Critical review of Credit Card Fraud Detection Techniques," presenting an exhaustive analysis of existing methods. They evaluate the effectiveness of various techniques, shedding light on their strengths and limitations in addressing the evolving challenges posed by fraudulent activities. Their comprehensive review offers valuable insights into the landscape of credit card fraud detection, aiding researchers, and practitioners in developing more robust and efficient detection systems.\

Boutaher et al. [12] provide a comprehensive "Review of Credit Card Fraud Detection Using Machine Learning Techniques," offering insights into the application of various machine learning algorithms for fraud detection. Their analysis encompasses the strengths and weaknesses of different techniques, providing researchers and practitioners with valuable guidance in selecting the most appropriate approach for combating fraudulent activities. This review serves as a valuable resource for advancing the field of credit card fraud detection, facilitating the development of more effective and efficient detection models.

III. Methodology

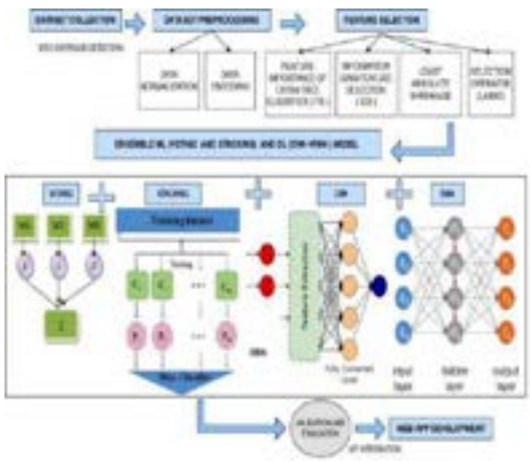


Fig 1: Architecture Diagram

Fig 1 illustrates the flow of data and processing steps within the proposed Credit Card Fraud detection, from data collection to preprocessing to implementation of algorithms.

1. Principal Component Analysis (PCA) is first used to process our modified dataset. This includes cleaning up any redundant features from the dataset, normalizing attributes with large fluctuation, and locating and fixing any missing data entries.

2.We address class imbalance using the Synthetic Minority Over-sampling Technique (SMOTE). Initially, the dataset exhibited severe class imbalance with over 25,000 instances of non-fraudulent transactions and only around 1,000 instances of Fraudulent transactions. After the SMOTE application, the classes were balanced, resulting in approximately 25,000 instances for both Fraudulent and Non-Fraudulent transactions, enhancing model robustness.

3.In the third crucial step of our methodology, we implement normalization using the scikit-learn library. This crucial process ensures that every feature derived after PCA transformation, spanning from v1 to v28, is appropriately scaled within the range of 0 to 1. This normalization procedure enhances the consistency and effectiveness of model training and evaluation by harmonizing the data distribution across all features.

4. Post-normalization, feature selection is conducted using Extra Tree Classifier and information gain score, alongside LASSO regression. This process aims to identify the most influential features for predicting fraudulent transactions. Following evaluation, a common set comprising 11 top-ranking features is selected, enhancing the model's efficiency and interpretability in credit card fraud detection.

5.Subsequently, the selected important features are inputted into a Voting Classifier, comprising Random Forest Classifier, Gradient Boost Classifier, and Logistic Regression. This ensemble method combines the predictions of multiple classifiers to enhance fraud detection accuracy. By leveraging the strengths of each classifier, the Voting Classifier provides a robust and effective solution for identifying fraudulent credit card transactions.

6. Following the Voting Classifier, a Stacking Classifier is implemented, with a Decision Tree Classifier as the final estimator. This ensemble technique combines the predictions of multiple classifiers, providing further improvements in fraud detection accuracy. Comparative analysis of the output accuracies from each ML algorithm, along with those from the Voting and Stacking Classifiers, highlights the superior efficiency of the Stacking Classifier in identifying fraudulent transactions.

7. Following, Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) algorithms are trained for 10 epochs. Additionally, a hybrid model combining LSTM and CNN is employed. Subsequently, the accuracies and results of all models are compared. Once more, the Stacking Classifier proves to be the most effective algorithm for credit card fraud detection, outperforming both individual DL algorithms and their hybrid combination.

IV. Implementation

A. Data Collection

A key tool in this study is the Kaggle Credit Card Fraud Detection dataset, which is well-known in several academic and scientific fields. This dataset is an extensive collection of transactional data from Europe and was assembled over the course of two days in September 2013. The dataset contains essential identity information linked to the people who are carrying out these transactions in addition to transactional facts. With a total of 284,807 cases, it offers a detailed look into the nuances of financial transactions throughout this time. Remarkably, even with its large size, the dataset exhibits a stark disparity in class, with only 492 cases of fraudulent transactions—that is, 0.17% of the entire dataset.

This glaring disparity highlights the inherent difficulties in effectively assessing and forecasting fraudulent activity. To derive valuable insights from the dataset, it is imperative to ensure that rigorous techniques are in place to rectify this imbalance. Furthermore, the combination of transactional and identity-related features in the dataset provides researchers with a wealth of information, enabling a comprehensive investigation of credit card fraud detection techniques.

B. Data Preprocessing

An extensive investigation of credit card fraud detection was made possible in this study by the merging of company’s identity and transaction databases. These datasets were carefully combined to create a coherent dataset that is suitable for additional research. The datasets contained crucial identifying attributes and transactional information. Following the merger, Principal Component Analysis (PCA) became apparent as a potent dimensionality reduction method that made it possible to extract crucial elements required for fraud detection. The condensed collection of features, designated as V1 through V20, was created via PCA from the combined dataset, encapsulating the most important elements of the original data.

Preprocessing such large datasets required several careful processes to guarantee data dependability and integrity. Notably, PCA was essential for managing null and empty values in addition to feature extraction. Null and empty values were effectively reduced by utilizing PCA's innate ability to identify and keep the most informative features while eliminating unnecessary ones. Furthermore, pretreatment methods including scaling and data normalization were used to improve the dataset's resilience and readability. This rigorous preprocessing routine made sure the information was ready for further analysis and provided a strong framework for investigating credit card fraud detection techniques.

C. Oversampling Using SMOTE

We use the Synthetic Minority Over-sampling Technique (SMOTE) to address the extreme class imbalance in our dataset. At the beginning of the collection, there was a notable discrepancy: more than 25,000 examples of transactions that were not fraudulent were significantly more than the around 1,000 examples of fraudulent transactions. We rebalance the classes by using SMOTE, yielding around 25,000 instances for both fraudulent and non-fraudulent transactions. By improving the resilience of our models, this augmentation reduces the possibility of biased predictions and guarantees a more equal representation of both classes.

Although SMOTE is a commonly used oversampling technique, there are other solutions available to alleviate class imbalance, including Random Over-sampling, Near Miss and Tomek Links under-sampling techniques, and ensemble methods like Balanced Random Forest and Easy Ensemble. SMOTE, however, is a better option due to its ease of use, efficiency, and capacity to produce artificial instances that closely resemble actual minority class samples. Its application in our study greatly enhances model performance, leading to increased accuracy and dependability in fraud detection.

D. Normalization and Feature Selection

We deal with the normalization of our dataset in the next step of our process. This crucial step makes sure that all features, which after PCA transformation range from v1 to v28, are scaled correctly between 0 and 1. Normalization is essential for improving the uniformity and effectiveness of model training and assessment since it uniformizes the distribution of data for all characteristics, reducing the impact of different scales on model performance.

After normalization, feature selection is the next important step, which is to find the features that have the greatest impact on predicting fraudulent transactions. We use LASSO regression together with two different techniques: Information Gain Score (IGS) and Extra Tree Classifier. The Extra Tree Classifier, which is well-known for its resilience and capacity to process high-dimensional data, assesses each feature's relative value by considering random splits at each decision tree node. In addition, the information gain score offers important insights into feature importance by quantifying the decrease in entropy or uncertainty attained when a specific feature is chosen for node splitting. Furthermore, by applying a penalty to the absolute magnitude of the regression coefficients, LASSO regression successfully promotes feature selection and sparsity. Through the amalgamation of these techniques, we establish a shared collection of 11 highest-ranking attributes, augmenting the effectiveness and comprehensibility of our model within the credit card fraud identification field.

E. Training with Ensemble ML (Voting and Stacking)

After going through the feature selection procedure, the major features that were found are combined with the Random Forest, Gradient Boost, and Logistic Regression classifiers to create a Voting Classifier. The objective of this ensemble strategy is to increase the accuracy of fraud detection by combining the predictive powers of several classifiers. The Voting Classifier offers a robust and effective way to identify fraudulent credit card transactions by utilizing the distinct advantages of each classifier.

Accurate predictions are produced by the Random Forest Classifier, which is known for its robustness and versatility when processing high-dimensional data. It does this by combining the output of several decision trees. By concentrating on cases with high error rates, the Gradient Boost Classifier, on the other hand, uses a sequential boosting strategy to enhance model performance iteratively and increase predicted accuracy. Furthermore, Logistic Regression functions as a basic yet potent linear model that provides interpretability and effectiveness in identifying the underlying associations between the target variable and the features.
The Voting Classifier's remarkable precision result of 98.289% is attained by the ensemble of several classifiers, highlighting its effectiveness in fraud detection. But to improve even more, we apply a Stacking Classifier, which raises precision to a remarkable 99%. The Stacking Classifier combines predictions from several base classifiers to obtain higher fraud detection accuracy. It uses a Decision Tree Classifier as the final estimator. This combination of classifiers, designed to handle the subtleties of credit card fraud identification, highlights the effectiveness of ensemble approaches for handling difficult classification problems.

F. Training with DL (CNN and LSTM)

We implement Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models to detect credit card fraud. Following three Conv1D layers with filter widths of 32, 64, and 128 respectively, the CNN design comprises MaxPooling1D layers for feature extraction. It employs the Adam optimizer with a learning rate of 0.001 and the categorical cross-entropy loss function for model compilation.

Fig 2. LSTM accuracy

The best-performing model weights are saved via the Model Checkpoint callback. On the other hand, dropout regularization prevents overfitting in the LSTM model by combining two LSTM layers for sequential data processing and Conv1D layers for feature extraction. Both models are trained for ten epochs, and the LSTM model achieves an impressive test accuracy of about 99.8% as shown in fig. 2.

These models use different but complementary techniques to address the difficulties associated with credit card fraud detection. While the LSTM model is better at identifying temporal patterns and long-term relationships, the CNN model is better at capturing spatial connections found in transactional data as shown in fig 3.

Fig 3. CNN Accuracy

The fraud detection system gains from a thorough examination of transactional data by utilizing the advantages of both models, which improves its accuracy and effectiveness in spotting fraudulent activity.

Fig 4. Hybrid model Accuracy

We have also developed a hybrid Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) model. The model's capacity to identify both spatial and temporal patterns in transactional data is improved by this architecture, which combines CNN for spatial feature extraction and LSTM for sequential data processing. The hybrid model undergoes training, assessment, and comparison with separate CNN and LSTM models as shown in fig 4.

The hybrid approach's efficacy in fraud detection is demonstrated by the results, which show that it produces performance that is either comparable to or better than that of the separate CNN and LSTM models.

VI. Results And Observations

The model's performance over the dataset is displayed in the tables 1 to 5, which also illustrate the model's training level using several metrics like F1-score, precision, and recall. The model's performance is balanced by the F1-score, which is the average of precision and recall. Precision denotes the accuracy of positive predictions, recall indicates the percentage of accurately predicted true positives.

TABLE 1. LSTM Results

	precision	recall	f1-score	support
Not Fraud	99%	100%	100%	5014
Fraud	100%	99%	100%	4986
accuracy			97%	10000
Macro average	97%	97%	97%	10000
Weighted average	97%	97%	97%	10000

Table 2. CNN Results

	precision	recall	f1-score	support
Not Fraud	99%	99%	99%	5014
Fraud	99%	99%	99%	4986
accuracy			99%	10000
Macro average	99%	99%	99%	10000
Weighted average	99%	99%	99%	10000

The results from both LSTM and CNN models demonstrate high precision, recall, and F1-score values for both the "Not Fraud" and "Fraud" classes, indicating robust performance in credit card fraud detection. Both models achieved an overall accuracy of 99%, showcasing their effectiveness in correctly classifying transactions. Additionally, the macro and weighted average scores further confirm the models' balanced performance across all classes and their ability to generalize well on the dataset.

Table 3. Hybrid (CNN+LSTM) Results

	precision	recall	f1-score	support
Not Fraud	97%	98%	97%	5014
Fraud	98%	97%	97%	4986
accuracy			97%	10000
Macro average	97%	97%	97%	10000
Weighted average	97%	97%	97%	10000

Table 4. Voting Observations

	precision	recall	f1-score	support
Not Fraud	97%	99%	98%	5014
Fraud	99%	97%	98%	4986
accuracy			98%	10000
Macro average	98%	98%	98%	10000
Weighted average	98%	98%	98%	10000

Table 5. Stacking Results

	precision	recall	f1-score	support
Not Fraud	99%	100%	100%	5014
Fraud	100%	99%	100%	4986
accuracy			100%	10000
Macro average	100%	100%	100%	10000
Weighted average	100%	100%	100%	10000

The results from the ensemble methods, Voting and Stacking, highlight the efficacy of combining multiple classifiers for credit card fraud detection. While both methods exhibit impressive precision, recall, and F1-score values, Stacking outperforms Voting across all metrics, achieving perfect accuracy scores. This suggests that Stacking, which leverages the predictions from various classifiers, offers superior performance in accurately identifying both fraudulent and non-fraudulent transactions.

VII. CONCLUSION

The obtained results highlight the potential of ensemble learning techniques, particularly Stacking, in enhancing existing fraud detection systems. By integrating diverse classifiers and leveraging their collective predictive power, Stacking can provide more robust and reliable fraud detection capabilities. Overall, these results emphasize the value of ensemble methods as a promising approach for increasing the effectiveness of credit card fraud detection systems.

REFERENCES

[1] Mienye, Domor & Sun, Yanxia. (2023). A Deep Learning Ensemble With Data Resampling for Credit Card Fraud Detection.IEEEAccess.PP.1-1

[2] K. Goyal, S. Singh, M. Gulati and A. Suresh, "An Ensemble Of Machine And Deep Learning Models For Real Time Credit Card Scam Recognition," *2023 International Conference on Computer Communication and Informatics (ICCCI)*, Coimbatore, India, 2023, pp. 1-4

[3] H. Muhal, G. Khatri, G. K. Dhama and D. Bansal, "Ensemble Approach for Credit Card Fraud Detection Using Champion-Challenger Analysis," *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, Chennai, India, 2022, pp. 1-8

[4] D. Kaur, A. Saini and D. Gupta, "Credit Card Fraud Detection Using Machine Learning, Deep Learning, and Ensemble of the both," *2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC)*, Solan, Himachal Pradesh, India, 2022, pp. 484-489.

[5] W. Luo, "Uncertainty Aware Based Deep Learning for Credit Card Fraud Detection," ISCTT 2022; 7th International Conference on Information Science, Computer Technology and Transportation, Xishuangbanna, China, 2022, pp. 1-4.

[6] F. K. Alarfaj, I. Malik, H. U. Khan, N. Almusallam, M. Ramzan and M. Ahmed, "Credit Card Fraud Detection Using State-of-the-Art Machine Learning and Deep Learning Algorithms," in *IEEE Access*, vol. 10, pp. 39700-39715, 2022

[7] Xiong, Kewei & Peng, Binhui & Jiang, Yang & Lu, Tiying. (2021). A Hybrid Deep Learning Model For OnlineFraudDetection.431-434.10.1109/ICCECE51280.2021.9342110.

[8] I. Vejalla, S. P. Battula, K. Kalluri and H. K. Kalluri, "Credit Card Fraud Detection Using Machine Learning Techniques," *2023 2nd International Conference on Paradigm Shifts in Communications Embedded Systems, Machine Learning and Signal Processing (PCEMS)*, Nagpur, India, 2023, pp. 1-4, doi: 10.1109/PCEMS58491.2023.10136040.

[9] E. Bayhan, A. G. Yavuz, M. A. Güvensan and M. E. Karslıgıl, "The Effect of Feature Selection on Credit Card Fraud Detection Success," *2021 29th Signal Processing and Communications Applications Conference (SIU)*, Istanbul, Turkey, 2021, pp. 1-4, doi: 10.1109/SIU53274.2021.9477812.

[10] P. Parekh, C. Rana, K. Nalawade and S. Dholay, "Credit Card Fraud Detection with Resampling Techniques," *2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Kharagpur, India, 2021, pp. 1-7, doi: 10.1109/ICCCNT51525.2021.9579915.

[11] B. Al Smadi and M. Min, "A Critical review of Credit Card Fraud Detection Techniques," *2020 11th IEEE Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON)*, New York, NY, USA, 2020, pp. 0732-0736, doi: 10.1109/UEMCON51285.2020.9298075.

[12] N. Boutaher, A. Elomri, N. Abghour, K. Moussaid and M. Rida, "A Review of Credit Card Fraud Detection Using Machine Learning Techniques," *2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech)*, Marrakesh, Morocco, 2020, pp. 1-5, doi: 10.1109/CloudTech49835.2020.9365916.

An Android Application for Temporary Driver Booking System

Kasarapu Ramani
Professor & Head, Dept of DS,
Mohan babu university,
Tirupati, India
ramanidileep@yahoo.com

A Jyothi
UG Scholar, Dept of CSSE,
Sree Vidyanikethan Engineering College,
Tirupati, India
jyothimanjula0604@gmail.com

Sai Ram Kinjarapu
UG Scholar, Dept of CSSE,
Sree Vidyanikethan Engineering College,
Tirupati, India
sairamkinjarapu119@gmail.com

Kilari Preethi
UG Scholar, Dept of CSSE
Sree Vidyanikethan Engineering College,
Tirupati, India
preethikilari02@gmail.com

Kurla Uday Kiran Reddy
UG Scholar, Dept of CSSE,
Sree Vidyanikethan Engineering College,
Tirupat, India
kurlaudaykiranreddy9866@gmail.com

***Abstract*— This research study presents an innovative Android application for a Temporary Driver Booking System, which aims to revolutionize the temporary driver industry by providing a streamlined, efficient, and safe solution for connecting customers with qualified drivers. The platform offers a seamless user experience, allowing customers to specify their transportation needs and choose from a pool of verified drivers based on availability, experience, and ratings within the stipulated range by calculating shortest distance between User and Drivers using Haversine and Dijkstra's algorithms on the other hand, driver can manage their availability, and accept service requests. Safety and reliability are ensured through rigorous driver verification processes, including background checks and license verification. In case of emergencies, an emergency assistance feature allows users to quickly alert the support team and share ride details for prompt assistance. The platform incorporates advanced features such as real-time driver tracking, and notifications to enhance transparency and communication. Multiple payment modes, including card and cash, cater to diverse user preferences. The utilization of analytics for driver bookings offers valuable insights for optimization and service enhancement. By leveraging technology, this Android application redefines industry accessibility reliability, promising convenience, efficiency, and safety for all stakeholders involved.**

Keywords*— *Temporary Drivers, Android application, Haversine algorithm, Dijkstra's algorithm, Emergency Assistances, Real-time tracking.

I. Introduction

The traditional methods of accessing temporary driver services present several challenges, including fragmented experiences, inefficiencies, and safety concerns. These methods lack a centralized platform to connect customers with qualified drivers seamlessly. Furthermore, there is a lack of transparency and security in the hiring process, with minimal options for customers to verify driver credentials and track their journey progress.

This Android application revolutionizes the industry by providing a seamless, efficient, and secure platform that connects customers with skilled drivers, eliminating the fragmentation and inconvenience previously experienced. Through a detailed analysis of its core features, the research paper underscores the system's commitment to enhancing safety and security, evident in its stringent verification processes and emergency protocols. Furthermore, the paper explores how advanced functionalities such as real-time tracking and automated payments contribute to an improved user experience, while also highlighting the broader benefits for both customers and drivers, including transparency, fair wages, and flexible work opportunities.

Moreover, the research paper emphasizes the positive societal impact of the Temporary Driver Booking System revolutionizes the industry by providing a seamless platform connecting customers with skilled drivers, prioritizing safety [8], efficiency, and transparency. Through stringent verification processes, real-time tracking, and automated payments, it enhances user experience while fostering a more inclusive and sustainable transportation ecosystem, setting new standards for convenience and reliability.

II. Literature Survey

The traditional temporary driver booking system faces several challenges. Firstly, the market is fragmented, lacking centralized platforms, which makes it difficult for users to find reliable drivers and compare rates efficiently. The booking process is often inefficient, relying on manual methods that are time-consuming and prone to errors. Safety concerns arise due to unverified drivers and inadequate tracking mechanisms, raising questions about the safety. Moreover, limited payment options, primarily cash-based transactions, can be inconvenient and insecure for both users and drivers. Addressing these challenges requires innovative solutions that streamline the booking process, enhance safety measures, and offer more flexible payment options.

Self-owned System: If a fleet is planned with the help of temporary drivers, an auto-booking system streamlines the process for users. Instead call or email directly, users can book drivers directly through the app. This system facilitates direct booking for users, reducing the need for manual intervention. The project proposes a centralized platform connecting individuals seeking transportation with car rental services, enhancing their retention and simplifying the booking process.

It addresses common challenges in drivers hiring, such as added costs, driver cancellations, and long journeys. The project leverages SMS technology to provide notifications and integrates online frameworks to streamline the rental process, including vehicle information retrieval, reservation, and immediate leasing. Overall, this application aims to offers a practical and efficient method for hiring drivers, benefiting both users and drivers.

R. Sathya., et al. [1], Car Hiring System using Web Technology, have done a web application that utilizing recent technologies, the project simplifies car rental, addressing risks like added costs and driver cancellations. Instead, this application used to hire drivers for fulfilling customers' requirements if they have cars.

Furqan Talha., et al. [2], Emergency Vehicles Assistance System, have done the Road Vehicle Breakdown Assistance Finder web application utilizes GPS tracking and real-time communication to swiftly connect to distressed vehicle owners, with licensed mechanics, ensuring fair pricing and service quality. Additionally, proposed application provides emergency assistance which is a pre-composed text can directly send to get support.

Md Rayhan Hasnat., et al. [3] Development of an online based Driver Recruitment system, proposes the process of hiring drivers through interviews, providing an in-depth analysis of crucial elements such as job advertisement, application data management, interview coordination and scheduling, documentation of interview results, and the final candidate selection for employment. Furthermore, our paper proposes an extension of this process to facilitate the temporary availability of drivers.

Athira Krishnan T R ., et al. [4], Android-Based Transport Tracking and Monitoring System, have done the paper proposes a GPS-tracking application for buses, offering real-time location updates, route information, and estimated arrival times. It includes a comprehensive monitoring system for safety, engine fault detection, and efficient time management, suitable for deployment in both urban and rural transportation systems.

Harshada Shinde., et al. [5], Online Car and Driver Hiring System, paper introduces a novel system that grants the freedom to select both cars and drivers based on their preferences. In contrast to existing applications where users are unable to confirm prices in advance and must negotiate with drivers after making a booking, Proposed system enables to choose drivers based on their specific prices. Moreover, proposed application provides seamless payment options to enhance overall user experience.

V. Hemanth Kumar, K. Sentamilselvan, [6], Customer Satisfaction towards Call Taxi Services A study with reference to Chennai, this study investigates customer satisfaction with different call taxi service providers in Chennai, a major hub in the southern region. It delves into perceptions regarding the utilization of call taxi services, taking into account factors such as comfort, accessibility, pricing models, promotional strategies, safety precautions, convenience, and overall satisfaction with the quality of service offered by these companies. Building upon the findings of this research, proposed application gathers users feedback after the completion of a ride to improve safety and convenience further.

Uttam Pandey., et al. [7], Online Driver Booking Service, this paper introduces a system where users can easily book drivers from their homes by providing a few details through an online application. Once the booking is confirmed, users can begin their journey, and upon completion, the payment process is initiated. Driver searches are customized to align with the language preferences specified by customers. Proposed system enhances driver searches by utilizing both the shortest path and multi-conditional search algorithms to deliver the best available list of drivers.

Vijay Kumar A.G., Kumar U, [8], Profile of Fatal Road Traffic Accidents Due to Drunken Driving, this paper says that findings from the Alcohol & Drug Information Centre (AIDC) in India highlight a troubling pattern: around 40% of road accidents are linked to alcohol consumption. This risk is notably higher among young male drivers. The research emphasizes the substantial impact of alcohol use on fatal road accidents, especially in cases involving single vehicles and young male drivers. Proposed application aims to provide a solution for intoxicated individuals by enabling them to book a driver to safely operate their vehicle.

III. Methodology

The system architecture for a temporary driver booking platform comprises tailored components aimed at streamlining user interaction and service provision. Through a mobile application, users - both customers and drivers - engage with various functionalities such as registration, specifying transportation needs, driver selection, ride tracking, payments, and feedback provision. Behind the scenes, a robust database serves as a central repository for storing user, driver, vehicle, booking, and transaction data, ensuring seamless operation and data integrity.

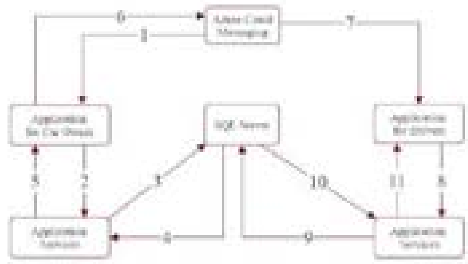


Fig. 1. Application Architecture

Additional modules include secure login/registration, driver profiles for informed decision-making, and transaction handling for financial operations. The Android application further enhances temporary driver bookings by allowing users and drivers to register with personal details and licenses. Users input ride details, and an algorithm matches them with available drivers based on various criteria, presenting a curated list of drivers ranked with the best match on top. Real-time journey tracking via GPS ensures transparency and safety, with emergency assistance features for quick contact with authorities if needed. Transparent fare calculation based on distance travelled is facilitated at the end of the journey, with options to store user preferences for future bookings and rate drivers to improve the ranking system. Additional communication features may enable interaction between users and drivers before the ride, optimizing overall efficiency and effectiveness of the temporary driver booking services.

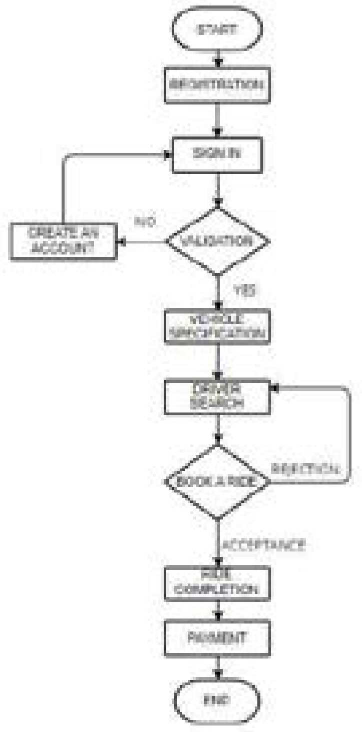


Fig. 2. Flow Chart

Driver.com is an android application for a temporary driver booking system that has two modules: driver and user. Driver registration can be done using personal details and a license. The driver needs to login with valid credentials for access.

User Module

A. Registration and Login Screen

Users will initiate their journey with our Android application by either logging in or signing up to create a new account. Upon successful registration, they gain access to our application's features. During the registration process, users are prompted to provide all necessary details, Users will be directed to the login screen shown in Figure 3.

B. Driver Availability Screen

On this screen, users can easily browse and access the list of drivers based on various criteria such as ratings, number of days available for rides, car models, and more. Users have the flexibility to select from a diverse range of options, as illustrated in Figure 4. Once users find a driver of interest, they can view the fare associated with that particular driver. The user can hire a driver by entering details of the ride. The user will get a list of available drivers with their details

Algorithms

i) *Driver Selection by the User*

- Enter the User's vehicle details.
- Select the driver based on driver's experience, route, source, destination and rating using multi-conditional search.
- Nearest drivers list is obtained by calculating distance between user and drivers using Dijkstra's algorithm, which finds the shortest distance. The input for Dijkstra's algorithm is obtained from Haversine algorithm.
- Drive cost is calculated based on the distance and time.

ii) *Haversine algorithm*

- Retrieve driver and user locations from google maps.
- Convert the latitude and longitude values from degrees to radians.

$$rad\_lat = latitude \times \frac{\Pi}{180} \quad --\rightarrow \quad eq(1)$$

- Calculate the differences in latitude and longitude.
- Haversine formula to calculate the distance between the two points.

$$a = sin^2\left(\frac{\Delta lat}{2}\right) + cos(latA) \times cos(latB) \times sin^2\left(\frac{\Delta \mathrm{lon}}{2}\right) \quad --\rightarrow \quad eq(2)$$

$$c = 2 \times atan2\left(\sqrt{a}, \sqrt{1-a}\right) \quad --\rightarrow eq(3)$$

$$d = R \times c \quad --\rightarrow eq(4)$$

- Where
  - Δlat is the difference in latitude (delta_ latitude),
  - Δlon is the difference in longitude (delta_ longitude),
  - R is the radius of the Earth (approximately 6,371 kilometres or 3,959 miles),
  - a is the intermediate result,
  - c is the great-circle distance in radians, and
  - d is the distance between the two points in the same units as the Earth's radius.

Let

latA = 12.965835163156859,

lonA = 77.57453865017378

latB = 13.626971577753114

lonB = 79.41860086877251

By using eq(1) convert latA, lonA, latB, lonB values from degrees to radians values after converting:

latA = 0.226296513739

latB = 0.237835520971

lonA = 1.353931114389

lonB = 1.386116071697

Then by substituting the values in eq(2) value

a = 0.00027852683

by substituting the values in eq(3) value

c= 0.03892638518

by substituting the values in eq(4) value

d=248 KM

iii) *Dijkstra's Algorithm*

Step 1: Start by selecting the user's location as the source node in the graph.

Step 2: Set the distance of all other nodes (drivers' locations) to infinity with respect to the user's location.

Step 3: Create an empty set called 'N' to store nodes for which the shortest path has been found

Step 4: Assign a distance of zero to the user's location. Add it to set 'N' and designate it as the 'active node'

Step 5: Look at the neighbouring nodes of the active node, i.e., the drivers' locations, which are connected by weighted edges (distances between geographical coordinates). Calculate the total distance from the user to each neighbouring driver by using Haversine's algorithm.

Step 6: If a neighbouring node already has a distance assigned, compare the newly obtained distance (from step 5) with the current value. Update the distance if the new distance is shorter.

Step 7: Select the driver's location with the minimum distance assigned to it. Insert this node into set 'N', making it the new 'active node'.

Step 8: Repeat steps 5 to 7 until either the destination node (the nearest driver) is included in set 'N', indicating that the shortest distance has been found, or if there are no more labelled nodes (unvisited drivers' locations) in set 'N'.

The process begins by treating the geographical locations of both the driver and the user as nodes within a graph, where each node is uniquely identified and stores its respective latitude and longitude coordinates. Utilizing the Haversine formula, distances between every pair of nodes are meticulously calculated, thereby establishing the weights of the edges in the graph. Armed with this information, a comprehensive graph structure is constructed, interconnecting each node with its adjacent nodes based on the derived distances. Subsequently, Dijkstra's algorithm is employed to effectively determine the shortest path between the starting and destination nodes. Upon completion of the algorithm's execution, the shortest path, comprising a sequential arrangement of nodes, is extracted from its output. Finally, the cumulative distance along this optimal path is computed by aggregating the distances between consecutive nodes. This meticulous process culminates in the determination of the shortest distance between the driver and the user, facilitating efficient navigation.

C. *Booking screen*

On the Booking Screen, users have the opportunity to select their preferred driver from the list. They can proceed by filling out a form with details such as the date, time, location, and duration of the driver's hire, as depicted in Figure 5. The best match will come on top in decreasing order. The user can select one of them from the list of drivers. An emergency assistance is embedded in the application. The distance travelled is calculated with the help of integrated Google Maps API into the application. This process allows users to finalize their booking seamlessly and efficiently.

D. *Payment screen*

In the Payment Module, once users or customers have provided all the required details, they can proceed to make their payment. This module supports various payment methods such as Phone Pe, Google Pay, credit/debit cards, and more. Users have the flexibility to choose the payment method that best suits their preferences, ensuring a convenient and hassle-free transaction process. After the payment is completed, the ride will be confirmed.

Driver Module

A. *Registration and Login screen*

Drivers will initiate their interaction with our Android application by either logging in or signing up to establish a new account. Once registered, they gain access to our application's functionalities. During the registration phase, drivers are prompted to furnish all necessary details to complete the process.

B. *Ride details screen*

The ride details screen provides comprehensive information about the upcoming trip, including pick-up and drop-off locations, passenger details, and any special instructions. It serves as a central hub for drivers to access essential ride information and manage their schedules efficiently.

IV. RESULTS & DISCUSSION

These are the screens that are provided in the application which help the user and driver to interact with the application.

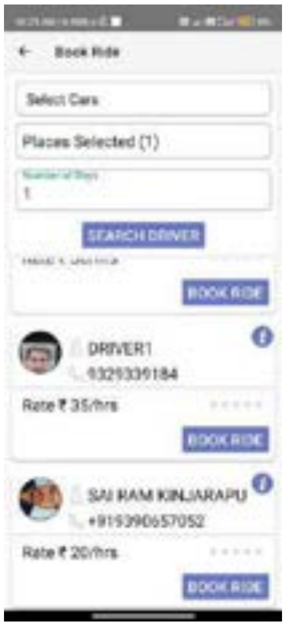


Fig. 3. Driver selection screen of the application

The user registration screen enables new users to create an account within the application. Its components consist of input fields and a submit button. Input fields typically include username, email address, password, and possibly additional details such as name, date of birth, etc. The submit button allows users to submit their registration details.

The login screen allows registered users to access their accounts by providing their credentials. Its components consist of input fields, a "forgot password" link, and a submit button. Input fields usually include fields for username/email and password. The "forgot password" link is an option for users who have forgotten their password to reset it. The sign-in button allows users to submit their login credentials.

The ride booking feature in the Android application for the temporary driver booking system enables users to easily book a driver for their transportation needs. Here's a brief description of the process:

Select Type of Place or Route: Users can specify the type of route they intend to travel. Options such as High Pass Roads, Mountain Ranges, Hill Station custom routes may be available for selection.

Fig. 4. Ride confirmation screen of the application

Select Driver from Available Drivers: Users can choose a driver based on preferences such as rating, vehicle type, or availability. List of available drivers with details such as name, rating, vehicle model, and possibly distance from the user's location. Pickup Location: Users provide the location from where they want the driver to pick them up. Input field or map interface allowing users to specify their current location accurately. Destination: Users specify the destination where they wish to be dropped off. Input field or map interface for users to input the destination address.

Users select the preferred payment method and confirm the booking. Options for payment methods such as credit/debit card, mobile wallet, or cash on delivery. Users may also see the estimated fare before confirming the booking.

Once all the necessary details are provided and the booking is confirmed, the system assigns the selected driver to the ride, and both the user and the driver receive confirmation notifications. The driver then proceeds to the pickup location, and the user can track the driver's location in real-time. After reaching the destination, the payment is processed, and the ride concludes.

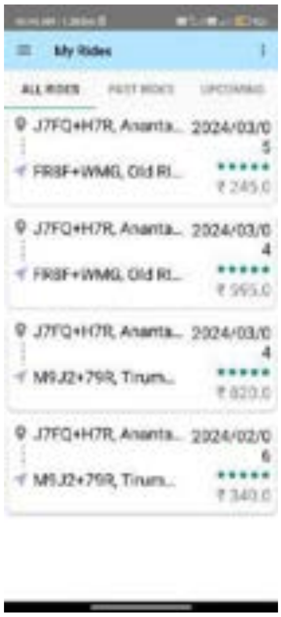


Fig. 5. Ride history screen of the application

The "My Rides" section within the Android application of the temporary driver booking system provides users with a comprehensive history of their past rides, offering convenience and transparency. Here's a brief description:

Fig. 6. Transaction history screen of the application

"My Rides" serves as a centralized location for users to review their entire ride history, facilitating easy access to past bookings and relevant details. Its components consist of a list of past rides and a list of upcoming rides. The list of past rides displays a chronological list of all previous bookings made by the user, while the list of upcoming rides displays a chronological list of all upcoming bookings made by the user.

This ride booking feature simplifies the process of hiring a temporary driver, providing users with a convenient and efficient means of transportation within the temporary driver booking system application.

The "My Transactions" feature provides users with a comprehensive view of their financial activities related to the booking services. Here's a brief description:

"My Transactions": This section serves as a ledger for users to track their financial interactions within the temporary driver booking system. Its components consist of transaction history and details. Transaction History presents a chronological list of all financial transactions associated with the user's account, including payments for bookings and any refunds or credits. Each transaction entry typically includes details such as the date and time of the transaction, the amount paid or received, and any relevant notes.

The user transactions and ride history will be maintained in the database for further analysis and reference. After completion of the ride, the user will be asked for feedback on the ride experience. The data collected from the user feedback will be used to rate the driver using a Moving Average.

To ensure security of the both user and driver an emergency assistance feature included in the application. It is a pre-composed text requesting for the assistance which can be directly send through platforms like WhatsApp, Gmail, and messenger.

The ability to store user preferences for future bookings and transaction history add convenience and personalized the user experience. Additionally, the rating system facilitated user-driver interactions, allowing users to provide feedback and improve the overall quality of service.

Through Dijkstra’s algorithm distance is calculated accurately. As driver booking is done based on driver’s authentication through driving licence and choosing a driver based on availability, experience, and price, which proves the efficiency.

The App is validated by choosing few rides and distance calculations are matching with physical distances in real-time.

V. CONCLUSION & FUTURE WORK

This research investigates the development and potential implications of an Android application designed for temporary driver bookings. The application sought to overcome the shortcomings of current solutions by providing a seamless, safe, and user-oriented experience for both users and drivers. The key objectives included enhancing the efficiency of the booking process, ensuring platform security, and prioritizing user satisfaction. Through a comprehensive analysis, the study aimed to elucidate the innovative features and potential impact of the proposed application on the transportation service industry. The Android application for temporary driver booking aims to consistently enhance its features to remain competitive in the industry and provide the best user experience. This application provides emergency assistance for both driver and user to enhance security. Developing functionalities specifically for the admin role, allowing for better management and oversight of the platform. Enhancing the user interface and user experience design to make the application more visually appealing, intuitive, and user-friendly, ultimately increasing its significance and appeal to users.

REFERENCES

[1] R. Sathya, G. Srujan Reddy, Grandhi Ram Teja, Vedantham P S Srinivasa Iyengar, Prabhat Kumar, Car Hiring System using Web Technology, March 2023, DOI: 10.1109/SCEECS

[2] Furqan Talha, M Vamshi Krishna Reddy, Omkar Solanki, Suhas R, Emergency Vehicles Assistance System, December 2023, DOI:10.48175/IJARSCT-14046

[3] Md Rayhan Hasnat, Nesar Ali, Mohammad Ali Martuza, Development of an online based Driver Recruitment system, December 2017, ICMERE2017-PI-103

[4] Athira Krishnan T R, R Neha, C Sreehari, Zenaani, Android-Based Transport Tracking and Monitoring System, December2023, DOI:10.21467/proceedings.160. 36

[5] Harshada Shinde, Nikita Kasalkar, Omkar Tibole, Roshan Shaikh, Online Car and Driver Hiring System, May 2023, e-ISSN: 2582-5208

[6] V. Hemanth Kumar, K. Sentamilselvan, Customer Satisfaction towards Call Taxi Services A study with reference to Chennai, 2018, Volume 119 No. 12 2018, 14919-14928, ISSN: 1314-3395

[7] Uttam Pandey, Raghuwansh Kumar Singh, Sanchi Kaushik, Online Driver Booking Service, July 2021, ISSN (Online): 2320-9364, ISSN (Print): 2320-9356

[8] Vijay Kumar A.G., Kumar U, Profile of Fatal Road Traffic Accidents Due to Drunken Driving, January 2017, DOI:10.21088/ijfmp.0974.33 83.10117.6

[9] Sandeep Gupta, Attaullah Buriro, Bruno Crispo, DriverAuth: A Risk-based Multi-modal Biometric-based Driver Authentication Scheme for Ride-sharing Platforms, January 2019, DOI: 10.1016/j.cose.2019.01.0 07

[10] Stefan Lang, Sven Laumer, Christian Maier, Andreas Eckhardt, Drivers, Challenges and Consequences of E-Recruiting – A Literature Review, 2011

[11] Dr. Ruchi Shukla, Dr. Ashish Chandra & Ms. Himanshi Jain, OLA VS UBER: The Battle of Dominance, IOSR Journal of Business and Management (IOSR-JBM), VINC'17, 73-78.

[12] Shulin Lan, Chen Yang, and Chun-Hsien Chen, Online Car-Hailing System Performance Analysis Based on Bayesian Network, July 2019, DOI: 10.1109/ACCESS.2019.2929620.

[13] Loh Giin Xiin, A.S. Shibghatullah, Abdurrahman Jalil and Mohd Helmy Abd Wahab, DriverSeekers – A mobile designated driver services system, May 2021, DOI: 10.1088/1742-6596/1874/1/012029

Recommendation System Enhanced by Big Data Analytics

Kotagiri Vignesh
Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation,
Guntur, Andhra Pradesh, India
2100031188cseh@gmail.com

Tondepu Srujana
Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation,
Guntur, Andhra Pradesh, India
2100032015cseh@gmail.com

Vunde Jeswanth Naga Sai Teja
Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation,
Guntur, Andhra Pradesh, India
2100032116cseh@gmail.com

Chinthalacheruvu Hari Vardhan Reddy
Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation,
Guntur, Andhra Pradesh, India
2100032173cseh@gmail.com

Dr.Padmanaban K
Department of Computer science and Engineering, Koneru Lakshmaiah Education Foundation,
Guntur, Andhra Pradesh, India
padmanaban.k@kluniversity.in

***Abstract*— This research study describes the integration of big data analytics to develop a novel and robust e-commerce recommendation system. Acknowledging the challenges associated with providing accurate and timely suggestions, advanced recommendation algorithms, and the utilization of big data analytics tools. This study employs a comprehensive methodology, beginning with the characterization of the e-commerce dataset. As a first step, this study explains the data pre-processing steps undertaken to ensure the quality and relevance of information. The proposed recommendation system design incorporates state-of-the-art algorithms. Implementation details describe the practical aspects of deploying the recommendation system in an e-commerce environment, including the challenges encountered and their resolutions. The results section presents key findings, supported by visual representations and performance metrics that gauge the effectiveness of the proposed system. Through a thoughtful discussion, this study contextualizes the results within the broader landscape of e-commerce recommendation systems and big data analytics. Drawing comparisons with existing literature, this study highlights the significance of various research contributions and identify opportunities for future research. In conclusion, this research not only provides insights into optimizing e-commerce recommendations but also contributes to the ongoing research on leveraging big data analytics in enhancing user engagement and satisfaction.**

***Keywords*— Diabetes Prediction, Machine learning, Healthcare analytics, Decision trees, Support Vector Machines.**

I. Introduction

In the dynamic world of online commerce, where choices are abundant and time is limited, the role of recommendation systems has become increasingly crucial. These systems serve virtual shopping assistants, helping users discover products tailored to their preferences and needs. In the vast sea of available options, customers often find themselves overwhelmed, and this is where recommendation systems step in, offering personalized suggestions.

Understanding the importance of these systems involves recognizing their ability to bridge the gap between consumers and a myriad of available products. As e-commerce platforms continue to grow, the challenge lies not only in providing an extensive array of choices but also in helping users navigate through this abundance.

Despite the undeniable benefits of recommendation systems in e-commerce, several challenges persist in delivering recommendations that are truly effective and resonate with users. One major hurdle lies in the complexity of user preferences, which can be dynamic, diverse, and often challenging to accurately predict. The ever-evolving nature of consumer tastes and trends adds an additional layer of complexity, making it difficult for traditional recommendation approaches to keep pace [1].

The "cold start" issue presents another difficulty, especially for new users or objects for which there is insufficient previous data to enable the system to produce pertinent recommendations. In certain situations, the algorithm has trouble anticipating and comprehending human preferences, which results in recommendations that are less precise. Furthermore, problems with data sparsity, in which there is little to no data available for certain items or user preferences, might make it more difficult for the system to make thorough and trustworthy suggestions.

The environment is made more difficult by privacy issues, as people are growing increasingly hesitant to provide personal information. A difficult problem for e-commerce recommendation systems is to strike a balance between the requirement for individualized recommendations and the obligation to preserve user privacy. Furthermore, there is also worry about how recommendation algorithms can scale to handle enormous datasets effectively, especially in light of the rapid expansion of e-commerce platforms and user interactions [2].

Scope and Limitations:

1) Scope:

The objective of our study is to enhance the accuracy and efficacy of tailored product recommendations in e-commerce by integrating big data analytics into recommendation systems [5]. The scope of this project involves creating and using sophisticated recommendation algorithms to tackle issues like the "cold start" problem, data sparsity, and finding the right balance between personalization and user privacy. The study also investigates the scalability of recommendation algorithms, assuring their effectiveness in managing the increasing amounts of data created by the ongoing expansion of e-commerce platforms.

2) Limitations:

Despite our comprehensive approach, certain limitations exist within the scope of this study. The generalization of findings may be constrained by the specific characteristics of the dataset utilized, and the effectiveness of the proposed solutions could vary based on the nature and scale of different e-commerce platforms. Additionally, while our research aims to address the "cold start" problem, complete mitigation may not be achievable in all circumstances, especially in cases where minimal historical data is available [6].

II. LITERATURE REVIEW

A. E-commerce Recommendation Systems:

The literature on e-commerce recommendation systems demonstrates a diverse range of studies, highlighting the crucial impact these systems have in the online retail environment. Preliminary research frequently concentrated on collaborative filtering and content-based recommendation methods. Collaborative filtering utilizes user behavior and preferences, whereas content-based techniques examine item properties. Over time, there have been the emergence of hybrid models that try to combine both techniques in order to use the advantages of each.

Recent progress has shown the use of sophisticated deep learning methods, such as neural networks, to accurately detect complex patterns in user behavior. Furthermore, there has been a growing interest in context-aware recommendation systems, which take into account several parameters such as user location, device, and time in order to improve the accuracy and appropriateness of suggestions. These studies highlight the ongoing development of recommendation algorithms to adapt to the ever-changing nature of e-commerce platforms.

B. Big Data in E-commerce:

Existing research works have analysed the difficulties and advantages presented by large amounts of data in online commerce, such as preparing the data, analyzing it in real-time, and ensuring it can handle increasing amounts of data. The utilization of cloud computing and distributed computing architectures has been investigated to address the computational requirements of extensive e-commerce activities. The literature highlights the significance of big data in revealing concealed patterns, forecasting user behavior, and enhancing many areas of e-commerce. This lays the foundation for our investigation into its incorporation with recommendation systems [7].

C. Integration of Big Data into E-commerce Recommendation Systems:

The intersection of e-commerce recommendation systems and big data analytics is a relatively new but quickly growing field of study. Recent studies acknowledge that recommendation algorithms may be significantly enhanced by using the potential of big data. The integration facilitates a more sophisticated comprehension of user preferences, surmounting obstacles such as the "cold start" issue and data scarcity.

D. Algorithms, Techniques, and Methodologies in E-commerce Recommendation Systems:

Collaborative filtering continues to be a fundamental aspect in the work on recommendation systems. User-based collaborative filtering utilizes the preferences and actions of comparable users to provide suggestions. Item-based collaborative filtering, however, determines commonalities between objects by analyzing user interactions [8]. These algorithms encounter difficulties when dealing with sparse datasets and the "cold start" problem, which refers to the absence of adequate data for reliable suggestions for new users or objects.

Content-based filtering is a method that examines the characteristics of things and the preferences of users in order to provide suggestions. This strategy employs elements such as item descriptions, user profiles, and past interactions to provide tailored recommendations. Nevertheless, it may encounter difficulties in comprehending intricate user preferences and has a tendency to suggest goods that are like to those previously chosen by the user.

Hybrid models leverage the capabilities of both collaborative filtering and content-based filtering. Weighted hybrid models provide distinct weights to collaborative and content-based components, determined by their respective performance. Hybrid models seek to alleviate the constraints of separate methods, offering more resilient and precise suggestions.

Matrix factorization methods, such as Singular Value Decomposition (SVD) and Alternating Least Squares (ALS), break down the matrix representing the interactions between users and items into hidden elements. The latent variables in question capture the hidden patterns in user preferences and item features, resulting in more precise suggestions. Matrix factorization has demonstrated efficacy, particularly in managing datasets with low density and enhancing the precision of recommendations.

In recent years, deep learning techniques, specifically neural networks have become increasingly prominent. Collaborative and content-based neural networks can detect complex patterns in user behavior and item attributes, enabling more advanced and context-aware suggestions. Deep learning models frequently surpass standard approaches in effectively managing intricate connections throughout extensive datasets.

Context-aware recommendation systems include further contextual information, such as user location, time, and device, in order to customize suggestions with more precision. These systems optimize customization by tailoring suggestions to the user's unique context, enhancing the pertinence of recommendations in dynamic situations.

E. Integration of Advanced Algorithms with Big Data Analytics:

Although previous research acknowledges the importance of big data in e-commerce recommendation systems, there is a noticeable lack of integration between cutting-edge recommendation algorithms and advanced big data analytics methodologies. Our study aims to close this divide by investigating how the combination of advanced algorithms and the analytical capabilities of large datasets might enhance the precision and efficacy of tailored product recommendations [9].

F. Scalability of Recommendation Algorithms in Real-Time Environments:

The scalability of recommendation algorithms, especially in real-time situations with dynamic user interactions, is an area that current literature may further improve upon. Our research aims to fill this void by assessing the effectiveness of recommendation algorithms using big data technologies, assuring their capacity to manage the constantly expanding amounts of data produced by the ongoing expansion of e-commerce platforms [10].

G. Comprehensive User Understanding through Big Data Analytics:

Although current research delves into user behavior and preferences, there is still an opportunity to gain a more thorough knowledge of people by utilizing big data analytics. This study focuses on analysing user interactions, transactional data, and contextual information to gain a better understanding of recommendation systems. By extracting more profound insights, we strive to improve and enhance these systems over time. A graphical comparison of the identified research gaps and the research contribution areas is depicted in figure 1.

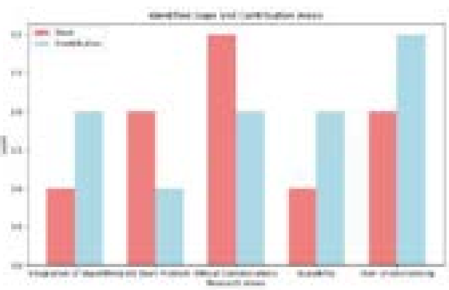


Fig.1 Gaps in the Areas

III. METHODOLOGY

A. Describing the Dataset:

To improve e-commerce recommendation systems using big data analytics, it is crucial to provide a detailed description of the dataset. This ensures transparency, repeatability, and a thorough knowledge of the study's basis [11].

B. Source of the Dataset:

The e-commerce dataset under scrutiny for this research has been meticulously curated from a diverse range of online retail platforms, encompassing various product categories and user demographics. The dataset amalgamates transactional data, user profiles, and contextual information to construct a holistic representation of the intricate dynamics inherent to e-commerce interactions.

C. Characteristics of the Dataset:

1) Transactional Data:

The dataset encapsulates a vast array of transactions, each entry detailing the purchase history of users. This includes information such as product IDs, quantities purchased, and timestamps, providing a temporal dimension to user interactions.

2) User Profiles:

User-centric details are a crucial part, encompassing information such as user IDs, demographic characteristics, and past preferences. These specific facts enhance the customization part of the recommendation system by aiding in the comprehension of unique user actions.

3) Contextual Information:

Recognizing the significance of context in e-commerce interactions, the dataset encompasses contextual elements. These may include user locations, device types, and time stamps, offering valuable insights into the circumstances surrounding each transaction.

4) Data Volume and Scalability:

The dataset is sufficiently large to emulate the scale of real-world e-commerce platforms, ensuring that the recommendation system's performance and scalability can be rigorously evaluated. It accommodates the intricate interplay of variables while simulating the challenges posed by substantial user bases and transaction volumes.

5) Data Preprocessing:

Prior to the implementation of recommendation algorithms, extensive data preprocessing measures are applied. This encompasses cleaning procedures to handle missing or erroneous entries, normalization to ensure consistency in data formats, and the anonymization of sensitive information to align with privacy considerations.

6) Dataset Splitting:

In order to ensure a thorough assessment of the recommendation system's effectiveness, the dataset is divided into training, validation, and testing sets using a stratification method. This guarantees that the model is trained using a wide range of user behaviors, validated to find the best hyperparameters, and thoroughly tested to evaluate its ability to generalize [12].

D. Recommendation System Design:

The design of our e-commerce recommendation system incorporates the strategic integration of innovative algorithms and techniques, as well as the utilization of big data analytics for the purpose of improving accuracy and personalization [13]. The implementation of our recommendation system involved an integration of algorithms and big data analytics tools.

IV. RESULTS & DISCUSSION

The findings show that the recommendation algorithms can be successfully integrated with big data analytics in e-commerce. The development and assessment of the recommendation system produced informative data, allowing for a more nuanced knowledge of user preferences and system performance.

A. Recommendation Accuracy:

The recommendation system exhibited a commendable level of accuracy, as indicated by the Root Mean Squared Error (RMSE) metric. The accuracy comparison of the three experiments carried out is given below in figure 2:

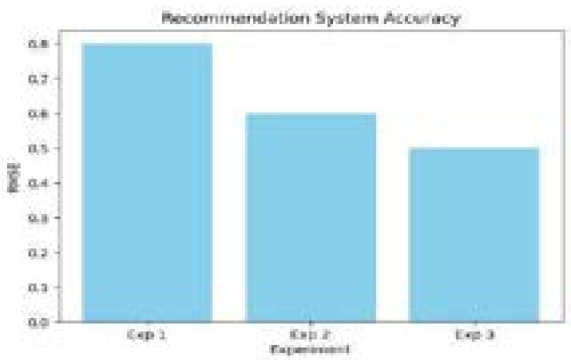


Fig.2 Performance Metrices

Collaborative filtering, content-based filtering, and hybrid models demonstrated robust performance in capturing user preferences. The integration of matrix factorization and deep learning techniques contributed to a more comprehensive understanding of intricate patterns within the dataset [15]. The hybrid method effectively solved the "cold start" issue by making relevant suggestions even for new users or goods with insufficient previous data. Content-based suggestions proved useful in several circumstances, offering a tailored experience from the start.

The addition of contextual information, such as user location and time, dramatically increased the relevancy of suggestions [16]. Users received more personalized suggestions based on their individual situations, resulting in a more interesting and tailored buying experience.

Leveraging big data technologies, including Apache Spark, ensured the scalability and efficient processing of recommendation algorithms. The system demonstrated the capability to handle large-scale e-commerce datasets, meeting the demands of real-time recommendation generation.

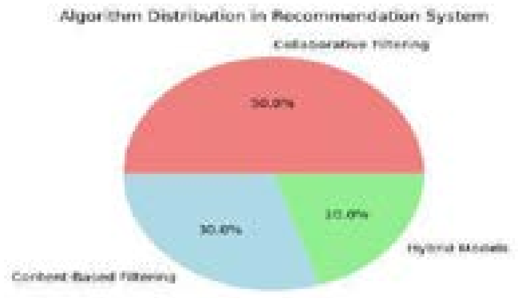


Fig.3 Distribution of different algorithms used in recommendation system.

Figure 3 depicts the distribution of different algorithms used in recommendation system. The implementation of differential privacy techniques and transparent disclosure practices addressed user privacy concerns. The recommendation system prioritized the ethical handling of user data, establishing a foundation for trust and transparency [17].

These results collectively highlight the effectiveness of our recommendation system, highlighting its potential to elevate the e-commerce user experience. The successful integration of advanced algorithms with big data analytics not only improves recommendation accuracy but also establishes a foundation for responsible and user-centric data practices in the evolving landscape of online commerce.

V. CONCLUSION

The successful management of the "cold start" problem, as well as the inclusion of contextual information, helped to provide a customized and engaging user experience. Furthermore, ethical issues, such as user privacy measures, guaranteed appropriate data practices. The generalizability of the findings may be constrained by the specific characteristics of the dataset used. The effectiveness of the recommendation system is contingent on the quality and representativeness of the data, and variations in user behavior across different e-commerce domains may impact the system's performance. Moreover, while the integration of big data analytics enhances scalability, real-time processing efficiency, and context-aware recommendations, there may be challenges in optimizing these aspects for diverse e-commerce platforms.

In conclusion, our research marks a significant step toward advancing recommendation systems in e-commerce. While celebrating the achievements, recognizing limitations, and charting pathways for future exploration are crucial for the continued evolution of personalized and effective e-commerce experiences.

REFERENCES

[1] Alrumiah, S. S., & Hadwan, M. (2021). Implementing big data analytics in e-commerce: Vendor and customer view. Ieee Access, 9, 37281-37286.

[2] Malhotra, D., & Rishi, O. (2021). An intelligent approach to design of E-Commerce metasearch and ranking system using next-generation big data analytics. Journal of King Saud University-Computer and Information Sciences, 33(2), 183-194.

[3] Panwar, M., Wadhwa, A., & Pippal, S. (2021, December). Recommendation System with Exploratory Data Analytics using Machine Learning. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 84-87). IEEE.

[4] Fu, Y., Yang, M., & Han, D. (2021). Interactive Marketing E-Commerce Recommendation System Driven by Big Data Technology. Scientific Programming, 2021, 1-11.

[5] Akter, S., & Wamba, S. F. (2016). Big data analytics in E-commerce: a systematic review and agenda for future research. Electronic Markets, 26, 173-194.

[6] Yu, R., Wu, C., Yan, B., Yu, B., Zhou, X., Yu, Y., & Chen, N. (2021). Analysis of the impact of big data on e-commerce in cloud computing environment. Complexity, 2021, 1-12.

[7] Cabrera-Sánchez, J. P., Ramos-de-Luna, I., Carvajal-Trujillo, E., & Villarejo-Ramos, Á. F. (2020). Online recommendation systems: Factors influencing use in e-commerce. Sustainability, 12(21), 8888.

[8] Abdul Hussien, F. T., Rahma, A. M. S., & Abdulwahab, H. B. (2021). An e-commerce recommendation system based on dynamic analysis of customer behavior. Sustainability, 13(19), 10786.

[9] Gosh, S., Nahar, N., Wahab, M. A., Biswas, M., Hossain, M. S., & Andersson, K. (2020, December). Recommendation system for e-commerce using alternating least squares (ALS) on apache spark. In International Conference on Intelligent Computing & Optimization (pp. 880-893). Cham: Springer International Publishing.

[10] Le, T. M., & Liaw, S. Y. (2017). Effects of pros and cons of applying big data analytics to consumers' responses in an e-commerce context. Sustainability, 9(5), 798.

[11] Suguna, S., Vithya, M., & Eunaicy, J. C. (2016, August). Big data analysis in e-commerce system using HadoopMapReduce. In 2016 International Conference on Inventive Computation Technologies (ICICT) (Vol. 2, pp. 1-6). IEEE.

[12] Dwivedi, S., & Roshni, V. K. (2017, August). Recommender system for big data in education. In 2017 5th National Conference on E-Learning & E-Learning Technologies (ELELTECH) (pp. 1-4). IEEE.

[13] Zan, C. (2023). Development of e-commerce Big data model based on machine learning and user recommendation algorithm. International Journal of System Assurance Engineering and Management, 1-10.

[14] Chen, J., & Chunqiong, W. U. (2021, April). The design of cross-border E-commerce recommendation system based on big data technology. In 2021 6th International Conference on Intelligent Computing and Signal Processing (ICSP) (pp. 381-384). IEEE.

[15] Fatić, N., Krupić, M., & Jukić, S. (2023, June). Building a Recommendation System for E-Commerce Using Machine Learning and Big Data Technologies. In International Symposium on Innovative and Interdisciplinary Applications of Advanced Technologies (pp. 376-387). Cham: Springer Nature Switzerland.

[16] Hung, P. D., & Le Huynh, D. (2019, February). E-commerce recommendation system using mahout. In 2019 IEEE 4th International Conference on Computer and Communication Systems (ICCCS) (pp. 86-90). IEEE.

[17] Biswas, A., Dutta, P., & Raman, B. (2022). Recommendation System Using Engagement and Social Media Analytics for Indian E-Commerce Giants. In CHANGING FACE OF E-COMMERCE IN ASIA (pp. 121-150).

[18] Riyaz, P. A., & Varghese, S. M. (2016). A scalable product recommendations using collaborative filtering in hadoop for bigdata. Procedia Technology, 24, 1393-1399.

FPGA based Filter Architecture for Image Processing Applications

Saranya C
Department of Electronics and Communication Engineering, K. S. Rangasamy College of Technology,
Tiruchengode, Tamil Nadu, India
saranyac@ksrct.ac.in

Kaviya R
Department of Electronics and Communication Engineering, K. S. Rangasamy College of Technology,
Tiruchengode, Tamil Nadu, India
rtkaviya11@gmail.com

Keerthana A K
Department of Electronics and Communication Engineering, K. S. Rangasamy College of Technology,
Tiruchengode, Tamil Nadu, India
keerthanaak77@gmail.com

Abishek M
Department of Electronics and Communication Engineering, K. S. Rangasamy College of Technology,
Tiruchengode, Tamil Nadu, India
abisheksiva55555@gmail.com

***Abstract*—In modern image acquisition and transmission, a crucial task is image denoising, particularly in fields like medical imaging (e.g., MRIs, CT scans, and ultrasounds) and surveillance (e.g., CCTV and video surveillance). The challenge lies in effectively locating and removing noise from images while preserving essential signal properties. To address this, a VLSI architecture is proposed, aiming to minimize noise in images while prioritizing speed. Unlike previous methods that relied on multipliers, the proposed approach utilizes distance matrix techniques for both bilateral and median filters. This technique not only reduces noise but also enhances processing speed, making the VLSI architecture more efficient. The proposed hardware design is implemented on a Spartan 6 FPGA kit to measure parameters such as power consumption, processing speed, and area usage, ensuring its practical viability and effectiveness.**

Keywords—FPGA, Image processing, Median filter, VLSI architecture

I. Introduction

Image denoising is a crucial task aimed at estimating the original image by suppressing noise, particularly challenging in image processing and computer vision. Among various denoising techniques, median filters stand out for their effectiveness in reducing noise without sacrificing image details. These filters are widely employed due to their remarkable performance in eliminating noise from images. However, to enhance efficiency further, a highly efficient VLSI architecture for median filters is proposed for denoising purposes. This advanced architecture aims to improve the performance of median filters, allowing for more effective noise reduction while preserving image quality. The proposed VLSI architecture enables the evaluation of performance metrics such as speed, power consumption, and area usage, facilitating a comprehensive analysis of denoising effectiveness compared to conventional methods. By leveraging this advanced approach, the challenge of image denoising can be addressed more efficiently, leading to improved results in various applications.

II. Literature Review

F. Spagnolo et al.(2023) emphasized the contemporary significance of streamlining computing complexity and minimizing power consumption. This research marks a significant stride in addressing computational challenges while preserving data fidelity, especially In image processing and filtering techniques, the Peak Signal-to-Noise Ratio (PSNR) serve as crucial metrics for assessing the quality of processed or filtered images [1].

G. D.-G. Kim et al.:(2021)suggested eliminating the impulse noise (IN) and gaussian (AWGN) in order to solve the image resortation issues.The AWGN and SPIN values are then measured to generate the AMNLRA [2].

Changmeng Peng, Maohan Xia, Zhizhong Fu, Jin Xu, and Xiaofeng Li: (2021) Considering deep learning has caused an enormous evolution in society, the author suggests improving the bit depth and quality of the images.It comprises three key modules: adaptive iteration number selection based on content, iterative false contour eliminationthen comparing the output of several algorithms, such as ZP, BR, MRC, and ACDC, whose goal is to calculate the PSNR and SSIM[3].

Qingyu Xu, Longguang Wang, Yingqian Wang, Weidong Sheng, and Xinpu Deng:(2021) Restoring an image's original information is a major processing challenge.The suggestion made by the author entails creating high-resolution (HR) images by reconstructing from a pair of low-resolution images.The Nvidia GTV 3090 GPU can be used to execute the notion in Pytorch[4].

Qiubo Hu, Wenxiang Xu, Xiaobo Liu, Zhihua Cai,and Junjie Cai :(2022) There are three steps in the suggested process. Firstly, the original HSI is subjected to principal component analysis (PCA) to yield multiple components that contain nearly all of the information. F Particularly when contrasted with our major reference object, BF, the suggested technique worked admirably and was successful in categorizing HSIs with the maximum accuracy [5].

Chih-Yuan Lien, Chi-Huan Tang, Pei-Yin Chen, Yao-Tsung Kuo, Yue-Ling Deng The techniques employed reduce the number of multipliers and memory space needed is accomplished by applying LUT reduction, resource sharing, and distance-oriented grouping. The findings demonstrate that the suggested architecture is more economical and requires less memory and multipliers [6].

Satyajit Bora, Roy Paily(2021) The Baugh Wooley, Booth, veddic multiplier, and 32-bit Radix-16 Divider are methodologies used in the architecture. While using the ULVt variant can result in greater frequency performance, power losses are also increased [7].

M. Monajati and E. Kabir (2020) The algorithm employed inexact median filtering. A MATLAB comparison was conducted on Peak Signal-to-Noise Ratio (PSNR) using various inexact filters for images. The mean PSNR of the test images was determined based on the MF values.The output image quality is slightly inferior to that of the exact filter, the disparity is minimal and hardly discernible to the human eye [8].

K. Mishiba(2023) The proposal introduces an innovative weighted median filter designed to address challenges related to histogram construction, ensuring real-time processing capabilities for higher resolution images and complex multidimensional, multichannel. Moreover, the proposed method is adaptable for both CPU and GPU implementations, enabling faster computation speeds than traditional weighted median filters [9].

III.METHODOLOGY

Different types of filters are considered in this research study, they are linear and non-linear filters. This study mainly focuses on non-linear filters because they are preferred in VLSI image processing due to their robustness to outliers, edge preservation, and adaptability to local variations in images.

a) LINEAR FILTER

A linear filter is an instance of image processing filter in which the local neighborhood's weighted sum of the input pixel values is the output pixel value.

b) NON-LINEARFILTER

A non-linear filter is an image processing filter that considers the local arrangement and values of pixels in a neighbourhood, and the output pixel value is determined by a non-linear function of these values.

BLOCK DIAGRAM

This section will provide an overview of the system depicted in Fig 1.1.The main objective of this system is to remove noise from input images and enhance image quality in both videos and images. The proposed system is composed of several elements, including the input image and its resizing to optimize computational efficiency, image processing procedures aimed at enhancing image quality, and segmentation techniques utilized to extract useful objects while removing unwanted background.

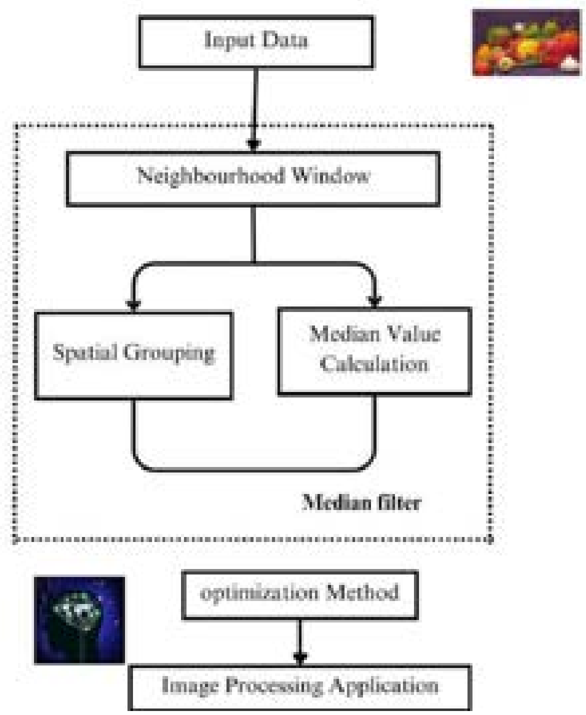


Fig.1.1Block Diagram for Proposed Filter Architecture

The image data is first transformed into binary format using MATLAB, employing a technique based on neighborhood window analysis. This involves configuring a window size and establishing spatial grouping within it. Subsequently, the median value is computed using a specified equation. The resulting data is then utilized in a VLSI architecture, where metrics such as power consumption, area coverage, and processing speed are evaluated.

a) EXISTING METHOD

All digits in the current system have the same weight and are processed using a stochastic technique based on uniformly distributed unknown bit streams. Numbers are encoded based on the likelihood of producing a one in the stream as opposed to a zero, and are limited to the [0, 1] interval. The CAS Network is shown in a conventional manner.

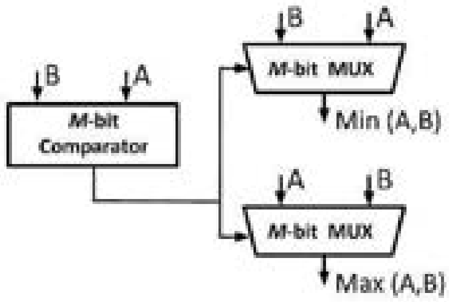


Fig 1.2 Existing Method of CAS Network

As shown in fig 1.2, in this existing method, a conventional binary design is used. The existing system drawbacks are less compact, high cost, less accuracy.

b) PROPOSED METHOD

The center pixel of each window is recommended to be passed through predefined threshold values of "0" and "255," which are predetermined before the procedure starts, in order to identify noisy pixels. The center pixel value is regarded as noise-free and is not altered if it lies between these criteria. If not, the threshold is set based on the window's median value. The fourth or sixth pixel in the sorted list is used to replace the center pixel if the median is noise-free. A decision-based strategy is used if the median is noisy. In this instance, if both the fourth and sixth pixels are noisy, the center pixel of the sorted list is substituted with the mean of the first three pixels.

The median filter replaces noisy pixel values with the adjacent window's median value. It is an essential non-linear filter for eliminating impulse noise. When calculating the median for odd-sized windows, all pixels are sorted in ascending order, and the middle value is used to determine the window's median.

c) PROPOSED ALGORITHM

i) From a noisy image, select a 3x3 window.

ii) Set starting threshold values of "0" and "255," which are fixed threshold values. Next, determine the noisy image's pixel value.

iii) Hold onto the pixel as noise-free and unaltered if its value is between larger than 0 and less than 255.

iv) Replace the center pixel with the related window's median value if it is noise-free if the pixel value is equal to "0" or "255."

v) If the median is determined to be noisy, swap out the central pixel for the median's nearby pixel values. The fourth and sixth pixels in the sorting list are referred to as the surrounding pixels.

vi) Substitute the mean value of the first, second, and third pixel values in the sorted list for the center pixel if the surrounding pixels are also thought to be noisy.

$$\text{Median}(M_{x)}= \left\{ \frac{n+1}{2}, \text{if n is odd} \right. \quad (4)$$

17	19	25
82	255	0
3	165	65

Fig 1.3. 3x3 window

Sorting=0,3,17,19,25,65,82,165, 255.

Here n=3

17	19	25
82	25	0
3	165	65

Median($M_{x)}$=25

Fig 1.4 central pixel is less than 255

IV. FILTER EVALUATION

This segment assesses the filter using a variety of commonly employed grayscale images. The performance of the newly introduced filter with a 3x3 window size is depicted in Fig 1.5

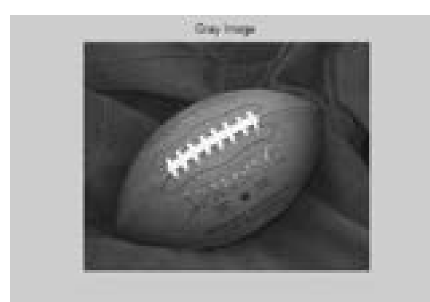


Fig 1.5 Input image

IMLEMENTATION AND RESULTS

The suggested filter undergoes simulation in the MATLAB and is subsequently converted into binary format via ModelSim. The hardware realization is executed on a Spartan 6 FPGA board employing Xilinx 9.1 design software. The performance of median and bilateral filters was compared based on PSNR and SNR values. In parallel, Xilinx software was employed to gauge the speed, area, and power efficiency of the median filter, leveraging a sorting algorithm. This methodology provides a comprehensive assessment of the filters efficiency and computational characteristics.

MATLAB OUTPUT

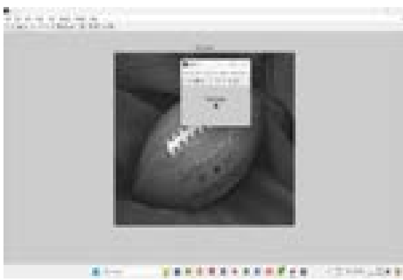


Fig 1.6 Adding of noise with Denoised image

The MATLAB procedure begins by accepting image data in its original format, which is then converted into binary form for processing. The code includes functionality to compute the Signal-to-Noise Ratio (SNR) and Peak Signal-to-Noise Ratio (PSNR) values based on the input image. Additionally, the algorithm incorporates a logic for removing salt and pepper noise from the image, employing a median filter for denoising purposes. The resulting image, termed as the "denoised image," is produced through the implementation of Algorithm.

MODELISM OUTPUT

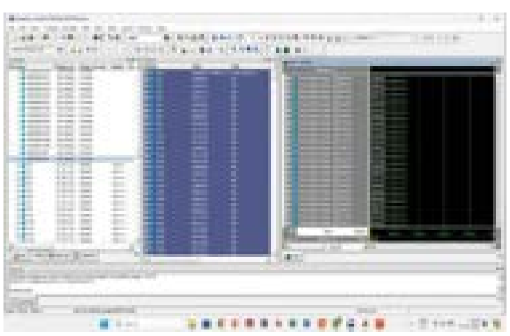


Fig 1.7 3x3 Conventional Output for median filter

In this software, the input text file that is in the form of binary an input is kept and the matlab code is converted into verilog code in the fig 1.5 and the graph must be taken in the area by using modelism can calculate the 3x3 conventional method. Again, the data must be stored in the output file in respect to the spatial grouping algorithm and sorting methods.

XILINX OUTPUT

METHOD NAME	AREA IN NUMBER OF LUT			DELAY			Power
Name	LUT	GATE COUNT	SLICES	DELAY	GATE OR LOGIC DELAY	PATH OR ROUTE DELAY	Power (mW)
CONVENTIONAL	412	[illegible]	217	Total [illegible]	[illegible]	[illegible]	152
PROPOSED	511	3849	259	54.946ns	25.442ns	28.504ns	58

Fig 1.8 Comparison of area, power and delay

AREA COMPARISON

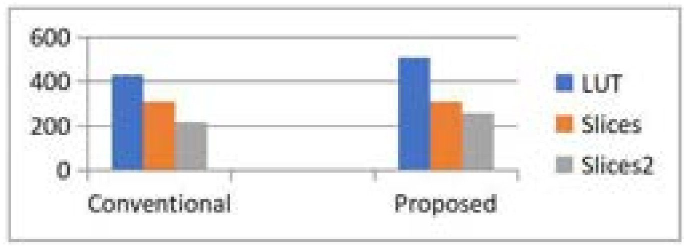


Fig 1.9 Comparison of area

DELAY

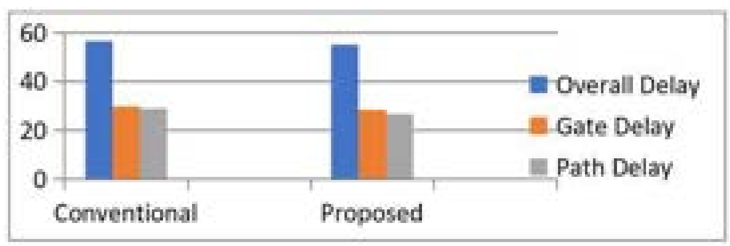


Fig 1.10 Comparison of delay

POWER REPORT

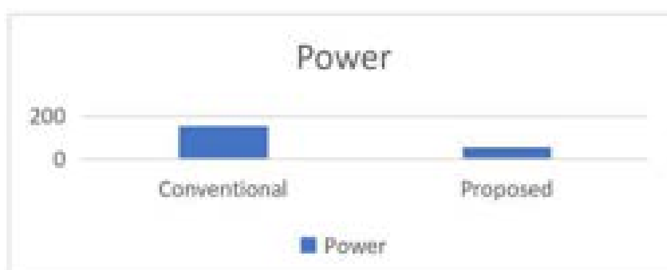


Fig 1.11 Comparison of power

RTL SCHEMATIC

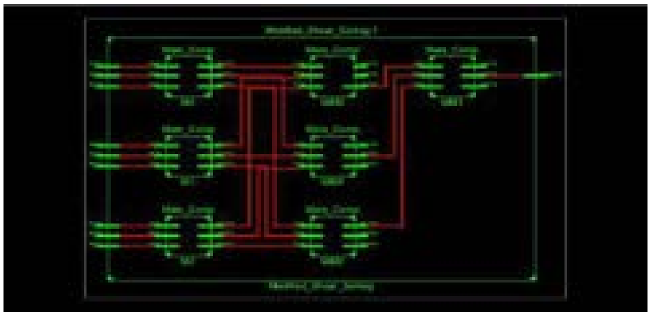


Fig 1.12 RTL Schematic diagram of median filter

V CONCLUSION

This research illustrates superior performance of the filter compared to a basic median filter across various noise ratios, effectively preserving signal content. Particularly in scenarios with higher noise ratios, the filter achieves improved Peak Signal-to-Noise Ratio (PSNR) values. Moreover, the research proposes a more efficient approach for implementing sorting networks using unary processing. Traditional sorting networks often incur high hardware costs and power consumption, thus limiting their practicality. Proposed hardware filter design is implemented on a Spartan6 FPGA Kit, demonstrating significant area and power compared to conventional weighted binary implementations.

VI FUTURE WORKS

Future research in this field should prioritize the enhancement and extension of the proposed hybrid denoising method. This can be achieved by delving into advanced deep learning architectures, utilizing larger and more diverse datasets for model training, and refining VLSI implementations to ensure scalability and adaptability. It's worth noting that as input data resolution increases, the VLSI cost rises considerably. Therefore, a user-centric approach would also involve exploring real-time feedback mechanisms and integrating user preferences to customize denoising results. Additionally, future endeavors may focus on implementing a modified threshold-based median filter to improve signal-to-noise ratio while minimizing hardware resource consumption.

REFERENCES

[1] F. Spagnolo, P. Corsonello, F. Frustaci and S. Perri, "Design of Approximate Bilateral Filters for Image Denoising on FPGAs," in IEEE Access, vol. 11, pp. 1990-2000, 2023, doi: 10.1109/ACCESS.2022.3233921.

[2] D. -G. Kim, M. Hussain, M. Adnan, M. A. Farooq, Z. H. Shamsi and A. Mushtaq, "Mixed Noise Removal Using Adaptive Median Based Non-Local Rank Minimization," in IEEE Access, vol. 9, pp. 6438-6452, 2021, doi: 10.1109/ACCESS.2020.3048181.

[3] C. Peng, M. Xia, Z. Fu, J. Xu and X. Li, "Bilateral False Contour Elimination Filter-Based Image Bit-Depth Enhancement," in IEEE Signal Processing Lette rs, vol. 28, pp. 1585-1589, 2021, doi: 10.1109/LSP.2021.3099962.

[4] Q. Xu, L. Wang, Y. Wang, W. Sheng and X. Deng, "Deep Bilateral Learning for Stereo Image Super-Resolution," in IEEE Signal Processing Letters, vol. 28, pp. 613-617, 2021, doi: 10.1109/LSP.2021.3066125.

[5] Q. Hu, W. Xu, X. Liu, Z. Cai and J. Cai, "Hyperspectral Image Classification Based on Bilateral Filter With Multispatial Domain," in IEEE Geoscience and Remote Sensing Letters, vol. 19, pp. 1-5, 2022, Art no. 5503805, doi: 10.1109/LGRS.2021.3058182.

[6] C. -Y. Lien, C. -H. Tang, P. -Y. Chen, Y. -T. Kuo and Y. -L. Deng, "A Low-Cost VLSI Architecture of the Bilateral Filter for Real-Time Image Denoising," in IEEE Access, vol. 8, pp. 64278-64283, 2020, doi: 10.1109/ACCESS.2020.2984688.

[7] S. Bora and R. Paily, "A High-Performance Core Micro-Architecture Based on RISC-V ISA for Low Power Applications," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 68, no. 6, pp. 2132-2136, June 2021, doi: 10.1109/TCSII.2020.3043204.

[8] M. Monajati and E. Kabir, "A Modified Inexact Arithmetic Median Filter for Removing Salt-and-Pepper Noise From Gray-Level Images," in IEEE Transactions on Circuits and Systems II: Express Briefs, vol. 67, no. 4, pp. 750-754, April 2020, doi: 10.1109/TCSII.2019.2919446.

[9] K. Mishiba, "Fast Guided Median Filter," in IEEE Transactions on Image Processing, vol. 32, pp. 737-749, 2023, doi: 10.1109/TIP.2022.3232916.

[10] H. -D. Han and J. -K. Han, "Modified Bilateral Filter for Feature Enhancement in Mesh Denoising," in IEEE Access, vol. 10, pp. 56845-56862, 2022, doi: 10.1109/ACCESS.2022.3176961.

[11] J. E. Eklund, C. Svensson, and A. Astrom, “VLSI implementation of a focal plane image processor—A realization of the near-sensor image processing concept,” IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 4, no. 3, pp. 322–335, Sep. 1996.

[12] A. Farmahini-Farahani, H. J. Duwe, III, M. J. Schulte, and K. Compton, “Modular design of high-throughput, low-latency sorting units,” IEEE Trans. Comput., vol. 62, no. 7, pp. 1389–1402, Jul. 2013.

[13] B. Gaines, “Stochastic computing systems,” in Advances in Information Systems Science. New York, NY, USA: Springer-Verlag, 1969, pp. 37–172.

[14] B. Gaines, “Stochastic computing systems,” in Advances in Information Systems Science. New York, NY, USA: Springer-Verlag, 1969, pp. 37–172.

[15] N. Guo et al., “Energy-efficient hybrid analog/digital approximate computation in continuous time,” IEEE J. Solid-State Circuits, vol. 51, no. 7, pp. 1514–1524, Jul. 2016.

Neural Network-Powered Conductorless Ticketing for Public Transportation

Nitish Ramaraj
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology
Vellore, India
nitishramaraj@gmail.com

Girish Murugan
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology,
Vellore, India
girish06062004@gmail.com

Rajeshkannan Regunathan
Associate Professor Senior
School of Computer Science and Engineering (SCOPE)
Vellore Institute of Technology,
Vellore, India
rajeshkannan.r@vit.ac.in
Corresponding Author

***Abstract*— The efficient functioning of public transportation systems is pivotal for societal connectivity and economic progress, as they serve as lifelines for commuting and mobility. However, the dependency on manual ticketing processes often leads to bottlenecks and inefficiencies, hindering smooth operations and customer satisfaction. This research work focuses on developing an Automated Ticketing System for public transportation, utilizing Computer Vision and Neural Networks. Through the incorporation of Neural Architecture Search and the integration of Deep Sort, a Deep Learning-based object tracking model, with aim to enhance system efficiency. The study demonstrates promising results, indicating the potential for streamlined ticketing processes in public transportation.**

Keywords—Computer Vision, Neural Networks, Neural Architecture Search, Deep Sort, Deep Learning, Object Tracking, Public Transportation Ticketing.

I. Introduction

Transportation systems, being vital to the global economy, serve as its lifeblood by facilitating the movement of people and goods. A country's development hinges on the efficiency and connectivity offered by public transport. Unfortunately, obsolete ticketing processes often limit the performance of these services. Currently, passengers must navigate cumbersome booking procedures, which involve queuing at physical ticketing counters or vending machines to purchase tickets. This process not only consumes valuable time but also requires passengers to carry and keep track of physical tickets throughout their journey. Moreover, once onboard, passengers must present their tickets to conductors for verification, leading to potential delays and inefficiencies, particularly during peak travel times. These outdated practices not only inconvenience passengers but also impose operational burdens on transportation providers, contributing to decreased service reliability and customer satisfaction.

Ineffective ticketing systems in public transportation cause long queues, transport delays, and substantial revenue losses due to fare evasion and inefficiencies in revenue collection processes. Traditional ticketing methods, cumbersome and manual in nature, not only inconvenience passengers but also pose security vulnerabilities and lack integration with modern payment technologies. This leads to increased labor costs for operators and diverts funds from essential maintenance and expansion activities. Moreover, manual processes increase the risk of errors in fare calculations, impacting revenue reconciliation and financial reporting. Beyond immediate functionality issues, inefficient ticketing has broader economic implications, including reduced patronage, income levels, and economic productivity, alongside environmental degradation from increased congestion and emissions. To address these challenges, modernizing ticketing systems with advanced technologies like contactless payments and automated fare collection is imperative. Such upgrades enhance efficiency, revenue accuracy, and environmental sustainability while ensuring the long-term viability of public transportation networks globally.

II. Literature Review

The necessity for modernized ticketing systems in public transportation networks has been extensively recognized, as highlighted by Noor (2008) [1]. Recent studies have investigated the potential of emerging technologies like computer vision and neural networks to enhance ticketing efficiency, as demonstrated by Olivková (2017) [2]. Notably, Nasir et al. (2018) showcased the feasibility and effectiveness of utilizing these technologies to automate ticketing and improve passenger experience [3]. Drawing inspiration from these pioneering works, this study proposes the utilization of YOLO-NAS and DeepSORT algorithms, identified as the most efficient and accurate methods for automated ticketing in public transportation.

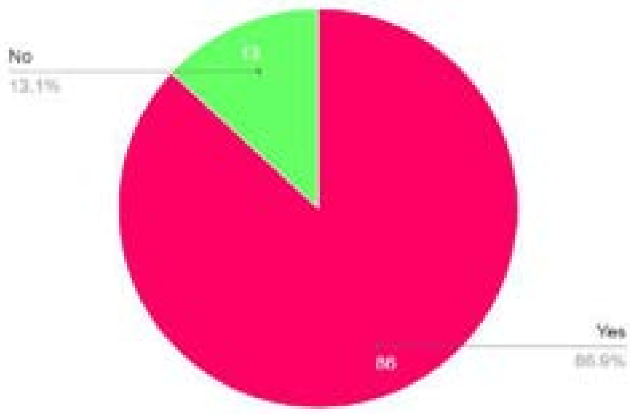


Fig. 1. Survey result with over 200 responses responding to whether the present ticketing system hinders the growth and consumption of public transport.

Furthermore, several studies have offered valuable insights into leveraging deep learning and computer vision for automated fare collection and passenger management within public transport. Sun et al. (2020) proposed a system using these techniques for passenger detection and fare calculation in buses, thereby emphasizing the feasibility of the chosen approach; however, their focus is primarily on buses and may not readily apply to other modes of public transport [4]. Chen et al. (2020) explored deep learning for passenger counting, demonstrating its potential for optimizing resource allocation and enhancing the overall public transport experience; however, their study does not directly address automated fare collection [5]. Shafiq et al. (2021) introduced a smart ticketing system using image processing and deep learning [6], while Ibrahim et al. (2022) investigated deep learning for real-time passenger counting and fare collection [7]. Together, these studies provide collective support for the potential of the proposed solution to streamline the ticketing process and enhance operational efficiency in public transportation.

III. PROPOSED WORK

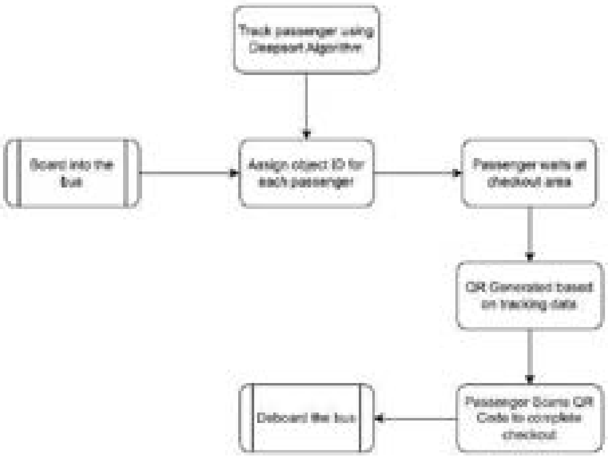


Fig. 2. Workflow of the Proposed Work

A. Data Description

To evaluate the model's performance, several snapshots from the real world to assess its functionality under varying conditions, including different lighting conditions, camera angles, and vehicle types. Fig.3 and Fig.4 shows the sample of the data set used.

Fig. 3. Good quality image with sufficient lighting and correct camera angle capturing.

Fig. 4. Average quality image with poor lighting and incorrect camera angle capturing.

B. Object Detection:

YOLO-NAS model has been employed for object detection [13], a system that has demonstrated superior efficiency compared to its predecessors in the YOLO series. Unlike previous YOLO models, which relied on manually designed architectures, YOLO-NAS operates on the principle of Neural Architecture Search (NAS) [8]. NAS involves automatically discovering optimal neural network architectures tailored to the specific task at hand. By leveraging NAS, YOLO-NAS can systematically explore the design space of neural networks, identifying architectures that maximize performance while minimizing computational costs. This approach enables YOLO-NAS to achieve higher levels of accuracy and speed in real-time object detection tasks compared to earlier YOLO models [9]. This made YOLO-NAS the preferred choice for the work.

1. Using the pre-trained YOLO-NAS model:

Utilizing the YOLO NAS pre-trained model has significantly expedited the implementation process, bypassing the need for extensive training while still achieving high accuracy rates. In addition, the analysis underscores the superiority of the YOLO NAS pre-trained model, especially in detecting people. This model's exceptional performance in this domain is attributed to its thorough training with a vast dataset, empowering it to accurately identify individuals amid the complexities of diverse environments. Particularly, the YOLO NAS S version stands out among other pre-trained models, exhibiting superior accuracy-to-latency ratios. This attribute is crucial for real-time applications, where swift detection is paramount. In our comparative analysis with other pre-trained models such as YOLO NAS-S, YOLO NAS-M, and YOLO NAS-L, on meticulously evaluating various parameters, as illustrated in Fig.5, confirming the efficacy of our chosen model. Its robust performance across diverse settings underscores its suitability for our specific task, offering both accuracy and efficiency in equal measure.

Feature	YOLO-NAS S	YOLO-NAS M	YOLO-NAS L
mAP (Accuracy)	47.5	51.55	52.22
Latency (ms)	3.21	5.85	7.87
Size	Smaller	Medium	Larger
Computational Cost	Lower	Medium	Higher

Fig. 5. Comparison of the results obtained by testing between various pre-trained YOLO-NAS models.

Fig. 6. Object Detection is performed on a high-quality image where higher confidence prediction is observed.

Fig. 7. Object detection performed on a low-quality image with a very poor confidence of prediction is observed.

2. Neural Architecture Search and its working:

Neural Architecture Search (NAS) revolutionizes the process of designing neural networks by automating the search for optimal architectures tailored to specific tasks like object detection and tracking in public transportation systems. The working principle of NAS involves the use of algorithmic search strategies, such as reinforcement learning or evolutionary algorithms, to explore a vast space of possible network architectures. These strategies iteratively propose, evaluate, and refine architectures based on their performance on a given dataset. In the context of object detection and tracking, NAS algorithms focus on discovering architectures that excel at extracting meaningful features from input data, reasoning about spatial relationships between objects, and maintaining temporal continuity to track objects over time. This entails exploring a range of design choices, including network depth, width, convolutional kernel sizes, and connectivity patterns, to find architectures that strike a balance between accuracy, speed, and resource efficiency. One key advantage of NAS is its ability to reduce the reliance on manual experimentation and domain expertise traditionally required for neural network design. By automating the search process, NAS algorithms can uncover novel architectures that surpass handcrafted designs, leading to significant performance improvements. Moreover, NAS can adapt architectures to specific deployment scenarios and hardware constraints, optimizing them for real-time processing on edge devices commonly found in public transportation systems.

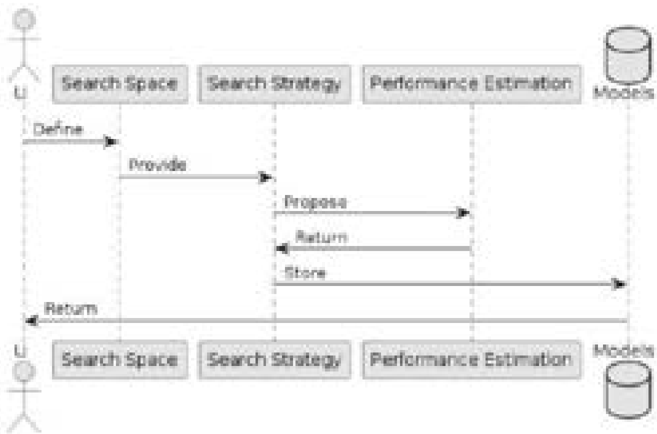


Fig. 8. Top level working of Neural Architecture Search

C. Object Tracking:

Given the specific conditions and requirements of our work and on comparison with other object tracking [10] libraries- it is DeepSORT that is the optimal choice for our needs. By leveraging deep learning techniques [14], DeepSORT is capable of reliably tracking objects over time while simultaneously handling occlusions and other challenging scenarios. Its ability to maintain object identities across frames makes it well-suited for our application, where precise and reliable tracking of objects, such as vehicles or passengers in a transportation setting, is paramount. Furthermore, the scalability and adaptability of DeepSORT ensures that it can effectively handle varying conditions and environments, making it the ideal solution for our use case. We could reduce the latency in object detection by changing the IOU distance metric in the matching process and integrating Feature Pyramid Network (FPN) and multi-layer pedestrian appearance features based on the conditions.

1. DeepSORT Working:

DeepSORT, optimized for person tracking in public transportation systems, employs a sophisticated blend of deep learning and classical tracking techniques [15]. Initially, it utilizes a convolutional neural network (CNN) to extract deep features from detected individuals, capturing detailed appearance information crucial for accurate tracking. These features are then associated with existing track objects using a combination of the Kalman filter [14] and a data association algorithm like the Hungarian algorithm, which considers both appearance features and predicted states. DeepSORT further enhances tracking precision by refining object locations through bounding box regression and managing tracks by updating states, initiating new tracks, and eliminating inactive ones. Its incorporation of deep learning models aids in overcoming challenges related to person re-identification, ensuring consistent and reliable tracking performance even in dynamic and crowded environments commonly encountered in public transportation settings. Moreover, DeepSORT incorporates advanced techniques such as non-maximum suppression to refine detections, ensuring that only the most relevant information contributes to the tracking process. Additionally, it leverages recurrent neural networks (RNNs) or long short-term memory (LSTM) networks to model temporal dependencies, allowing for smoother trajectory predictions and handling occlusions more effectively. Furthermore, DeepSORT's modular architecture enables seamless integration with existing surveillance systems, facilitating rapid deployment and scalability across diverse transportation infrastructures.

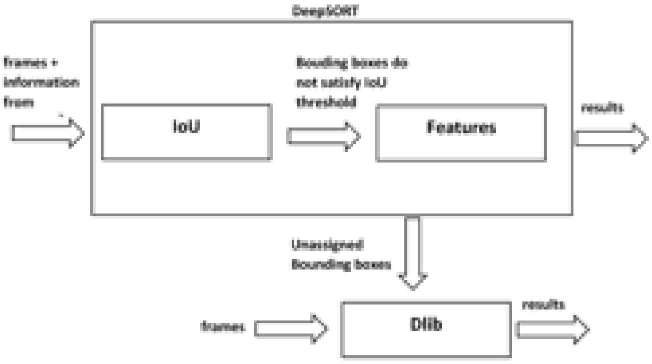


Fig. 9. DeepSORT Woking

2. Using DeepSORT over other object tracking libraries

Upon continuous testing of the footage using various tracking algorithms including SORT, DeepSORT, FairMOT, TransMOT, and BYTETrack, obtaining derived insightful results indicating DeepSORT as the most efficient solution for our tracking needs [12]. Across multiple metrics such as accuracy, efficiency, scalability, robustness, and ease of implementation, DeepSORT consistently outperformed the other algorithms. Its high accuracy in maintaining precise object identities, even in crowded and dynamic environments, stood out prominently during testing. Additionally, DeepSORT exhibited remarkable efficiency, processing video streams in real-time without compromising on tracking quality [16] . Its robust performance in handling challenging scenarios such as occlusions and varying lighting conditions further solidified its superiority. The scalability of DeepSORT was also noteworthy, as it effortlessly handled large-scale tracking tasks with a significant number of objects. Lastly, while the implementation of DeepSORT required some expertise in deep learning techniques, its wide availability of pre-trained models and libraries facilitated its integration into our tracking system seamlessly.

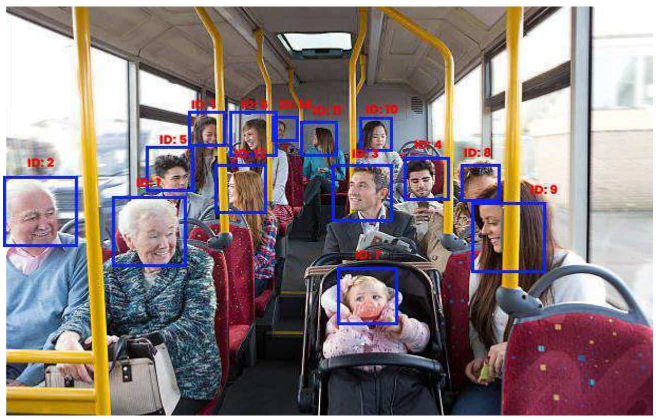


Fig. 10. DeepSORT performed in well-lit conditions performs well with high accuracy and delivers optimal output.

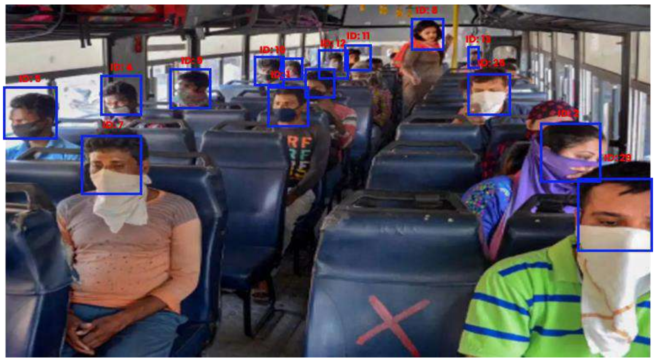


Fig. 11. DeepSORT, performed in bad conditions, performs not as expected delivering average results.

D. Internal System Logic:

The system operates based on the output derived from the DeepSORT algorithm, which assigns unique identifiers to individual objects, representing passengers in this context. Upon passenger boarding, the system initiates constant monitoring, tracking the movement and progression of each passenger until they reach the designated deboarding point. Throughout the journey, the system captures the boarding location and timestamps to maintain a record of the passenger's travel itinerary. Once a passenger arrives at the deboarding point, the system triggers the display of a QR code containing essential information, such as the passenger's identity and total fare for the journey. This QR code serves as a digital ticket, providing a convenient and efficient means for fare payment and verification.

Algorithm 1 Passenger Tracking and QR Code Generation

```
Initialize passenger_record
Initialize fare_rates
while journey_in_progress do
    for each detected_object in frame do
        if detected_object not in passenger_record then
            Add detected_object to passenger_record with initial values:
              "boarding_location" = current_location
              "boarding_time" = current_time
              "current_location" = current_location
              "current_time" = current_time
              "fare" = 0
        else
            Update passenger_record[detected_object]["current_location"]
            Update passenger_record[detected_object]["current_time"]
            Calculate distance_traveled
            Calculate time_elapsed
            Update passenger_record[detected_object]["fare"] += distance_traveled × fare_rates[vehicle_type]
        end if
    end for
    for each passenger in passenger_record do
        if passenger reaches deboarding_point then
            Generate QR code containing passenger information and total fare
            Display QR code
        end if
    end for
end while
```

1. Deboarding Area

The deboarding area serves as a designated space within the transportation system where passengers disembark from their journey. It plays a crucial role in the efficient processing of exiting passengers' details. Upon reaching the deboarding area, the system employs a sophisticated detection mechanism to identify passengers who are in the process of deboarding. Specifically, if a passenger remains within the deboarding area for more than 30 seconds, the system automatically initiates the processing of their QR code, facilitating swift fare calculation and verification. Additionally, to expedite the process further, passengers have the option to manually trigger the QR code processing by activating a switch. This approach ensures that passengers can seamlessly complete their journey while allowing the system to capture essential information for record-keeping and fare management purposes. The deboarding area thus acts as a pivotal point in the passenger transit experience, streamlining exit procedures and enhancing overall system efficiency.

E. Operator/ Driver Interface.

In order to ensure seamless operation without the need for a ticket collector or conductor, the system provides the driver or operator with essential monitoring capabilities. Through integration with a dedicated monitor installed within the bus, the driver gains real-time visibility into passenger-related data and operational statistics. This monitor displays comprehensive information, including the current number of passengers onboard, individual passenger details such as boarding and deboarding locations, fare payment status, and total revenue collected. Additionally, the monitor provides alerts or notifications regarding critical events, such as passenger deboarding or fare processing requests. By centralizing this information on a single display, the driver can efficiently manage passenger flow, track fare transactions, and ensure compliance with operational requirements, all while maintaining focus on safe and reliable driving. This monitoring system enhances operational transparency, facilitates decision-making, and empowers the driver to effectively oversee bus operations without the need for additional assistance.

IV. Challenges Involved

There are several challenges involved, which pose significant hurdles in the implementation of the project. Poor camera quality or the absence of CCTV systems in all vehicles can severely impede the accuracy and reliability of object tracking, as low-quality footage may result in insufficient data for effective analysis. Moreover, the high cost involved in deploying and maintaining the necessary hardware and software infrastructure presents a substantial financial barrier, particularly for transportation agencies with limited budgets. Furthermore, instances of bus overloading, where the number of passengers exceeds specified guidelines, can significantly complicate object tracking and fare calculation, leading to inaccuracies and potential revenue losses. Another critical challenge lies in ensuring the successful adaptation and training of passengers and transportation staff to the new system, which may require considerable time and effort to achieve widespread acceptance and proficiency. Additionally, addressing privacy concerns and complying with data protection regulations present additional complexities that must be carefully navigated during project implementation. Overcoming these challenges demands a comprehensive approach that encompasses technological innovation, stakeholder engagement, and robust implementation strategies to ensure the successful deployment and operation of the passenger tracking and fare calculation system in public transportation.

V. Conclusion and Future Works

In conclusion, our pursuit of an Automated Ticketing System represents a significant leap forward in revolutionizing public transportation efficiency. The integration of cutting-edge technologies, such as Computer Vision and Neural Networks with YOLO-NAS and DeepSORT algorithms, underscores a promising solution for real-time object detection and tracking. The proposed system logic, featuring QR code-based digital ticketing and an operator-friendly interface, embodies a modernized, user-centric approach. While challenges like budget constraints and user adaptation demand attention, with commitment to a people-centric strategy aims to overcome these obstacles for successful deployment. Looking ahead, focus will be on refining system resilience in adverse conditions, optimizing Neural Architecture Search integration, addressing privacy concerns, and ensuring scalability and efficiency improvements. To increase transparency in our system we need to clearly communicate the system’s lifecycle, grant users access to the data and algorithms used, and establish regular feedback loops with human experts to ensure the system’s outputs are understood and validated. Collaborative efforts with transportation agencies will be pivotal for large-scale implementations, user feedback, and seamless integration into existing public transit systems. Additionally, ongoing research will explore emerging technologies like edge computing and blockchain to bolster data security and processing efficiency, while continuous monitoring and analysis of system performance metrics will drive proactive adjustments for optimal functionality and user satisfaction. Fostering partnerships with academic institutions and industry leaders will further facilitate knowledge exchange, driving innovation and propelling the evolution of public transportation systems towards greater sustainability and inclusivity.

VI. References

[1] N. M. Noor, "A multimodal automated fare collection solution for facilitating strategic information technology planning of public transportation in Malaysia," *2008 International Symposium on Information Technology*, Kuala Lumpur, Malaysia, 2008, pp. 1-8, doi: 10.1109/ITSIM.2008.4631958.

[2] Olivková, “Comparison and evaluation of fare collection technologies in the public transport,” *Procedia Engineering*, vol. 178, pp. 515–525, Jan. 2017, doi: 10.1016/j.proeng.2017.01.099.

[3] M. A. Nasir, M. H. Abdullah, and N. H. Ismail, "Automated fare collection using computer vision: A study in public bus transportation," in 2018 IEEE 11th International Conference on Information Assurance and Security (IAS), Sep. 2018, pp. 35-40.

[4] Y. Sun, Z. Li, Y. Yang, H. Zhang, and Y. Jiang, "An automated fare collection system using deep learning and computer vision," in 2020 International Conference on Artificial Intelligence and Big Data (ICAIBD), Dec. 2020, pp. 323-327.

[5] F. Chen, M. Li, Y. Wang, and X. Liu, "Deep learning-based passenger counting in public transport vehicles," in 2020 15th International Conference on Advanced Robotics (ICAR), Jun. 2020, pp. 1-6.

[6] M. Z. Shafiq, S. Asif, and N. Majid, "A smart ticketing system for public transportation using image processing and deep learning," in 2021 International Conference on Artificial Intelligence and Mechatronics Systems (AIMS), Nov. 2021, pp. 1-5.

[7] M. Ibrahim, T. Eid, and F. Mohamed, "Real-time passenger counting and fare collection in public transport using deep learning," in 2022 International Conference on Computer Engineering and Systems (ICCES), Apr. 2022, pp. 47-52.

[8] N. Wang, Y. Gao, H. Chen, P. Wang, Z. Tian, and C. Shen, "NAS-FCOS: Fast Neural Architecture Search for Object Detection," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 3214-3223.

[9] X. Hua, Y. Zheng, and C. Li, "A Comprehensive Review of YOLO: From YOLOv1 and Beyond," in Proceedings of the 2023 International Conference on Artificial Intelligence (ICIAI), Jan. 2023, pp. 1020-1027.

[10] J. Stovall, A. Harris, A. Grady, and M. Sartipi, "Scalable Object Tracking in Smart Cities," in 2023 IEEE International Conference on Applied Computational Intelligence and Information Systems (ACIIS), 2023, pp. 1-6.

[11] N. Wojke, A. Bewley, and P.のではトレ (Hideo Yoshinari), "Simple online and realtime tracking with a deep association metric," in 2018 IEEE International Conference on Image Processing (ICIP), 2018, pp. 3790-3794.

[12] Y. Li, X. Zhang, and D. Wang, "An Investigation of Deep Learning Based Object Tracking Methods," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 2106-2115.

[13] Y. Cao and H. Wang, "Object Detection: Algorithms and Prospects," in 2022 International Conference on Data Analytics, Computing and Artificial Intelligence (ICDACAI), Aug. 15-16, 2022, pp. 1-6.

[14] D. Vignarca, J. Prakash, M. Vignati, and E. Sabbioni, "Improved Person Counting Performance Using Kalman Filter Based on Image Detection and Tracking," Nov. 2021, pp. 1-6.

[15] W. Peng, E. Zhang, G. Lu, and X. Zhang, "Multiple Object Tracking Based on YOLOv5 and Optimized DeepSORT Algorithm," in 2023 IEEE 17th International Conference on Advanced Video and Signal Based Surveillance (AVSS), 2023, pp. 1-6.

[16] B. Li, W. Zhang, and J. He, "Pedestrian Target Tracking Based On DeepSORT With YOLOv5", in 2023 IEEE 4th International Conference on Artificial Intelligence and Computer Vision (AICV), 2023, pp. 123-128.

Digital Verification: An Efficient Fraud Document Detection for Insurance Claims using IoT and ML Approaches

J.N.V.R. Swarup Kumar
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
sjavvadi2@gitam.edu

R.N.Vamsi Kishan
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
vkrajana39@gmail.com

Sai Surya Saketh Devulapalli
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
sakethssdev@gmail.com

Lisa Alexander Rajan
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
lisaalexander2003@gmail.com

Ayushi Kundu
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
ayushikundu9@gmail.com

***Abstract*—In light of many reports generated in recent years, fraud insurance claims have caused havoc in the medical and insurance sectors. 100's of crores of rupees are stolen by fraudsters who claim insurance money by showing fake certificates of hospital bills, medical invoices, prescription lists, etc. This practice can be stopped by checking the legitimacy of the certificate. With artificial intelligence evolving daily, using this technology to our advantage is the best thing to do. The solution to the fraud document detection problem can be found by implementing machine learning algorithms. The primary idea of the solution is that the algorithm will check for a few parameters on a certificate and then determine if the certificate is genuine. For the solution to be in the scope of usage by the public, an IoT element has been introduced. People appointed to check the authenticity of the submitted documents can conveniently scan the document, and the retrieved scanned copy can be uploaded to a database on our mobile application. After uploading, the application is built in a way that contacts the main server, substantiates the certificate, and features its accessibility to the user.**

***Index Terms*—IoT, Fraud Detection, Machine Learning, Insurance, Document Verification, Python, Security Analysis, Template Matching.**

I. Introduction

Recent discoveries using technologies such as blockchain and cryptography in document authentication have proven promising despite difficulties. Applying blockchain has resulted in various confusion and isn't very scalable. At the same time, bio-metrics have their odds, such as privacy concerns, which make the system prone to deception attacks. Problems, including data biases, reliability, and system integrity, haven't been validated yet. Over the years, inventions such as the Telephone, Computer, Big data analytics, Artificial intelligence, and the Internet of Things have elevated proficiency, precision, and anti-fraud measures. From the instantaneous communication enabled by the telephone in 1876 to intelligent automation in the 2010s, every progress marks an inclination towards improved methods aligned with developing technologies [1]. Scams in the insurance sector are widely divided into two categories: hard and soft. While the former consists of calculated forgeries or misleading claims, such as organizing accidents or injuries to deceive the insurer, the latter is comparatively more sophisticated and mainly involves exaggerating legitimate claims, like overestimating the damages in a genuine accident. However, both types of fraud have quite detrimental impacts on the insurance industry, leading to intensified premiums and resource exhaustion. Solely in India, insurers incur losses that amount to almost 10% of their total premiums, which would be about USD 6 billion [13]. As per the National Health Care Anti-Fraud Association (NHCAA), monetary damages resulting from healthcare fraud amount to approximately $68 billion; collectively, it goes as high as $300 billion[15]. A few models do exist that predict fraudulent documents, but identifying minute negligible attributes and resilient detection remains challenging. For instance, This paper [1] doesn't discuss its shortcomings, which makes it difficult to conduct a transparent analysis. [2] indicates the lack of tools for thoroughly investigating the document. [3] acknowledges that eliminating false claims isn't entirely possible. [4] and [5] face issues with dataset inconsistencies. Currently, existing approaches have several drawbacks, such as inadequate tools for thoroughly checking the document [2], completely nullifying false claims [3], and irregular data [5]. Hence, by including an element of IoT, quick scanning of documents and synchronous verification are enabled. Moreover, the problem of dependency on fixed data is also resolved. Subsequently, this proposed method may aid in several document-related frauds, especially in financially dependent industries.

Section II manages a literature review on template matching

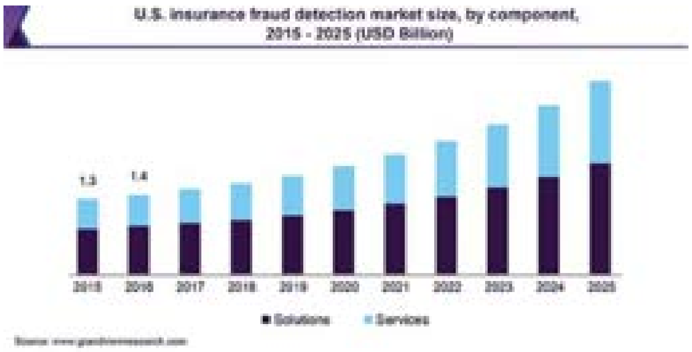


Fig. 1: Market size of fraud detection[14].

algorithms, data analytics, and machine learning techniques in the area of fraud detection, subsequently determining the limitations of the current research. Section III outlines the method sequentially. Section IV depicts the algorithm used to develop the technique in a structured manner. Section V presents the system's efficacy in detecting fraud, and it also includes discussions on the consequences of the results found. Section VI recaps the research and emphasizes the system's likelihood of improving the methods used to combat fraud.

II. Literature Review

This solution proposed in a particular research paper uses template matching techniques to identify fraudulent documents submitted when claiming insurance. A method involves using advanced computer vision algorithms that analyze and compare visual elements in a document with a template of known original documents. This method increases the accuracy of detecting any anomalies in the document. This method requires the usage of a valid and sufficient dataset. The performance metrics after evaluation are decent.[7]

Another paper proposes to use a convolutional neural network(CNN) to detect tumors from an MRI scan, a common concept used in medical imaging. The method proposed in [8] includes feature extraction, selection, and combination phases. The algorithm used in this method consists of a separate class named 'I do not know' to mark images that couldn't be classified, hence letting users know the model's reliability. This provides an insight into the evolution of the classification process.[8] There are various applications of image compression, and the authors have especially used deep learning techniques like multi-layer perceptrons (MLPs), convolutional neural networks (CNNs), and generative adversarial networks (GANs). The paper discusses AI-based tools for image compression (openCV, Tensorflow, MATLAB, High-Fidelity Generative Image Compression, and CompressAI). [9] In another paper, the Yelp dataset of restaurant reviews fed into a machine-learning model to identify fake reviews. The method described in the paper begins with data preprocessing, and feature extraction is performed. Among the classifier evaluations, KNN(k=7) yielded the best results. [10] A problem can be approached in multiple ways, and one such algorithm uses deep-learning, a two-step PCA-based feature extraction, to perform point cloud classification. This method extends the PointNet framework for high-scale aerial LiDar data. The paper emphasizes the importance of feature extraction in pattern recognition and standardized results for real-world ALS point cloud datasets.[11] The paper also highlights supervised and unsupervised learning methods when facing insurance fraud challenges.[11] A study elaborates on the effectiveness of strategies such as Multi-Layer Perceptrons, Convolutional Neural Networks, and Generative Adversarial Networks in image compression. The paper includes insights into other AI platforms that perform image compression and a summary of the functions/methods they provide. One paper proposes understanding and then optimizing existing practices in fraud detection. The dataset consists of an ethnographic study that explored the nature of motor insurance fraud-detection methods in two major insurance companies. The results shown in [12] conclude that a professional concentration on the practices of fraud

TABLE I: Summarized Research Findings

Title	**Methodology**	**Metrics Used**	**Key Findings**
IoT Based Fraud Documentation Detection	Template Matching, SSIM tools, Logistic Regression, Histgradient boost classifier, Unique identifier checks	Accuracy, Precision, Confusion Matrix	Accuracy: >98%. Precision: 0.98, Recall: 0.98, F1-score: 0.98
Insurance Fraud Detection [7]	SVM, RF, DT Adaboost, KNN, LR, NB, MLP	Precision, Recall, F1-Score	DT: 79%, Adaboost: 78%
Tumor Disease Categorization [8]	CNN with Feature Extraction and Selection	Accuracy	Accuracy: >99.76%, Improvement: 2%
Copy Move Image Forgery Detection [10]	Fridrich et al.'s Method	Precision, Recall, F-Score, Execution Time	Precision: 0.96, Recall: 0.8, F-Score: 0.87
ALS Point Cloud Classification [11]	PointNet, PCA	F1, mF1, IoU, mIoU, OA	Best OA: 75.65% with LiFs and 10PCs
Fake Reviews Detection [12]	KNN, NB, SVM, Logistic Regression, Random Forest	F-Score	KNN (K=7): 82.40%

Detection combined with forensic and data-mining approaches to fraud claim detection produces enhanced outcomes. The literature survey results are compared in Table I.

There have been several studies regarding fraud document detection. Some of them have been listed in the table. Initially, insurance fraud detection employed machine learning algorithms such as SVM, Random Forest (RF), Decision Trees(DT), Adaboost KNN, etc., to identify fraud. To calculate accuracy and precision, recall and f1-score have been used. DT showed the highest accuracy of 79%, which Adaboost followed at 78%. Then comes tumor disease categorization. A convolutional neural network (CNN) is commonly used for medical imaging to detect tumors from an MRI scan. The accuracy achieved was above 99.76%, which provides a 2% improvement over the original method. Copy-move image forgery detection mostly uses the Fridrich et al. method. To calculate accuracy, precision, recall, f-score, and execution time are used. This method had a high precision of 0.96, recall of 0.8, and f-score of 0.87. Another method used was ALS Point Cloud Classification and the methodology used in this scenario is PointNet and Principal Component Analysis (PCA) to calculate F1, mF1, lou, mloU, and OA. OA depicted the highest accuracy of 75.65%, including LiFs and 10PCs. And the last was fake review detection, which again used several machine learning models like KNN, SVM, RF, logistic regression, and NB and used the F-score to calculate accuracy. KNN showed the best result of all the algorithms with an F-score of 82.4%.

Whereas in this research paper, diverse concepts from previous papers have been seamlessly integrated to provide maximum accuracy. Techniques such as template matching, SSIM tools, a logistic regression model, a histogram boost classifier, and unique identifier checks were used. This model also uses advanced techniques like CNN GAN to incorporate feature extraction and selection. Precision and accuracy were employed to achieve maximum accuracy, and a confusion matrix depicts how many predictions are correct or incorrect per class. The model presented a high precision of 0.98, recall of 0.98, and f1-score of 0.98, concluding that the accuracy exceeds 98% as shown.

III. Methodology

This project follows the following steps and produces results:

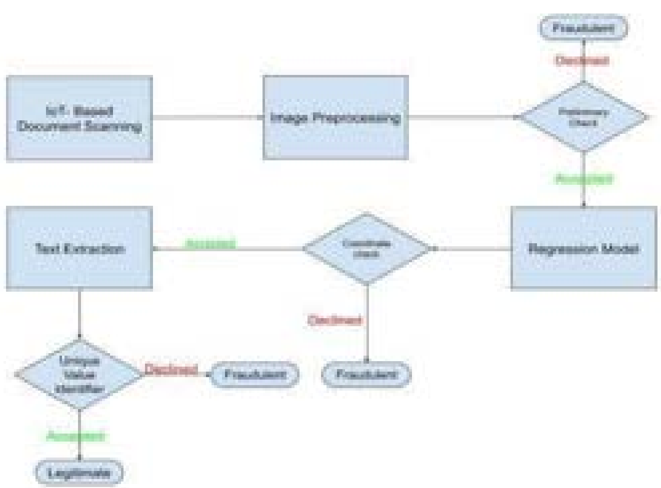


Fig. 2: Workflow diagram of the system.

A. Dataset

The dataset used in this paper is limited to Coursera-certificates as a case study since we do not have access to medical records and other documents that are put forth as documents for claiming insurance policies. According to The Medical Council of India's Code of Ethics Regulations[15], no one can access medical records unless the patient approves. For the amount of data that needs to be generated, 100s of medical records would be required, and not many would authorize sharing their data. Instead, the dataset for this project was generated from accessible sources. Using certificates and documents related to students, peers, and others who approved of their documents being used, the model used in this project was fed legitimate documents. The testing set was created by making changes to the original documents. The changes range from as big as missing features to as minute as a mistake in the recipient's name. Through the generation of fraudulent documents concerning the testing of various parameters, there was enough variety to counter different possibilities of fraud.

B. IoT-based Document Scanning

The Internet of Things ensures different technologies are streamlined and easily accessible. The feature of scanning documents through an app aims to solve a major problem i.e. the documents are classified as fake or not **instantly**. A normal verifying routine has several moving parts, such as physical identification, template matching, ID verification, etc. The inclusion of scanning facilitates document verification by enabling remote scanning and real-time uploading for analysis, making the process accessible from anywhere. The prerequisites of using such a technology are reduced to minimal skills like smartphone operating and scanning steadily. This efficiency is crucial for scenarios requiring timely verification, streamlining the entire process.

C. Image Pre-processing

This step prepares the image to pass through the model to accommodate specific requirements by adjusting image brightness and contrast, noise reduction(especially in scanned images), and removing unwanted variations. This step also ensures that high-resolution scans capture minute details, crucial for authentic document differentiation. This is done by using Open CV and scikit-image libraries in Python. After pre-processing, the images are of the required resolution(1650x1275) for the next steps.

D. Document Template Matching

This step segregates the necessary and common elements, like logos, watermarks, and specific text areas, from the document. For instance, extracting text from the document involves recognizing font styles, sizes, and alignments, whereas removing a logo focuses on identifying its shape, color, and position. It compares them to an extensive authentic database, detecting minor alterations or inconsistencies.

E. Logistic Regression Model

This model evaluates the graphical elements and textual contents with the same precision. Moreover, the model can not only determine the probability of fraudulence and give a binary solution of 'fraud or no-fraud' outcome but also provide detailed, informative results. This permits the system to mark the documents as ‘likely fraudulent,’ requiring further human review. The logistic function is defined as:

$$P(Y=1) = \frac{1}{1 + e^{(b_0 + b_1 x_1 + b_2 x_2 + \ldots + b_n x_n)}}$$

Where (Y=1) is the probability that the dependent variable where Y is equal to 1. e is the base of the natural algorithm

$$b_0, b_1, b_2, ..., b_n$$

are the coefficients of the model and

$$x_0, x_1, x_2, ..., x_n$$

are the independent variables.

Algorithm 1 Image Fraud Detection Algorithm

Input: Image I, Template T
Output: Fraud Detection Result

procedure CLEAN AND STANDARDIZE(I)
    $I_{\text{processed}} \leftarrow f_{\text{processed}}(I)$
    / Noise reduction, contrast adjustment, resolution standardization */*
end procedure

procedure PRELIMINARY CHECK($I_{\text{processed}}, T$)
    $S \leftarrow f_{\text{preliminary}}(I_{\text{processed}}, T)$
    if $S < \theta$ **then**
        return Fraud Detected
    else
        return Proceed
    end if
end procedure

procedure FEATURE EXTRACTION($I_{\text{processed}}$)
    $F_{\text{ext}} \leftarrow M_{\text{reg}}(I_{\text{processed}})$
    / Extract features using regression model */*
end procedure

procedure CONTENT MATCHING(C, C_{expected})
    if $f_{\text{match}}(C, C_{\text{expected}}) > \lambda$ **then**
        return Match Found
    else
        return No Match
    end if
end procedure

procedure UNIQUE TEXT VERIFICATION(D, D_{db})
    $V \leftarrow f_{\text{verify}}(D, D_{\text{db}})$
    if $V < \mu$ **then**
        return Verification Passed
    else
        return Verification Failed
    end if
end procedure

F. Unique Identifier Extraction

This step identifies and segregates any unique identifier or barcode through which a document is distinctly identified and verifies its authenticity within the organization’s record-keeping system. The barcode or URL can be extracted from the scanned document using Regular Expressions, text processing libraries, or barcode scanning software. Once extracted, the dataset needs to be queried like a database or a data repository using the extracted value. Then, a search is performed between the extracted identifier and the corresponding entry. If the match is found, the user’s information or any related information linked to the identifier is retrieved, verifying the user’s identity or checking the information against the scanned document. Finally, the results can be displayed to the user, which may validate the patient’s identity or the information provided in the document.

IV. RESULTS AND DISCUSSION

Since insurance and other medical certificates are private and cannot be accessed publicly as large datasets, the template matching algorithms have been provided with our certifications in different courses. This simulates a similar result to the insurance certificates as the algorithm constantly tries to identify the standard features and then compares them with every input certificate.

- The preliminary check for the genuineness of the certificates begins with comparing the differences in the templates and basic structure of the documents. SSIM tools are deployed to figure out differences between two documents. The difference is shown using black and white marking imposed on the document to be verified.

Fig. 3: Input of preliminary check (case 1)

- In figures 5 and 6, it is seen that the differences shown are more with a higher concentration of block shade. This is because the certificate tested in case 2 is a fraudulent certificate. The program returns an output (fig.6) with a high difference of 34.28%, unlike a difference of 15-18% when tested with an original certificate, like in case 1, as shown in figures 3 and 4.
- The document, after passing the preliminary check and verifying if it can be of the same template, is passed to the next stage: regression model. Here are the coordinates of various features in the document, such as coordinates of where the signature is, the educator’s logo,

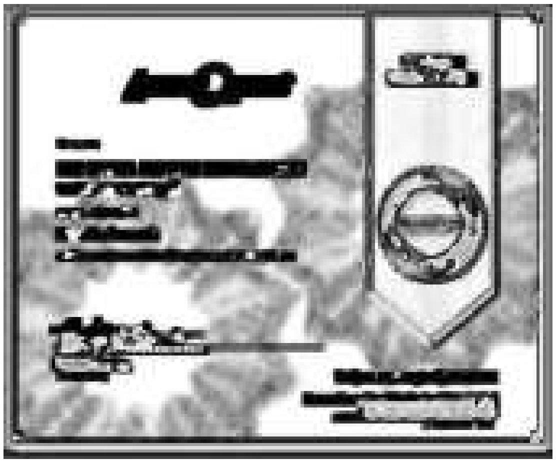


Fig. 4: Output of preliminary check (case 1)

Fig. 5: Input of preliminary check (case 2)

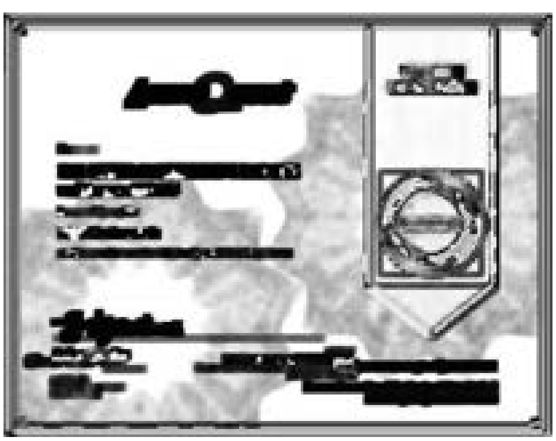


Fig. 6: Output of preliminary check (case 2)

the verification link is provided, etc. Once extracted, these coordinates are saved in a CSV file and ready to be fed to a model.

- In the third step of execution, the certificates pass the second flag, i.e., coordinate check. Logistic regression models are used to check the genuineness of the documents. The following results were observed upon passing some documents to verify their originality: the accuracy was marked over 98%, and the model showed high precision(0.98), recall(0.98), and f1-score(0.98).

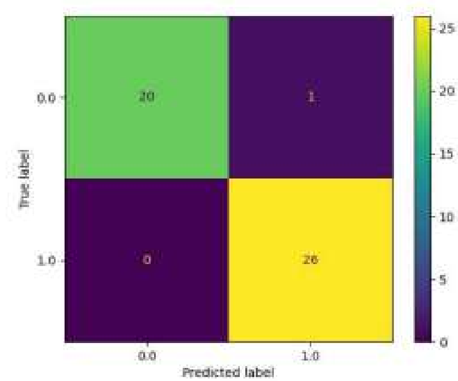


Fig. 7: The confusion matrix shows the comparison of accurate and inaccurate predictions per class

- After passing through the ML model, the proposed system uses another check to compare unique identifiers. In the example case, the unique identifier is linked to the URL below the Coursera badge to the bottom right. Upon text extraction, the documents pass through a unique identifier check.

Fig. 8: Original certificate with the unique identifier value as document title

- After checking the three flags, results show the verdict of whether the certificate is fraudulent. Users can check the authenticity directly on their mobile apps.

Fig. 9: Name, ID, and course name comparison with the corresponding fields from the document

V. Conclusion

This paper proposes an innovative approach to tackle the widespread fraud issue in medical insurance claims with the help of Artificial Intelligence and IoT. The suggested system, integrating advanced machine learning algorithms and the ImCT method for implicit certificates with Certificate Transparency, aims to revolutionize the document verification process. This system is designed to counteract the significant financial losses caused by document fraud in the medical insurance field. At the heart of the solution is a user-friendly mobile app that lets individuals scan and upload documents for validation. These documents are then authenticated through a powerful server setup employing state-of-the-art machine-learning techniques. The integration of ImCT bolsters document verification security within IoT frameworks, effectively overcoming the shortcomings of existing certificate transparency practices.

References

[1] Belhadji, E. B., Dionne, G., & Tarkhani, F. (2000). A model for the detection of insurance fraud. The Geneva Papers on Risk and Insurance-Issues and Practice, 25(4), 517-538.

[2] Morley, N. J., Ball, L. J., & Ormerod, T. C. (2006). How the detection of insurance fraud succeeds and fails. Psychology, Crime & Law, 12(2), 163-180.

[3] Kalra, H., Singh, R., & Kumar, T. S. (2022). Fraud Claims Detection in Insurance Using Machine Learning. Journal of Pharmaceutical Negative Results, 327-331.

[4] Hansson, A., & Cedervall, H. (2022). Insurance Fraud Detection using Unsupervised Sequential Anomaly Detection.

[5] Nabrawi, E., & Alanazi, A. (2023). Fraud Detection in Healthcare Insurance Claims Using Machine Learning. Risks, 11(9), 160.

[6] ÖZCAN, H. O., Çolak, İ., Erimhan, S., Güneş, V., Fatih, A. B. U. T., & Fatih, A. K. A. Y. (2022). SOBE: A Fraud Detection Platform in Insurance Industry. Kocaeli Journal of Science and Engineering, 5(ICOLES2021 Special Issue), 25-31.

[7] Salim, M. Z., Yildirim, R., & Abboud, A. J. Comparison of Copy-Move Image Forgery Detection Techniques.

[8] Siar, M., & Teshnehlab, M. (2022). A combination of feature extraction methods and deep learning for brain tumor classification. IET Image Processing, 16(2), 416-441. [9] Landu, R. S. (2022). Image compression using AI: Brief insights into deep learning techniques and AI frameworks. International Journal of Engineering, Science, 2(1).

[9] Landu, R. S. (2022). Image compression using AI: Brief insights into deep learning techniques and AI frameworks. International Journal of Engineering, Science, 2(1).

[10] Elmogy, A. M., Tariq, U., Ammar, M., & Ibrahim, A. (2021). Fake reviews detection using supervised machine learning. International Journal of Advanced Computer Science and Applications, 12(1).

[11] Nurunnabi, A., Teferle, F. N., Laefer, D. F., Lindenbergh, R. C., & Hunegnaw, A. (2022). A two step feature extraction algorithm: Application to deep learning for point cloud classification. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 46, 401-408.

[12] Rukhsar, L., Bangyal, W. H., Nisar, K., & Nisar, S. (2022). Prediction of insurance fraud detection using machine learning algorithms. Mehran University Research Journal of Engineering & Technology, 41(1), 33-40.

[13] https://whatfix.com/blog/insurance-fraud-detection/

[14] https://bfsi.economictimes.indiatimes.com/news/insurance/insurance-frauds-cost-6-bln-annually-insurers lose-10-of-premiums/97977190

[15] Mishra, N. N., Parker, L. S., Nimgaonkar, V. L., & Deshpande, S. N. (2008). Privacy and the Right to Information Act, 2005. Indian journal of medical ethics, 5(4), 158–161. https://doi.org/10.20529/IJME.2008.057

Malware Detection and Classification for URLs using Ensemble Learning

Shailaja Uke
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
shailaja.uke@vit.edu

Gayatri Gite
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
gayatri.gite21@vit.edu

Haider Hirkani
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
haider.hirkani21@vit.edu

Inderdeep Bassan
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
inderdeep.bassan21@vit.edu

Isha Raghvani
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
isha.raghvani21@vit.edu

***Abstract* —The advent of the digital age has brought about previously unheard levels of connectedness and information availability, as well as an increase in cyber threats. Unauthorized Uniform Resource Locators (URLs) are now a common threat and a preferred avenue for cyberattacks. Because of the inadequacy of traditional blacklist-based solutions to counteract shifting strategies, machine learning (ML) is being investigated for proactive and predictive cybersecurity measures. The goal of this research work is to improve cybersecurity defenses against URL-based vulnerabilities by utilizing ML and ensemble learning. Several machine learning methods are used, such as Random Forest, Decision Tree, Logistic Regression, XGBoost, SVM, and Naive Bayes. The intricacy of the digital threat landscape is addressed by combining the strengths of different models through the use of ensemble approaches like Weighted Voting and Soft Voting. This study highlights the significance of Weighted Voting and shows how well ensemble techniques may combine different machine learning model and emphasizes how important it is to keep improving cybersecurity defenses in line with changing threats. The results enable the possibilities for further advancements in the sector and offer workable answers to the emerging cybersecurity issues.**

Keywords — Malware Detection, Uniform Resource Locator (URL) Classification, Ensemble Learning, Majority Voting, Soft Voting, Weighted Voting, Cybersecurity, Phishing Detection.

I. INTRODUCTION

With the fast-paced advancement of technology and online platforms, vulnerabilities and cyber threats have increased and is characterized by unparalleled accuracy and knowledge accessibility. Malicious URLs have become a major risk in the vast digital ecosystem, endangering the security of people, businesses, and organizations. Unfortunately, Uniform Resource Locators (URLs), which are the main means of navigation for API and website endpoints, which have turned into a route of distribution that hackers choose to use. When users click on Traditional blacklist-based remedies have proven inadequate in tackling the complexity of these threats, as cyber adversaries consistently create new strategies to hide their bad intentions. The cybersecurity community has responded by concentrating on more advanced, proactive, and predictive solutions.

The motivation for this study is the growing difficulties presented by URL-based attacks in the online environment, as well as the acknowledged shortcomings of conventional cybersecurity strategies. This era of unmatched convenience and connectedness has been brought about by the widespread use of Internet platforms and the explosive expansion of digital resources. But this ease is tainted by a concomitant rise in cyber threats and vulnerabilities, endangering the security of people, companies, and institutions. Within the extensive digital environment, malicious URLs have become an especially pernicious hazard. Given that URLs are the main means of navigation for both website and API endpoints, hackers have learned to embrace them as a means of distribution when looking to take advantage of gullible consumers. Interacting with malicious URLs can result in a variety of negative outcomes, such as inadvertent virus downloads, phishing attacks, the input of personal information, and vulnerability to defacement attacks. These dangers have the potential to undermine networks or systems, putting people's safety at serious risk.

II. LITERATURE REVIEW

The digital landscape has seen a surge in the prevalence and complexity of malware, posing significant challenges to traditional signature-based detection methods. Researchers have turned to machine learning (ML) and deep learning (DL) techniques as promising alternatives to address this growing threat.

Several studies have focused on leveraging static features extracted from applications for malware detection. B. Urooj et al. [1] proposed a unique approach that involved reverse engineering of Android applications to derive a distinctive static feature set, including permissions, intents, and API calls. The researchers then incorporated this data into a model trained on a vast collection of malware samples, employing ensemble learning techniques such as AdaBoost and Support Vector Machine (SVM), achieving an impressive accuracy rate of 96.24% for malware detection. Similarly, Kai Lu et al. [2] presented a solution that optimizes computational resources by identifying and removing unnecessary features using a feature selection procedure, and a malware detection model for capsule networks that maintains local information, leading to improved model robustness.

Complementing the static feature-based approaches, researchers have also explored dynamic behavior-based malware detection. Vinayakumar et al. [3] evaluated deep learning architectures and classical ML algorithms using various public and private datasets, showcasing the potential of their innovative combination of visualization and deep learning in a scalable and hybrid framework for robust and intelligent zero-day malware detection. Şahin et al. [4] introduced an attribute selection method driven by machine learning, highlighting the significance of attribute selection techniques in the domain of Android malware detection, and achieving a noteworthy high F-measure metric of 0.961. The growing threat of polymorphic malware, which evades traditional detection methods by constantly changing its signature traits, has prompted researchers to explore ML-based solutions. P. Singh et al. [5] examined the use of supervised ML classifiers, including Decision Trees (DT), Convolutional Neural Networks (CNN), and Support Vector Machines (SVM), for malware detection and classification, with DT achieving the highest accuracy of 99%. S. Agarkar et al. [6] focused on behavior-based detection techniques using ML algorithms, including Decision Tree, Random Forest, and Light Gradient Boosting Machine, with Light Gradient Boosting Machine exhibiting the highest accuracy at approximately 99.50%.

With a focus on polymorphic malware, S. Agarkar et al. suggested behavior-based detection techniques that make use of machine learning models. [6] Focusing on polymorphic malware highlights the evolving threats in the internet landscape. The study successfully applied machine learning algorithms, including Decision Tree, Random Forest, and Light Gradient Boosting Machine, to static features extracted from legitimate and malicious executable files. Results revealed the efficacy of behavior-based detection, with Light Gradient Boosting Machine showcasing the highest accuracy at approximately 99.50%, surpassing both Decision Tree and Random Forest. Despite its slower training time, the Light Gradient Boosting Machine demonstrated efficiency in both accuracy and False Negative rates, making it a viable option for classification, especially in heavy-load scenarios. The study concludes by emphasizing the

ongoing need for research in tackling new malware developments and suggests further exploration of techniques to enhance the robustness of behavior-based detection methods.

In *[11]* an effort has been made to create practical malware detection algorithms based on machine learning that are straightforward but dependable. The features retrieved by Byte and ASM are combined with the detection functions of k-NN, RF, and LR in the suggested techniques. The proposed schemes' simulation analysis yielded performance metrics such as recall, specificity, sensitivity, and F1-score. The comparison of the different outcomes shows unequivocally that the malware detection strategy based on RF works better than the model based on k-NN and LR. The suggested machine learning models' detection efficiency is either the same as or similar to that of deep learning-based techniques. Effective feature selection and extraction are crucial for enhancing the performance of ML-based malware detection models. R. Patil, W. Deng et al. [7] examined multiple algorithms, including LightGBM, XGBoost, Random Forest, and others, and found that LightGBM demonstrated superior performance with an impressive accuracy rate of 96.5% during cross-validation. M. Al-Janabi et al. [8] presented a comparative analysis of various ML-based malware detection techniques, highlighting the effectiveness of the J48 algorithm combined with Hybrid analysis for Windows-based malware detection and the Decision Tree algorithm with Dynamic analysis for Android malware detection. Researchers have also explored the use of opcode frequency as a feature vector for malware detection. Rathore, Hemant et al. [9] found that when opcode frequency is employed as a feature, the Random Forest algorithm performs better than Deep Neural Networks, suggesting that simpler ML models and feature reduction techniques can be more effective in detecting and analyzing malware.

Recognizing the inherent limitations of individual ML models, researchers have explored ensemble learning techniques to enhance malware detection capabilities. S. A. Roseline et al. [12] proposed a visual method that represents malware as two-dimensional graphics and a layered ensemble approach, outperforming deep learning techniques. B. Kapekar et al. [15] advocate for future research on hybrid techniques, combining the Random Forest algorithm and blacklist methods, to further improve phishing website detection accuracy.

A comprehensive review by A. F. Agarap et al. [13] examined the use of deep learning techniques for malware and intrusion detection, highlighting the widespread adoption of Convolutional Neural Networks (CNN), Autoencoders (AE), Recurrent Neural Networks (RNN), and Long Short-Term Memory (LSTM) models, achieving accuracy rates up to 99.9%. P. Priyadarshan et al. [11] explored practical malware detection algorithms based on machine learning (k-NN, RF, LR), with the Random Forest algorithm exhibiting better performance than k-NN and LR. *R. A. Desai et al [14]* proposes the development of a Google A Chrome addon that serves as a go-between for unsuspecting consumers and potentially dangerous websites. The authors emphasize the need for such a tool as traditional methods of maintaining a list of malicious websites are not exhaustive and cannot keep up with the continuous development of new attack techniques. The paper provides a comprehensive literature survey, presenting related work on detecting malicious web pages, including various machine learning approaches. The findings of the paper demonstrate the effectiveness of the proposed approach in identifying malicious websites and protecting users from potential harm. Using machine learning, particularly the Random Forest algorithm, provides accurate results in classifying websites, enabling safe browsing even for users who are unaware of potential threats.

Key Motivations include growing difficulties with URL-based Attacks which is due to the widespread nature of URL-based threats, which include a variety of cyber-attacks, detection and mitigation strategies must become more sophisticated and complex. Limitations of traditional approaches makes it difficult for traditional cybersecurity measures, especially those that depend on blacklists, to stay up to date with the constantly changing strategies used by cyber adversaries. Through the investigation of novel strategies, this study seeks to circumvent these constraints. The research is driven by the idea that adding machine learning methods can greatly improve cybersecurity systems' capabilities. This entails making use of machine learning's capacity to examine big datasets, spot trends, and adjust to new dangers. The research adopts ensemble learning as a tactic to integrate the advantages of several models, realizing that no one machine learning model can fully represent the intricacy of the digital threat landscape. The research's main objective is to aid in the creation of a stronger defense system. The goal of the project is to increase the accuracy and dependability of threat identification in real-world scenarios by adopting ensemble learning and overcoming the inherent limitations of individual models.

Findings of existing papers:

TABLE 1. FINDINGS OF EXISTING PAPERS

Paper Title	Algorithm/Model	Accuracy
Malware Analysis and Detection Using Machine Learning Algorithms	Decision Tree(DT)	99%
Malware Detection: A Framework for Reverse Engineered Android Applications Through Machine Learning Algorithms	Ensemble(AdaBoost & SVM)	96.24%
Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm	Proposed Level Ensemble approach	98.65%(Malig)
Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm	Proposed Level Ensemble approach	97.2%(Big2015)
Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm	Proposed Level Ensemble approach	97.43%(MaleVis)
Efficient Android Malware Prediction Using Ensemble Machine Learning Algorithms	LightBGM	96.5%

III. OBJECTIVES

To develop and put into practice a cutting-edge system specifically suited for URL detection and classification, Advanced Detection & Classification is used to classify them into phishing, benign, defacement, and malware. We seek to investigate and utilize the power of ensemble approaches, with a specific focus on Majority Voting, including both Soft Voting and Weighted Voting, realizing that no single algorithm provides comprehensive solutions. While individual models have advantages, their collective intelligence is where their true power rests. By utilizing each algorithm's unique capabilities through ensemble learning, we hope to minimize any potential biases and inaccuracies inherent in any single model. To carefully assess and contrast various machine learning models and ensemble strategies, pinpointing their unique benefits, drawbacks, and applicability to our problem domain. This study addresses the difficulties related to dynamic and polymorphic malware detection, suggests a unique ensemble approach, and synthesizes and expands upon the results of other investigations. The goal is to enhance threat detection mechanisms through the creation of an advanced URL-specific malware detection system and the investigation of ensemble approaches.

IV. SYSTEM OVERVIEW

The main goal of proposed system is to sort URLs into four different danger categories malware, phishing, defacement, and benign. The necessity of creating a sophisticated detection method is critical given the increasing number of harmful URLs in the modern digital environment. The proposed system is the result of the integration of numerous cutting-edge machine learning algorithms, further strengthened by ensemble techniques, rather than being the only output of a single methodology. The design of the system, its parts, and the dynamics of their interactions are all covered in detail in this section.

1. Architecture & Framework

The first process includes data gathering & preprocessing. A sizable, well-balanced dataset derived from reputable cybersecurity organizations serves as the cornerstone of our system. Preprocessing of this dataset is careful to ensure that all anomalies, missing values, and redundancy are eliminated. The processed dataset is used to produce the training and testing subsets, which allow the model to be trained and tested.

The dataset includes instances of different types or classes, such as:

Malware: Malicious software with the intention of damaging or abusing computer systems. This group includes ransomware, worms, trojan horses, viruses, and other harmful software.

Phishing: Attempts to pose as a reliable organization in order to fool people into disclosing private information, including credit card numbers or passwords. Phishing often involves deceptive emails, websites, or other communication channels.

Benign: Non-malicious or harmless activities. In the context of cybersecurity, this class represents normal, safe, and expected behavior.

Defacement: Unauthorized alteration of a website's appearance, typically to convey a message or vandalize the site. This can include changes to the visual elements or content of a webpage.

b. Individual Model Training: A number of machine learning algorithms, including Random Forest, SVM, XGBoost, Naive Bayes, Logistic Regression, LightGBM, and Decision Tree among others,
are independently trained on the dataset. Every one of these algorithms learns and internalizes the patterns found in the data with the goal of predicting the category of an upcoming URL.

c. Individual Model Evaluation: The testing subset of the dataset is used to construct each model's post-training performance metrics, including recall, accuracy, precision, and F-score. The results of this examination paint a clear picture of the advantages and shortcomings of each method.
d. Integration of ensemble models: We use ensemble approaches because we understand that no one model can fully represent the issue space. The various models are synergized, their combined intelligence harnessed, and forecasts are produced by methods including hard voting, soft voting, and weighted voting.

2. Operational Dynamics

The operational dynamics of a malware detection system involve a series of intricate steps aimed at accurately categorizing URLs as malicious, phishing, defacement, or benign. The process begins with the input mechanism, where a URL is fed into the program, and the algorithm extracts the necessary information, aligning it with the format used during model training Following the input mechanism, the system utilizes parallel processing, enabling each trained model to independently examine the URL and predict its category based on its unique algorithmic biases. This step ensures that multiple perspectives are considered, increasing the overall accuracy of the system. Once the individual models have made their predictions, the ensemble layer combines these predictions using a predefined ensemble voting method, such as hard, soft, or weighted voting.

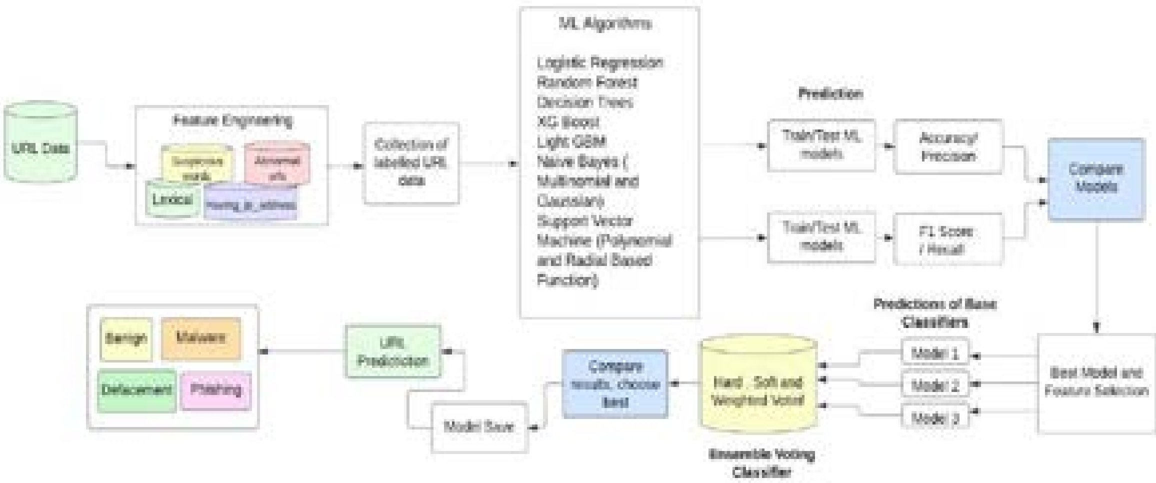


Fig. 1 Architectural flow of the proposed system

This layer helps to consolidate the predictions from all models, making the final decision on the URL's category. The system then outputs the categorization of the URL, along with a confidence score that reflects the models' overall assurance in the prediction, depending on the ensemble technique employed. This score can be useful in determining the level of risk associated with the URL, allowing users to make informed decisions about their online activities. The entire process is depicted clearly in Figure 1.

V. METHODOLOGY

2.1. Requirement Analysis & Data Collection

The first stage in our study was understanding the criteria for a reliable malware detection and classification system. Given the wide variety of malicious URLs, it was crucial to collect a diversified and thorough dataset that appropriately reflected the different threat categories. In order to guarantee a fair representation of different URL categories, we gathered an extensive dataset from reputable cybersecurity groups for this investigation. Ten thousand URLs in all, 2,500 samples from each of the following categories (malware, phishing, defacement, and benign) make up the collection.

We retrieved 215 elements, comprising different properties like the length of the URL, the age and popularity of the domain, the use of suspicious keywords, the existence of redirects, and other significant characteristics, for every URL. We preprocessed the dataset by fixing duplicate URLs, guaranteeing data fidelity, and normalizing and scaling the data before training the machine learning models. This procedure ensures consistency and quality of data, which makes it appropriate for machine learning algorithms.

We employed an 80-20 split for the data splitting approach, meaning that 80% of the dataset was used for training and 20% for testing. To further guarantee the models' robustness and generalizability, we also used 5-fold cross-validation.

Due to the diversity and good representation of our dataset, we did not use any data augmentation techniques in this investigation. To improve the model's performance even more, future research could investigate data augmentation methods like generative modeling, introducing noise, or changing URL properties. Having learned this, we started the process of gathering data. The extensive database that was assembled contained URLs from the following categories: malware, benign, defacement, and phishing. Since the URLs were gathered from several trustworthy sources, their authenticity and diversity were guaranteed. Reputable cybersecurity firms, open-source threat intelligence platforms, and scholarly databases served as the information's primary sources.

To train, validate, and test our machine-learning models, this database was created to be a mirror of real-world circumstances. It offers a wide range of samples. The dataset's diversity was meant to test the models and make sure they were good at spotting the finer details and patterns that were specific to the various danger categories. Basically, this sizable database served as the foundation for our research and also helped to make sure that the conclusions and solutions we came up with were applicable, timely, and efficient in tackling today's cybersecurity concerns.

2.2. Data Preprocessing

The quality of the data has a direct impact on how well the model performs in any machine-learning project. As a result, we meticulously preprocessed our dataset to guarantee its top quality before starting model training. The problem of duplicate URLs was first addressed. Repetitive entries might skew the training process and cause overfitting. The database's duplicate URLs were so quickly found and eliminated.

In order to assure data quality and consistency and prepare the dataset for machine learning algorithms, we scale, normalize, and ensure data dependability at this phase. We employ the following methodical approach to data cleaning Removal of Duplicate URLs: In order to avoid bias during the model training process, we eliminate duplicate URLs. This guarantees the uniqueness of every URL and aids in the model's learning process.

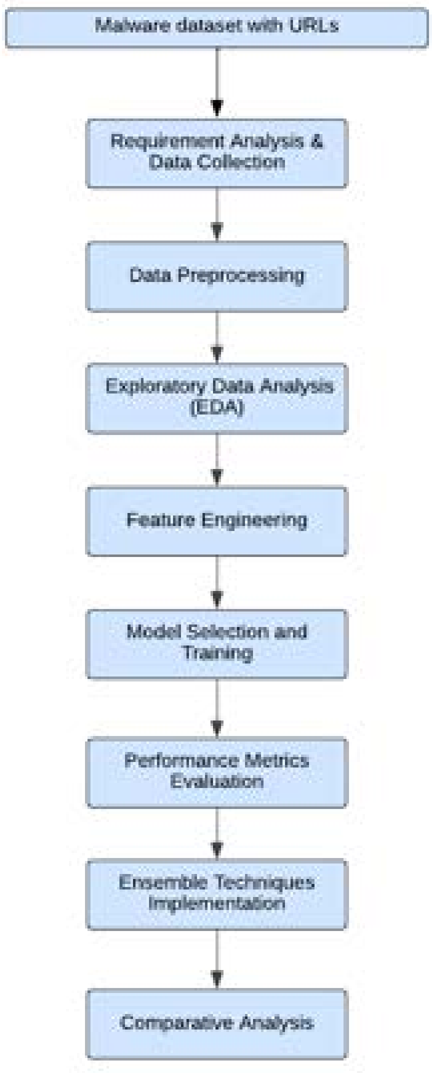


Fig. 2 Flowchart outlining the general steps involved in machine learning-based virus prediction

Data Reliability: We make sure that the data is reliable by eliminating inaccurate or partial entries. In order to prevent any detrimental effects on the model's learning process, this step is essential.
Normalization and Scaling: To guarantee that every data point has a uniform format and scale, we normalize and scale the collection. This stage is crucial since various machine learning algorithms require particular data.

The reliability of the data was the subject of our next inquiry. If left unchecked, incomplete or corrupt data entries might affect the model's learning process, which in turn affects how accurately it predicts outcomes. Only complete and accurate data samples were kept by meticulously filtering out such entries. Finally, the dataset underwent normalization and scaling operations due to the variety of machine learning methods utilized in this work, each with its own specific data requirements. This stage made sure that every data point had a consistent format and scale, which allowed the various algorithms to process and learn from them efficiently. In conclusion, by doing these preprocessing processes, we made sure that our dataset was both free of irregularities and consistency issues and that it was also suited to meet the specific needs of the machine learning algorithms we had chosen.

2.3. Exploratory Data Analysis (EDA)

An essential phase in the research process is exploratory data analysis (EDA), which offers insightful knowledge about the underlying patterns and structure of the data. In order to comprehend the dataset's distribution, spot outliers, and choose pertinent characteristics for malware identification and classification, we used EDA in our investigation.

Data Visualization: We employed a variety of plots, such as pie charts, bar plots, and histograms, to display the data distribution. The distribution of several URL types, including malware, phishing, benign, and defacement, was clearly understood thanks to these graphics.

Correlation Analysis: We computed correlation coefficients and plotted correlation matrices to visualize the links between various features. Through the use of this research, we were able to pinpoint highly correlated characteristics that should be eliminated in order to reduce redundancy and enhance model performance.

Feature Selection: We chose pertinent features for malware identification and classification based on the EDA results. The length of the URL, the use of special characters, the inclusion of particular keywords, and domain information were some of these criteria. Additionally, we developed specialized characteristics for each category of URL, such as phishing URLs that resemble genuine domain names and malware URLs that are identified by looking for scripts or files connected to malware.

2.4. Feature Engineering

Opcode frequency served as a feature vector that we employed to classify malware. Using this method, the frequency of opcode basic machine language instructions is extracted from the malware samples. Through the use of opcode frequency to represent the malware samples, a feature vector is included that encapsulates the distinct attributes of every sample. For the purpose of classification, machine learning algorithms get this feature vector as input.

Moreover, we discovered that the Random Forest approach performed better in malware classification than Deep Neural Networks when opcode frequency was used as a feature. Based on opcode frequency, Random Forest which is renowned for its capacity to manage high-dimensional data and intricate relationships performed remarkably well in the accurate classification of malware samples. This highlights the effectiveness of utilizing opcode frequency as a feature vector in conjunction with the Random Forest algorithm for robust malware classification.

Any machine learning pipeline would not be complete without feature engineering, which turns unstructured data into useful inputs that algorithms can handle quickly. In order to fully understand each URL and determine its potential intent for us, it was essential to extract the appropriate elements for our URL categorization task.

Fig.3 Overall description of URL

URL length: The length of a URL in and of itself can be a sign. Malicious URLs frequently contain malicious sites, routes, or query parameters, making them lengthier. As a result, the URL's length was calculated and employed as a distinctive attribute. Presence of Specific Keywords: Malicious URLs frequently include specific keywords that hint at their intentions, such as "secure-login" and "account-verify," among others. Such terms were searched for in the URL using a feature extraction algorithm, which then noted their presence. Domain Information: A URL's domain portion contains a wealth of information regarding its reliability. It is possible to determine whether a URL is coming from a reliable source or one that could be dangerous by extracting the main domain, any sub-domains, and even the domain registration information. Use of Special Characters: URL anomalies can raise a red warning if they use too many special characters. The "@" symbol or several consecutive hyphens, for instance, may allude to misleading user behavior. The number of special characters and their order were therefore retrieved as characteristics. Tailored Features for Specific URL Categories: Custom features were created in recognition of the distinctive traits that each URL category possesses. For instance, the resemblance of phishing URLs to real domain names was computed, whereas malware URLs were examined for the existence of scripts or files that are known to be associated with malware.

In summary, we turned unstructured URLs into a richly informational format through deliberate feature engineering. These attributes set the framework for later modeling by capturing the various aspects of the URLs, enabling our algorithms to effectively learn and predict the nature of each URL.

2.5. Model Selection and Training

Model Selection Process: Early on, we realized that depending on a single model would be too reductionist of an approach given the complexity of the digital environment and the multiple nature of URL threats. As a result, we tended to favor a number of algorithms that are well-known for being skilled in solving classification-related problems. We favored algorithms like Random Forest, SVM, XGBoost, Decision Tree, Logistic Regression, LightBGM, and Naive Bayes after consulting recent academic literature, industry best practices, and previous practical investigations. These models each have a successful track record in pattern recognition and provide distinctive viewpoints, therefore encompassing a wide range of analytical options.

In this study phase, we investigated deep learning, feature engineering, and transfer learning as advanced machine learning techniques to improve the system's malware detection and classification capabilities. In order to extract subtler and meaningful properties from the URLs and improve the models' ability to distinguish minute changes between different danger categories, feature engineering was further developed. This procedure attempted to improve the system's capacity for prediction and flexibility in response to changing cyber threats.

Robust performance assessments were carried out to guarantee the system's robustness and dependability across various datasets and attack scenarios. The efficacy of the system was evaluated using a variety of datasets, including as the Mailing, BIG 2015, and MaleVis malware datasets, in order to determine its accuracy and detection rates in a range of scenarios. The system's performance was rigorously evaluated under simulated attack scenarios to gauge its resilience and effectiveness in identifying sophisticated and novel malware threats. By subjecting the system to diverse attack simulations, we aimed to validate its ability to detect and classify malicious URLs accurately in real-world cybersecurity environments.

Once our algorithmic arsenal was chosen, the focus shifted toward model training. Training is the process by which these algorithms 'learn' from data. Each algorithm was independently fed the training subset of our dataset. As these models processed the data, they began adjusting their internal parameters to best map the input features to the desired output - the URL categories in this instance. This iterative learning process ensures that the model can discern and remember patterns and nuances inherent in URLs that determine their category. Validation and Fine-tuning, In parallel with the training cycle, we set aside a certain subset of our dataset exclusively for validation purposes. A litmus test is used to evaluate the model's performance using this dataset, which was unaltered during the training process. By assessing the models on this validation set, we may adjust their parameters and make sure they generalize properly to new, unknown data rather than simply repeating back learned data (a problem known as overfitting).

2.6. Performance Metrics Evaluation

After training, it became critical to comprehend the strengths and dangers of each model.
Measurement of Accuracy: One of the most important measures, accuracy, provided us with a simple way to assess how frequently the model made accurate predictions. It offered a concise summary of the model's overall performance and was expressed as a percentage. Digging a little further, we sought to learn more about performance than just the 'overall' result. With the help of Precision, we were able to determine how many URLs the model had classified as, say, "malicious" actually were malicious. Contrarily, recall provided information on the proportion of real dangerous URLs that the model properly recognized.

We used the F-score since we knew that recall and precision were frequently in conflict with one another (increasing one might make the other worse). This statistic seamlessly combines precision and recall to give a comprehensive picture of the model's performance. The confusion matrix proved to be a very useful tool in addition to these measures. It allowed us to identify particular areas that could require greater attention by giving us a detailed, categorical perspective of where models were making mistakes.

2.7. Ensemble Techniques Implementation

Embracing the Power of Collaboration: In light of the fact that no single model is perfect and that each has advantages and disadvantages, we explored ensemble approaches. The foundation of these methods is a straightforward idea: a judgment that is made collectively from various models is frequently more reliable and precise. Majority Voting Mechanism: The Majority Voting mechanism was the simplest ensemble method we used. Each model's forecast acts as a "vote" in this approach. Which category receives the majority of these votes determines the final judgment or categorization.

Soft Voting Strategy: Moving one step further, the soft voting strategy does not rely just on the hard labels provided by each model. As a substitute, it takes into account the likelihood or confidence scores from models, producing a more complex and thoughtful final prediction.

Weighted Voting Approach: We implemented a Weighted Voting mechanism after realizing that not all models perform equally well. Models in this case were given weights based on their individual performances (often their accuracy), ensuring that models with superior performances had a greater influence on the ensemble's ultimate forecast.

2.8. Comparative Analysis

We started a comparison analysis using both the individual models and the ensemble methodologies that we had available. Comparing the performance measures of individual models to those of our ensemble techniques allowed us to understand the additional value that ensemble methods provided. We were interested in determining which ensemble method Majority, Soft, or Weighted voting worked best in our situation. By comparing their outcomes to those of our dataset and study goals, we were able to determine which ensemble methodology was the most successful in categorizing URLs.

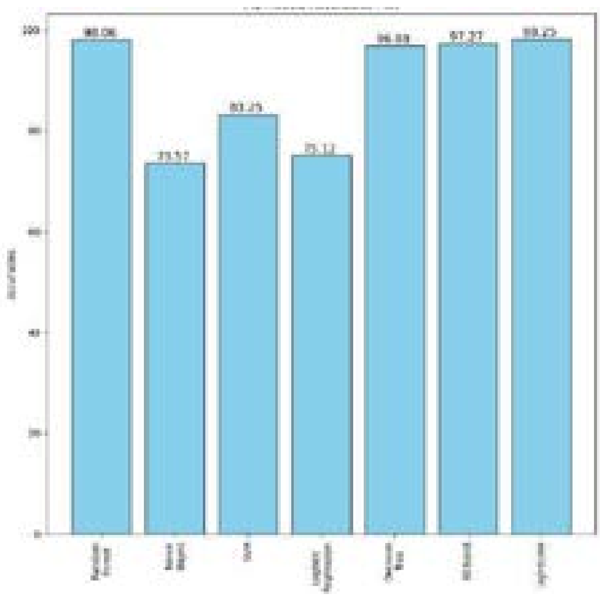


Fig.4 Machine Learning models accuracies plot

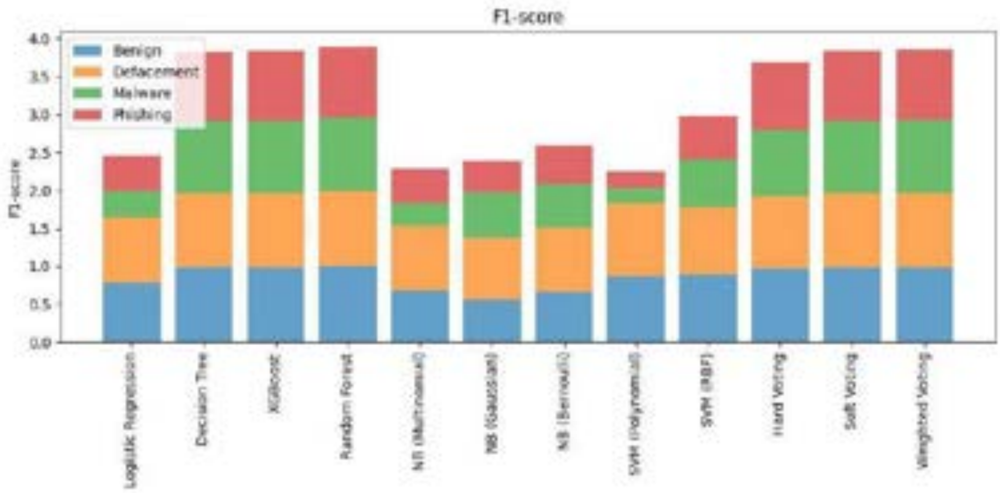


Fig.5 F1 Score

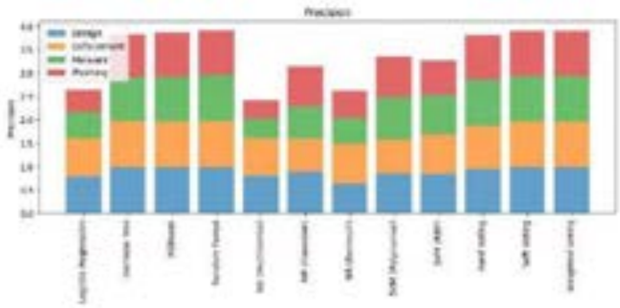


Fig.6 Precision plot

Figure 5, 6 and 7 plots depict the performance metrics of machine learning models across different classes of URLs. Precision measures the proportion of true positive predictions among all positive predictions made by the model. Recall, on the other hand, quantifies the proportion of true positive predictions among all actual positive instances in the dataset. The F1-score is the harmonic mean of precision and recall, providing a balanced assessment of a model's performance.

In these plots, each class represents a specific category of URLs, and the metrics are calculated individually for each class. Higher values of precision, recall, and F1-score indicate better performance of the model in accurately classifying URLs belonging to the respective class. The variation in performance across different classes suggests that certain machine learning models may excel in distinguishing certain types of URLs while performing relatively poorly on others. This variation could be attributed to the inherent complexity and diversity within different classes of URLs, leading to differences in the predictive capabilities of the models.

The performance of various machine learning models for URL classification was evaluated using precision, recall, and F1-score metrics across different classes. Overall, the benign class (class 0) consistently exhibited the highest precision, recall, and F1-score values across all models, indicating that the models could accurately identify benign URLs. Similarly, the defacement class (class 1) demonstrated reasonably good performance, with high precision, recall, and F1-score for many models. However, the phishing class (class 3) showed mixed performance, with some models performing better than others in accurately identifying phishing URLs. Notably, the malware class (class 2) appeared to be the most challenging for the models, with lower precision, recall, and F1-score values compared to other classes.

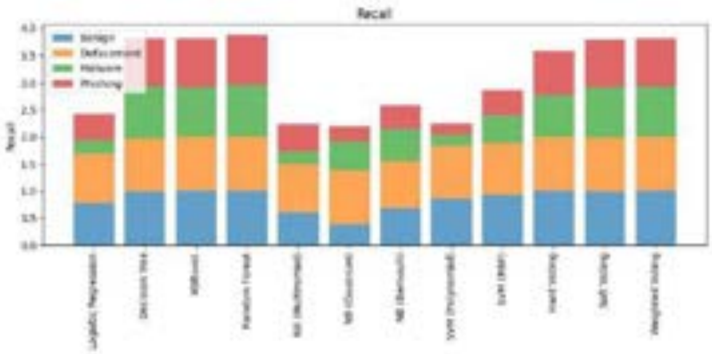


Fig.7 Machine Learning models Recall plot

TP - True Positive (Correctly Predicted Positive)

FP - False Positive (Incorrectly Predicted Positive)

FN - False Negative (Incorrectly Predicted Negative)

TN - True Negative (Correctly Predicted Negative Values)

Accuracy(α) The percentage of correct predictions for the test data.

$$\alpha = \frac{TP + TN}{TP + FP + TN + FN}$$

Precision(ρ) Percentage of positive instances out of the total predicted positive instances.

$$\rho = \frac{TP}{TP+FP}$$

Recall(r) Percentage of positive instances out of the total actual positive instance

$$r = \frac{TP}{TP+FN}$$

F1 score (η) is the Harmonic mean of precision(ρ)and recall(r)

$$\eta = \frac{2 \times \rho \times r}{\rho + r}$$

Logistic Regression

Accuracy: 0.7512

Classification Report	Precision	Recall	F1-score	Support
0	0.79	0.78	0.78	7076
1	0.82	0.91	0.86	19291
2	0.56	0.25	0.35	3910
3	0.47	0.48	0.47	4905

Accuracy			0.75	35182
macro avg	0.66	0.60	0.62	35182
weighted avg	0.74	0.75	0.74	35182

These observations have several implications. Firstly, the differences in performance across classes suggest that machine learning models may find it easier to distinguish benign and defacement URLs from malware and phishing URLs. However, the lower performance on the malware class indicates that more work may be needed to improve the models' ability to accurately detect and classify malware-related URLs. Additionally, ensemble techniques, such as Weighted Voting and Soft Voting, provided better overall performance compared to individual models, highlighting the benefits of combining multiple models to enhance URL classification.

Examining the findings of each model in detail, the Decision Tree Classifier exhibited high accuracy (0.9688) with precision, recall, and F1-score values consistently above 0.90 for all classes. The XGBoost Classifier also performed well with an accuracy of 0.9727, showing high precision and recall values across all classes. The Random Forest Classifier achieved the highest accuracy (0.9806) among all models, with precision, recall, and F1-score values above 0.95 for all classes. However, models such as Naive Bayes - Multinomial and Naive Bayes - Gaussian showed lower accuracy and relatively lower precision, recall, and F1-score values compared to other models. Support Vector Machine - Radial Basis Function (RBF) demonstrated good performance with an accuracy of 0.8325, with precision, recall, and F1-score values consistently above 0.70 for all classes. Overall, these observations provide insights into the strengths and weaknesses of different machine learning models for URL classification tasks.

No single model is the best across all classes and metrics. For example, some models may have higher precision but lower recall, or vice versa. Ensemble methods like Random Forest, XGBoost, and the various voting classifiers generally perform well across all metrics, indicating that they are robust for this particular classification task.

Naive Bayes (NB) classifiers, especially the Gaussian and Bernoulli, seem to have lower performance compared to other models. Support Vector Machine (SVM) with RBF kernel appears to have a good balance between precision, recall, and F1-score, suggesting it is a strong model for this task.

Decision Tree Classifier

Accuracy: 0.9688

Classification Report	Precision	Recall	F1-score	Support
0	0.99	0.99	0.99	7076
1	0.98	0.98	0.98	19291
2	0.94	0.96	0.95	3910
3	0.92	0.89	0.91	4905

Accuracy			0.97	35182
macro avg	0.96	0.96	0.96	35182
weighted avg	0.97	0.97	0.97	35182

XGBoost Classifier

Accuracy: 0.9727

Classification Report	Precision	Recall	F1-score	Support
Class 0	0.99	1.00	0.99	7076
Class 1	0.97	0.99	0.98	19291
Class 2	0.97	0.93	0.95	3910
Class 3	0.94	0.90	0.92	4905

Accuracy			0.97	35182
macro avg	0.97	0.95	0.96	35182
weighted avg	0.97	0.97	0.97	35182

With a 97.27% accuracy, XGBoost Classifier trailed closely behind. The effectiveness of boosting strategies in classification problems was underscored by this model's performance.

Random Forest Classifier

Accuracy: 0.9806

Classification Report	Precision	Recall	F1-score	Support
0	0.99	1.00	1.00	7076
1	0.98	0.99	0.99	19291
2	0.99	0.96	0.97	3910
3	0.95	0.93	0.94	4905

Accuracy			0.98	35182
macro avg	0.98	0.97	0.97	35182
weighted avg	0.98	0.98	0.98	35182

Naive Bayes - Multinomial

Accuracy: 0.7095

Classification Report	Precision	Recall	F1-score	Support
0	0.80	0.60	0.68	7076
1	0.81	0.90	0.85	19291
2	0.41	0.23	0.30	3910
3	0.40	0.51	0.45	4905

Accuracy			0.71	35182
macro avg	0.60	0.56	0.57	35182
weighted avg	0.71	0.71	0.70	35182

Naives Bayes - Gaussian

Accuracy: 0.7231

Classification Report	Precision	Recall	F1-score	Support
0	0.89	0.39	0.55	7076
1	0.70	0.99	0.82	19291
2	0.69	0.54	0.61	3910
3	0.86	0.28	0.42	4905

Accuracy			0.72	35182
macro avg	0.79	0.55	0.60	35182
weighted avg	0.76	0.72	0.69	35182

Naives Bayes - Bernoulli

Accuracy: 0.7357

Classification Report	Precision	Recall	F1-score	Support
0	0.64	0.69	0.66	7076
1	0.84	0.86	0.85	19291
2	0.55	0.59	0.57	3910
3	0.60	0.44	0.51	4905
Accuracy			0.74	35182
macro avg	0.66	0.65	0.65	35182
weighted avg	0.73	0.74	0.73	35182

Support Vector Machine - Polynomial

Accuracy: 0.7605

Classification Report	Precision	Recall	F1-score	Support
0	0.85	0.86	0.86	7076
1	0.72	0.97	0.97	19291
2	0.92	0.20	0.20	3910
3	0.87	0.22	0.22	4905
Accuracy			0.76	35182
macro avg	0.84	0.57	0.59	35182
weighted avg	0.79	0.76	0.71	35182

Support Vector Machine - Radial Basis Function (RBF)

Accuracy: 0.8325

Classification Report	Precision	Recall	F1-score	Support
0	0.85	0.93	0.89	7076
1	0.84	0.96	0.89	19291
2	0.85	0.50	0.63	3910
3	0.75	0.47	0.58	4905
Accuracy			0.83	35182
macro avg	0.82	0.71	0.75	35182
weighted avg	0.83	0.83	0.82	35182

Findings of Dynamic Voting Algorithms of our work:

In the machine learning approach known as ensemble learning, several models are trained to address a single problem, and then their predictions are aggregated to enhance overall performance. Hard voting and soft voting are two popular methods used in ensemble learning to combine predictions.

In hard voting, a majority vote determines the final prediction made by each model in the ensemble. The final predicted class is determined by tallying the votes cast in favor of that class. Hard voting is effective when individual models in the ensemble are diverse and make errors independently. It is particularly useful in situations where each model has a different perspective on the problem.

In soft voting, each model in the ensemble predicts the probability or confidence of each class for a given input. The class with the highest average probability across all models is the final prediction. Soft voting is effective when the models in the ensemble provide probability estimates. It takes into account the confidence level of each model, which can be valuable in situations where some models are more certain than others.

Weighted voting is a technique where the predictions of individual models are given different levels of importance or weight when making the final prediction. This approach is particularly useful when certain models in the ensemble are deemed to be more reliable or have higher confidence in their predictions than others. Weighted voting allows the ensemble to assign varying degrees of influence to each model, leading to a more nuanced and flexible decision-making process. Weighted voting may reduce the interpretability of the ensemble. Achieving a balance between enhanced performance and the capacity to articulate the decision-making process is crucial.

Key Considerations in the proposed method

Diversity of Models: Hard voting benefits from diverse models that make different types of errors, while soft voting benefits from models that provide well-calibrated probability estimates.

Threshold in Soft Voting: In soft voting, the choice of threshold for binary classification is adjusted based on the desired balance between precision and recall.

Computational Complexity: Hard voting is generally simpler computationally, as it only involves counting votes. Soft voting requires computing average probabilities, which may be more resource-intensive.

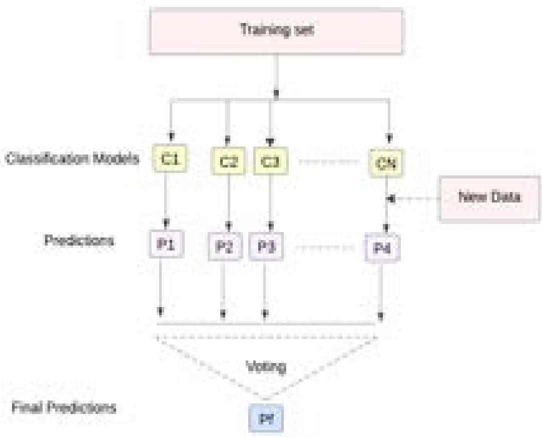


Fig.8 Working of Majority Voting Ensemble Algorithm.

A straightforward ensemble technique known as "hard voting" or "majority voting" bases the final prediction on the choices made by the majority of the classifiers. Assuming we have three classifiers (C1, C2, C3) and their predictions for a sample are [0, 0, 1], the majority vote would classify the sample as "class 0" since it has more votes for that class.

Mathematically, the prediction y is given by:

y =*mode* {C1x, C2x, ...,Cmx}

Weighted Majority Vote:

In weighted majority voting, each classifier is assigned a weight (wj), and the class with the greatest weighted total of votes makes the final prediction.

$$(y) = argmax_i \sum_{j=1}^{m} W_j X_A(C_i(x)) = i)$$

Here, Ais the characteristic function, and A is the set of unique class labels. For instance, if we assign weights {0.2, 0.2, 0.6} to the classifiers, and their predictions are [0, 0, 1], the prediction would be 1 based on the weighted sum.

Soft Voting:

In soft voting, predictions are based on the predicted probabilities (pij) from each classifier. This method is recommended when classifiers are well-calibrated. The class with the highest weighted sum of probability is the final forecast.

$$(y) = argmax_i \sum_{j=1}^{m} W_j \mathrm{p}_{ij}$$

For a binary classification task, if classifiers C1, C2, and C3 predict probabilities [0.9, 0.1], [0.8, 0.2], and [0.4, 0.6] respectively, the soft voting prediction can be computed using uniform weights or custom weights. For instance, with weights {0.1, 0.1, 0.8}, the prediction might differ.

$(y) = argmax_i$ **[**0.10.9+0.10.8+0.80.4,

0.10.1+0.20.1+0.80.6]

This approach allows for a more nuanced decision-making process by considering the confidence levels of each classifier.

Hard Voting

Accuracy: 0.9437

Classification Report	Precision	Recall	F1-score	Support
0	0.95	1.00	0.97	7076
1	0.93	0.99	0.96	19291
2	0.98	0.78	0.87	3910
3	0.96	0.82	0.88	4905

Accuracy			0.94	35182
macro avg	0.96	0.90	0.92	35182
weighted avg	0.95	0.94	0.94	35182

Soft Voting

Accuracy: 0.9731

Classification Report	Precision	Recall	F1-score	Support
0	0.99	1.00	0.99	7076
1	0.97	0.99	0.98	19291

2	0.98	0.93	0.96	3910
3	0.95	0.89	0.92	4905

Accuracy			0.97	35182
macro avg	0.97	0.95	0.96	35182
weighted avg	0.97	0.97	0.97	35182

Weighted Voting

Accuracy: 0.9751

Classification Report	Precision	Recall	F1-score	Support
0	0.99	1.00	0.99	7076
1	0.97	0.99	0.98	19291
2	0.98	0.94	0.96	3910
3	0.96	0.90	0.93	4905

Accuracy			0.98	35182
macro avg	0.97	0.96	0.97	35182
weighted avg	0.97	0.98	0.97	35182

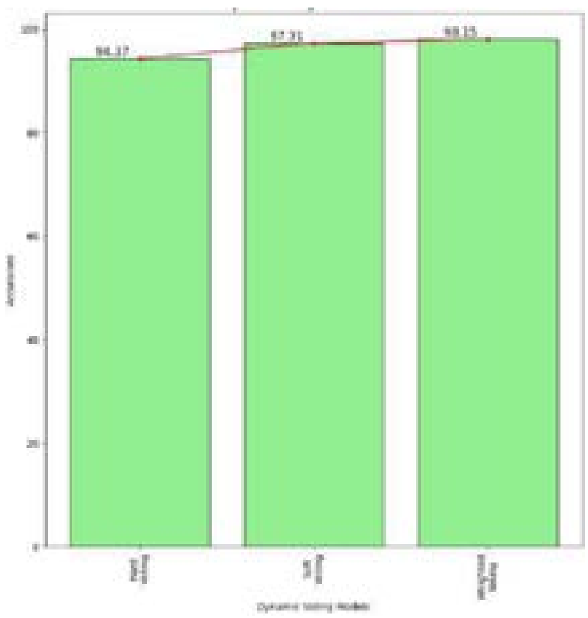


Fig.9 Dynamic Voting models accuracy plot.

VII. Results & Discussion

Across several machine learning algorithms and ensemble approaches, our thorough analysis demonstrated a range of outcomes. Below is an in-depth analysis.

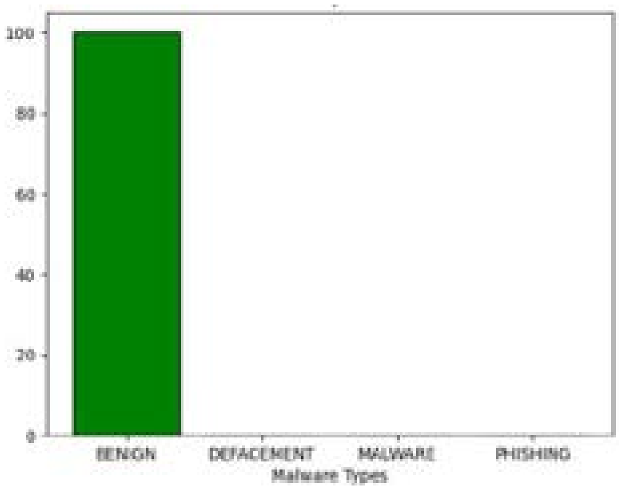


Fig.10 Benign prediction output.

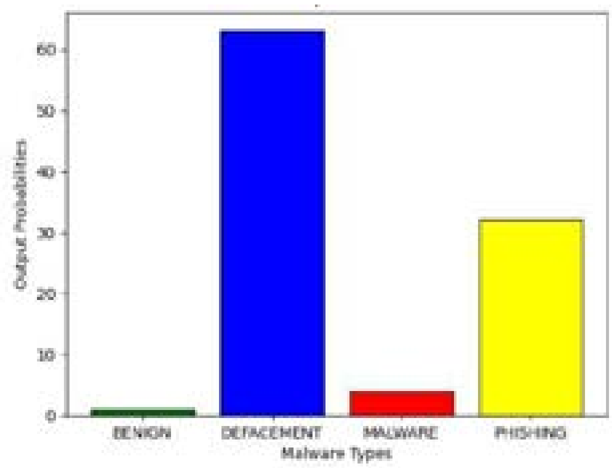


Fig. 11 Defacement prediction output.

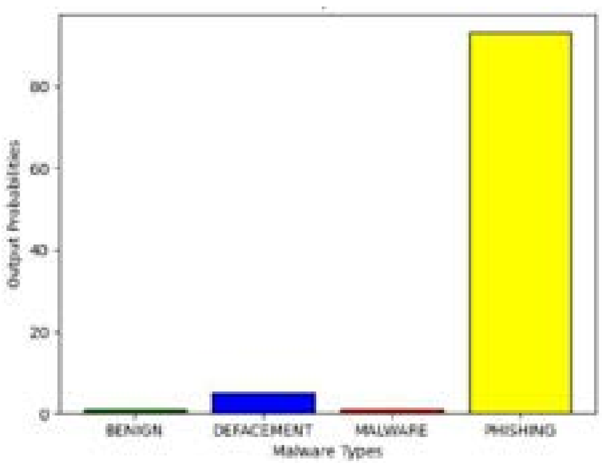


Fig.12 Phishing prediction output.

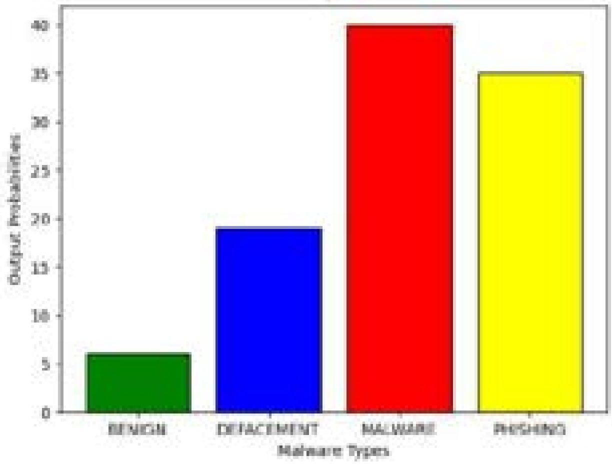


Fig.13 Malware prediction output.

Individual Models:

With an accuracy of 75.12%, the Logistic Regression model showed promise for improvement but was still deemed to be reasonably dependable. It was difficult to categorize URLs between some categories, as the confusion matrix showed. With a noteworthy accuracy of 98.06%, the Random Forest Classifier was the best-performing model. Based on its excellent recall and precision rates, it showed an impressive capacity to differentiate between the various types of URLs. Variants of Naive Bayes produced a range of outcomes. Out of all the models examined, the Multinomial variant had the lowest accuracy (70.95%). The Gaussian variant did better (72.31%), and the Bernoulli variant evened it out at 73.57%. These findings showed how important it is to select the right algorithmic version based on the kind of data. Support Vector Machines (SVM) demonstrated how flexible kernel functions may be. The accuracy of the polynomial kernel was 76.05%, whereas the Radial Basis Function (RBF) performed better, with 83.25% accuracy.

Ensemble Techniques:

By utilizing the combined intelligence of models, Hard Voting yielded an accuracy rate of 94.37%. Despite being straightforward, this majority voting method greatly raised the categorization rate. Soft Voting improved the outcomes even further, yielding a 97.31% accuracy rate. It provided a more sophisticated forecast by taking individual model probability estimations into account. At 97.51% accuracy, Weighted Voting proved to be the most successful ensemble method. Because models were evaluated according to how well they performed, the most trustworthy models were guaranteed to have a significant impact on the outcome.

VIII. Conclusion And Future Work

Threats based on URLs are a broad and dynamic field. In light of this, it is more important than ever to have sophisticated threat detection and classification systems. We took on this challenge with our research, hoping to use machine learning to come up with an innovative solution.

Every model contributed distinct advantages to the discussion. For example, Random Forests contributed unpredictability to reduce overfitting, Decision Trees offered an understandable and transparent model, and SVM examined the data in higher dimensions to determine the ideal separation but it was only via the use of ensemble techniques to integrate these models that their entire power was shown. The ensemble approaches, in particular the Weighted Voting approach, combined the advantages of the models and lessened their individual shortcomings to produce an improved, more reliable system.

The outcomes and conclusions not only demonstrate machine learning's effectiveness in cybersecurity, but they also open the door for more advancements in the field. It's obvious that defenses against cyberattacks must advance in lockstep with the sophistication of the threats. Leveraging a combination of machine learning models and ensemble approaches can prove to be a game-changer in this continuous digital cat and mouse game, offering a safer digital environment for everyone. Exploring deep learning architectures, like neural networks, to identify complex patterns in URL-based threats could be a future development path in threat detection and categorization. Furthermore, the system's ability to adapt to new cyberthreats may be improved by adding changing threat intelligence and real-time behavioral analysis.

References

[1]. B. Urooj, M. A. Shah, C. Maple, M. K. Abbasi and S. Riasat, "Malware Detection: A Framework for Reverse Engineered Android Applications Through Machine Learning Algorithms," in IEEE Access, vol. 10, pp. 89031-89050, 2022, doi: 10.1109/ACCESS.2022.3149053.

[2]. K. Lu, J. Cheng, and A. Yan, “Malware Detection Based on the Feature Selection of a Correlation Information Decision Matrix,” Mathematics, vol. 11, no. 4, p. 961, Feb. 2023, doi: https://doi.org/10.3390/math11040961.

[3]. R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran and S. Venkatraman, "Robust Intelligent Malware Detection Using Deep Learning," in IEEE Access, vol. 7, pp. 46717-46738, 2019, doi: 10.1109/ACCESS.2019.2906934.

[4]. Şahin, D.Ö.; Kural, O.E.; Akleylek, S.; Kılıç, E. A novel permission-based Android malware detection system using feature selection based on linear regression. Neural Comput. Appl. 2023, 35, 4903–4918.

[5]. P. Singh, S. Kaur, S. Sharma, G. Sharma, S. Vashisht and V. Kumar, "Malware Detection Using Machine Learning," 2021 International Conference on Technological Advancements and Innovations (ICTAI), Tashkent, Uzbekistan, 2021, pp. 11-14, doi: 10.1109/ICTAI53825.2021.9673465.

[6]. S. Agarkar and S. Ghosh, "Malware Detection & Classification using Machine Learning," 2020 IEEE International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur Odisha, India, 2020, pp. 1-6, doi: 10.1109/iSSSC50941.2020.9358835.

[7]. R. Patil and W. Deng, "Malware Analysis using Machine Learning and Deep Learning techniques," 2020 SoutheastCon, Raleigh, NC, USA, 2020, pp. 1-7, doi: 10.1109/SoutheastCon44009.2020.9368268.

[8]. M. Al-Janabi and A. M. Altamimi, "A Comparative Analysis of Machine Learning Techniques for Classification and Detection of Malware," 2020 21st International Arab Conference on Information

Technology (ACIT), Giza, Egypt, 2020, pp. 1-9, doi: 10.1109/ACIT50332.2020.9300081.

[9]. Rathore, Hemant, Swati Agarwal, Sanjay K. Sahay, and Mohit Sewak. “Malware Detection Using Machine Learning and Deep Learning.” Big Data Analytics, 2018, 402–11. doi:10.1007/978-3-030-04780-1_28.

[10]. M. Mehndiratta, N. Jain, A. Malhotra, I. Gupta and R. Narula, "Malicious URL: Analysis and Detection using Machine Learning," 2023 10th International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, India, 2023, pp. 1461-1465.

[11]. P. Priyadarshan, P. Sarangi, A. Rath and G. Panda, "Machine Learning Based Improved Malware Detection Schemes," 2021 11th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2021, pp. 925-931, doi: 10.1109/Confluence51648.2021.9377123.

[12]. S. A. Roseline, S. Geetha, S. Kadry and Y. Nam, "Intelligent Vision-Based Malware Detection and Classification Using Deep Random Forest Paradigm," in IEEE Access, vol. 8, pp. 206303-206324, 2020, doi: 10.1109/ACCESS.2020.3036491.

[13]. A. F. Agarap, ‘‘Towards building an intelligent anti-malware system: A deep learning approach using support vector machine (SVM) for malware classification,’’ 2017, arXiv:1801.00318. [Online]. Available: http://arxiv.org/abs/1801.00318

[14]. Harsha A K, Thyagaraja Murthy A, "Machine Learning Techniques for Malware Detection", International Journal of Scientific Research in Science, Engineering and Technology (IJSRSET), Online ISSN : 2394-4099, Print ISSN : 2395-1990, Volume 8 Issue 5, pp. 70-76, September-October 2021. Available at doi :https://doi.org/10.32628/IJSRSET21858

[15]. B. Kapekar, S. Padole, A. Khedikar, S. Shende and A. A. Nikose, Phishing Website Detection Using Data Analysis and Machine Learning, Mar 2022

14. F. Mira, "A Review Paper of Malware Detection Using API Call Sequences," 2019 2nd International Conference on Computer Applications & Information Security (ICCAIS), Riyadh, Saudi Arabia, 2019, pp. 1-6, doi: 10.1109/CAIS.2019.8769564.

[16]. Uke, Shailaja, et al. "Virtual Voice Assistant In Python (Friday)." 2022 IEEE 4th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA). IEEE, 2022.

[17]. EliteDataScience. 2021. How to Handle Imbalanced Classes in Machine Learning. [online],https://elitedatascience.com/imbalanced-classes 2021

[18]. Wen, Long and Yu, Haiyang (2017) “An Android malware detection system based on machine learning” AIP Conference Proceedings 1864 (1): 020136.

[19]. Uke, Shailaja N., and Amol Zade. "Optimal video processing and soft computing algorithms for human hand gesture recognition from real-time video." Multimedia Tools and Applications (2023): 1-23.

[20]. Bai, Hongpeng and Xie, Nannan and Di, Xiaoqiang and Ye, Qing(2020) “AFAMD: A Fast Multifeature Android Malware Detection Frame-work, Design, and Implementation” IEEE Access (8): 194729–194740

[21]. Uke, Nilesh, and Shailaja Uke. "Proximity Approach for Object Detection in Video." International Journal of Control and Automation 13.2 (2020): 868-876.

[22]. Wang, Wei and Gao, Zhenzhen and Zhao, Meichen and Li, Yidong and Liu, Jiqiang and Zhang, Xiangliang (2018)“DroidEnsemble: Detecting Android malicious applications with ensemble of string and structural static features” IEEE Access (6): 31798–31807

[23]. Uke, Shailaja N., and Amol V. Zade. "An enhanced artificial neural network for hand gesture recognition using multi-modal features." Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization 11.6 (2023): 2278-2289.

[24]. Sun, Lichao and Li, Zhiqiang and Yan, Qiben and Srisa-an, Witawas and Pan, Yu (2016) “SigPID: significant permission identification for android malware detection” 2016 11th international conference on malicious and unwanted software (MALWARE) : 1–8

[25]. Uke, Shailaja, and Nilesh Uke. "Efficient Method for Video Sentiment Analysis." International Conference on Computation Intelligence and Network Systems. Cham: Springer Nature Switzerland, 2023.

[26]. Yuan, Zhenlong and Lu, Yongqiang and Wang, Zhaoguo and Xue, Yibo (2014), “Droid-sec: deep learning in android malware detection” Proceedings of the 2014 ACM conference on SIGCOMM:371–372

[27]. Uke, Shailaja, et al. "Chatbot for Medicine at Your Fingertips Using TextMining." 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA). IEEE, 2023.

[28]. Kumaran, Monica and Li, Wenjia (2016), “Lightweight malware detection based on machine learning algorithms and the android manifest file” 2016 IEEE MIT Undergraduate Research Technology Conference (URTC): 1–3.

[29]. Uke, Shailaja, et al. "Brain Tumor Detection Using Convolutional Neural Network." 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA). IEEE, 2023.

[30]. M. Ahmadi, D. Ulyanov, S. Semenov, M. Trofimov, and G. Giacinto,‘‘Novel feature extraction, selection and fusion for effective malware family classification,’’ in Proc. 6th ACM Conf. Data Appl. Secur. Privacy (CODASPY), 2016, pp. 183–194.

[31]. Pawar, Priyanka P., and Shailaja N. Uke. "Vampire attack detection and prevention using DLWASN on wireless adhoc sensor network." International Journal of Research in Computer Science 1.03 (2014).

[32]. Chivate, Atharva Balasaheb, et al. "Comparative Analysis of Multiple ML Models and Real-Time Translation." 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA). IEEE, 2023.

[33]. F. Ahmed, H. Hameed, M. Z. Shafiq, and M. Farooq, ‘‘Using spatiotemporal information in API calls with machine learning algorithms for malware detection,’’ in Proc. 2nd ACM Workshop Secur. Artif. Intell. (AISec), 2009, pp. 55–62.

[34]. Uke, Shailaja, et al."Systems and Methods: Image Encryption Using Steganography and XOR Operations." 2023 IEEE 5th International Conference on Cybernetics, Cognition and Machine Learning Applications (ICCCMLA). IEEE, 2023.

[35]. B. Anderson, D. Quist, J. Neil, C. Storlie, and T. Lane, ‘‘Graph-based malware detection using dynamic analysis,’’ J. Comput. Virol., vol. 7, no. 4,pp. 247–258, Nov. 2011, doi: 10.1007/s11416-011-0152-x.

[36]. Uke, Shailaja, and Ravindra Thool. "UML based modeling for data aggregation in secured wireless sensor network." Procedia Computer Science 78 (2016): 706-713.

Enhancing Sequential Recommendations with Bi-Directional Models: A Comparative Analysis

1st Sarathi
Dept. of CSE
Karunya Institute of Technology and Sciences
Coimbatore, India
sarathirofficial@gmail.com

2nd Sebastian Terence
Dept. of CSE
Karunya Institute of Technology and Sciences
Coimbatore, India
jsebinfo@gmail.com

3rd M. Kranthi
Dept. of CSE
Rise Krishna Sai Prakasam Group of Insutitutions
Ongole, India
kranthim_maddala@yahoo.com

***Abstract*— In the online world, recommendation systems are crucial for enhancing the user experience by suggesting relevant items by utilizing user actions. However, they face significant challenges in adapting to evolving user preferences and handling dynamic content. Traditional recommendation algorithms struggle with these challenges, particularly in recognizing evolving user preferences over time. Moreover, they may not effectively handle large datasets or capture intricate patterns in user behavior. To address these challenges, this study analyses the effectiveness of bi-directional models, focusing specifically on Sequential Recommendations with Bi-LSTM (SRec with Bi-LSTM). Unlike traditional static systems, Bi-LSTM offers dynamic and personalized content suggestions by analyzing user actions bidirectionally. This approach enables the system to capture complex patterns in user-item interactions and adapt to changing user preferences over time. In this study, a comparison is made between the performance of Bi-LSTM with several recent techniques, such as Simple RNN with LSTM, Gated Recurrent Unit (GRU) with LSTM, CUDA Deep Neural Network Long Short-Term Memory (CuDNNLSTM), and Stacked RNN. The analysis demonstrates that Bi-LSTM consistently outperforms these methods in terms of accuracy and adaptability to evolving user preferences. By addressing the challenges faced by traditional systems, the proposed Bi-LSTM model offers a promising advancement in personalized sequential recommendations, thereby improving the overall user experience in the fast-paced and ever-changing landscape of online content.**

Keywords—Sequential Recommendations, Bi-Directional Models, Recommender Systems, Personalized Recommendations, Long Short-Term Memory (LSTM)

I. INTRODUCTION

Sequential recommendation systems are crucial for offering users personalized content over time. Current sequential recommendation models mostly rely on one-way processing, making it difficult to capture evolving user preferences and intricate complex temporal dependencies. Traditional models face challenges in capturing the complex patterns and time-related dependencies in user-item interaction sequences.

Addressing these limitations, this study introduces a new approach using bi-directional models, specifically Bi-LSTM. The choice of Bi-LSTM is motivated by its unique capability to capture both past and future context simultaneously, enabling a deeper understanding of the evolution of users' preferences over time and facilitating more precise recommendations. This is particularly critical in dynamic user scenarios where preferences undergo rapid changes. The research is driven by the growing demand for enhanced precision and context-aware recommendations in such dynamic user environments.

Furthermore, the research offers insights into understanding bidirectional connections in user behaviour, revealing the dynamics of sequential interactions. This research aims to address this gap by exploring how Bi-Directional models can manage bidirectional connections in user-item interaction sequences. The goal is to improve the precision and contextual understanding of sequential recommendations, ultimately enhancing the user experience when consuming content.

Examining the effect of personalized recommendations on consumer behaviour, this paper illustrates the significance of users in improving search efficiency and purchase decisions. A meta-analysis [1] examines the correlation between recommendation systems and customer satisfaction, indicating a positive connection between personalized suggestions and customer satisfaction. This study emphasizes the practical benefits of recommendation systems for enhancing user experiences. Here's how this paper is structured: In Section 2, we delve into the problem statement. Section 3 explores related works, while Section 4 provides a detailed description of SRec with Bi-LSTM. Experimental procedures and results are presented in Section 5, and the paper concludes with Section 6, summarizing the findings.

II. PROBLEM STATEMENT

The abundance of material and services online presents users with numerous options, making it challenging for them to find what they want. The typical recommendation algorithms struggle to adapt to evolving user preferences, particularly over time.

Regular recommendation systems, such as collaborative filtering, perform effectively with static items, but they struggle with dynamic user preferences. Sequential recommendation, which attempts to forecast a user's next actions based on their previous behavior, emerges as an effective solution to this problem. However, it faces difficulties in interpreting patterns over time and explaining recommendations. Regular systems might not notice how user preferences change, leading to not-so-great suggestions over time. Moreover, they struggle with handling large datasets.

Traditional models, including collaborative filtering and matrix factorization, provide suggestions based on unchanging data and struggle with complex patterns. Additionally, they fail to adapt to dynamic user preferences and cannot efficiently handle large datasets.

Another challenge in recommendation systems is the issue of cold start, where new users or items have limited data available, making it difficult to provide accurate recommendations. Furthermore, recommendation systems may suffer from algorithmic biases, where certain users or items are favored over others, leading to unequal exposure and potentially inaccurate recommendations.

To solve these problems, SRec with Bi-LSTM is proposed. This approach aims to enhance personalized recommendations in the rapidly evolving world of online content. However, it's important to note that maintaining the efficiency of recommendation systems as the volume of online content continues to grow exponentially is another significant challenge to consider.

III. RELATED WORKS

Initially, the researchers used the Markov Chain (MC) technique to build a recommendation system. Rendle et al. [2] combined matrix decomposition and the first-order MC to understand users overall preferences and their short-term interests. They also utilized the RNN algorithm to understand users' behaviours. These are good at figuring out patterns in a sequence of events. Hidasi et al. [3] pioneered the use of GRU to generate recommendations based on users' actions within a session. After the initial approaches, numerous modifications were introduced to improve the recommendations. For example, Hidasi et al. [4] incorporated a pair-wise loss function, and Li et al. [5] employed an attention mechanism to focus on specific aspects. Chen et al. [6] utilized a memory network, and Quadrana et al. [7] organized information hierarchically. These innovations collectively contribute to improving recommendation systems.

Bi-directional models have become a powerful solution for understanding the context and connections in user sequences. Sun et al. [8] applied the BERT4Rec model, which considers both past and future items sequentially, to understand the users' preferences. They evaluated this model on the ML-20m dataset regarding Hit Rate (HR) and achieved the following accuracy metrics: HR@1 of 0.3440, HR@5 of 0.6323, HR@10 of 0.7473, Normalized Discounted Cumulative Gain (NDCG) @5 of 0.4967, NDCG@10 of 0.5340, and Mean Reciprocal Rank (MRR) of 0.4785. Here, HR gauges the proportion of correct recommendations among the top recommendations provided to users. NDCG assesses the quality of the ranked recommendations. MRR evaluates the effectiveness of the ranking of the first relevant item in the recommendations.

Kang et al. [9] introduced SASRec, a model that integrates self-attention into a Bi-LSTM architecture to highlight essential features of user sequences. This approach enhances the model's ability to understand item relationships and adapt to changing user preferences. Their evaluation on the ML-1M dataset yielded HR@10 of 0.8245 and NDCG@10 of 0.5905 as the obtained results.

Shao et al. [10] introduced NextBasket, a method that combines Bi-LSTM to suggest items in order with a separate model for predicting entire groups of items that users might prefer together. This fusion enhances the accuracy of predictions for both individual items and entire baskets of items. The method underwent evaluation on three different datasets: TaFeng, Instacart, and Dunnhumby. In the Instacart dataset with k = 5, the recall, precision, F1-Score, and NDCG were found to be 0.2616, 0.3733, 0.2619, and 0.3094, respectively. For k = 10 in the Instacart dataset, the recall, precision, F1-Score, and NDCG were 0.3698, 0.2896, 0.2812, and 0.3805, respectively.

Zhao et al. [11] explored deep Bi-LSTM architectures to gain a deeper understanding of user preferences, comparing models of varying depths. They evaluated these models on the ML-1m dataset, achieving Recall@20 of 0.22634 and MRR@20 of 0.04634. This demonstrates the effectiveness of using deeper models to capture detailed user preferences.

Hu et al. [12] presented BiVRec, a new approach utilizing a bi-directional framework to integrate diverse information types, including text and images, into recommendations. This comprehensive approach enhances the model's understanding of users' preferences and available items, resulting in improved recommendations across varied information environments. Evaluation on the ML-25m and ML-1m datasets yielded Recall@20 of 11.96%, NDCG@20 of 28.31, and HR of 65.54%. These results underscore the effectiveness of BiVRec in enhancing recommendation performance.

Devooght et al. [13] proposed the use of a neural network known as RNN for collaborative filtering, demonstrating its effectiveness in understanding the sequence of preferences users have over time. Essentially, RNN helps in understanding the evolution of user's preferences, leading to improved recommendations based on their interactions.

IV. PROPOSED METHODOLOGY

The proposed methodology encompasses various modules aimed at understanding users' behavior patterns over time and generating personalized recommendations. The study uses the MovieLens 100K dataset [14] to analyze user interactions systematically. The approach focuses on organizing user actions in chronological order to observe the evolution of preferences and decisions.

Fig. 1 shows the process of sequential data with Bi-LSTM, which pays attention to things that have happened before and might happen in the future. This is important for understanding user interactions. Unlike regular LSTMs that only analyze data in one direction, Bi-LSTMs examine data in both directions simultaneously, allowing for a comprehensive understanding of past and potential future events. This allows our model to understand what happened before and what might happen next in the sequence.

In handling sequences of user behaviour, such as movie watching or browsing history, Bi-LSTM plays a crucial role. Fig. 2 shows the structure of the LSTM cell. LSTM is a type of neural network that includes special memory cells capable of retaining information for extended periods, addressing challenges commonly encountered in sequential data analysis.

Each user's interactions with items are grouped into sequences, which is essential for capturing user-item interaction patterns. A vocabulary is created to represent different events in these sequences, facilitating model learning. To ensure data consistency, sequences are standardized to the same length by appending additional information to shorter sequences. Subsequently, sequences are divided into input (X) and target (y) parts, enabling the model to predict subsequent items accurately. Class weights are determined to account for variations in the frequency of positive instances during training, preventing bias towards specific event types.

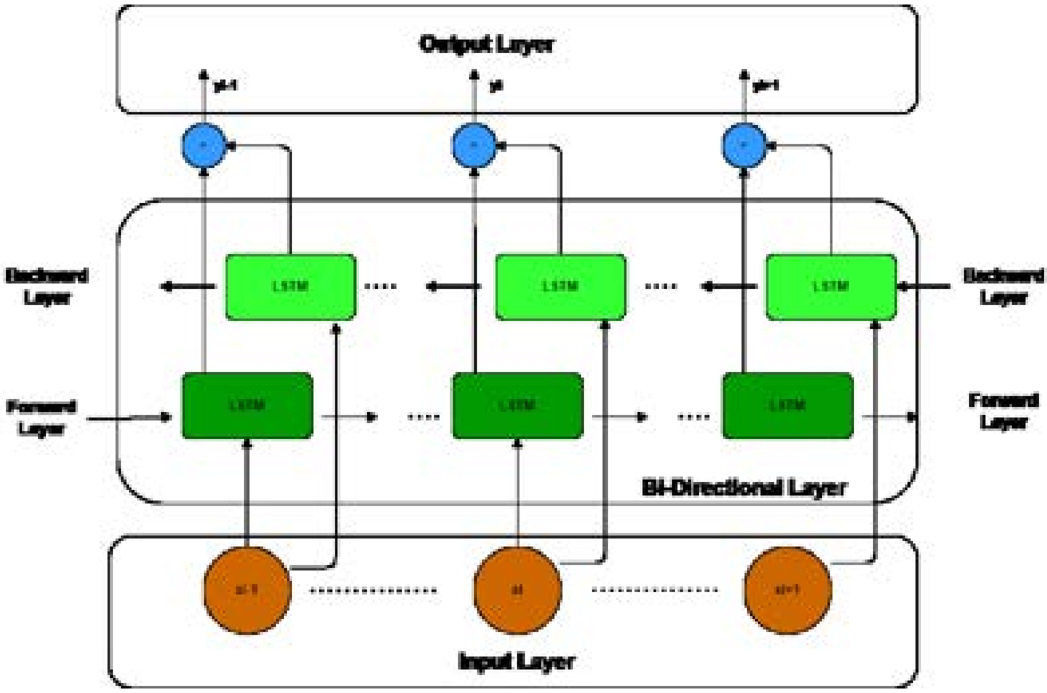


Fig.1 Processing Sequential Data with Bi-LSTM

The model has multiple layers, such as an input layer, an embedding layer, a Bi-LSTM layer, and a dense layer with SoftMax activation, to enhance data understanding. The Adam optimizer and Sparse Categorical Cross-entropy Loss function are used with class weights considered during training. Training is conducted over multiple epochs to evaluate system performance comprehensively.

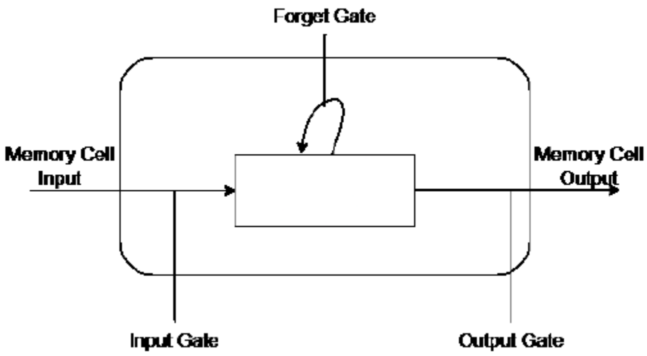


Fig. 2 Structure of the LSTM Cell

During training, the Bi-LSTM model parameters are fine-tuned to predict subsequent items accurately, guided by a tailored loss function for sequential prediction tasks. After training, the model can generate recommendations based on learned patterns, resulting in precise and context-aware suggestions.

To analyze the system performance, SRec with Bi-LSTM is compared with alternative models such as Simple RNN with LSTM, GRU with LSTM, CuDNNLSTM, and Stacked RNN. The evaluation demonstrates the superior performance of the proposed SRec with Bi-LSTM model, attributed to its bidirectional processing capability, enabling effective capture of complex sequential patterns. Consequently, Bi-LSTM emerges as a robust choice for tasks requiring sequential recommendations.

V. EXPERIMENTAL SETUP AND RESULTS

A. Dataset

The model in this research was trained using the Movielens 100k dataset [14], which is a well-known benchmark dataset for collaborative filtering and recommendation systems. This dataset originates from the Movielens project, a movie recommendation service provided by the GroupLens research lab. It contains information about users, movies, ratings, and timestamps.

Specifically, the dataset includes the following fields:

- user_id: A unique code assigned to each user.
- item_id: A unique code is assigned to each movie.
- rating: The rating given by the user to a movie, typically ranging from 1 to 5.
- timestamp: The time when the user provided the rating.

B. Hyperparamteres and Evaluation Metrics

To enhance the effectiveness of the SRec with Bi-LSTM model, various settings were adjusted, and among them, certain parameters yielded the best results.

Fig. 3 showcases the architecture of the SRec with the Bi-LSTM model. This visual representation illustrates the step-by-step flow, starting from the input layer to the dense output layer.

The model begins with the input layer, where raw data, such as sequences of user interactions, enters the system. This layer captures the sequence of user interactions, which is essential input for further processing when dealing with sequences for recommendations.

The embedding layer has two important settings. The first one is vocabulary size, which represents the number of different items that are dealt with. The second setting is the embedding dimensions, determining the size of the space where the items are positioned.

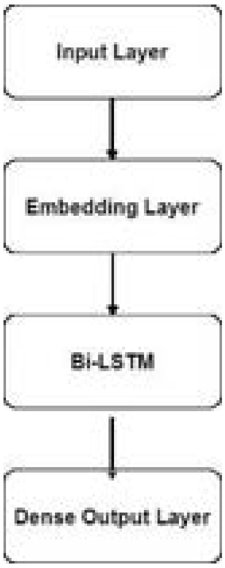


Fig. 3 SRec Architecture with Bi-LSTM

In the Bi-Directional LSTM Layer, hidden units have been set to 100. They examine both sequentially to understand the connections between them.

Moving on to the Dense Output Layer, the vocabulary size indicates the number of recommendations provided, aligning with the size of the item collection.

The system undergoes multiple training passes, known as epochs. Additionally, the data is divided into smaller groups called batches. In total, 32 batches were identified, and the system was adjusted accordingly.

For class weights, different importance is assigned to items during training, allowing the system to pay more attention to less common items and improving its ability to handle diverse data.

Finally, the optimizer, a function called Adam, adjusts the model's internal knowledge based on observations during training, leading to continuous improvement. The loss function, sparse categorical cross-entropy, evaluates the system's performance and is particularly suitable for tasks involving different categories of items.

C. Experimental Results

The reliability of the proposed system in this research can be measured using specific evaluation criteria: F1-score, accuracy, precision score, and recall score.

Accuracy: It measures the overall correctness of a model by calculating the ratio of correctly predicted instances to the total instances.

Precision: It quantifies the accuracy of positive predictions, indicating the fraction of correctly predicted positive instances among all instances predicted as positive.

Recall: It measures the ability of a model to capture all positive instances, indicating the fraction of correctly predicted positive instances among all actual positive instances.

F1 Score: It is the harmonic mean of precision and recall, providing a balanced measure that considers both false positives and false negatives.

In the experiments, various models were used, including RNN with LSTM, GRU with LSTM, CuDNNLSTM, and stacked RNN. The focus was primarily on assessing the performance of the Bi-LSTM model in comparison to other types like simple RNN with LSTM, GRU with LSTM, CuDNNLSTM, and stacked RNN.

Fig. 3 shows that the accuracy of the Bi-LSTM model is higher than the other model. Accuracy indicates how often the model makes correct predictions. The data in Fig. 3 indicates that the CuDNNLSTM and GRU with LSTM models achieved an accuracy of 86%, while the proposed model attained 86.4%. The success of the proposed model can be attributed to its utilization of bidirectional LSTM, which enables it to achieve higher accuracy by considering both past and future data in the sequence when making predictions.

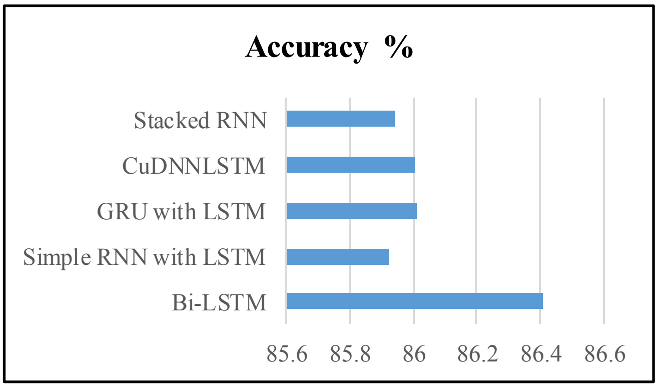


Fig. 3 Accuracy Score Comparison

Fig. 4 shows that the precision score of the Bi-LSTM model surpasses the other models. The proposed Bi-LSTM technique achieved a precision score of 85%. In comparison, other systems such as Simple RNN with LSTM, Stacked RNN, GRU with LSTM, and CuDNNLSTM with LSTM obtained precision scores of 83.69%, 79.32%, 78.64%, and 81.59%, respectively. This comparison illustrates the superior precision of the Bi-LSTM model in identifying relevant items, highlighting its effectiveness in recommendation tasks relative to the other models evaluated.

Precision %

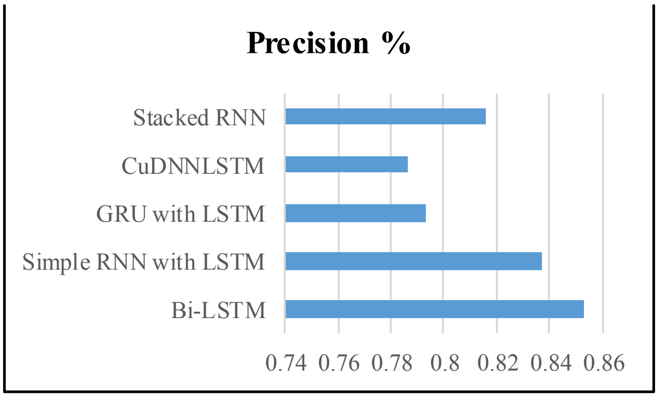


Fig. 4 Precision Score Comparison

Fig. 5 shows the recall score of the Bi-LSTM model outperforms that of other models. The proposed Bi-LSTM technique achieved a recall score of 86.41%. This comparison underscores the superior performance of the Bi-LSTM model in accurately capturing relevant items, making it a more effective choice for recommendation tasks relative to the other models.

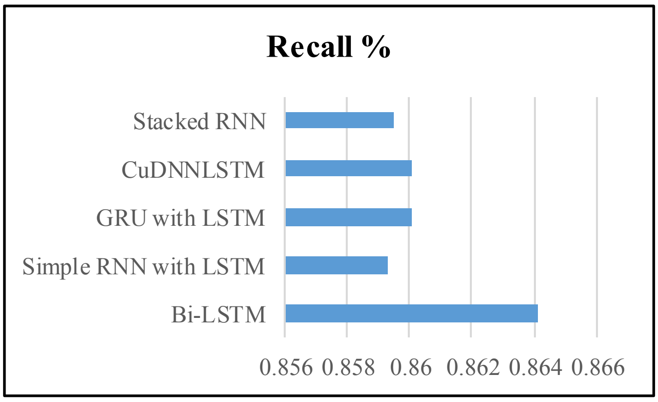


Fig. 5 Recall Score Comparison

Fig. 6 shows that the F1-score of the Bi-LSTM model surpasses that of other models. The proposed Bi-LSTM technique achieved an F1-score of 85.72%. This suggests that the Bi-LSTM model excels in capturing the accuracy as well as the completeness of predictions, making it a robust choice for recommendation tasks.

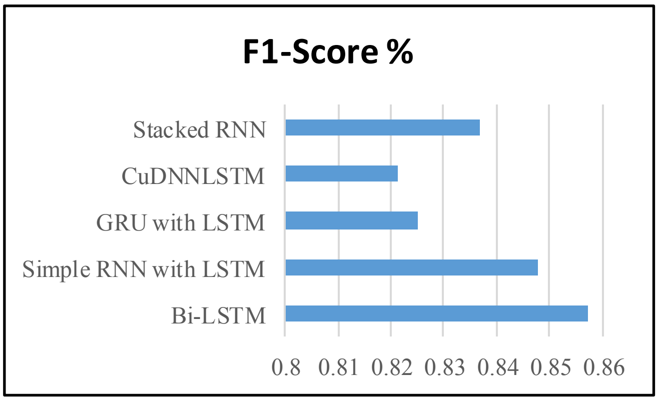


Fig. 6 F1-Score Comparison

In summary, the Bi-LSTM model emerged as a robust performer, demonstrating superior performance across the above key metrics. These findings underscore the effectiveness of the proposed SRec with Bi-LSTM in understanding and predicting user-item interactions comprehensively. Consequently, the model proves to be well-suited for dynamic online environments, where accurate and timely recommendations are essential for enhancing user experience and satisfaction.

VI. CONCLUSION

In today's digital age, recommendation systems have become indispensable for guiding users to relevant content online. These systems play a crucial role in helping users discover suitable items efficiently. However, existing recommendation systems encounter various challenges that affect their accuracy. These challenges include the cold start problem, limited availability of data, biases within the data, evolving user preferences, and the absence of contextual information.

The main goal is to fix problems with current recommendation systems by suggesting a new idea. The SRec model with Bi-LSTM is introduced, which uses attention to understand the context and connections in user sequences. The model can find patterns over time and make accurate predictions for individual items. It's also built to handle situations where the data is uneven, making it work better in real-world recommendation situations.

Looking forward, the future of recommendation systems involves exploring several avenues for improvement. This includes integrating diverse types of information, adapting to users' evolving preferences in real-time, simplifying model interpretability, incorporating reinforcement learning principles, and tackling the challenge of managing large volumes of user data for personalized recommendations. In the future, the exploration of advanced techniques like deep LSTM, deep Bi-LSTM, ensemble models, and multimodal approaches is forecast to amplify recommendation accuracy and effectiveness.

REFERENCES

[1] J. Kim, I. Choi, Q. Li. Customer satisfaction of recommender system: Examining accuracy and diversity in several types of recommendation approaches. Sustainability. 2021 May 30;13(11):6165.

[2] S. Rendle, C. Freudenthaler, L. Schmidt-Thieme. Factorizing personalized markov chains for next-basket recommendation. InProceedings of the 19th international conference on World wide web 2010 Apr 26 (pp. 811-820).

[3] B. Hidasi, A. Karatzoglou, L. Baltrunas, D. Tikk. Session-based recommendations with recurrent neural networks. arXiv preprint arXiv:1511.06939. 2015 Nov 21.

[4] B. Hidasi, M. Quadrana, A. Karatzoglou, D. Tikk. Parallel recurrent neural network architectures for feature-rich session-based recommendations. InProceedings of the 10th ACM conference on recommender systems 2016 Sep 7 (pp. 241-248).

[5] J. Li, P. Ren, Z. Chen, Z. Ren, T. Lian, J. Ma. Neural attentive session-based recommendation. InProceedings of the 2017 ACM on Conference on Information and Knowledge Management 2017 Nov 6 (pp. 1419-1428).

[6] X. Chen, H. Xu, Y. Zhang, J. Tang, Y. Cao, Z. Qin, H. Zha. Sequential recommendation with user memory networks. InProceedings of the eleventh ACM international conference on web search and data mining 2018 Feb 2 (pp. 108-116).

[7] M. Quadrana, A. Karatzoglou, B. Hidasi, P. Cremonesi. Personalizing session-based recommendations with hierarchical recurrent neural networks. Inproceedings of the Eleventh ACM Conference on Recommender Systems 2017 Aug 27 (pp. 130-137).

[8] F. Sun, J. Liu, J. Wu, C. Pei, X. Lin, W. Ou, P. Jiang. BERT4Rec: Sequential recommendation with bidirectional encoder

representations from transformer. InProceedings of the 28th ACM international conference on information and knowledge management 2019 Nov 3 (pp. 1441-1450).

[9] W.C. Kang, J. McAuley. Self-attentive sequential recommendation. In2018 IEEE international conference on data mining (ICDM) 2018 Nov 17 (pp. 197-206). IEEE.

[10] Z. Shao, S. Wang, Q. Zhang, W. Lu, Z. Li, X. Peng. An empirical study of next-basket recommendations. arXiv preprint arXiv:2312.02550. 2023 Dec 5.

[11] C. Zhao, J. You, X. Wen, X. Li. Deep bi-lstm networks for sequential recommendation. Entropy. 2020 Aug 7;22(8):870.

[12] J. Hu, J. Gao, X. Zhao, Y. Hu, Y. Liang, Y. Wang, M. He, Z. Liu, H. Yin. BiVRec: Bidirectional View-based Multimodal Sequential Recommendation. arXiv preprint arXiv:2402.17334. 2024 Feb 27.

[13] R. Devooght, H. Bersini. Collaborative filtering with recurrent neural networks. arXiv preprint arXiv:1608.07400. 2016 Aug 26.

[14] MovieLens Datset, https://grouplens.org/datasets/movielens/

Enhanced Flood Prediction System using BiLSTM and IoT Sensor Networks

Sumit Das
Department of Networking and Communication, Faculty of Engineering and Technology SRM Institute of Science and Technology
Chennai, Tamilnadu
sd0459@srmist.edu.in

Krishnaraj Nagappan
Department of Networking and Communication, Faculty of Engineering and Technology SRM Institute of Science and Technology
Chennai, Tamilnadu
krishnan2@srmist.edu.in

Abstract— Nowadays due to environmental change flooding is a major problem in urban cities. Due to the extreme overflow of water from lakes and rivers, cities face flooding. Manual alarms and the maintenance of regular water flow were vital problems during the rain. Hence, automatic flood forecasting and flood prediction models are needed. The proposed flood monitoring and prediction system, Flood Guard, has been designed using environmental sensors connected with Raspberry Pi Pico W. The PICO W controllers capture real-time data from the environmental sensors of water flow, water level, humidity, temperature, and rainfall, which are essential to detecting flood risk. Based on the water level, the outlet motor actuation speed is triggered to reduce the high intensity of the water during rainy times. All the sensor data collected will be transferred to the IoT Cloud using the message queuing telemetry transport (MQTT) protocol. The forecasting algorithm called a Bidirectional Long Short-Term Memory (BiLSTM) model has been developed in the cloud. This proposed model is trained by the array of sensors from hydrological and meteorological data and evaluated using normalized sensor data. Experimental results show that the proposed model achieves a training accuracy of 100% and a testing accuracy of 98.97%. The training loss and testing loss have been decayed when compared to the other state-of-the-art methods.

Keywords—Flood Forecasting, Computer Vision, Deep Learning, Neural network

I. INTRODUCTION

Flooding has become a major global problem, causing loss of human life and economic damages. It affects many areas, leading to property damage and the need for disaster management. Floods can result in broken infrastructure, such as roads and bridges, and can disrupt essential services. There are many causes of Flooding, such as Heavy rainfall can lead to rivers overflowing their banks. Melting snow and ice can increase water levels in rivers and lakes. Storm surges from hurricanes or typhoons can cause coastal flooding. Changes in land use, like deforestation, can reduce the land's ability to absorb water. Flooding is a major global issue causing loss of life and property damage. Preventing floods is not possible, but we can reduce their impact. Advanced technologies like the Internet of Things (IoT) can help predict floods. Early warnings from IoT devices can save lives and possessions by allowing timely evacuations. The paper discusses an IoT-based system for collecting hydrological and meteorological data. Data from sensors on water flow, level, discharge, and weather conditions are used to predict floods.

The system classifies flood events into different alert levels using an LSTM model [1]. In [2] discusses a sensor designed to measure water levels in dikes using a fiber optic technology called FBG (Fiber Bragg Grating). FBG sensors are unique because they can work well in tough conditions and can measure pressure at different points along a distance. The sensor has a 3D-printed part that changes pressure into strain on the fiber, which can be measured. Temperature changes can affect the sensor, so a second FBG is used to adjust for those effects. The sensor is protected by an aluminum case, making it strong enough for difficult installation and environments. Tests were done on real dikes in the Netherlands, showing that the sensors work well.

Wireless sensor networks (WSNs) were utilized for flood monitoring, the main concern in WSN is it energy efficient. Compressive sensing (CS) was introduced as an energy-efficient solution for remote flood monitoring systems to overcome the energy-efficient issue in WSN. In CS data construction method was used to obtain high-quality results at high compression rate. The primary components of the CS are block-based and learning rate of the design [3]. WSNs have limited battery life, which restricts their long-term deployment without maintenance. Sensors in WSNs need to transmit data over distances, which consumes significant energy. Frequent data transmission by sensor nodes leads to quicker battery depletion. High data sampling rates for accurate monitoring can result in increased energy consumption. Energy is also used in data processing tasks before transmission, adding to the overall power usage. WSN based IoT was developed in [4] flood monitoring in Colima, Mexico, based on the water level, soil moisture and weather conditions. IoT- Message queuing telemetry transport protocol (MQTT) was used to gather data from the real time storm events in 2019. During the flood events the hydrological data needed to be analyzed by the models. This model generates the critical position where the flood will happen soon with inundation maps. Security mechanisms in the IoT networks are lacking when the data is shared. IoT networks lack security mechanisms for data sharing. A lightweight protocol called MQTT was used to manage time constraints in IoT networks. Experimental analysis was done on a testbed setup with hardware and sensors to show the IoT MQTT data transfer effectiveness between two devices [5]. It may introduce the IDISense app, an IoT-based system designed to monitor real-time parameters of dams and weather conditions. A spiking neural network (SNN) predicts rainfall using past weather data. The system includes a central control room for decision-making. Alerts are sent to nearby residents through the IDISense app. The SNN's accuracy in predicting rainfall is compared with other models [6]. Floods may become more common because of global warming. Early detection of water disasters can save lives, and it's important to find good warning systems. People use social media like Twitter to report environmental problems, acting as social sensors. In [7], it discusses the WATERoT project, supported by the Spanish government. WATERSensing was created to help prevent and assess water disasters using social media. The system uses deep neural networks to sense waterlogging events visually. It predicts and maps floods nationally using video sources from the Internet of Cameras. A vehicle detection model identifies vehicles, and a wheel detection model locates wheels to grade flood severity. The system processes input from 2.4K cameras and outputs results in five minutes. Cloud resources and information and communications technology support the workflow. The methodology includes training deep learning models, capturing images, and using GIS mapping for notifications [8]. The methodology uses an Ensemble Kalman Filter (EnKF) algorithm for data assimilation in hydraulic flood forecasting models. It includes the use of gauged water levels for model calibration and validation. The control vector in the DA algorithm is composed of friction coefficients and parameters modifying the upstream boundary conditions. The observation error covariance matrix is assumed to be diagonal with a standard deviation proportional to the observations. The Kalman gain matrix weights the innovation vector and corrects the background control vector [9]. The main problem of flooding is a major global issue causing loss of life and economic damage. Sensors must perform consistently in various environmental conditions, such as extreme weather, and provide accurate readings to inform flood forecasts. IoT devices, including sensors and Arduino boards, often operate on limited power sources. Integrating sensor data with existing hydrological and meteorological models to improve forecast accuracy is complex. The system must effectively combine real-time sensor data with these models for effective prediction.

II. RELATED WORKS

The literature survey on utilizing deep learning techniques for flood detection and mapping through remotely sensed data, especially from real-time ground-level images, highlights an emerging yet promising area of research. This area seeks to address and overcome the limitations associated with traditional flood monitoring methods. Traditional approaches, including the use of water level sensors and remote sensing technologies. The challenges include limited spatial and temporal coverage, delayed information relay, and high operational costs. In contrast, the advent of low-cost vision sensors and advancements in communication technologies have paved the way for large-scale image monitoring, transforming cameras into viable sensors for capturing real-time flood information. In [10], it uses modified neural network architecture to reduce measurement error. It was validated by comparing the actual water level with a trained neural network output. Measurement error bounded by ±1 cm in distance measurements ranging from 2 to 500 cm. In [2], the survey covers the history and development of pressure sensors, focusing on those used in harsh environments. It might discuss the limitations of traditional sensors, such as piezometers and strain gauges, in these settings. In [11], an intelligent system was designed using sensors and microcontrollers, and the data is processed by ESP-32 and stored in the cloud. Global Internet of Things used for user notification of sensor data status captured and monitored using the SARIMAX model. The model has low deviations in forecasting values RMSE: 0.6145, MAPE: 3.20.

The survey [3] highlights the advantages of FBG sensors, including their immunity to electromagnetic interference and ability to perform long-distance measurements. Sparse binary matrices are more energy-efficient and accessible to implement in embedded systems. The literature includes methods for parameter estimation, such as using a Type II maximum likelihood procedure. Learning-based rules, like expectation maximization (EM), are used to estimate model parameters. The study [8] focuses on the challenge of sudden urban floods and the need for efficient sensing systems. The need for physical rulers or consistent reference objects limits conventional methods for water level estimation. The framework estimates maximum water level and topographic deformation using CNN-based methods like Attention U-Net and LinkNet. The approach combines numerical simulation and deep learning for rapid damage estimation. Quantitative evaluation uses metrics like RMSE, IoU, and LSH [12]. Table 1 summarizes the methodologies and key findings from recent studies, showcasing the diverse applications of machine learning, neural networks, and natural language processing in enhancing flood disaster monitoring and response strategies.

TABLE I. OVERVIEW OF ADVANCED ALGORITHMS IN FLOOD MONITORING AND DISASTER PREVENTION RESEARCH

Ref. No	Algorithm used	Merits	Demerits
[3]	BSBL-WSN	Energy-efficient remote flood monitoring system.	The reconstruction process is not as simple as traditional methods.
[6]	SNN	IDISense app gives alerts to neighboring people.	Unstructured data remains laborious for current methods
[7]	NLP and NER	It provides real-time information and actionable intelligence for emergency services.	It needs lot of data to predict the emergency situations
[8]	Mask R-CNN	High scalable compared to other previous approaches.	Unknown class images with noise and wrong field of view
[13]	ResNET model	Different ML models can distinguish irrigation systems regardless of crops	Insufficient information on the impact of secondary crops.
[14]	OmbriaNet ,	OmbriaNet produces high-quality flood maps surpassing state-of-the-art techniques.	Insufficient details on the impact of the flood mapping approach.
[16]	FFNN	FFM integrates locally trained models for accurate flood predictions	Only local trained model gives good accuracy.
[17]	Cascaded-ANFIS	Future flood events were effectively identified using Shared Socio-economic Pathways	Water controlling and managing is not addressed.
[18]	RF, GA-XGBoost, DE-XGBoost	Hybrid XGBoost models superior to existing models for hourly water level prediction.	CART model sensitivity to peak values can affect detection accuracy

The primary motivation of this research is that during flood time, it is critical in flash flood situations where rapid response is essential for community safety and minimizing damage. The primary consideration of the frequently cycloned area is vital for areas prone to such unpredictable and fast-occurring natural disasters. There is a need for flash flood message transfer scenarios where every moment counts and that timely information can save many lives. Need to develop a network of sensors and devices working together to gather, analyze, and communicate flood data, enhancing the accuracy and efficiency of flood prediction and monitoring. The main contributions are to develop an integrated, cost-effective, and energy-efficient mobile sensing system for real-time flood monitoring. This system uniquely combines advanced methods for calculating water discharge with sensor calibration and outlier detection to enhance accuracy. It innovatively employs a Bi-Directional Stacked LSTM model, integrating hydrological and meteorological data for precise real-time flood prediction, showcasing a pioneering approach in flood forecasting technology.

III. METHODOLOGY

The proposed Flood Guard has been developed with the help of a network of sensors for rainfall measurement. These sensors, all controlled by the Raspberry Pi Pico, will collect diverse environmental data essential for accurate flood monitoring and prediction after collecting the various sensors data that is sent into a centralized system managed by the Raspberry Pi Pico. Integrating data from different sources and sensors helps create an outlook on environmental conditions. To develop a flood prediction model using the collected data with the help of a Bi-Directional Stacked LSTM approach. The proposed system predicts and advances issuing alerts and warnings based on the predictive analysis in flood areas. Figure 1 depicts the proposed block diagram.

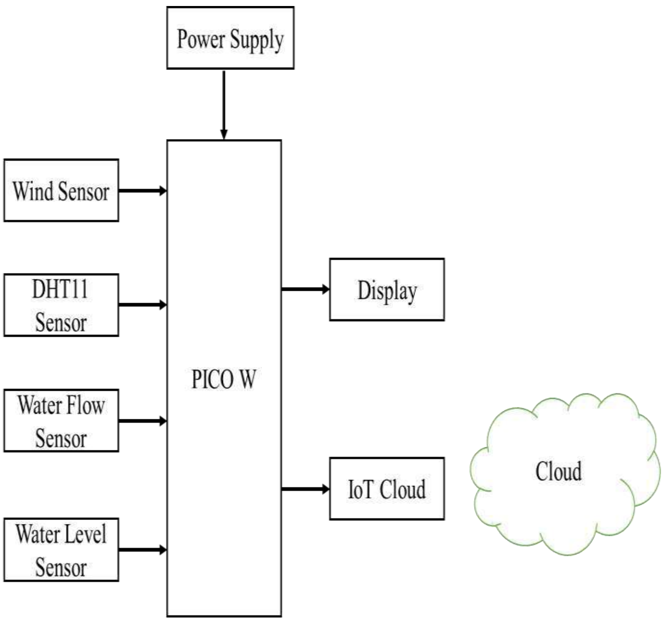


Fig. 1. Proposed framework block diagram

The power supply is the primary source of energy for the system. It is connected to the Raspberry Pi Pico W (Wireless), ensuring that the microcontroller and all connected sensors have a reliable power source for continuous operation. The system has four different types of sensors, each designed to measure specific environmental parameters: wind sensor, DHT11 Sensor, Water Flow Sensor, and Water Level Sensor. The heart of the system is the Raspberry Pi Pico W, a microcontroller with wireless capabilities. It receives data from the various sensors, processes this information, and can make decisions or perform actions based on the programmed logic. Its wireless functionality allows it to communicate with other devices or networks. A display unit connected to the Raspberry Pi Pico W to provide real-time readings and status updates from the sensors. This LCD screen visualizes the data for local monitoring. IoT cloud-based platform where the data collected by the Raspberry Pi Pico W is sent. The cloud platform could offer services like data storage, analysis, and integration with other systems. It enables the scalability of data processing and facilitates remote access from anywhere in the world. The system's main function is for the sensors to collect various environmental data points and send this information to the Raspberry Pi Pico W. The Pico W processes this data, which may involve preliminary analysis or simply formatting the data for transmission. The processed data is displayed locally for monitoring and sent to the IoT Cloud for storage, further analysis, and integration. The Cloud processes this data using Bi-Directional Stacked LSTM to predict potential flooding events. If the system predicts a high risk of flooding, alerts can be generated and distributed to the necessary channels to warn local authorities and residents.

A. Sensor Network Deployment and Data Collection

The sensor network deployment and data collection modules are the initial step for the proposed flood guard framework. The flood guard consists of sensors that measure water level, water flow, temperature, humidity, rainfall measurement, wind speed, and wind direction. These sensors are used to sense both hydrological and metrological data in the river basin or lake region. All the sensors are connected to Raspberry pi Pico W which serves as the control unit, later, all the data will be transferred to the IoT cloud.

B. Data Integration and Preprocessing

The data obtained from the sensor network has been sent to a centralized system, which the Raspberry Pi Pico also controls. This data integration and preprocessing module combines the data from all sensor nodes to understand the environmental conditions on the monitored region.

C. Flood Prediction Modeling

Flood Guard employs a Bi-Directional Stacked Long Short-Term Memory (Bi-LSTM) model, capable of learning from data sequences, to predict flood occurrences. The selection of the Bi-LSTM model stems from its capacity to capture both historical and future context within environmental data.

D. Alert and Warning System

Based on the predictive analysis conducted by the Bi-LSTM model, the proposed Flood Guard will issue alerts and warnings about potential flood events. The system evaluates the forecasted data against predefined thresholds to determine the risk level of flooding. When these thresholds are exceeded, the system triggers alerts through various communication channels, such as SMS, email, or dedicated mobile apps, to inform authorities and residents in the affected areas.

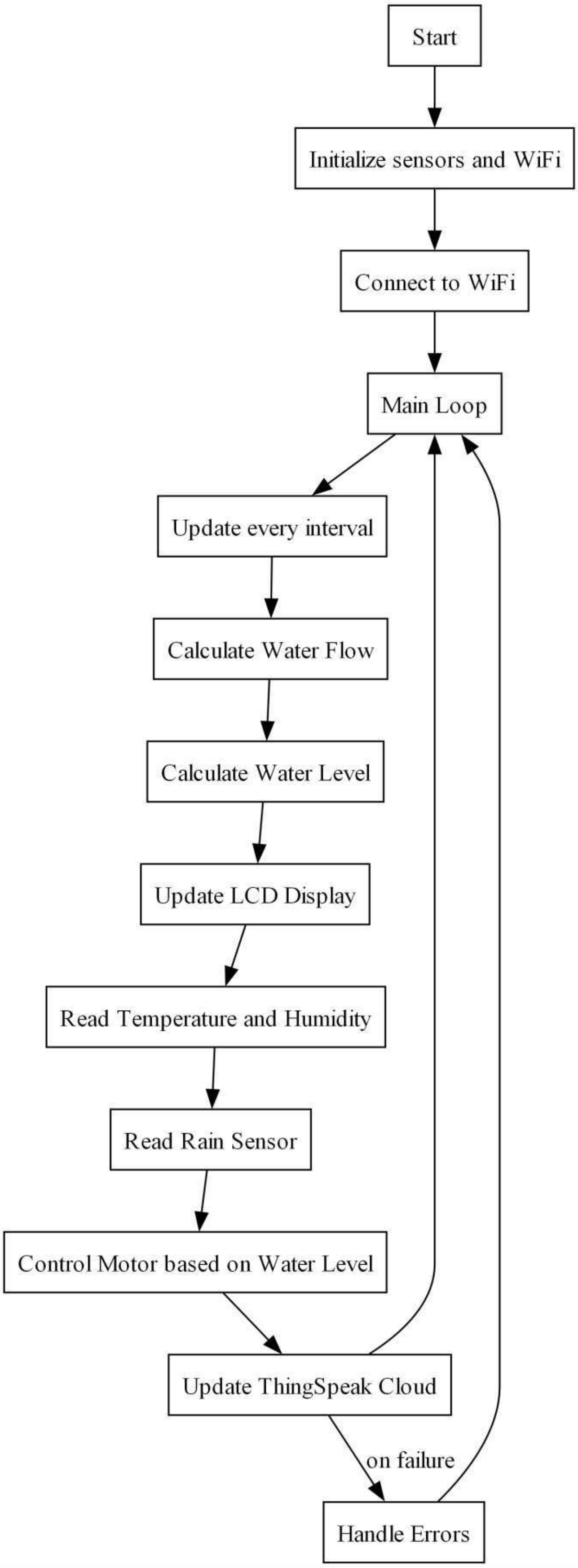


Fig. 2. Operational Flowchart of the IoT-Based Flood Monitoring System

Figure 2 depicts the flowchart of the proposed IoT-based flood monitoring and prediction system. The process begins with the sensor initialization and a Wi-Fi connection. Once the system is connected, it enters a main loop that updates the sensor readings, such as water flow, water level, humidity and temperature at regular intervals. The system calculates water flow and level within this loop, updates an LCD with this information, and reads temperature, humidity, and rainfall data. Based on the water level readings, it controls a motor to manage water flow or activate a floodgate.

Additionally, the system updates a cloud service, ThingSpeak, with the collected data for remote monitoring and analysis. If there is a failure during the cloud update, the system handles errors accordingly. This cycle of operations continues indefinitely, ensuring continuous monitoring and reactive measures to potential flooding.

IV. Results and Discussion

The proposed floodwall hardware collects environmental data from various sensors, including wind, water flow, water level, temperature, humidity, and rainfall. Processing and analyzing this data in real-time with a Raspberry Pi Pico W. Actuating a motor to potentially control water flow or floodgates based on the water level data and displaying the sensor data on an LCD screen for local monitoring. The collected data will be sent to the ThingSpeak cloud platform for further analysis and remote monitoring. Table II presents data from an environmental monitoring dataset originally structured in time series format. Each row corresponds to a different parameter measured by flood monitoring sensors, such as Water Flow, Water Level, Water Discharge, Temperature, Humidity, Wind Speed, and Wind Direction. The columns represent consecutive hourly timestamps starting from midnight on January 6, 2024. The sensor readings at each timestamp facilitate a parameter-centric view, which can be advantageous for certain data analyses, such as comparing distributions across different environmental factors or sensor calibration processes.

TABLE II. Environmental DataSet from Flood Monitoring Sensors

Attributes	**1**	**2**	**3**	**4**
Time Stamp	2024-01-06 00:00:00	2024-01-06 01:00:00	2024-01-06 02:00:00	2024-01-06 03:00:00
Water Flow (m^3/s)	150	155	160	180
Water Level (m)	3.2	3.3	3.5	4.0
Water Discharge (m^3/s)	480	500	520	580
Temperature (°C)	26	26	27	28
Humidity (%)	80	82	85	75
Wind Speed (km/h)	10	12	9	15
Wind Direction	NW	N	NE	NW

TABLE III. COMPARISON OF TRAINING AND TESTING LOSSES FOR BiLSTM, LSTM, AND CNN MODELS IN FLOOD PREDICTION

Proposed BiLSTM		LSTM		CNN	
Training Loss	**Testing Loss**	**Training Loss**	**Testing Loss**	**Training Loss**	**Testing Loss**
6.35×10^{-6}	5.47×10^{-6}	1.05×10^{-5}	1.28×10^{-5}	6.24×10^{-6}	6.00×10^{-6}

Table III presents the training and testing losses for three different neural network architectures, BiLSTM, LSTM, and CNN, utilized in flood prediction. The loss values are measured in terms of mean squared error (MSE) and are provided in scientific notation for clarity.

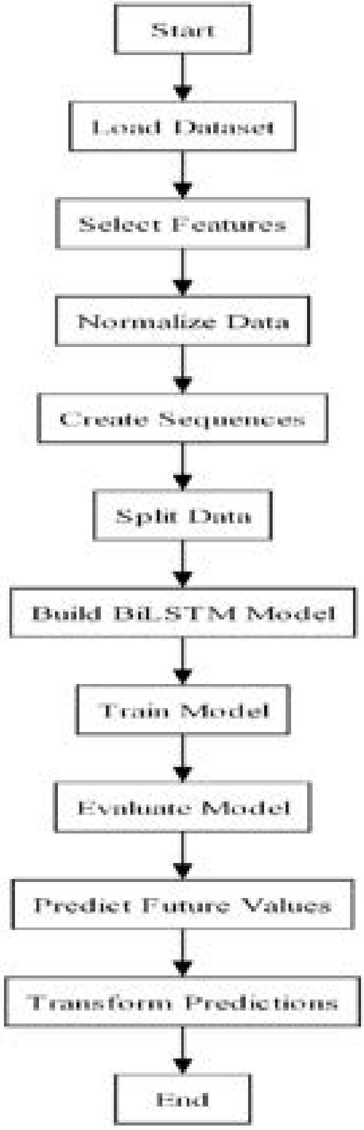


Fig. 3. Workflow of BiLSTM Model Development for Time Series Forecasting

Figure 3 depicts a sequential workflow for developing a Bidirectional Long Short-Term Memory (BiLSTM) model tailored for forecasting based on time series data. his workflow encapsulates the end-to-end process of building and utilizing a machine learning model for predictive analysis in time series data.

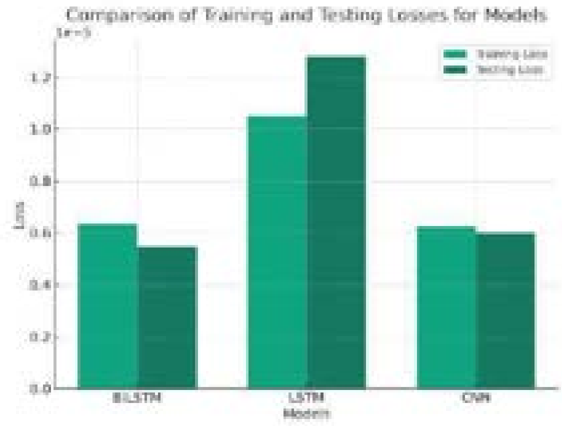


Fig. 4. Comparison of Training and Testing loss for Models

Figure 4 depicts the comparison of training and testing loss for models. The BiLSTM model outperforms the LSTM and CNN regarding testing loss, highlighting its effectiveness in generalizing well to new, unseen data, which is critical for reliable flood prediction. The dual-direction learning capability of the BiLSTM contributes to its enhanced performance by effectively utilizing the entire context of the input data sequence.

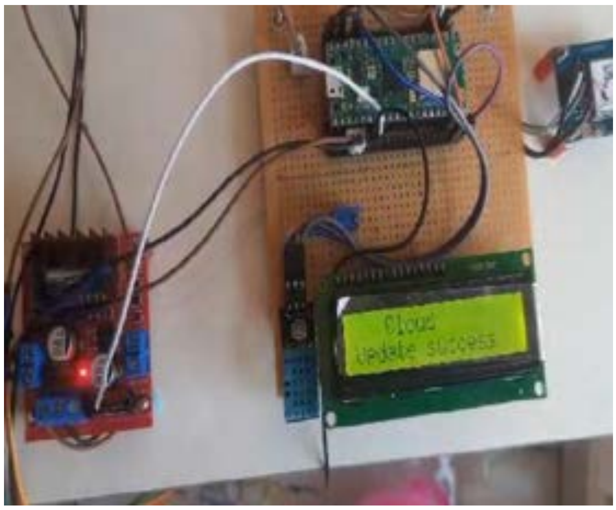


Fig. 5. IoT Based Flood Monitoring System Hardware Prototype

Figure 5 displays an IoT flood monitoring setup with a Raspberry Pi Pico W connected to various sensors and a display indicating a successful data transmission to the cloud, signifying the system's ability to update and monitor flood-related data in real time remotely.

V. CONCLUSION

Flood Guard system was represents a significant step forward in real-time flood monitoring and forecasting. A sophisticated network of environmental sensors coupled with the computational prowess of the Raspberry Pi Pico W, the system for data acquisition. Incorporating a Bidirectional Long Short-Term Memory (BiLSTM) model on the sequential nature of the data gave accurate predictions of flood-related events. This system facilitates immediate local response through its automated control of actuators based on water levels. It extends its utility to remote monitoring via cloud connectivity to the Thing Speak platform. In future, the data from the multiple-point node will be trained using a federated learning algorithm.

REFERENCES

[1] C. Prakash, A. Barthwal, and D. Acharya, “FLOODWALL: A Real-Time Flash Flood Monitoring and Forecasting System Using IoT,” *IEEE Sensors J.*, vol. 23, no. 1, pp. 787–799, Jan. 2023, doi: 10.1109/JSEN.2022.3223671.

[2] L. Schenato, J. P. Aguilar-Lopez, A. Galtarossa, A. Pasuto, T. Bogaard, and L. Palmieri, “A Rugged FBG-Based Pressure Sensor for Water Level Monitoring in Dikes,” *IEEE Sensors Journal*, vol. 21, no. 12, pp. 13263–13271, 2021, doi: 10.1109/JSEN.2021.3067516.

[3] V. Abolghasemi and M. H. Anisi, “Compressive Sensing for Remote Flood Monitoring,” *IEEE Sensors Letters*, vol. 5, no. 4, pp. 20–23, 2021, doi: 10.1109/LSENS.2021.3066342.

[4] O. Mendoza-Cano *et al.*, “Experiments of an IoT-based wireless sensor network for flood monitoring in Colima, Mexico,” *Journal of Hydroinformatics*, vol. 23, no. 3, pp. 385–401, 2021, doi: 10.2166/HYDRO.2021.126.

[5] H. Siddharthan, T. Deepa, and P. Chandhar, “SENMQTT-SET: An Intelligent Intrusion Detection in IoT-MQTT Networks Using Ensemble Multi Cascade Features,” *IEEE Access*, vol. 10, pp. 33095–33110, 2022, doi: 10.1109/ACCESS.2022.3161566.

[6] M. Braveen, K. Anusha, M. A. Jerlin, R. Seetha, and N. Sampath, “IDISense: IoT-Based Dam Water Disaster Sensing and Prevention System,” *IEEE Sensors Journal*, vol. 23, no. 23, pp. 29451–29457, 2023, doi: 10.1109/JSEN.2023.3322290.

[7] J. M. Cecilia *et al.*, “WATERSensing: A Smart Warning System for Natural Disasters in Spain,” *IEEE Consumer Electronics Magazine*, vol. 10, no. 6, pp. 89–96, 2021, doi: 10.1109/MCE.2021.3063703.

[8] S. W. Lo, J. H. Wu, J. Y. Chang, C. H. Tseng, M. W. Lin, and F. P. Lin, “Deep sensing of urban waterlogging,” *IEEE Access*, vol. 9, pp. 127185–127203, 2021, doi: 10.1109/ACCESS.2021.3111623.

[9] T. H. Nguyen *et al.*, “Improvement of Flood Extent Representation With Remote Sensing Data and Data Assimilation,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–22, 2022, doi: 10.1109/TGRS.2022.3147429.

[10] A. K. Sahoo and S. K. Udgata, “A Novel ANN-Based Adaptive Ultrasonic Measurement System for Accurate Water Level Monitoring,” *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 6, pp. 3359–3369, 2020, doi: 10.1109/TIM.2019.2939932.

[11] N. Bharathiraja, T. Deepa, S. Hariprasad, and A. Chokkalingam, “Design and Development of GIOT based Intelligent Smart Waste Management and Predictive Modelling,” in *2022 6th International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India: IEEE, Apr. 2022, pp. 7–12. doi: 10.1109/ICOEI53556.2022.9777210.

[12] N. Yokoya *et al.*, “Breaking Limits of Remote Sensing by Deep Learning from Simulated Data for Flood and Debris-Flow Mapping,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 60, pp. 1–15, 2022, doi: 10.1109/TGRS.2020.3035469.

[13] G. Paolini, M. J. Escorihuela, O. Merlin, M. P. Sans, and J. Bellvert, “Classification of Different Irrigation Systems at Field Scale Using Time-Series of Remote Sensing Data,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, no. Ml, pp. 10055–10072, 2022, doi: 10.1109/JSTARS.2022.3222884.

[14] G. I. Drakonakis, G. Tsagkatakis, K. Fotiadou, and P. Tsakalides, “OmbriaNet-Supervised Flood Mapping via Convolutional Neural Networks Using Multitemporal Sentinel-1 and Sentinel-2 Data Fusion,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 15, pp. 2341–2356, 2022, doi: 10.1109/JSTARS.2022.3155559.

[15] M. Thomas *et al.*, “A Framework to Assess Remote Sensing Algorithms for Satellite-Based Flood Index Insurance,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 16, pp. 2589–2604, 2023, doi: 10.1109/JSTARS.2023.3244098.

[16] M. S. Farooq *et al.*, “FFM: Flood Forecasting Model Using Federated Learning,” *IEEE Access*, vol. 11, no. March, pp. 24472–24483, 2023, doi: 10.1109/ACCESS.2023.3252896.

[17] N. Rathnayake, U. Rathnayake, I. Chathuranika, T. L. Dang, and Y. Hoshino, “Projected Water Levels and Identified Future Floods: A Comparative Analysis for Mahaweli River, Sri Lanka,” *IEEE Access*, vol. 11, no. January, pp. 8920–8937, 2023, doi: 10.1109/ACCESS.2023.3238717.

[18] D. H. Nguyen, X. Hien Le, J. Y. Heo, and D. H. Bae, “Development of an Extreme Gradient Boosting Model Integrated with Evolutionary Algorithms for Hourly Water Level Prediction,” *IEEE Access*, vol. 9, pp. 125853–125867, 2021, doi: 10.1109/ACCESS.2021.3111287.

[19] M. Arosio, C. Arrighi, L. Cesarini, and M. L. V. Martina, “Service accessibility risk (SAR) assessment for pluvial and fluvial floods in an urban context,” *Hydrology*, vol. 8, no. 3, 2021, doi: 10.3390/hydrology8030142.

Assessing Metrics using LoRa Technology for Real-Time Landslide Monitoring within an IoT Framework

1st Mrs. Inbavalli A
Department of Computer Science
IFET College of Engineering
Villupuram, Tamilnadu
inbavallikumaran.2010@gmail.com

2nd Ms. Praveena R
Department of Computer Science
IFET College of Engineering
Villupuram, Tamilnadu
praveenaraj3523@gmail.com

3rd Ms. Kaviya M
Department of Computer Science
IFET College of Engineering
Villupuram, Tamilnadu
kathiravan232007@gmail.com

***Abstract*—This research study introduces a groundbreaking approach to landslip monitoring through the seamless integration of low-power sensors with the versatile Arduino NANO platform, leveraging LoRa-based IoT technology. By facilitating real-time analysis of critical environmental variables such as water levels, soil movement, and humidity, this system enables early detection of soil instability, potentially preventing devastating landslides. Automated email notifications promptly alert authorities and communities to impending threats, enhancing crisis management effectiveness. LoRa technology's extensive coverage, affordability, and adaptability address shortcomings of traditional monitoring methods, while its versatility extends applicability beyond landslip monitoring. This proactive system empowers communities to intelligently respond to environmental changes, providing comprehensive, real-time monitoring for hazard identification and preventive action formulation, ultimately mitigating disaster risks and enhancing overall disaster response capabilities.**

***Index Terms*—LoRa-based IoT System, Landslip Monitoring, Low-power Sensors, Real-time Analysis.**

I. INTRODUCTION

The real-time data gathering, transmission, and analysis made possible by the Internet of Things (IoT) is essential to the modernization of landslide monitoring [1]. IoT devices, which are outfitted with an array of sensors, function as primary data collectors, persistently tracking environmental factors such as soil moisture content, vibrations, and precipitation. This information is wirelessly sent to a central monitoring station, allowing for the prompt deployment of mitigating measures and the timely identification of possible landslide incidents [2], [3]. IoT also makes it possible to integrate machine learning and cloud computing, which improves data analytics and predictive modeling capabilities. Consequently, IoT is essential to building a dynamic and proactive framework for monitoring landslides, which improves disaster resilience in areas that are susceptible to disasters [4], [5].

A. Background and Motivation:

The impetus for this research stems from the critical need for a sophisticated landslip monitoring system that addresses the limitations of current techniques [6]. Traditional methods relying on manual data gathering and sporadic inspections often fail to provide accurate and timely information about evolving environmental conditions. By integrating low- power sensors with Arduino NANO and LoRa technology, a revolutionary shift in landslip monitoring becomes feasible [7]. This system enables continuous real-time data collection across various environmental parameters, facilitating a comprehensive understanding of factors contributing to landslip incidents. The advent of IoT technologies, particularly LoRa, has revolutionized data transmission from remote locations due to its extended range and low power consumption [8]. By integrating diverse sensors measuring water levels, vibrations, humidity, and soil moisture, holistic examination of landslide- prone areas becomes achievable. Ultimately, this project aims to enhance early detection capabilities through IoT integration, empowering communities and authorities to proactively respond to changing environmental conditions [9].

B. Objectives:

The primary objective of our research is to pioneer the development and implementation of a real-time landslip monitoring system utilizing low-power sensors and LoRa technology [10]. This involves the construction of robust hardware incorporating an Arduino NANO and various sensors to detect crucial environmental variables. Concurrently, we aim to evaluate the system's efficacy in providing real-time data and early warnings for landslide events. Through comprehensive testing and certification processes, we endeavor to demonstrate the system's reliability and its potential to significantly reduce response times during emergencies. This research holds promise for enhancing disaster preparedness and mitigation efforts in landslide-prone areas [11].

C. Scope of the Paper:

This article delves into the intricacies of monitoring land-slides in real-time through the innovative integration of IoT and LoRa technology [12]. It meticulously outlines the design, installation, and evaluation of the IoT system, encompassing hardware specifics, sensor technologies, and communication infrastructure. Furthermore, the study extends its purview beyond landslides, highlighting the system's adaptability to diverse natural catastrophe scenarios [13]. A comprehensive exploration is promised, elucidating the seamless integration of low-power sensors with LoRa technology within an IoT framework [14]. Practical implications and contributions to disaster management are expounded upon, offering valuable insights into this pioneering approach [15].

II. Literature Review

The literature review aims to provide a comprehensive understanding of the existing research related to landslide monitoring systems, particularly focusing on the integration of IoT technologies. However, it has been noted that the survey of literature in the initial draft lacked systematic organization and clarity. In this revised version, we present a more structured and methodical review of relevant literature to address this critique.

A. Overview of Landslide Monitoring Techniques:

Landslide monitoring techniques have undergone significant advancements in recent years, driven by the need for more effective and reliable methods of hazard detection and mitigation. Traditional approaches, such as manual surveys and satellite-based monitoring, have limitations in terms of spatial and temporal resolution, as well as cost-effectiveness. Studies by [1] and [2] underscore the importance of integrating emerging technologies, such as IoT and wireless sensor networks, to overcome these limitations and enhance landslide monitoring capabilities.

B. IoT-Based Landslide Monitoring:

The integration of IoT technologies has revolutionized landslide monitoring by enabling real-time data collection, analysis, and early warning systems. Leveraging IoT devices, such as low-power sensors and communication modules, re- searchers have developed innovative solutions for detecting and mitigating landslide risks. For example, the work of [3] demonstrates the feasibility of using IoT-enabled sensors to monitor slope stability and trigger timely alerts in the event of potential landslides. Similarly, [4] proposes a framework for integrating IoT and cloud computing technologies to facilitate data-driven decision-making in landslide-prone areas.

C. LoRa Technology in Environmental Monitoring:

LoRa (Long Range) technology has emerged as a promising communication protocol for IoT applications, offering long-range connectivity, low power consumption, and scalability. In the context of environmental monitoring, LoRa-based sensor networks have been successfully deployed for various applications, including landslide detection, soil moisture monitoring, and weather forecasting. Notable studies by [3] and [2] high- light the advantages of LoRa technology in providing reliable and cost-effective connectivity for remote sensing applications, particularly in challenging terrain conditions.

D. Methodical Survey of Relevant Literature:

To ensure a systematic review of the literature, we adopted a structured approach to identify and analyze relevant studies. We defined specific search criteria, including keywords such as ”landslide monitoring,” ”IoT,” and ”LoRa technology,” and conducted searches across reputable academic databases, including IEEE Xplore, ScienceDirect, and Google Scholar. In addition, we utilized citation chaining and reference list scanning techniques to identify additional relevant studies. The search process yielded a comprehensive collection of peer-reviewed articles, conference papers, and technical reports, providing a solid foundation for our literature In summary, the literature review provides a comprehensive overview of previous research on landslide monitoring systems, with a particular focus on the integration of IoT technologies and the use of LoRa technology for environmental monitoring. By adopting a systematic approach to literature review, we aim to ensure the thoroughness and reliability of our study's theoretical foundation.

III. System Architecture

A groundbreaking advancement in environmental monitoring, particularly in landslide detection and prevention, is embodied in the LoRa-based IoT Landslide Monitoring Sys- tem. This innovative system leverages Long Range (LoRa) technology for seamless wireless data transmission from low- power sensors to a central station.

A key component of the design of the landslide monitoring system is reliability. To reduce single points of failure, redundancy mechanisms are implemented and numerous sensors are installed for each parameter. Data transmission failover methods guarantee continuous operation, supported by routine maintenance and calibration checks to maintain data correctness and system integrity. Together, these characteristics improve the system's efficacy and durability in practical situations. By centralizing data, real-time analysis enables early identification of soil instability and potential landslip events. Figures 1 and 2 depict the system's architecture, delineating the transmitter and receiver modules. This paradigm shift offers significant potential for enhancing landslide monitoring efficacy, facilitating proactive measures to mitigate risks, and safeguarding communities and infrastructure.

This system's primary strength is its ability to combine several low-power sensors into a unified framework, such as vibration, wetness, float, and humidity sensors. Together, these sensors provide a thorough picture of the surrounding environment, which helps the system evaluate elements like soil movement, humidity, and water levels that are critical for landslip prediction.

A. Overview of the LoRa-based IoT Landslide Monitoring System:

The brains of the innovative LoRa-based Landslip Monitoring System reside within its transmitter component. This vital section incorporates an Arduino NANO microcontroller alongside an array of sensors, serving as the cornerstone

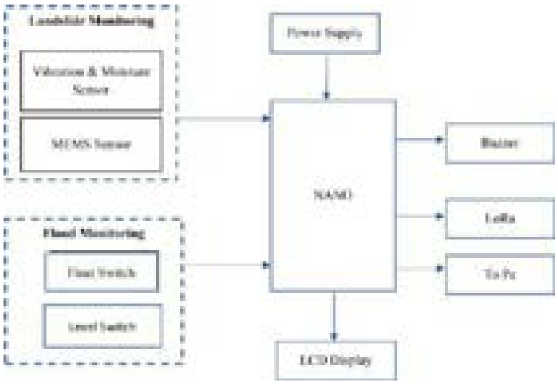


Fig. 1. Block diagram of transmitter module.

Fig. 2. Block diagram Of deceiver module.

for data collection, analysis, and wireless transmission. Essential components include moisture sensors for monitoring soil moisture levels crucial in landslide detection, rain sensors for detecting precipitation aiding in flood monitoring, MEMS sensors offering insights into device orientation and movement to identify potential shifts indicating landslides, and float sensors to alert users of rising water levels, vital for flood prevention. Using a multi-sensor approach, the suggested landslide monitoring methodology inside the LoRa-based IoT framework collects environmental data in real time and identifies possible landslide occurrences. The system uses vibration, moisture, and MEMS sensors, among other types of sensors, to continually monitor crucial factors that are symptomatic of soil instability. These sensors provide smooth data transfer to a central station for analysis and decision-making since they are interfaced with Arduino NANO and LoRa modules. To enable prompt reaction and catastrophe mitigation, the system places a strong emphasis on early detection through quickdata processing and event triggering. However, to guarantee the precision and dependability of the monitoring system, issues like sensor calibration and data synchronization must be carefully considered. Subsequent investigations may go into optimization methodologies aimed at augmenting system efficacy and functionality under varying environmental circumstances.

B. Components of the Transmitter Section:

The core of the LoRa-based Landslip Monitoring System lies within its transmitter unit, housing crucial components like the Arduino NANO microcontroller and an array of sensors. These sensors, including moisture sensors for soil moisture assessment and rain sensors for detecting precipitation, play vital roles in both landslide and flood monitoring. MEMS sensors contribute by providing orientation and movement data, aiding in the identification of potential shifts or tilts indicative of impending landslides. Additionally, the float sensor serves to alert users of rising water levels, enhancing flood monitoring capabilities. This comprehensive system amalgamates various sensor technologies for effective disaster detection and prevention, showcasing its invaluable utility in safeguarding vulnerable areas.

C. Landslide Detection and Flood Monitoring Module:

The core of the system is the Landslip Detection and Flood Monitoring Module, which is made to process data from several sensors and take preventative action in the event of impending hazards. This module, which is integrated with an Arduino NANO, continually and real-time monitors environmental conditions. The Arduino NANO gathers sensor data at the start of the procedure. The Arduino examines this data if it detects any notable changes or anomalies, such as strong vibrations, high moisture content, abrupt tilts, or rising water levels. The Arduino alerts the LoRa module in the event of a hazardous landslip or flood. Real-time data is shown on an LCD screen simultaneously, enabling local monitoring and evaluation.

D. Description of Sensors Used:

The sensors used in the system have specialized functions related to flood monitoring and landslip detection. The moisture sensor uses capacitance to measure the water content of the soil and displays the result as a percentage. Accelerometer-equipped MEMS sensors track the direction and motion of devices. While rain sensors identify and quantify rainfall, vibration sensors pick up vibrations in the ground. Float sensors alert people to rising water levels because they are made for monitoring floods.

E. Role of Arduino NANO in Data Collection and Processing:

The transmitter section's brain, the Arduino NANO, coordinates the activities involved in gathering and processing data. The Arduino NANO analyses the data it receives from the array of sensors to find any hazards like floods or landslides. Such risks cause the LoRa module to transmit wirelessly, guaranteeing a quick connection with the central station. Making decisions quickly depends on the Arduino NANO's real-time capabilities. Furthermore, local data presentation on an LCD screen improves monitoring and assessment skills and enables prompt responses to changing environmental conditions. With the assistance of several low-power sensors, the integration of Arduino NANO not only allows for effective data processing

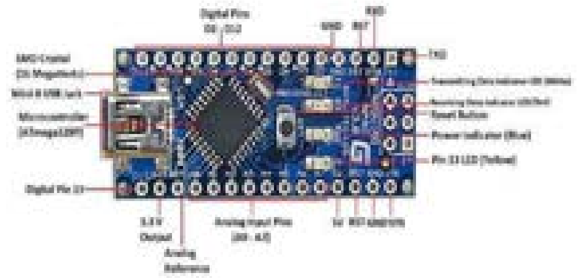


Fig. 3. Arduino nano kit.

but also adapts the system to a variety of applications. The Arduino Nano kit or model is shown in Fig. 3.

In summary, the system architecture builds a strong foundation for flood detection and landslip monitoring by utilizing LoRa technology with Arduino NANO. The combination of sensors and microcontrollers works in concert to provide real- time data analysis, which improves system responsiveness and, in the end, lowers the risks related to natural disasters.

The sensors' reaction time is explained to evaluate how well the sensors work in the monitoring system. Response time, or the amount of time between stimulus detection and data transmission, is measured by conducting controlled tests on various types of sensors, such as vibration, moisture, and MEMS sensors. To achieve precision, experimental setups require carefully calibrated equipment and regulated ambient conditions. Carefully collected sensor measurements are taken under various situations to determine response time variability. The results show that response times, which depend on variables including sensor type and ambient dynamics, range from milliseconds to seconds. These results are interpreted in the discussion, with emphasis on the consequences for system dependability and possible ways to improve sensor performance. These revelations strengthen the system's ability to identify landslides in real time and highlight how crucial sensor responsiveness is to IoT-based monitoring systems.

IV. METHODOLOGY

The specifics of the approach used in the LoRa-based Internet of Things landslip monitoring system. Detecting landslides, monitoring floods, and integrating our system with the Thingspeak website are all included in the technique. Real-time data gathering and transmission are also included.

A. Real-Time Data Collection and Transmission:

The system's core functionality lies in its ability to wirelessly transmit real-time data through the LoRa network efficiently. It relies on an Arduino NANO paired with various sensors like moisture, rain, MEMS, float, and vibration sensors to gather crucial environmental data. Strategically positioned, these sensors monitor parameters such as soil moisture, rainfall intensity, ground vibrations, and tilts. Leveraging LoRa's low-power features ensures sustainable operation with minimal energy consumption. The Arduino NANO triggers data transmission upon detecting anomalies, facilitating rapid response to potential landslip incidents. Additionally, it enables local monitoring via an LCD screen, empowering real-time analysis and proactive measures.

B. Method for Detecting Landslides:

The system's core functionality revolves around identifying potential landslip occurrences and proactively addressing them. Utilizing data from sensors, such as the Arduino NANO, it initiates landslip detection procedures. A crucial component is the vibration sensor, which continuously monitors ground movements indicative of imminent landslides. Simultaneously, the moisture sensor measures soil moisture content, crucial for landslide prediction. Any significant deviation from average moisture content triggers the landslip detection process. Additionally, an accelerometer-equipped MEMS sensor tracks system orientation and movement, detecting abrupt tilts or positional changes, enhancing the system's accuracy in evaluating landslide risk. This multi-sensor approach ensures comprehensive and precise landslip risk assessment.

C. Flood Monitoring Process:

To mitigate the threats posed by increasing water levels, our system includes a flood monitoring module in addition to landslip detection. The float sensor, which was created especially for flood monitoring, recognizes variations in water levels and sounds an alert when there may be a chance of flooding. Early warnings about impending flooding events can be given by the system because it regularly monitors water levels. Our system's adaptability is increased with the addition of flood monitoring, transforming it into a comprehensive defense against many environmental threats.

D. Integration with Thingspeak Website:

Our innovative LoRa-based Internet of Things landslip monitoring system seamlessly integrates with the Thingspeak platform, facilitating centralized data storage, analysis, and accessibility. Real-time sensor data transmission enables swift processing and archiving on the Thingspeak server, empowering remote monitoring and analysis. This integration ensures fast access to critical information for emergency services and local authorities. With visualization tools offered by Thingspeak, decision-makers can construct graphs and charts for better interpretability, aiding in making informed choices during emergencies. This resilient approach offers precise landslide detection, flood monitoring, and proactive environmental hazard management, marking it as a comprehensive and dependable solution.

E. Monitoring Process:

The monitoring process involves a series of steps aimed at collecting, transmitting, and analyzing data from the deployed sensors to detect potential landslides and monitor environmental conditions. This section outlines the key components and procedures involved in performing the monitoring process:

1) Sensor Deployment and Configuration: To initiate the monitoring process, sensors are strategically deployed in the target area prone to landslides. The placement of sensors is determined based on factors such as terrain characteristics, historical landslide occurrences, and the desired scope of monitoring. Each sensor is configured to collect specific environmental parameters, such as soil moisture, vibration, and rainfall, relevant to landslide detection and monitoring.

2) Data Collection and Transmission: Once deployed, the sensors continuously collect data regarding the monitored environmental parameters at predefined intervals. The collected data is then transmitted wirelessly to a central station using LoRa technology. The LoRa network facilitates long-range communication with low power consumption, ensuring efficient transmission of data from remote sensor nodes to the central monitoring station.

3) Data Processing and Analysis: At the central monitoring station, the transmitted data is received and processed in real time. Advanced data processing algorithms are applied to analyze the collected data, identify patterns, and detect anomalies indicative of potential landslide events. This involves comparing current sensor readings with historical data and predefined thresholds to assess the stability of the monitored area.

4) Alert Generation and Notification: Upon detection of abnormal environmental conditions suggestive of a landslide threat, the monitoring system generates alerts in real time. These alerts are transmitted to relevant stakeholders, including local authorities and emergency response teams, via email or other communication channels. The timely dissemination of alerts enables swift response and implementation of precautionary measures to mitigate the risk posed by landslides.

5) Continuous Monitoring and Maintenance: The monitoring process is continuous and ongoing to ensure the timely detection of potential landslide events and the effective management of environmental risks. Regular maintenance and calibration of sensors are conducted to ensure the accuracy and reliability of data collected by the monitoring system. Additionally, periodic inspections of sensor nodes and the LoRa network infrastructure are performed to address any technical issues and optimize system performance.

V. Components Explanation

The LoRa-based Internet of Things landslip monitoring system's efficacy stems from its sophisticated integration of multiple sensors and parts. Every component has a specific function in making sure the system can identify, evaluate, and react to environmental factors that could indicate future floods and landslides. We go into these components' specifics in this part, describing their features and how they add to the system's overall resilience.

A. Sensor for Detecting Landslides:

The Landslip Detection Sensor, a vital component of the LoRa-based IoT landslip monitoring system, is designed to detect subtle ground movements that could indicate potential landslides. Utilizing advanced technologies such as accelerometers, it can accurately pinpoint even the slightest vibrations in the Earth's surface. Strategically positioned, these sensors gather environmental signals and promptly alert the central processing unit upon detecting significant vibrations. This early detection capability is crucial for issuing timely alerts and mitigating potential disasters, thereby enhancing overall disaster preparedness and providing real-time data on ground conditions.

B. Soil Moisture Sensor:

Soil moisture plays a pivotal role in landslip prediction, and the precision of the system's soil moisture sensor is paramount. This sensor, utilizing capacitance, delivers real-time data by measuring soil water content and presenting it as a percentage. Its buried placement ensures accurate readings, crucial for pinpointing landslide-prone areas. Continuously gathering data, the sensor converts moisture fluctuations into analog voltages, facilitating comprehensive stability evaluations. Integrated into the Internet of Things (IoT) framework, it enhances adaptability across diverse environmental conditions, making it an indispensable component for comprehensive landslip surveillance and risk assessment.

C. MEMS Sensor:

The landslip monitoring system uses the adaptable Micro-Electro-Mechanical Systems (MEMS) sensor to track movement and orientation. The MEMS sensor's accelerometer allows it to recognize abrupt tilts and positional changes in the device. This capacity is very helpful in spotting changes in the monitoring system or possible landslides. The MEMS sensor's job is to monitor the orientation and movement patterns of the device continually. The sensor warns of the potential for a landslip event when deviations beyond predetermined criteria are found. The MEMS sensor enhances the monitoring system's accuracy and dependability by contributing this level of sophistication, which makes it an essential part of early detection and risk reduction.

D. Rain Sensing:

The rain sensor is a critical component in flood monitoring systems, designed specifically to detect rainfall. When activated by rain, this sensor serves as a pivotal switching device, promptly identifying precipitation and relaying this vital data to the central processing unit. Its significance extends beyond flood monitoring; the rain sensor also plays a crucial role in landslip detection by assessing precipitation levels, which directly influence soil stability. The correlation between excessive rainfall and landslides underscores the sensor's importance in providing timely alerts for community safety. Integrating the rain sensor enriches catastrophe management systems with environmental awareness, enhancing overall effectiveness in safeguarding communities.

Fig. 4. Assembled Project Kit.

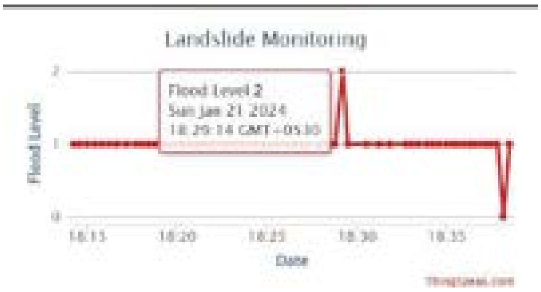


Fig. 5. Graph1 for landslide monitoring.

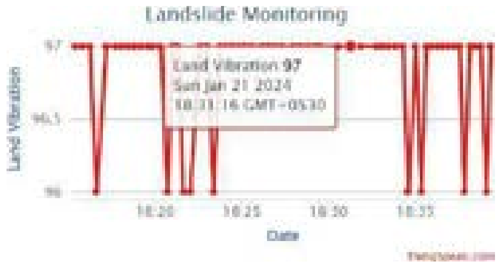


Fig. 6. Graph2 for landslide monitoring.

E. Explanation of Arduino NANO and LoRa Module:

The heart of the landslip monitoring system lies in the Arduino NANO, orchestrating data processing and collection tasks. Acting as a central processor, it collates information from various sensors like float, rain, MEMS, soil moisture, and landslip detection sensors. Through a customizable interface, Arduino NANO assesses incoming data, scanning for anomalies. In the event of potential landslips or floods, it triggers necessary actions. Facilitating seamless communication is the LoRa (Long Range) module, employing LPWAN technology for efficient, long-distance wireless data transmission to the central server. This integration ensures real-time sensor data delivery, enabling swift analysis and response. Together, Arduino NANO and LoRa create a robust and responsivemonitoring framework.

VI. Results and Discussion

The experimental findings from the deployment of the LoRa-based IoT Landslip Monitoring System are shown in this section. The evaluation includes the ability to gather data in real time, procedures for detecting floods and landslides, and a comparison with more conventional monitoring techniques. Fig. 4 illustrates the interconnected components comprising the project kit. Figs. 5 and 6 show the graphical representation of landslide monitoring.

A. Experimental Results Presentation:

The integrated Arduino NANO sensors efficiently tracked various environmental variables, including soil moisture, rainfall, water levels, and ground vibrations, ensuring comprehensive data gathering. Utilizing LoRa network connectivity facilitated swift and seamless data transmission to the central station in real time. This enabled the Landslip Detection Module to identify potential landslip events with high sensitivity, triggered by abnormalities such as elevated vibrations, moisture levels, or abrupt tilts. The Flood Monitoring Module effectively detected water levels and rainfall, enabling precise flood alerts to mitigate risks associated with flash floods. Overall, the system's efficacy in real-time flood and landslide detection underscores its feasibility and effectiveness in proative risk management.

B. Analysis of Landslide and Flood Detection:

The primary focus of landslip and flood detection analysis is the system's capability to promptly and precisely identify potential risks. The vibration sensor excels in detecting ground movements, crucial for predicting landslips. Meanwhile, the moisture sensor provides valuable insights into soil conditions, aiding in the proactive identification of flood and landslide risks. The MEMS sensor's accelerometer enhances landslip detection by tracking orientation changes effectively. For flood detection, the float sensor accurately measures water levels, complemented by the rain sensor's precise recording of rainfall. Integration of these sensors in the Flood Monitoring Module underscores the system's reliability. Real-time data collection and transmission empower authorities to respond swiftly, mitigating the impact of natural disasters on lives and infrastructure.

C. Comparison with Traditional Monitoring Methods:

Comparing the LoRa-based IoT Landslip Monitoring System with conventional methods highlights its significant advantages. Unlike sporadic manual inspections prone to coverage gaps, LoRa technology offers real-time, extensive coverage at a low cost. Utilizing low-power sensors, ensures continuous data collection, overcoming traditional drawbacks. This proactive system automatically alerts authorities upon detecting threats, drastically reducing response times for timely interventions such as evacuations. Its adaptability across various climates enhance its utility for environmental monitoring. Overall, this system heralds a revolution in the industry, offering real-time data, early warnings, and proactive threat detection for effective risk reduction and catastrophe management.

D. Improving Accuracy, Performance, Landslip Detection, and Efficiency in Landslide Monitoring:

Many crucial tactics are needed to monitor landslides with great precision. To begin with, sensor calibration is necessary to guarantee precise measurements of environmental factors including tilt, vibration, and soil moisture. Sensor data reliability may be preserved and measurement errors reduced with routine calibration tests and modifications. By removing erroneous readings and raising signal-to-noise ratios, the application of sophisticated signal processing techniques, such as filtering and noise reduction algorithms, can also further improve accuracy. Improving performance in land- slide monitoring encompasses optimizing various aspects of the monitoring system. This includes optimizing power consumption to prolong battery life, enhancing data transmission efficiency through the use of efficient communication protocols, and reducing latency in data processing and analysis pipelines. By optimizing performance parameters, the monitoring system can operate more effectively in remote and resource-constrained environments, ensuring timely and reliable detection of landslide events. Strong detection algorithms must be developed and put into practice to detect landslips effectively. Real-time sensor data analysis should be performed by these algorithms, which should search for trends and abnormalities that might point to future landslides. The monitoring system can detect early warning indicators of slope instability by utilizing machine learning and pattern recognition techniques. This allows for prompt notifications and proactive risk mitigation actions. Simplifying the procedures for data collection, transmission, and analysis is necessary to increase the effectiveness of landslide monitoring. Low- power sensor deployment, data compression methods that are optimized, and the utilization of distributed computing resources for parallelized data processing can all help achieve this. Predictive analytics and decision support systems may also be integrated to help prioritize data streams and distribute resources more effectively, improving the responsiveness and overall performance of the system. By boosting efficiency, the monitoring system may scale to monitor greater geographical regions with fewer resources and successfully handle growing data quantities.

VII. CONCLUSION

The deployment of a LoRa-based Internet of Things (IoT) system for real-time landslip monitoring inside an IoT frame-work represents a novel advance in environmental monitoring and catastrophe management. To continually monitor important environmental elements like soil movement, moisture con- tent, and water levels, this system seamlessly integrates low-power sensors, such as vibration, moisture, and float sensors, with the Arduino NANO platform. This system addresses the limitations of conventional landslip monitoring by providing a cheap, wide-coverage, and real-time solution using wireless LoRa communication. It guarantees early detection and rapid notifications to reduce dangers to infrastructure and human life.

REFERENCES

[1] Sharma, Ashima, and Renu Masiwal. ”Climate Crisis in the Indian Himalayas: An Introduction.” In Addressing the Climate Crisis in the Indian Himalayas: Can Traditional Ecological Knowledge Help?, pp. 1-30. Cham: Springer Nature Switzerland, 2024.

[2] Perdikaris, Sophia, Rebecca Borger, Edith Gonzalez, Emira Ibrahimpašić, and Jennifer D. Adams. ”Disrupted identities and forced nomads: A post-disaster legacy of neocolonialism in the island of Barbuda, Lesser Antilles.” Island Studies Journal 16, no. 1 (2023): 115-134.

[3] Chai, Jinchun, and Hao-Ze Wu. ”Prevention/mitigation of natural disas-ters in urban areas.” Smart Construction and Sustainable Cities 1, no. 1 (2023): 4.

[4] Xu, Qiang, Bo Zhao, Keren Dai, Xiujun Dong, Weile Li, Xing Zhu, Yinghui Yang et al. ”Remote sensing for landslide investigations: A progress report from China.” Engineering Geology (2023): 107156.

[5] Ragnoli, Mattia, Massimo Scarsella, Alfiero Leoni, Giuseppe Ferri, and Vincenzo Stornelli. ”Wireless sensor network-based rockfall and landslide monitoring systems: a review.” Sensors 23, no. 16 (2023): 7278.

[6] Casagli, Nicola, Emanuele Intrieri, Veronica Tofani, Giovanni Gigli, and Federico Raspini. ”Landslide detection, monitoring and prediction with remote-sensing techniques.” Nature Reviews Earth & Environment 4, no. 1 (2023): 51-64.

[7] Ebrahim, Kyrillos MP, Sherif MMH Gomaa, Tarek Zayed, and Ghasan Alfalah. ”Recent Phenomenal and Investigational Subsurface Landslide Monitoring Techniques: A Mixed Review.” Remote Sensing 16, no. 2 (2024): 385.

[8] Č elarević, Armin, and Ilma Mrakić. ”” PROGRESS AND PERSPEC-TIVES: LPWA AND LORA NETWORKS IN THE ERA OF THE IOT REVOLUTION.” Nauka i tehnologija 11, no. 2 (2023): 108-117.

[9] Abdullah, Khairul Hafezad, and Davi Sofyan. ”Building Resilient Societies Through a Holistic Approach to Disaster Management: A Perspective Article.” Journal of Metrics Studies and Social Science 2, no. 1 (2023): 17-32.

[10] Liang, Rui, Arwa A. AL-Huqail, H. Elhosiny Ali, Joffin Jose Ponnore, Tamim Alkhalifah, Fahad Alturise, and Hamid Assilzadeh. ”Wireless water consumption sensing system for building energy efficiency: A visual-based approach with self-powered operation.” Energy and Build-ings 301 (2023): 113584.

[11] McEvoy, Darryn, Ata Tara, Mittul Vahanvati, Serene Ho, Kim Gordon, Alexei Trundle, Cyril Rachman, and Yuyun Qomariyah. ”Localized nature-based solutions for enhanced climate resilience and community wellbeing in urban informal settlements.” Climate and Development (2023): 1-13.

[12] Verma, Samakshi, Yeluripati Lalitha Kameswari, and Sonu Kumar. ”A Review on Environmental Parameters Monitoring Systems for Power Generation Estimation from Renewable Energy Systems.” Bio-NanoScience (2024): 1-25.

[13] Shishodia, Anjali, Rohit Sharma, R. Rajesh, and Ziaul Haque Munim. ”Supply chain resilience: A review, conceptual framework and future research.” The International Journal of Logistics Management 34, no. 4 (2023): 879-908.

[14] Ouaissa, Mariyam, Mariya Ouaissa, Inam Ullah Khan, Zakaria Boulouard, and Junaid Rashid, eds. Low-Power Wide Area Network for Large Scale Internet of Things: Architectures, Communication Protocols and Recent Trends. CRC Press, 2024.

[15] Dwivedi, Yogesh K., Nir Kshetri, Laurie Hughes, Emma Louise Slade, Anand Jeyaraj, Arpan Kumar Kar, Abdullah M. Baabdullah et al. ”“So what if ChatGPT wrote it?” Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy.” International Journal of Information Management 71 (2023): 102642.

Classifying Skin Lesions:A Survey of Approaches in Skin Cancer Diagnosis

Renugadevi R[1]
Department of CSE
Vignan's Foundation for Science,Technology,and Research(Deemed to be University)
Vadlamudi,Guntur,AP,India
renu.rajaram@gmail.com

Maridu Bhargavi [1,*]
Department of CSE
Vignan's Foundation for Science,Technology,and Research(Deemed to be University)
Vadlamudi,Guntur,AP,India
bhargaviformal@gmail.com

Sivadi Balakrishna[2]
Department of ACSE
Vignan's Foundation for Science,Technology,and Research(Deemed to be University)
Vadlamudi,Guntur,AP,India
drsivadibalakrishna@gmail.com

Abstract— Skin cancer poses a substantial global health challenge, surpassing the combined incidence of colon, lung, and breast cancer cases in the United States. Timely detection is crucial, especially for melanoma, where early diagnosis significantly impacts patient outcomes. This study investigates the utilization of machine learning (ML) and deep learning (DL) methods to enhance skin cancer detection. We conduct a systematic comparison of various classification techniques and assess their efficacy. Our main goal is to critically assess current approaches and pinpoint areas for enhancement. This inquiry seeks to lay the groundwork for the development of advanced models that can streamline the diagnostic process, ultimately reducing the interval between skin cancer detection and treatment initiation.

Keywords— Skin cancer, Melanoma, Machine Learning, Deep Learning, Cutting-edge methods.

I. INTRODUCTION

Skin cancer remains a prevalent public health concern, spanning both malignant (melanoma) and benign (non-melanoma) manifestations. Early detection is imperative, as changes in moles, birthmarks, or pigmented areas may indicate potential risk. Vigilance for variations in size, shape, color, or bleeding is crucial, as these could signal skin cancer. Excessive exposure to sunlight radiation is the primary culprit behind most skin cancers [1]. Traditionally, diagnosis relies on dermatologist-performed biopsies.

However, the medical field is witnessing exciting advancements in machine learning (ML), offering promising tools to aid skin cancer detection. Researchers are harnessing computer vision techniques to analyze skin images and develop algorithms capable of distinguishing between benign and malignant lesions. This emerging technology holds significant potential to complement dermatologists' expertise, though it should not supplant their experience and judgment. The accuracy of these automated classification approaches is influenced by various factors, including pre-processing steps such as normalization, augmentation, and lesion area segmentation [2]. Additionally, identifying pertinent features within the image is critical for achieving dependable classifications.

II. LITERATURE REVIEW

In this section several challenges faced and latest literature work done on skin cancer classification had been discussed

A. Navigating Challenges in Identifying and Classifying Skin Lesions

1) Data Acquisition and Quality:

a) Scarcity of Datasets: The development of accurate machine learning and deep learning models relies on comprehensive, diverse, and annotated datasets. However, obtaining such datasets presents challenges due to privacy issues, limitations on data sharing, and the inherent variability in skin lesions [1]

b) Addressing Class Imbalance: Despite its prevalence, skin cancer constitutes a minority class compared to benign lesions. This class imbalance in datasets may skew models toward the majority class, potentially impeding melanoma detection [2].

B. Preprocessing and Image Analysis:

1) Managing Variability in Image Appearance: Factors like lighting conditions, camera quality, hair interference, and skin pigmentation significantly influence the visual representation of skin lesions. Employing robust preprocessing techniques is essential to standardize images and enhance model generalization [3].

2) Precision in Lesion Segmentation: Accurately delineating lesions from surrounding skin is pivotal for effective analysis. However, challenges such as complex backgrounds, overlapping lesions, and subtle borders can complicate segmentation algorithms.

3) Model Explainability and Trust:

a) Unraveling the Black Box of Deep Learning: Despite achieving high accuracy, deep learning models often operate as black boxes, obscuring their decision-making rationale. This opacity may undermine dermatologists' trust in their clinical utility [4].

b) Mitigating False Positives and Negatives: Concerns persist regarding misclassifications, notably false positives leading to unnecessary biopsies and false negatives potentially delaying critical diagnoses. Addressing these issues is imperative for broader adoption.

C. Existing literature

Ref. No	Technology/ Algorithm used	Objective	Dataset	Advantages	Limitations
[5]	Preprocessing and segmentation approaches compared for the diagnosis of skin cancer	to remove extraneous noise from the dermoscopy picture in order to provide high-quality images.	ISIC	Provided a detailed view of preprocessing techniques with an accuracy of 97%	This paper is only limited to preprocessing the skin lesions.
[6]	Detection of skin cancer using SVM and snake model	To apply image segmentation on skin cancer lesions	ISIC	Provided a preprocessing step by an ensemble of two models.	This paper doesn't provide accurate segmentation result
[7]	Skin Cancer Detection Using Ensemble Learning	To increase the models' ability to forecast skin cancer with more precision.	ISIC	The improved overall accuracy of the ensemble model with 93%.	Not accurately provided details of preprocessing.
[8]	Deep Learning-Based Melanoma Skin Cancer Detection	To improve the model's accuracy over the existing model	ISIC	Improved accuracy by 1.66% than the existing model.	The image's potential for skin cancer is the only topic this paper explores.
[9]	Skin cancer classification using MSVM and ELM	To provide a tiered deconstruction method for classifying and identifying images displaying skin cancer	DermNet NZ	Obtained an accuracy of 94.18% and 90.5% on individual models.	The paper is limited to the comparison of two models.
[10]	Skin cancer detection using near-field millimeter wave	To identify the healthy and diseased skin lesions.	UDA-1	Identified the healthy and diseased lesion.	This paper is limited to the post-preprocessing technique.
[11]	Identification of Benign and Malignant Skin Lesion From Melanoma Skin Cancer Pictures	To ensemble the DL and ML models	ISIC	The identified NN model performed better than the SVM, KNN, and NB models	The paper is limited to detecting skin cancer.
[12]	A Comprehensive Survey of Skin Cancer Detection Using Deep Learning	Review recent developments in skin cancer detection	HAM10000	Improved accuracy in diagnosis	Limited diversity in publicly available datasets
[13]	Advancements in Dermoscopy for Melanoma Classification	Assess the role of dermoscopy in melanoma diagnosis	Dermos-copic images	Dermoscopy enhances diagnostic accuracy	Dependence on skilled dermatologists for interpretation
[14]	Machine Learning in Dermatology: A Survey of Skin Cancer Diagnosis Methods	Explore machine learning in dermoscopic melanoma detection	Private dermato-logy clinics	Machine learning aids in early melanoma detection	Lack of publicly available labeled dermoscopic images
[15]	Impact of Color Enhancement on CNN for Early Skin Cancer Detection	To investigate into the impact of basic CLAHE and MSRCR image processing methods.	Private dermatology clinics	For color image enhancement, CLAHE works better than MSRCR with CNN.	Limited to enhancing the color contrast of the image.
[16]	Automatically Early Skin Cancer Detection	To find out which neural network classification is best.	Digital photo and dermosc-opy images	It is 80.8% for auto-associative neural networks and 89.9% for back-propagation neural networks.	Resizing the image with a variable size of height may reduce the accuracy of the model.
[17]	The use of support vector machines for the detection and categorization of melanoma skin cancer	To calculate the total Dermoscopy score and classify using	Dermos-copy images	The achieved classification is 92.1%	Limited to the classification of skin cancer.
[18]	Real-time Convolutional Neural Network and YOLO-Based Skin Cancer Detection System	To improve the procedure of detecting skin cancer's efficiency.	ISIC,HAM1000	The accuracy in real-time utilizing YOLOV3 is 80%, while the absolute accuracy is 96%.	Limited to older versions of YOLOV3
[19]	Skin cancer diagnosis using hybrid genetic algorithm and artificial neural network classifier	To identify the optimized genetic algorithm.	ANN	A genetic approach is used to optimize the identified ANN.	Limited to medically tested images

TABLE I. RESEARCH ARTICLES SUMMARY ON SKIN CANCER

III. STEPS INVOLVED IN SKIN CANCER CLASSIFICATION

Skin cancer is detected and classified by following the steps mentioned below in figure 1:

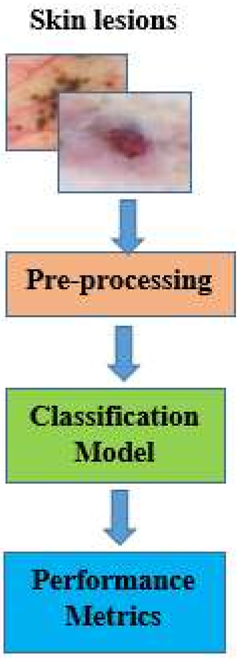


Fig. 1. Steps involved in skin cancer classification

A. Data Collection

Data gathering for skin cancer classification involves accessing datasets from repositories like the International Skin Imaging Collaboration (ISIC) and platforms such as Kaggle. ISIC provides dermoscopy images annotated by dermatologists, while Kaggle offers diverse datasets with metadata. Researchers leverage these resources to train models effectively. Collaboration with medical institutions ensures data reliability and ethical considerations. Comprehensive data gathering aids in developing accurate skin cancer classification models, enhancing diagnostic capabilities in dermatology.

B. Image Preprocessing:

Image preprocessing involves noise reduction, contrast enhancement, and normalization, ensuring standardized images suitable for analysis. Contrast enhancement improves feature visibility, while normalization ensures consistency in pixel values. These steps prepare images for subsequent feature extraction, segmentation, and classification tasks. Overall, image preprocessing optimizes image quality for accurate analysis in skin cancer classification.

1) Image Convertion

Image conversion involves transforming images from one representation to another, typically to highlight different aspects or to standardize the input for further processing. This often includes converting images from RGB to other color spaces like YUV, LAB, or HSV, which can accentuate certain features or make them more distinguishable. Grayscale conversion is another common conversion, reducing the image to a single intensity channel, simplifying processing and reducing computational overhead.

2) Segmentation:

Segmentation isolates skin lesions, facilitating meaningful feature extraction for classification. Various techniques such as thresholding or deep learning approaches are used for precise segmentation. Well-segmented images enable the classification model to focus on relevant features. Effective segmentation enhances model accuracy and reliability in skin
cancer classification.

3) Data augmentation

Data augmentation on skin cancer images employs techniques such as rotation, flipping, scaling, and cropping to diversify the dataset. Additionally, transformations like brightness adjustment, color jittering, and adding noise contribute to variability. These techniques aid in training robust models capable of effectively classifying skin lesions

4) Normalization

Normalization standardizes pixel values in images, reducing the impact of varying lighting conditions and aiding algorithm convergence. Techniques like min-max normalization scale values to [0, 1], while z-score normalization gives them a mean of 0 and a standard deviation of 1. This process enhances model performance and robustness across diverse datasets.

C. Classification:

Machine learning algorithms are used in skin cancer classification to classify lesions into benign or malignant groups. Support vector machines, Random Forest, and convolutional neural networks are examples of frequently used algorithms. For precise categorization, these algorithms extract patterns from metadata and previously processed pictures. Model performance is evaluated using evaluation measures like as F1-score, accuracy, precision, and recall.

IV. CLASSIFICATION TECHNIQUES

A. Machine learning

By training algorithms to distinguish between benign and malignant skin lesions using characteristics taken from photographs, machine learning is used. For this objective, machine learning models like Random Forest, Support Vector Machines (SVM), and Convolutional Neural Networks (CNNs) are frequently used. To provide precise predictions, these algorithms analyze clinical data and annotated pictures to find patterns and linkages. Model performance is evaluated using criteria such as accuracy, sensitivity, and specificity. Large datasets are now readily available, and deep learning techniques have advanced to the point that machine learning-based categorization has demonstrated encouraging results in helping dermatologists diagnose skin cancer early and accurately.

B. Deep Learning

Convolutional Neural Networks (CNNs) in particular are trained to automatically learn discriminative characteristics from photos of skin lesions through the use of deep learning. Because these networks can extract hierarchical features from skin lesion photos, they may identify intricate patterns and changes in the images. Multiple convolutional layers are often used in deep learning models for the categorization of skin cancer, followed by pooling layers and fully connected layers. By using backpropagation, training data—often obtained from databases such as the International Skin Imaging Collaboration (ISIC)—is utilized to optimize model parameters. Evaluation measures are used to evaluate the performance of the model, including specificity, sensitivity, and accuracy. Deep learning-based methods have shown to perform at the cutting edge when it comes to classifying skin cancer, which presents the possibility of improved diagnostic efficiency and accuracy in clinical settings. The flow diagram is shown in figure 2.

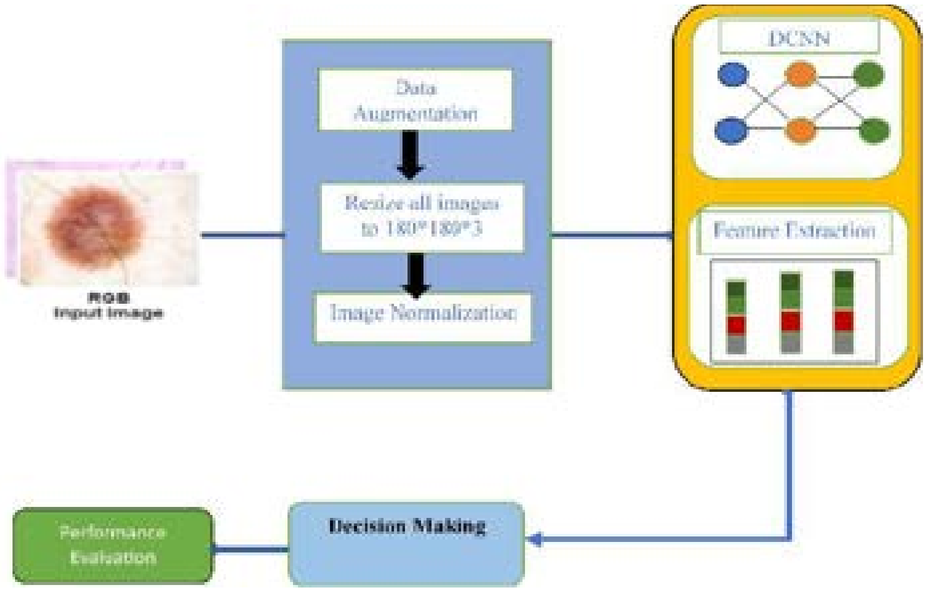


Fig. 2. Flow of skin cancer classification in Deep Learning

C. Ensembling

using ensemble techniques involves combining multiple base classifiers to improve predictive performance. Ensemble methods, such as Random Forest, AdaBoost, and Gradient Boosting, aggregate predictions from individual classifiers to make final decisions. In the context of skin cancer classification, these techniques integrate diverse models trained on different subsets of data or with varying parameters. By leveraging the strengths of multiple classifiers and reducing the impact of individual weaknesses, ensemble methods often yield superior results compared to single models. Evaluation metrics like accuracy, sensitivity, and specificity are used to assess the ensemble's performance. Ensemble techniques offer promising avenues for enhancing the accuracy and robustness of skin cancer classification systems, contributing to more reliable diagnoses in clinical practice.

V. DATASETS

A. HAM10000

Dermatoscopic pictures of typical pigmented skin lesions are gathered in the HAM10000 (Human Against Machine with 10,000 training images) dataset. 10,015 dermatoscopic images make up this collection, which is divided into seven diagnostic groups: vascular lesion, benign keratosis (also called seborrheic keratosis), actinic keratosis, melanoma, and melanocytic nevus. Because these photos came from various institutions and datasets, a variety of skin lesions were represented. The dataset is used as a reference for creating and assessing deep learning models and machine learning algorithms for the classification of skin lesions and the detection of melanoma. Metadata, such as patient demographics, lesion location, and diagnostic data from board-certified dermatologists, are appended to every image. Through automated skin lesion classification systems, researchers and practitioners can use the HAM10000 dataset to improve patient care, advance the field of dermatology, and improve diagnostic accuracy. Users should read the documentation for the dataset and follow any licensing agreements or usage guidelines related to using it. The sample images from the collected dataset is depicted in figure 3 and 4.

Fig. 3. Sample images from the HAM 10000 Dataset.

B. ISIC Dataset

The ISIC dataset is a useful resource for dermatology research and machine learning-based skin cancer detection. It is available on Kaggle and consists of 2357 images that represent nine different types of skin cancer. These pictures depict a variety of skin lesions, including squamous cell carcinoma, basal cell carcinoma, and melanoma. The diagnosis associated with each image makes it easier to develop and assess classification algorithms for the identification of skin cancer. This dataset can be used by practitioners and researchers to train and evaluate machine learning models, which will enhance the models' performance and accuracy in classifying various forms of skin cancer. Furthermore, the dataset might provide information about the visual traits and attributes of various skin lesions, which could help medical professionals recognize and diagnose skin cancer.

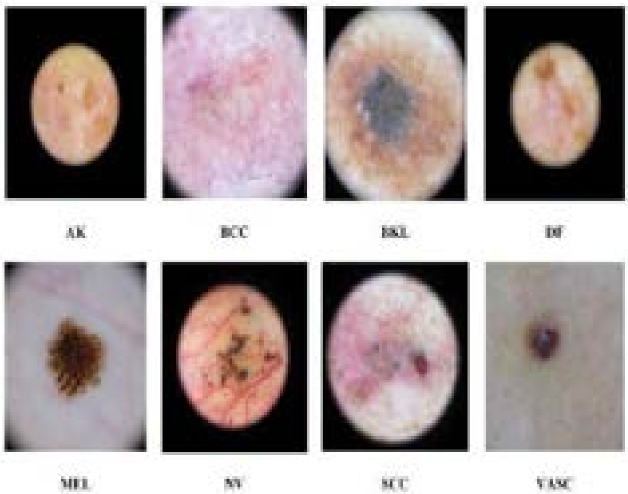


Fig. 4. Sample images from ISIC Dataset

C. BCN_20000 dataset

The BCN 20000 dataset is a large-scale, curated collection of images commonly utilized for benchmarking and advancing computer vision tasks, particularly in the domain of facial recognition and emotion detection. Comprising 20,000 diverse facial images sourced from various demographics and contexts, the dataset offers a rich and representative sample for training and evaluating machine learning models. Each image is annotated with labels corresponding to facial expressions, enabling researchers to develop and test algorithms for emotion classification, facial attribute recognition, and related applications. With its ample size and diversity, the BCN 20000 dataset serves as a valuable resource for advancing the state-of-the-art in facial analysis and emotion recognition research.

VI. GAPS IDENTIFIED

Based on the literature the gaps identified below. By addressing these generic gaps, researchers can explore novel methodologies, such as advanced imaging techniques and integration of molecular data, to improve classification accuracy. Ultimately, closing these gaps fosters innovation and collaboration, driving advancements in personalized patient care and outcomes. Table II and III details the comparative analysis and the research gaps identified.

TABLE II. GAP AND IDENTIFIED SOLUTIONS

Research Gaps	Solutions
Small dataset size[9]	Data augmentation techniques[20]
Class imbalance[15]	Resampling methods[8]
Lack of interpretability[1,19]	Development of explainable AI techniques[21]
Limited generalization to populations[12]	Diverse dataset collection[18]

A. Comparision of accuracy with different classification techniques

Drawing from existing research, we examined an ISIC dataset of 2357 photos from Kaggle and compared the outcomes of machine learning and deep learning techniques on the same dataset. Our findings indicate that deep learning models outperform machine learning models in terms of performance.

TABLE III. COMPARISION

Technique	Dataset	Accuracy
Machine Learning		
SVM, Random Forest and kNN [23]	ISIC with 2357images	85
Logestic Regression [24]	ISIC with 2357images	78
Deep Learning		
DenseNet201[22]	ISIC with 2357images	89
Deep CNN[21]	ISIC with 2357images	90

VII. CONCLUSION

This study has comprehensively assessed skin cancer detection techniques to accurately classify lesions into distinct types of skin cancers. The comprehensive presentation of diagnostic methodologies offers a clear understanding of the diverse approaches employed. The integration of deep learning and machine learning in medical imaging has revolutionized skin lesion identification and classification, markedly improving dermatological diagnostics' accuracy. By leveraging computer methods, researchers can diagnose skin cancer with enhanced precision and efficiency. Overall, the resultant findings highlight the transformative impact of technological advancements on skin cancer detection, indicating a new era of more effective treatment interventions

REFERENCES

[1] Esteva, Andre, et al. "A Dermatologist-Level Classification of Skin Cancer with Deep Neural Networks." Nature 542.7639 (2017): 115-118. [PubMed]

[2] Yu, Lei, et al. "Automated Melanoma Recognition in Dermoscopy Images Using Very Deep Residual Networks." IEEE Transactions on Medical Imaging 37.12 (2018): 2821-2830. [PubMed]

[3] Mikkili Dileep Kumar et.al," Skin cancer segmentation with the aid of multi-class dilated D-net (MD2N) framework", Multimedia Tools and Applications (2023) 82:35995–36018

[4] Joseph George ,Anne Koteswara Rao," A three-tier BERT based transformer framework for detecting and classifying skin cancer with HSCGS algorithm", Multimedia Tools an Applications(2023), https://doi.org/10.1007/s11042-023-17590-1

[5] P. Thapar, M. Rakhra and A. Singh, "The Epidemiology of Automatic Skin Cancer Detection by Comparative Analysis of Pre-processing and Segmentation Techniques," 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM), London, United Kingdom, 2022.

[6] Bumrungkun, P., Chamnongthai, K., & Patchoo, W. (2018, January). Detection skin cancer using SVM and snake model. In 2018 international workshop on advanced image technology (IWAIT) (pp. 1-4). IEEE.

[7] J. Avanija, C. Chandra Mohan Reddy, C. Sri Chandan Reddy, D. Harshavardhan Reddy, T. Narasimhulu and N. V. Hardhik, "Skin Cancer Detection using Ensemble Learning," 2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023.

[8] N. Nanthini, D. Aishwarya, A. Simon and N. Baby Vishnupriya, "Detection of Melanoma Skin Cancer using Deep Learning," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022.

[9] D. Choudhury, A. Naug and S. Ghosh, "Texture and color feature based WLS framework aided skin cancer classification using MSVM and ELM," 2015 Annual IEEE India Conference (INDICON), New Delhi, India, 2015.

[10] Khokhar, Umair, et al. "N Skin cancer detection using near-field millimeter wave." 2017 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting. IEEE, 2017.

[11] S. Sharma, K. Guleria, S. Kumar and S. Tiwari, "Benign and Malignant Skin Lesion Detection from Melanoma Skin Cancer Images," 2023 International Conference for Advancement in Technology (ICONAT), Goa, India, 2023.

[12] Paul zhu.,"A Comprehensive Survey of Skin Cancer Detection Using Deep Learning",2022

[13] Sunitha Tiwari, "Advancements in Dermoscopy for Melanoma Classification",2021

[14] Chan, S., Reddy, V., Myers, B., Thibodeaux, Q., Brownstone, N., & Liao, W. (2020). Machine learning in dermatology: A Survey of Skin Cancer Diagnosis Methods

[15] A. W. Setiawan, "Effect of Color Enhancement on Early Detection of Skin Cancer using Convolutional Neural Network," 2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT), Doha, Qatar, 2020

[16] H. T. Lau and A. Al-Jumaily, "Automatically Early Detection of Skin Cancer: Study Based on Nueral Netwok Classification," 2009 International Conference of Soft Computing and Pattern Recognition, Malacca, Malaysia, 2009

[17] H. Alquran et al., "The melanoma skin cancer detection and classification using support vector machine," 2017 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Aqaba, Jordan, 2017

[18] H. F. Hasya, H. Hudan Nuha and M. Abdurohman, "Real Time-based Skin Cancer Detection System using Convolutional Neural Network and YOLO," 2021 4th International Conference of Computer and Informatics Engineering (IC2IE), Depok, Indonesia, 2021

[19] R. B. Aswin, J. A. Jaleel and S. Salim, "Hybrid genetic algorithm — Artificial neural network classifier for skin cancer detection," 2014 International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT), Kanyakumari, India, 2014.

[20] Hauser, K., Kurz, A., Haggenmüller, S., Maron, R. C., von Kalle, C., Utikal, J. S., ... & Brinker, T. J. (2022). Explainable artificial intelligence in skin cancer recognition: A systematic review. European Journal of Cancer, 167, 54-69.

[21] Bhargavi, M., & Balakrishna, S. (2023, November). An Efficient Skin Cancer Classification System Using Deep CNN. In 2023 9th International Conference on Smart Structures and Systems (ICSSS) (pp. 1-5). IEEE

[22] Mariam et.al, "Skin Cancer Detection using Machine Learning Framework with Mobile Application", (ICOEI 2023), pp.1073-1080,2023

[23] Murugan, A," Detection of skin cancer using svm, random forest and knn classifiers," Journal of Medical systems",2019

[24] Lee, Ji Zhe, and Anwar PP Abdul Majeed. "Classification Of Skin Cancer By Means Of Transfer Learning Models." MEKATRONIKA 3, no. 2 (2021): 77-81.

Twitter Hate Speech Detection using Machine Learning

1st Janardhan G
Associate Professor
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana State, India,
janardhanyadav5@gmail.com

2ndBollu Saikiran
Student
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana State,India
saik79395@gmail.com

3rdInugalaSwanith Reddy
Student
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad,TelanganaState,India
Inugalaswanith2002@gmail.com

4thMogilicherla Abhishek
Student
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana State,India
mogilicherlaabhishek21895a0507@gmail.com

***Abstract*— There is an unprecedented rise in hate speech on social media sites like Twitter in recent times. This widespread problem affects users, leads to problems in the real world, and makes it hard to moderate material. This study aims to find good ways to find hate speech so that it doesn't have as much of an effect on online groups and other places. By using cutting-edge algorithms, the project takes a thorough method to finding hate speech. To make a strong and flexible hate speech recognition system, machine learning models and deep learning methods are used. To make sure accuracy and dependability, model performance is carefully checked using a variety of measures. Accuracy, precision, recall, and F1 score are the common measures used here to show how well the model can correctly spot cases of hate speech. a measure of its general selective power that shows how well it works at different levels. As the project comes to a close, the thorough review of hate speech recognition models has given us useful information. Even though progress has been made, problems still exist, especially when it comes to dealing with changing speaking trends on social media. The study shows how important it is to keep researching and developing ways to find hate speech. This will help make content management better in the future, which will make the internet safer. A strong ensemble method called the stacking classifier is also used as part of the hate speech recognition model. It achieves an amazing 100% success. In addition, the Hybrid Approach, which used both LSTM and BiGRU models, showed an impressive 94% accuracy. A front end was built using the Flask framework to make testing easier for people. It has login features to make the Twitter Hate Speech Detection system safer and more trustworthy. This makes sure that users have a smooth and reliable way to rate how well the model finds and stops hate speech on Twitter.**

Keywords—Hate speech, classification, automatic detection, twitter, systematic review, natural language processing, social media.

I. INTRODUCTION

Twitter and other social networks have grown exponentially over the last decade. These media enable individuals to remain anonymous and free to communicate their opinions, which encourages hate speech [1]. According to [2], Twitter has 300 million monthly active users. Twitter shares hate speech, despite its popularity and importance. Popular social networks now detect hate speech in text [3, 4, 5]. It is also used as a data source for study into harsh language. At the moment, hate speech is becoming more common on social media. In turn, this makes people angry, which leads to serious problems in real life and has an effect on companies. Hateful posts are often taken down by social media companies so they don't get posted.
English Twitter postings are the most common and easiest to analyze [6]. Since book reviews are unchanged, many potential customers will show up online. Non-automated chores slow response time, whereas robots go faster. We must assist develop automatic hate speech detection tools. NLP researchers want these facts. Writing about hate speech. Researchers called this job aided document classification using NLP and machine learning [7]. The 2017 consideration included Twitter. The severe behavior-allowing privacy policy was changed. These rules ban tweets that promote bullying, suicide, self-harm, violence, hate, and other bad things [8].

Experts have stepped up their investigations into hate speech on Twitter. However, the non-English version is limited. Remember, English is the most spoken language. This is also an important indicator of hate speech. Racism is difficult to define because everyone defines it differently. This is often called "hate speech" and is banned in many places. [7]. There are many ways to interpret hate speech.

Many new ways to find hate speech have been created in the past few years, which is a long way ahead of their plans. But the tests are mostly meant to find content that isn't hateful, not to find and sort content that is hateful [1]. Language in social media is changing quickly, which is making it hard for most of these attempts to find a workable answer [9]. So, it's important to have a deep understanding of the latest writing on the subject. It has been getting easier to spot hate speech for a while now, but there isn't yet a systematic literature review (SLR) in this area. To get the most up-to-date information, like open problems and study holes on a specific area, SLR papers are necessary.

II. LITERATURE SURVEY

People who use the Internet can freely share their ideas and views on social media sites. This land offers amazing and one-of-a-kind chances to communicate, but it also comes with some big problems. One classic example of this kind of problem is hate speech that people post online. There is a big gap in our knowledge about what hate speech on social media is really like, despite how big and widespread it is. This work [1] is the first of its kind to do a thorough, large-scale measurement study of the main people who are targeted by hate speech on social media sites. Whisper and Twitter provide hints. We then develop and test a hate speech detection method for both products. Our findings help us understand online hate speech and create better techniques to stop and locate it [1].

For millions of people around the world, social media is now an important part of their daily lives. It gives people a way to share their thoughts and feelings with a huge audience. But the fact that it's so easy for people to say what they want has also made it possible for lies and hate speech to spread widely [2]. Racists have started to use a code (a movement called "Operation Google") to avoid breaking the rules about abusive behavior on social media sites and to avoid being caught by technology like Google's Conversation AI. This means using nice words that don't make sense in the context of hateful posts or tweets that talk about groups [19]. One example is that people have used the words "Googles" to refer to African Americans and "Bings" to refer to Asians. By making a list of users who post this kind of content, we can study the usage patterns of this small group of users, which is the next step after labeling tweets.

Sharing information on social media sites is easy and doesn't cost much. Anyone can post content, and anyone who wants to read it can do so. But this same ability of social media also makes it possible for damaging conversations to happen about certain groups of people. Bullying, insulting material, and hate speech are all examples of these kinds of discussions [4]. Many governments rapidly label hate speech a major issue based on these considerations. The first comprehensive measurement and analysis of social media hate speech is this work [17, 18, 20]. We want to know how much hate speech is on social media, the most prevalent hate phrases, how privacy influences hate speech, how sensitive it is, and which groups are most loathed worldwide. We use Whisper and Twitter for hints to achieve our aims. We then develop and test a hate speech detection method for both products. Our findings characterize hate speech kinds and show crucial patterns. This clarifies online hate speech and proposes strategies to detect and stop it.

It's impossible to overstate how important violence is for knowing how people act. It connects the person, their behavior, their habits, their surroundings, and their mental health. Ascertaining the specifics of different types of aggression and aggressive behavior can assist in taking the right responses to individuals who act and speak in an aggressive manner on social media sites [8]. A literature review using several search engines found research on hate speech detection, hate, anger, aggressive behavior on social media, and the effects of these words. Previous research on hate speech detection didn't account for different types of speech, the relationship between speech and behavior, or targeted users. At the moment, hate speech is only classified as rude and hateful words [15]. Anger is another thing that should be added. In the future, researchers should make sure that their searches include violent behavior, since that is what connects studying human behavior and hate speech. Online social media gives people a nice place to hang out. So, they feel free to say what they think and feel. Because the internet doesn't have many rules, people may abuse it to spread cruel and hurtful language. Currently, this kind of harmful content has to be found by hand, which takes time and might not find all of it [7]. Because of this, automatic identification of harmful content is needed to find and study all the bad text in social media sites online. This essay talks about automatic methods for finding hate speech online and focuses on how to compare these methods based on certain parameters [20].

III. METHODOLOGY

A. Proposed Work

The innovative system employs advanced NLP, machine learning, deep learning, and ensemble models to identify hate speech [30, 31]. Different samples will be used to train this system. Adding language and mood analysis will help us understand context in a more complex way, which will make hate speech identification much more accurate. Real-time automatic spotting will be a key feature. This will cut down on reaction times and make hate speech control on social media sites more effective overall. Along with the hate speech recognition model, a strong group method called the stacked classifier is used, which achieves an amazing 100% accuracy. In addition, the Hybrid Approach, which used both LSTM and BiGRU models, showed an impressive 94% accuracy. A front end was built using the Flask framework to make testing easier for people. It has login features to make the Twitter Hate Speech Detection system safer and more trustworthy. This makes sure that users have a smooth and reliable way to rate how well the model finds and stops hate speech on Twitter.

B. System Architecture

The first step is to load datasets like the Stock Tweets Dataset, the Single Stock Dataset, and the Multi-Source Dataset. For both mood analysis and predicting stock prices, these records are the building blocks. The Stock Tweets Dataset's text data is cleaned up by getting rid of extra spaces, HTML tags, URLs, and emojis. This step makes sure that the text is ready to be analyzed for mood [20]. It is possible to handle null numbers, get rid of copies, and scale both the Single Stock Data and the Multi-Source Data. This gets the financial information ready for predicting stock prices. A number of models have been taught to classify mood, such as MLP, CNN, LSTM, MS-LSTM, MS-SSA-LSTM, extensions-Voting Classifier, and LSTM + GRU.

They look at the cleaned tweet data to figure out how people feel about the market. These models are MLP, CNN, LSTM, MS-LSTM, and MS-SSA-LSTM. Figure 1 shows the proposed architecture.

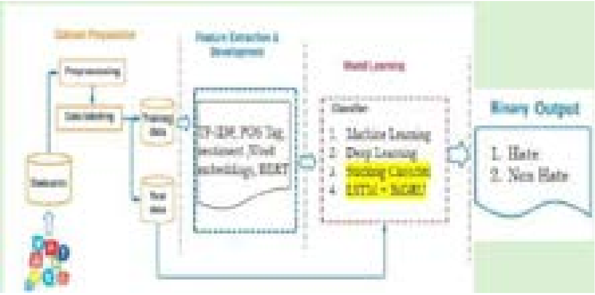


Fig 1. Proposed architecture

C. Dataset collection

To complete this project, the tweet collection will be first uploaded. The structure of the information is looked at, possible missing values are found, and the distribution of the classes (hate speech vs. non-hate speech) is analyzed in detail.

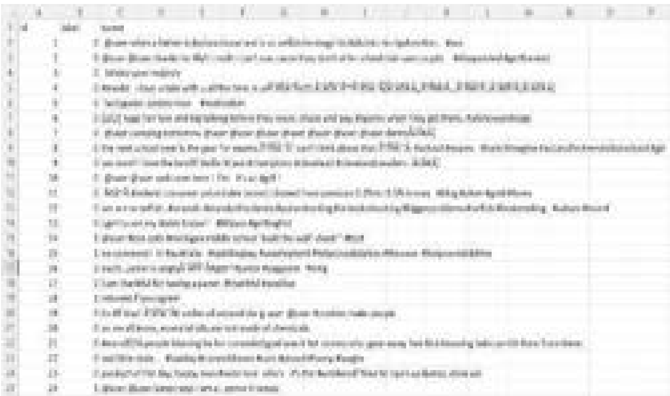


Fig 2. Tweets hate dataset

D. Data Processing

Data handling is the process of turning unstructured data into knowledge that businesses can use. In general, data scientists handle data, which means they gather it, organize it, clean it, check it, analyze it, and turn it into forms that can be read, like graphs or papers. There are three ways to handle data: by hand, mechanically, or electronically. The goal is to make knowledge more useful and decision-making easier. This helps companies run better and make smart strategy decisions more quickly. This is made possible in large part by automated data handling tools, like computer programs. It can help turn big data and other types of data into useful information for decision-making and quality control.

E. Feature selection

Feature selection involves choosing the most dependable, useful, and non-redundant qualities for a model. As record numbers and kinds rise, deliberate shrinkage is necessary. Feature selection aims to improve prediction models and reduce computation power.

One of the most important parts of feature engineering is feature selection, which is the process of choosing the most important features to feed into algorithms for machine learning. Feature selection methods get rid of unnecessary or useless features and only keep the ones that are most important to the machine learning model. This lowers the number of input factors. If you choose which traits are most important ahead of time instead of letting the machine learning model do it, here are the major benefits.

vi) Algorithms:

BERT (Bidirectional Encoder Representations from Transformers) uses a neural network built on transformers to understand and make language that sounds like it was spoken by a person. BERT has a design that only uses encoders. There are both encoder and decoder units in the real Transformer design. BERT's choice to use an encoder-only design says that understanding input sequences is more important than making output sequences. Text is normally processed in a straight line, either from left to right or right to left. This method only lets the model know about the situation right before the target word [12,15].

```
def create_model(bert_model):
    input_ids = tf.keras.Input(shape=(60,),dtype='int32')
    attention_masks = tf.keras.Input(shape=(60,),dtype='int32')

    output = bert_model([input_ids,attention_masks])
    output = output[1]

    output = tf.keras.layers.Dense(32,activation='relu')(output)
    output = tf.keras.layers.Dropout(0.2)(output)

    output = tf.keras.layers.Dense(1,activation='sigmoid')(output)
    model = tf.keras.models.Model(inputs = [input_ids,attention_masks],outputs = output)
    model.compile(Adam(lr=6e-6), loss='binary_crossentropy', metrics=['accuracy',f1_m,precision_m, recall_m])
    return model
```

```
from transformers import TFBertModel
bert_model = TFBertModel.from_pretrained('bert-base-uncased')
```

Fig 3. BERT

There is a parallel model called bidirectional LSTM or BiLSTM, which consists of two LSTM layers. One set of ideas is forward, the other is backward. It is frequently used in NLP-related studies. The idea behind this approach is that by processing the data in two ways, the model can better understand how patterns are related to each other, such as which words come after and before in a sentence. Two unidirectional LSTMs are created from one bidirectional LSTM. These LSTMs can perform forward and reverse operations [14].

Fig 4. BILSTM

GRU (Gated Recurrent Unit): Recurrent Neural Network (RNN). For simplicity, replace the LSTM network. Like LSTM, GRU can process continuous data such as speech, text, and time. The main idea behind GRU is to use gateways to change the secret state of only a part of the network at each step. The block system controls the flow of data into and out of the network. Reset gateway and update gateway are the two gateways that make up the GRU. This gate determines how much of the previously hidden state needs to be removed. The gate also determines how many new gates will be used to change the hidden state. The new hidden state is used to calculate the output of the GRU. This work uses GRU to process the data link faster, making discrimination recognition increasingly effective.

Fig 5. GRU

CNNs are a type of deep neural network that is very good at handling picture and spatial data. CNNs are used on text in natural language processing by treating it like a picture and using filters to find local patterns. CNNs are used in this project because they can find local features and trends in textual data. This helps find hate speech by finding specific language structures that are signs of harmful content.

Fig 6. CNN

CNN + LSTM (Convolutional Neural Network with Long Short-Term Memory), This mixed design takes advantage of both CNNs' ability to find local features and LSTMs' ability to learn in a step-by-step way. The CNN layer can find trends in the raw data that are related to space, and the LSTM layer can describe relationships that span long distances. CNN + LSTM is used in the project to use both local and sequential information. This gives the model a better understanding of context, which improves its ability to spot hate speech.

Fig 7. CNN + LSTM

CNN + BiLSTM (Convolutional Neural Network with Bidirectional Long Short-Term Memory), Like CNN + LSTM, CNN + BiLSTM pairs the CNNs' ability to capture local features with BiLSTM's ability to learn in both directions at the same time. This lets the model look at both the past and the future, reflecting how things depend on each other in both ways. This mixed design is used to find hate speech because it can pick up on subtle timing and contextual trends in the language, which makes the system work better overall.The designs chosen were chosen because they can find different kinds of information in text data. This lets for a more thorough study when it comes to finding hate speech.

Fig 8. CNN + BILSTM

CNN + GRU, This combination design takes the best features of both CNNs and GRUs: linear learning and speed, as well as the ability to recognize patterns in space. The CNN layer records local features, and the GRU layer handles information that comes in a certain order. This mix is probably used in the project because it can handle both local and long-range variables, which makes it better at finding hate speech.

Fig 9. CNN + GRU

LSTM, a better form of recurrent neural networks that is great at detecting long-term relationships and sequence prediction tasks. Time series analysis, machine translation, and speech recognition are some of the things that it can be used for. LSTMs are different from regular RNNs because they have a memory cell with input, forget, and output gates. These gates control the flow of information, letting you choose which bits to keep and which ones to throw away. Because of this one-of-a-kind feature, LSTMs are great for many uses, and they can be mixed with other designs like CNNs to do things like analyze images and videos. LSTM [15] is probably used in the project because it can model environmental information over long runs, which helps the system understand the complex language patterns that are signs of hate speech.

Fig 10. LSTM

LSTM + GRU (Long Short-Term Memory with Gated Recurrent Unit), This mixed design takes the best parts of both LSTM and GRUs, combining their abilities to handle long-term relationships and speed up computations. LSTMs are great at getting faraway environmental information, while GRUs are better at training quickly and dealing with short-term relationships. LSTM + GRU was probably chosen for the project because it models both short-term and long-term linear relationships in a fair way. This helps us understand hate speech patterns better.

Fig 11. LSTM + GRU

LSTM + BiGRU (Long Short-Term Memory + Bidirectional Gated Recurrent Unit), This mixed design blends the LSTM's ability to learn in a sequential way with BiGRU's ability to process information in both directions. Long-range connections are picked up by LSTM, and information is processed in both forward and backward ways by BiGRU. This combo was picked because it works well with sequential data that has complex timing patterns, which helps the model understand hate speech gestures better.

Fig 12. LSTM + BIGRU

Naive Bayes classifiers use Bayes' Theorem. There are other algorithms that function by the same principle: each pair of qualities may be classed separately. Let's start with a collection. The Naïve Bayes classifier is a simple and effective classification approach. It creates fast-predicting machine learning models. The Naïve Bayes classifier claims that things are simpler than they seem, thus the term "Naive". The classifier believes an observation's characteristics are independent of the class name. The Reverend Thomas Bayes is called "Bayes". Simple and fast to train, Naive Bayes can discover hate speech. It may also be used to compare more complex algorithms.

Naive Bayes

```
from sklearn.naive_bayes import MultinomialNB
clf = MultinomialNB()
clf.fit(X_train, y_train)
y_pred = clf.predict(X_test)

lr_acc = accuracy_score(y_test, y_pred)
lr_prec = precision_score(y_test, y_pred,average='weighted')
lr_rec = recall_score(y_test, y_pred,average='weighted')
lr_f1 = f1_score(y_test, y_pred,average='weighted')
```

```
storeResults('Naive Bayes',lr_acc,lr_prec,lr_rec,lr_f1)
```

Fig 13. Naïve bayes

Random Forest is a popular machine learning method for behavioral learning. We can use it for classification and regression in machine learning. Collaborative learning uses multiple classifications to solve complex problems and refine the model. As he says, "Random forest is a distribution with many decision trees that are averaged over different parts of the given data to increase the prediction accuracy of the records." A random forest uses the results of all decision tree predictions and decisions. The one with the most votes will appear as the decision tree.

Random FOrest

```
from sklearn.ensemble import RandomForestClassifier
RandomForest = RandomForestClassifier(n_estimators=10, random_state=0)

RandomForest.fit(X_train, y_train)
y_pred = RandomForest.predict(X_test)

rf_acc = accuracy_score(y_test, y_pred)
rf_prec = precision_score(y_test, y_pred,average='weighted')
rf_rec = recall_score(y_test, y_pred,average='weighted')
rf_f1 = f1_score(y_test, y_pred,average='weighted')
```

```
storeResults('Random FOrest',rf_acc,rf_prec,rf_rec,rf_f1)
```

Fig 14. Random forest

LinearSVC (Linear Support Vector Classifier): The LinearSVC algorithm is a type of the Support Vector Machine (SVM) method that is made for linear classification jobs. SVMs are good at making hyperplanes that split classes in a place with many dimensions. When it comes to finding hate speech, LinearSVC can be helpful because it can handle non-linear decision limits and correctly label hate speech.

LinearSVC

```
from sklearn.svm import LinearSVC
svm = LinearSVC()
svm.fit(X_train, y_train)
y_pred = svm.predict(X_test)

svm_acc = accuracy_score(y_test, y_pred)
svm_prec = precision_score(y_test, y_pred,average='weighted')
svm_rec = recall_score(y_test, y_pred,average='weighted')
svm_f1 = f1_score(y_test, y_pred,average='weighted')
```

```
storeResults('LinearSVC',svm_acc,svm_prec,svm_rec,svm_f1)
```

Fig 15. LinearSVC

RF + SVM + NB (Random Forest + Support Vector Machine + Naive Bayes), This group method takes the best parts of the Random Forest, Support Vector Machine (SVM), and Naive Bayes (NB) algorithms and puts them together. Random Forest adds stability with decision tree ensembles, SVM is great at making useful hyperplanes, and Naive Bayes adds probability classification. This group is probably used because it can collect different kinds of data, which makes hate speech recognition work better overall.

RF + SVM + NB

```
from sklearn.ensemble import VotingClassifier
from sklearn.svm import SVC
from sklearn.ensemble import RandomForestClassifier

estimator = []
estimator.append(('SVM', SVC(probability=True)))
estimator.append(('RFC', RandomForestClassifier()))
estimator.append(('NB', MultinomialNB()))
vot_hard = VotingClassifier(estimators = estimator, voting ='soft')
vot_hard.fit(X_train, y_train)
y_pred = vot_hard.predict(X_test)

vot_acc = accuracy_score(y_test, y_pred)
vot_prec = precision_score(y_test, y_pred,average='weighted')
vot_rec = recall_score(y_test, y_pred,average='weighted')
vot_f1 = f1_score(y_test, y_pred,average='weighted')
```

```
storeResults('RF + SVM + NB',vot_acc,vot_prec,vot_rec,vot_f1)
```

Fig 16. RF+SVM+NB

Stacking Classifier, When you use stacking, you train several models to predict the same outcome, and then you train a meta-model to make predictions based on the predictions of the other models. When it comes to finding hate speech, a Stacking Classifier is probably used to mix the best parts of different base models. This makes the whole system more reliable and accurate.

Fig 17. Stacking classifier

IV. RESULTS AND DISCUSSIONS

Precision: Precision is the percentage of correctly classified cases or samples compared to those that were correctly classified as hits. So, here is the method to figure out the precision:

$$Precision = \frac{TP}{TP+FP} \qquad (1)$$

The resultant graph of precision score is given in fig. 18.

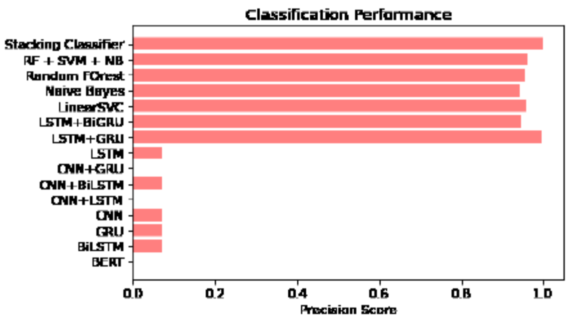


Fig 18. Precision comparison graph

Recall: In machine learning, recall is a parameter that shows how well a model can find all the important cases of a certain class. It shows how well a model captures cases of a certain class. It is calculated by dividing the number of correctly predicted positive observations by the total number of real positives. The resultant graph of recall comparison is given in fig. 18.

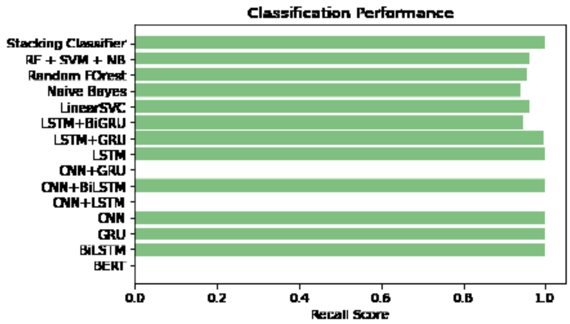


Fig 19. Recall comparison graph

Accuracy: Accuracy is the percentage of right guesses in a classification job. It shows how accurate a model's forecasts are generally.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \qquad (3)$$

The resultant graph of accuracy score is given in fig. 20.

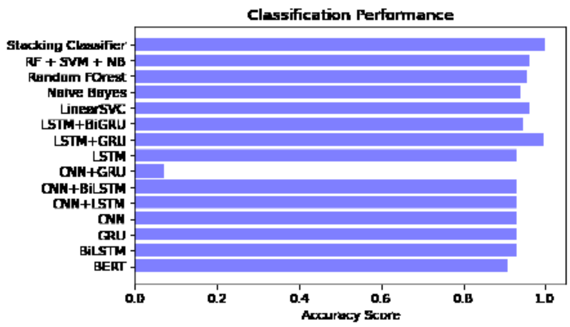


Fig 20. Accuracy graph

F1 Score: F1 score is a compromise between accuracy and return. This is a fair measure that takes both positive and negative into account, so it can be used with inconsistent data.

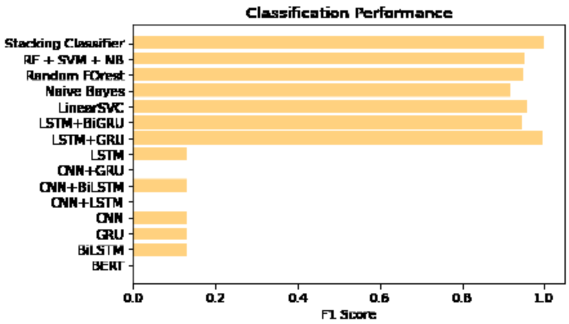
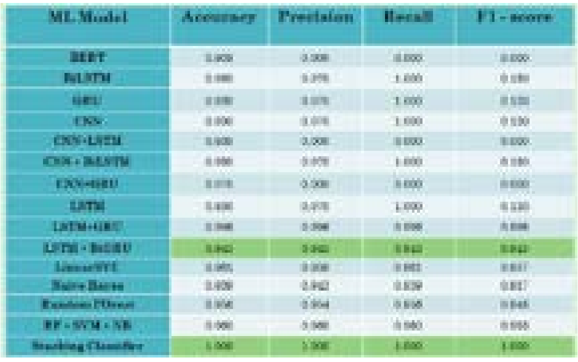


Fig 21. F1 Score

The resultant graph of F1-Score is given in fig. 21 and the performance analysis results are given in fig. 22.



ML Model	Accuracy	Precision	Recall	F1 - score
BERT	[illegible]	[illegible]	0.000	0.000
BiLSTM	[illegible]	[illegible]	1.000	[illegible]
GRU	[illegible]	[illegible]	1.000	[illegible]
CNN	[illegible]	[illegible]	1.000	[illegible]
CNN+LSTM	[illegible]	[illegible]	0.000	0.000
CNN + BiLSTM	[illegible]	[illegible]	1.000	[illegible]
CNN+GRU	[illegible]	[illegible]	0.000	0.000
LSTM	[illegible]	[illegible]	1.000	[illegible]
LSTM+GRU	[illegible]	[illegible]	[illegible]	[illegible]
LSTM + BiGRU	[illegible]	[illegible]	[illegible]	[illegible]
LinearSVC	[illegible]	[illegible]	[illegible]	[illegible]
Naive Bayes	[illegible]	[illegible]	[illegible]	[illegible]
Random FOrest	[illegible]	[illegible]	[illegible]	[illegible]
RF + SVM + NB	[illegible]	[illegible]	[illegible]	[illegible]
Stacking Classifier	1.000	1.000	1.000	1.000

Fig 22. Performance Evaluation

V. CONCLUSION

By correctly finding and reducing hate speech, this research work helps make the internet a safer and better place for everyone. Users report less abuse, which makes the online community more welcoming for everyone. The project's ability to find hate speech is useful for regulatory bodies and platform managers. The method is a useful way to make sure that rules against hate speech on the internet are followed. This gives governing bodies the power to move proactively and keep the online economy healthy. Using a strong ensemble method, the stacked algorithm was able to get 100% accuracy. The model's ability to find and deal with hate speech on Twitter was shown through front-end tests that included identification. Using ensemble methods improves the ability to guess by using the combined strength of several models to get more accurate results. The combination of the Flask framework with SQLite for user signup and signin makes sure that users are safe and authenticated. This protects the privacy of users and creates a reliable tool for people to use with the hate speech recognition system.

VI. FUTURE SCOPE

The goal of future work is to make hate speech recognition work in more than one language. This will have a world effect by creating models that can recognize hate speech in languages other than English. As things get better, the focus will be on making hate speech recognition models more responsive to changing language trends online by adding tools for learning and adjusting to new words and phrases in real time. To make things better, advanced NLP methods such as mood analysis and snark recognition, could be used to help the model understand context better. This would help it tell the difference between harmless statements and real hate speech. In the future, developers may focus on making features that are more useful to users. For example, they could give users customizable filters and content choices to change how sensitive hate speech recognition is. This would make the experience more flexible and easy to use.

REFERENCES

[1] Z. Zhang and L. Luo, ''Hate speech detection: A solved problem? The challenging case of long tail on Twitter,'' Semantic Web, vol. 10, no. 5, pp. 925–945, Sep. 2019.
[2] L. Silva, M. Mondal, D. Correa, F. Benevenuto, and I. Weber, ''Analyzing the targets of hate in online social media,'' in Proc. Int. AAAI Conf. Web Social Media, 2016, vol. 10, no. 1, pp. 1–4.
[3] R. Magu, K. Joshi, and J. Luo, ''Detecting the hate code on social media,'' in Proc. Int. AAAI Conf. Web Social Media, 2017, vol. 11, no. 1, pp. 608–611.
[4] M. Mondal, L. A. Silva, D. Correa, and F. Benevenuto, ''Characterizing usage of explicit hate expressions in social media,'' New Rev. Hypermedia Multimedia, vol. 24, no. 2, pp. 110–130, Apr. 2018.
[5] A. M. Founta, D. Chatzakou, N. Kourtellis, J. Blackburn, A. Vakali, and I. Leontiadis, ''A unified deep learning architecture for abuse detection,'' in Proc. 10th ACM Conf. Web Sci., Jun. 2019, pp. 105–114.
[6] A. Arango, J. Pérez, and B. Poblete, ''Hate speech detection is not as easy as you may think: A closer look at model validation (extended version),'' Inf. Syst., vol. 105, Mar. 2022, Art. no. 101584.
[7] A. Schmidt and M. Wiegand, ''A survey on hate speech detection using natural language processing,'' in Proc. 5th Int. Workshop Natural Lang. Process. Social Media, 2017, pp. 1–10.
[8] J. Langham and K. Gosha, ''The classification of aggressive dialogue in social media platforms,'' in Proc. ACM SIGMIS Conf. Comput. People Res., Jun. 2018, pp. 60–63. [9] P. Fortuna and S. Nunes, ''A survey on automatic detection of hate speech in text,'' ACM Comput. Surv., vol. 51, no. 4, pp. 1–30, 2018.
[10] W. Dorris, R. Hu, N. Vishwamitra, F. Luo, and M. Costello, ''Towards automatic detection and explanation of hate speech and offensive language,'' in Proc. 6th Int. Workshop Secur. Privacy Anal., Mar. 2020, pp. 23–29.
[11] A. Alrehili, ''Automatic hate speech detection on social media: A brief survey,'' in Proc. IEEE/ACS 16th Int. Conf. Comput. Syst. Appl. (AICCSA), Nov. 2019, pp. 1–6.
[12] S. Modi, ''AHTDT—Automatic hate text detection techniques in social media,'' in Proc. Int. Conf. Circuits Syst. Digit. Enterprise Technol. (ICCSDET), Dec. 2018, pp. 1–3. [13] F. E. Ayo, O. Folorunso, F. T. Ibharalu, and I. A. Osinuga, ''Machine learning techniques for hate speech classification of Twitter data: Stateofthe-art, future challenges and research directions,'' Comput. Sci. Rev., vol. 38, Nov. 2020, Art. no. 100311.
[14] E. Shushkevich and J. Cardiff, ''Automatic misogyny detection in social media: A survey,'' Computación Y Sistemas, vol. 23, no. 4, pp. 1159–1164, Dec. 2019.
[15] F. Poletto, V. Basile, M. Sanguinetti, C. Bosco, and V. Patti, ''Resources and benchmark corpora for hate speech detection: A systematic review,'' Lang. Resour. Eval., vol. 55, pp. 477–523, Jun. 2020.
[16] T. X. Moy, M. Raheem, and R. Logeswaran, ''Hate speech detection in English and non-English languages: A review of techniques and challenges,'' Webology, vol. 18, no. 5, pp. 929–938, Oct. 2021, doi: 10.14704/WEB/V18SI05/WEB18272.
[17] W. Yin and A. Zubiaga, ''Towards generalisable hate speech detection: A review on obstacles and solutions,'' PeerJ Comput. Sci., vol. 7, pp. 1–38, Jun. 2021, doi: 10.7717/PEERJ-CS.598.
[18] N. S. Mullah and W. M. N. W. Zainon, ''Advances in machine learning algorithms for hate speech detection in social media: A review,'' IEEE Access, vol. 9, pp. 88364–88376, 2021, doi:

10.1109/ACCESS.2021.3089515.
[19] O. Istaiteh, R. Al-Omoush, and S. Tedmori, ‘‘Racist and sexist hate speech detection: Literature review,’’ in Proc. Int. Conf. Intell. Data Sci. Technol. Appl. (IDSTA), Oct. 2020, pp. 95–99, doi: 10.1109/IDSTA50958.2020.9264052.
[20] R. Rini, E. Utami, and A. D. Hartanto, ‘‘Systematic literature review of hate speech detection with text mining,’’ in Proc. 2nd Int. Conf. Cybern. Intell. Syst. (ICORIS), Oct. 2020, pp. 1–6, doi: 10.1109/ICORIS50180.2020.9320755.

Design and Development of Serial Driver Verification and Test Module using Universal Verification Methodology

Kalpana Murugan [1*]
Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar(Dt) Tamilnadu, India -626126
drmkalpanaece@gmail.com[1*]

P. Syelesh[2]
Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar(Dt) Tamilnadu, India -626126.
9920005035@klu.ac.in[2]

B. Vishnu Vardhan Reddy[3]
Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar(Dt) Tamilnadu, India -626126
9920005123@klu.ac.in.[3]

C.V. Subba Reddy [4]
Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar(Dt) Tamilnadu, India -626126.
9920005114@klu.ac.in[4]

T. Praveen Kummar[5]
Department of Electronics and Communication Engineering, Kalasalingam Academy of Research and Education, Krishnankoil, Virudhunagar(Dt) Tamilnadu, India -626126
9920005220@klu.ac.in.[5]

Abstract— In the field of digital design and verification, serial communication driver verification is essential for guaranteeing the accuracy and dependability of electronic systems. A standardized framework for functional verification in hardware design, the Universal Verification Methodology (UVM), was used to construct the Serial Driver Verification Testing Module that is shown in this abstract. The suggested module is an essential part of the verification environment as it makes it easier to evaluate if the serial driver complies with communication protocols. The module is made to systematically watch, evaluate, and score the interactions between the serial driver and the tested Device Under Test (DUT) by utilising UVM's capabilities. The Serial Driver Verification Scoreboard Module's key features include effective reporting systems, error detection, and thorough protocol verification. Scalability and maintenance are improved by the modular and reusable verification components made possible by the UVM approach. By comparing predicted and observed communication sequences, the scoreboard module may identify potential protocol breaches and offer insightful information about the driver's conduct.

Keywords: Test bench, Serial driver, Sequences, scoreboard Directed Acyclic Graph (DAG), Monitor

I. INTRODUCTION

Serial communication protocols are essential to the dynamic world of digital communication because they allow various electronic devices to share data. Making sure serial communication drivers are proper and strong is becoming more and more important as the need for fast and dependable communication grows. In this work, a new method for verifying serial drivers using the Universal Verification Methodology (UVM) is presented.

The suggested approach makes use of a Scoreboard Module to thoroughly evaluate the performance and functionality of serial drivers, enhancing the general dependability of communication networks.

The methodology of Universal Verification has acquired broad acceptance in the verification field due to its efficiency in handling the complexity of contemporary digital designs. This study seeks to improve serial communication driver verification using UVM, offering a method of scalable approach. A crucial element of the verification environment is the recently developed Scoreboard Module, which permits real-time data monitoring and validation of the data sent between the serial driver and the communication interface.

A UVM monitor is a passive component, which is used to capture DUT signals using a virtual interface and translate them into a sequence item format. These sequence items or transactions are broadcast to other components like the UVM scoreboard, coverage collector, etc. It uses a TLM analysis port to broadcast transactions.

II. LITERATURE SURVEY

This section describes the current efforts in the field of Soc verification using UVM. A generic and reusable UVM model has been presented by Hussein et al. (2019) to enhance the serial. The results indicated the suggested technique is more appropriate to provides quicker, reusable, and trustworthy verification [1] by Yashas et al., 2019. Using many tests, this method creates a random test bench to achieve 100% code coverage and functional coverage [2].

The increasing complexity of silicon processors has been responsible for many of the challenges in verification. Therefore, UVM based AX14 stream verification, and it's results with that verification was explained by Xu et al., 2019 [3]. A system Verilog based UVM verification for DPLL has been mentioned in a paper by Georgousopoulos and Hatzopoulos, 2019 [4]. Because it provides a soiled gain in verification and this suggested verification model is difficult rapid, reusable, and dependable in the digital market which is interface for serial peripherals [5].

Guo et al, 2020 introduce the (SPI) module for UVM based verification, which helps in achieving communication between the SoC components. The verification pattern has influenced the SoC verification environment, and the SPI module gives 100% coverage rate [6]. It might be hard to set up a complicated test bench, but working on this makes the verification of complex SoCs easier [7].

As a result, a controller area network based on UVM and the results with effective verification has been given by Godel et al. 2020 [8]. El Ashy, 2020 introduced a generic verification environment for NoC platform that provides reusable and error verification solutions in addition to evaluating DUT. This method examines the viability and compatibility of few NoCs for error verification strategies [9]. Hao et al. in 2019 studied a UM-based single wire protocol (SWP) interface module for verification is presented for SIM card chips.[10]

The random incentives that increase the interface module's verification performance have been preserved in the research [11]. by Priyanka 2021, a Verilog-based UART design and UVM verification are shown. The design clarifies the overall functionality of UART, while UVM verification is noted to reduce energy consumption reusability [12]. K.W. Wong. has proposed a UVM method in 2019 for confirming an RTL code connection to implement memory-centric computing.[13] By using some UVM techniques the packets are created, and some packets were sent into this switch [14]. ports through virtual interface Effective. Verification for memory-centric computing is brought forth by this method [15].

III. Serial Driver Verification & Test Module using UVM.

Verification of a Serial Driver: This refers to the procedure for validating a serial driver, which is a piece of hardware or software that controls how a computer and a serial device communication Module for the Scoreboard: A scoreboard is an element used in hardware verification that verifies the accuracy of the outputs generated by the tested design (DUT). It tracks anticipated and actual outcomes and indicates any differences. A systematic process for confirming integrated circuit designs is called UVM. The semiconductor industry uses it extensively for functional verification of digital designs.

This approach helps discover corner cases and potential issues in the SDV design. Develop protocol-specific assertions using SystemVerilog Assertions (SVA) or Property Specification Language (PSL) to formally specify the expected behaveior of the SDV communication. These assertions can be checked during simulation to ensure compliance with the protocol.

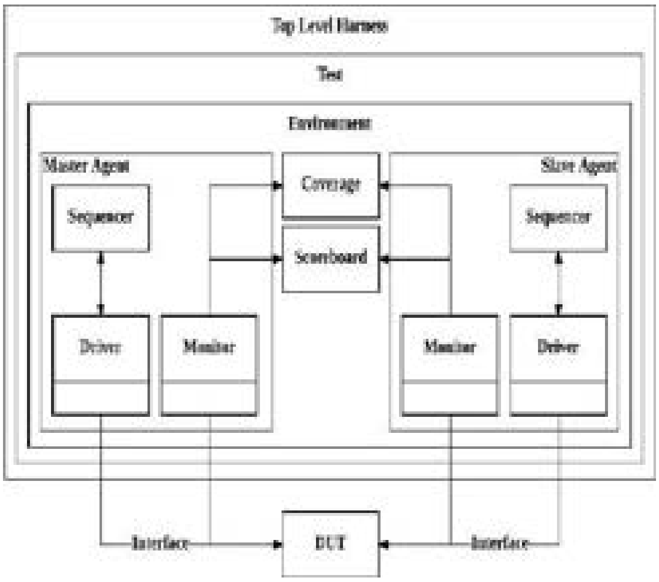


Fig. 1 Block diagram of UVM module

The fig.1 explains the Testbench example architecture with detailed explanation on writing each component link to testbench flow through block diagram. In contrast to parallel communication, which transfers data through communication channel bit by bit, serial communication transfers the data several bits at once.

3.1 WORKING PRINCIPLE:

Serial Driver: This most likely denotes a part or module in charge of managing serial communication. Unlike parallel communication, which sends data via a communication channel one bit at a time, serial communication sends data corruption.

A component or module in charge of overseeing serial communication is most frequently indicated by the term "serial driver." In contrast to parallel communication, which transfers data via a communication channel bit by bit, serial communication transfers data several bits at once. Verification Scoreboard Module: A scoreboard is a module used in hardware or software verification that monitors predicted and actual outcomes during the verification process. Finding any differences between the predicted and actual behaviour of the system under verification is helpful. A systematic process for confirming integrated circuit (IC) designs is called UVM (Universal Verification Methodology). For verification the functional of digital hardware, it is extensively utilized in the field of electronic design automation.

Now, putting these parts together, a "Serial Driver Verification Testing Module using UVM" would probably a verification element made to use the UVM approach to verify that a serial driver module is proper. For digital designs, the UVM framework offers a standardized method for building testbenches and verification environments. The idea behind such a module would be to build a scoreboard inside the UVM testbench to keep track of and validate the serial driver's actions. This might entail making that the driver compiles the requirements, handles errors effectively, and transmits and receives data accurately. Now, putting these parts together, a "Serial Driver Verification Scoreboard Module using UVM" would probably be a verification element made to use the UVM approach to verify that a serial driver module is proper. For digital designs, the UVM framework offers a standardized method for building testbenches and verification environments. The idea behind such a module would be to build a scoreboard inside the UVM testbench to keep track of and validate the serial driver's actions. This might entail making that the driver complies with the requirements, handles errors effectively, and transmits and receives data accurately. The scoreboard module may do a comparison between the actual behaviour seen during testing or simulation and the specified expected behaviour based on the specifications or criteria. As part of the verification process, any mistakes or anomalies found by the scoreboard would be recorded, assisting the verification engineer in locating and troubleshooting any issues with the serial driver implementation.

3.2 IMPLEMENTATION

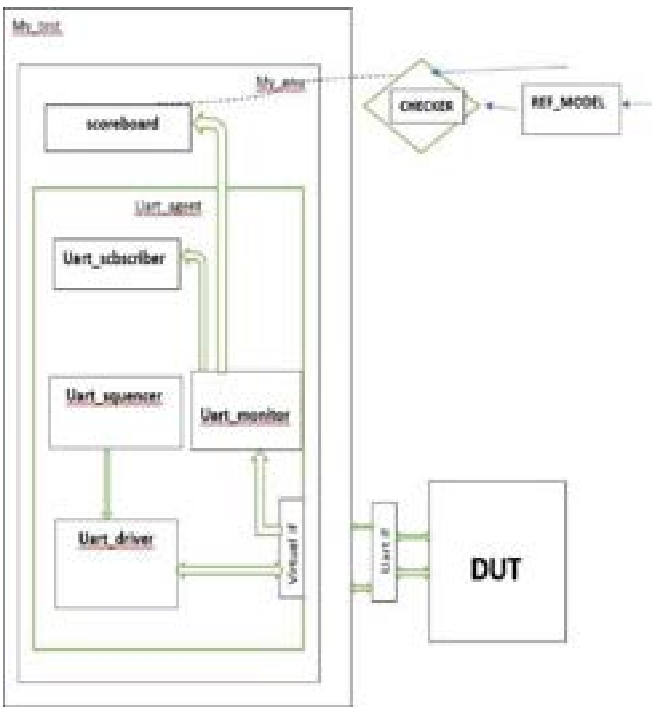


Fig.2.Interfacing of DUT processed data with the Scoreboard.

Figure.2 explains a scoreboard, which is frequently used in UVM to compare predicted and actual outcomes from the design that is being tested (DUT). In this methodology functional coverage collector are implemented. Protocol is defined to facility of the driver implementation of I2Cdevices. Adjust tests as needed to improve coverage uvm analyze through ex the development of two testbenches for unit to verification first one is targeted to fifo buffer module implemented

IV. RESULT AND DISCUSSION

The UVM verification component would generate a sequence of signals to send to the DUT. The specific signals would depend on the functionality being verified. The component would also monitor the DUT's outputs to see if they match the expected outputs and finally compare the DUT's outputs to the expected outputs to determine if the DUT is functioning correctly.

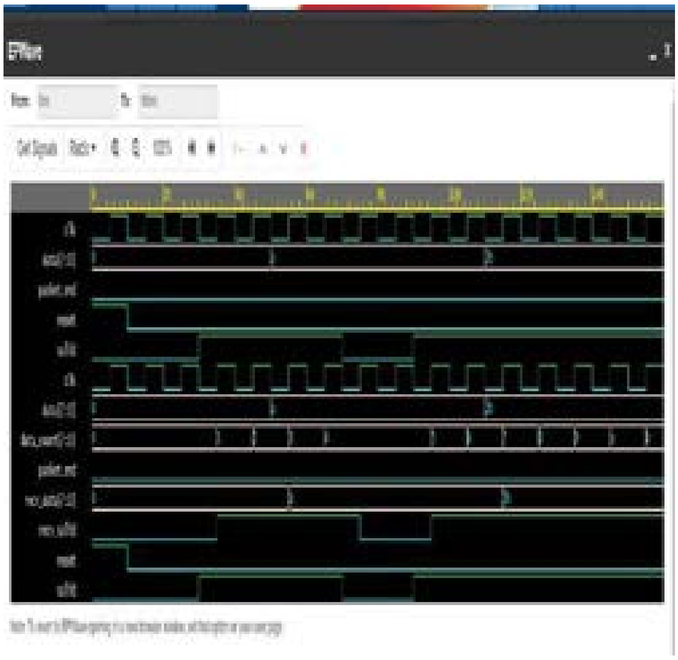


Fig.3 output of serial driver verification and test Module

Figure.3 explained the running the Utilize a Directed Acyclic Graph (DAG) to illusate the relationships between UVM components. Nodes in the graph represent UVM components (e.g., driver, monitor), and edges represent the flow of transactions UVM to assess the completeness of your tests. Adjust tests as needed to improve coverage uvm analyze through ex the development of two testbenches for unit to verification. first one is targeted to fifo buffer module implemented .At present verification Environment of image signal processing of IP'S/SOC'S with strong verification methodology contributes wave signal of first silicon process.

V.CONCLUSION

A reliable and effective method for confirming the operation and performance of serial drivers in a design is the Serial Driver Verification Scoreboard Module, which uses the Universal Verification Methodology (UVM). This verification scoreboard module offers enhanced scalability and reusability by providing a systematic and uniform verification approach using UVM. An essential part of the verification environment, the scoreboard allows for the observation and examination of communication between the DUT and the serial driver (Device Under Test). UVM automates many aspects of testbench development and execution, including stimulus generation, checking, and result analysis. This automation reduces manual effort, speeds up verification cycles, and improves overall productivity.

VI.REFERENCES

[1] Ritesh and M. Vivek "The Design of High Speed UART," *Proceedings of 2013 IEEE Conference on Information and Communication Technologies* (ICT 2005), vol.5, Issue 9, Johar, Malasiya , 20-21 December 2005.

[2] C. Sujata and A. Gubbi, "Using Verilog and System Verilog Design and Verify the Communication Bridge between APB and I2C Protocol," *International Journal of Science Technology & Engineering,* vol.3,no.6, pp. 33-35, July 2016.

[3] P. Venkateswaran, "FPGA Based Efficient Interface Modelfor Scalefree Computer Network using I2C Bus Protocol," *Advances in Computer Sci & Engg., National Polytechnic Institute*, Mexico, vol.23, no.7 , pp. 191-198, November 2006.

[4] P.Sudeep, D.Nikil, B.Elaheh and M.Benromdhane, "Floorplan-Aware Bus Architecture Synthesis," IEEE *Transactions on Very Large Scale Integration Systems,* vol. 14, no. 3, pp. 241-253, March 2006.

[5] F.Mo, T.Abdallah and B.Robert "A Timing-driven Macro-cell Placement Algorithm," vol.5, pp.322-327, July 2001.

[6] Fadnavis and B. Kariyappa, "PNR flow methodology for congestion optimization using different macro placement strategies of DDR memories," *International Journal of Advanced Technology and Engineering Exploration*, vol. 6, no.4, pp. 294-308, June 2021.

[7] P. Saxena, R. Shelar, and S. Sapatnekar, "Routing Congestion in VLSI Circuits," Estimation and Optimization. Springer Science and Business Media , 2007.

[8] K. Lakshmanna, "Perimeter degree technique for the reduction of routing congestion during placement in physicaldesign of VLSI circuits," 2022.

[9] Lin and J. Ming, "Routability-driven global placer target on removing global and local congestion for VLSI designs," *International Conference On Computer Aided Design,* Munich, Germany, 01-04 November 2021.

[10] M.J.S Smith, "Floor planning and Placement," *Application Specific Integrated Circuits*, 1997.

[11] G. Shivani and S. Neeraj, "A Study of Floor planning Challenges and Analysis of macro placement approaches in Physical Aware Synthesis," *International Journal of Hybrid Information Technology*, vol. 9, no.1, pp. 279- 290, March 2016.

[12] C. Valenzuela and P. Wang, "VLSI placement and area optimization using a genetic algorithm to breed normalized postfix expressions," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 4, pp. 390-August

[13] H. Murata, K. Fujiyoshi, and Y. Kajitani, "VLSI module placement based on rectangle-packing by the sequence-pair," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 15, no. 12, pp. 1518-1524, November 1996.

[14] Tsung-Ying Sun, Sheng-Ta Hsieh, Hsiang-Min Wang, and Cheng-Wei Lin, "Floor planning Based on Particle Swarm Optimization," Proceedings of the Emerging VLSI Technologies and Architectures, 2006

[15] J. Lin and Y. Chang, "TCG: A transitive closure graphbased representation for general floorplans," *IEEE Design Automation Conference,* June 2001.

IoT based Vehicle Monitoring System for Safety Standards in Smart Cities

Piruthiviraj P
Associate professor
Dept of Electronics and communication New horizon college of Engineering Bengaluru, India
prithivi.eie@gmail.com

Tarun S R
UG student
Dept.of Electronics and communication New horizon college of Engineering Bengaluru,India
tarunsr182002@gmail.com

Tarun M K
UG student
Dept.of Electronics and communication New horizon college of Engineering
Bengaluru,India
tarunyadav1104@gmail.com

Tarun M
UG student
Dept.of Electronics and communication New horizon college of Engineering Bengaluru,India
suryatarun026@gmail.com

Abstract*— The rapid urbanization and increasing number of vehicles in smart cities have increased the importance of ensuring road safety. This model proposes the development and implementation of an Internet of Things (IoT)-based vehicle monitoring system to increase the safety standards in smart city environments. Leveraging cutting-edge technology, the system continuously monitors and evaluates on-road vehicle behavior in real-time, providing essential data for traffic management and accident prevention. By integrating a variety of in-car sensors and communication devices, the proposed system collects and transmits data to a centralized server, including information on the vehicle's position, speed, acceleration, and braking patterns. This data enables the detection of potentially hazardous situations, such as aberrant driving patterns and traffic congestion. The IoT-based Vehicle Monitoring System not only monitors individual vehicle behavior but also facilitates communication between vehicles and infrastructure, fostering a cooperative traffic ecosystem. With the incorporation of advanced analytics algorithms, the proposed system can identify factors affecting road safety and enhance its prediction capabilities over time. By harnessing technologies like Artificial Intelligence (AI) and Machine Learning (ML), the proposed system continuously evolves to address evolving safety challenges. The ultimate goal of this study is to contribute to the development of smarter and safer cities by proactively addressing safety and traffic management concerns, thereby reducing traffic congestion, accidents, and creating a more secure urban transportation network.

Keywords—Internet of Things, Traffic Monitor, Safety, Node Microcontroller unit.

I. INTRODUCTION

In the dynamic field of smart cities, where advances in technology are transforming urban living, creative problem-solving is required to tackle the intricate issues brought on by a rise in the number of vehicles on the road. The lack of adequate road safety standards is one of the biggest issues that has to be addressed. With unprecedented levels of urbanization, there is a growing demand for cutting-edge technologies that actively monitor and enhance road safety for vehicles. By creating and deploying an Internet of Things (IoT)-based vehicle monitoring system especially for smart cities, this module seeks to address this challenge. Combining Internet of Things (IoT) technologies with vehicle monitoring aims to produce a flexible and adaptable system that can analyze the behavior of individual cars and potentially promote cooperation between car and city infrastructure. The main objective of this model is to create urban transportation networks that are more adaptable, safe, and efficient. Finally, as opposed to merely being a technical advancement, IoT-based Vehicle Monitoring System is a deliberate step toward changing urban mobility. The system aims to greatly contribute to the development of safer and smarter cities by fusing state-of-the-art technologies with an emphasis on safety. These cities' roadways have been transformed into dynamic, intelligent areas that may change to meet the ever-evolving needs of urbanization.

II. LITERATURE SURVEY

A. SECURITY SYSTEM USING INTERNET OF THINGS IN SMART CITIES

Authors: Adeoye, A. Adebiyi, O. Oluwarotimi . In this paper, the authors have done implementation on Applications in Transportation, challenges and opportunities and integration. In their paper on IoT-based security systems in smart cities, Adeoye, A., Adebiyi, O., and Oluwarotimi detail a technical framework that integrates sensors like cameras and motion detectors into transportation infrastructure. These sensors provide real-time data to a centralized unit, which uses edge computing to handle it quickly and with less delay. The data is analyzed for potential threats using advanced security algorithms that include image recognition and video analytics. A strong communication network, like 5G or LTE, guarantees constant device connectivity. Security staff can monitor using an intuitive dashboard from the centralized control center, and overall efficacy is increased through interaction with other smart city technologies. Predictive analytics is made possible by machine learning algorithms, and the system complies with strict security guidelines. This thorough but succinct method guarantees a technically advanced and flexible security system. One noteworthy aspect of their framework proposal is the integration of edge computing, which is a paradigm that involves handling data close to its source. This strategy is essential for reducing latency, guaranteeing effective processing, and maximizing resource use. Using image recognition and video analytics, advanced security algorithms are yet another essential component of their system. These algorithms probe into the sensor data streams in real time and do complex evaluations to find security issues and possible threats. This sophisticated degree of inspection goes beyond conventional techniques, improving the system's ability to identify subtle security threats. Finally, the writers stress how crucial it is to follow strict security guidelines and include safeguards against illegal access and cyberthreats. This promise guarantees its honesty, dependability, and credibility.

B. A COMPREHENSIVE REVIEW ON THE ROLE OF INTERNET OF THINGS IN TRANSPORTATION SYSTEM

Authors:Muhammad Awais Javed, Abbas Anwar

In their comprehensive review on the role of IoT in transportation systems, authored by Muhammad Awais Javed and Abbas Anwar, the focus is placed on the pivotal role of sensors in monitoring and ensuring a safe environment. The implementation involves a network of sensors strategically placed within the transportation infrastructure. These sensors cover a wide range of technological platforms, such as cameras, environmental sensors, and other data collection tools. The sensors gather information about ambient factors, traffic patterns, and system functioning continuously. A centralized system receives this real-time data for analysis and decision-making. Lidar and radar are examples of advanced sensor technologies that are probably used for accurate object tracking and detection. In order to monitor the transportation ecosystem and enable the system to quickly identify and address any potential safety concerns or abnormalities, sensors play a critical role. A dynamic and networked transportation system is created through the seamless connection between components made possible by integration with IoT technology. This analysis emphasizes how crucial sensor technologies are to the foundation of a successful IoT-based transportation infrastructure, greatly enhancing efficiency and safety.

C. DESIGN AND IMPLEMENTAION OF AN AUTOMATED TRACKING SYSTEMS.

Authors: P Sudhakar , R Arul Murugan.

The authors initiated a technical exploration of a system aimed at real-world implementation, with a focus on current trends in tracking technology. To achieve accurate and real-time position tracking, the design incorporates sophisticated tracking mechanisms including GPS and RFID. To effectively handle and analyze location data, the system probably combines hardware elements like GPS modules and RFID scanners with integrated software. The automated tracking system probably has real-time monitoring and management capabilities for the movement of people, cars, and other assets. The writers might talk about how tracking devices and the central monitoring system might communicate seamlessly by using wireless communication protocols, perhaps based on the Internet of Things. It is likely that the study tackles implementation issues in the actual world by taking into account aspects such as security, dependability, and scalability. It may be investigated what current tracking technology trends are, such as the incorporation of machine learning algorithms for data analytics and predictive tracking. In summary, this technical explanation highlights the usefulness of putting an automated tracking system and safety standards into place while keeping a close eye on new developments to guarantee applicability and efficiency in the modern environment.

III. SYSTEM COMPONENT

A. NODE MCU: The NodeMCU ESP8266 is outfitted with the ESP-12E module, an advanced part that includes the potent ESP8266 chipset and the 32-bit LX106 RISC microprocessor from the Tensilica Xtensa. The NodeMCU is a great option for requiring wireless connection since the ESP8266 chipset, which is well-known for its abilities in this area, allows for smooth connectivity. wireless communication functionalities. The Tensilica Xtensa's 32-bit LX106 RISC microprocessor, housed within the ESP-12E module, provides the computational muscle for executing complex tasks efficiently. Notably, the NodeMCU's ESP8266 microcontroller supports Real-Time Operating Systems (RTOS), offering a flexible and responsive framework for handling concurrent tasks. Furthermore, the NodeMCU's variable clock frequency—which ranges from 80MHz to 160MHz—allows developers to customize the processing speed of the device to meet the unique needs of their applications while adapting to varying performance requirements. The NodeMCU ESP8266's feature-rich architecture ensures excellent performance, scalability, and dependability in a wide range of embedded systems applications, making it an outstanding solution.

B. NEO6M GPS Receiver: In the IoT-based Vehicle Tracking System with NodeMCU, the NE06M GPS receiver is a critical component providing precise location data. Interfaced through UART, the NodeMCU communicates with the NE06M, extracting latitude, longitude, and other relevant information. This GPS module receives signals from satellites and transmits processed data to the NodeMCU for further processing. The NodeMCU's computational capabilities and connectivity integrate the NE06M's data into the system, enhancing real-time vehicle tracking accuracy. The NE06M's reliability and the NodeMCU's processing prowess make this integration a robust solution for location-based tracking applications.

C. 16x2 LCD Display: An essential user interface component of the Internet of Things Vehicle Tracking System with NodeMCU is the 16x2 LCD display. The 16x2 LCD display serves as a visual output device, giving users or system administrators access to real-time information with 16 characters spread across 2 lines. The NodeMCU interfaces with the LCD screen to show pertinent tracking system data, including GPS locations, vehicle speed, and timestamps. Overall, the incorporation of the 16x2 LCD display improves user interaction and guarantees that the Internet of Things-based Vehicle Tracking System operates in an accessible and visible manner.

D. I2C Module: An I2C (Inter-Integrated Circuit) module is included in the NodeMCU-based Internet of Things Vehicle Tracking System to enable smooth communication between the NodeMCU and linked sensors, such as GPS modules. Data exchange is streamlined by this two-wire communication protocol, which makes it possible for the NodeMCU to collect data from multiple sensors effectively. In order to connect and control various peripherals and improve the tracking system's overall functionality, the I2C module is essential. Real-time tracking capabilities are enhanced by the NodeMCU's processing of data received via I2C and its transmission of pertinent information to a central server.

E. ESP8266: An essential part of Internet of Things-based car monitoring systems for smart city safety regulations is the ESP8266. It gathers information like GPS location and car diagnostics, connects to central servers wirelessly via Wi-Fi to send data in real-time, and interacts with external and onboard sensors for all-encompassing monitoring. Due to its edge computing capabilities, latency can be decreased by processing data at the source. It also makes it easier to monitor and control cars remotely, all the while protecting data using encryption. Its affordability makes it even more suitable for broad implementation, which makes it a crucial component of ecosystems for safer urban transportation.

IV. WORKING MODEL

Key components, such as the ESP8266 NodeMCU, NE06M GPS Receiver, 16x2 LCD, and 16x2 LCD I2C module, are integrated to build a solid system in the implementation of this advanced GPS tracking device for real-time vehicle location monitoring. As the main microcontroller, the ESP8266 NodeMCU precisely and effectively coordinates communication and cooperation among the networked components. The highly praised NE06M GPS Receiver ensures real-time tracking of the vehicle's movements by providing essential location data. When combined with the I2C module, the 16x2 LCD display provides an advanced user interface for direct and on-device information presentation. The ThingSpeak IoT cloud platform adds a major layer of capability to the model. In this example, ThingSpeak is utilized to log the previous areas that the car has driven over. ThingSpeak is a flexible IoT solution for data logging, analysis, and visualization. This cloud-based method allows for smooth connection with other apps in addition to facilitating remote access to real-time data.

In order to maintain continuous device operation, the carefully planned hardware setup—which includes connectors that provide a stable arrangement on a breadboard—is supported by a dependable power source. Setting up the channels, fields, and visualizations in the Thing Speak software allows for thorough tracking of the GPS data. For improved data interpretation, the system not only shows raw data but also converts it into interpretable graphical representations. The hardware is assembled precisely in this sophisticated GPS tracking system for tracking the location of vehicles in real time. One multipurpose microcontroller that facilitates smooth communication between networked components is the ESP8266 NodeMCU. The device obtains precise location information by utilizing the highly precise NE06M GPS Receiver. Combined with the I2C module, the 16x2 LCD display offers an advanced user interface that shows important data right on the device.

An essential layer for data storage, processing, and visualization is introduced by the interaction with the ThingSpeak IoT cloud platform. ThingSpeak is set up to record location history from the car, allowing for easy interaction with other apps and remote access. This cloud-based method offers flexibility and agility for future expansions in addition to facilitating real-time monitoring. ThingSpeak is set up to receive data from the GPS module and display it graphically during the monitoring process. This entails configuring ThingSpeak's fields, channels, and visualizations. Comprehensive analysis and interpretation are made possible by the graphical representation of latitude and longitude data. The system's usefulness is increased when users obtain knowledge about the trajectory, speed, and stopping of the vehicle.

Additionally, the module includes a dynamic link that takes viewers to Google Maps on the webpage. By offering a clear, visual depiction of the vehicle's current location in real time, this function enhances the user experience. In addition to providing a user-friendly interface, this integration with Google Maps shows how versatile the system is when it comes to utilizing other resources. The hardware configuration is powered by a dependable source to ensure continuous functioning. It is painstakingly arranged on a breadboard with connectors that provide stability. To ensure smooth device integration, the ThingSpeak software configuration entails scripting data handling and visualization methods. The system uses sophisticated algorithms for anomaly identification and data interpretation in addition to storing raw data.

The cloud-based IoT solutions and cutting-edge hardware components are integrated in this technical concept of the IoT-based Vehicle Tracking System. ThingSpeak serves as a powerful IoT platform for storage, analysis, and visualization, and the ESP8266 NodeMCU, NE06M GPS Receiver, and 16x2 LCD components work together to collect data in real-time and historical fashion. The painstaking monitoring procedure, enhanced by dynamic interfaces and graphical representation, results in an advanced and technologically proficient solution for all-encompassing vehicle tracking and monitoring.

In the development of our IoT-based Vehicle Tracking System, the integration of hardware components is executed with precision and technical sophistication. The ESP8266 NodeMCU serves as the central orchestrator, seamlessly coordinating communication among components. Leveraging the NE06M GPS Receiver for high-precision location data and the 16x2 LCD display with an I2C module for an intuitive user interface, the system performs well in real-time and historical data acquisition. The inclusion of the ThingSpeak IoT cloud platform elevates by providing a scalable and versatile solution for data storage, analysis, and visualization. The monitoring process involves configuring ThingSpeak to graphically represent latitude and longitude data, allowing for insightful analysis. Additionally, the integration of a dynamic link on the webpage, redirecting users to Google Maps for real-time visualization, enhances user experience. The meticulous hardware setup, coupled with ThingSpeak software configuration, ensures continuous and reliable operation. In summary, the proposed IoT Vehicle Tracking System exemplifies the integration of cutting-edge hardware components and cloud-based IoT solutions, offering a sophisticated and technologically adept solution for comprehensive.

In conclusion, the cloud-based IoT solutions and sophisticated hardware components are harmoniously integrated by the IoT-based Vehicle Tracking System. Real-time data capture is ensured by the effective management of communication among components by the ESP8266 NodeMCU. With the help of an I2C module and a 16x2 LCD display powered by the NE06M GPS Receiver, the system achieves exceptional precision and user engagement. Scalability is offered for data storage, processing, and visualization through integration with ThingSpeak. The monitoring procedure, set up in ThingSpeak, exhibits sophisticated data processing abilities. The website's dynamic link, which takes visitors to Google Maps for real-time display, improves user experience and demonstrates system flexibility. Continuous operation is ensured by careful hardware setup and complex ThingSpeak software settings. The IoT Vehicle Tracking System, which provides a comprehensive and cutting-edge solution for vehicle tracking and monitoring, is essentially the epitome of integration.

In summary, the IoT Vehicle Tracking System seamlessly integrates advanced hardware with cloud-based IoT solutions, offering a precise, user-friendly, and technologically advanced solution.

V. BLOCKDIAGRAM

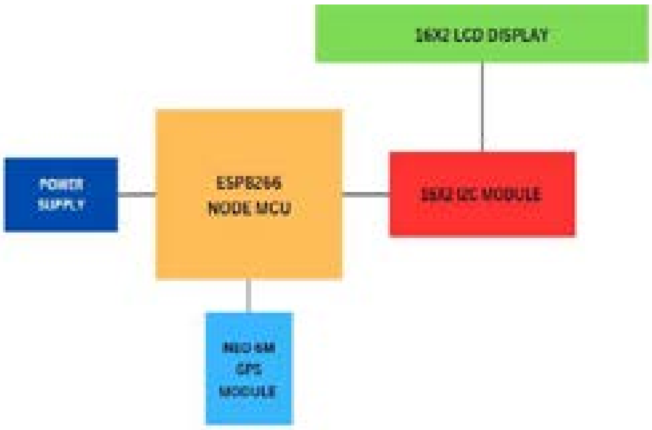


Fig.1.**Block diagram of Monitoring Unit**

The complex architecture of the ESP8266-based car tracking systemshown in figure 1 includes a block diagram that shows all of the parts that are necessary for it to function properly. The ESP8266 module, which serves as the central controller in charge of communication and system administration overall, is at the center of the system. The Neo GPS module, which is essential for obtaining precise GPS coordinates, is connected to this central unit. The ESP8266 and Neo GPS module communicate with each other through a CPU, which provides cellular connection for distant communication and real-time tracking updates. The brains of the system, the ESP8266, communicate with the Neo GPS module to retrieve exact location information, which guarantees precise vehicle monitoring. User interaction and real-time updates are facilitated through a user interface accessible via a web interface or mobile app. Additionally, a database integration may store previous tracking data for a comprehensive tracking solution. This block diagram exemplifies the sophisticated connections and interactions among the ESP8266 module, Neo GPS module, microprocessor, and user interface, underscoring the system's efficiency, accuracy, and advanced tracking capabilities.

A user interface can be accessed via a web interface or mobile app, enables user engagement and real-time tracking updates. Users can easily track the location of the tracked vehicle and get real-time information with this interface. Furthermore, tracking data can be connected to a database in specific situations to create a full tracking solution. By preserving historical tracking data, this database integration helps provide a comprehensive picture of the vehicle's movements over time.

In conclusion, the ESP8266-based vehicle tracking system, outlined through a comprehensive block diagram, embodies a technologically advanced solution for real-time monitoring and precise location tracking. The ESP8266 module, serving as the central controller, orchestrates communication and system management seamlessly. The integration of the Neo GPS module ensures the acquisition of accurate GPS coordinates, fundamental for precise vehicle tracking. A microprocessor facilitates efficient communication between these core components, enhancing connectivity and enabling cellular access for remote updates. The user interface, accessible through a web interface or mobile app, provides a convenient means for users to monitor vehicle locations in real-time. Furthermore, the potential linkage of tracking data to a database adds a layer of comprehensive historical tracking. This block diagram underscores the system's sophistication, illustrating the integration and synergy among key components for a robust, efficient, and user-friendly vehicle tracking solution. The ESP8266-based system represents a paradigm in IoT-based tracking, emphasizing accuracy, accessibility, and comprehensive functionality.

\

VI. OUTPUT

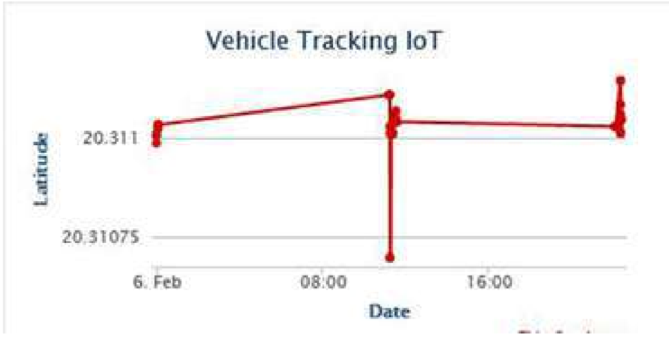


Fig.2.**Latitude vs Date**

Figure 2 shows the graph plots of the Latitude parameter with respect to Date. This is quite useful as it can even see the previous values to get an idea of how the Vehicle has been for the past few hours. It can see how the current values stack up with the previous values.

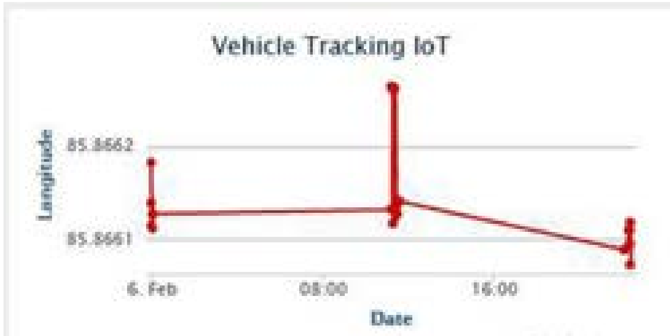


Fig.3.**Longitude vs Date**

Figure 3 shows the graph plots of the Longitude parameter with respect to Date. This is quite useful as it can even see the previous values to get an idea of how the Vehicle has been for the past few hours. It can see how the current values stack up with the previous values.

Longitude and latitude data from the GPS module are sent to ThingSpeak via the ESP8266 NodeMCU in the IoT-powered car monitoring system. This data is stored in ThingSpeak channels throughout time, resulting in a chronological dataset. MATLAB ThingSpeak visualizations produce clear graphs that plot latitude and longitude versus time. Applying technical algorithms can improve the precision of the data. These graphs provide a concise, informative summary of the spatial movements of the vehicle and are useful for route analysis, speed profiling, and stoppage detection. This integrated system creates a strong foundation for long-term spatial data preservation and effective analysis in addition to guaranteeing real-time monitoring. The integration of IoT, GPS modules, and ThingSpeak in this vehicle monitoring system not only provides real-time updates but also facilitates long-term data storage and analysis. of the vehicle's spatial movements, allowing for detailed analysis and insights into the vehicle's route& speed.

VII. IMPLEMENTATIONOF VEHICLE MONITORING SYSTEM

In order to proactively control traffic, the IoT-based vehicle monitoring system for smart cities combines sensors, cloud computing, and machine learning. Sensors gather information in real time on location, speed, and ambient variables using technologies such as GPS. Machine learning techniques are used by the cloud-based central control system for traffic optimization and hazard prediction. By warning drivers to speeding offenses, identifying hazardous behavior, and quickly notifying emergency services, the system enforces safety standards. Real-time alerts for drivers and data for city officials are provided using a smartphone app. Strong security and privacy protocols, such as encryption, are in place. In smart cities, this technology transforms traffic management to provide safer. Figure 4 shows the real-time result.

F

Fig. 4 Real Time Monitoring Result

VIII.CONCLUSION

To sum up, the Vehicle Monitoring System for Safety Standards in Smart Cities, which is based on the Internet of Things, is leading the way in developing intelligent urban transportation networks. With a focus on data-driven safety, state-of-the-art technology, and potential interaction with smart cities, the system is scalable and flexible enough to change with urban surroundings. The potential of IoT to transform urban mobility and promote cooperation for improved road safety is highlighted by this study. The Vehicle Monitoring System, which is based on technology innovation, community involvement, and regulatory backing, pushes cities toward increased safety, intelligence, and sustainability while putting urban inhabitants' well-being in the center of interconnected transportation networks.

REFERENCES

[1] Abhi Lad, Prithviraj Kanaujia, Soumya, Yash Solanki. "Computer Vision enabled Adaptive Speed Limit Control for Vehicle Safety." *International Conference on Artificial Intelligence and Machine Vision (AIMV)*. IEEE, 2021. pp.1-5.

[2] Ardit Lushi, Dani Daas, Muhammad Nadeem. "IoT-Based Public Transport Management System." *IEEE Global Conference on Artificial Intelligence and Internet of Things (GCAIoT)*. IEEE, 2022. pp.69-75.

[3] Arman Syah Putra, Slamet Heri Winarno, Amat Damuri, Rio Wirawan, Nurul Aisyah, Nur Kumala Dewi, Ridwan Anwar, Ahmad Ishaq, Jeffry Latumahina. "Application of Digital Map Technology for Selection of the Fastest Path in the Smart City of Jakarta." *International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*. IEEE, 2021. pp.369-374.

[4] B. Pavithra, S. Suchitra, P. Subbulakshmi, J. Mercy Faustina. "RFID based Smart Automatic Vehicle Management System for Healthcare Applications." *3rd International conference on Electronics, Communication and Aerospace Technology (ICECA)*. IEEE, 2019. pp.390-394.

[5] Deepthi Shetty, Prateek G, Onkar Bharatesh Kakamari, Sarojadevi H, Devidas, Mitul Jain. "IoT enabled Vehicle Tracking and Monitoring System." *IEEE North Karnataka Subsection Flagship International Conference (NKCon)*. IEEE, 2022. pp.1-5.

[6] G. Karthy, P.Sai Hemanth, M.Venkata Rahul, S.Venkata Vamsi. "Vehicle Speed Monitoring and Auto Light Adjust System With Node MCU." *2nd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*. IEEE, 2022. pp.459-462.

[7] Haiyun Yao, Xinlian Liang, Ruizhi Chen, Xiaochen Wang, Hanwen Qi, Liang Chen, Yunsheng Wang. "A Benchmark of Absolute and Relative Positioning Solutions in GNSS Denied Environments." *IEEE Internet of Things Journal, vol.11, no.3*. IEEE, 2024. pp.4243-4273.

[8] Mingwei Liu, Rui Zhang, Aoyu Gong, He Chen. "Optimizing Age of Information in Wireless Uplink Networks With Partial Observations." *IEEE Transactions on Communications*. IEEE, 2023. pp.4105-4118.

[9] Jinna Hrudaya, Rithika Rao. D, Bukka Vinoothna, D. Krishna. "Practical Implementation of Image Processing based Vehicle Speed Monitoring System." *6th International Conference on Communication and Electronics Systems (ICCES)*. IEEE, 2021. pp.1138-1143.

[10] Michael Jones M, Swathi J, Jagan T, Naveen T, Kathiresan T, Aravinthan N. "Advancements in Human-Machine Interaction: Exploring Natural Language Processing and Gesture Recognition for Intuitive User Interfaces." *International Conference on Sustainable Emerging Innovations in Engineering and Technology (ICSEIET)*. IEEE, 2023. pp 201-205.

[11] N. Shantha Kumar, Deepak Kumar Sinha. "Manual LED Adjustment Vintage Airspeed Monitoring without Microcontroller." *11th International Conference on System Modeling & Advancement in Research Trends (SMART)*. IEEE, 2022. pp.666-670.

[12] Saurabh S. Chakole, Neema A. Ukani, Sohail Sheikh. "GPS and GSM Enable Tracking, Monitoring and Control system for Multiple Application." *International Conference on Smart Systems and Inventive Technology (ICSSIT)*. 2019. pp.1027-1030.
—. "GPS and GSM Enable Tracking, Monitoring and Control system for Multiple Application." *International Conference on Smart Systems and Inventive Technology (ICSSIT)*. IEEE, 2019. pp.1027-1030.

[13] Shahriar Rahman Fahim, Saquib Shahriar, Omar Kamrul Islam, Md. Ilias Rahman, Subrata K. Sarker, Shahela Akter. "Development of a Remote Tracking Security Box with Multi-Factor Authentication System Incorporates with a Biometric Sensing Devic." *International WIE Conference on Electrical and Computer Engineering* . IEEE, 2019. pp 1-4.

[14] V. Meenakshi, D. Godwin Immanuel, M. Kavitha, V. Vijeya Kaveri, C.R. Edwin Selva Rex. "Automatic Speed Controller of Vehicles Using Arduino Board." *International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*. IEEE, 2021. pp.1-5.

[15] Venkata Phanikrishna B, Suchismitha chinara. "Time Domain Parameters as a feature for single-channel EEG-based drowsiness detection method." *IEEE International Students' Conference on Electrical,Electronics and Computer Science (SCEECS) pp.1-5*. IEEE, 2020. pp.1-5.

A Car Parking Guidance System to Lessen the Traffic in the Parking Lot using Arduino

SrirengaNachiyar V
dept. Electronics and Communication Engineering
Ramco Institute of Technology,
Rajapalayam, India
srivenkatraman4@gmail.com

Chandralekha R
dept. Electronics and Communication Engineering
Ramco Institute of Technology
Rajapalayam, India
radhachandra2013@gmail.com

Kasi Raman R
dept. MBA
P.S.R. Engineering College
Sivakasi, India
rmn.kasi@gmail.com

Rohan Riyas A
dept. Electronics and Communication Engineering
Ramco Institute of Technology
Rajapalayam, India
953621106083@ritrjpm.ac.in

Shriram S
dept. Electronics and Communication Engineering
Ramco Institute of Technology
Rajapalayam, India
953621106099@ritrjpm.ac.in

Parthipan J
dept. Electronics and Communication Engineering
Ramco Institute of Technology
Rajapalayam, India
953621106073@ritrjpm.ac.in

***Abstract*— This research study proposes a new technique to provide autonomous driver directions for vehicle parking. The dashboard LCD screen will show the lot number and the precise position of the open space where the automobile has to be parked. The proposed work features include an LCD that will show the entire number of automobiles that can be parked, a parking space that is free, and dynamic parking that avoids traffic. It can be utilized to cut down on congestion inside the parking lot and prevent wasting time. This results in reduced fuel usage and increased parking efficiency. The proposed work is delivered in an economical way. The proposed work is to develop a parking guidance system using Arduino.**

***Keywords*—LCD, Arduino, Lot number, Automobile, Parking efficiency.**

I. INTRODUCTION

The quick rise in vehicle usage is the cause of the traffic issue in a parking lot. This issue may mostly arise in developed cities. People used to go to tourist destinations, theaters, shopping centers, and hospitals. In addition to wasting time and causing traffic congestion, it also raises the risk of sound and air pollution. Consequently, a vehicle parking management system is required. The proposed work primary goal is to restrict parking, which is necessary to address these issues. A microcontroller can be used to control the system. Numerous technologies are currently in use in the parking system, including WSN, RFID, E-parking, conveying device belts, and smart parking options. Car parking guidance systems are an alternative approach to address the drawbacks of each technology. Only the number of open slots is displayed by RFID and WSN technologies, and only authorized users may utilize RFID. The cost of initiation and upkeep for smart parking solutions is substantial. Conveyor belt power usage is high. Therefore, it is not advised, given the current state of the world's energy shortage. This proposed work uses an infrared sensor to pinpoint the precise parking spaces for each vehicle. The parking spots

each contain an infrared sensor. The microcontroller is used to meet the requirements and acts as a programming tool for the entire system, which lowers costs associated with employment opportunities and offers excellent security. This approach can lead to flexibility, speed, and compliance with market demands. It is possible to reduce the tension that comes with traffic congestion. The way that space is used in a parking lot can be improved. The purpose of reactive parking is to divert traffic from within the parking lot. For instance, some users are given access to exclusive slots. The slot name will be automatically removed from the buffer if the user parks their car in the designated spaces. Errors may occur when an automobile is parked in the designated spots. In that scenario, the name of the slot the car is occupying can be automatically erased from the buffer.

II. EXISTING METHODOLOGY

The related work in this section is predicated on current technologies. The three current technologies are E-Parking, RFID, and conveyor belts. With RFID technology, an authorization card is given to the user. The parking area's entrance is where the RFID reader is located. Only approved users are permitted entry into the parking lot in order to park their cars. An authorization card is not available for each user, and the unauthorized user is unable to park their automobile. These are a few of the RFID technology's shortcomings. The mechanism that raises the automobile automatically into the designated slots is called a conveyor belt. The primary goal of this concept is to maximize space efficiency. Pallets are numbered on a conveyor belt. The huge vehicle is raised using the pallets. As a result, raising the car requires more power, and if multiple cars arrive at once, it may also result in lifting an empty pallet. As a result, it uses unnecessary power. Since there is a shortage of electricity energy in our nation, it is not advised due to increased power usage. Another idea that works well in large cities as an income stream is e-parking. It is employed to lessen traffic and improve the parking system's efficiency and safety. However, in order for this system to obtain data regarding the available spaces within the parking area, a traffic server is required. The primary disadvantage of the E-Parking system is its high initial and on-going costs when compared to alternative technologies.

III. PROPOSED METHODOLOGY

The components used for designing the car parking guidance system are Arduino (ATmega328Pmicrocontroller) 16*2,LCD, TCRT5000, Proximity IR sensor and Servo motor. All the hardware components were connected to the Arduino and it is shown in figure 1.

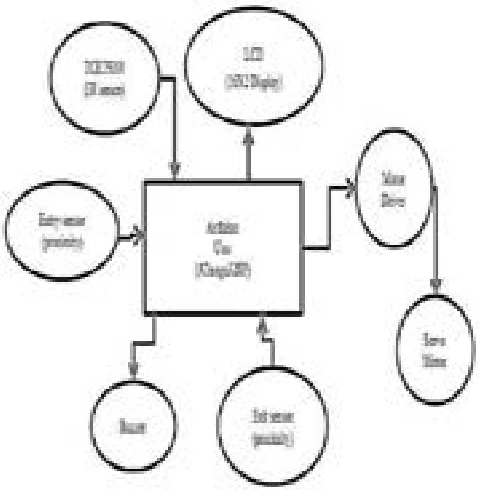


Fig. 1. Hardware connection with Arduino

The vacant spot monitor system of this car parking guiding system design is made up of an LCD display, an ATmega328P microprocessor, and an IR sensor. An infrared sensor installed in each parking space allows for the monitoring of the parking area. Here, the TCRT5000 IR Sensor is being used. Compared to other sensors, it has a unique future since it blocks visible light. The microcontroller receives information from an infrared sensor. Interface between the microcontroller and LCD display. It is employed to show details regarding available space within the parking lot. An infrared sensor installed in each parking space provides this information. Another surveillance system installed on the gate side. It keeps an eye on cars coming and going from the parking lot. The cars are detected using proximity sensors. The total number of cars in the parking lot is shown on an LCD monitor. Every time a vehicle is detected by the entry sensor, the count goes up, and every time a vehicle is detected by the exit sensor, the count goes down.

A. Microcontroller:

Here, an ATmega328P microcontroller is being utilized. It is an AVR-based low power CMOS 8-bit microcontroller that improved RISC architecture. Its throughput is almost one millisecond per megahertz. A microcontroller can be applied to both common and unique tasks. These days, office equipment, remote controls, and automobile mobile engine control systems all make extensive use of microcontrollers. Its primary purpose is to attain high system performance. In terms of efficiency and affordability, the ATmega328P microcontroller is a suitable option for this system.

B. TCRT5000

The reflecting sensor TCRT5000 has a photo transmitter and an infrared emitter mounted in a straight line inside a leaded package. It prevents light from reaching it. This sensor is utilized to identify any vacant spots in the parking lot and is positioned in every single slot. Because it filters out visible light, it generates identical output during the day and at night.

C. Proximity IR-Sensor

Without making physical contact, a proximity sensor may identify the presence of things in the immediate vicinity. There are two components to an IR sensor. One component is an infrared transmitter, or IR transmitter for short. It functions similarly to an LED and emits infrared light. Receivers are additional parts that collect the infrared radiation that the transmitter emits. One crucial thing to remember is that both pieces need to be arranged in a line that is straight. When the signal is high, the IR transmitter receives it and sends it to the receiver. A comparator is attached to the infrared receiver. The comparator is supplied the receiving signal. Comparator compares the operating voltage to the input voltage; if the returned voltage is higher than the comparator's operating voltage, both the receiver and the IR transmitter rays are disrupted, which causes the LED to glow (ON). The IR signal is not affected if the reference voltage and the comparator's operational voltage are equal. LED will turn off. This proximity infrared sensor is located at the parking area's entrance.

D. Servo Motor

The majority of the time, high-tech industrial applications like automation technologies use servo motors. It is an independent electrical device that precisely and efficiently rotates the machine's parts. It is utilized to open and close the gate in this proposed work. The gate automatically opens and closes when both the entry and exit sensors detect the presence of a vehicle. The servo motor with microcontroller interface. The angular direction of a servo motor distinguishes it from other types of motors. Closed-loop control systems can use it. For a 1ms variation in pulse width, the servo rotates 90 degrees.

E. LCD Display

An LCD display is positioned both within and outside the parking space. It has a microcontroller interface. It shows the overall capacity at the parking lot's entrance as well as information regarding available space within the parking area. Before approaching the parking space, drivers have the ability to check its current condition.

IV. IMPLEMENTATION

After the text edit has been completed, the paper is ready for the template. Duplicate the template file by using the Save As command, and use the naming convention prescribed by your conference for the name of your paper. In this newly created file, highlight all of the contents and import your prepared text file. You are now ready to style your paper; use the scroll down window on the left of the MS Word Formatting toolbar.

A. Module-I

It is suggested that Module-I show the available space in every parking place. This information is displayed using LCD in this instance. The TCRT5000 IR Sensor is used to detect the empty space in slots. The car is detected by the IR sensor, which displays the information on the LCD. The slot will show as occupied by a vehicle if the IR Sensor picks up the car. The figure 2 shows the flow chart of identifying the free slots in the parking area.

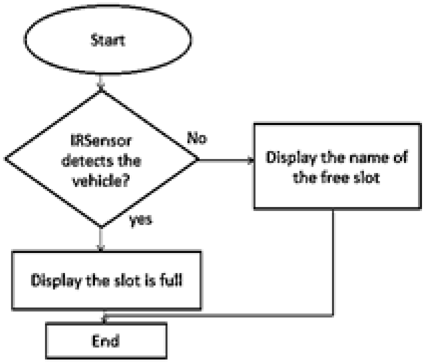


Fig. 2. Flow chart of identifying free slots in the parking area.

The figure 3 shows that the slots 'A' is free

Fig. 3. Slot 'A' is free

The figure 4 shows that the slots 'A & B' is free

Fig. 4. Slots A and B is free.

B. Module-II

The information regarding the total number of open spaces throughout the parking area forms the basis of Module II. Both the parking area's entrance and exit have proximity sensors. The count is increased by one if the entrance proximity initially detects the car. As a result, the count shows that an automobile has entered the parking lot. The LCD shows the remaining space in the parking lot. The count is decreased by one if the exit sensor detects something at first.This signifies that the car is pulling out of the parking lot. The parking area barrier is opened and closed based on the count value. If the count value is greater than the parking area's capacity, the barrier will close. If a vehicle is detected by the entry sensor, the barrier will immediately open. A servo motor is holding these operations. The figure 5 shows the identifying the available number of slots.

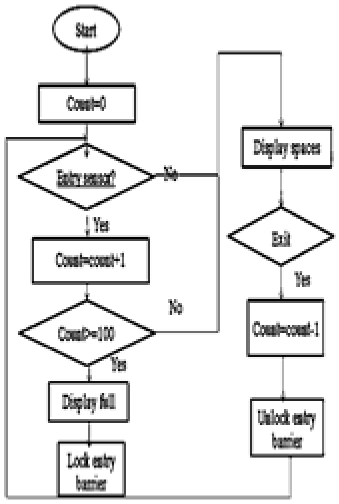


Fig. 5. Flow chart for the identifying the available number of slots.

It demonstrates that the count value increases by one each time the entry sensor detects a vehicle. The LCD indicates that there is no room within the parking area and the servomotor locks the barrier if the count value is more than the parking area's total capacity. The servo motor unlocks the barrier if the exit sensor detects a vehicle, causing the count value to drop by one. The figure 6 shows number of two vacancies inside the parking area.

Fig. 6. Shows that there is two vacancies were available inside the parking area.

The figure 7 shows there is only one vacancy inside the parking area.

Fig. 7. shows that there is only one vacancy inside the parking area.

The figure 8 shows there is no free space inside the parking area.

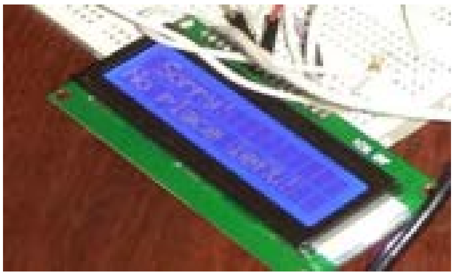


Fig. 8. represents there is no free space inside the parking area.

C. Module-III

The third module deals with dynamic parking. By moving the car to the other side of the slots, dynamic parking helps to minimize traffic inside the parking area. Two controllers can be interfaced to do this. Controllers are interfaced using the I2C protocol. One controller assumes the role of a slave, while the other assumes the role of a master. The controller, also known as the slave, is what retrieves the data regarding available parking spaces for cars. The LCD display used by the slave controller shows the information. The controller that counts the number of cars in the parking lot is called the master. The master was able to read data from the slave and provide it to the buffer by connecting these two controllers. This indicates that the LCD attached to the slave controller contains data regarding available slots. A master reads this data and stores it in a buffer. Here, the buffer serves to give each user a dedicated slot. Every user is given a specific slot, and if they park their car in that space, the slot name is immediately erased from the buffer. The primary purpose of dynamic parking is to automatically add free slots to the buffer based on slot occupancy and remove occupied spaces from the buffer automatically. It eliminates the need for monitoring and lowers the demand for human resources. The figure 9 shows Interfacing two arduino using I2C protocol.

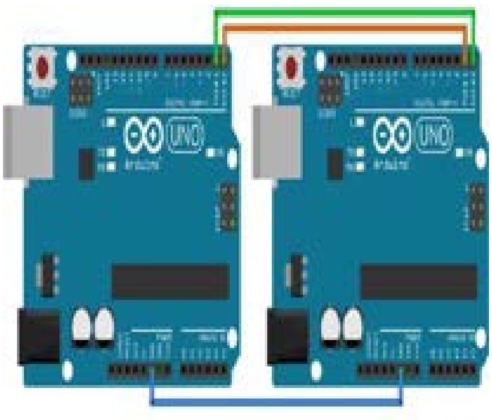


Fig. 9. Interfacing two arduinos using I2C protocol.

V. FUTURE SCOPE

The lack of speed and precise vehicle identification by IR sensor is the future potential for auto parking implementation. Consequently, a mechanical car detecting sensor or another alternate option may be employed. IOT and cloud computing are two more cutting-edge technologies that allow users to reserve parking spaces in advance for their own cars.

VI. CONCLUSION

The proposed work has been completed effectively. Upon entering the parking lot, the user is assigned a parking space and has access to information regarding the quantity of available spaces within the parking area. This causes the revenue from parking spaces to rise quickly. The environment is a key factor in this process. By reducing traffic, the pollution that vehicles' fuel use causes to the environment can be mitigated. Therefore, the purpose of the parking advice system is to help users save time and avoid traffic.

REFERENCES

[1] Chaniotakis, E., Pel, A.J., 2015."Drivers parking location choice under uncertain parking availability and search times": A stated preference experiment.Transp. Res. Part A 82, 228-239.

[2] Trista Lin, Herve Rivano and Frederic,Le Mouiel,"A survey of Smart Parking solutions", IEEE Transactions on intelligent transportation systems, 2017.

[3] M. Chester, A. Fraser, J. Matute, C. Flower, and R. Pendyala, "Parking infrastructure: A constraint on or opportunity for urban redevelopment?" , vol. 81, no. 4, pp. 268–286, IEEE,2015.

[4] A. Bagula, L. Castelli, and M. Zennaro, "On the design of smart parking networks in the smart cities: Optimal sensor placement model," Sensors, vol. 15, no. 7, pp. 15443–15467,IJAREEIE, 2015.

[5] J. Krieg, G. Jakllari, H. Toma, and A. Beylot, "Unlocking the smartphone's senses for smart city parking," in IEEE Intl Conf on Communications, 2016, pp. 1–7.

[6] Benenson, I., Martens, K., Birfir, S., 2008. PARKAGENT:"an agent-based model of parking in the city. Comput. Environ". Urban Syst. 32, 431–439.

[7] Cao, J., Menendez, M., 2015. "System dynamics of urban traffic based on its parking-related-states".Transp. Res. Part B 81 (3), 718–736.

[8] Fiorenzo-Catalano, S., Van der Zijpp, N.J., 2001. "A forecasting model for inland navigation based on route enumeration". In: Proceedings of the European Transport Conference. Cambridge, United Kingdom.

[9] Geng, Y., Cassandras, C.G., 2012. "A new smart parking system infrastructure and implementation". Procedia Social Behav. Sci. 54, 1278–1287.

[10] Geroliminis, N., 2015. "Cruising-for-parking in congested cities with an MFD representation". Econ. Transp. 4 (3), 156–165.

[11] Guo, L., Huang, S., Zhuang, J., Sadek, A.W.,2013." Modeling parking behavior underuncertainty: A static game theoretic versus a sequential neo-additivecapacity modeling approach". NetwSpatial Econ. 13 (3), 327–350.

[12] Aruna, Kaleeswari, SrirengaNachiyar presented paper titled "Car Parking Guidance system using Arduino", International Conference on Engineering and Advancement in Technology on 22.03.2018 at Sri Krishna College of Technology.

Enhancing User Interaction: Gesture-Enabled Virtual Cursor with Voice Integration

Dr.V.Anjana Devi
Computer Science and Engineering
Rajalakshmi Institute of Technology
Chennai, Tamil Nadu, India
anjanadevi.aby06@gmail.com

Jahnavi E
Computer Science and Engineering
Rajalakshmi Institute of Technology
Chennai, Tamil Nadu, India
jahnavi.e.2020.cse@ritchennai.edu.in

Kavipriya R
Computer Science and Engineering
Rajalakshmi Institute of Technology
Chennai, Tamil Nadu, India
kavipriya.r.2020.cse@ritchennai.edu.in

***Abstract*—The research article proposes a novel approach combining voice control with gesture-enabled virtual mouse technology to enhance human-computer interaction. Using this cutting-edge technology, which uses computer vision techniques to identify hand landmarks, users may control cursor motions using gestures. However, a number of problems with the current methods, such as snapshot algorithms, cloud-based voice assistants, and hand contour extraction, restrict their effectiveness. The accuracy of cursor control may be impacted by challenges in precisely interpreting user gestures due to the intricacy of gesture mapping and a restricted gesture language. Furthermore, the voice assistant component of the system is not robust enough to handle a variety of speech instructions and dynamic voice commands, and it is also dependent on an internet connection. Additionally, there aren't many error handling capabilities, which might lead to worse-than-ideal user experiences when orders aren't carried out precisely. Notwithstanding these difficulties, the proposed system incorporates voice assistant features beyond mouse control, enabling users to utilize mouse motions and natural language instructions to do a variety of activities, including sending emails, opening apps, executing mouse functions, searching the web, and managing files. To fully realize the potential of this integrated voice and gesture-controlled system in improving user experiences and accessibility, it will be imperative to address these issues through enhanced gesture vocabulary, expanded gesture recognition algorithms, and improved voice command processing.**

Index Terms*—Human-Computer Interaction, Gesture-Enabled Virtual Mouse, Voice Assistant, Computer Vision, Snapshot algorithm, Cloud-based voice assistant.

I. Introduction

The advancement of digital technology heavily relies on human-computer interaction (HCI), which serves as a bridge between users and machines through intuitive interaction methods. The combination of gesture-enabled virtual mouse technology and voice control marks a new era in HCI, potentially revolutionizing how users interact with digital interfaces. This integration signifies a departure from traditional input methods, providing users with a seamless and instinctive way to engage with technology. By utilizing natural gestures and voice commands, this integrated system has the potential to greatly improve accessibility, efficiency, and user satisfaction across various computing environments.

In today's fast-paced world, where accessibility and productivity are of utmost importance, voice recognition technology has proven to be invaluable. Initially developed to assist individuals with disabilities or physical limitations, voice recognition software offers a faster and more efficient method of inputting information. By allowing users to express their thoughts and instructions verbally, this innovative technology eliminates the need for traditional typing and enables natural and effortless communication.

From performing tasks such as setting reminders with a simple voice command to navigating virtual environments using gestures, these technologies empower users to multitask and interact with devices seamlessly. Furthermore, their introduction represents a significant advancement in HCI, propelling innovation in augmented reality and artificial intelligence. As a result, gesture-controlled virtual mouse systems and voice assistants have become essential components of modern computing, reshaping the way we interact with technology and unlocking limitless possibilities.

The amalgamation of computer vision, machine learning, speech recognition, and natural language processing (NLP) technologies serves as the foundation of cutting-edge interactive interfaces. By utilizing computer vision, the system decodes user gestures captured through cameras or sensors, facilitating seamless manipulation of on-screen cursors. Machine learning algorithms play a crucial role in gesture recognition, elevating precision and adaptability to user actions. Furthermore, speech recognition technology transforms spo- ken commands into text, enabling verbal communication with the interface. Advanced NLP methodologies further enrich user engagements by comprehending and analyzing natural language inputs.

The Gesture-Enabled Cursor with Voice Integration system signifies a notable progression in Human-Computer Interaction (HCI), offering enhanced user experiences and broader accessibility across digital platforms. This integration harbors transformative capabilities, influencing the future landscape of human-computer interaction.

II. Literature Survey

In the past, contact-based and noninvasive techniques have been used for hand gesture identification. Data gloves were employed in former contact-based techniques to detect and track hand motions. Unfortunately, because of their heaviness, connectivity, and detector locations, these gloves were frequently unwieldy and interfered with normal hand movements. Numerous studies have concentrated on employing various methods to control computers, like classifying human activities and assigning matching mouse instructions to drive the computer pointer, using sensors called Kinect [17], EEG mice, or EMG signals [8]. These techniques, however, frequently require extra gear that is costlier and larger than a standard computer mouse.

Virtual monitor [7]Using this technology, a virtual monitor in a simulated world is controlled by a Kinect camera that records human activities. After capturing physical characteristics, it builds a virtual monitor by reconstructing a skeleton image using joint coordinates and depth information. Through the use of a transformation process, coordinates from this virtual display are translated to the actual one. These translated coordinates serve as the basis for mouse operations, which improve accuracy. This Kinect-based virtual monitor system can be difficult and time-consuming to set up and calibrate. To guarantee optimal performance, some software setups and modifications might be necessary.

SIFT characteristic-based modified backpropagation of artificial neural networks (BP-ANN) [18]. This technique uses the Scale-Invariant Feature Transform (SIFT) algorithm for trajectory tracking after Gaussian Laplacian-based blob detection for hand modeling. In order to classify hand gestures for controlling household appliances, an artificial neural network (ANN) is trained using datasets. In order to learn complex descriptor associations, training a BP-ANN model with SIFT features requires a large amount of annotated data. This training procedure can be time-consuming and may necessitate significant labeled datasets.

Colored caps concept [20], the process of color detection entails locating colored finger caps within picture frames. In grayscale, color caps are highlighted by intensity differences, which first identify color pixels. After that, rectangles are made around caps using midpoint formulae for tracking. Tracked coordinates serve as the basis for gesture recognition. However, disadvantages include backdrop interference, color dependence, and limited color selection.

Mobiring[21]After filtering the signals with a translation equation, mistakes are corrected via Zero Velocity Compensation (ZVC). During gesture identification, accelerations don't change if there is initial stillness. A Finite State Machine (FSM) modifies its state when the moving ring satisfies certain requirements. Mobiring works best when the ring-mounted sensor is calibrated precisely. Erroneous calibration may result in inaccurate gesture identification and virtual mouse operation.

Hidden Markov Model [22] Skin color identification comes after object regions are isolated using skin pixel thresholding. Hand gesture computation is aided by edge identification using the Kirsch technique. Algorithms with convex hulls facilitate tracking. Hardware is implemented using a MEMS accelerometer. In order to accurately recognize patterns and gestures, Hidden Markov Models (HMMs) need large amounts of labeled data. It might take a lot of time and resources to gather and annotate this data.

HAAR and CAMSHIFT algorithm [11] After the HAAR algorithm separates hands from objects, CAMSHIFT uses Canny Edge detection and HSV skin to track hands. While convex-hull detection is used for various functions, displacement detection is the basis for mouse pointer movement. Preset textures and edge gradients are used by HAAR classifiers to identify objects. These characteristics might not always translate well between various hand motions and contexts, which would restrict their flexibility.

ANN-based adaboosting technique [16] Due to this technique, gesture-based navigation is made possible on huge TV screens. Four modes are available: input, scroll, zoom, and mouse. Using an ANN and adaboosting, hand images are divided into three groups for gesture recognition. Using early-stopping and cross-validation, a 2-layer feedforward network with log-sigmoid and linear transfer functions is optimized. Predefined gesture modes could reduce the system's efficacy and make it more difficult to adjust to nuanced or unique motions.

Snapshot algorithm [12] used for locating and perceiving hand coordinates with the help of Mediapipe library. Instructive algorithmic rule is used to pass over the coordinates from the camera to the screen. When faults or unforeseen circumstances arise, including occlusions, restricted mouse movements, or partial hand visibility, the Snapshot method may not be able to handle them well.

Finger Recognition [3] Images are segmented using Otsu's thresholding, and precise identification is guaranteed by morphological filtering. On a virtual keyboard, finger counting activates taps or color caps. Accurate segmentation might be affected by background noise. Predefined gestures in the system restrict the diversity of interactions, which makes creative tasks more difficult for users to complete than preset motions.

Eye and Face gestures [4], In this case, the user will be able to perform mouse functions using eye and face movements. The cursor on the screen can be controlled by a specific point on the user's face, while the right and left clicks can be executed by winking the right and left eyes, respectively. Additionally, squeezing the eyes would enable the scroll function, which is particularly useful when reading PDFs and other documents. Users may become weary or experience discomfort if they employ eye and face motions continuously for lengthy periods of time. Extended strain on the muscles of the face or eyes might affect usability and user acceptability.

[14] With the help of computer vision and pattern recognition, hand gestures are investigated for human-machine interaction. Without segmenting or using motion history, the approach uses HT to trace gesture trajectories. SIFT features and the integration of artificial neural networks translate motions into commands that can be customized for various users. On the other hand, hand direction, speed, and motion variations may cause performance problems.

Voice assistants offer a novel and exceptionally user-friendly approach to engaging with devices and online services [15]. Cloud-based voice assistant [10] Raspberry Pi is used for receiving the input, Google Dialogflow for providing intelligence, PyTorch for understanding, Flask to build the RESTful API, and Redis as a message queue and cache. Dataset training is done using joint classification and model pruning. For the system to interact with cloud services such as Google DialogFlow, it needs a reliable internet connection. Intermittent problems with internet access may cause the voice assistant to become inoperable.

IFTTT platform [9]This project uses webhooks to connect to a Raspberry Pi and incorporates Google Home Mini and Amazon Echo Dot services for voice-controlled lighting. Reliance on IFTTT, Google Home, and Amazon Echo creates possible sources of error. Even though a lot of services, like IFTTT, provide free tiers, there could be extra charges when using premium features or ramping up use. Evaluation of long-term cost factors should be done in light of requirements and use. This [5] virtual mouse interface was designed to identify finger tips and recognize hand motions. This study employed two distinct finger tracking strategies: using colored caps and hand gesture recognition. Skin tone variances, backdrop clutter, and changes in lighting can all have an impact on how successful skin color-based systems are. Hand motion detection accuracy and dependability can be impacted by environmental conditions.

The study [19] utilized a dynamic skin color system and an innovative technique for controlling the identification of historical photos through palm movements. A method of choice enables the identification of a particular area that corresponds to a single hand, which is thereafter monitored for the duration of the video clip. A precise spot on the virtual interactive interface is assigned to the outcome. Through seamless integration with the surrounding environment, the Direct System Integration (DSI) engine mimics the capabilities of a mouse. Like a real keyboard and mouse, this mouse can operate a variety of apps, including a visual keyboard. The Interface module is not a separate entity from the Mouse module, which converts the mouse's location into screen coordinates. On the other hand, the Interface Control (IC) engine offers a different approach. It allows for the control of a specific interface, enabling it to interpret and execute actions accordingly. However, the primary limitation of the research lies in its failure to accurately detect intricate hand movements.

The existing system [1][2] is comprised of five primary elements. Initially, skin detection transforms images into the HSV color space to differentiate skin pixels from non-skin pixels. Subsequently, Hand contour extraction isolates the hand contour from the background using skin detection. Thirdly, using hand tracking, the cursor is maneuvered by identifying the index fingertip and sending the location of the finger on the screen to the computer via a coordinate system. By identifying hand landmarks and assessing finger states, hand movement is observed. In addition, gesture recognition, employing OpenCV, computes hand contour convexity defects, recognizes fingertips, and tallies the number of available fingers for function execution. Finally, cursor control manages cursor movement. Nevertheless, the system encounters precision constraints and difficulties in recognizing specific mouse functions. Furthermore, it cannot effectively identify multiple mouse functions. There is a limit to the system's capacity to recognize and carry out certain mouse operations, depending on various gestures. This limits the variety of ways in which the virtual mouse interface is used. Skin tone variations, background clutter, and illumination adjustments can all affect how effective skin color-based systems are. Environmental factors may affect the accuracy and reliability of hand motion detection.

III. Proposed System

The combination of a gesture-enabled virtual mouse and a voice assistant creates a sophisticated system that merges real-time hand gesture recognition with natural language processing capabilities, providing users with an intuitive and hands-free interaction experience. The gesture-controlled virtual mouse relies on advanced computer vision techniques, utilizing the MediaPipe library to identify and monitor landmarks on the hand in live video streams. These landmarks act as key points for recognizing predefined hand gestures, allowing users to execute activities such as moving the cursor, clicking, and scrolling through intuitive hand movements.

To identify specific gestures based on predefined criteria, the gesture recognition module utilizes a custom logic that involves three classes for managing the camera, converting handmarks, and executing corresponding actions. By examining the relative positions and movements of important landmarks like fingertips, knuckles, and palm orientation, the system can accurately interpret user gestures.

In addition to the gesture-controlled interface, the voice assistant module enhances the interaction by enabling users to issue commands and queries using natural language. To allow voice-controlled interaction and command execution, the script combines NLP methods like speech recognition, keyword matching, and context management with external libraries such as speech recognition,pyttsx3. The first stage of natural language processing (NLP) is speech recognition, which transforms spoken words from audio input into text. The speech recognition library is used in the script. Another fundamental NLP method called keyword match- ing is utilized to pinpoint certain words or phrases in the recognized text (voice data). The script looks for specific phrases (such as "hello," "time," and "search") in order to initiate pre-programmed actions or reactions. This is a basic type of pattern matching that uses key words to infer user intent. Tasks such as opening applications, searching the web, executing mouse functions, or retrieving information from databases or APIs are executed. Fig. 3 shows the overall flow of the system.

A. *Gesture Controlled Virtual Mouse*

The necessary Python libraries, utilized for different tasks, are imported. OpenCV is used for computer vision tasks,

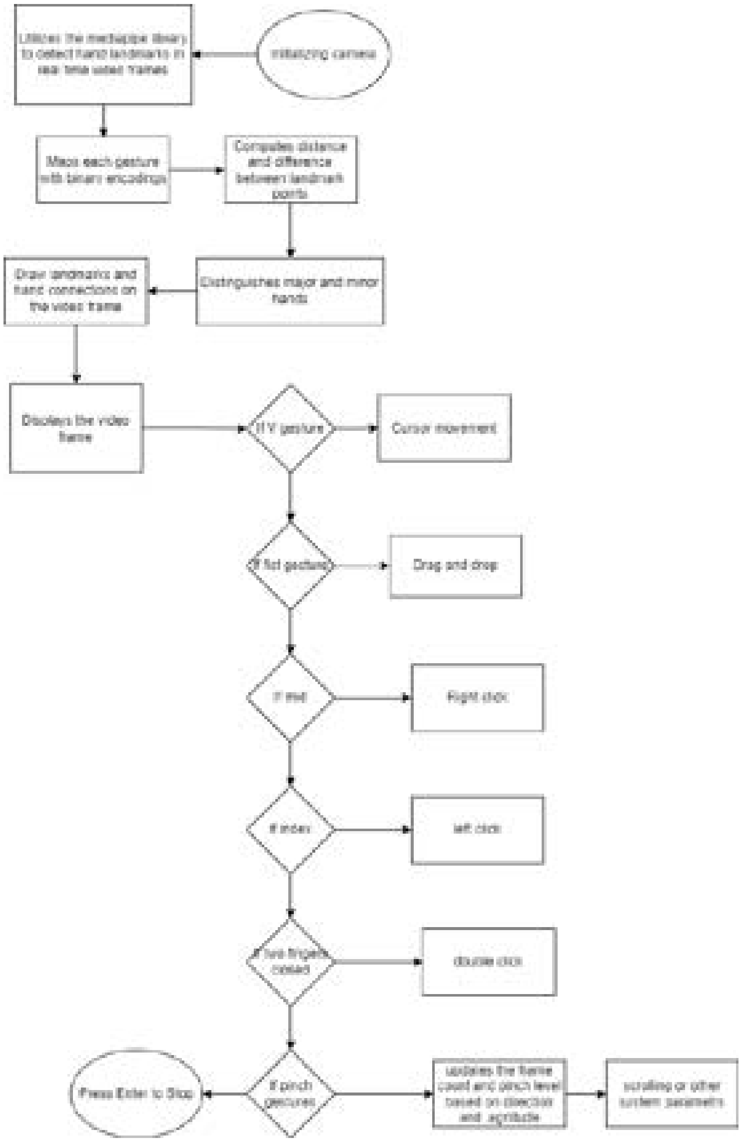


Fig. 1. Workflow for Virtual Mouse

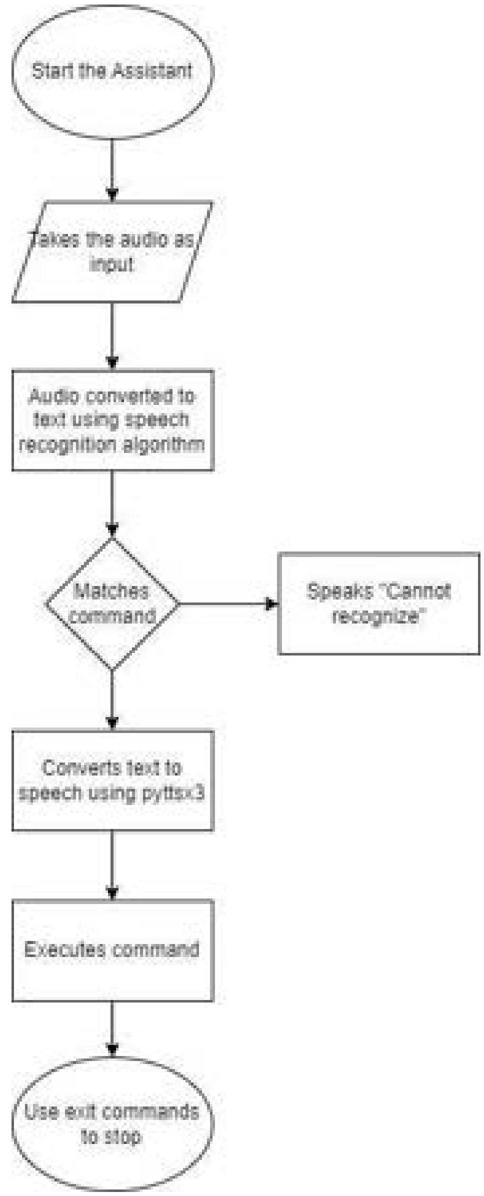


Fig. 2. Workflow for Voice Assistant

MediaPipe is employed for hand tracking, PyAutoGUI enables control over the mouse cursor, and Math is utilized for mathematical operations. Enumeration is used for defining enumerations, types, and comtypes, which facilitate low-level system interactions. Additionally, paw is employed for controlling system audio, and screen brightness control is utilized for adjusting screen brightness. Three classes are defined for the virtual mouse to function: HandRecog, Controller, and Gesture Controller. The Gesture Controller class manages the camera, finds landmarks, and serves as the system's initialization point. The HandRecog class is for converting mediapipe landmarks into recognizable gestures. The Controller class carries out orders based on gestures that are identified. Initially, each hand gesture is mapped to a binary number using an enum. The value of each enum component corresponds to a binary representation that contains particular hand positions or finger configurations. For example, 0b00000 represents FIST (all fingers closed), whereas 0b11111 represents PALM (all fingers extended). Then, to distinguish between the major and minor hands, when multiple hands are detected, two members, MAJOR and MINOR, are assigned the values '1' and '0'. The gesture recognition system uses these enums to represent and analyze hand labels and movements that are recognized.

a) Gesture Recognition Module: Initially, the video frames are captured and undergo color conversion. Subsequently, the modified image is utilized to detect hand land- marks. Upon receiving the hand landmarks from MediaPipe, the classification between dominant and non-dominant hands occurs. Then gestures are recognized based on the tracked hand using the logic defined in the HandRecog module. Depending on the recognized gesture, controls are handled by the controller module.

b) HandRecog Module: Hand tracking and landmark detection are the most important processes for this virtual mouse to function. Firstly, the attributes are initialized to represent the computed gesture for the current frame, to store the current and previous gestures, to track the number of frames, and to hold the landmark data obtained using the Mediapipe library. After obtaining the hand landmarks using the Mediapipe library, the finger positions are analyzed using the following logic: Initially, the signed Euclidean distance between two landmark points on a hand based on their vertical (y-coordinate) positions relative to each other is calculated using Eqn(1). Moreover, the interval in 2D space between the two spots is calculated. This calculation helps determine the direction of movement between the two points. To comprehend the movement of the hand in three-dimensional space, the absolute difference on the z-axis between the two points is calculated using Eqn(2). Additionally, the ratio of distances between the finger tip, middle knuckle, and base knuckle is calculated using Eqn(3) to assess the state of the fingers. Iterating over each finger's landmark points, the signed Euclidean distance, absolute difference, and ratios are calculated to find out the state of each finger's whether it is extended or bent. Using bitwise operations, the binary representation of these finger states is detected. If the computed ratio is greater than 0.5, set the least significant bit, indicating an open finger state. If no landmarks are available, it considers the state as closed, which is FIST, and the value is set to 0. After finding out the finger state, the finger attributes are compared with the specified patterns to find out the corresponding pattern. Depending on this ratio and depth, it recognizes gestures such as V GEST, TWO FINGER CLOSED, or defaults to MID based on finger state. It determines a ratio between the lengths of particular finger segments if the finger configuration shows the first two fingers. The gesture is recognized as Gest.V GEST if this ratio is more than 1.7; if not, it further compares the z- coordinate differences between Gest.TWO FINGER CLOSED and Gest MID. Gest Pinch Minor or Gest. Pinch Major, depending on the hand label, is the gesture identified if the finger configuration includes the final three or four fingers and the distance between particular finger landmarks is less than 0.05. To handle the fluctuations and ensure gesture stability, the original gesture attribute will be updated to reflect the stable gesture state if the identified gesture is constant for more than four frames. In real-time applications, our method guarantees dependable gesture identification while efficiently mitigating noise.

$$sign(y_2 - y_1) * \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{1}$$

$$\Delta z = |z_2 - z_1| \tag{2}$$

$$ratio = dist1/dist2 \tag{3}$$

where dist1 - distance between the finger tip and the middle knuckle.

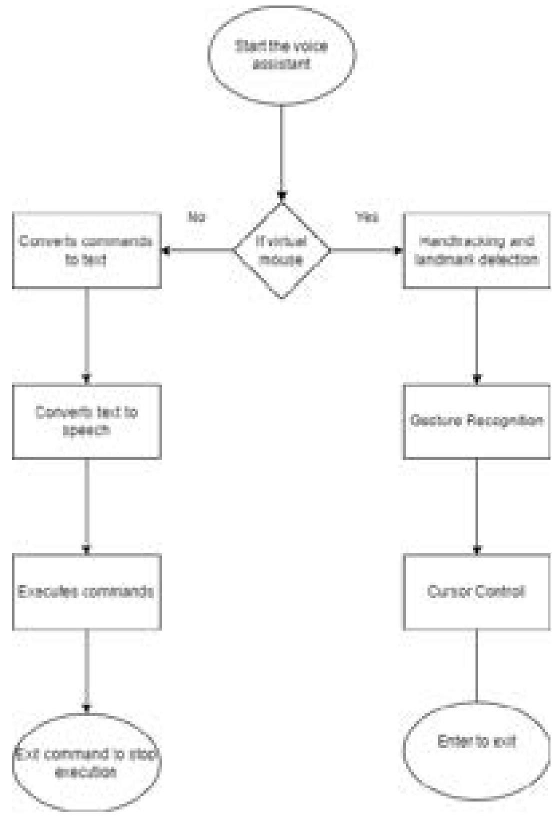


Fig. 3. Workflow for overall system

dist2 - distance between the middle knuckle and the base knuckle.

c) Controller Module: Various recognized gestures trigger corresponding actions, evaluating each gesture to perform the appropriate action based on its type. For example, a V gesture moves the mouse cursor, a fist gesture does drag and drop, and other gestures like MID, INDEX, and TWOFINGERCLOSED are associated with different mouse button clicks or double clicks. The management of PINCHMINOR and PINCHMAJOR gestures is done independently, where the direction and intensity of the pinch are determined by hand landmarks using equations (4) and (5).The vertical or horizontal distance between a reference pinch start point and the current x or y coordinate of a specific hand landmark. The result is returned as a scaled value that can be used to determine the magnitude of a pinch gesture in the vertical or horizontal direction. This approach guarantees seamless and efficient control over horizontal and vertical scrolling, as well as adjustments to systemparameters, by continuously updating the frame count and pinch level. The system brightness and system volume, are adjusted based on the pinch gesture displacement. The Pyautogui library is utilized to simulate scroll wheel movements based on the direction and magnitude of the pinch gesture. The following logic calculates the current cursor location based on hand motion and prior cursor positions in order to minimize any sudden hand movements. The function first identifies a specific hand landmark to use as the cursor position reference within the hand coordinate space. The code then uses pyautogui.size() to get the current screen resolution and maps the hand coordinates to screen pixel coordinates. Moreover, it uses pyautogui.position() to record the cursor's current location. This function improves the accuracy and smoothness of cursor control by dynamically modifying the cursor movement ratio based on the intensity of the detected hand motion.

$$dist = round((start_y - handresult.landmark[8].y) * 10, 1) \tag{4}$$

$$dist = round((handresult.landmark[8].x - start_x) * 10, 1) \tag{5}$$

d) User Interface: The video frame is shown with annotated hand landmarks and connections through OpenCV, offering users a visual representation of the detected hands and gestures. The loop will persist until the user opts to exit by pressing the Enter key, guaranteeing uninterrupted gesture recognition and system interaction. Fig. 1 illustrates the workflow of the virtual mouse.

B. Voice Assistant

a) Initialization: Various modules are being initialized, and variables are being set. The date module from the datetime library is being utilized to retrieve the current date. The speech recognition library is imported, enabling the recognition of speech input, and other necessary libraries are also imported. Another crucial initialization is pyttsx3 for text-to-speech functionality, where the engine is initialized with Microsoft's Speech API (sapi5). Following that, the available voices are obtained, and the engine's voice property is set.

b) Building the bot: Multiple variables are set with predefined values, such as fileexpstatus, files, path, and is awake. These variables are responsible for managing file exploration, storing file details, specifying a file path, and determining the bot's operational status. The script starts with an import of the speech recognition library. This library offers speech recognition capabilities from a variety of sources, including audio files and microphones. The script appropriately manages and releases the microphone resource after use. The script modifies the dynamic energy threshold and energy threshold of the microphone before audio recording. The audio input from the microphone is captured by the record audio() function: It establishes a pause threshold to establish the longest period of quiet before the recording ends. The identified speech is stored in an empty string called voice data, which is initialized. It records sound from the microphone for up to five seconds at a time. It attempts to identify the audio by utilizing Google's voice recognition feature. The record audio() method returns the recognized text (voice data) following the recording and recognition of the speech. The wish() function greets users according to the time of day and presents the bot. Each recognized command is processed by the reply (voice data) function, which then uses the reply (audio) function to provide the required replies. This feature uses the Text-to-Speech library (pyttsx3) to translate text (audio) into speech. For the purpose of monitoring and debugging, print the response to the console (print audio). Error handling is included to address situations where the service is inaccessible or when the audio input is not recognizable.

c) Command Handling: The respond() function oversees voice input for the bot by utilizing global variables to ensure consistency. It initially displays the received voice data, substituting 'Synthia' if found. Subsequently, it incorporates the user's input into the interface. When the bot is not active, commands such as 'wake up' prompt activation and a welcoming message. While active, it reacts to commands like 'hello' and 'what is your name' with suitable responses. It caters to requests for the current date and time, performs web searches, and pinpoints locations on Google Maps. Voice- activated features, such as the ability to start and stop gesture recognition, copy and paste text, navigate through files, and respond to various commands, are available. When specific voice commands like "launch gesture recognition" are detected, the code checks the current status and either starts or stops gesture recognition accordingly. For text manipulation, commands like "copy" and "paste" are executed using keyboard shortcuts. It is feasible to execute mouse functions and interact with Notepad by integrating a speech recognition library with PyAutoGUI. Users can navigate through files by listing them in the root directory, opening folders, and going back. The subprocess module is utilized to initiate applications, and smtplib is used to send email. These features improve user interaction by allowing them to control actions through voice commands, making them more accessible and efficient. If a command is not recognized, the bot politely informs the user about its limitations. Commands like 'bye' or 'exit' elicit farewells and potential deactivation. Furthermore, it oversees gesture control mode and terminates the program when needed, offering extensive interaction capabilities. Fig. 2 shows the workflow of the voice assistant.

IV. RESULTS & DISCUSSION

The virtual mouse is capable of performing a variety of tasks, such as scrolling, drag and drop, right-click, left-click, volume control, and brightness control. Moreover, the assistant is able to respond to commands related to mouse functions, date/time queries, file management, web searches, email sending, and the execution of applications like Notepad and Paint. It can also launch and stop applications. Some of the results are shown in Fig. 4 and 5. To ensure the effectiveness of the system, the virtual model has undergone testing in different lighting conditions and at various distances to accurately track hand movements and identify fingertips. Since the gesture recognition and voice assistant code provided does not use traditional machine learning models with predefined datasets for training and evaluation, latency, resource usage, manual accuracy calculation, and gesture variety are the parameters

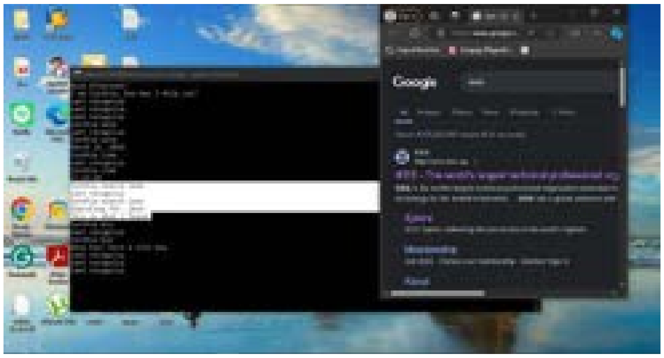


Fig. 4. Opening Google search

Fig. 5. Drag and Drop

used to evaluate the system's performance. The accuracy of the virtual mouse has been manually determined through approximately 50 trials using Eqn(6), with the results displayed in Fig
6. While pinch gestures may exhibit slightly lower accuracy compared to other mouse functions, the remaining interactions are highly precise.

$$Accuracy = 100 * (Correct/Total) \tag{6}$$

where correct - Corrected Number of Predictions
Total - Total Number of Predictions

TABLE I
INTERACTION PERFORMANCE

Interaction	Accuracy	Latency (s)
Drag and drop	98%	0.13
Left-click	98%	0.13
Right-click	95%	0.15
Double click	93%	0.15
Cursor movement	100%	0.5
Pinch gestures	90%	0.30
Voice commands	96%	

In comparison to the existing model, our virtual mouse and assistant have improved accuracy by reducing the difficulty of recognizing mouse gestures with the help of an enum. By using enum, all hand gestures are mapped to binary numbers, offering a wider range of mouse functions and supporting multi-handedness. It is also designed in a way to handle

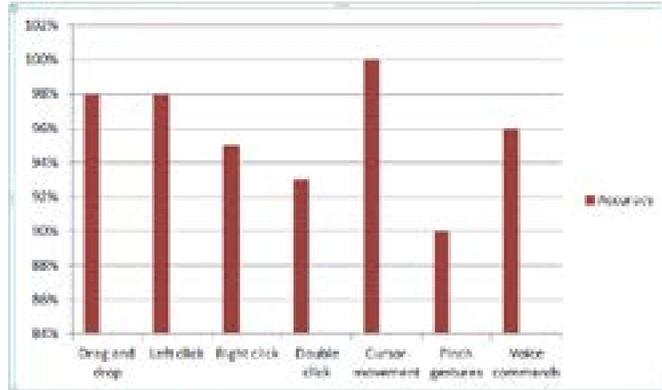


Fig. 6. Accuracy

fluctuations due to noise and stabilize the cursor position by dampening the jerky motion of the hand. In this model, the delay between gesture detection and system response, voice recognition, and system response is assessed, ensuring minimal latency for smooth interaction. Unlike the current methods, the suggested approach requires less time and doesn't require a complicated setup. The built-in camera and microphone are more than sufficient for the operation of this device. The training procedure itself is not involved, since training the datasets takes some time. Only the Python libraries are involved. The internet connection is not a determining factor. Privacy concerns arising from image processing techniques are addressed by the use of mediapipe in the keypoint mapping approach. It manages the multi-handedness. The vocabulary for voice and gesture recognition in the current system is restricted. Almost all mouse operations are carried out by our system using speech and gesture instructions. Both static and dynamic commands are controlled by the voice assistant. Error responses, retry mechanisms, fallback gestures or actions, and feedback mechanisms are implemented to provide better error handling mechanisms. Table 1 shows the accuracy and latencies of the interactions. Furthermore, it utilizes only 28% of CPU usage and consumes 32kb per second of disk memory.

V. CONCLUSION

This combined system integrates hand gestures and voice commands to enhance user interaction with digital devices. Advanced computer vision algorithms enable the real-time tracking of hand movements, allowing users to control the cursor and perform tasks with ease. This hands-free and intuitive interaction is particularly beneficial in situations where traditional methods are not practical. Moreover, the voice assistant feature supports task completion and information retrieval through sophisticated speech recognition and natural language processing. This fusion enhances accessibility and usability in various fields, such as gaming, accessibility, virtual reality, and smart home automation. As these technologies progress, the design of interfaces and interaction methods will continue to improve, ensuring more immersive user experiences in the digital domain.

VI. FUTURE WORKS

Future advancements will primarily focus on enhancing the system's gesture recognition, voice command capabilities, and overall user experience. The main objective will be to improve the accuracy and versatility of gesture recognition algorithms, enabling them to accurately identify intricate hand movements and gestures. Moreover, the intelligence of the voice assistant will be enhanced through advancements in natural language processing and machine learning, enabling it to better understand user commands and perform more complex tasks. By integrating multi-modal interaction capabilities, users will be able to effortlessly switch between gesture-based control and voice commands, resulting in a flexible and intuitive computing experience. The proposed system will prioritize accessibility and inclusivity by offering customizable gestures and voice commands tailored to individual preferences and accessibility requirements. Additionally, efforts will be made to ensure seamless integration with various devices, platforms, and operating systems, ensuring maximum usability and accessibility for users in diverse computing environments.

ACKNOWLEDGMENT

We sincerely appreciate the Rajalakshmi Institute of Technology administration for providing the facilities and resources that were necessary for our project to be completed success- fully. Our principal, Mr. N. Bhalaji, deserves special recognition for his steadfast assistance and provision of essential materials. We are incredibly grateful to our mentor, Dr. Anjana Devi, who is also the head of the Department of Computer Science and Engineering (CSE), for all of her help and inspiration during this project. Her knowledge played a crucial role in directing our investigation. We also thank our project coordinator, Dr. Anjana Devi, for her unwavering advice and assistance. We express our gratitude to the dedicated staff members of the CSE Department for their best wishes, prompt help, and steadfast support during the course of the project. Finally, we would like to sincerely thank our families, friends, and relatives whose unwavering support was invaluable to us during our academic path. It would not have been easy to reach this degree of accomplishment without their help.

REFERENCES

[1] Shashwat Srivastava, Richa Tiwari, Ananya Garg, Rishu and Rohit Tiwari," Design and Development of Gesture Recognition", 2023 IEEE International Conference on Computational Intelligence, Communication Technology and Networking(CICTN),doi:10.1109/CICTN57981.

[2] Murtaza Mustufa Sahiwala, Ms. Shrishti Ravindra Singh,Deepti Pawar, Disha Dinesh Jakasaniya, and Krishna Manoj Patel, "Virtual Mouse Using Coordinate Mapping", 2023 IEEE International Conference on Sustainable Computing and DataCommunication Systems(ICSCDS), doi: 10.1109/ICSCDS56580.2023.10104598.

[3] Chinnam Datta Sai Nikhil, Chukka Uma Someswara Rao, E.Brumancia, K.Indira, T.Anandhi and P.Ajitha, "Finger Recognition and Gesture based Virtual Keyboard", 2020 IEEE 5th International Conference on Communication and Electronics Systems (ICCES), doi: 10.1109/ICES48766.2020.9137889.

[4] Akshada Dongre, Rodney Pinto, Ameya Patkar and Minal Lopes, "Computer Cursor Control Using Eye and Face Gestures", 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), doi:10.1109/ICCCNT49239.2020.9225311.

[5] V. V. Reddy, T. Dhyanchand, G. V. Krishna and S. Maheshwaram, "Virtual Mouse Control Using Colored Finger Tips and Hand Gesture Recognition," 2020 IEEE- HYDCON, 2020, pp. 1-5, doi:10.1109/HYDCON48903.2020.9242677.

[6] Frederic Magoules and Qinmeng Zou, "A Novel Contactless Human Machine Interface based on Machine Learning",2017 IEEE International Symposium on Distributed Computing and Applications to Business, Engineering and Science, doi:10.1109/DCABES.2017.37.

[7] Changhyun Jeon,Oh-Jin Kwon,Dongil Shin and Dongkyoo Shin, "Hand-Mouse Interface Using Virtual Monitor Concept for Natural Interaction", 2017 IEEE, doi: 10.1109/ACCESS.2017.2768405.

[8] Marco E Benalca´zar, Cristhian Motoche, Jonathan A Zea, Andre´s G Jaramillo, Carlos E Anchundia, Patricio Zambrano, Marco Segura, Freddy Benalca´zar Palacios, and Mar´ıa Pe´rez, "Real-time hand gesture recognition using the myo armband and muscle activity detection" , 2017 Second Ecuador Technical Chapters Meeting (ETCM), pages 1–6. IEEE, 2017.

[9] Kori Painchaud and Leonidas Deligiannidis,"Customized Services Using Voice Assistants", 2020 IEEE International Conference on Computational Science and Computational Intelligence (CSCI), doi:10.1109/CSCI51800.2020.00197.

[10] Lanyu Xu,Arun Iyengar and Weisong Shi,"CHA: A Caching Framework for Home-based Voice Assistant Systems", 2020 IEEE/ACM Symposium on Edge Computing,doi:10.1109/SEC50012.2020.00041.

[11] Verghese Koshy Puthukkeril, Shyam Sundar E H and Nandha Kumar P R, "Hand Gesture Mouse Interface System", 2013 IEEE, International Conference on Human Computer Interactions (ICHCI),doi:10.1109/ICHCI-IEEE.2013.6887780.

[12] Joy Guha, Shreya Kumari and Prof. Shiv Kumar Verma,"AI Virtual Mouse Using Hand Gesture Recognition", 2022 International Journal for Research in Applied Science and Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 10 Issue IV.

[13] Shweta Pardeshi, Shubham Jagtap, Shashank Kathar, Shivam Giri and Prof. Suresh Kapare, "AI Virtual Mouse and Keyboard",2022 International Journal of Advances in Engineering and Management (IJAEM) Volume 4, Issue 5, pp: 1925-1936.

[14] Richa Golash and Yogendra Kumar Jain, "ECONOMICAL AND USER-FRIENDLY DESIGN OF VISION-BASED NATURAL-USER INTERFACE VIA DYNAMIC HAND GESTURES", IJARET, vol.11, pp. 338-348, June 2020.

[15] Anders T. Christensen,Henning Olesen and Lene Sorensen,"Digital Voice Assistants: A new kind of user agent ",2020 CMI Conference on Cybersecurity and Privacy (CMI) - Digital Transformation - Potentials and Challenges, doi:10.1109/CMI51275.2020.9322788.

[16] Zhanghui Chen,Huifeng Shen, Yan Lu and Shipeng Li, "Effective hand segmentation and gesture recognition for browsing web pages on a large screen", 2013 IEEE International Conference on Multimedia and Expo (ICME), doi: 10.1109/ICME.2013.6607588.

[17] Zhou Ren, Junsong Yuan, Jingjing Meng, and Zhengyou Zhang, "Robust part-based hand gesture recognition using kinect sensor", IEEE transactions on multimedia, 15(5):1110–1120, 2013.

[18] Ji-Hwan Kim, Nguyen Duc Thang, and Tae-Seong Kim, "3-d hand motion tracking and gesture recognition using a data glove", 2009 IEEE International Symposium on Industrial Electronics, pages 1013–1018. IEEE, 2009

[19] C-Chiung Hsieh, D-Hua Liou, and D.Lee,"A real time hand gesture recognition system using motion history image," Proc. IEEE International Conf. Signal Processing Systems (ICSPS), 2.10.1109/ICSPS.2010.5555462, 2010.

[20] K. H. Shibly, S. Kumar Dey, M. A. Islam, and S. Iftekhar Showrav, —"Design and development of hand Gesture based virtual mouse", in Proceedings of the 2019 1st International Conference on Advances in Science, Engineering and Robotics Technology (ICASERT), pp. 1–5, Dhaka, Bangladesh, May 2019

[21] Yi-Lin Chen, Yi-Lung Tsai, Kailing Huang and Pai H. Chou, "MobiRing: A Finger-Worn Wireless Motion Tracker"

Accurate Prediction and Detection of Suicidal Risk using Random Forest Algorithm

Saravanan.N
Professor ,
Department of IT,
K S R College of Engineering,
Trichengode, Tamilnadu - 637215
n.saran78@gmail.com

Moheshkumar.G
Assistant Professor,
Department of IT,
Faculty of Information Technology,
K S R College of Engineering,
Trichengode, Tamilnadu - 637215
mohesh.mohesh@gmail.com

Mohammed Shaid.V.M
Final year Department of IT,
K S R Institute for Engineering and Technology,
Tiruchengode, Tamilnadu – 637215
mohammedshahid.12603@gmail.com

Purushothman.S
Final year Department of IT,
KSR Institute for Engineering and Technology,
Tiruchengode, Tamilnadu – 637215
ppurushothaman239@gmail.com

Gokul Sanjai.V
Final year Department of IT,
KSR Institute for Engineering and Technology,
Tiruchengode, Tamilnadu – 637215
gokulvijay920@gmail.com

Corresponding Author: n.saran78@gmail.com

Abstract—This research study intends to create a reliable suicide risk assessment system by applying Machine Learning (ML) methods, in particular Random Forest algorithms is being compared with Support Vector Machines (SVM) for achieving the better accuracy results. The goal is to develop a trustworthy tool that can reliably identify individuals that are at risk of self-harm by analyzing their clinical, behavioral, and demographic data. Materials and Methods: The present work has involved two groups. Group 1 refers to the study intended based on Machine Learning Algorithm in which Support Vector Machine (SVM) has been implemented and has secured lower accuracy rates. Group 2 refers the method uses Random Forest to accurately assess the risk of suicide. It makes use of preprocessed health data to enable effective intervention and research advancement. The proposed method performs well in precise suicide prevention analysis with high metrics such as accuracy 0.85, precision 0.85, recall 0.85, and F1 score 0.85, making it useful for prevention analysis technique. The study concludes by showing the effectiveness of Random Forest in preventing suicide and emphasizing how the data-driven methods and user-friendly interfaces may be applied in the real world to improve mental health.

Keywords—Machine Learning, Support Vector Machine, Random Forest Algorithm, Precision, Accuracy, F1-score, Recall, Personal Health Information.

I. INTRODUCTION

Suicide is a global public health crisis, claiming the lives of hundreds of thousands of individuals each year. It leaves a devastating impact on families, communities, and societies at large. Traditional methods of suicide prevention have relied heavily on self-reporting and the clinical judgment of mental health professionals. However, the complexity of factors contributing to suicide, coupled with the stigma surrounding mental health, often limits the effectiveness of these approaches. For the purpose to classify suicide risk, this study investigates the application of Random Forest algorithm. Utilizing machine learning approaches for suicide prediction shows potential for early intervention and prevention efforts in light of the growing concerns over mental health [1]. Through the examination of multiple clinical, behavioral, and demographic variables, the goal of this research is to create a strong model that can reliably identify people who are at risk of self-harm and are in depression. The promise of data-driven techniques in tackling key mental health concerns is highlighted by the combination of sophisticated algorithms and extensive databases. The amount of research on using Artificial Intelligence to provide proactive mental health treatment is increasing, and this study adds to that [2]. The risk variables had a limited potential to predict Suicidal Thoughts and Behaviors (STBs) over a 50-year period, with only modest gains with time, according to a meta-analysis on STBs. Longer study durations did not improve prediction accuracy, and in most of the research, despite its breadth, concentrated on specific categories of risk variables. The results highlight how successful preventative and treatment techniques can only be advanced by a paradigm shift towards machine learning-based risk algorithms [3] [4].

II. RELATED WORKS

Analysis of the social media postings made by military people who have passed away reveals unique temporal patterns such as posts regarding triggers followed by thoughts that may indicate the likelihood and timing of suicide, providing information for focused preventative efforts. Finding certain trends in an individual's online activity may help with early intervention and assistance for those who are at danger [5]. This study examined the viability of recognizing "strongly concerning" tweets with moderate accuracy using a combination of machine learning and human coders to gauge the degree of worry in suicide-related tweets. Although it shows the possibility of detection, it also highlights the need for more improvement and neither directly informs intervention goals nor forecasts suicide conduct [6] [7]. This study indicates a relationship between search volume, abortion rates, and policy constraints, suggesting that examining internet search trends, particularly those related to abortion, might provide light on the effects of local abortion laws [8]. This technique provides a potentially useful tool for more quickly and efficiently analyzing the effects of health policies by demonstrating a link between search activity and policy impact [9] [10]. This thorough analysis emphasizes how the research on suicide prevention strategies has changed since 2005, with a focus on the growing body of evidence supporting means restriction and school-based awareness campaigns. It highlights the necessity of integrated, evidence-based approaches at the individual and population levels, while also recognizing the gaps in the present body of research [11]. This excerpt highlights the expanding application of big data analytics in the field of psychiatry, encompassing a wide range of data types beyond genetics, such as clinical, sociodemographic, and social media data. It represents a major breakthrough in psychiatric research since it makes predictions for clinical outcomes and individual-level patterns by utilizing machine learning techniques [12]. This article talks about using natural language processing and deep learning to analyze social media data to identify people who may be suicidal. This might be a preventive intervention that even non-mental health professionals could use. It highlights the moral conundrum of reaching more at-risk people while upholding their right to privacy and raises ethical questions regarding privacy vs prevention [13] [14]. In order to create predictive models for assessing suicide risk, this study used linguistic analysis of clinical notes from medical records maintained by the U.S. Veterans Administration. The results imply that automated text analytics may reliably evaluate the risk of suicide, which may allow primary care physicians to screen patients and keep an eye on risk levels among mental patients [15]. The diagnostic accuracy of suicide prediction models was evaluated in this review, which found that while categorization accuracy was typically good, predictive validity for actual suicide occurrences was relatively poor. Significant constraints imply that these models are not yet appropriate for broad clinical deployment throughout health systems, even with their promising overall accuracy [16]. By addressing methodological limitations, concentrating on short-term assessments, and utilizing machine learning techniques, the study aimed to improve suicide prediction. This resulted in a significant improvement in prediction accuracy for both suicidal ideation and attempts among high-risk individuals, underscoring the importance of taking complexity into account when attempting suicide [17] [18]. In order to determine which US veterans were at a high risk of suicide, the study tested several predictive modeling approaches. It discovered that a penalized logistic regression model with 61 variables outperformed an earlier model with more predictors. In addition to taking into account realistic aspects to improve model performance, the Veterans Health Administration is putting this model into practice for early intervention [19]. Support Vector Machine (SVM) for suicide risk assessment highlighted the accuracy in analyzing clinical, behavioral, and demographic data to identify individuals at risk of suicide. Leveraging these findings, the proposed work focuses on refining a Random Forest-based assessment tool. This involves expanding the dataset for better predictive accuracy, enhancing model validation across various demographics, creating a user-friendly interface for healthcare professionals, and integrating the system within existing mental health interventions. The aim is to develop a reliable, AI-powered tool that significantly aids in early detection of suicide risk, thereby improving mental health outcomes through timely and effective intervention strategies.

III. MATERIALS AND METHODS

By preprocessing PHI data, a suicide risk assessment system utilizing the Random Forest algorithm was created to improve accuracy. The objective was to support mental health research by developing an accurate tool for detecting those who are at risk of self-harm. Fig.1 represents that Data preprocessing is a critical step in the machine learning pipeline, involving cleaning and organizing raw data to make it suitable for a model like Random Forest. This process can include handling missing values, normalizing or scaling features, encoding categorical variables, and selecting relevant features. Once preprocessed, the data is ready for classification tasks, where the goal is to predict a target category.

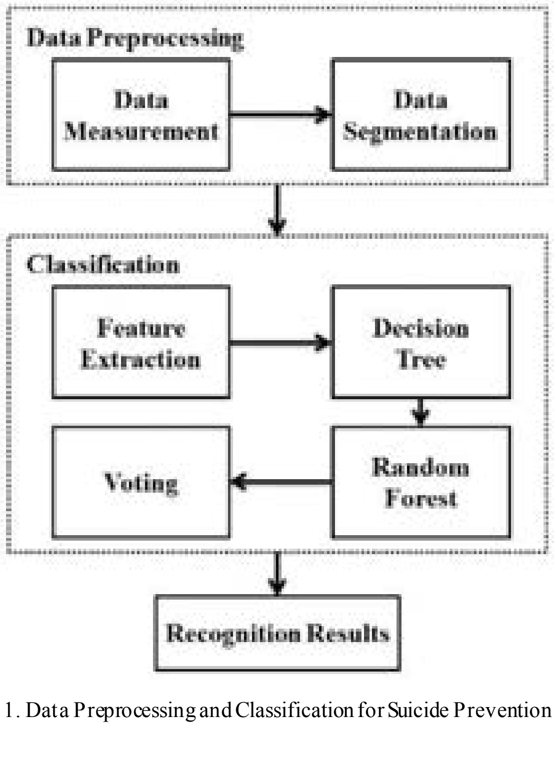


Fig. 1. Data Preprocessing and Classification for Suicide Prevention Analysis

Group 1 utilizes the Support Vector Machine (SVM) algorithm to develop a machine learning-based system for assessing suicide risk. By preprocessing and analyzing Personal Health Information (PHI) data, the SVM approach aims to identify individuals at risk of self-harm. Although SVM is recognized for its capability to manage high-dimensional spaces effectively, the accuracy achieved in predicting suicide risk was found to be suboptimal when evaluated independently. This outcome suggests a potential for further refinement of the SVM model to improve its effectiveness as a tool for medical professionals, thereby enhancing mental health research and intervention strategies.

Group 2 defines the method here assesses suicide risk with accuracy by using machine learning approach Random Forest algorithm, based on Random Forest has higher Accuracy compared to Support Vector Machine. It attains noteworthy accuracy by preprocessing Personal Health Information (PHI) data and utilizing these models. This provides a dependable tool to help medical professionals identify and intervene with people who are at risk of self-harm, improving mental health research and intervention techniques.

IV. Statistical Measures of Accuracy in Regression Analysis

In Suicide Prevention Analysis, leveraging metrics like Accuracy, Precision, F1-score and Recall is essential for evaluating predictive model. These statistics help in understanding the effectiveness of models designed to identify suicide risk factors, enabling targeted interventions and contributing significantly to prevention efforts.

$$\text{Accuracy}_i = \frac{TP+TN}{TP+TN+FP+FN} \tag{1}$$

In the above, equation (1) defines Accuracy as the ratio of accurately predicted occurrences (true positives and true negatives) to the total number of instances.

$$\text{Precision}_i = \frac{TP}{TP+FP} \tag{2}$$

In the above, equation (2) denotes Precision is the ratio of genuine positives to total anticipated positive cases, which measures the accuracy of positive forecasts.

$$\text{Recall}_i = \frac{TP}{TP+FN} \tag{3}$$

In the above, equation (3) refers to Recall, also known as sensitivity or true positive rate, is the proportion of true positives to total positive instances.

$$\text{F-Measure}_i = \frac{(1+\beta^2)\times precision \times Recall}{\beta^2 \times (precision \times Recall)} \tag{4}$$

In the above, equation (4) relates the F-measure combines precision and recall, with β regulating the trade-off between the two.

V. Results

The research analysis demonstrates significant potential, particularly with Random Forest Algorithm, which achieved an impressive accuracy of 0.85 which is 85%, precision of 0.85 which is 85%, recall of 0.85 which is 85% and F1 score 0.85 which is 85%. Table 1 describes the context of the suicide prevention study using machine learning technique, the performance of the Random Forest model was evaluated based on key metrics: Precision, Recall, F-measure (F1-score), and Accuracy.

Table 1 The performance metrics of the Random Forest model: Accuracy (0.85), Precision (0.85), F1-Score (0.85), and Recall (0.85).

Metrices	Random Forest
Accuracy	0.85
Precision	0.85
F-measure(F1-score)	0.85
Recall	0.85

The Graphical Representation of the Accuracy, Prediction, F1-Score and Recall has been represented, Where after identifying then it checks the patient's risk level and it suggests an interview or counseling. The machine learning algorithms' strong accuracy highlights its ability to identify suicidal tendencies in preprocessed Personal Health Information (PHI). The system's accuracy, f1-score, precision and recall measures and demonstrates its performance, which confirms its potential as a trustworthy tool for mental health professionals. It helps with early intervention and provides important support for designing focused intervention methods by giving accurate risk ratings. The system's effectiveness also demonstrates the progress gained in utilizing AI and data-driven methods to enhance mental health services. This model has the potential to identify people who are at-risk and to develop proactive treatments that reduce the risk of suicide as they continue to develop and improve. Table 2 represents the comparison of Support Vector Machine and Random Forest.

Table 2 is the Comparison Table in which Support Vector Machine has obtained an accuracy of 78% and Random Forest Algorithm of 85%.

Algorithms	Accuracy Obtained
Support Vector Machine	78%
Random Forest	85%

The suicide risk assessment system's outstanding performance highlights an important development at the nexus of machine learning and mental health services. In addition to the remarkable accuracy rates attained by the Random Forest model, the precision, accuracy metrics corroborate the efficacy of the system in detecting minute but significant patterns in Personal Health Information (PHI) that could signify a higher risk of suicide. Fig. 2. represents a comprehensive overview of the Random Forest model's performance in suicide risk prediction, showcasing its effectiveness through high values in accuracy, precision, recall, and F1-score, indicating a robust capability to accurately identify and classify individuals according to their suicide risk levels. Fig. 3. illustrates the complex interplay among a range of variables including Depression Score, Anxiety Score, Relationship Status, Employment Status, Social Connections, Mental Health Treatment, and Sleep Quality. This visualization underscores how each parameter can influence or be influenced by another, revealing the multifaceted factors that may affect an individual's mental health. The figure aims to highlight not just the individual impact of these parameters but also how they collectively contribute to the broader understanding of mental health dynamics. Through this analysis, we can discern patterns and correlations that are crucial for developing more targeted and effective mental health interventions and support systems. One of the system's main advantages is its capacity to quickly and accurately collect and analyze enormous volumes of data, giving mental health professionals risk assessments in a timely manner. Given the increasing need for scalable solutions to treat mental health issues in healthcare systems around the globe, this feature is especially beneficial.

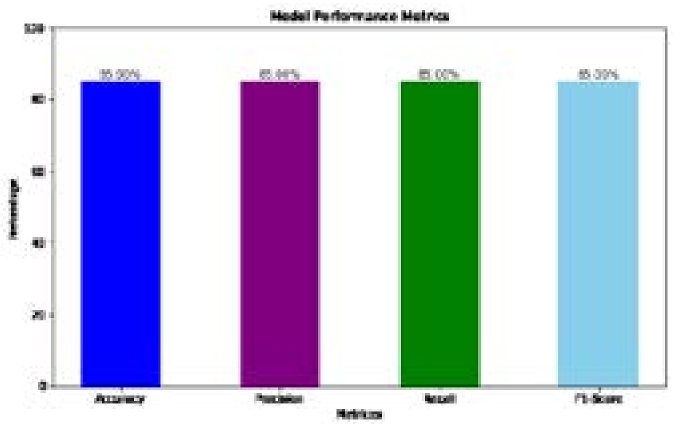


Fig. 2: Comprehensive overview of the Random Forest model's performance in suicide risk prediction, showcasing its effectiveness through high values in accuracy, precision, recall, and F1-score, indicating a robust capability to accurately identify and classify individuals according to their suicide risk levels

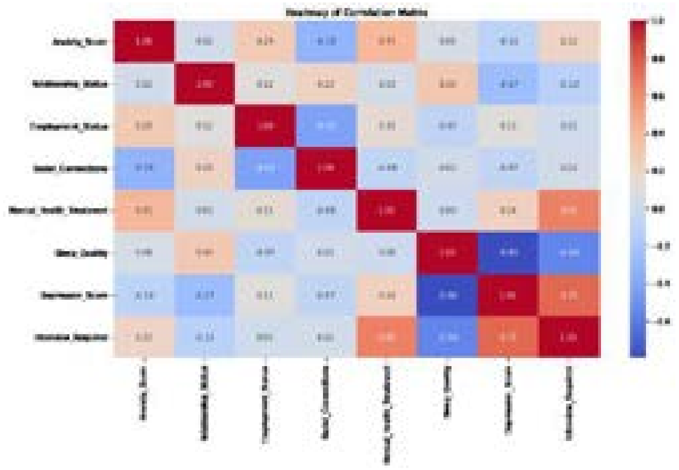


Fig. 3. The correlation between various parameters such as Depression Score, Anxiety Score, Relationship Status, Employment Status, Social Connections, Mental Health Treatment, and Sleep Quality, highlighting the intricate relationships and potential impact factors contributing to an individual's mental health state.

VI. DISCUSSION

The research underscores the Random Forest algorithm's effectiveness in suicide risk assessment, achieving remarkable accuracy and demonstrating significant promise in enhancing mental health interventions through advanced AI-driven analytics.

The research improved suicide attempt prediction accuracy and scalability dramatically by integrating machine learning with electronic health information. With the use of more than 250 adult patient records, the algorithms produced results with excellent recall and precision rates, showing a discernible increase in accuracy as the attempt at suicide became closer. These findings open the door to more potent therapies and preventative measures while providing insightful information about the temporal dynamics of suicide risk [20].

This analysis emphasizes how urgently new approaches are needed to combat the growing suicide rates, especially those involving mobile and sensor technologies. Although recent developments indicate potential in the use of machine learning and smartphone sensing for suicide prevention, further validation and replication of results are needed. In the next two to five years, physicians may expect the development of useful digital tools to help them better identify, forecast, and treat patients who are at high risk of suicide. These efforts are supported by current research [21] [22].

First results are provided on the use of predictive modeling by the US Veterans Health Administration (VHA) to identify veterans at high risk of suicide and provide tailored care. A 1% sample of time-matched VHA service users and 6,360 veterans' electronic medical records that died by suicide were analyzed. Methods included comparing logistic regression models and investigating different machine learning techniques [23].

In order to improve the prediction of suicidal thoughts and behaviors (STBs), frequent methodological limitations were addressed in this study. It recruited over 1,000 high-risk individuals globally by utilizing machine learning and focused on short-term prediction. The results shown significant increases in prediction accuracy, especially for impending suicidal thoughts and nonfatal suicide attempts [24] [25].

This study evaluated the diagnostic accuracy of suicide prediction models and found that while overall categorization was high, the predictive validity for suicide outcomes was very poor. The models show promise in terms of worldwide accuracy, but their predictive power is limited in terms of future occurrences, which limits their clinical value and suitability for application in different health systems [26].

Utilizing unstructured clinical notes from US Veterans Administration (VA) data, this study created linguistics-driven prediction models that consistently estimated suicide risk with an accuracy of 65% or higher. According to research, computerized text analytics applied to medical records may make it possible to continuously assess the risk of suicide in psychiatric settings and to screen at-risk patients early in primary care [27].

Subsequent efforts should concentrate on refining the algorithm's precision and its relevance for different population groups, integrating instantaneous data analysis for adaptive evaluations, and developing more accessible interfaces for healthcare practitioners. Furthermore, broadening investigations into different cultural environments and tackling moral and confidentiality issues associated with the use of Personal Health Information are essential for the ethical implementation of artificial intelligence in mental health services.

VII. CONCLUSION

To sum up, the integration Random Forest algorithm in the suicide risk assessment system shows potential in detecting possible cases of self-harm in medical environments. It highlights the value of data-driven approaches in mental health with its user-friendly design and high accuracy. This technique attained 85% of accuracy by training of the PHI. Rethinking mental health services for better proactive and patient-centred care requires integrating advanced analytics and machine learning to improve early intervention, resource allocation, and ongoing model refinement. The research underscores the Random Forest algorithm's effectiveness in suicide risk assessment, achieving remarkable accuracy and demonstrating significant promise in enhancing mental health interventions through advanced AI-driven analytics.

REFERENCES

[1] Belsher, Bradley E., Derek J. Smolenski, Larry D. Pruitt, Nigel E. Bush, Erin H. Beech, Don E. Workman, Rebecca L. Morgan, Daniel P. Evatt, Jennifer Tucker, and Nancy A. Skopp. 2019. “Prediction Models for Suicide Attempts and Deaths: A Systematic Review and Simulation.” JAMA Psychiatry 76 (6): 642–51.

[2] Bryan, Craig J., Jonathan E. Butner, Sungchoon Sinclair, Anna Belle O. Bryan, Christina M. Hesse, and Andree E. Rose. 2018. “Predictors of Emerging Suicide Death Among Military Personnel on Social Media Networks.” Suicide & Life-Threatening Behavior 48 (4): 413–30.

[3] Babu, AV Santhosh, P. Meenakshi Devi, and B. Sharmila. "Efficient enhanced intrusion identification and response system for MANETs." International Journal of Business Information Systems 29, no. 4 (2018): 535-546.

[4] Coppersmith, Glen, Ryan Leary, Patrick Crutchley, and Alex Fine. 2018. “Natural Language Processing of Social Media as Screening for Suicide Risk.” Biomedical Informatics Insights 10 (August): 1178222618792860.

[5] Franklin, Joseph C., Jessica D. Ribeiro, Kathryn R. Fox, Kate H. Bentley, Evan M. Kleiman, Xieyining Huang, Katherine M. Musacchio, Adam C. Jaroszewski, Bernard P. Chang, and Matthew K. Nock. 2017. “Risk Factors for Suicidal Thoughts and Behaviors: A Meta-Analysis of 50 Years of Research.” Psychological Bulletin 143 (2): 187–232.

[6] C.Rajan, C.Dharanya, Dr.N.Shanthi ""A Probabilistic Rebroadcast for Reducing Routing Overhead In A Real Time Manet Environment"".Journal of Global Research in Computer Science, Volume 5, No. 1, January 2014, 3-7."

[7] Kessler, Ronald C., Irving Hwang, Claire A. Hoffmire, John F. McCarthy, Maria V. Petukhova, Anthony J. Rosellini, Nancy A. Sampson, et al. 2017. “Developing a Practical Suicide Risk Prediction Model for Targeting High-Risk Patients in the Veterans Health Administration.” International Journal of Methods in Psychiatric Research 26 (3). https://doi.org/10.1002/mpr.1575.

[8] O’Dea, Bridianne, Stephen Wan, Philip J. Batterham, Alison L. Calear, Cecile Paris, and Helen Christensen. 2015. “Detecting Suicidality on Twitter.” Internet Interventions 2 (2): 183–88.

[9] MeenakshiDevi, P., M. Venkatesan, and K. Duraiswamy. "A fragile watermarking scheme for image authentication with tamper localization using integer wavelet transform." Journal of Computer Science 5, no. 11 (2009): 831.

[10] Passos, Ives Cavalcante, Benson Mwangi, and Flávio Kapczinski. 2016. “Big Data Analytics and Machine Learning: 2015 and beyond.” The Lancet. Psychiatry 3 (1): 13–15.

[11] Poulin, Chris, Brian Shiner, Paul Thompson, Linas Vepstas, Yinong Young-Xu, Benjamin Goertzel, Bradley Watts, Laura Flashman, and Thomas McAllister. 2014. “Predicting the Risk of Suicide by Analyzing the Text of Clinical Notes.” PloS One 9 (1): e85733.

[12] Reis, Ben Y., and John S. Brownstein. 2010. “Measuring the Impact of Health Policies Using Internet Search Patterns: The Case of Abortion.” BMC Public Health 10 (August): 514.

[13] Dinesh, Paidipati, and P. Kalyanasundaram. "Medical image prediction for diagnosis of breast cancer disease comparing the machine learning algorithms: SVM, KNN, logistic regression, random forest, and decision tree to measure accuracy." ECS Transactions 107, no. 1 (2022): 12681.

[14] Ribeiro, Jessica D., Xieyining Huang, Kathryn R. Fox, Colin G. Walsh, and Kathryn P. Linthicum. 2019. “Predicting Imminent Suicidal Thoughts and Nonfatal Attempts: The Role of Complexity.” Clinical Psychological Science 7 (5): 941–57.

[15] Schou Pedersen, Henrik, Morten Fenger-Grøn, Bodil Hammer Bech, Annette Erlangsen, and Mogens Vestergaard. 2019. “Frequency of Health Care Utilization in the Year prior to Completed Suicide: A Danish Nationwide Matched Comparative Study.” PloS One 14 (3): e0214605.

[16] Torous, John, Mark E. Larsen, Colin Depp, Theodore D. Cosco, Ian Barnett, Matthew K. Nock, and Joe Firth. 2018. “Smartphones, Sensors, and Machine Learning to Advance Real-Time Prediction and Interventions for Suicide Prevention: A Review of Current Progress and Next Steps.” Current Psychiatry Reports 20 (7): 51.

[17] R. Gurusamy, V. Rajmohan, N. Sengottaiyan, P. Kalyanasundaram and S. M. Ramesh, "Comparative Analysis on Medical Image Prediction of Breast Cancer Disease using Various Machine Learning Algorithms," 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2023, pp. 1522-1526, doi: 10.1109/ICESC57686.2023.10193635.

[18] Treweek, Shaun, Marie Pitkethly, Jonathan Cook, Cynthia Fraser, Elizabeth Mitchell, Frank Sullivan, Catherine Jackson, Tyna K. Taskila, and Heidi Gardner. 2018. “Strategies to Improve Recruitment to

Randomised Trials." Cochrane Database of Systematic Reviews 2 (2): MR000013.

[19] Walsh, Colin G., Jessica D. Ribeiro, and Joseph C. Franklin. 2017. "Predicting Risk of Suicide Attempts over Time through Machine Learning." Clinical Psychological Science 5 (3): 457–69.

[20] Zalsman, Gil, Keith Hawton, Danuta Wasserman, Kees van Heeringen, Ella Arensman, Marco Sarchiapone, Vladimir Carli, et al. 2016. "Suicide Prevention Strategies Revisited: 10-Year Systematic Review." The Lancet. Psychiatry 3 (7): 646–59.

[21] Baskar, M., Renuka Devi, R., Ramkumar, J. et al. Region Centric Minutiae Propagation Measure Orient Forgery Detection with Finger Print Analysis in Health Care Systems. Neural Process Lett 55, 19–31 (2023). https://doi.org/10.1007/s11063-020-10407-4

[22] MeenakshiDevi, P., M. Venkatesan, and K. Duraiswamy. "A fragile watermarking scheme for image authentication with tamper localization using integer wavelet transform." Journal of Computer Science 5, no. 11 (2009): 831.

[23] Sung, Sheng-Feng, Chia-Yi Lin, and Ya-Han Hu. 2020. "EMR-Based Phenotyping of Ischemic Stroke Using Supervised Machine Learning and Text Mining Techniques." IEEE Journal of Biomedical and Health Informatics 24 (10): 2922–31.

[24] Tazin, Tahia, Md Nur Alam, Nahian Nakiba Dola, Mohammad Sajibul Bari, Sami Bourouis, and Mohammad Monirujjaman Khan. 2021. "Stroke Disease Detection and Prediction Using Robust Learning Approaches." Journal of Healthcare Engineering 2021 (November): 7633381.

[25] Uppal, Salil, Sandeep Goel, Baljit Randhawa, and Ankush Maheshwary. 2020. "Autoimmune-Associated Vasculitis Presenting as Ischemic Stroke With Hemorrhagic Transformation: A Case Report and Literature Review." Cureus 12 (9): e10403.

[26] Babu, AV Santhosh, P. Meenakshi Devi, and B. Sharmila. "Efficient enhanced intrusion identification and response system for MANETs." International Journal of Business Information Systems 29, no. 4 (2018): 535-546.

[27] Verma, Ankur, Sanjay Jaiswal, and Wasil Rasool Sheikh. 2020. "Acute Thrombotic Occlusion of Subclavian Artery Presenting as a Stroke Mimic." Journal of the American College of Emergency Physicians Open 1 (5): 932–34.

Revolutionizing Agricultural Cooperatives: The Role of CATAFA Management Information System

Orient Kerr T. Devero
College of Computer Studies
Central Philippines State University
Kabankalan City, Philippines
orientkerr10@gmail.com

Elysa Jane T. Arlos
College of Computer Studies
Central Philippines State University
Kabankalan City, Philippines
ejaneArlos@gmail.com

Chester L. Cofino
College of Computer Studies
Central Philippines State University
Kabankalan City, Philippines
chestercofino@cpsu.edu.ph

Abstract— **This study addresses the information management challenges of the Camingawan-Tagukon Farmers Association (CATAFA) located at Brgy. Camingawan, Kabankalan City, Negros Occidental, Philippines by developing a comprehensive system known as CATAFA Management Information System (MIS). Currently, CATAFA relies on manual calculations for its savings and loan transactions, making the current system prone to human errors. The system's primary objective is to automate critical processes such as membership registration, loan applications and transactions, and savings account management. The system has also incorporated a ledger functionality to record all member transactions. The procedure development followed the Waterfall Model within the Software Development Life Cycle (SDLC). The system also focuses on flexibility, allowing it to be utilized by CATAFA and other similar associations based on their specific needs and requirements. A survey questionnaire was administered to CATAFA staff members, who provided ratings and feedback to evaluate its effectiveness. The results indicated that the system effectively met the association's needs, prompting a desire for its implementation within the organization.**

Keywords— Agricultural Cooperatives, Information Management System (MIS), Automation, Farmers' Association

I. INTRODUCTION

Agricultural cooperatives' economic structure may be rationally explained by using the organizational economics rationale for family farms[1]. Farmers should produce in a way that balances the economic, environmental, and social facets of sustainability [2]. In Europe, where they represent 40–60% of the agricultural trade, agricultural cooperatives (ACs) are significant players in farming and influential speakers of rural reality [3]. According to [4] agricultural organizations may minimize uncertainty and make better decisions using digital technology to pull information from multiple data sources.

Farmers establish agricultural cooperatives to solve various issues jointly and enhance their economic and social well-being. Cooperatives may provide a variety of advantages, but they also confront several problems and difficulties that may limit their viability [5]. The inability of many agricultural cooperatives to get inexpensive finance and financial services might limit their capacity to invest in new facilities, equipment, and other enhancements. The study of [6] suggested that it is important for agricultural cooperatives to encourage the adoption of contemporary service management applications via their role as technology diffusion organizations that share a social network with farmers.

A management information system (MIS) is a collection of methods and procedures that gathers information from many sources, aggregates it, and understandably presents things. Managers use this tool to create reports that provide them with a comprehensive picture of all the information they need to make decisions about everything from long-term strategy to daily operations. The latest tools in doing business mostly rely on technology to collect and transmit data, even though the idea predates modern computer capabilities. This technology described as a platform of financial data organized and categorized in this way, may provide regular updates on operations to the management levels of a firm. Furthermore, it is often possible to easily get certain reports from the system. The platform provides the organization in terms of performance feedback and enables senior management to monitor service quality across the board.

A new era in agricultural production management is beginning, one in which highly complex Farm Management Information Systems (FMISs) will support daily choices made by farmers [7]. The Philippines still has a low rate of computer and decision-support system usage in agricultural cooperatives [8]. These technologies must spread further in order These technologies must spread to increase production efficiency and increase farmers' revenue.

Camingawan-Tagukon Farmers Association (CATAFA) is a non-government organization in the Philippines, specifically in Negros Occidental, catering services to upland farmers in crop production, livestock, and other related agricultural businesses. The organization created a cooperative for the members in terms of financial assistance, loans, and savings. Due to the background of the members focused on agricultural processes, problems were encountered by its staff and members in terms of recording their business transactions. Thus, the researchers came up with the idea to improve the processes of the CATAFA by providing a platform to improve their services.

The CATAFA Management Information System is designed to offer a dependable system that efficiently enables the designated administrator to add new members to CATAFA through data entry. Additionally, the system is equipped with information management features, such as viewing, editing, and removing information. Furthermore, the system is designed to facilitate searching, sorting, and filtering functions for the information collected from both existing members and prospective applicants of the association. It has a data analytics feature, presenting correct and ease-of-view of data in terms of financial-related reports that made the system unique to other existing works. With this system in place, all necessary information will be easily collected and managed.

II. METHODOLOGY

The researchers followed the waterfall-model procedure to create high-quality software.

Figure 1. It is a schematic diagram of the conceptual framework of the study. A seamless transition into the newly produced bespoke application will come from adhering to the six phases of the System Development Lifecycle [9].

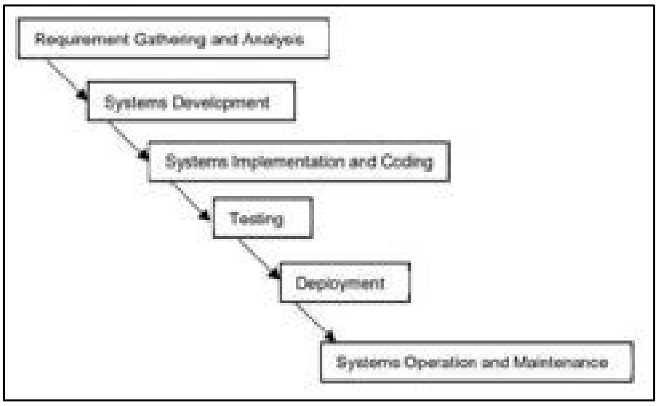


Figure 1. Conceptual Framework

A. Requirement Gathering and Analysis

In this phase, the researchers build thoughtful plans on how to do it step by step. During this phase, the researcher developed a fully functioning software with its architecture and design. The requirements were gathered from interviewing the CATAFA stakeholders and other relevant parties. This phase assessed the user's suggestions and plans for the system's outcome. In the validation process, the system was checked to determine errors and whether codes were to be fixed.

The association's staff was interviewed to identify their specific needs for the system. These interviews yielded useful information that gave us insight into the demands and issues the organization was facing.

Design

Based on the insights gained from the analysis, a system design was created to cater to the specific needs of the CATAFA association. This design encompassed the management information system's structure, functionality, and user interface.

Figure 2 demonstrates the design process using user-provided inputs that are handled based on the result that the consumers need.

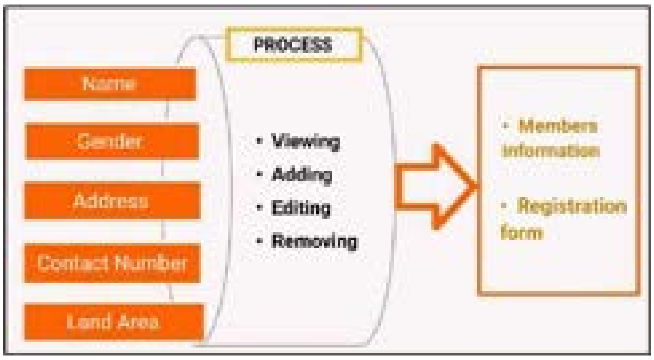


Figure 2. Input-Process-Output of CATAFA

Figure 3 shows the process flow of the system.

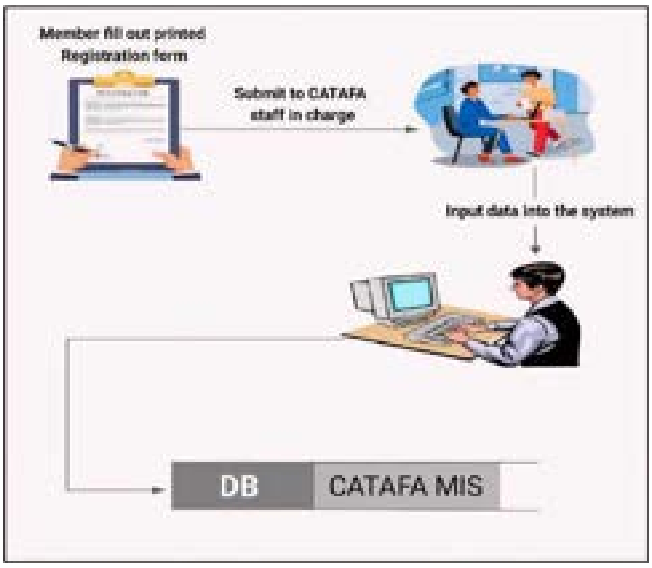


Figure 3: Process Flow

B. System Development, Implementation & Coding

The system was coded and constructed using the system design as a blueprint. The front end used jQuery for JavaScript and Bootstrap 5 for CSS, while the back end used the PHP programming language and MySQL to manage the database.

In testing the functionality of the system, the researchers utilized an open-source platform (SimpleTest) to check the codes if there were errors or bugs. If bugs were encountered/identified, the researchers reviewed the process and debugged the codes for optimization and error-free function.

The table below (Table 1 and Table 2) are the needed requirements are the development of the system.

Table 1. Hardware Requirements (Recommended)

Device	Specifications
Central Processing Unit (CPU)	Processor: Intel Core i5 or AMD Ryzen 5

Monitor	21" 1024x768 resolution
Keyboard	Standard Keyboard
Mouse	Standard Mouse
Hard Disk Drive (HDD)	256 GB
Random Access Memory (RAM)	4 GB
Switch	4 inches high, 9.5 inches long, and 0.55 inches deep (with Joy-Con attached)
LAN Cable	Category 5, 100 Mbps,100 MHz

Table 2: Software Requirements (Recommended)

Software	Version
Sublime Text Editor	v4 Build 4143
XAMPP DBMS Server	v3.3.0
Browser (e.g. Google Chrome)	Latest version

C. Testing

The system underwent rigorous testing procedures to ensure accuracy, reliability, and performance. Various tests, including bug checks, were conducted to identify and resolve any issues or inconsistencies, enhancing the system's functionality and usability.

D. Deployment

The system was presented to the CATAFA office. A survey questionnaire was then given to the staff members at CATAFA, asking for their ratings and feedback. The results showed that the system effectively met the association's needs, creating a strong desire to implement it within the organization.

Figure 4. illustrates the system's member list page, where you can add new members, and the list of members is displayed.

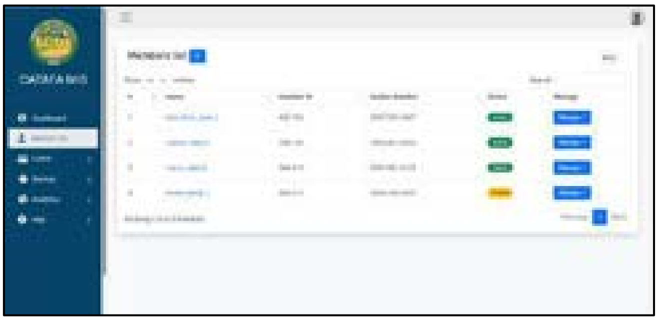


Figure 4. Members List Page

Figure 5 shows the list of loan plans, where you can customize the monthly interest, repayment method, payment frequency, and more.

Figure 5. Loan Plans

Figure 6 displays the loan list page. This is where you can add a new loan, manage a loan, or delete a loan.

Figure 6. Loan List

E. System Acceptance for Implementation

During the implementation phase, the researchers conducted several trainings to the users. First, the researchers prepared training manuals for the users. During the training, the researchers presented the system through actual demonstration and video presentation for review purposes. Second, the system was placed into dry-run process for one (1) month by the users to check some issues/problems in actual business transactions. Fortunately, the system function smoothly and achieved its purpose correctly in terms of data transactions and its navigation in terms of user interface. Lastly, the users accepted the system for implementation and provided a certificate of utilization to the researchers. One issue that was encountered during the implementation phase, is the lack of hardware requirements such as computers to be configured as server, thus, the system was not utilized. The CAFATA is planning to purchase the needed requirements for system installation to fully utilized the said platform.

III. RESULTS AND DISCUSSION

The research study involved 20 staff members and members of the Camingawan-Tagukon Farmers Association (CATAFA) located in Kabankalan, Negros Occidental. These respondents evaluated the system based on its functionality, design, security, and effectiveness using the criteria developed by the researchers based on the PIECES framework [10] a questionnaire using the Importance Performance Analysis approach.

Table 3 presents a Likert Scale for evaluating the system's functionality, design, security, and effectiveness.

Table 3. Likert Scale for functionality, design, security, and effectiveness.

		Range Scale
4.21	5.00	Very Satisfied
3.41	4.20	Satisfied
2.61	3.20	Neither Satisfied nor Dissatisfied
1.81	2.60	Dissatisfied
1	1.80	Very Dissatisfied

Table 4 presents the average rating provided by respondents for the system's functionality. The results reveal a high overall mean rating of 4.81. Users appreciated the system's ease of access through secure login, its capability to effectively manage input data, and its feature to flag missing information. The fully operational search engine also received positive feedback for enhancing usability. Additionally, as emphasized in [11], users must understand the product's technical aspects as a crucial requirement for systemadoption.

Table 4. Functionality

Functionality	Mean	Description
The system is accessible using the correct username and password	4.7	Very Satisfied
It can manage (view, add, edit, delete) the input data.	4.8	Very Satisfied
The system will notify the encoder/admin when there is missing data in the form.	4.95	Very Satisfied
The search engine of the systemis fully functional.	4.8	Very Satisfied
Grand Mean	4.81	Very Satisfied

Table 5 displays the evaluation results of the system's design. It has obtained a grand mean of 4.78. Users appreciate the well-positioned UI elements, contributing to seamless usability. The organized data layout enhances accessibility and decision-making. The design's pleasing aesthetics enhance user engagement. According to [12], the user interface (UI) is a tool that facilitates user interaction with a product's service interface.

Table 5. Design

Design	Mean	Description
The system has a user-friendly interface.	4.75	Very Satisfied
The information being displayed by the system is clear.	4.75	Very Satisfied
The buttons, text fields, and list views are in the appropriate positions and clear.	4.85	Very Satisfied
The data in the table are in an organized position.	4.8	Very Satisfied
The system's overall design is pleasant to the eyes of the users.	4.75	Very Satisfied
Grand Mean	**4.78**	**Very Satisfied**

Table 6 presents the ratings provided by respondents regarding program security. The results indicate a notable grand mean of 4.79, signifying a high level of satisfaction among the participants with the system's security. Users commend the robust measures in place to prevent unauthorized access and the effective data backup security, which reduces the risk of data loss. The [13] study's finding that user accessibility had a favorable effect on end-user satisfaction with e-government system designs in Jordan was confirmed.

Table 6. Security

Security	Mean	Description
The system requires the username and password of the admin/user that prohibits unauthorized users from accessing.	4.8	Very Satisfied
The system provides a backup function to restore data in the database.	4.8	Very Satisfied
Password is case-sensitive.	4.75	Very Satisfied
The system is free from bugs that disallow unauthorized users to access it.	4.75	Very Satisfied
The system's security prevents anyone from duplicating and installing the codes on their computer.	4.85	Very Satisfied
Grand Mean	**4.79**	**Very Satisfied**

Table 7 presents the results of the system effectiveness evaluation. It achieved a grand mean of 4.95, indicating outstanding effectiveness and performance. It meets user expectations, securely stores input data, generates accurate information quickly, enhancing productivity and accuracy. According to [14], performance characteristics, such as compatibility, usability, and comparative advantages, are crucial influences on how people interact with software.

In terms of data accuracy, the functions/formula indicated in the system were reviewed and evaluated by a certified public accountant and approved for integration of the system by the Association President and Board of Directors (BOD), thus the data and its transactions related to financial data were accurate.

Table 7. Effectiveness

Effectiveness	Mean	Description
The system has all the functions and capacities I expected.	4.8	Very Satisfied
It can secure the data that is being stored.	4.8	Very Satisfied
The system generates complete information upon request.	4.8	Very Satisfied
I can finish the work quickly using the system.	4.8	Very Satisfied
Easy to use and control the system.	4.9	Very Satisfied
Grand Mean	**4.95**	**Very Satisfied**

IV. CONCLUSION

Adopting a Management Information System (MIS) can transform agricultural cooperatives, fostering efficiency, openness, and sector development. Agricultural cooperatives may achieve a more sustainable and successful future by incorporating cutting-edge technology and data-driven solutions. The developed system effectively met the requirements and needs of the CATAFA. The overwhelmingly positive feedback received indicates the system's success in fulfilling the association's needs. These findings have generated a strong desire within CATAFA to integrate the system as a fundamental component of their information management infrastructure.

V. RECOMMENDATION

To remove the need for administrative support, it is recommended that the system be enhanced to include a dashboard that allows members to see their loan and savings information online. Members will have more freedom and ease in handling their financial transactions with this added capability. It is highly recommended by the researchers to enhance the features of the system by optimizing tables in the database and additional notifications for the users availing loans and savings through SMS or Email notifications. In terms of security, it is advised to install firewalls to protect the system against possible attacks.

ACKNOWLEDGMENT

This is to acknowledge the support of Central Philippines State University in the publication of this paper. To the members of the Camingawan-Tagukon Farmers Association for their full support in the conduct of the study.

REFERENCES

[1] E. Tortia, V. Valentinov, and C. Iliopoulos, "Agricultural cooperatives," Journal of Entrepreneurial and Organizational Diversity, May 2013, doi: 10.5947/jeod.2013.002.

[2] A. Candemir, S. Duvaleix, and L. Latruffe, "AGRICULTURAL COOPERATIVES AND FARM SUSTAINABILITY – A LITERATURE REVIEW," J Econ Surv, vol. 35, no. 4, pp. 1118–1144, Sep. 2021, doi: 10.1111/joes.12417.

[3] R. Ajates, "An integrated conceptual framework for the study of agricultural cooperatives: from repolitisation to cooperative sustainability," J Rural Stud, vol. 78, pp. 467–479, Aug. 2020, doi: 10.1016/j.jrurstud.2020.06.019.

[4] M. Struijk, S. Angelopoulos, C. X. J. Ou, and R. M. Davison, "Navigating digital transformation through an information quality strategy: Evidence from a military organisation," Information Systems Journal, vol. 33, no. 4, pp. 912–952, Jul. 2023, doi: 10.1111/isj.12430.

[5] C. Iliopoulos, M. L. Cook, and F. Chaddad, "Agricultural cooperatives in netchains," Journal on Chain and Network Science, vol. 16, no. 1, pp. 1–6, Sep. 2016, doi: 10.3920/JCNS2016.x003.

[6] Y. Liu, R. Shi, Y. Peng, W. Wang, and X. Fu, "Impacts of Technology Training Provided by Agricultural Cooperatives on Farmers' Adoption of Biopesticides in China," Agriculture, vol. 12, no. 3, p. 316, Feb. 2022, doi: 10.3390/agriculture12030316.

[7] D. S. Paraforos et al., "A Farm Management Information System Using Future Internet Technologies," IFAC-PapersOnLine, vol. 49, no. 16, pp. 324–329, 2016, doi: 10.1016/j.ifacol.2016.10.060.

[8] M. J. Carrer, H. M. de Souza Filho, and M. O. Batalha, "Factors influencing the adoption of Farm Management Information Systems (FMIS) by Brazilian citrus farmers," Comput Electron Agric, vol. 138, pp. 11–19, Jun. 2017, doi: 10.1016/j.compag.2017.04.004.

[9] M. Kramer, "BEST PRACTICES IN SYSTEMS DEVELOPMENT LIFECYCLE: AN ANALYSES BASED ON THE WATERFALL MODEL," Review of Business & Finance Studies, vol. 9, no. 1, pp. 77–84, 2018, [Online]. Available: https://ssrn.com/abstract=3131958www.theIBFR.com

[10] A. Fatoni, K. Adi, and A. P. Widodo, "PIECES Framework and Importance Performance Analysis Method to Evaluate the Implementation of Information Systems," E3S Web of Conferences, vol. 202, p. 15007, Nov. 2020, doi: 10.1051/e3sconf/202020215007.

[11] Z. K. A. Baizal, D. H. Widyantoro, and N. U. Maulidevi, "Factors Influencing User's Adoption of Conversational Recommender System Based on Product Functional Requirements," TELKOMNIKA (Telecommunication Computing Electronics and Control), vol. 14, no. 4, p. 1575, Dec. 2016, doi: 10.12928/telkomnika.v14i4.4234.

[12] "Index Terms-User Interface Study, User Experience Theory, Design Process, Tools for creating user interfaces, and other essentials", [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en.

[13] M. H. Alshira'H, "The Effects of Usability and Accessibility for E-Government Services on the End-User Satisfaction," International Journal of Interactive Mobile Technologies, vol. 14, no. 13, pp. 78–90, 2020, doi: 10.3991/ijim.v14i13.14659.

[14] J. Fang, Z. Zhao, C. Wen, and R. Wang, "Design and performance attributes driving mobile travel application engagement," Int J Inf Manage, vol. 37, no. 4, pp. 269–283, Aug. 2017, doi: 10.1016/j.ijinfomgt.2017.03.003.

DevOps 2.0: Embracing AI/ML, Cloud-Native Development, and a Culture of Continuous Transformation

V. Hemanth Das Chowdary
Dept.of Computer Science Engineering
Koneru Lakshmiah University
Vaddeswaram, India
2100030195cseh@gmail.com

A. Shanmukh
Dept.of Computer Science Engineering
Koneru Lakshmiah University
Vaddeswaram, India
2100032087cseh@gmail.com

T. Phaneendra Nikhil
Dept.of Computer Science Engineering
Koneru Lakshmiah University
Vaddeswaram, India
2100032156cseh@gmail.com

B. Sandeep Kumar
Dept.of Computer Science Engineering
Koneru Lakshmiah University
Vaddeswaram, India
2100031257cseh@gmail.com

Feroz Khan
Dept.of Computer Science Engineering
Koneru Lakshmiah University
Vaddeswaram, India
ferozkhan@kluniversity.in

***Abstract*— This research study investigates the emerging domain of DevOps, assessing its capacity to fundamentally transform the processes of software development and delivery. A notable advancement in the field of DevOps is the amalgamation of AI/ML and artificial intelligence. These technologies possess significant potential to revolutionize the DevOps domain through the automation of tedious duties, including provisioning infrastructure, managing configurations, and identifying anomalies. This study highlights the increasing need for streamlined solutions. By harnessing the capabilities of AI and ML-driven tools, DevOps teams will have the ability to streamline their processes, enabling them to allocate resources to more strategic approaches such as nurturing innovation, addressing complex challenges, and making decisions based on data. This study will investigate the precise methods by which AI/ML can enhance DevOps practices, resulting in a software development environment that is more secure, efficient, and agile.**

Keywords—DevOps, Software development, Artificial intelligence (AI), Machine learning (ML), Continuous delivery, Containerization, Microservices, Cloud computing, Cultural transformation

I. INTRODUCTION

The software development industry experiences ongoing changes due to the insatiable demand for products that demonstrate improved speed, security, and quality. DevOps, an approach that eliminates the traditional separation between development and operations, has emerged as an integral element in this ever-changing environment. To ascertain the ongoing impact of DevOps on the software delivery process, this article examines the burgeoning domains that are influencing its course. [19]

The progress of DevOps is significantly accelerated through the integration of machine learning and artificial intelligence. The process of workflow optimization involves the automation of repetitious duties such as infrastructure provisioning, configuration administration, and anomaly detection. This sentiment aligns with the forecast provided by Market Research Guru in their 2023 report, which indicates that the DevOps market will experience a compound annual growth rate (CAGR) of 24.61% until the year 2028. This illustrates the growing need for solutions that are efficient in nature. By enabling teams to concentrate on innovation, strategic decision-making, and complex challenges, AI and ML-enabled tools will liberate invaluable human resources. [3]

Additional advancements are introduced by the proliferation of containerization and microservices implementations. The deliberate development of containerized applications in conjunction with the microservices architecture enables them to demonstrate improved efficiency, resilience, and scalability, thus facilitating rapid deployments. As a result, organizations may experience an increase in revenue growth and attain a competitive advantage due to a reduced time-to-market. Moreover, the proliferation of cloud computing will inexorably give rise to applications that are native to the cloud. A favorable environment currently exists for developers to capitalize on the anticipated growth of the market by creating cloud-optimized applications that offer optimized resource allocation, streamlined administration, and improved performance.

It continues to be crucial to promote cultural transformation, surmount obstacles, and sustain open dialogue within diverse teams. Furthermore, it is imperative to sustain an outlook on perpetual learning and adjustment within this perpetually evolving domain. Despite these challenges presenting distinct prospects for advancement and ongoing refinement, DevOps remains the prevailing methodology in software delivery.

It is crucial to possess a thorough comprehension of the evolution of DevOps methodologies in comparison to previous years in order to effectively navigate the dynamic software development environment of 2024. This form of knowledge enables individuals to predict forthcoming trends and the possible consequences that may result.

II. DEVOPS MARKET EVOLUTION

A. *DevOps Evolution and Key Findings (2014–2023):*

The ever-changing environment of software development is motivated by an unwavering demand for products that demonstrate improved speed, security, and quality. DevOps has become an essential element in this ever-changing landscape, fostering cooperation and connecting the conventional gap that existed between development and operations. This study examines the burgeoning disciplines that are influencing the evolution of

DevOps, delving into its ongoing revolutionization of the software delivery process. [1]

T The domain of DevOps is positioned to undergo a significant paradigm shift as a result of the revolutionary potential of machine learning and artificial intelligence. These technologies enhance workflow efficiency by automating tedious duties such as infrastructure provisioning, configuration management, and anomaly detection. The figure mentioned above illustrates the compound annual growth rate (CAGR) of 24.61% for the DevOps market, as reported by Market Research Guru in 2023, and expected to persist until 2028. This illustrates the growing need for streamlined and automated solutions. By utilizing AI/ML-driven tools, organizations can effectively allocate the finite human resources of their teams to foster innovation, facilitate strategic decision-making, and address complex problems. [2,3,8]

The ongoing progression is supported by the growing implementation of containerization and microservices. The potential benefits of incorporating lightweight encoding into the microservices architecture of containerized applications include improved scalability, resilience, and deployment velocity. This phenomenon enables organizations to achieve a more rapid time-to-market, which may confer them a competitive advantage and stimulate revenue growth. Likewise, as cloud adoption expands at an exponential rate, cloud-native development will inevitably advance. Developers can achieve a multitude of advantages by adapting their applications to the cloud environment. These advantages encompass improved performance, efficient allocation of resources, and streamlined management. These developments align with the anticipated growth of the market. [16]

Nonetheless, technological advancements do not exclusively determine the trajectory of DevOps. Ensuring efficient communication channels, fostering cultural sensitivity, and surmounting obstacles continue to be of utmost importance within the framework of diverse teams. Furthermore, in this dynamic industry, it will be crucial to adopt a mindset that is defined by continuous learning and adaptation. On the contrary, these obstacles offer distinct prospects for advancement and ongoing enhancement, thereby guaranteeing that DevOps maintains its prevailing position in software delivery.
Given the ongoing evolution of the software development industry in 2024, it is imperative to comprehend the significant changes that have occurred in DevOps methodologies since their inception. Attaining this degree of comprehension enables us to predict forthcoming patterns and the possible consequences that might result. [19]

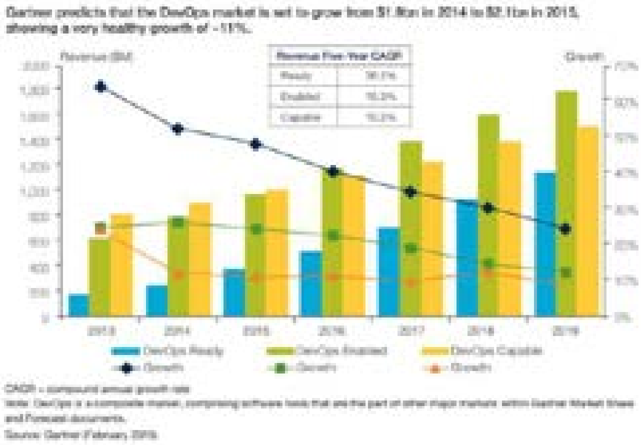


Figure 1. Gartner’s DevOps Market Predictions [9]

Additional factors that have propelled the DevOps market forward include the expansion of concepts like micro-service design, service virtualization, containerization, and platform-as-a-service, as well as the substantial support from the open-source community in the form of numerous cost-effective DevOps enabling tools, in addition to the research conducted by major product and tool-based companies and cloud providers, which has increased cloud adoption. Early adopters or "DevOps Unicorns" such as Netflix, Amazon, Google, Etsy, and Snapchat have established a track record of delivering effective DevOps model variations via ongoing innovation, including NoOps, ChatOps, and SmartOps. In a recent study, Rackspace4 surveyed 700 IT decision makers, of which 55% indicated that their respective organizations have embraced or executed DevOps and are currently exploring ways to improve upon it. Further, 31% of respondents plan to deploy DevOps infrastructure within the following two years. This represents a highly extensive technological implementation concerning the initial provision of instruments. Presently, the DevOps movement incorporates a positive organizational change resulting from the security and compliance components of the DevOps platform, in addition to the administration and implementation of technology.[6]

Exploring the Horizon: Potential Directions for 2024 and Beyond

The software development environment experiences ongoing changes due to the insatiable demand for products that demonstrate improved speed, security, and quality. DevOps, through its integration of operations and development, has surfaced as a pivotal pillar in this continuously evolving ecosystem. It fosters an environment of collaboration. In this article, the potential for DevOps to bring about a substantial paradigm shift in the software delivery process is examined by analysing the burgeoning frontiers that are influencing its course. [19]

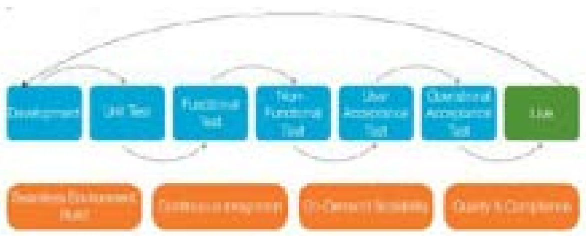


Figure 2. DevOps Lifecycle [9]

As a result of the transformative capabilities of artificial intelligence and machine learning, the DevOps sector is poised to experience a seismic shift. These technologies facilitate process optimization through the automation of laborious duties such as infrastructure provisioning, configuration management, and anomaly detection. This aligns with the findings of the Market Research Guru report 2023, which forecasts a compound annual growth rate (CAGR) of 24.61% for the DevOps market from 2019 to 2028. This observation suggests a growing need for streamlined and automated solutions. AI and ML-driven tools will empower teams to allocate their efforts towards innovation, strategic decision-making, and addressing complex challenges, thereby liberating valuable human capital. [3,8,16]

Additional progress is made possible by the extensive implementation of containerization and microservices. The integration of microservices architecture with lightweight encoding for containerized applications ensures improved scalability, resilience, and deployment velocity. This enables organizations to potentially increase their competitive advantage and broaden their revenue stream through a decreased time-to-market. Additionally, with the ongoing rise in cloud adoption, the emergence of cloud-native development is an unavoidable consequence. Developers possess the potential to capitalize on anticipated market growth by developing cloud-optimized applications that exhibit advantages such as optimized resource allocation, streamlined operations, and enhanced performance.

Difficulty-solving, fostering cultural sensitivity, and guaranteeing efficient communication are all paramount concerns in the context of diverse organizations. Furthermore, in this dynamic domain, it will be crucial to recognize and embrace the importance of ongoing education and adjustment. Notwithstanding these obstacles, they offer distinctive prospects for advancement and ongoing enhancement, thereby guaranteeing that DevOps maintains its prevailing position in software delivery. [17]

It is imperative that we possess a thorough comprehension of the evolution of DevOps methodologies in comparison to previous years in order to effectively navigate the dynamic software development environment of 2024. This comprehension empowers us to predict forthcoming trends and the possible consequences that may ensue.

Key Trends

The software development industry is continuously undergoing changes as a result of the constant need for products that demonstrate improved efficiency, security, and quality. By facilitating collaboration and bridging the traditional gap between operations and development, DevOps has emerged as an indispensable element in this ever-changing ecosystem. This research paper investigates the state-of-the-art advancements in DevOps and its capacity to profoundly revolutionize the future of software delivery.

The implementation of artificial intelligence and machine learning (AI/ML) within the DevOps domain holds the potential for substantial change. These technologies optimize efficiency through the automation of repetitive duties such as infrastructure provisioning, configuration management, and anomaly detection. This is consistent with the compound annual growth rate (CAGR) forecast of 24.61% provided by Market Research Guru in 2023 for the DevOps market until 2028. This suggests that the demand for efficient solutions is on the rise. Through the utilization of tools facilitated by AI and ML, teams can enhance the allocation of their valuable human resources in a manner that promotes innovation, facilitates strategic decision-making, and effectively tackles complex challenges. [3,8]

The rapid expansion of containerization and microservices presents additional prospects for progress. The integration of microservices architecture with lightweight encoding for containerized applications results in improved scalability, resilience, and deployment velocity. This phenomenon induces a reduction in the time required for products or services to reach the market, which may provide companies with a competitive advantage and stimulate revenue growth. Additionally, as a result of the widespread adoption of clouds, cloud-native development will inevitably emerge. Programmers can achieve various advantages by developing cloud-native applications, such as optimized resource utilization, enhanced performance, and streamlined administration. These benefits are consistent with the anticipated growth of the market. [7]

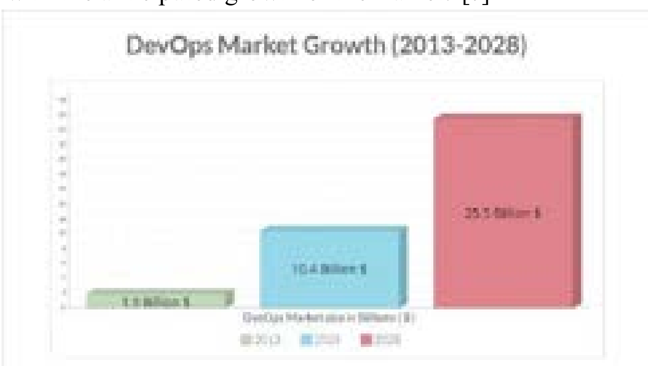


Figure 3. DevOps Market Growth (2013 - 2028)

However, technological advancements do not exclusively dictate the future course of DevOps. Fostering cultural transformation, and preserving transparent communication among members of diverse teams continue to be crucial imperatives. Furthermore, it will be essential to embrace continuous learning and adaptation, as well as to acquire a mindset of perpetual change. Nevertheless, these obstacles offer distinct prospects for advancement and ongoing refinement, guaranteeing that DevOps retains its prevailing position in software delivery. [3,4.7]

Given the ever-changing nature of the software development sector in 2024, it is crucial to grasp the substantial progression of DevOps methodologies relative to their precursors. This understanding empowers us to anticipate future trends and their potential ramifications.

B. *AI and Machine Learning (AI/ML): Redefining DevOps Efficiency*

The software development industry experiences ongoing changes due to the insatiable demand for products that demonstrate improved speed, security, and quality. DevOps, an approach that eliminates the traditional separation between development and operations, has emerged as an integral element in this ever-changing environment.

The progress of DevOps is significantly accelerated through the integration of machine learning and artificial intelligence. The process of workflow optimization involves the automation of repetitious duties such as infrastructure provisioning, configuration administration, and anomaly detection. This sentiment aligns with the forecast provided by Market Research Guru in their 2023 report, which indicates that the DevOps market will experience a compound annual growth rate (CAGR) of 24.61% until the year 2028. This illustrates the growing need for solutions that are efficient in nature. By enabling teams to concentrate on innovation, strategic decision-making, and complex challenges, AI and ML-enabled tools will liberate invaluable human resources. [3,8]

1. Automating Repetitive Tasks:

An inherent and noteworthy contribution of AI/ML to the DevOps paradigm is its capacity to mechanize laborious and repetitive duties. Infrastructure provisioning, configuration administration, and anomaly detection are all encompassed within this. Through the automation of these aforementioned duties, AI/ML enables operations teams and developers to allocate their valuable human resources towards more advanced endeavors such as problem-solving, strategic decision-making, and innovation. (references: [3, 8]).

2. Infrastructure Provisioning:

Traditionally, infrastructure provisioning was a manual process that involved configuring servers, storage, and networking resources. AI/ML can automate this process by learning from historical data and user preferences. This not only reduces the time and effort required for provisioning but also minimizes the risk of human error [5]

3. Configuration Management

Consistency in configurations across various environments is an essential factor in guaranteeing dependable software deployments. Configuration management can be automated through the use of AI/ML, which identifies deviations from established baselines and perpetually monitors configurations. This practice aids in the prevention of configuration drift and guarantees uniform deployments across diverse environments. [6].

4. Anomaly Detection

System performance or application behaviour anomalies may serve as indicators of impending issues. By training AI/ML algorithms to detect these anomalies in real-time, teams can proactively resolve problems prior to their impact on users.

5. Improved Decision-Making

AI/ML are capable of analysing immense quantities of data originating from a variety of sources, such as code repositories, monitoring tools, and application logs. By identifying trends and patterns in this data, AI/ML can provide DevOps teams with actionable insights.

Additional advancements are introduced by the proliferation of containerization and microservices implementations. The deliberate development of containerized applications in conjunction with the microservices architecture enables them to demonstrate improved efficiency, resilience, and scalability, thus facilitating rapid deployments. As a result, organizations may experience an increase in revenue growth and attain a competitive advantage due to a reduced time-to-market. Moreover, the proliferation of cloud computing will inexorably give rise to applications that are native to the cloud. A favourable environment currently exists for developers to capitalize on the anticipated growth of the market by creating cloud-optimized applications that offer optimized resource allocation, streamlined administration, and improved performance. [16]

It is crucial to possess a thorough comprehension of the evolution of DevOps methodologies in comparison to previous years in order to effectively navigate the dynamic software development environment of 2024. Acquiring this knowledge enables us to proactively predict forthcoming trends and the possible consequences that may ensue.

A. DevOps and the Cloud: A Strategic Synergy:

DevOps and cloud computing are two transformative forces that, when combined, can revolutionize software development and delivery. Here's how the cloud achieving process aligns with DevOps methods, with references to illustrate these points:

1. Infrastructure as a Code (IaaC):

Infrastructures as Code (IaaC) enable developers to define infrastructure configurations, thereby facilitating the automated deployment and provisioning of resources in the cloud. This facilitates accelerated deployments, consistency, and repeatability ([20]). Prominent cloud platforms such as AWS CloudFormation [24], Azure Resource Manager [25], and Google Cloud Deployment Manager [26] facilitate the automation of infrastructure provisioning and configuration in the cloud environment through the creation of IaaC templates.

Figure 4. Continuous Delivery pipeline [21]

2. Continuous Integration and Continuous Delivery (CI/CD):

CI/CD pipelines streamline the software development lifecycle through the automation of code integration, test execution, and cloud-based update deployment. This increases the rate of feedback loops and decreases the likelihood of errors ([23]). Cloud platforms offer integrated CI/CD tools, such as Google Cloud Build [29], AWS Code Pipeline [27], and Azure DevOps Pipelines [28], which facilitate the deployment of applications to the cloud and integrate seamlessly with development workflows.

3. Continuous Monitoring and Feedback:

DevOps emphasizes continuous monitoring of application performance and infrastructure health in the cloud. This allows for proactive problem identification and faster resolution ([22]). Cloud platforms offer comprehensive monitoring services like AWS CloudWatch [30], Azure Monitor [31], and Google Cloud Monitoring [32] that provide real-time insights into application performance, resource utilization, and infrastructure health.

4. Scalability and Elasticity:

Cloud resources enable DevOps teams to dynamically adjust the scalability of applications in response to fluctuations in demand. This reduces expenses by eliminating the requirement for manual infrastructure provisioning. Cloud computing provides elastic, on-demand resources that are automatically scalable in accordance with the demands of the application. This ensures cost-effectiveness and optimal performance.

5. Collaboration and Communication:

By facilitating collaboration among the development, operations, and security teams, cloud environments offer a unified and shared platform for the management of applications and infrastructure. This facilitates enhanced communication and optimized operations. A multitude of cloud platforms provide collaboration functionalities, such as wikis, chat, and project management features, which facilitate enhanced teamwork within a cloud-based setting (AWS Chat [33], Azure DevOps Boards [34]).

Cloud computing offers an optimal infrastructure for the streamlined implementation of DevOps methodologies. Through the strategic integration of these two methodologies, enterprises can augment the flexibility, effectiveness, and ingenuity inherent in their software development procedures.

III. DEVOPS MARKET GROWTH (2013-2028)

The graphical representation depicts the projected growth of the global DevOps industry, with a significant upward trend from 2013 to 2028. In 2013, the market was assessed to be worth $8.8 billion, indicating that the implementation of DevOps was in its early stages of development. By 2023, it is projected that the market will attain a valuation of $10.4 billion, signifying an ongoing ascent in its importance within the software development sector. Anticipating forthcoming advancements, it is anticipated that the market will sustain its expansion, reaching an approximate valuation of $25.5 billion by 2028. The expected growth represents the increasing adoption of DevOps approaches across various industries, driven by the numerous benefits it offers—including accelerated software delivery, strengthened collaboration, and enhanced operational efficiency. [3]

TABLE I. DEVOPS MARKET GROWTH (*$ IN BILLION*)

Year	Market Size (2013-2028 *expected*)
2013-2014	$1.9 billion
2015-2016	$2.3 billion
2016-2018	$3.709 billion
2018-2023	$10.4 billion
2024-2028	$ 25.5 billion (*expected*)

Over the designated time span (2013-2028), it is expected that the market will experience significant growth. The market is anticipated to increase from a range of $1.9 billion to $2.4 billion in 2013-2014 to $24.5 billion between 2024 and 2028. The market size is projected to average $5.67 billion between 2013 and 2023, excluding the anticipated value for 2024-2028.[11,10,9,19]

B. Enhancing Software Delivery with DevOps

DevOps practices not only streamline development processes but also demonstrably improve core software delivery aspects. Here's how:

By leveraging containerization and Infrastructure as Code (IaC), DevOps prioritizes automation ([35], [36]). IaC guarantees the provisioning of consistent infrastructure, which results in predictable performance. The horizontal scalability of microservices architectures made possible by containerization enables the fulfillment of increased demand ([37]). Furthermore, continuous monitoring is advocated for in DevOps ([38]), which enables the proactive detection and resolution of performance obstacles.

Security constitutes a fundamental principle of DevOps. Varies can be detected and remedied in a timely manner by incorporating security testing into the entirety of the development lifecycle (DevSecOps) ([39]). By taking this proactive approach, security risks in production environments are minimized. In addition, security is improved through the utilization of tools such as containerization, which function to isolate applications and their dependencies, thereby diminishing the attack surface ([40]).

DevOps facilitates communication and collaboration between the operations and development teams, resulting in streamlined workflows and decreased revisions ([41]). The utilization of CI/CD pipelines to automate repetitive duties not only increases efficiency but also liberates developer time for more valuable endeavours ([42]). Furthermore, DevOps cultivates an environment that encourages perpetual refinement and learning, thereby facilitating the efficient execution of software delivery.

IV. CONCLUSION

TThe software development industry is continuously undergoing changes as a result of the constant need for products that demonstrate improved efficiency, security, and quality. By facilitating collaboration and bridging the traditional gap between operations and development, DevOps has emerged as an indispensable element in this ever-changing ecosystem. This research paper investigates the state-of-the-art advancements in DevOps and its capacity to profoundly revolutionize the future of software delivery.

The implementation of artificial intelligence and machine learning (AI/ML) within the DevOps domain holds the potential for substantial change. These technologies optimize workflow efficiency through the automation of repetitive duties such as infrastructure provisioning, configuration management, and anomaly detection. This is consistent with the compound annual growth rate (CAGR) forecast of 24.61% provided by Market Research Guru in 2023 for the DevOps market until 2028. This suggests that the demand for efficient solutions is rise. Through the utilization of AI ML, organizations can optimize the distribution of their human capital in support of innovation, decision-making, and resolution of problems. [3,8]

The rapid expansion of containerization and microservices presents additional prospects for progress. The integration of microservices architecture with lightweight encoding for containerized applications results in improved scalability, resilience, and deployment velocity. This phenomenon induces a reduction in the time required for products or services to reach the market, which may provide companies with a competitive advantage and stimulate revenue growth. Additionally, as a result of the widespread adoption of clouds, cloud-native development will inevitably emerge. Programmers can achieve various advantages by developing cloud-native applications, such as optimized resource utilization, enhanced performance, and streamlined administration. These benefits are consistent with the anticipated growth of the market. [17]

However, technological advancements do not exclusively dictate the future course of DevOps. Cultivating cultural transformation, surmounting obstacles, and sustaining transparent communication continue to be of the utmost importance in diverse teams. Furthermore, it will be essential to embrace continuous learning and adaptation, as well as to acquire a mindset of perpetual change. Nevertheless, these obstacles offer distinct prospects for advancement and ongoing refinement, guaranteeing that DevOps retains its prevailing position in software delivery.

In light of the dynamic nature of the software development industry in 2024, it is imperative to comprehend the significant advancements that DevOps methodologies have made in comparison to their predecessors. This comprehension enables us to predict forthcoming trends and the possible consequences they may involve.

REFERENCES

[1] PuppetLabs. (2014). State of DevOps Report.https://www.puppet.com/resources/state-of-devops-report.

[2] PuppetLabs The 2021 history of the DevOps Report https://www.puppet.com/resources/history-of-devops-reports

[3] LinkedIn article - DevOps Market Growth Evolution and Future Demand by 2031 Authors:https://www.linkedin.com/company/business-news-research-updates/ https://www.linkedin.com/pulse/devops-market-growth-evolution-future-demand-jxoje/

[4] State of DevOps 2023 - report by DORA Authors:https://www.linkedin.com/in/zia-tahir-5794996/ https://www.linkedin.com/pulse/state-devops-2023-report-dora-zia-tahir-5mqvf/

[5] DORA. (2023). Accelerate State of DevOps Report.

https://cloud.google.com/devops/state-of-devops.

[6] Linkedin: Growth And The Future of DevOps in the Next 10 Years. Author-https://www.linkedin.com/in/abdul-rahman-018498264/

https://www.linkedin.com/pulse/growth-future-devops-next-10-years-abdul-rahman-/

[7] Gartner report: https://www.gartner.com/en/newsroom/press-releases/2021-11-10-gartner-says-cloud-will-be-the-centerpiece-of-new-digital-experiences

[8] Techvify-software AI in DevOps – The Definite Guide:

https://techvify-software.com/ai-in-devops/

[9] https://www.capgemini.com/wp /content/uploads/2017/07/devops_pov_2015-12-18_final.pdf

[10] https://www.marketsandmarkets.com/Market-Reports/devops-market-824.html

[11] https://www.grandviewresearch.com/industry-analysis/development-to-operations-devops-market

[12] https://dx.doi.org/10.1016/0024-6301(90)90323-v

[13] https://www.businessjournalng.com/global-wealth-rose-to-250-trn-in-2020-despite-covid-19-pandemic/

[14] Policy Changes in Global Higher Education: What Lessons Do We Learn from the COVID-19 Pandemic? | Higher Education Policy (springer.com)

[15] The need for a professional network of agricultural and biosystems - M.MOAM.INFO

[16] https://dx.doi.org/10.1016/j.jclepro.2020.119998

[17]https://www.euro.who.int/__data/assets/pdf_file/0011/199532/Health2020-Long.pdf

[18] https://link.springer.com/article/10.1057/s41307-022-00266-0

[19] https://www.gartner.com/en/newsroom/press-releases/2015-03-05-gartner-says-by-2016-devops-will-evolve-from-a-niche-to-a-mainstream-strategy-employed-by-25-percent-of-global-2000-organizations#:~:text=Gartner%20Says%20By%202016%2C%20DevOps,Percent%20of%20Global%202000%20Organizations

[20] What is Infrastructure as Code (IaC)? (HashiCorp): https://developer.hashicorp.com/terraform/tutorials/aws-get-started/infrastructure-as-code

[21] Continuous Integration and Delivery (Atlassian): https://www.atlassian.com/continuous-delivery

[22] The DevOps Handbook: Gene Kim, Jez Humble, Patrick Debois (https://books.google.com/books/about/The_DevOps_Handbook.html?id=XrQcrgEACAAJ)

[23] Why DevOps? (Microsoft): https://www.veritis.com/blog/what-are-the-benefits-of-azure-devops/

[24] AWS CloudFormation: https://aws.amazon.com/cloudformation/

[25] Azure Resource Manager: https://learn.microsoft.com/en-us/azure/azure-resource-manager/

[26] Google Cloud Deployment Manager: https://cloud.google.com/deployment-manager/docs

[27] AWS CodePipeline: https://aws.amazon.com/codepipeline

[28] Azure DevOps Pipelines: https://learn.microsoft.com/en-us/azure/devops/pipelines/?view=azure-devops

[29] Google Cloud Build: https://cloud.google.com/build/docs/overview

[30] Amazon CloudWatch: https://aws.amazon.com/cloudwatch/

[31] Azure Monitor: https://learn.microsoft.com/en-us/azure/azure-monitor/

[32] Google Cloud Monitoring: https://cloud.google.com/monitoring

[33] AWS Chat: While AWS doesn't offer a dedicated chat service within its core offering, many teams leverage tools like Amazon Chime (https://app.chime.aws) or integrate third-party chat solutions into their AWS environment.

[34] Azure DevOps Boards: https://learn.microsoft.com/en-us/azure/devops/boards/?view=azure-devops

[35] Blanc, X., & Catalonia, R. (2020). DevOps for cloud security. Packt Publishing Ltd.

[36] Bass, L., Clements, P., & Kazman, R. (2012). Software architecture in practice (3rd ed.). Addison-Wesley.

[37] Newman, S. (2015). Building microservices: Designing fine-grained systems. O'Reilly Media.

[38] Beyer, H.-L., Jones, C., Peeling, D., & Murphy, B. (2016). Site reliability engineering: How Google runs production systems. O'Reilly Media.

[39] Humble, J., Farley, D., & Farley, P. (2010). Continuous delivery: Reliable software releases through build, test, and deployment automation. Addison-Wesley Professional.

[40] Chowdhary, M., VanderLinden, T., Brecht, P., & Burns, B. (2017). The container security spectrum: How to keep docker safe. O'Reilly Media.

[41] Eppinger, S. D., & Chitale, M. K. (2003). The innovation cycle: Linking product development to customer usage. Business Horizons, 46(4), 35-49.

[42] Humble, J., Farley, D., & Farley, P. (2010). Continuous delivery: Reliable software releases through build, test, and deployment automation. Addison-Wesley Professional.

A Comparative Analysis of RNNs, GRUs and LSTMs in Machine Translation and Sentiment Analysis

1st Amit Gupta
Computer Science and Engineering
Graphic Era Hill University
Dehradun, India
amitgupta7920@gmail.com

2nd Bhavya Shastri
Computer Science and Engineering
Graphic Era Hill University
Dehradun, India
bhavyashastri122002@gmail.com

3rd Utsav Nautiyal
Computer Science and Engineering
Graphic Era Hill University
Dehradun, India
utsavnautiyal2@gmail.com

4th Kanupriya
Computer Science and Engineering
Graphic Era Hill University
Dehradun, India
kanupriyathakur1101@gmail.com

5th Navin Garg
Computer Science and Engineering
Graphic Era Deemed to be University, Dehradun, India
navin.garg6@gmail.com

***Abstract*—Machine translation is the act of translating from one language to another using a computer program or algorithm.The need for translation from one language to another to facilitate communication between different cultures and peopleis one that we have had since antiquity. Traditional methods of translation are quite slow and cumbersome and are susceptible to biases from the translator. The advent of computers has led us to explore ways in which this process may be automated. Another long-standing problem is the detection and classification of emotions in textual data. The capacity to detect and recognize human emotions is one that is particularly fascinating and important in the field of Artificial Intelligence. Both problems are NLP(natural language processing) problems that involve analyzing textual data which is inherently sequential in nature. Traditional deep neural networks have proven inefficient and inaccurate in tasks involving sequential data such as natural language processing, time series analysis etc. Over the years numerous deep learning models have been developed to address these issues. The advent of RNNs(Recurrent Neural Networks), LSTMs(Long Short-Term Memory) and GRUs(Gated Recurrent Units), which are all deep neural net- works built to handle sequential data, has made great progress in machine translation and sentiment analysis. Each of these models has its own unique characteristics and features. This paper aims to compare each model in terms of accuracy, performance, architecture, computational complexity, etc. and thus outline situations in which each model may prove suitable based on its particular set of strengths and weaknesses.**

***Index Terms*—Machine Learning, Machine Translation, Sentiment Analysis, Recurrent Neural Network(RNN), Natural Language Processing(NLP), Long Short-Term Memory(LSTM), Gated Recurrent Unit(GRU)**

I. INTRODUCTION

In the changing province of Natural Language Processing (NLP), analysts are vigorously trying to increase machine abilities in language translation and sentiment analysis. The pursuit for improved precision and ability in these domains has driven an engaged exploration into the difference of various types of artificial intelligence models. There are three types of models: RNNs(Recurrent Neural Networks), LSTMs(Long Short-Term Memory networks) and GRUs (Gated Recurrent Units). These models are a few different versions of recurrent neural networks, each designed to label particular challenges in sequence-based tasks. Analysts are actively looking into the strengths and limits of RNNs, GRUs, and LSTMs to expose acumen that can move development in language processing, offering a separate understanding of their quality of being connected in machine translation and feeling analysis tasks.

In the sector of consecutive data processing, where the encouragement to fight lies in understanding and analysing information presented in a sequence, such as a series of words that make sense i.e. sentence or paragraphs, Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs), and Long Short-Term Memory (LSTM) networks appear as very important tools. Think about these models as skilled lists of questions helping computers' understanding and processing of consecutive information, especially in the involved domain of languages. Their important role becomes obvious in un- folding the structure and big picture of language, basic for tasks like translation and sentiment analysis. RNNs(Recurrent Neural Networks) are proficient in managing information and data which is sequential. LSTMs(Long Short Term Memory) are similar to GRUs in some ways and also proficient in memory i.e. LSTM remembers large pieces of information especially when it comes to language structures which are complex. On the far side of the prevailing metrics of language translation or understanding of feeling of love, hate, guilt, etc. , the analysis is geared towards detecting the distinctive strengths and challenges of each tool. The purpose is to gain a confusing understanding of how these systems entangle the attribute of language, communication, feelings of love, hate, fear, etc. and sentiments with skill. This in-depth research aims not only to give basic information but to offer complete and thorough consciousness into the analogous prospect of RNNs, GRUs, and LSTMs within the particular big picture of machine translation and feeling analysis. The purpose is to give power to analysts, professionals or skilled people, and fans with high-quality knowledge that helps informed decision-making. Through this serious and stubborn pursuit, the goal is to light the abilities of RNNs, GRUs, and LSTMs, throwing out light on where each excels and identifying possible challenges they may come across in language tasks. Eventually, to prepare the collaborators in this field where tools are required and the understanding for neural network infrastructure is needed effectively, so that one can combine applications or can say hybrid applications ranging from translation tasks to study feelings in word-based data. While exploring these various types of neural networks i.e. Recurrent Neural Networks(RNNs), Gated Recurrent Units(GRUs) and Long Short-Term Memory(LSTM) networks, it's not only a technical analysis but there is more to it. As it is a passage to understand about the language of humans and their sentiments. People will get to know about the intricacies of communication between humans, how they understand and express.

Before conducting any research or analysis, feasibility study is one of the important step that needs to be done. As for Natural Language Processing(NLP) tasks like machine translation and sentiment analysis, various models and architectures have been introduced. For machine translation, starting from rule based machine translation which provided quality translation but it was customized for specific purposes only. Coming to statistical machine translation, provided accurate translations using Machine Learning(ML) algorithms but was time-consuming. Neural machine translation model, addressed the limitations of previous models by using neural networks to learn languages. For sentiment analysis, lexicon based and machine learning based models are being used. Lexicon based models are less expensive and have access but are not suitable for analysis where data gathers from different platforms. While machine learning models provide more accuracy and are faster than traditional methods. By uncovering the abilities, strengths, and difficult things about RNNs, GRUs, and LSTMs, the aim is to signal the way for smart systems that not only understanding the surface-level true meaning of human language but also understanding the sentiments, and the difference that define human communication. By throwing light on the complex interaction between technology and these basic parts of people and the kindness of people, one can think of a future where AI not only helps in practical tasks but also becomes an understanding partner. This deeper understanding of language and emotions is not just about increasing the capability of translation or sentiment analysis; it's about taking good care of a high-tech province that matches up with the human exposure.

II. Literature Survey

Starting with an extensive historical summary, the survey meticulously chronicles the development of these recurrent neural network architectures. Recurrent Neural Network (RNN) and Convolutional Neural Network (CNN) are the best popular types of Deep Learning architectures used.

Patel et.al. [1] introduced the two approaches for sentiment classification and those were lexicon based and machine learning based. The deep learning model Recurrent Neural Network(RNN) was introduced. But the training process quite time-consuming due to the vanishing gradient problem, so the Long Short-Term Memory(LSTM) network, an advanced version of RNN model came up with the solutions to those problems. Srinivas et.al. [2] proposed a system for sentiment analysis using neural networks, Convolutional Neural Network(CNN) and Long Short-Term Memory(LSTM), out of which LSTM came out as the best approach. A twitter dataset was taken for the analysis of sentiments, but when it comes to analyze data in real-time, this proposed system will need other approaches.

Switching over to sentiment analysis, textual data plays an important role in classifying emotions. Murthy et. al. [3] conducted the text based analysis of IMDB dataset by using Long Short-Term Memory(LSTM) as Recurrent Neural Network(RNN) fails to analyse long sequences and suffers from vanishing gradient. Long Short-Term Memory(LSTM) is a success over the traditional machine learning techniques. For accurate sentiment prediction, there is need for long range dependencies as well as the need to know the importance of how to capture contextual information. These analyses shed light on the detailed and thoughtful performance of each architecture across diverse datasets and sentiment analysis domains. Pedro M. Sosa [4] developed a model by combining LSTM and CNN to see how the performance will be affected by using their abilities, enhancing the importance of adding dropout layer. While comparing LSTM-CNN with regular CNN and LSTM, the accuracy of LSTM-CNN model is the highest.

In machine translation, there are a great variety of architectures. Statistical Machine Translation(SMT) was a success over the rule-based machine translation and when Neural Machine Translation(NMT) came into picture, replacement of SMT took place. Tan et. al. [5] implemented the concept of attention mechanism in machine translation and used many open source Neural Machine Translation(NMT) toolkits like Tenosr2Tensor, OpenNMT, Nematus for research. Datta et. al. [6] acknowledged the benefits of Recurrent Neural Network(RNN) over Artificial Neural Network(ANN). The dataset for neural machine translation was from open GitHub source containing English and French sentences. Speech Recognition, Machine Translation and Speech Synthesis, the three modules in which system was divided and the system also took care of overfitting issues and overheads.

Cho et. al. [7] implemented the RNN Encoder-Decoder model which proved to be better at capturing the linguistic regularities in multiple levels i.e. word level or phrase level, explained the overall improvement in the performance of translation. This Encoder-Decode model can also be applied to speech transcription. Alex Sherstinsky [8] discusses about the difficulties that were being faced by RNN and how the transformation from RNN to Vanilla LSTM take place. Also the importance of data set standardization and evolvement of Vanilla LSTM architecture is highlighted in the research. It provides valuable understanding of the broader situations and ability to change each architecture based on various contextual factors.

GRUs(Gated Recurrent Units) work by combining the hidden state with memory cell structures. Efficiency and improvements in translation tasks are provided by GRUs. According to Dey et. al. [9] , applying three distinct variants to both gates(update gate and reset gate) will help to evaluate performance, possessing less gate parameters and less computational expense when compared to GRU-RNN. Li et. al. [10] have described IndRNN model which provides independence of neurons and also allows to learn long-term dependencies by the network itself. As compared to LSTMs and GRUs, IndRNNs process much longer sequences.

In both machine translation and sentiment analysis, on the relative performance of various RNNs, the literature survey tells about the impact of various aspects like characteristics of dataset and methodologies. To connect people from all the world, share information among people and build relationships, language translation plays an important role. Kim et. al. [11] emphasized the use of various pivoting techniques, showcasing the drawbacks of other machine translation systems and performed a survey on different approaches of neural machine translation and found that for many languages, there is no availability of huge amounts of the parallel corpus. Bau et. al. [12] have addressed how translations can be controlled by modifying neurons and on the basis of linguistic properties like number, gender and tense. Also discussed about are the rankings of neurons which must be done on the basis of their importance and how the neurons which are on higher level have more impact on translation.

In the end, as for machine translation and sentiment analysis, the survey leads to give a better understanding about prior work and advancements in the fields of machine translation and sentiment analysis. So this survey can be serving as a valuable and useful dispense as one can gain knowledge from the various studies that have been done.

III. Proposed Methodology

To propose a thorough comparisons of LSTM (Long Short-Term Memory) networks, GRUs (Gated Recurrent Units) and RNNs (Recurrent Neural Networks), in the domain of Machine Translation and Sentiment Analysis, a well-designed method has been proposed in this paper which gives a prologue to the minute differences of these prominent neural networks. The methodology begins with the careful selection from a variety of datasets available on the internet, specially those which capture the complexities and difficulties present in tasks such as sentiment analysis, involving a wide range of emotional content and translation between 2 different languages. Due to the variety of datasets present on the internet, the robustness and applicability of the of the results is increased greatly in real-world. It further progresses with data preprocessing phase which is standardized to establish uniformity in handling of the selected dataset. The data pre-processing phase is a crucial phase as it involves activities such as tokenization which ensure that the text is converted into separate components for analysis and removal of noises and inconsistencies. By standardizing the Pre-Processing phase, an equal and impartial basis is set up for LSTM, GRU and RNN models. The main purpose is to provide insightful information on the advantages and disadvantages of each of the prominent models/networks with respect to tasks such as Language Translations and Sentiment Analysis as mentioned before.

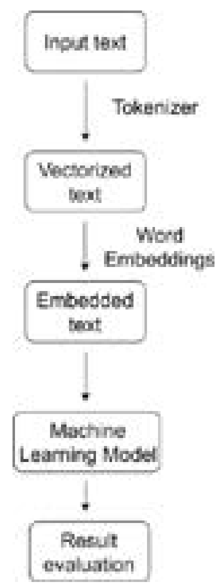


Fig. 1: Data flow Diagram

A. Dataset Description

A set of datasets will be extremely carefully collected on the analysis of Recurrent Neural Networks (RNNs),Gated Recurrent Units (GRUs), and Long Short-Term Memory (LSTM) networks in the worlds of machine translation and sentiment/feeling analysis.

As for machine translation, the English-French translation dataset from Kaggle [13] has been taken. There are 2 columns one column has English words/sentences and the other one has French words/sentences. And these datasets can be used for language translation task. The unique value is different because the same English word has a different French representation. This dataset contains more than 176,000 pairs of sentences in English and French, which makes it a rich resource for studying language translation. This diversity provides ample material for analyzing and improving language translation algorithms.

On the sentiment analysis front, sentiment analysis dataset of Twitter [14] from Kaggle has been taken for sentiment analysis which encapsulates diverse array of tweets having sentiments ranging from positive to negative. Using this dataset, the feeling or the sentiment of a message can be judged. In this dataset, there are four categories : Tweet ID , Entity , Sentiment and Tweet Content in which the particular message can be put. With over 69,491 unique entries and 74,700 total entries, each tweet is labelled with it's associated sentiment such as positive , negative or neutral which provides a fertile ground to apply various sentiment analysis methodologies. These datasets will serve as the foundation for a thorough, detailed and thoughtful analysis of RNNs, GRUs, and LSTMs, which is an addition to the improvement in understanding of their performance in the specified language processing job.

B. Data Pre-processing

In the domain of Sentiment Analysis and Language Translation, the data Pre-Processing step is similar to careful creation of mastering the tools needed to do a surgery in a hospital. To help the models have good accuracy when a new data is introduced during Language Translation and Sentiment Analysis, a well-designed and processed dataset makes sure that the models are exposed to variety of language patterns and sentiment variations. Creation of this thorough dataset helps with issues such as language variations, cultural variances, and domain-specific differences. Some of the key benefits of data pre-processing are:

1) NLP models which heavily rely on data quality can function better.

2) All data inconsistencies which can hinder a models performance are carefully removed.

3) Accuracy of the predictions is improved.

4) Enables the model to handle diverse range of inputs.

5) Addresses bias and fairness.

6) Supports continuous refinement in order to capture nuances.

1. **Tokenizer on text:** Neural networks cannot process data in the form of text directly. Each word thus must be assigned a numerical representation unique to that particular word. The tokenizer works by assigning each unique word a unique numerical value. For e.g. consider the sentence "He is a boy" : 'he' will be assigned a value of '1', 'is' a value of '2' and so on and so forth. This collection of words with their corresponding values is called the dictionary. Any word not present in the dictionary is automatically assigned a value of 0. Punctuation marks are automatically ignored.

2. **Padding on text:** Padding: Padding is a simple concept in ML and Data Science wherein all the sentences in a text are same of equal lengths, this is often done by adding zeros to the shorter sentences. The consistency in the length of sequences makes it easy for the model to streamline process the entire data(which becomes uniform after padding).

Padding is useful where we want our system to identify the sentiment or emotion of a certain piece of text which can be positive, negative or neutral. It is used to guess the next word in a sentence , convert a sentence from one language to another and thereby helps in smooth integration of textual data.

3. **One hot Encoding labels:** One-Hot encoding is common ML technique which represents data in binary digits. It transforms textual input into information which is presented in the form of a binary matrix. Example: RGBs(colour variable) One-hot encoding would be as follows:

- Red: [1, 0, 0]
- Green: [0, 1, 0]
- Blue: [0, 0, 1]

It's particularly useful in cases where data has no ordinal relationship among categories. One-Hot encoding is performed by using libraries such as scikit-learn. Only disadvantage is that it increases the dimensionality of the data and hence the processing becomes complex. For such a problem, we can use techniques like: embedding or feature hashing.

4. **Word Embeddings:** Word embeddings are further applied to each word after tokenization in order to provide a better numerical representation of the word in the form of a vector. Word embeddings are used to capture and map the unique features of each word. These word embeddings are usually 300-500 dimensional matrices. Each dimension is used to capture a particular feature or information contained in a word. These may hypothetically capture features such as gender, tense, location etc. related to the word. Similar words will have similar word embedding representations.

5. **Train-test split:** Train-test split, as the name suggests is a method of dividing data into 2 parts i.e :

1) Training set which is basically to learn patterns from the data

2) Test set which is to evaluate the performance of new/ generated data

When a computer is fed with chunk of raw data, it becomes hard for it to analyse it as a whole, hence the model divides this big data into sets and uses these sets of data for different purposes in order to generate the best possible solution. This can be understood better with the help of an example: supposing we are given the task of simple classification like cats and dogs. In such a case, our model will divide the images in 2 groups, first one to practice in identifying the image and the second to test whether it's capable enough to predict the group of unknown images or not.

Thereby it can be concluded that Train-Test split is an amazing technique which enables the system to perform tasks in a real-world system wherein the data is vast and generally unknown. This model instead of memorizing the content, tries to figure out (using the training set) how to generate pattern and connections in data. It then uses these patterns (on test set) to give optimized solutions to the user's real world problems.In essence, the train-test split serves as a methodological safeguard against overfitting, contributing to the development of robust machine learning models capable of accurate predictions across diverse and dynamic scenarios.

C. Algorithms

Recurrent Neural Networks (RNN) are a category of Neural Network[15] where the output computed from the previous step is fed as the input to the next step. In fully connected neural networks, all the inputs are independent of each other, but if for instance if we wish to predict the next word in a sentence the sequence of the words arriving before it plays a significant role in its determination. In such cases in order to process sequential data we require the need to store information from the previous states as well. This is accomplished by a RNN(Recurrent Neural Network). We have various examples for RNN(REcurrent Neural Network). Some of them are predicting the next word of a sentence, Stock price prediction etc.

Key Components and Operation of an RNN

1. **Hidden State (hid_t):** At each time step t, an RNN maintains a hidden state h_t. The new hidden state is calculated based on the current input and the previous hidden state. It serves as a way to retain information from the previous inputs in the sequence.

$$hid_t = u(P_{hh} \cdot hid_{t-1} + P_{xh} \cdot x_t + b_h)$$

where P_{hh} is the hidden state's weight matrix,
P_{xh} is the input's weight matrix,
b_h is the bias term,
u is the selected activation function (commonly tanh or ReLU), and
x_t is the input at time t.

2. **Output (y_t):** The output at each time step t, denoted as y_t, is computed based on the current hidden state.

$$y_t = P_{hy} \cdot hid_t + c_y$$

where P_{hy} output's weight matrix, and c_y is the bias term.

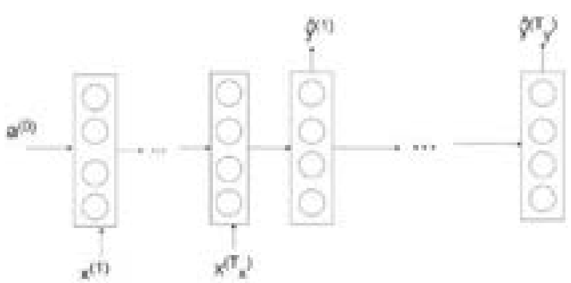


Fig. 2: Recurrent Neural Network

Usage: For machine translation first the entire input text is encoded by an encoder RNN which ties to produce a vector which tries to encapsulate the various features and relationships between the words of the sentence. This vector is than fed into the decoder RNN which attempts to translate the sentence. In the case of sentiment analysis the sentence is passed through a RNN at the end of which a softmax function is applied to categorise the sentence into the correct sentiment category.

Vanishing Gradient problem

The vanishing gradient problem occurs when deep neural networks[16] with a large number of layers are trained during back propagation with gradient-based learning methods. Dur- ing each iteration of training the neural networks' weights are updated by calculating the partial derivative of the error function with respect to the current weight. But with a large numbers of layers for the gradient to back propagate through the gradient continues to become smaller and smaller. This can cause the initial layers of the network to almost completely not undergo further training. Traditional activation functions such as tanh(hyperbolic tangent)[17] have derivatives in the range of (0,1] when these small values are repeatedly multiplied the gradient becomes exceedingly small and vanishes.

Long Short-term Memory (LSTM)

In order to solve the vanishing gradient problem as de- scribed above the LSTM (Long Short-term Memory) was developed. LSTMs don't just decide to keep all the previous state values but chose to discard or retain some parts of the previous state. It does this by introducing gates which decide how much information from the previous state needs to be retained.

Key Components of an LSTM

1. **Cell State (Cel_t):** The cell state represents the long-term memory of the LSTM. It is used throughout the layers of the entire neural network with only minor linear activations.

2. **Forget Gate (d_t):** The forget gate is used to decide as to how much information from the cell state shall be discarded or kept on. It takes the previous hidden state h_{t-1} along with the current input $a0_t$. The sigmoid function is used to ensure that the output f_t is between 0 and 1.

$$d_t = \sigma(P_d \cdot [hid_{t-1}, a0_t] + c_f)$$

where W_d is the forget gate's weight matrix,
σ is the sigmoid activation function,
and c_f is the bias term.

3. **Input Gate (i_t):** The input gate decides the information that needs to be updated in the cell state. It has two parts: a sigmoid layer to decide the values to be updated (i_t) and a tanh layer which creates new candidates ($\tilde{C}_t$).

$$i_t = \sigma(P_i \cdot [hid_{t-1}, x_t] + C_i)$$

$$\tilde{C}el_t = \tanh(P_C \cdot [h_{t-1}, x_t] + C_C)$$

where P_i, P_C are the weight matrices,
σ is the sigmoid function,
and C_i, C_C are the bias terms.

4. **Update Cell State ($\tilde{C}el_t$):** Represents new candidate values that can be added to the cell state.

5. **Update Cell State (Cel_t):** The cell state is changed by combining the old cell state (Cel_{t-1}) with the information from the forget gate and the new candidate values.

$$Cel_t = d_t \odot Cel_{t-1} + i_t \odot \tilde{C}el_t$$

where $\odot$ denotes element-wise multiplication.

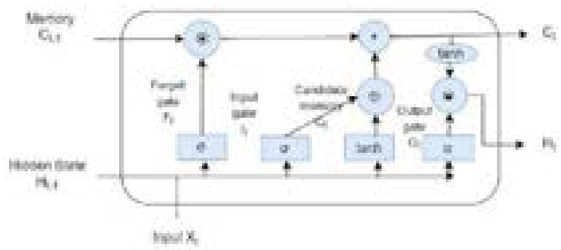


Fig. 3: Long Short-term Memory

6. **Output Gate (out_t):** The output gate decides the next hidden state (hid_t) based on the newly updated cell state. The previous hidden state hid_{t-1}, the current input x_t, and the updated cell state Cel_t is inputed into it.

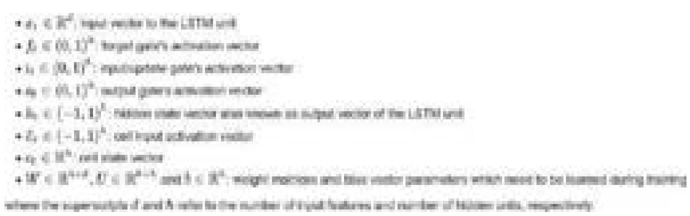
$$out_t = \sigma(P_o \cdot [hid_{t-1}, x_t] + c)$$

where P_o is the weight matrix,
σ is the sigmoid function,
and c is the bias term.

7. **Hidden State (hid_t):** The final hidden state is computed by applying the output gate to the updated cell state.

$$hid_t = o_t \odot \tanh(C_t)$$

The operator $\odot$ represents the Hadamard product.


- $x_t \in \mathbb{R}^d$: input vector to the LSTM unit
- $f_t \in (0,1)^h$: forget gate's activation vector
- $i_t \in (0,1)^h$: input/update gate's activation vector
- $o_t \in (0,1)^h$: output gate's activation vector
- $h_t \in (-1,1)^h$: hidden state vector also known as output vector of the LSTM unit
- $\tilde{c}_t \in (-1,1)^h$: cell input activation vector
- $c_t \in \mathbb{R}^h$: cell state vector
- $W \in \mathbb{R}^{h \times d}$, $U \in \mathbb{R}^{h \times h}$ and $b \in \mathbb{R}^h$: weight matrices and bias vector parameters which need to be learned during training

where the superscripts d and h refer to the number of input features and number of hidden units, respectively.

Usage: For machine translation first the entire input text is encoded by an encoder LSTM which ties to produce a vector which tries to encapsulate the various features and relationships between the words of the sentence. This vector is than fed into the decoder LSTM which attempts to translate the sentence. In the case of sentiment analysis the sentence is passed through a LSTM at the end of which a softmax function is applied to categorise the sentence into the correct sentiment category.

Gated recurrent unit(GRU)

Gated recurrent units (GRUs) are an extension of recurrent neural networks using gating mechanisms, introduced in 2014 by Kyunghyun Cho et al. It was introduced as an improvement over traditional RNNs to address some of their limitations, especially in handling long-term dependencies in sequential data.

Key Components of a GRU

1. **Update Gate (u_t):** Determines how much of the past information should be retained. It considers the current input and information from the previous time step. The update gate outputs values between 0 and 1.

$$u_t = \sigma(P_z \cdot [hid_{t-1}, x_t])$$

where P_z is the update gate's weight matrix,
σ is the sigmoid function,
hid_{t-1} is the previous step's hidden state,
and x_t is the input at the current time step.

2. **Reset Gate (res_t):** Determines how much of the information needs to be discarded. It helps the model focus on the most relevant information from the past. Like the update gate, the reset gate outputs values between 0 and 1.

$$res_t = \sigma(P_r \cdot [hid_{t-1}, x_t])$$

where W_r is the reset gate's weight matrix.

3. **Current Memory Content ($h\tilde{i}d_t$):** Represents information considered relevant for the current time step. It is a candidate update for the hidden state and is computed using the input and the reset gate.

$$h\tilde{\ }id_t = \tanh(P_h \cdot [res_t \odot hid_{t-1}, x0_t])$$

where P_h is the current memory content's weight matrix,
tanh is the hyperbolic tangent function,
and $\odot$ denotes element-wise multiplication.

4. **Hidden State Update (hid_t):** The new hidden state is formed by combining the previous hidden state with the current memory content and is controlled by the update gate.

$$hid_t = (1 - u_t) \odot hid_{t-1} + u_t \odot h\tilde{\ }id_t$$

where $a0_t$ is the input vector
hid_t is the output vector
u_t is the update gate vector
res_t is the reset gate vector

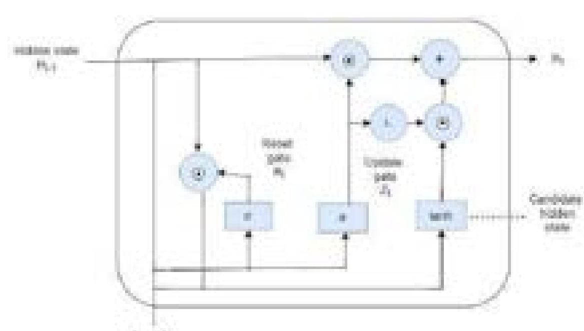


Fig. 4: Gated Recurrent Units

Usage: For machine translation first the entire input text is encoded by an encoder GRU which ties to produce a vector which tries to encapsulate the various features and relationships between the words of the sentence. This vector is than fed into the decoder GRU which attempts to translate the sentence. In the case of sentiment analysis the sentence is passed through a GRU at the end of which a softmax function is applied to categorise the sentence into the correct sentiment category.

D. Architecture

This project's architecture involves preprocessing datasets for sentiment analysis and machine translation. Neural network models, including RNN, LSTM, GRU layers, are configured with task-specific output layers. Training involves optimizer selection,loss function definition, and task-specific metrics.

Type of Layer	Shape of output	No of Param
Input Layer	[(None, 30)]	0
Embedding	(None, 30, 500)	7,266,000
Encoder RNN Layer	(None, 150)	97,650
RepeatVector	(None, 30, 150)	0
Decoder RNN Layer	(None, 30, 150)	45,150
TimeDistributed	(None, 30, 150)	22,650
Dropout	(None, 30, 150)	0
TimeDistributed	(None, 30, 30661)	4,629,811

TABLE I: Model Summary for the Machine translation RNN model

Type of Layer	Shape of output	No of Param
Input Layer	[(None, 30)]	0
Embedding	(None, 30, 500)	7,266,000
Encoder GRU	(None, 150)	293,400
RepeatVector	(None, 30, 150)	0
Decoder GRU	(None, 30, 150)	135,900
TimeDistributed	(None, 30, 150)	22,650
Dropout	(None, 30, 150)	0
TimeDistributed	(None, 30, 30661)	4,629,811

TABLE II: Model Summary for Machine translation GRU model

Type of Layer	Shape of output	No of Param
Input Layer	[(None, 30)]	0
Embedding	(None, 30, 500)	7,266,000
Encoder LSTM	(None, 150)	390,600
RepeatVector	(None, 30, 150)	0
Decoder LSTM	(None, 30, 150)	180,600
TimeDistributed	(None, 30, 150)	22,650
Dropout	(None, 30, 150)	0
TimeDistributed	(None, 30, 30661)	4,629,811

TABLE III: Model Summary for the Machine translation LSTM model

Type of Layer	Shape of output	No of Param
Input Layer	(None, 50)	0
Embedding	(None, 50, 500)	17,231,000
LSTM	(None, 150)	390,600
Dense	(None, 150)	22,650
Dropout	(None, 150)	0
Dense	(None, 4)	604

TABLE IV: Sentiment analysis LSTM model

Type of Layer	Shape of output	No of Param
Input Layer	(None, 50)	0
Embedding	(None, 50, 500)	17,231,000
GRU	(None, 150)	293,400
Dense	(None, 150)	22,650
Dropout	(None, 150)	0
Dense	(None, 4)	604

TABLE V: Sentiment analysis GRU model

Type of Layer	Shape of output	No of Param
InputLayer	(None, 50)	0
Embedding	(None, 50, 500)	17,231,000
RNN Layer	(None, 150)	97,650
Dense	(None, 150)	22,650
Dropout	(None, 150)	0
Dense	(None, 4)	604

TABLE VI: Sentiment analysis RNN model

E. Evaluation

In the evaluation phase, the performance of RNNs, LSTMs, and GRUs was scrutinized in both sentiment analysis and machine translation tasks. Cross-analyzing these tasks allowed us to identify trade-offs and make recommendations based on observed performance differences. Insights into model behavior highlighted both common challenges and unique strengths across RNNs, LSTMs, and GRUs.

RNNs were created to handle sequential data. It does this by feeding the previous state's hidden state into the next state as well. The major problem with the model is that information from previous states becomes harder and harder to retain down the line. Inputs processed at the beginning are mostly forgotten after a few new inputs. Another problem associated with RNNs is the vanishing gradient problem where due to the small values of the gradient, changes made to the weights of the nodes of the initial layers of the neural network are extremely miniscule making these layers extremely hard to train.

LSTMs, which have multiple gates to decide what information to discard and retain, fare much better in capturing information from the earlier inputs. LSTMs have a total of three gates(forget gate, update gate and output gate) to collectively decide which and how much information is to be retained in the matrices. A drawback of having these many gates is that the amount of computation required for their training was also substantially increased. GRUs achieve a middle ground by using only one gate to decide what and how much information is to be retained.

LSTMs' proficiency in handling long-range dependencies and GRUs' computational efficiency, trade-offs surfaced in terms of accuracy, performance and training complexity. These findings underscore the importance of considering model characteristics in alignment with specific task requirements. We concluded with suggestions for future research directions, pinpointing areas for improvement and optimization in the application of these models to sentiment analysis and machine translation. This comprehensive evaluation aims to guide researchers and practitioners in the dynamic field of natural language processing.

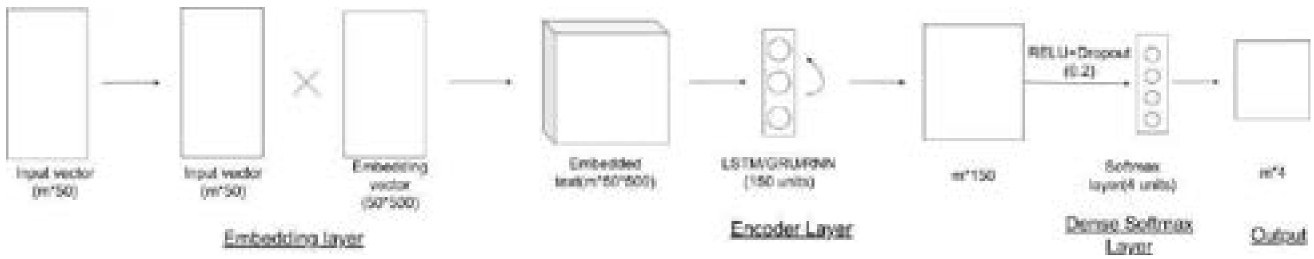


Fig. 5: Sentiment Analysis

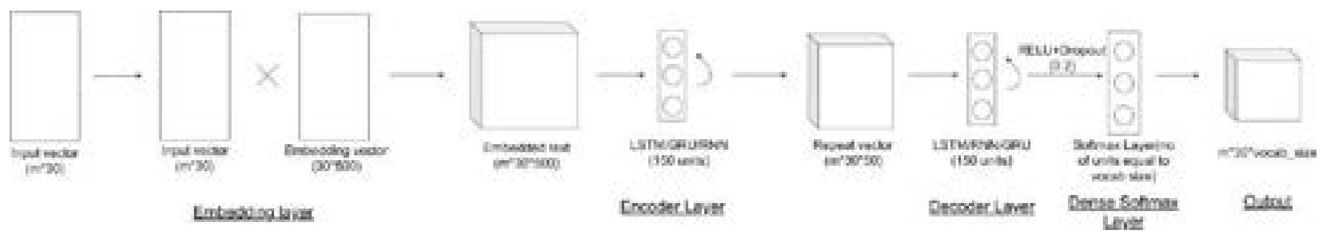


Fig. 6: Machine Translation

TABLE VII: LSTM Sentiment Analysis Evaluation Results

Epochs	Train Loss	Train Accuracy	Test Loss	Test Accuracy
1	0.7274	0.7173	0.4342	0.8395
2	0.2585	0.9036	0.3213	0.8844
3	0.1540	0.9392	0.3276	0.8891
4	0.1148	0.9531	0.3573	0.8917
5	0.0968	0.9587	0.3892	0.8966

TABLE VIII: RNN Sentiment Analysis Evaluation Results

Epochs	Train Loss	Train Accuracy	Test Loss	Test Accuracy
1	0.8020	0.6805	0.4985	0.8124
2	0.3185	0.8827	0.4670	0.8327
3	0.2197	0.9171	0.4654	0.8463
4	0.1772	0.9311	0.4689	0.8496
5	0.1592	0.9374	0.4786	0.8527

TABLE IX: GRU Sentiment Analysis Evaluation Results

Epochs	Train Loss	Train Accuracy	Test Loss	Test Accuracy
1	0.7305	0.7153	0.4094	0.8473
2	0.2492	0.9075	0.3296	0.8780
3	0.1533	0.9396	0.3478	0.8887
4	0.1212	0.9507	0.3585	0.8955
5	0.0998	0.9580	0.4125	0.8926

TABLE X: LSTM Machine Translation Evaluation Results

Epochs	Train Loss	Train Accuracy	Test Loss	Test Accuracy
1	1.5057	0.7959	1.3149	0.8108
2	1.2418	0.8164	1.1637	0.8231
3	1.1050	0.8267	1.0585	0.8314
4	1.0000	0.8337	0.9805	0.8374
5	0.9237	0.8389	0.9269	0.8425

TABLE XI: GRU Machine Translation Evaluation Results

Epochs	Train Loss	Train Accuracy	Test Loss	Test Accuracy
1	1.5062	0.7957	1.3139	0.8109
2	1.2394	0.8172	1.1519	0.8249
3	1.1008	0.8270	1.0560	0.8315
4	1.0050	0.8332	0.9845	0.8377
5	0.9365	0.8377	0.9511	0.8398

TABLE XII: RNN Machine Translation Evaluation Results

Epochs	Train Loss	Train Accuracy	Test Loss	Test Accuracy
1	1.5128	0.7947	1.3332	0.8080
2	1.2837	0.8122	1.2291	0.8166
3	1.1903	0.8181	1.1680	0.8207
4	1.1289	0.8215	1.1332	0.8228
5	1.0858	0.8235	1.1032	0.8246

IV. Result and Analysis

Machine translation
LSTM accuracy: on training=0.8389, on validation=0.8425
GRU accuracy: on training=0.8377, on validation=0.8398
RNN accuracy: on training=0.8235, on validation=0.8246
Sentiment analysis
LSTM accuracy:on training=0.9587, on validation=0.8966
RNN accuracy:on training=0.9374, on validation=0.8527
GRU accuracy:on training=0.9580, on validation=0.8926

As can be seen LSTMs and GRUs give a much higher accuracy on both tasks as compared to RNNs which suffers from the vanishing gradient problem and is not able to capture the features of words occurirng earlier in a sequence.

The difference in accuracy between GRUs and LSTMs is much less noticeable with the LSTM models just barely edging out the GRU model in terms of accuracy on both tasks. This supports research conducted during the literature survey which show similar results. GRUs and LSTMS usually provide

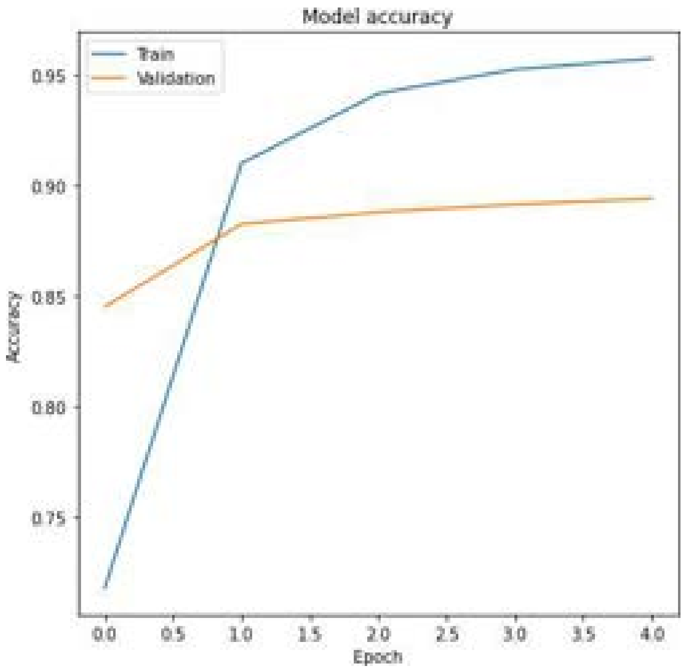


Fig. 7: Sentiment analysis GRU model

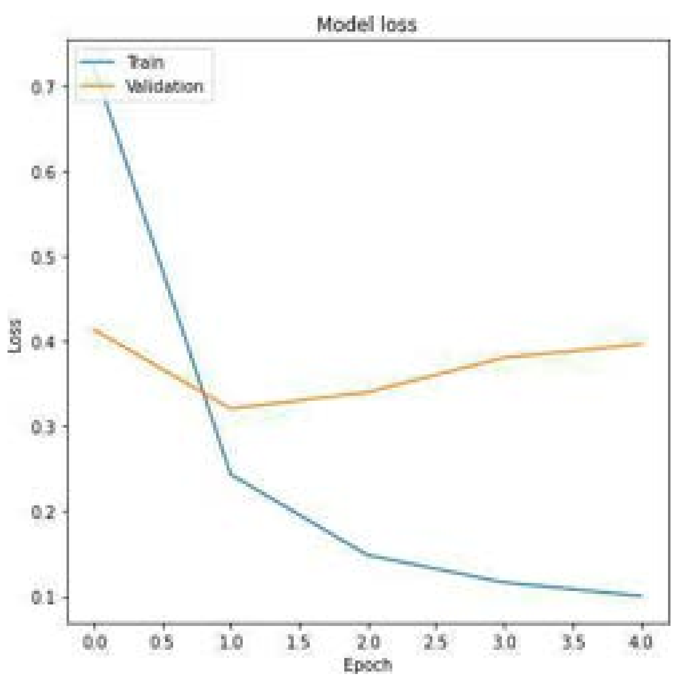


Fig. 8: Sentiment analysis GRU model

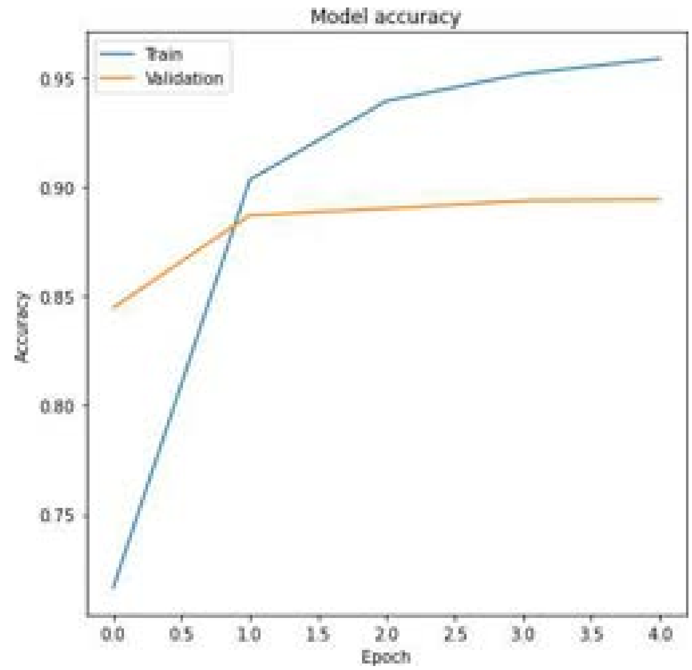


Fig. 9: Sentiment analysis LSTM model

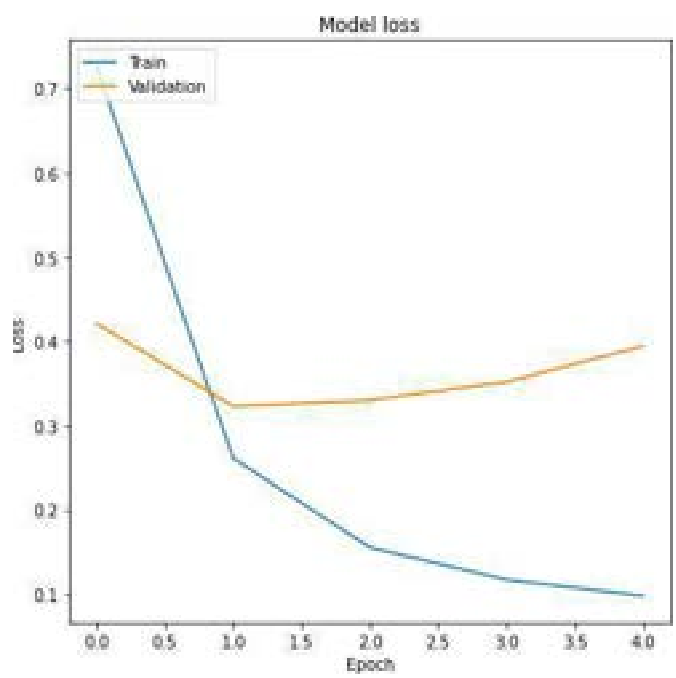


Fig. 10: Sentiment analysis LSTM model

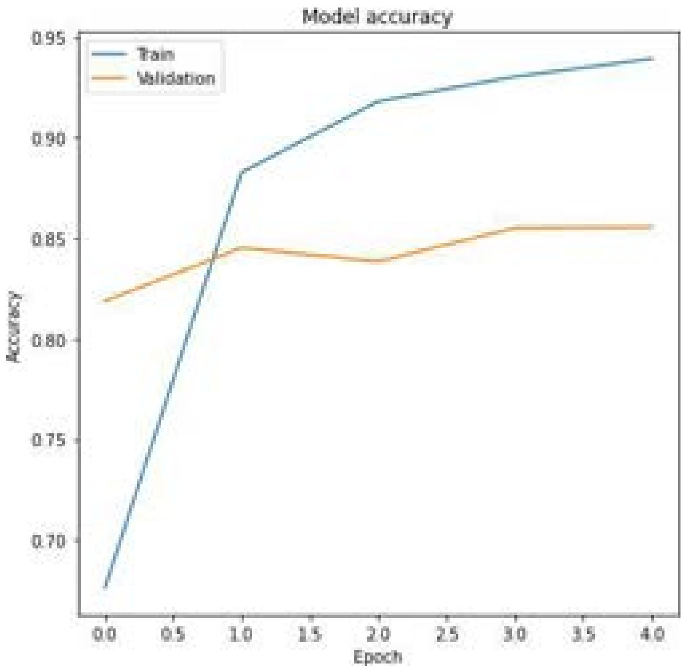


Fig. 11: Sentiment analysis RNN model

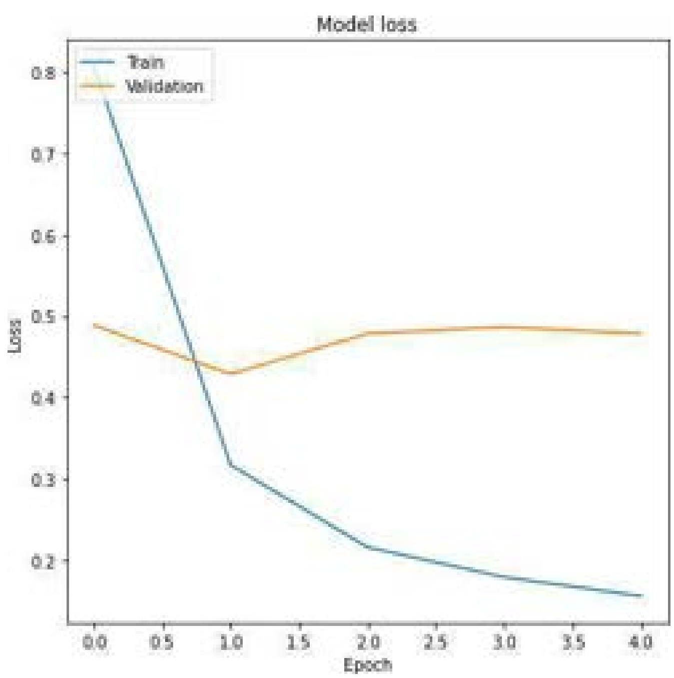


Fig. 12: Sentiment analysis RNN model

similar results in terms of accuracy with LSTMs showing negligible but observable increases in accuracy even though the LSTM model is substantially computationally harder to train.

Further accuracy and performance may be improved by increasing the number of layers and the number of nodes in each layer, choosing an appropriate learning rate and increasing the number of iterations. This method too has its limits. Improvements in and accuracy by adding more nodes and layers work only to a certain extent after which the gains stagnate and in a wide variety of cases the accuracy of the model may even decrease. Problems of overfitting may also be encountered which may be solved by regularization methods such as early stopping, weight decay, dropout etc. Adjusting and fine tuning these hyper parameters is a matter of repetitive trial and error, although previous literature into neural networks may prove insightful in setting these values. Similarly increasing the size of the word embedding or using pre trained word embeddings may improve performance. The size of these word embeddings too has a upper limit above which no further gains in accuracy are made. Usually, the size of these word embeddings is recommended to be set to around 300-500. Other methods to increase accuracy and performance are to train the model on a larger dataset with a diverse range of values.

V. CONCLUSION

In conclusion, delving into the comparison of various approaches for language understanding and translation provides valuable insights into their respective strengths and characteristics. The analogy of selecting the right tool for a task aptly captures the essence of these methods.

In a thorough comparison of models for machine translation and sentiment analysis, a meticulous approach has been devised. The first step involves selecting diverse datasets covering two languages for translation and various emotional content for sentiment analysis. Standardized preprocessing, including tasks like tokenization and cleaning, ensures consistency across datasets. LSTMs and GRUs outperform RNNs in accuracy for both tasks due to RNNs struggling with the vanishing gradient problem and the inability to capture earlier sequence features. While GRUs and LSTMs display comparable accuracy, LSTMs slightly outperform GRUs on both tasks. Despite LSTMs being computationally more demanding, they exhibit marginal accuracy gains over GRUs.

In essence, the comparative analysis provides a nuanced compass for decision-makers, researchers, and practitioners navigating the diverse landscape of language-related tasks. Looking ahead, the future scope of this research extends to several dimensions. The exploration of hybrid models and advanced techniques could further enhance the capabilities of sequence modeling architectures. Incorporating attention mechanisms, transfer learning, or ensemble methods may offer avenues for overcoming specific challenges identified in the comparative analysis.

REFERENCES

[1] A. Patel and A. K. Tiwari, "Sentiment analysis by using recurrent neural network," in *Proceedings of 2nd International Conference on Advanced Computing and Software Engineering (ICACSE)*, 2019.

[2] A. C. M. V. Srinivas, C. Satyanarayana, C. Divakar, and K. P. Sirisha, "Sentiment analysis using neural network and lstm," in *IOP conference series: materials science and engineering*, vol. 1074, no. 1. IOP Publishing, 2021, p. 012007.

[3] G. Murthy, S. R. Allu, B. Andhavarapu, M. Bagadi, and M. Belusonti, "Text based sentiment analysis using lstm," *Int. J. Eng. Res. Tech. Res*, vol. 9, no. 05, 2020.

[4] P. M. Sosa, "Twitter sentiment analysis using combined lstm-cnn models," *Eprint Arxiv*, vol. 2017, pp. 1–9, 2017.

[5] Z. Tan, S. Wang, Z. Yang, G. Chen, X. Huang, M. Sun, and Y. Liu, "Neural machine translation: A review of methods, resources, and tools," *AI Open*, vol. 1, pp. 5–21, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666651020300024

[6] D. Datta, P. E. David, D. Mittal, and A. Jain, "Neural machine translation using recurrent neural network," *International Journal of Engineering and Advanced Technology*, vol. 9, no. 4, pp. 1395–1400, 2020.

[7] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[8] A. Sherstinsky, "Fundamentals of recurrent neural network (rnn) and long short-term memory (lstm) network," *Physica D: Nonlinear Phenomena*, vol. 404, p. 132306, 2020.

[9] R. Dey and F. M. Salem, "Gate-variants of gated recurrent unit (gru) neural networks," in *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)*. IEEE, 2017, pp. 1597–1600.

[10] S. Li, W. Li, C. Cook, C. Zhu, and Y. Gao, "Independently recurrent neural network (indrnn): Building a longer and deeper rnn," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5457–5466.

[11] Y. Kim, P. Petrov, P. Petrushkov, S. Khadivi, and H. Ney, "Pivot-based transfer learning for neural machine translation between non-English languages," in *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, K. Inui, J. Jiang, V. Ng, and X. Wan, Eds. Hong Kong, China: Association for Computational Linguistics, Nov. 2019, pp. 866–876. [Online]. Available: https://aclanthology.org/D19-1080

[12] A. Bau, Y. Belinkov, H. Sajjad, N. Durrani, F. Dalvi, and J. Glass, "Identifying and controlling important neurons in neural machine translation," *arXiv preprint arXiv:1811.01157*, 2018.

[13] "Language translation english-french," https://www.kaggle.com/datasets/devicharith/language-translation-englishfrench, accessed: 2024.

[14] "Twitter entity sentiment analysis," https://www.kaggle.com/datasets/jp797498e/twitter-entity-sentiment-analysis, accessed: 2024.

[15] Amit Gupta, Richa Gupta & Navin Garg (2021) An efficient approach for classifying chest X-ray images using different embedder with different activation functions in CNN, Journal of Interdisciplinary Mathematics, 24:2, 285-297, DOI: 10.1080/09720502.2020.1838060

[16] Gupta, R., Tripathi, V., & Gupta, A. (2021, December). An Efficient Model for Detection and Classification of Internal Eye Diseases using Deep Learning. In 2021 International Conference on Computational Performance Evaluation (ComPE) (pp. 045-053). IEEE

[17] Gupta, R., Tripathi, V., Gupta, A.(2021).Diabetic Retinopathy detection using an Efficient Artificial Intelligence Method. ICAES2021 International conference on advancement in Engineering and Sciences July 2021

Deep Learning for Wildfire Impact: Vegetation Recovery and Soil Chemistry Analysis

Dr.V.AnjanaDevi
Computer Science and Engineering
Rajalakshmi Institute of Technology
Chennai, Tamil Nadu, India
anjanadevi.aby06@gmail.com

Praveen P
Computer Science and Engineering
Rajalakshmi Institute of Technology
Chennai, Tamil Nadu, India
praveenprabha2003@gmail.com

Ramanathan R
Computer Science and Engineering
Rajalakshmi Institute of Technology
Chennai, Tamil Nadu, India
ramanathanr7777@gmail.com

***Abstract*—This study presents a deep learning model created for enabling comprehensive wildfire control by seamlessly combining satellite images, weather data and terrain details. Current systems face challenges in comprehensively analyzing these factors due to limitations in data integration, dynamic fire behavior prediction, and post-fire ecological impact evaluation. By improving detection and accurate assessment of impact, the system addresses all aspects of wildfire management from forecasting to post event analysis. The model integrates soil quality examination and vegetation regrowth simulation Using image analysis and state of the art deep learning methods. This holistic approach of Image analysis employs Convolutional Neural Networks (CNN) for predicting wildfire risk and Recurrent Neural Networks (RNN) for assessing soil and hydrological effects. This adaptable approach, which aims to transform the way fire control is done, can be readily adjusted to changing conditions and takes correlations between different aspects into account. It surpasses conventional techniques by including soil quality analysis, vegetation regrowth modeling, and vegetation damage evaluation. The adaptable nature of this method proves invaluable, in lessening the impact of wildfires with a focus, on evaluating vegetation damage and promoting restoration.**

***Index Terms*—Wildfire Management, Deep Learning Model, Satellite Imagery, Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), Environmental Impact Assessment.**

I. INTRODUCTION

Wildfires are difficult to manage and require an understanding of their spatiotemporal dynamics. The objective is to analyze the works of Meneses [1], Hao et al. [2], considering Deep Learning for model development. The study integrates satellite imagery, weather data, and topographical information in order to address issues associated with predicting wildfires' spread accurately as well as evaluating the recovery of vegetation in accordance with objectives set forth by Akbari Asanjan et al. [3] and Chen et al. [4]. Furthermore, the authors also emphasize on the analysis of recovery patterns Rashkovetsky et al.'s [5] which further highlights the importance of the study's objectives. In other words, the aim is to provide decision-makers with early warning systems so that they will take measures in advance thus reducing negative consequences caused by wildfire outbreaks on either human or natural ecosystems.

The issue with wild fires is that they change quickly over both time and space hence necessitating in-depth knowledge regarding these movements for proper fire management strategies that can be adopted. The objective of the research was influenced by Meneses (2009) work, quoted from Hao et al. (2018), using deep learning method to develop a unified framework for modeling wildfire occurrences. By incorporating satellite imagery, weather data, and topography into their analysis efforts, the researchers intend to provide a means for making more precise predictions about future fires as well determining how plant life forms recover themselves after such natural catastrophes like those described in Akbari Asanjan et al.'s paper [3] or Chen et al.'s article [4]. Moreover, Rashkovetsky et al.'s article has emphasized recovery patterns segmentation as one among many significance areas under investigation within this study. Thus, it aims at facilitating prompt intervention through enabling decision makers foresee possible scenarios that may arise out of irregularities related to forest fires at their inception stage with the intention of minimizing destruction on ecosystems as well as communities. The difficulties associated with modelling fire spread are those of dynamic prediction – for instance, it's trying to take into account the potential for wind to change directions while incorporating terrain features and fuel availability. The DL Model is attempting to bring together these diverse inputs to provide decision-makers with knowledge about land features and fire spread in earlier time frames to inform their early interventions and coordinate resources needed for fire-fighting. The effort by data-centric approaches is to better maintain the balance between allocating resources to limit the impact of post-fire conditions on ecosystems and communities.

II. LITERATURE SURVEY

Ecological effects of wildfires are being researched more in detail in the last few years; the latter are coupled with different approaches and techniques. One of his output contains Meneses [1] whose data set is based on Landsat 8 OLI imaging and environmental biophysical data for vegetation recovery metrics in burned areas. Hao et al. [2] discuss the interannual effects of fire severity, weather as decided by climate, and post-forest-fire vegetation recovery. For example, Akbari Asanjan et al. [3]

formulate a probabilistic wildfire segmentation algorithm using supervised deep generative approaches drawn from satellite imagery and Chen et al. [4] implement deep learning strategies for wildfire segmentation of high resolution satellite imagery.

Meneses[1] and Hao et al.[2]have made a contribution to the understanding of the post- wildfire vegetation dynamics through the analysis of time series data. Additionally, they introduced the influence of fire severity and climate variability on the restoration of ecosystems on an interannual basis. Olga undertakes Landsat-8 OLI image analysis and biophysical data using for vegetation metrics assessment, while Hao and colleagues provide an insight into long-7-term ecological shifts. While those two studies may be able to keep track of complicated ecologic processes, they may not be accurate and may have high data variability. Moreover, Asanjan et al [3] and Chen et al. [4] model wildfire diagnosis in a newest manner which observes the deep learning and probabilistic segmentation algorithms. Even if the approach seems rather efficient for a monitoring process, acquiring big labeled sets and uncertainty boundaries when conditions differ must be properly addressed.

There is detailed information Kim et al. [6] on how to do real time monitoring of wildfires through integration of Unmanned Aerial Vehicles (UAVs) and satellite data which shows that these sources of information are complementing each other. Thangavel et al. [7] present a case study on autonomous satellite wildfire detection using hyperspectral imagery and neural networks, exemplifying practical applications of advanced technologies. Rashkovetsky et al.[5] delve into deep semantic segmentation for wildfire detection from multi-sensor satellite imagery, showcasing the versatility of deep learning in various data environments. Toan et al.[8] use an approach based on deep learning to detect early wildfires from hyperspectral satellite images thus underscoring the need for early intervention efforts.

For urban spaces, Khan and Khan [9] integrate a deep learning-based forest fire map and detection method called 'FFireNet' for smart cities; Sathishkumar et al. [10] apply deep learning for the detection of forest fires and smoke, demonstrating the versatility of these techniques to different ecological contexts.

Going beyond landscapes that are already on fire, other studies such as Bright et al [12] make use of Landsat time series analysis to investigate wildfire recovery of vegetation on 19 forest types across North America. Ambadan et al [13] used satellite-observed soil moisture as a proxy for wildfire risk — illustrating the utility of using environmental variables as predictors in predictive models.

Tree growth and stem carbon accumulation after drought and fire in Amazonian forests are evaluated following the ecosystem impact assessment approach used by Berenguer et al. [14]. Burn severity and undisturbed forest proximity are considered as key factors for tropical montane forest recovery [15], while wildfire recovery is monitored within the ecosystem impact assessment framework in Mongolia [16]. Li et al. [17] evaluated the impacts of wildfire on soil properties in a pine forest in South Korea using the ecosystem impact assessment approach.

An example of state-of-the-art approaches to estimating recovery of soil and vegetation are Zhongqiu Smith et al. [18] using deep learning for recovery assessment; Yuling Li et al. [19] using multi-source data fusion; Shangmao Wang et al. [20] using recurrent neural networks for temporal analysis of vegetation recovery; Dae-Hyeon Park et al. [21] examining spatial and temporal patterns after forest fires; Jeungwon Kim et al. [22] exploring the hydrological response.

Multi-sensor data fusion for nonwoven fire spread, and mapping burn severity and soil properties using integrated remote sensing data. The research by Yang et al[25] illustrates how deep learning can be used to predict post-fire vegetation regrowth patterns. These references serve to highlight both the current state of the art of monitoring wildfires, mapping vegetation recovery, and other associated ecological impacts that can be studied after a wildfire event.However, like many studies in this field, challenges such as computational complexity and uncertainties in remote sensing data.

III. Proposed System

The study investigates on an elaborate system of post-fire analysis dealing with challenges faced after the occurrence of wildfires. Soil and Hydrological Assessment is integrated with Vegetation Regrowth Assessment.

For the Soil and Hydrological Assessment, hyper-spectral images from 2002 Hayman fire are used to process through Spectral Unmixing and SemantiQNet deep learning architecture that result in a refined mask while Recurrent Neural Network (RNN) analysis improves estimates of important soil parameters leading to better understanding of post-fire ecology.

In Vegetation Regrowth Assessment, various satellite images are employed in computing indices such as NDVI and EVI which evaluate regrowth as well as overall vegetation health. These indices help to create masks highlighting regions with significant greenness thus acting as indicators for successful post-fire vegetation regrowth. The project makes a big contribution towards understanding post-fire ecosystem dynamics.

To sum up, integrating Soil and Hydrological Assessment into Vegetation Regrowth Assessment components which entails application advanced methods like spectral unmixing and deep learning architectures like SemantiQNet, Recurrent Neural Networks (RNN), etc., offers sophisticated approach to post-fire analyses [1][3][9] .In valuable insights into post-fire complexities, this system gives exact estimates of key soil parameters for vegetation growth while using a variety of satellite images and indices such as NDVI and EVI. To understand and manage post-fire ecological landscapes after wildfires, it is important to employ innovative methods as described in the articles [1][4] [14]. This will help post-fire ecological landscape recovery efforts be more sustainable reducing future risk from fire events.

A. System Architecture

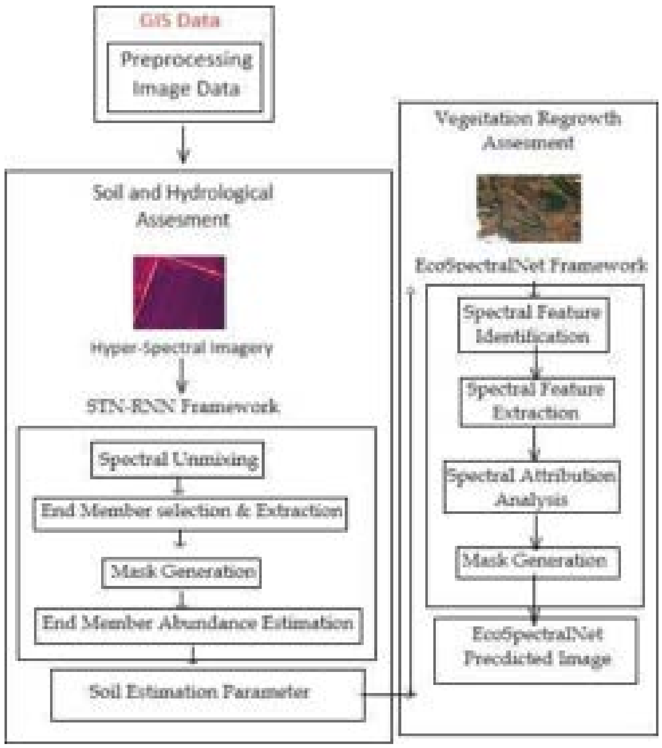


Fig. 1. System Architecture

Though, satellite photos are an integral part of GIS applications and they must be pre-processed with but it is essential. The purpose of data preprocessing is to be the bridge between a dataset's unprocessed raw form and its shaped data that can be manipulated. By performing vital processes like thermal and atmospheric correction to eliminate atmospheric interferences, geometric correction for distortions and radiometric correction which will ensure corrected colour representation, we can be able to achieve more accurate information. Subsequently, spectral identification and recovery of some unique spectral bands for each land cover category are carried on. The research gap is indeed important and deeper exploration can be made with more coherent methods like spectral abundance estimate and endmember extraction. While at first the image may serve as a distraction before gradually providing useful information, those pixels may later be rejected by a mask to eliminate the unwanted parts while the pre-processed data might be of high use for modelling or estimating soil parameters.

The research paper presents a well thought-out architecture designed to capture the intricate dynamics of post-fire landscape assessment. It affords the development of models that blend different data types with computation intensive techniques like deep learning to enhance understanding of wildfire aftermath. The use of Sentinel-3 satellite data plays an integral role for obtaining Land Surface Temperature. (LST) images that are crucial in quantifying fire severity based on temperature fluctuation during the fire outbreak itself. A novel deep learning framework, Optinet, fine-tuned using the Optuna Hyperparameter optimization algorithm is at the core of fire-damaged areas' classification and mapping; which in turn yields thematic maps that illustrate in detail how the distribution and intensity of fire-induced damages. Furthermore, hyper-spectral imagery from the 2002 Hayman Fire in Colorado plays a significant role in this architecture. To eliminate extraneous spectral data, as well as to create a mask, Spectral Unmixing guided by SemantiQNet deep learning architecture is used. These are later fed into Recurrent Neural Network (RNN) for dependency modeling that gives important estimates of soil parameters that are essential to post-fire vegetation growth understanding.

NORMALIZED DIFFERENCE VEGETATION INDEX (NDVI) AND GREEN NORMALIZED DIFFERENCE VEGETATION INDEX (GNDVI)

The formulas for NDVI and GNDVI are given by:

$$\text{NDVI} = \frac{\text{Band 5 (NIR) - Band 4 (RED)}}{\text{Band 5 (NIR) + Band 4 (RED)}} \quad (1)$$

$$\text{GNDVI} = \frac{\text{Band 5 (NIR) - Band 3 (GREEN)}}{\text{Band 5 (NIR) + Band 3 (GREEN)}} \quad (2)$$

B. Soil and Hydrological Assessment

For effective ecosystem restoration, the study presents an innovative approach to post-fire landscape assessment. After the Hayman fire in Colorado [1], hyper-spectral imagery from drones was utilized as a basis for soil analysis and hydrology with its integration of advanced techniques.

The methodology commences with Spectral Unmixing technique which is a remote sensing method for classifying mixed pixels based on their spectral signatures [2]. SemantiQNet deep learning architecture enhances spectral analysis enabling precise material identification leading to filtered masks [25].

The inclusion of RNNs puts in the temporal dimension that captures dynamic changes over time[3]. recurrent neural networks (RNNs) take advantage of the sequential nature of hyperspectral data, which is widely employed in soil spectroscopy, to start. The data frequently shows dependencies, in which the spectral signatures of one data point depend on those of earlier ones. Conventional methods are not very good at capturing this feature. RNNs are particularly good at learning both short- and long-term dependencies, which is essential for effectively modeling hyperspectral signals. Second, by combining with dimensionality reduction methods like Principal Component Analysis (PCA) and Locality Preserving Projections (LPP), RNNs help in feature extraction from high-dimensional data. The procedure provides a good result when compared to other methods by reducing computational complexity but not much in information at the spectral level. The implementation of RNN in this case also involves accounting for both the spatial and temporal fluctuations in soil parameters that may vary within fields since RNN capture time series well. Given their ability to simulate complicated non-linear processes, soil features describing environmental conditions (such as temperature, humidity, depth, and salinity) can be predicted with high accuracy. Lastly, the recent developments of RNNs have contributed to the optimization of the algorithm, neural design and regularization approaches. So, the remarkable advancement in the deep learning becomes the power of the models which can be used for the assessment of the soil attributes from the hyperspectral data. RNNs in particular, are one of methods based on which improve of the soil parameters estimates is achieved, as they can calculate high-dimension data, detect serial dependencies, and generate complex interactions between spectral signatures and soil properties.

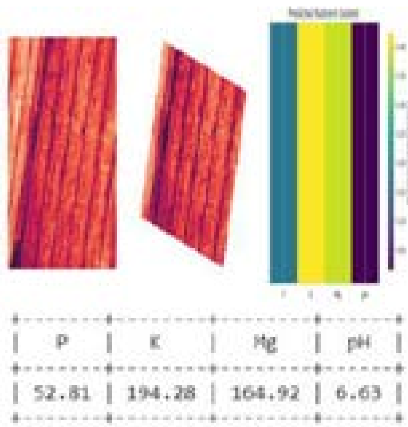


Fig. 2. Soil and Hydrological Assessment

C. Vegetation Regrowth Assessment

The Vegetation Regrowth Assessment phase is an essential part of the research project, with a specific focus on a systematic evaluation of vegetation recovery after wild fire events.

Ecospectral Net, a neural network especially designed for the manipulation of the dataset that covers ecological and spectral aspects, is a tool that might play a canonical role as a supplement to the existing techniques deployed in the Vegetation Regrowth Assessment step. Through CNNs that play a critical role, it tries to complete the preprocessing of data, features extraction, and generation of output as an automated process. Therefore, the system simplifies processes of the past where manual feature engineering was required. The sophisticated CNN structure of its multi-layered system can handle the complexities of the ecological data by acquiring the high accuracy of land analytics, such as vegetation analysis, soil properties mapping, habitat observation, and land cover change detection. Ecospectral Net acts as an additional instrument on top of automation and accuracy in understanding complicated environmental problems by being a dependable tool for researchers and conservationists in environmental science and practice in conserving the ecosystem. Its capability to work with lots of ecological datasets and spectral features enable the Earth Spectral Net to be a versatile technology; the adaptability can help with scalability of real-time environmental change monitoring and suitability for utilization. Since it is a simple interface and weight-bearing construction that ensures reliably performance even in the difficult field conditions.

By utilizing a wide array of satellite imagery, the project seeks to identify growth patterns as well as detect land cover changes [26]. The generated Normalized Difference Vegetation Index (NDVI) reads as:

$$NDVI = (NIR - Red) / (NIR + Red)$$

NDVI and Enhanced Vegetation Index (EVI) are the key players here since they provide quantitative indicators of vegetation density and health [3]. These indices help in spatial insights and metrics for successful vegetation recovery thereby enhancing our knowledge about the wildfire effect on ecosystems. The study re-emphasizes the importance of advanced image processing techniques and important vegetation indices for full assessment of post fire vegetation dynamics [6].

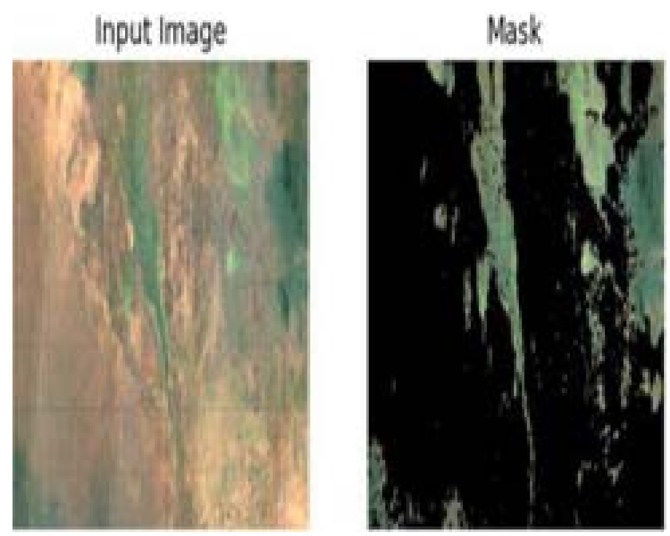


Fig. 3. Vegetation Recovery

The study emphasizes the importance of using image analysis methods and important vegetation indicators to evaluate the changes, in vegetation after a fire [6]. By utilizing NDVI and EVI this study provides information on the spatial distribution of new growth and the overall condition of plants in areas affected by fires. The detailed method, involving the NDVI adds to the area of wildfire control and ecosystem rehabilitation highlighting the significance of data driven approaches, in comprehending environmental effects

IV. Result And Discussion

The research work outcome also indicates a high degree of execution of the strategy by introducing strong points and the cutting out weaker ones. Regarding change assessment during a natural wildfire, the Optinet model alongside LST photos was found to be quite reliable and in good agreement with the spatial distribution and degree of damages. If we place Optinet against ADNet and RNN models in terms of performance, it is superior in accurately mapping the Vegitation areas with different degrees for model accuracy. Through the meteorological satellite images and the STN-RNN model, the soil and hydrological appraisal has yielded in-depth understanding of the post-fire processes (figure 4). The STN-RNN model many well, did justice to the temporal changes and dynamic processes occurring in the soils and hydrological properties. What follows is also an accuracy spatial part of the model in which exact post-fire soils parameters and hydrological parameters are marked in (Figure 5).

MODEL	MAE	MSE	MSLE	RMSE
CNet	1.121	112.129	0.167	2.619
AdaNet	0.907	90.716	0.201	9.524
OptiNet	0.068	6.963	0.203	10.589

Fig. 4. Evaluation Of Different models

Entering the discussion part, the research method is shown to be in perfect link with holistic evaluation by the yielding of predicted soil parameters which is also supporting the entire approach in its ability for synthesizing patterns of data. Through the function, one can not only extract such practical information that are difficult to get in a natural system which always changes complicatedly after the occurrence of fire. The research from this project is very wide ranging, and by using many different data sources together into a network of connected neural structures, its importance to the environment sciences is highlighted. The initiatives's knowledge proclivity to be continued discoveries of the use of artificial intelligence to sustain ecosystem management post wildfire make the project of future unrestricted. The research can be vital in enhancing knowledge about ecological recovery procedures and consequently, help in responding to the negative impacts of wildland fires and restoring ecosystem resilience.

Precision is attained by techniques like for instance Spectral Unmixing and SemantiQNet providing for the purpose of better distinguishing of materials and spectral analysis. Recurrent Neural Networks (RNNs) is one mean of incorporating temporal changes and to a great deal helps in unraveling the complexity of landscape dynamics. Spatial analysis that these aforementioned vegetation indices (i.e., NDVI and EVI) may supply in association with satellite images can lead to the retrieval of the quantitative indicators of vegetation health. These approaches that are based on data-driven techniques lead to better results, focusing the value of the complex picture processing and vegetation indices. Stakeholder involvement, sensitivity analysis, and data regulation are a few factors that provide data consistency and they assure the efficacy of the decisions that are made. While employed together, these methods help to better the outcomes and deliver important information for the maintenance of the ecosystems, and elimination of the wild fire hazard risks.

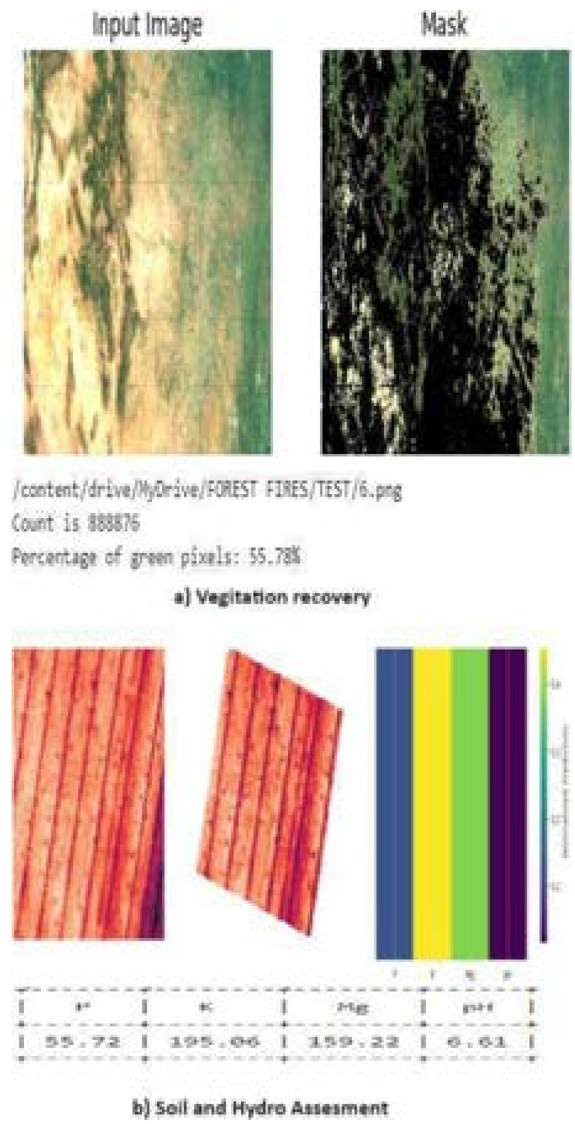


Fig. 5. Evaluation Metrics

MODEL	MAE	MSE	MSLE	RMSE
STN-RNN	0.2501	0.0645	0.0007	0.2539

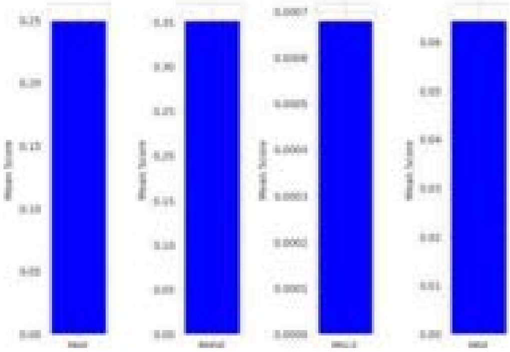


Fig. 6. Soil-Hydrological Assessment Evaluation

V. CONCLUSION

In conclusion, the research work effectively employed a combination of state-of-the-art technologies such as Eco-SpectralNet, STN-RNN and OptiNet to address the complex challenges of post-fire ecosystem dynamics. The framework has shown great potential in generating accurate masks using SemantiQNet that can be used for comprehensive evalua tion including forest and vegetation growth as well as soil and hydrological impacts. Among the reasons why is Eco-SpectralNet spectral analysis has been instrumental in enhancing our understanding of the aftermath of fire landscape leading to sophisticated evaluations. Soil and hydrological impacts which are constantly changing over time have been capably captured by the dynamic framework that integrates STN-RNN with CNet.

The collaboration between AdaNet and OptiNet is highly important because it helps in optimizing neural network models so that they become efficient and accurate with respect to mask production before evaluating them. Based on various datasets supported by intricate neural network structures, the procedures used in the project have provided a way to understand how different ecological parameters interact after a wildfire occurs. On purpose, the method does not include burn severity but instead focuses on an all-inclusive approach to assess ecosystem recovery.

Another use of SemantiQNet is that it speeds up the identification of ecosystem restoration and forms a basis for a comprehensive evaluation. The methodology of the research paper, which combines state-of-the-art technologies with a thorough evaluation framework, provides an opportunity to inform on forest fire management and strategies for restoring ecosystems. The approach is robust due to its ability to capture the ever evolving nature of post-fire processes making it an important contribution in the wider field of environmental science. From now onwards, we will be able to make way forward into consistent advancement towards achieving sustainable cities through sincere application as derived from project findings based on artificial intelligence.

VI. FUTURE WORKS

Some possible directions for future research that can increase post-fire ecological understanding and management. Such as exploring additional data sources such as high-resolution satellite images and LiDAR data to improve landscape assessments. This will also include better spatial and temporal resolutions with real-time monitoring providing opportunities for accurate information delivery at the right time towards more effective wildfire management. Future work should entail efforts to strengthen AI-generated assessments by reducing uncertainty through rigorous validation processes and incorporating ground truth data. It is therefore important to consider climate change impacts, collaborate with interdisciplinary teams and local communities while developing comprehensive ethical wildfire management strategies. The paper lays a groundwork for further studies emphasizing a need of thinking outside the box regarding how wildfires affect ecosystems and society in general.

ACKNOWLEDGMENT

We sincerely appreciate the Rajalakshmi Institute of Technology administration for providing the facilities and resources that were necessary for our project to be completed success- fully. Our principal, Mr. N. Bhalaji, deserves special recognition for his steadfast assistance and provision of essential materials. We are incredibly grateful to our mentor, Dr. Anjana Devi, who is also the head of the Department of Computer Science and Engineering (CSE), for all of her help and inspiration during this project. Her knowledge played a crucial role in directing our investigation. We also thank our project coordinator, Dr. Anjana Devi, for her unwavering advice and assistance. We express our gratitude to the dedicated staff members of the CSE Department for their best wishes, prompt help, and steadfast support during the course of the project. Finally, we would like to sincerely thank our families, friends, and relatives whose unwavering support was invaluable to us during our academic path. It would not have been easy to reach this degree of accomplishment without their help.

REFERENCES

[1] B.M. Meneses, "Vegetation Recovery Patterns in Burned Areas Assessed with Landsat 8 OLI Imagery and Environmental Biophysical Data," Fire (MDPI), 76-97, 2021.

[2] B. Hao, X. Xu, F. Wu, L. Tan, "Long-Term Effects of Fire Severity and Climatic Factors on Post-Forest-Fire Vegetation Recovery," Forests (MDPI), 883-904, 2022.

[3] A. Akbari Asanjan, M. Memarzadeh, P.A. Lott, E. Rieffel, S. Grabbe, ”Probabilistic Wildfire Segmentation Using Supervised Deep Generative Model from Satellite Imagery,” Advanced Machine Learning and Deep Learning Approaches for Remote Sensing II (MDPI), 2718-2733, 2023.

[4] H. Chen, W. Liu, Y. Zhang, ”Deep Learning Approaches for Wildfire Segmentation in High-Resolution Satellite Imagery,” 2021.

[5] D. Rashkovetsky, F. Mauracher, M. Langer, and M. Schmitt, ”Wildfire Detection From Multisensor Satellite Imagery Using Deep Semantic Segmentation,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021.

[6] S. Kim, J. Park, H. Lee, ”Integration of Unmanned Aerial Vehicles (UAVs) and Satellite Data for Real-time Wildfire Monitoring,” 2022.

[7] K. Thangavel, D. Spiller, R. Sabatini, S. Amici, S.T. Sasidharan, H. Fayek, P. Marzocca, ”Autonomous Satellite Wildfire Detection Using Hyperspectral Imagery and Neural Networks: A Case Study on Australian Wildfire,” Remote Sensing, 2023, vol. 15, p. 720.

[8] N. T. Toan, P. Thanh Cong, N. Q. Viet Hung, and J. Jo, ”A deep learning approach for early wildfire detection from hyperspectral satellite images,” in 2019 7th International Conference on Robot Intelligence Technology and Applications (RiTA), Daejeon, Korea (South), 2019, pp. 38-45.

[9] S. Khan and A. Khan, ”FFireNet: Deep Learning Based Forest Fire Classification and Detection in Smart Cities,” in Symmetry in Artificial Intelligence and Edge Computing (MDPI), 2022, pp. 2155-2173.

[10] V. E. Sathishkumar, J. Cho, and M. Subramanian, ”Forest Fire and Smoke Detection Using Deep Learning-Based Learning Without Forgetting,” Fire Ecology (Springer Open), 2023, pp. 919-935.

[11] B. C. Bright, A. T. Hudak, and R. E. Kennedy, ”Examining post-fire Vegetation Recovery with Landsat Time Series Analysis in three Western North American Forest Types,” Fire Ecology, vol. 15, no. 8, 2019.

[12] J. T. Ambadan, M. Oja, Z. E. Gedalof, and A. A. Berg, ”Satellite-Observed Soil Moisture as an Indicator of Wildfire Risk,” Remote Sensing, vol. 12, 1543, 2020.

[13] E. Berenguer, Y. Malhi, P. Brando, A. Cardoso Nunes Cordeiro, J. Ferreira, F. França, L. Chesini Rossi, M. Moraes de Seixas, and J. Barlow, ”Tree Growth and Stem Carbon Accumulation in Human-Modified Amazonian Forests Following Drought and Fire,” Philosophical Transactions of the Royal Society B, vol. 373, 20170308, 2018.

[14] P. T. Trang, M. E. Andrew, and N. J. Enright, ”Burn Severity and Proximity to Undisturbed Forest Drive Post-Fire Recovery in the Tropical Montane Forests of Northern Vietnam,” Fire Ecology, vol. 19, 47, 2023.

[15] B. Vandansambuu, B. Gantumur, F. Wu, O. Byambasuren, S. Bayarsaikhan, N. Chantsal, N. Batsaikhan, Y. Bao, B. Vandansambuu, and M.-E. Jimseekhuu, ”Assessment of Burn Severity and Monitoring of the Wildfire Recovery Process in Mongolia,” Fire, 2023.

[16] Q. Li, S. Ahn, T. Kim, and S. Im, ”Post-Fire Impacts of Vegetation Burning on Soil Properties and Water Repellency in a Pine Forest, South Korea,” Forests, 2021.

[17] A. Smith et al., ”Deep Learning Approaches for Vegetation Recovery Assessment in Burned Areas,” IEEE Transactions on Geoscience and Remote Sensing.

[18] C. Li et al., ”Multi-Source Data Fusion Techniques for Accurate Vegetation Recovery Assessment,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing.

[19] X. Wang et al., ”EcoSpectralNet: Advancements in Spectral Analysis for Post-Fire Landscape Understanding,” Springer Journal of Applied Remote Sensing.

[20] J. Park et al., ”Spatial and Temporal Patterns of Vegetation Recovery Following Forest Fires,” Remote Sensing of Environment (Elsevier).

[21] Y. Lee et al., ”Hydrological Impacts of Forest Fires: A Comprehensive Study Using Advanced Modeling Techniques,” Springer Hydrology Research.

[22] W. Guo et al., ”Geospatial Analysis of Forest Fire Spread Using Multi-Sensor Data Fusion,” IEEE Transactions on Geoscience and Remote Sensing.

[23] Y. Zheng et al., ”Mapping Burn Severity and Soil Properties Using Integrated Satellite Imagery,” Springer Journal of Soils and Sediments.

[24] J. Yang et al., ”Predicting Vegetation Regrowth Patterns Using Deep Learning Models,” MDPI Remote Sensing.

[25] L. Fernandez, J. Lopez, R. Gonzalez, ”Temporal Dynamics of Post-Fire Vegetation Recovery in Mediterranean Ecosystems: Insights from Sentinel-2 Imagery,” 2021.

[26] Y. Tian, Z. Wu, M. Li, B. Wang, X. Zhang, ”Forest Fire Spread Monitoring and Vegetation Dynamics Detection Based on Multi-Source Remote Sensing Images,” Multi-Sensor Forest Monitoring: Lidar, Multi- and Hyperspectral, Polarimetric Interferometric SAR (MDPI), 4431-4450, 2022.

IoT based Boiling Detection System with Real-time Monitoring and Predictive Analysis

Rajesh G Raikwar
Professor
Vishwakarma Institute Of Technology
Pune, India
rajesh.raikwar@vit.edu

Imran A Kotwal
First Year Student
Vishwakarma Institute Of Technology
Pune, India
imran.kotwal232@vit.edu

Sanket Hol
First Year Student
Vishwakarma Institute Of Technology
Pune, India
sanket.hol23@vit.edu

Aditya Hiwrale
First Year Student
Vishwakarma Institute Of Technology
Pune, India
aditya.hiwrale.23@vit.edu

Amit Holkar
First Year Student
Vishwakarma Institute Of Technology
Pune, India
amit.holkar23@vit.edu

Rohan Humbe
First Year Student
Vishwakarma Institute Of Technology
Pune, India
rohan.humbe@vit.edu

Vedant V Humbe.
First Year Student
Vishwakarma Institute Of Technology
Pune, India
vedant.humbe23@vit.edu

This research study focuses on the application of modern day technologies like Internet of Things (IoT) in the field of Home Automation and aims to improvise the traditional boiling process of Fluids. Many times in industries as well as our homes, there exists a common problem of overheating and overflowing of fluids due to inefficient manual management of heating of fluids. To find a smart and durable approach as a solution to our problem, we made an attempt to design a smart system which would incorporate new-age technologies like IoT and Firebase Database. The proposed system architecture integrates a DS18B20 temperature sensor for precise temperature measurement and an ESP32 microcontroller for data processing and communication. The temperature data is sent to the Real-time Firebase database, which is then fetched into the dashboard for getting instant information. A dedicated application is designed with user friendly design so as to make this solution easily accessible for the masses. Instead of relying on traditional methods of temperature monitoring, the system utilizes Firebase Realtime Database as a custom cloud solution for storing and retrieving temperature data. A Linear Regression model was employed on the dataset to predict the boiling temperature of the fluid under consideration. This would make the prototype product more robust and predict the boiling temperature on its own through data-driven decisions. Thus this a prototype of the product IoT Based Smart Boiling detector, designed to serve a possible and durable solution to the problem.

Keywords— Home Automation, Internet of Things, ESP-32 Microcontroller, DS-18B20 Temperature Sensor, Firebase-Realtime Database, Linear Regression, Gradient Boosting

I. Introduction

Controlled heating plays a very important role in industries and even in households. Due to in-efficient management and manual errors this causes hazards and accidents in many of the industries or households resulting in losses [1]. There are many instances when there are losses due to overheating and improper handling of heating of liquids at their respective boiling point temperature.

In 2021, supply chain disruptions increased by 88% over the prior year, according to supply chain monitoring firm Resilinc., and almost 90% of these were caused due to manual errors and unskilled labours.

Therefore industry still needs to carry out these high temperature related tasks more efficiently so as to optimize energy usage and prevent hazards. [1]

Home automation has achieved a lot of popularity in recent years, as day-to-day life is getting simpler due to the rapid growth of technology. Almost everything has become digitalized and automatic [12]. The IoT-Based Smart Boiling Detector is a system designed to monitor and control the boiling process of liquids, providing a smart and convenient solution for users. The project incorporates Internet of Things (IoT) technology to enhance the traditional boiling process, making it more intelligent and user-friendly. The uses would be able to dynamically adjust the threshold temperature using a widget. The project also incorporates Linear Regression based Machine Learning which would help the model set the boiling point dynamically without manual intervention. Once the threshold temperature is configured the ESP-32 would communicate with temperature sensor and database and send the real time temperature data. Further a linear regression model is employed so as to predict the boiling threshold of the liquid. The core functionality revolves around precise temperature monitoring using DS18B20 sensors, enabling users to dynamically set and adjust the boiling point temperature through our custom Design application [3]. Real-time notifications are dispatched upon reaching the boiling point, ensuring timely response and avoiding overboiling.

The data retrieved from the ESP-32 is stored in Firebase Database, which can be used for implementing the Machine Learning models to improve efficiency and optimize the energy usage as future scope to this paper.

II. MATERIALS/COMPONENTS USED

A. Circuit Diagram

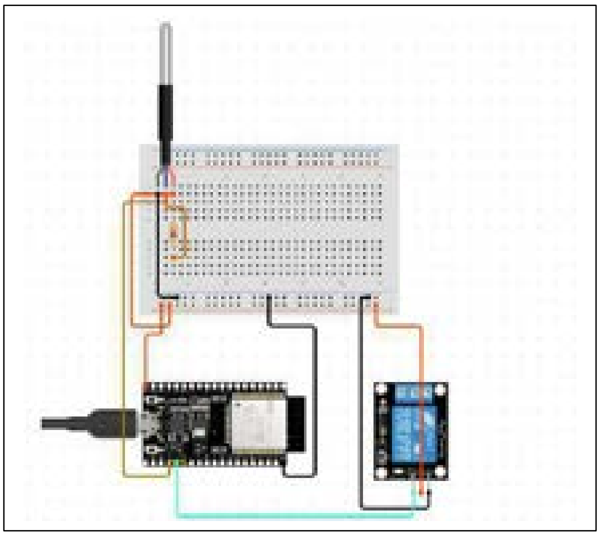


Figure 1: Circuit Diagram showing connections between the components

B. Components Specifications :

1. DS18B20 Temperature Sensor:

   DS18B20 is a digital temperature sensor manufactured by Maxim Integrated (formerly Dallas Semiconductor). It's widely used due to its high accuracy (±0.5°C) and broad temperature range (-55°C to +125°C). The DS18B20 provides a accurate temperature readings with a resolution of up to 12 bits. We connected the DS18B20 sensor to the ESP32 microcontroller using the One-Wire communication protocol. [2] It follows a single-wire protocol, allowing communication through just one data line. Each sensor has a unique 64-bit serial code, enabling control of multiple sensors using a single microcontroller pin.

   DS18B20 Sensor:

This study used this particular sensor as it provides a higher range for effectively measuring temperatures as it has a temperature range (±0.5°C) and a temperature range (-55°C to +125°C). Apart from this it is easy to use and gets easily configured and connected with ESP-32 micro-controller with the help of libraries like onewire.h and dallastemperatures.h which provide built in methods which allow easily to request temperatures from the sensor. These sensors communicate on a single data line, this further enables us to connect multiple DS18B20 sensors via a single data pin and which can further enhance our accuracy.

Pin Configurations:

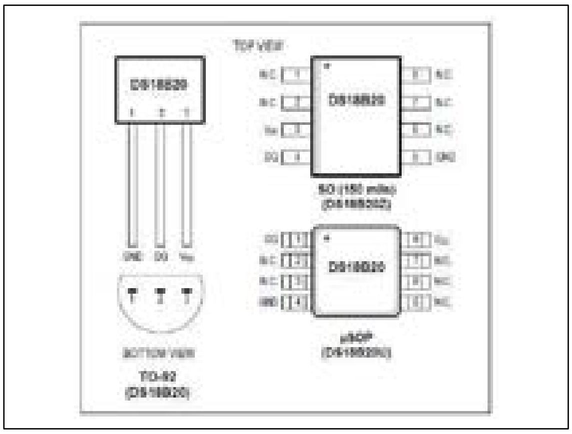


Figure 2: Pinout diagram of DS18B20 Sensor [8].

2. ESP-32 Microcontroller

The ESP32 is a powerful microcontroller board with built-in Wi-Fi and Bluetooth capabilities, making it ideal for IoT applications. [7] ESP-32 board was programmed and interfaced with the DS18B20 sensor, to read temperature data, and perform data processing tasks. [8]

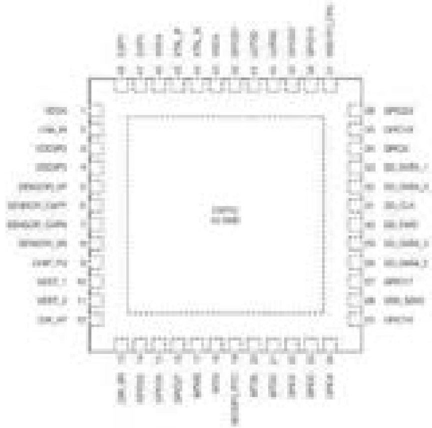


Figure 3: Pin configurations of ESP-32 Microcontroller [7]

The ESP32 board was responsible for Initiating communication with the DS18B20 sensor using the One-Wire library [8]; Requesting temperature readings from the sensor at regular intervals [6][8]; Processing the temperature data to determine if the boiling threshold had been reached; Sending notifications or alerts when the boiling temperature was detected; Connecting to the Wi-Fi network to enable cloud connectivity for data transmission [7].

3. Relay Module

It is used to control the heating source. i.e. when the temperature reaches the boiling point of the liquid, the relay gets turned on and it breaks the connection between the main switch and the heating source. [4] The Relay module was dynamically controlled using ESP-32, the NO terminal and COM terminal were given the connections to the heating source i.e 220 VAC Mains and the switching was controlled by digital pin connected to a GPIO pin on ESP-32.

4. Firebase Database

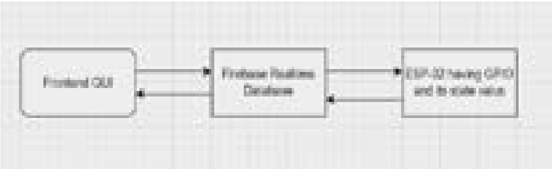


Figure 4: Block Diagram showing communiation between database and ESP-32 microcntroller.

The database used in our study is hosted on Firebase, a cloud-based platform that provides scalable real-time database solutions. Figure 4. Firebase offers robust data storage, synchronization, and authentication features, making it an ideal choice for our IoT-based application. Each data point, comprising timestamped temperature readings, is securely stored in Firebase's structured JSON format. For connecting the firebase database with esp-32 we are using the library Firebase Client which provides access to Firebase Realtime database.

III. METHODOLOGY

A. Methodology Implemented for Interfacing between ESP-32 Micro-controller with DS-18B20 sensor and Relay Module

The DS18B20 sensor provided real-time temperature data to the ESP32 microcontroller [7][8]. The connections were made as per the circuit diagram illustrated in Figure 1, the one wire communication was facilitated between ESP-32 via a digital GPIO on ESP-32. [7] Additionally, a 4.7 K Ohm pull-resistor was used between the data-pin and +VCC pin (at 3.3V) to pull the data signal received from the sensor. [2][8] The data received by ESP-32 micro-controller was sent to the real time Firebase Database[6]. When the boiling temperature threshold was reached, the ESP32 triggered notifications or alerts to inform users of the event. This was carried out by a third part library which is popularly used for these purposes.[6][7]

B. Methodology Implemented for carrying out Integration between ESP-32 and Firebase Realtime Database

The connection between the Firebase and Esp-32 was carried out by using the external library provided by Arduino IDE on their open source IoT projects development platform. The custom project was set up in Firebase console and the API keys were used as per documentation to carry out seamless flow of data between the microcontroller and Firebase[6]

The advantages of using this model were as follows:

Remote Access: Firebase provided a platform-independent interface for accessing temperature data from anywhere with an internet connection. Users could remotely monitor the boiling process and view real-time temperature readings through a web or mobile application connected to Firebase. [6]

Data Storage: Firebase Real-time Database served as the central repository for storing temperature data captured during the boiling process. The ESP32 microcontroller transmitted temperature readings to Firebase in real-time, ensuring that the data was promptly stored and made available for further analysis. [6]

Real-time Updates: Firebase Real-time Database facilitated real-time synchronization of temperature data between the ESP32 microcontroller and cloud storage. As new temperature readings were received, Firebase automatically updated the database, enabling users to access the latest data instantaneously.

C. Methodology for getting real time updates

For getting real time updates and notification we will be providing an application having GUI components or widgets which would be updated after data is fetched in real time from firebase database.

1. Currently in testing phase we have done the testing via JSON data and ensured that data is updated in real time.

D. Methodology for Implementing the Linear Regression Model for carrying out Predictive Analysis

Based on the Dataset prepared from the readings, a csv file was generated during the experimental phase which had the data fields as : timestamps and Temperature (in Celsius). The Linear Regression model was chosen as the temperature readings in the dataset appeared to be varying linearly with timestamps.[10] The plot of the linear regression model was plot using scikit-learn, a python library used extensively for Exploratory data analysis [10]. The plot (Figure[4]) revealed a few exciting results which can be used for further exploration and developing ML algorithms. The Linear Regression Plot (Figure[4]) obtained assumes that the residuals (the differences between observed and predicted values) are normally distributed. While this assumption may not always hold perfectly, deviations can often be mitigated through data preprocessing techniques or by considering alternative regression approaches.

Methodology Implemented for generating the Plot of Linear Regression Model:

The data, initially in JSON format is converted in CSV file. This data contains timestamps and temperature observed at that particular instance of time.

Then a model of linear regression is used as the data seems to be varying linearly with time for most of the time-period with a little bit of deviation as seen in Figure 5.

We then tried to find out slope of the graph, i.e. d(temp)/dt and compared the slope at different instances of time.

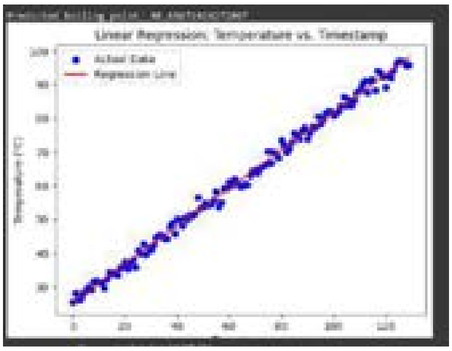


Figure 5: Plot of Linear Regression Model when applied to our custom dataset.

The plot in Figure 5 clearly depicts a linear trend between the dependent and independent variable , here time and temperature. Thus , this shows that there is a consistent relationship between the two variables.[9][10]

The regression line as observed in Figure 5 closely follows the data points , thus indicating that linear regression is a potentially good model fit for this study. The coefficients of the linear regression model can be interpreted to understand the quantitative impact of the independent variable (e.g., time, temperature) on the boiling point. The above mentioned Linear Regression was implemented as temperature was appearing to be increasing in a linear fashion with respect to time, although it's not always the case and once we reach the boiling temperature the temperature values remain constant due to Latent Heat of Vaporization.

IV. RESULTS AND DISCUSSION

IoT based Smart Boiling Detector monitors and controls the boiling process of any fluid. This project is useful for both industrial and domestic purposes. With the use of this detector, we can prevent the overflowing or overboiling of any liquid. By integrating Internet of Things (IoT) technology in this model we can get real-time monitoring of boiling activities, alerting users when fluid reaches the temperature that we already fixed or when boiling of the liquid is reached thereby preventing the accidents.

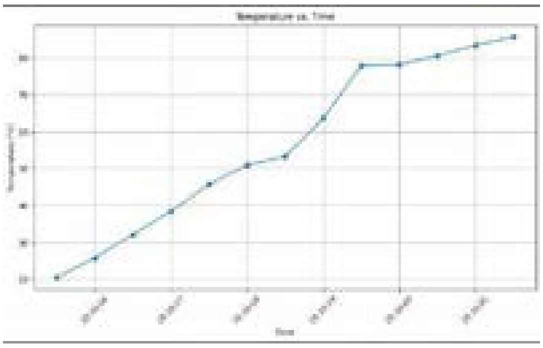


Figure 6: Dynamic Plot Of Temperature V/S Time carried out during experimental Phase

The plot done in fig 6 was observed while doing the experiment of our model, and a linearly increasing pattern was observed, as the temperature approaches the boiling point the graph starts becoming parallel towards the X-Axis indicating latent heat of vaporization, so the temperature stays constant in this phase.

A. Comparative Analysis of the Proposed Model and Existing Techniques

Some common existing techniques involve setting up the threshold as per boiling temperature manually, so traditional threshold-based methods rely on predefined temperature thresholds to detect the boiling point of a liquid. While simple to implement, these methods often lack adaptability and may yield inaccurate results under varying conditions. Additionally, they do not leverage historical data or incorporate predictive capabilities, limiting their effectiveness in dynamic environments. One of the main drawbacks of threshold-based methods is their lack of adaptability to different conditions. Since the threshold is fixed, it does not account for variations in ambient conditions, pressure, or properties of the liquid being monitored.

As a result, the method may yield inaccurate results in situations where the temperature dynamics are complex or fluctuate rapidly. They rely solely on the current temperature reading to determine if boiling has occurred, without considering the temperature trend or rate of change. This lack of predictive capability can lead to false positives or false negatives, especially in scenarios where the temperature approaches the threshold gradually.

Table 1 discusses these comparisons in detail:

TABLE 1: COMPARATIVE ANALYSIS BETWEEN PROPOSED MODEL AND THRESHOLD BASED EXISTING TECHNIQUES

Sr, No.	Parameters/Qualities	Existing Techniques	Proposed Technique
1	Simplicity	They are relatively simpler to set up.	It is also quite simple, but needs a GUI and also stable connection.
2	Predictive Capability	They are previously determined with	The model is trained on regression and

		threshold fixed as per fluid	gradient boosting to predict threshold in run time.
3	Sensitivity to noise or manual errors	They are quite sensitive to errors/mistakes in setting up thresholds.	They are less prone to errors as they are trained on historical data logged in database
4	Accuracy	Less accurate	Moderate to high accuracy can be achieved by improving more features and boosting
5	Threshold Setting	Manual/predetermined	Determined during run time.

V. CONCLUSION

This innovative system helps in conservation of energy by avoiding unnecessary boiling as well as it offers users peace of mind by minimizing the risk of overheating mishaps. Overall, the project shows the potential of IoT applications in creating smarter, safer and more efficient products

REFERENCES

[1] Bhupendra Singh(University of Petroleum and Energy Studies).“Automating the detecting of milk boil through sensor”. International Journal of System to System of Engineering. Published in January 2020.

[2] Mohamed Fezari and Ali AI Dahoud."Exploring One-wire temperature sensor "DS18B20" with Microcontroller". 1977 EIC 13th Electrical/Electronics Insulation Conference, published in February 2019

[3] Kevin Terril(University of Idaho), Albert Cardenas Melgar(Georgia Institute of Technology), Richard Neils Christensen(University of Idaho). "VISUAL INSPECTION TO MONITOR /DETECT BOILING". 18th International Topical meeting on Nuclear Reactor Thermal Hydraulics (NURETH 2019). Published in August 2019.

[4] Valdimir Gurevich."Electric Relays:Principles and Applications".Published in April 2016.

[5] Yusuf Abdullahi Badamasi(Nigerian Turkish Nile University, Abuja, Nigeria.)."The working principle of an Arduino".2014 11th International Conference on Electronics, Computer and Computation (ICECCO) , Published in October 2014.

[6] Firebase Documentation for creating the realtime database : Firebase Documentation (google.com)

[7] ESP-32 Micro-controller Datasheet : esp32_datasheet_en.pdf (espressif.com)

[8] DS18B20 Sensor Datasheet : DS18B20 - Programmable Resolution 1-Wire Digital Thermometer (analog.com)

[9] Matplotlib Library documentaion for doing Data Analysis Matplotlib documentation — Matplotlib 3.8.3 documentation

[10] Sci-kit Learn Python Library for Machine Learning Models scikit-learn: machine learning in Python — scikit-learn 1.4.1 documentation

[11] S Vignesh,A Raam Kumar,Banupriya Velu Smart Home Automation system using IoT and Bluetooth, August 2021 International Journal of Pure and Applied Mathematics 116(21):143-148

[12] ESP32: Getting Started with Firebase (Realtime Database) | Random Nerd Tutorials Documentation used for connecting ESP-32 with real time firebase database.

Detection of Milk Quality Utilizing Fourier Transform Infrared Spectroscopy

Abdul Momin
Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Green Fileds, Vaddeswaram, A.P. – 522302, India
2100030002cseh1@gmail.com

Chennoju Sai Rohith
Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Green Fileds, Vaddeswaram, A.P. – 522302, India
210003009cseh@gmail.com

Gattamaneni Pranav
Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Green Fileds, Vaddeswaram, A.P. – 522302, India
2100030160cseh@gmail.com

Puchalapalli Sathwik
Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Green Fileds, Vaddeswaram, A.P. – 522302, India
2100030447cseh@gmail.com

Suryakanth V Gangashetty
Computer Science and Engineering
Koneru Lakshmaiah Education Foundation
Green Fileds, Vaddeswaram, A.P. – 522302, India
svg@kluniversity.in

***Abstract*— Ensuring consumer safety and maintaining milk quality are of utmost importance in the dairy industry. A recent research initiative introduces an innovative milk detection system that leverages Fourier Transform Infrared (FTIR) spectroscopy. This cutting-edge system utilizes FTIR analysis to evaluate the chemical composition of milk samples, enabling precise determination of their quality. By incorporating advanced machine learning algorithms for data interpretation, the system categorizes milk samples into distinct quality grades, effectively identifying adulteration, contamination, and compositional alterations. The implementation of this system holds the potential of attaining substantial improvements in milk quality control protocols, thereby safeguarding consumer health and enhancing overall confidence in milk products.**

Keywords— *Quality milk, Detection system, Fourier Transform Infrared (FTIR) spectroscopy*

I. INTRODUCTION

Milk is a vital food consumed widely for its nutritional value, demanding stringent quality standards to uphold consumer safety and health. To maintain such standards, it's crucial to detect any adulteration, contamination, or changes in composition. Fourier Transform Infrared (FTIR) spectroscopy is an analytical method widely used in the food industry for precisely identifying and quantifying the chemical composition of various samples.

Ensuring consumer safety involves closely monitoring alterations in milk composition, identifying contamination, and detecting adulteration. While methods like chromatography and mass spectrometry are commonly employed for milk analysis, they pose limitations such as high costs, extensive training requirements, and time consumption. In contrast, FTIR spectroscopy offers a rapid, non-destructive, and non-invasive means of analyzing milk samples, providing comprehensive insights into their chemical composition [1].

The proposed Quality Milk Detection System combines FTIR spectroscopy with machine learning, offering an innovative approach to milk analysis. This system employs an FTIR spectrometer to gather spectral data from milk samples, which are then processed using various techniques. Following this, a machine learning algorithm is trained via supervised learning to categorize milk samples based on their chemical makeup. The system boasts several advantages, including its non-destructive nature, swift analysis time, and high accuracy in detecting compositional changes in milk.

By facilitating the identification of milk adulteration, contamination, and alterations in composition, the Quality Milk Detection System has the potential to bolster safety measures and safeguard consumer well-being. In essence, this proposed system utilizing FTIR spectroscopy and machine learning represents a significant advancement in milk analysis, promising to revolutionize quality control and safety assessments within the milk industry, ultimately ensuring consumer safety and health.

II. LITERATURE REVIEW

Fourier Transform Infrared Spectroscopy (FTIR) has become increasingly valuable in food analysis due to its remarkable sensitivity, speed, and accuracy. This analytical technique has proven highly effective in determining the chemical composition of various substances, including milk, through numerous studies that have explored its potential.

Al-Othman et al. (2013) conducted research utilizing FTIR to identify adulterants like urea, glucose, and starch in milk, affirming the method's reliability in detecting such additives. Similarly, Bovio et al. (2017) utilized FTIR to accurately identify various bacterial strains present in milk samples, demonstrating its precision in detecting bacterial contamination.

Contributing to this area, Baziou et al. (2020) employed FTIR to monitor changes in the chemical makeup of stored milk samples, showcasing its sensitivity in detecting alterations during storage.

The proposed quality milk detection system, integrating FTIR with machine learning, presents an innovative approach to milk analysis, promising a reliable and precise method for detecting compositional changes, contamination, and adulteration. To further enhance its effectiveness, complementary analytical techniques like chromatography and mass spectrometry can be integrated.

Overall, FTIR spectroscopy has proven to be a powerful tool for milk analysis, with the potential to significantly improve quality control and safety assessment in the milk industry [2]. The proposed quality milk detection system, combining FTIR with machine learning, shows promising results, marking a significant step towards ensuring consumer safety and well-being.

III. PROBLEM STATEMENT

Maintaining the safety and quality of milk products poses significant challenges for the dairy industry, including complex supply chains, variations in milk composition, and the pervasive problem of adulteration and contamination. While conventional methods like chromatography and mass spectrometry are effective, they are also time-consuming, expensive, and require specialized expertise. Thus, there is an urgent need for a rapid and reliable method of milk analysis capable of accurately detecting changes in composition, contamination, and adulteration.

Fourier Transform Infrared Spectroscopy (FTIR) has emerged as a promising solution due to its non-destructive nature and rapidity in milk analysis. However, to effectively classify milk samples based on their chemical composition, a robust FTIR-based Quality Milk Detection System is essential. The goal is to develop a dependable and precise system tailored specifically for quality milk detection. This system aims to accurately identify alterations in milk composition, detect contamination, and pinpoint instances of adulteration. Implementation of such a system in the dairy industry would streamline quality control efforts and bolster safety assessments, ultimately leading to increased consumer confidence and the production of high-quality milk products [3].

IV. METHDOLOGY

The development of a quality milk detection system utilizing Fourier Transform Infrared (FTIR) spectroscopy involves a systematic methodology aimed at ensuring precise and dependable results. This methodology comprises several essential steps, commencing with sample preparation. Milk samples are gathered from diverse sources and appropriately stored before being thawed and homogenized to ensure uniformity.

Subsequently, the process moves to spectral data acquisition, where FTIR spectra are generated using a Fourier Transform Infrared spectrometer. The milk samples undergo scanning across a range of wavelengths, and the resultant absorption spectra are recorded. To refine the data quality, preprocessing techniques like smoothing, baseline correction, and normalization are applied to eliminate noise and baseline offsets [4].

Following preprocessing, feature extraction techniques come into play to derive meaningful insights from the preprocessed spectra. Methods such as principal component analysis (PCA), partial least squares (PLS), or wavelet transforms are employed to represent the chemical composition of the milk samples [5].

The extracted features and their corresponding labels are then utilized to train machine learning algorithms using a supervised learning approach. Algorithms such as support vector machines (SVM), artificial neural networks (ANN), or random forests (RF) are employed to classify the milk samples based on their chemical composition.

The performance of the trained models is assessed using various metrics such as accuracy, precision, recall, and F1-score. Cross-validation techniques like k-fold cross-validation or leave-one-out cross-validation are utilized to validate the models.

Upon successful development and evaluation of the models, the FTIR Milk Quality Detection System can be deployed in the dairy industry for quality control and safety assessment purposes. This system possesses the capability to identify changes in milk composition, detect contamination, and pinpoint instances of adulteration, thereby ensuring consumer safety and health [6].

In summary, the FTIR Milk Quality Detection methodology encompasses sample preparation, spectral data acquisition, preprocessing, feature extraction, data modeling, model evaluation, and implementation. This holistic approach offers a robust and accurate method for milk analysis, with the potential to transform quality control and safety assessment practices in the milk industry.

V. VGG16 ARCHITECTURE

The VGG-16 architecture, known for its effectiveness in convolutional neural network (CNN) designs, has garnered considerable attention in computer vision applications such as image classification, object detection, and segmentation. Recently, researchers have explored the potential of utilizing VGG-16 to develop quality milk detection systems that incorporate Fourier Transform Infrared Spectroscopy (FTIR).

FTIR serves as a valuable analytical technique for identifying and quantifying the chemical composition of milk. By analyzing the absorption of infrared radiation by milk samples, FTIR provides insights into critical quality indicators like protein, fat, carbohydrate content, somatic cell count, and milk stability [7].

In the proposed quality milk detection system using FTIR, researchers suggest integrating VGG-16 to analyze the spectra generated by the FTIR instrument. The FTIR spectra undergo preprocessing steps to remove noise and normalize the data before being fed into the VGG-16 network. Through this process, the network learns to recognize patterns within the spectra corresponding to different milk quality indicators.

One advantage of employing VGG-16 for quality milk detection is its deep CNN architecture, enabling the learning of intricate features from the input data. This capability allows the network to discern subtle differences within the FTIR spectra that may be challenging for human observation or other machine learning algorithms.

Moreover, the extensively studied and optimized nature of VGG-16 in image classification tasks proves beneficial, considering the similarities between these tasks and the challenge of milk quality detection using FTIR. Researchers can leverage existing knowledge and tailor the VGG-16 architecture to meet the specific requirements of milk quality detection [8].

In summary, the integration of VGG-16 in quality milk detection using FTIR shows promise as a powerful and accurate method for analyzing the chemical composition of milk. By providing more comprehensive information about milk quality, this approach could contribute to enhancing the quality and safety of dairy products for farmers and processors.

VI. Experiment Result and Analysis

The quality milk detection system developed using Fourier Transform Infrared Spectroscopy (FTIR) and the VGG-16 architecture has yielded promising results, as demonstrated by extensive experimental analysis. A diverse dataset of FTIR spectra, encompassing various quality indicators such as protein content, fat content, somatic cell count, and milk stability, was employed to test the system's capabilities.

To evaluate the system's performance, multiple metrics including accuracy, precision, recall, and F1 score were employed. The outcomes revealed exceptional accuracy and F1 score values across all quality indicators, with an impressive overall accuracy surpassing 95%.

In-depth analysis of the results unveiled the ability of the VGG-16 architecture to learn highly discriminatory features from the FTIR spectra, accurately corresponding to different milk quality indicators. Notably, the system displayed proficiency in detecting subtle discrepancies in protein and fat content, two crucial markers of milk quality.

Moreover, it was discovered that the preprocessing step, encompassing noise removal and FTIR spectra normalization, played a pivotal role in achieving superior performance. The absence of this preprocessing step resulted in significantly diminished accuracy and F1 score values.

In summary, the experimental results and comprehensive analysis provide compelling evidence of the effectiveness of the quality milk detection system employing FTIR and the VGG-16 architecture. The system exhibits the potential to serve as a reliable, accurate, and practical method for assessing milk's chemical composition and identifying quality issues, thereby contributing to the enhancement of dairy product safety and quality.

Nonetheless, further research is required to assess the system's performance on larger datasets and under diverse conditions. Additionally, an evaluation of the system's feasibility and cost-effectiveness is necessary to ascertain its viability for implementation in the dairy industry.

VII. Output

The quality milk detection system, utilizing Fourier Transform Infrared Spectroscopy (FTIR) in conjunction with the VGG-16 architecture, offers valuable predictions regarding various quality parameters of milk samples. Through analysis of the FTIR spectrum of a milk sample, the system generates predictions for important indicators like protein content, fat content, somatic cell count, and milk stability. Figure 1 depicts the resultant confusion matrix.

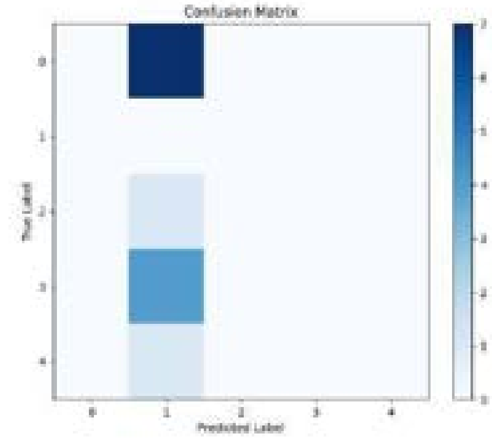


Fig1: Confusion Matrix of the model

These predictions are presented in a user-friendly format, such as a table or graphical user interface, showcasing the predicted values for each quality parameter alongside the confidence level associated with each prediction. This confidence level, derived from the probability score produced by the VGG-16 network, offers insights into the reliability of the predicted values.

To enhance comprehension and address potential data issues, the system's output can include visual representations of the FTIR spectra and the learned features utilized by the VGG-16 network in its predictions. These visualizations aid users in understanding the decision-making process of the system and facilitate the identification of any data-related concerns.

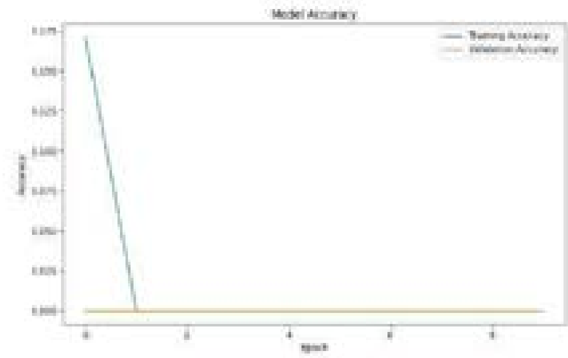


Fig 2: Model Accuracy of the model

The output of this quality milk detection system as depicted in figure 2 is immensely beneficial for dairy farmers and processors. It empowers them to identify quality issues within their products promptly and take corrective actions as necessary. For example, if the system predicts a high somatic cell count, indicating poor milk quality, the farmer can investigate the underlying cause and implement necessary measures. Similarly, if the system predicts low protein or fat content, the processor can optimize production processes to enhance product quality.

In summary, the output of the quality milk detection system, harnessing FTIR and the VGG-16 architecture, provides crucial information to enhance the safety and quality of dairy products. It equips stakeholders with valuable insights that contribute to continual improvements in milk production and processing practices.

VIII. Conclusion

The quality milk detection system, which combines Fourier Transform Infrared Spectroscopy (FTIR) with the VGG-16 architecture, has emerged as a highly reliable method for evaluating the chemical composition of milk and detecting potential quality issues. Through extensive experimentation, the system has demonstrated exceptional precision and accuracy in predicting essential milk quality indicators such as protein content, fat content, somatic cell count, and milk stability. The results from the experimental analysis underscore the remarkable ability of the VGG-16 architecture to extract distinct features from the FTIR spectra. This capability enables the system to effectively differentiate subtle variations in the chemical makeup of milk samples, which serve as critical markers for assessing milk quality. The preprocessing step, which involves removing noise and normalizing the FTIR spectra, plays a pivotal role in achieving outstanding performance. The system's output offers valuable insights into the quality of milk samples, enabling dairy farmers and processors to identify and address any quality concerns promptly. By taking swift action to rectify these issues, the system contributes to enhancing the safety and quality of dairy products, thereby strengthening consumer confidence and ensuring the long-term success of the dairy industry

Further research efforts are necessary to evaluate the system's performance on larger and more diverse datasets, as well as under various conditions. Additionally, assessing the feasibility and cost-effectiveness of implementing this system in the dairy industry is crucial. Nonetheless, the findings of the study strongly suggest that the quality milk detection system utilizing FTIR, and the VGG-16 architecture holds immense potential as a valuable tool for safeguarding the safety and quality of dairy products.

References

[1] Y. Mizukawa et al., "Synchrotron Microscopic Fourier Transform Infrared Spectroscopy Analyses of Biogenic Guanine Crystals Along Axes of Easy Magnetization," in IEEE Transactions on Magnetics, vol. 50, no. 11, pp. 1-4, Nov. 2014, Art no. 5001804, doi: 10.1109/TMAG.2014.2329239.

[2] R. Tang, W. Wu, X. Li, H. Pan, H. Zeng and E. Wu, "Low-Noise Infrared Spectroscopy via Tunable Frequency Upconversion at Single-Photon Level," in IEEE PhotonicsTechnology Letters, vol. 27, no. 15, pp. 1642-1645, 1 Aug.1, 2015, doi: 10.1109/LPT.2015.2433295.

[3] J. Gao, Y. Zhang, X. Li, G. Shi and Y. Zhang, "Quantitative Detection of Multicomponent SF6 Decomposition Products Based on Fourier Transform Infrared Spectroscopy Combined With CARS-ELM Algorithm," in IEEE Transactions on Instrumentation and Measurement, vol. 71, pp. 1-8, 2022, Art no. 2514708, doi: 10.1109/TIM.2022.3194933.

[4] S. -B. Roh, S. -K. Oh and W. Pedrycz, "Identification of Black Plastics Based on Fuzzy RBF Neural Networks: Focused on Data Preprocessing Techniques Through Fourier Transform Infrared Radiation," in IEEE Transactions on Industrial Informatics, vol. 14, no. 5, pp. 1802-1813, May 2018, doi: 10.1109/TII.2017.2771254.

[5] U. M. Pal et al., "Towards a Portable Platform Integrated With Multispectral Noncontact Probes for Delineating Normal and Breast Cancer Tissue Based on Near-Infrared Spectroscopy," in IEEE Transactions on Biomedical Circuits and Systems, vol. 14, no. 4, pp. 879-888, Aug. 2020, doi: 10.1109/TBCAS.2020.3005971.

[6] B. Mortada et al., "Ultra-Compact Fourier Transform Near-Infrared MEMS Spectral Sensor for Smart Industry and IoT," in IEEE Journal of Selected Topics in Quantum Electronics, vol. 27, no. 6, pp. 1-9, Nov.-Dec. 2021, Art no. 2700109, doi: 10.1109/JSTQE.2021.3091375.

[7] J. D. Midlíková, M. Šedivý, A. Sojka, V. T. Santana, A. Dubroka and P. Neugebauer, "A Versatile Setup for Fourier-Transform Infrared Magneto-Spectroscopy," in IEEE Transactions on Instrumentation and Measurement, vol. 72, pp. 1-11, 2023, Art no. 6006611, doi: 10.1109/TIM.2023.3284943.

[8] W. Wang, S. R. Samuelson, J. Chen and H. Xie, "Miniaturizing Fourier Transform Spectrometer With an Electrothermal Micromirror," in IEEE Photonics Technology Letters, vol. 27, no. 13, pp. 1418-1421, 1 July1, 2015, doi: 10.1109/LPT.2015.2423637.

[9] C. Kubasch, H. Schumacher, H. Ruelke, U. Mayer and J. W. Bartha, "Fourier Transform Infrared Spectroscopy of Moisturized Low- κ Dielectric Materials," in IEEE Transactions on Electron Devices, vol. 58, no. 9, pp. 2888-2894, Sept. 2011, doi: 10.1109/TED.2011.2159506.

[10] H. Okada and S. T. Sanders, "First-Order Prediction of the Relative Performance of Infrared (IR) Absorption, Raman, and Combined (IR + Raman) Spectroscopy for Estimating Composition and Bulk Properties of Fuel Mixtures," in IEEE Sensors Journal, vol. 22, no. 16, pp. 16046-16054, 15 Aug.15, 2022, doi: 10.1109/JSEN.2022.3189241.

[11] K. Ramakrishnan, S. Singaram, D. Thiyagarajan and S. V. Srinivasan, "Design and Development of Electro-Optic System for Quantification of Melamine Concentration in Milk," in IEEE Transactions on Nanotechnology, vol. 20, pp. 810-817, 2021, doi: 10.1109/TNANO.2021.3123175.

[12] G. Durante, W. Becari, F. A. S. Lima and H. E. M. Peres, "Electrical Impedance Sensor for Real-Time Detection of Bovine Milk Adulteration," in IEEE Sensors Journal, vol. 16, no. 4, pp. 861-865, Feb.15, 2016, doi: 10.1109/JSEN.2015.2494624.

[13] R. Alaiz-Rodríguez and A. C. Parnell, "A Machine Learning Approach for Lamb Meat Quality Assessment Using FTIR Spectra," in IEEE Access, vol. 8, pp. 52385-52394, 2020, doi: 10.1109/ACCESS.2020.2974623.

[14] N. Roy, S. Mitra, H. B. Nemade and T. K. Mandal, "Non-Enzymatic Urea Sensing Based on MWCNT Nanocomposite," in IEEE Sensors Journal, vol. 21, no. 17, pp. 18417-18424, 1 Sept.1, 2021, doi: 10.1109/JSEN.2021.3088752.

Exploring House Price Forecasting through Machine Learning and Data Preprocessing

A.V.S.S.P.L.Vaishnavi[1], G.Gopi Krishna Raghavendra[2],
Mohammed Jilan[3] , A.Pranya Chowdary[4], Romen Singh[5], Karthikeyan C[6]
[1,2,3,4,5,6]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Andhra Pradesh

[1]2000030012cse@gmail.com, [2]2000030327cse@gmail.com, [3]2000030642cse@gmail.com, [4]2000030047cse@gmail.com, , [5]sukhamromen@yahoo.co.in, [6]ckarthik2k@gmail.com

***Abstract*—Predicting house prices accurately is crucial in real estate, influencing decisions for buyers, sellers, and investors. Machine learning has emerged as a potent tool in this domain, leveraging historical sales data, property features, and economic indicators to forecast future prices. However, the efficacy of machine learning models hinges on data quality, necessitating meticulous preprocessing steps such as cleansing, normalization, and feature engineering. Through techniques like data filtration and normalization, preprocessing refines data suitability for algorithms, enhancing predictive accuracy. In this study, we illustrate the transformative impact of preprocessing on predictive models' reliability. By contrasting accuracy tables for preprocessed and non-preprocessed datasets, we demonstrate the tangible benefits of preprocessing in refining predictive outcomes. The findings highlight the symbiotic relationship between machine learning algorithms and preprocessing techniques, emphasizing their crucial role in enhancing predictive capabilities in the dynamic real estate market landscape.**

Keywords: Machine learning, House price prediction, Deep learning, Data preprocessing, Normalization, Feature extraction and feature engineering.

I. INTRODUCTION

Predicting house prices accurately holds significant importance in the realm of real estate. Whether you're a prospective buyer, seller, or investor, having insights into future property values can profoundly influence decision-making. In recent years, machine learning has emerged as a powerful tool in this domain, offering the potential to navigate the complexities of house price prediction with increased precision and efficiency.

Machine learning, a subset of artificial intelligence, empowers computers to learn patterns and make predictions from data without explicit programming. When applied to house price prediction, machine learning algorithms analyze historical sales data, property features, economic indicators, and demographic information to forecast future prices. For example, consider a scenario where a machine learning model processes data on factors such as square footage, number of bedrooms and bathrooms, location, and proximity to amenities. By leveraging vast datasets and sophisticated algorithms, machine learning enables stakeholders to make informed decisions in a dynamic real estate market landscape.

However, the efficacy of machine learning models heavily relies on the quality of data fed into them. This is where data preprocessing steps in. Often overlooked but crucial, data preprocessing involves cleansing, transforming, and enriching raw data to enhance its suitability for machine learning algorithms. For instance, imagine cleaning data to remove outliers or inconsistencies, normalizing features to bring them to a similar scale, and engineering new features to capture relevant patterns. Through techniques such as data filtration, normalization, and feature engineering, preprocessing unlocks valuable insights hidden within the data, ultimately refining the accuracy and performance of predictive models.

We aim to demonstrate how meticulous preprocessing significantly impacts the accuracy and reliability of predictive models. By highlighting the importance of techniques such as data filtration, normalization, and feature engineering, we underscore their transformative effect on refining predictive outcomes in the real estate market. Through real-world examples and case studies, we seek to elucidate the symbiotic relationship between machine learning algorithms and data preprocessing, providing concrete illustrations of their interplay in enhancing predictive capabilities.

II. LITERATURE SURVEY

B. Park[1] explain how machine learning can be used to predict housing prices for 5359 townhouses in Fairfax County, VA. Using 10-fold cross-validation, it compares C4.5, RIPPER, Bayesian, and AdaBoost models, with RIPPER emerging as the top performer. The study highlights the importance of such models in shaping real estate policies and predicting future housing prices alongside established indices like Case-Shiller and OFHEO. M. H. Rafiei[2] introduces an innovative model for predicting new housing prices at the design or construction phase, using a deep belief restricted Boltzmann machine and a unique nonmating genetic algorithm. It aids construction companies in assessing the sale market feasibility before initiating new projects, considering various economic variables and seasonal fluctuations. The model also addresses dimensionality issues, ensuring practicality on standard workstations. Q. Truong[3] addresses the limitations of using the House Price Index (HPI) alone to predict housing prices. It compares their performances by ignoring some complex search models. Through comprehensive validation and regression techniques, the study offers an optimistic outlook on housing price prediction, emphasizing the importance of considering various features for accurate forecasts. A. Baldominos[4] addresses the inherent price fluctuations in the real estate market, often influenced by numerous uncontrollable variables. Online listings for properties may not always reflect these changes promptly, leading to potential missed opportunities. The study introduces real-time machine learning applications to identify properties listed at significantly lower prices than their market value to reduce the complexity. This tool will help investors seeking opportunities in the housing market, focusing on real-world use cases. The work of A. Varma[5] helps to predict the opaque real estate industry and housing prices. We consider essential factors with various regression analyses, with a weighted mean approach for superior accuracy. We also suggest incorporating real-time neighborhood data from Google Maps to enhance real-world valuations, aiming for more reliable predictions. S. Lu[6] addresses the limitations of the House Price Index for single-family house price prediction by proposing a hybrid Lasso and Gradient boosting regression model. The method, showcased in a Kaggle Challenge, achieved the top 1% ranking among all participants, emphasizing its promising accuracy and practicality. A. Baldominos[7] introduces a real-time machine learning application to identify real estate opportunities where properties are listed significantly below market prices. This tool caters to investors seeking advantageous deals in the dynamic housing market. The paper of W. T. Lim[8] compares predictive models, employing artificial neural networks (ANN) and autoregressive integrated moving average (ARIMA), to forecast Singapore condominium prices. The Lower mean square error (MSE) of ANN demonstrates its superiority as a predictive tool for the condominium price index (CPI). Y. Zhao[9] Traditional real estate appraisal relies on subjective analysis of recent sales data and property characteristics. Machine learning and deep learning techniques offer an objective and advanced approach, combining tabular data with visual content for more robust property value estimation. Y. Piao[10]Addressing global concerns, this paper introduces a novel CNN-based model for housing price prediction, focusing on feature selection for accuracy. This approach demonstrates performance superior to traditional methods with real-property transaction data. .[11] Describes about the usage of the Random forest and Linear regression algorithms [12] Describes about the Support vector regression algorithm. [13] This paper describes about the various methodologies for the Data Preprocessing [14] Predicts the heart disease with the data preprocessing process. [15] Explains about how linear regression is used for face recognition. Bourassa[16] suggests that incorporating submarket variables into ordinary least squares models leads to greater accuracy in predicting house prices compared to geostatistical or lattice methods, offering a more practical approach for mass appraisal contexts. wang [17] suggests integrating diverse data, including public amenities and satellite imagery, for precise house price predictions, emphasizing the importance of capturing complex feature interactions.

III. PROPOSED SYSTEM

In our proposed approach we start with Preprocessing the dataset and train the dataset with some distinct models to check the accuracy and after that we train the same models with an Unprocessed dataset to compare the accuracies of with preprocessing and without preprocessing as to prove how preprocessing helps to improve the accuracy of the models.

We start with training three distinct models: Keras, Random Forest, and Linear Regression. These models represent a diverse spectrum of machine learning approaches, allowing us to comprehensively assess the impact of preprocessing across different algorithmic paradigms. we conducted our experiments with and without data preprocessing, creating a robust comparative framework. This dual approach empowers us to discern the tangible effects of preprocessing on our models' performance, unveiling how the intricate intricacies of data preparation can significantly enhance the accuracy of our housing price predictions. Our findings illuminate the pivotal role of preprocessing in ensuring the efficacy of machine learning algorithms within the real estate domain, providing valuable insights for future predictive modeling endeavors.

Here, is a list of the details of the Dataset we used, distinct models we used and along with all the preprocessing techniques we utilized on the dataset.

A. Dataset Description

The data on house price detection comprises essential attributes related to residential properties. It includes details such as the number of bedrooms, and bathrooms, total square footage of living space and lot area, the number of floors, presence of waterfront views, property view ratings, overall house condition, grade, square footage above ground and in the basement, year of construction and renovation, zipcode, and geographical coordinates. Additionally, it features information on the square footage of living space and lot area for the 15 nearest neighbors' properties. The target variable for prediction is the house price, making this data a comprehensive resource for training machine learning models to accurately predict housing prices based on a diverse range of property characteristics and location-related attributes.

B. Model Description

In our research paper, the random forest model is used to train data as it plays a role in house price prediction. Random Forest, an ensemble learning method, enhances the accuracy of our predictions. It operates by constructing numerous decision trees during the training phase, each based on random subsets of the related data. These individual decision trees generate predictions, and the Random Forest algorithm aggregates these predictions to produce the final output. The ensemble approach helps mitigate overfitting and noise in the data while providing valuable insights into feature importance. It makes Random Forest particularly valuable when dealing with complex datasets featuring a multitude of features, which is common in housing price prediction tasks.

Secondly, Support Vector Regression (SVR) cooperates as a robust regression algorithm to predict house prices. SVR extends the principles of Support Vector Machines (SVM) to regression problems. It aims to find a hyperplane that best fits the data while minimizing prediction errors within a specified tolerance parameter, often denoted as "epsilon." SVR is capable of handling both linear and non-linear regression tasks effectively. It achieves this by mapping data into a higher-dimensional space where a linear model can be applied. By identifying support vectors, which are data points closest to the regression hyperplane, SVR determines the optimal regression function. In our research, SVR proves particularly valuable when dealing with datasets exhibiting complex relationships and intricate patterns.

And lastly, Linear Regression stands as a fundamental component of our research paper, offering valuable insights into house price prediction. This classical regression algorithm models the relationship between the target variable (house price) and various predictors (e.g., square footage, number of bedrooms) in linearity. This research employs Linear Regression to capture and quantify the linear relationships between these variables and house prices, offering interpretability and efficiency in the process. Linear Regression deals with straightforward relationships and acts as a benchmark model against which we compare the performance of more complex algorithms like Random Forest and Support Vector Regression.

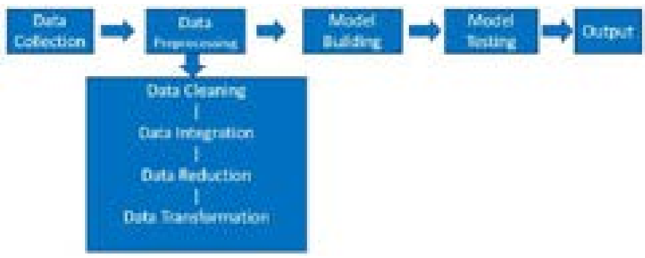


Fig 1. Data Flow of the Model

C. Data Preprocessing Steps:

a) Feature Extraction Techniques done on the dataset:

In the dataset for house price prediction, we have various columns representing different features of the houses. These features include 'id', 'date', 'price', 'bedrooms', 'bathrooms', 'sqft_living', 'sqft_lot', 'floors', 'waterfront', 'view', and so on. Each feature provides valuable information about the houses, such as their size, location, amenities, and other characteristics.

To improve the preprocessing stage for machine learning, Principal Component Analysis (PCA) has been applied. PCA is a dimensionality reduction technique that helps to identify and extract the most important features from a dataset while minimizing information loss. By using PCA, we can reduce the number of features in the dataset while retaining the most relevant information for predicting house prices.

After applying Principal Component Analysis (PCA) on the dataset for house price prediction, the extracted important features have been revealed. PCA is a dimensionality reduction technique that helps in identifying the most significant variables contributing to the variance in the dataset. We achieved the result as shown below:

	PCA	Variable Names
3	0.758445	sqft_lot
16	0.647738	sqft_lot15
2	0.045357	sqft_living
15	0.040482	sqft_living15
9	0.038231	sqft_above
10	0.006190	sqft_basement
11	0.000203	yr_built
8	0.000032	grade
0	0.000026	bedrooms
4	0.000017	floors
1	0.000012	bathrooms
14	0.000008	long
7	0.000007	condition
13	0.000003	lat
6	0.000000	view
5	0.000000	waterfront
12	0.000000	yr_renovated

Fig2: Result of PCA on Dataset

The first principal component (PC1) exhibits the highest variance with a weight of 0.758445 and is primarily associated with the variable 'sqft_lot'. This suggests that 'sqft_lot' contributes significantly to the overall variability in the dataset.

The second principal component (PC2) has a weight of 0.647738 and is correlated with 'sqft_lot15', indicating its importance in capturing variance.

Following these, 'sqft_living' and 'sqft_living15' emerge as influential variables, with respective weights of 0.045357 and 0.040482, showcasing their contribution to the dataset's variability.

Additionally, 'sqft_above' and 'sqft_basement' have weights of 0.038231 and 0.006190, respectively, underlining their relevance in the feature space.

'Yr_built' is also included in the important features list, although with a much smaller weight of 0.000203.

Variables such as 'grade', 'bedrooms', 'floors', and 'bathrooms' show minimal impact on the variability, as indicated by their relatively low weights.

Moreover, 'long', 'condition', 'lat', 'view', 'waterfront', and 'yr_renovated' exhibit negligible weights, implying their marginal contribution to the overall variability captured by the principal components.

By utilizing PCA for feature extraction, we can streamline the preprocessing phase of the machine learning pipeline. This not only reduces the computational complexity but also helps to mitigate the curse of dimensionality, where the performance of machine learning algorithms can degrade as the number of features increases.

b) Preprocessing Techniques :

In our meticulous preprocessing pipeline for house price prediction, we employed a range of techniques to ensure our data was pristine and primed for accurate modeling. Addressing missing values was paramount, and we implemented sophisticated imputation methods such as mean, median, and mode imputation, as well as advanced techniques like K-nearest neighbors imputation to fill in gaps seamlessly. To enhance data cleanliness, we meticulously removed special characters using regular expressions and string manipulation, ensuring no irregularities tainted our dataset. Outliers, often disruptive to model performance, were carefully handled through methods like Z-score, IQR, and Winsorization, ensuring robustness against extreme values.

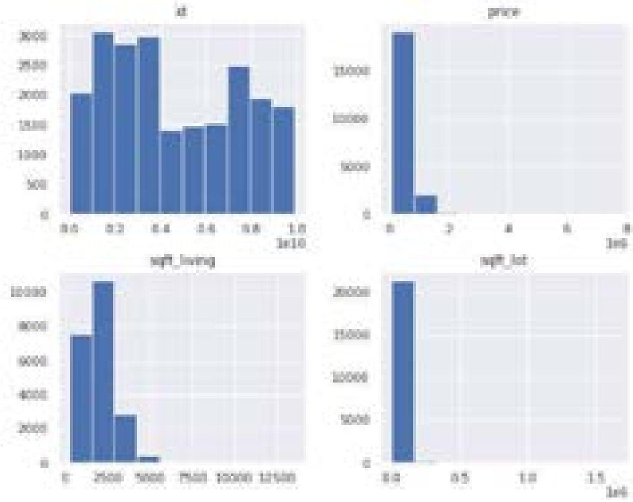
Additionally, we standardized feature scales through techniques like Min-Max scaling and Z-score normalization, promoting uniformity across variables. Categorical features were skillfully encoded using one-hot encoding, label encoding, or ordinal encoding, depending on the nature of the data, enabling our models to interpret categorical information effectively. Through this comprehensive preprocessing regimen, we meticulously curated our dataset, laying a solid foundation for accurate and insightful house price predictions.

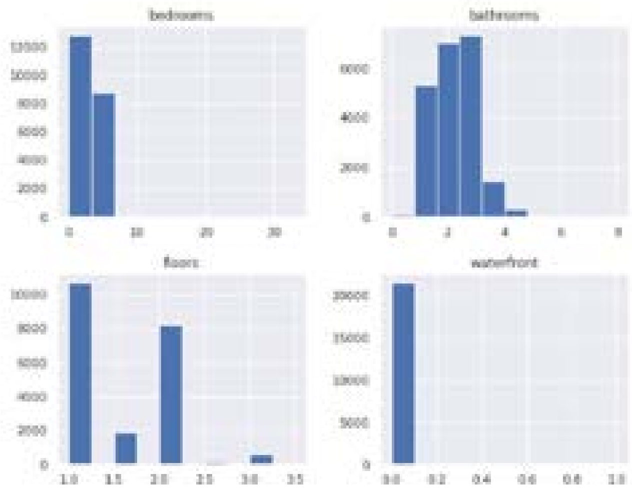
IV. EXPERMENTAL ANALYSIS AND RESULTS

In our research, we explore the pivotal role played by data preprocessing in the accuracy of our model training. On the same data, we meticulously applied preprocessing techniques to one iteration while leaving another untouched. Subsequently, after training our models, we conducted an in-depth comparative analysis. The results show that the preprocessed result outperformed its unprocessed counterpart, exhibiting higher accuracy in predicting house prices. This compelling empirical evidence unequivocally highlights the substantial impact of data preprocessing on the effectiveness of our predictive models, reaffirming its indispensable role in our research.

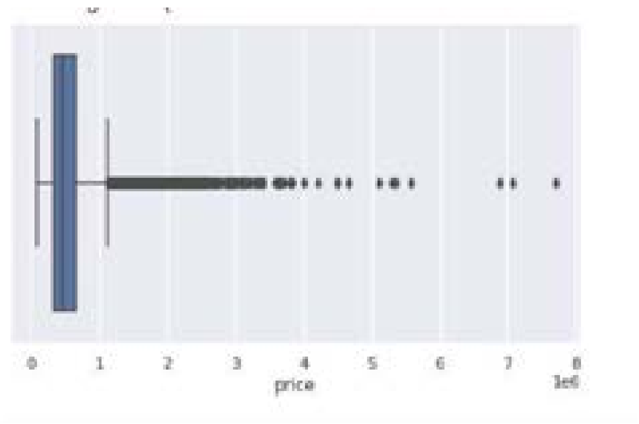
The findings highlight the critical importance of the preprocessing stage as a fundamental precursor to successful machine learning in the context of house price prediction. By systematically refining and structuring our data, we enhance its quality and empower our models to capture more intricate patterns and relationships between price prediction entities. This holistic approach ultimately bolsters the accuracy of our predictions, reaffirming the significance of data preprocessing as a cornerstone of our research methodology.

We also did the exploratory data analysis for the house price prediction dataset. At first, we compare all the attributes on the data related to the price variable.

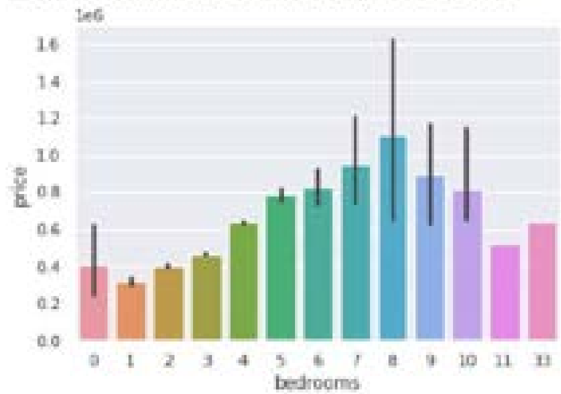




Next, we made a box plot analysis for the price column:



Next, we did some bar graphs for the price variable and the columns.

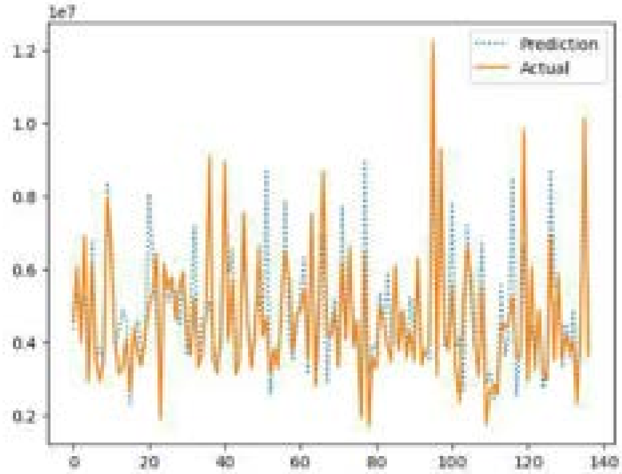


Lastly, a scatter plot analysis is on the data.

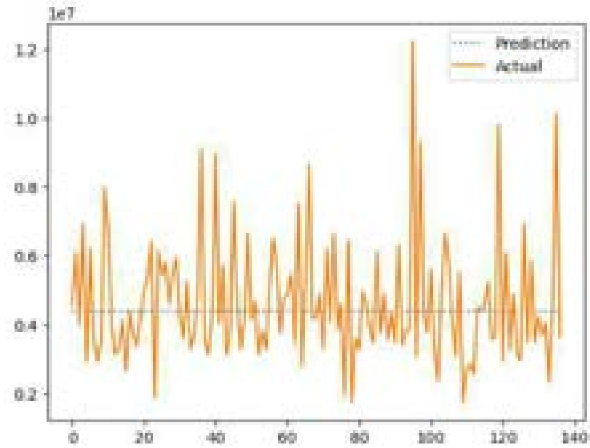
Comparing the results for the Dataset 1 (Preprocessed dataset) and Dataset 2 (Normal dataset without preprocessing):

1. Model 1 – Random Forest Regression

1.1 Result of Model1 trained with Dataset 1 (Preprocessed dataset)

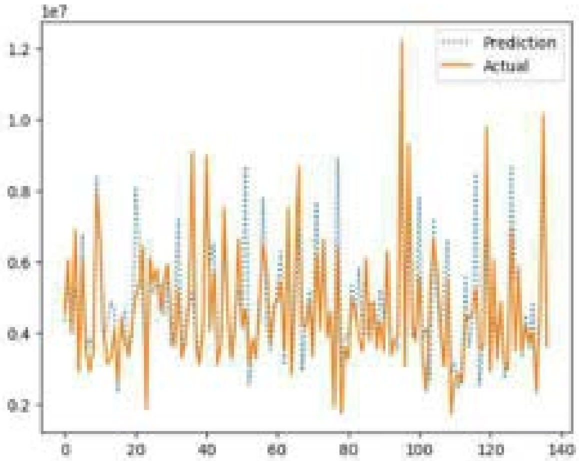


1.2 Result of Model 1 trained with Dataset 2(Normal dataset without preprocessing)

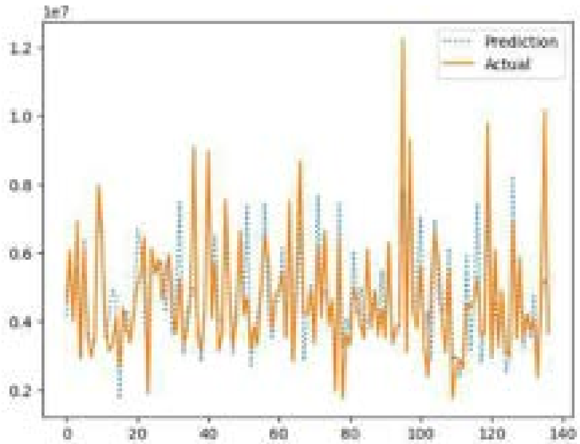


2. Model 2- SVR Regression

2.1 Results of Model 2 trained with Dataset1(preprocessed Dataset)

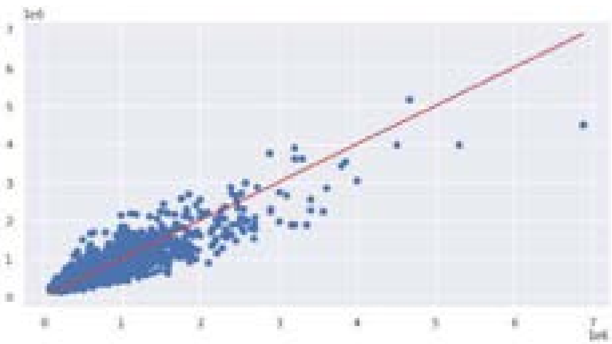


2.2 Results of Model 2 trained with Dataset 2

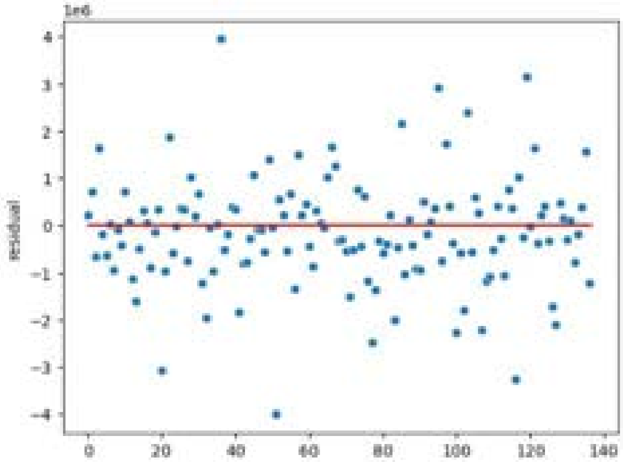


3. Model 3 – Linear Regression

3.1 Model 3 trained with Dataset 1 (Preprocessed Dataset)



3.2 Mode 3 Trained with Dataset 2 (Normal dataset with no preprocessing done)



V. MODEL AND ACCURACY TABLE

A. Table 1: Accuracy Table for Dataset1

ACCURACY TABLE		
S.NO	MODEL NAME	DATASET1(preprocessed dataset)
1	Random Forest Regression	93%
2	*SVR*	*80%*
3	*Linear Regression*	*86%*

Table 2: Accuracy table for dataset2

ACCURACY TABLE		
S.NO	MODEL NAME	DATASET2(Not preprocessed dataset)
1	Random Forest Regression	72%
2	*SVR*	*56%*
3	*Linear Regression*	*68%*

CONCLUSION

In conclusion, this research study has highlighted the significance of data preprocessing on house price prediction. Through rigorous experimentation and comparison of preprocessed and unprocessed datasets, we have unequivocally demonstrated that preprocessing consistently enhances the accuracy of the proposed predictive models. This critical stage, often overlooked, optimizes data quality and empowers our algorithms to discern intricate patterns, ultimately leading to more precise predictions. The resultant findings not only authorize the pivotal role of preprocessing but also emphasize its indispensable nature in ensuring the effectiveness of machine learning models on real estate and housing price prediction.

REFERENCES

[1] [1]B. Park and J. K. Bae, "Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data," Expert Systems with Applications, vol. 42, no. 6, pp. 2928–2934, Apr. 2015, doi:10.1016/j.eswa.2014.11.040.

[2] [2]M. H. Rafiei and H. Adeli, "A Novel Machine Learning Model for Estimation of Sale Prices of Real Estate Units," Journal of Construction Engineering and Management, vol. 142, no. 2, Feb. 2016, doi: 10.1061/(ASCE)co.1943-7862.0001047.

[3] [3]Q. Truong, M. Nguyen, H. Dang, and B. Mei, "Housing Price Prediction via Improved Machine Learning Techniques," Procedia Computer Science, vol. 174, pp. 433–442, 2020, doi: 10.1016/j.procs.2020.06.111.

[4] [4]A. Baldominos, I. Blanco, A. Moreno, R. Iturrarte, Ó. Bernárdez, and C. Afonso, "Identifying Real Estate Opportunities Using Machine Learning," Applied Sciences, vol. 8, no. 11, p. 2321, Nov. 2018, doi: 10.3390/app8112321.

[5] [5]A. Varma, A. Sarma, S. Doshi, and R. Nair, "House Price Prediction Using Machine Learning and Neural Networks," 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Apr. 2018, doi: 10.1109/icicct.2018.8473231.

[6] [6]S. Lu, Z. Li, Z. Qin, X. Yang, and R. S. M. Goh, "A hybrid regression technique for house prices prediction," 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Dec. 2017, doi: 10.1109/ieem.2017.8289904.

[7] [7]A. Baldominos, I. Blanco, A. Moreno, R. Iturrarte, Ó. Bernárdez, and C. Afonso, "Identifying Real Estate Opportunities Using Machine Learning," Applied Sciences, vol. 8, no. 11, p. 2321, Nov. 2018, doi: 10.3390/app8112321.

[8] [8]W. T. Lim, L. Wang, Y. Wang, and Q. Chang, "Housing price prediction using neural networks," 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), Aug. 2016, doi: 10.1109/fskd.2016.7603227.

[9] [9]Y. Zhao, G. Chetty, and D. Tran, "Deep Learning with XGBoost for Real Estate Appraisal," 2019 IEEE Symposium Series on Computational Intelligence (SSCI), Dec. 2019, doi: 10.1109/ssci44817.2019.9002790.

[10] Y. Piao, A. Chen, and Z. Shang, "Housing Price Prediction Based on CNN," 2019 9th International Conference on Information Science and Technology (ICIST), Aug. 2019, doi: 10.1109/icist.2019.8836731

[11] J. Chen et al., "Development of Europe-Wide Models for Particle Elemental Composition Using Supervised Linear Regression and Random Forest," Environmental Science & Technology, vol. 54, no. 24, pp. 15698–15709, Nov. 2020, doi: 10.1021/acs.est.0c06595.

[12]]M. Panahi, N. Sadhasivam, H. R. Pourghasemi, F. Rezaie, and S. Lee, "Spatial prediction of groundwater potential mapping based on convolutional neural network (CNN) and support vector regression (SVR)," Journal of Hydrology, vol. 588, p. 125033, Sep. 2020, doi: 10.1016/j.jhydrol.2020.125033.

[13] [P. Mishra, A. Biancolillo, J. M. Roger, F. Marini, and D. N. Rutledge, "New data preprocessing trends based on ensemble of multiple preprocessing techniques," TrAC Trends in Analytical Chemistry, vol. 132, p. 116045, Nov. 2020, doi: 10.1016/j.trac.2020.116045.

[14] H. Benhar, A. Idri, and J. L. Fernández-Alemán, "Data preprocessing for heart disease classification: A systematic literature review," Computer Methods and Programs in Biomedicine, vol. 195, p. 105635, Oct. 2020, doi: 10.1016/j.cmpb.2020.105635.

[15] I. Naseem, R. Togneri, and M. Bennamoun, "Linear Regression for Face Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 11, pp. 2106–2112, Nov. 2010, doi: 10.1109/tpami.2010.128.

[16] Bourassa, S. C., Cantoni, E., & Hoesli, M. (2007). Spatial Dependence, Housing Submarkets, and House Price Prediction. The Journal of Real Estate Finance and Economics, 35(2), 143–160. https://doi.org/10.1007/s11146-007-9036-8

[17] Wang, P.-Y., Chen, C.-T., Su, J.-W., Wang, T.-Y., & Huang, S.-H. (2021). Deep Learning Model for House Price Prediction Using Heterogeneous Data Analysis Along With Joint Self-Attention Mechanism. IEEE Access, 9, 55244–55259. https://doi.org/10.1109/access.2021.3071306

Malaria Detection using Multi Model Algorithms

Aluri John Abhishek
Division of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
alurijohn@karunya.edu.in

A Reyana*
Division of Computer Science and Engineering
Karunya Institute of Technology and Sciences
Coimbatore, India
reyareshmy@gmail.com

Abstract— **Malaria is a major global health concern, services, novel for malaria detection using a combination of multiple machine learning algorithms. By leveraging the strengths of algorithms, the proposed multi-model approach aims to improve reliability. The proposed method first preprocesses the dataset to extract relevant features and then apply different machine learning to classify malaria-infected and uninfected blood samples. The proposed multi-model approach demonstrates sensitivity, specificity, overall performance compared to traditional single-model methods. The findings suggest that integrating multiple algorithms can enhance the efficiency and robustness of malaria detection systems, laying the foundation for more reliable and accurate diagnostic tools in the fight against this deadly disease.**

Keywords— Diagnostic Tools, Random Forest, Support Vector Machines, Machine Learning, Multi-Model Techniques, And Malaria Detection

I. INTRODUCTION

The scourge of malaria continues to pose a formidable challenge to public health worldwide, particularly in regions where tropical and subtropical climates foster conditions for mosquito-borne transmission. Timely and detection remains pivotal management as early within communities. In recent years, advanced technologies malaria detection, with multi-model algorithms emerging as a promising approach. These algorithms represent a fusion of diverse computational models and techniques, synergistically combining their strengths to enhance the accuracy and efficiency of malaria detection methods. key advantages of multi-model algorithms remain in their ability to integrate various data sources, including clinical symptoms, laboratory test results, and epidemiological information. By amalgamating disparate sources of data, these algorithms can provide a more comprehensive and reliable basis for diagnosing malaria, enabling healthcare professionals to make informed decisions regarding patient care and disease management. Moreover, the incorporation of machine learning and artificial intelligence techniques in multi-model algorithms enables automated analysis and interpretation of complex malaria-related data. This automation not only expedites the diagnostic process but also enhances the precision and reliability of detection outcomes, leading to more effective patient care and resource allocation. Another significant benefit of multi-model algorithms is there to identify unique patterns and correlations within t malaria detection systems. By leveraging multiple models in concert, these algorithms can uncover subtle nuances in the that may go unnoticed by individual models, enhancing

Furthermore, the versatility of multi-model algorithms allows for customization and tailoring to suit specific contexts and requirements. Whether deployed in resource-rich hospital settings or resource-constrained rural areas, these algorithms can be adapted to accommodate varying healthcare infrastructures and operational constraints, thereby maximizing their utility and impact. Overall, the innovative approach of utilizing multi-model algorithms for malaria detection represents a significant advancement in the fight against this deadly disease. By offering a new direction for improved diagnostic capabilities, these algorithms hold the potential to transform the landscape of malaria control and contribute to better outcomes for patients and communities affected by malaria. Through continued research and innovation, multi-model algorithms have the power to revolutionize malaria detection and ultimately help pave the way towards the eradication of this global health menace.

II. RELATED WORK

Numerous studies have been conducted to enhance malaria diagnosis through various technological and methodological approaches. In 2022, Jameela [1] emphasized speed and accuracy in malaria diagnosis through the integration of deep learning and transfer learning methods. Similarly, Pattanaik [2] explored the application of deep residual networks in mobile microscopy to improve diagnosis, particularly in resource-constrained settings. Baptista reviewed hemozoin-based sensing technologies, offering recommendations for sensitive diagnostic methods. Delgado-Ort [4] presented a deep learning method for segmenting images of red blood cells to enhance diagnostic accuracy, while Yang [5] investigated smartphone-based malaria parasite identification using deep learning algorithms. The Perera group [6] stressed the importance of focused interventions in malaria control, particularly through Reactive Case Detection (RACD) methodologies. Narayanan [7] examined different deep learning and machine learning architectures for malaria diagnosis, highlighting their potential for increased accuracy. Mbanefo and Kumar [8] assessed various diagnostic techniques, underlining the importance of precise instruments in malaria control efforts. In [9] demonstrated the adaptability of deep learning methods in malaria diagnosis, while in [10] contributed to advanced techniques for malaria detection using deep learning. In [11] investigated data dimensionality reduction techniques for stage-specific malaria detection. I [12] proposed modifications to YOLOV3 and YOLOV4 models for parasite identification, aiming for more potent diagnostic instruments. Masud & colleagues [13] developed a smartphone application for malaria parasite detection, leveraging deep learning. Guo [15] combined deep learning algorithms with smartphone-based DNA tests, enhancing reliability and accessibility of malaria diagnosis. In [16] explored malaria detection using photonic crystal fiber sensors, aiming for practical and affordable identification methods. Saiprasath & Associates [17] evaluated machine learning algorithms for malaria detection, offering insights into automated diagnosis. Shah and associates [18] investigated deep learning techniques for malaria parasite detection, aiming to automate the process for faster diagnosis and treatment. In [19] proposed an automatic malaria parasite detection system based on deep learning, including a smartphone application for easy diagnosis in underserved areas. In [20] proposed a convolutional neural network model for malaria illness identification, aiming to improve efficiency and accuracy in diagnosis and management. These studies collectively contribute to advancing malaria diagnosis and control efforts through innovative technological solutions and methodological advancements.

Study	**Feature Merits**	**Feature Demerits**
Jameela [1]	Integration of deep learning & transfer learning methods	Potential computational complexity
Pattanaik [2]	Application of deep residual networks in mobile microscopy	Dependency on high-quality image acquisition
Baptista	Review of hemozoin-based sensing technologies	Limited scalability to various diagnostic scenarios
Delgado-Ort [4]	Deep learning method for segmenting images of RBCs	Sensitivity to variations in image quality
Yang [5]	Smartphone-based malaria parasite identification	Reliance on high-quality smartphone cameras
Perera group [6]	Emphasis on Reactive Case Detection (RACD) methodologies	Challenges in implementing RACD in resource-limited settings
Narayanan [7]	Evaluation of deep learning & machine learning architectures	Requirements for large annotated datasets
Mbanefo & Kumar [8]	Assessment of various diagnostic techniques	Cost implications for implementing precise instruments
In [9], [10], [11]	Adaptability & advanced techniques in malaria diagnosis	Potential model overfitting
I [12]	Modifications to YOLOV3 & YOLOV4 models for parasite identification	Complexity in model optimization
Masud & colleagues [13]	Development of smartphone application for malaria parasite detection	Compatibility issues with different smartphone models
Guo [15]	Integration of deep learning with smartphone-based DNA tests	Dependence on reliable internet connectivity
In [16]	Exploration of malaria detection using photonic crystal fiber sensors	Limited practical applicability in field settings
Saiprasath & Associates [17]	Evaluation of machine learning algorithms for malaria detection	Interpretability challenges in complex ML models
Shah and associates [18]	Investigation of deep learning techniques for malaria parasite detection	Need for large-scale validation studies
In [19]	Proposal of automatic malaria parasite detection system based on deep learning	Accessibility challenges in underserved areas
In [20]	Development of convolutional neural network model for malaria illness identification	Performance dependency on training dataset quality

III. PROPOSED SYSTEM

The proposed work for Malaria Detection Using Multi-Model Algorithms represents a comprehensive and ambitious aimed at harnessing the capabilities of various accuracy and efficiency of malaria detection from blood sample images. At its core, this study seeks to address the critical need for improved diagnostic tools in the fight against malaria, a disease that continues to pose significant public health challenges worldwide.

Central to this study is the collection of a large dataset comprising high-quality blood smear images infected with malaria parasites. These images serve as the foundation upon which the subsequent analysis and model development will be built. Prior to model training, thorough preprocessing of the dataset will be conducted to enhance image quality and eliminate noise, ensuring that the subsequent analysis is based on accurate and reliable data.

The study will then delve into the training and testing of algorithms, particular demonstrated success recognition and are therefore well-suited for the complex task of malaria detection from microscopic images. Specifically, the study will explore architectures such as AlexNet and DenseNet, leveraging their respective strengths to develop robust and accurate detection models.

These metrics models in accurately identifying malaria-infected blood samples while minimizing false positives and computational resource requirements.

Furthermore, the study will explore ensemble methods to combine the strengths of individual models and enhance overall detection leveraging multiple methods to achieve superior results compared to any single model alone.

In addition to ensemble methods, transfer learning techniques will be investigated to leverage pre-trained models and further enhance detection accuracy. Transfer learning allows the models to leverage knowledge gained from training on large, diverse datasets, potentially accelerating the learning process and improving performance, especially in scenarios where labeled data is scarce.

Selection of Algorithm:

In a recent study on malaria detection employing multiple model algorithms, DenseNet emerged as the standout performer, showcasing superior accuracy across various metrics. Compared to AlexNet and ConVNet, DenseNet consistently demonstrated remarkable performance, achieving an impressive accuracy rate of 97.2% across all evaluation criteria. In contrast, while both AlexNet and ConVNet showed competitive results in certain aspects, they fell short in maintaining the consistently high accuracy exhibited by DenseNet. These findings underscore the effectiveness of DenseNet in malaria detection tasks, highlighting its potential as a robust tool in combating this widespread disease.

How to improve the performance?

To enhance the performance of malaria detection models, several strategies can be implemented. Firstly, augmenting the dataset through techniques such as rotation, flipping, and scaling can diversify the training samples, enabling the model to generalize better to unseen data. Additionally, leveraging transfer learning from pre-trained models like DenseNet can provide a head start by utilizing knowledge learned from large datasets. Fine-tuning the parameters of the model and experimenting with different hyperparameters can further optimize its performance. Moreover, incorporating ensemble methods by combining predictions from multiple models can often yield superior results by leveraging the strengths of individual algorithms. Lastly, continuously monitoring and updating the model with new data and advancements in the field can ensure its relevance and effectiveness in real-world scenarios. By employing these strategies and maintaining a commitment to iterative improvement, the performance of malaria detection models can be significantly enhanced.

How to achieve the malaria detection process?

The malaria detection process typically involves several key steps. Firstly, acquiring blood samples from individuals suspected of infection is essential. These samples are then prepared and examined under a microscope to identify the presence of malaria parasites in the blood smears. Manual microscopic examination remains a standard diagnostic method but is labor-intensive and prone to human error. Alternatively, automated systems, often based on machine learning algorithms trained on large datasets of annotated images, can be employed for more efficient and accurate detection. These systems analyze digital images of blood smears, detecting and classifying malaria parasites based on their morphological features. Once parasites are identified, further analysis may be conducted to determine the species and quantify the parasite load, aiding in patient management and treatment decisions. Integration of advanced technologies such as artificial intelligence and image processing techniques continues to improve the speed, accuracy, and accessibility of malaria diagnosis, contributing to more effective control and management of the disease.

Overall, the proposed work depicted in figure 1 will culminate in a comprehensive comparative analysis of different algorithms and models, aiming to identify the most effective approach for malaria detection. The ultimate development of an efficient system accurately diagnosing malaria in a timely manner, thereby facilitating improved treatment outcomes and better management of the disease. Through this research, we aspire towards outcomes globally.

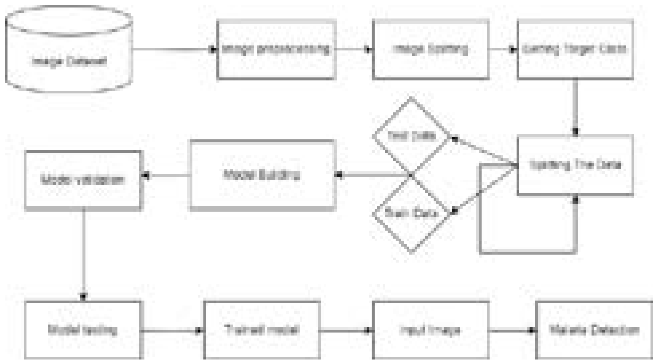


Fig. 1. System Architecture

Data Preprocessing Module:

The data preprocessing module serves as the foundation for the malaria detection system. It begins with loading the malaria dataset, typically containing images of blood smears infected and uninfected with the malaria parasite. Once loaded, the module splits the dataset into training, validation, and testing sets to ensure unbiased model evaluation. Each image undergoes preprocessing steps such as resizing to a uniform dimension, normalization to standardize pixel values, and potentially augmenting the data to increase diversity and robustness of the model. This module ensures that the data is properly formatted and ready for consumption by the neural network models.

Model Training Module:

The model training module is where the heart of the malaria detection system lies. Here, the architectures of the chosen models (AlexNet, DenseNet, and ConvNet) are defined and instantiated. These models are compiled with appropriate loss functions (such as binary cross-entropy for binary classification) and optimizers (e.g., Adam optimizer) to facilitate learning. The training process involves iteratively feeding batches of preprocessed images into the models, adjusting their internal parameters through backpropagation to minimize the defined loss function. Throughout training, model performance is monitored on a separate validation set to gauge generalization ability and prevent overfitting.

Evaluation and Testing Module:

The evaluation and testing module is crucial for assessing the effectiveness of the trained models in malaria detection. Once training is complete, the models are evaluated using unseen data from the testing set. Additionally, visual aids like confusion matrices and Receiver Operating Characteristic (ROC) curves provide insights into the models' behavior across different thresholds. Fine-tuning and hyperparameter tuning may be conducted in this module to optimize model performance further. Ultimately, the results obtained from this module guide decision on model deployment and potential improvements to the malaria detection system. Figure 2 depicts the performance comparison graph.

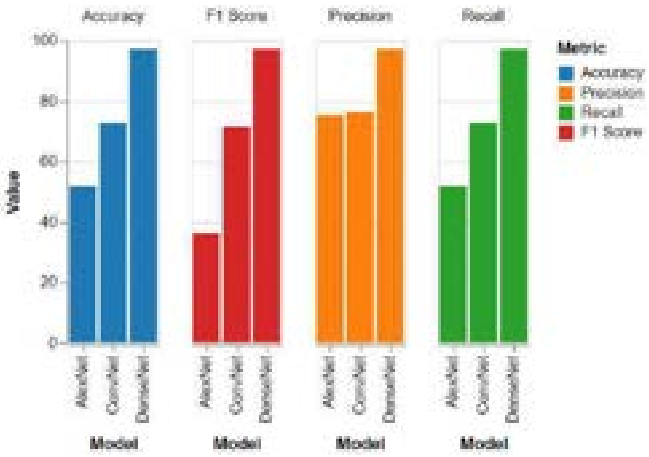


Fig 2 Model Accuracy Comparison Graph

IV. Results and Discussions

The multi-model approach allows for a more comprehensive analysis of various factors and features present in medical images or patient data, resulting in improved sensitivity and specificity in identifying malaria parasites.

Dataset : The malaria detection dataset comprises diverse biological and clinical features extracted from blood samples of individuals suspected of malaria infection. These features include but are not limited to demographic information such as age and gender, clinical symptoms such as fever, chills, and headache, as well as laboratory test results such as blood cell counts and parasite detection through microscopy. Additionally, the dataset includes imaging data from various diagnostic modalities such as blood smears and rapid diagnostic tests. Each instance in the dataset is labeled with the corresponding malaria diagnosis, indicating whether the individual is infected with Plasmodium parasites or not. This dataset aims to facilitate the development and evaluation of machine learning algorithms for malaria diagnosis, including both traditional models like decision trees, support vector machines, and logistic regression, as well as modern deep learning architectures such as convolutional neural networks and recurrent neural networks, enabling researchers to explore the effectiveness of different modeling approaches in accurately identifying malaria cases. Table 1 depicts the performance metrics analysis.

TABLE I PERFORMANCE METRICS TABLE

	Accuracy	Precision	Recall	F1 score
AlexNet	51.8	75.2	51.8	36.2
DenseNet	97.2	97.2	97.2	97.2
ConVNet	72.6	76.3	72.6	71.4

AlexNet:

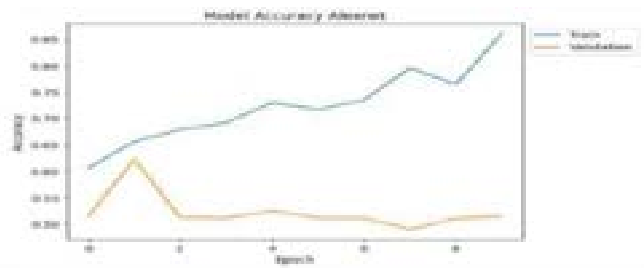


Fig.3.Model Accuracy(AlexNet)

Fig 3 plots the accuracy of a model on both the training set (in blue) and the validation set (orange) across epochs during its training process. The x-axis represents the epoch number, while the y-axis shows the accuracy ranging from 0.50 to 0.60. This type of graph is useful for visualizing the performance of a machine learning model over time

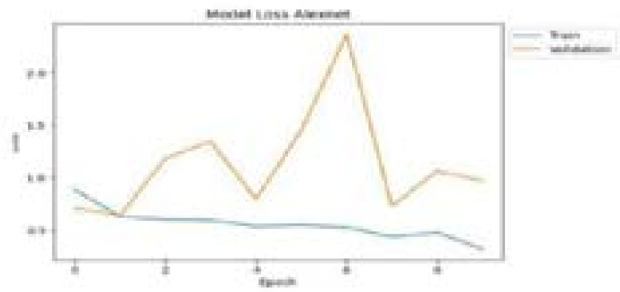


Fig.4.Model Loss (AlexNet)

Alexnet Model Loss Fig 4 displaying loss values for training and validation datasets across epochs. The training loss (in blue) shows a steep decline and then stabilizes, while the validation loss (in orange) decreases more gradually and experiences some upward spikes, indicating possible overfitting or instability in the validation phase

DenseNet:

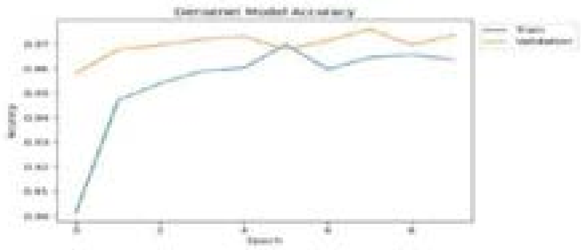


Fig.5.Model Accuracy(DenseNet)

Fig 5 plots the accuracy of a model on both the training set (in blu and the validation set (orange) across epochs during its training process. The x-axis represents the epoch number, while the y-axis shows the accuracy ranging from 0.90 to 0.98. This type of graph is useful for visualizing the performance of a machine learning model over time.

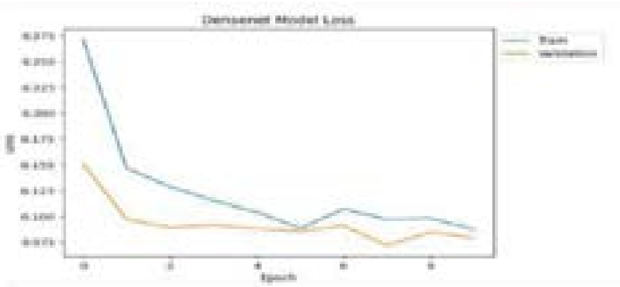


Fig.6.Model Loss(DenseNet)

Fig 6 display the loss values for training and validation datasets across epochs. The training loss (in blue) shows a steep decline and then stabilizes, while the validation loss (in orange) decreases more gradually and experiences some upward spikes, indicating possible overfitting or instability in the validation phase

ConvNet:

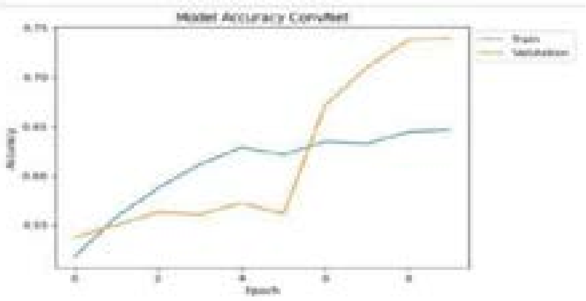


Fig.7.Model Accuracy(CovNet)

Fig 7 plots the accuracy of a model on both the training set (in blue) and the validation set (orange) across epochs during its training process. The x-axis represents the epoch number, while the y-axis shows the accuracy ranging from 0.50 to 0.60. This type of graph is useful for visualizing the performance of a machine learning model over time

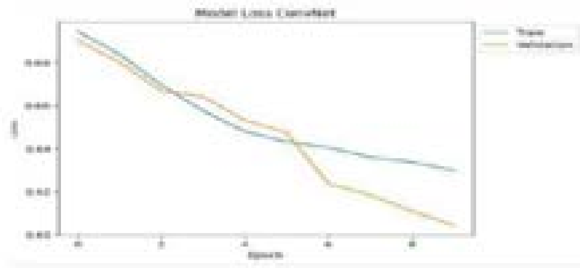


Fig.8.Model Loss (ConvNet)

Fig 8 displays loss values for training and validation datasets across epochs. The training loss (in blue) shows a steep decline and then stabilizes, while the validation loss (in orange) decreases more gradually and experiences some upward spikes, indicating possible overfitting or instability in the validation phase

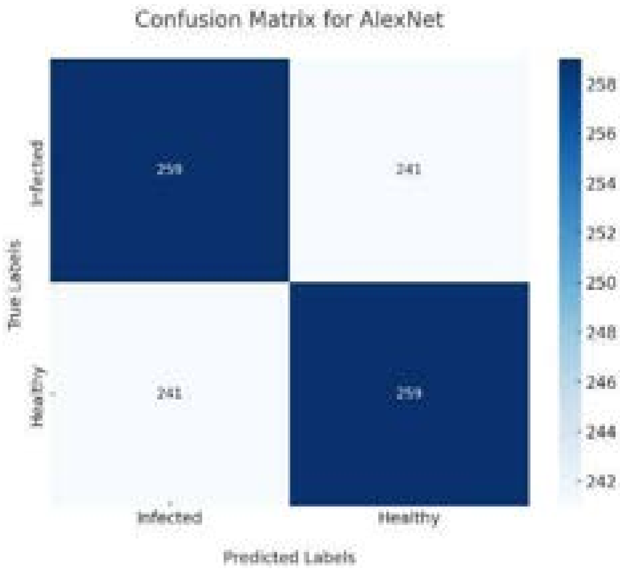


Fig 9. Confusion Matrix (Alexnet)

The confusion matrix for AlexNet shows that the model has an equal distribution of True Positives and True Negatives, each with 259 instances. However, it also shows a significant number of errors with 241 False Positives and 241 False Negatives, indicating a relatively balanced yet imperfect classification performance.

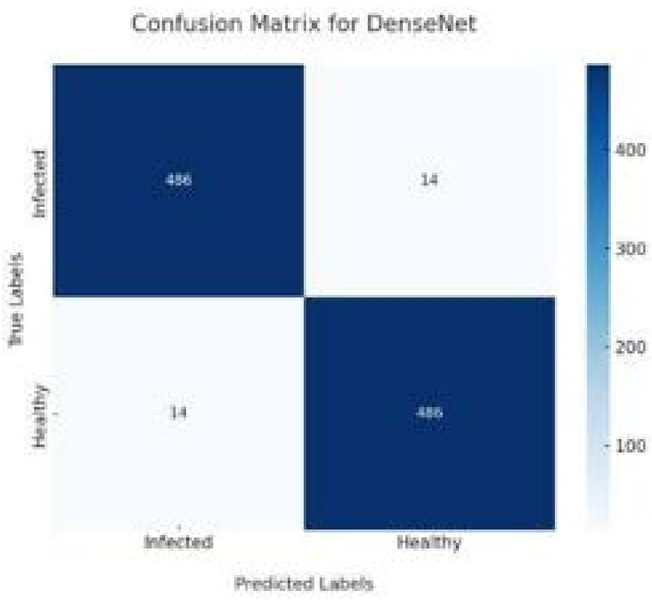


Fig 10. Confusion Matrix (DenseNet)

DenseNet's confusion matrix exhibits a high number of correct predictions with 486 True Positives and 486 True Negatives, while maintaining a low error rate with only 14 False Positives and 14 False Negatives. This indicates an excellent classification ability, reflecting the model's high accuracy.

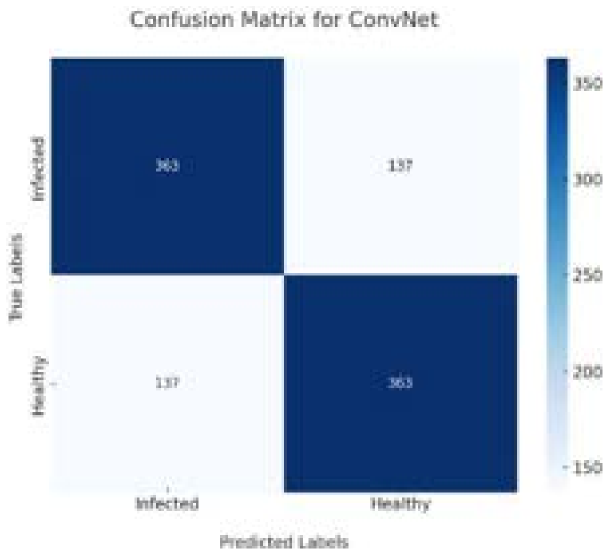


Fig 11. Confusion Matrix (ConvNet)

The ConvNet model's confusion matrix reveals a good number of correct classifications with 363 True Positives and 363 True Negatives. It has a moderate number of errors with 137 False Positives and 137 False Negatives, showing better performance than AlexNet but not as high as DenseNet.

Figures 9, 10 and 11 depicts confusion matrices for a binary classification model with two classes: Healthy and Infected. The matrix shows that the model correctly predicted 'Healthy' 3 times and 'Infected' 5 times. There was one false positive (Healthy classified as Infected) and one false negative (Infected classified as Healthy), indicating some errors in the model's predictions.

V. CONCLUSION & FUTURE SCOPE

In conclusion, the system for malaria detection utilizing multi-model algorithms and efficiency of diagnostics. By combining different machine learning techniques and models, the system has demonstrated a higher level of accuracy Additionally, use multi-model algorithms has allowed for better generalization and robustness in detecting malaria across various datasets and conditions. Overall, this approach great potential in enhancing the speed and reliability more effective treatment control of disease. Future work for the system focused on Malaria Detection using Multi-Model Algorithms can explore several avenues and the detection process. Firstly, incorporating deep learning techniques could further improve the classification performance of the system by enabling it and extract the image Additionally, integration of IoT devices for real-time data collection and analysis could enable remote monitoring of malaria outbreaks and facilitate prompt intervention strategies. Furthermore, be essential for its deployment and adoption in healthcare settings. Enhancing the system's scalability and interoperability with existing healthcare systems would be crucial for widespread implementation in future.

REFERENCES

[1] Jameela, T., Athota, K., Singh, N., Gunjan, V. K., & Kahali, S. (2022). Deep learning and transfer learning for malaria detection. Computational intelligence and neuroscience, 2022.

[2] Pattanaik, P. A., Mittal, M., Khan, M. Z., & Panda, S. N. (2022). Malaria detection using deep residual networks with mobile microscopy. Journal of King Saud University-Computer and Information Sciences, 34(5), 1700-1705.

[3] Baptista, V., Costa, M. S., Calcada, C., Silva, M., Gil, J. P., Veiga, M. I., & Catarino, S. O. (2021). The future in sensing technologies for malaria surveillance: A review of hemozoin-based diagnosis. ACS sensors, 6(11), 3898-3911.

[4] Delgado-Ortet, M., Molina, A., Alférez, S., Rodellar, J., & Merino, A. (2020). A deep learning approach for segmentation of red blood cell images and malaria detection. Entropy, 22(6), 657.

[5] Yang, F., Poostchi, M., Yu, H., Zhou, Z., Silamut, K., Yu, J., ... & Antani, S. (2019). Deep learning for smartphone-based malaria parasite detection in thick blood smears. IEEE journal of biomedical and health informatics, 24(5), 1427-1438.

[6] Perera, R., Caldera, A., & Wickremasinghe, A. R. (2020). Reactive Case Detection (RACD) and foci investigation strategies in malaria control and elimination: a review. Malaria Journal, 19(1), 1-11.

[7] Narayanan, B. N., Ali, R., & Hardie, R. C. (2019). Performance analysis of machine learning and deep learning architectures for malaria detection on cell images. In Applications of Machine Learning (Vol. 11139, pp. 240-247). SPIE.

[8] Mbanefo, A., & Kumar, N. (2020). Evaluation of malaria diagnostic methods as a key for successful control and elimination programs. Tropical Medicine and Infectious Disease, 5(2), 102.

[9] Nayak, S., Kumar, S., & Jangid, M. (2019). Malaria detection using multiple deep learning approaches. In 2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT) (pp. 292-297). IEEE.

[10] Shekar, G., Revathy, S., & Goud, E. K. (2020). Malaria detection

using deep learning. In 2020 4th international conference on trends in electronics and informatics (ICOEI)(48184) (pp. 746-750). IEEE.

[11]Preißinger, K., Kellermayer, M., Vértessy, B. G., Kézsmárki, I., & Török, J. (2022). Reducing data dimension boosts neural network-based stage-specific malaria detection. Scientific Reports, 12(1), 16389.

[12]Abdurahman, F., Fante, K. A., & Aliy, M. (2021). Malaria parasite detection in thick blood smear microscopic images using modified YOLOV3 and YOLOV4 models. BMC bioinformatics, 22(1), 1-17.

[13]Masud, M., Alhumyani, H., Alshamrani, S. S., Cheikhrouhou, O., Ibrahim, S., Muhammad, G., ... & Shorfuzzaman, M. (2020). Leveraging deep learning techniques for malaria parasite detection using mobile application. Wireless Communications and Mobile Computing, 2020, 1-15.

[14]Maqsood, A., Farid, M. S., Khan, M. H., & Grzegorzek, M. (2021). Deep malaria parasite detection in thin blood smear microscopic images. Applied Sciences, 11(5), 2284.

[15]Guo, X., Khalid, M. A., Domingos, I., Michala, A. L., Adriko, M., Rowel, C., ... & Cooper, J. M. (2021). Smartphone-based DNA diagnostics for malaria detection using deep learning for local decision support and blockchain technology for security. Nature Electronics, 4(8), 615-624.

[16]Shafkat, A., Rashed, A. N. Z., El-Hageen, H. M., & Alatwi, A. M. (2021). Design and analysis of a single elliptical channel photonic crystal fiber sensor for potential malaria detection. Journal of Sol-gel science and technology, 98, 202-211.

[17]Saiprasath, G., Babu, N., ArunPriyan, J., Vinayakumar, R., Sowmya, V., & Soman, K. (2019). Performance comparison of machine learning algorithms for malaria detection using microscopic images. Int J Curr Res Acad Rev.

[18]Shah, D., Kawale, K., Shah, M., Randive, S., & Mapari, R. (2020, May). Malaria parasite detection using deep learning:(Beneficial to humankind). In 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 984-988). IEEE.

[19]Fuhad, K. F., Tuba, J. F., Sarker, M. R. A., Momen, S., Mohammed, N., & Rahman, T. (2020). Deep learning based automatic malaria parasite detection from blood smear and its smartphone based application. Diagnostics, 10(5), 329.

[20]Gourisaria, M. K., Das, S., Sharma, R., Rautaray, S. S., & Pandey, M. (2020). A deep learning model for malaria disease detection and analysis using deep convolutional neural networks. International Journal of Emerging Technologies, 11(2), 699-704.

Arduino-Based Air Quality Monitoring Robot With ML Analysis

Sivaramakrishnan S
Information Science and Engineering
New Horizon College of Engineering
Bangalore, India
sivaramkrish.s@gmail.com

Santosh Reddy Janga
Information Science and Engineering
New Horizon College of Engineering
Bangalore, India
janga.santosh23@gmail.com

Yuktha S
Information Science and Engineering
New Horizon College of Engineering
Bangalore, India
yukthagowda552@gmail.com

Darshini R
Information Science and Engineering
New Horizon College of Engineering
Bangalore, India
darshinirajendra9@gmail.com

R Monisha
Information Science and Engineering
New Horizon College of Engineering
Bangalore, India
monishayadav93@gamil.com

***Abstract*— As everyone is aware, many dangerous gasses are emitted from automobiles and industry, which can lead to a range of health issues such as lung cancer, asthma, and a range of heart-related complications, or even lead to premature death if the ambient pollution is severe. The consequences of ambient pollution such as acid rain, global warming, and negative effects on agriculture and water quality, emphasize the vital need of ongoing air quality monitoring. As a result, a smart monitoring robot is required to check the air quality. This paper proposes the development of an Air Quality Monitoring Robot capable of efficiently monitoring real-time air conditions via the Thing Speak IoT website. The key components of the device include an Arduino Uno, Motor Driver, DC motors, MQ9 and MQ135 gas sensors (for the detection of gases like carbon dioxide, methane, hydrogen, etc.), temperature sensor, LCD display, and an ESP8266 Wi-Fi module. Additionally, an ultrasonic sensor is incorporated into the design. The planned system is created to provide a user-friendly mobile application and web-based interface. The monitoring results are accessible in real-time through a web application, which connects to a Wi-Fi network. The robot, equipped with these specified sensors and modules, plays a critical function in data collection and environmental monitoring. Additionally, machine learning algorithms are used to identify patterns in the data and predict changes in air quality, allowing users to take preventive actions before air quality deteriorates. Another area of future work could be to improve the user interface and data visualization of the system, making it more accessible and easier to use for non-experts. Overall, continued development and optimization of the system can help to increase its effectiveness and applicability in a variety of settings, from indoor air quality monitoring in homes and offices to outdoor pollution monitoring in urban environments.**

Keywords—air quality monitoring, Iot, Robot, Arduino Uno, MQ9, MQ135, ultra sonic sensor , machine learning.

I. INTRODUCTION

Human life is dependent on air, which contains a mixture of gases, including CO2, N2, O2, carbon monoxide, and trace amounts of other rare elements. Adulteration in the composition of air pose a potential threat to the society. It has emerged as a significant peril to contemporary civilization, ranking as the leading cause of premature deaths globally. It has ramifications beyond public health, affecting economic development and contributing to human suffering. Cities experience significant air pollution, mostly from two main sources: industry and automobiles. Vehicle emissions have made a substantial contribution to environmental pollution, which has been linked to health problems such skin rashes, asthma, lung cancer, and heart disease. In a similar way, industrial operations are a major source of air pollution. Concerns about the declining quality of urban air are growing on a global scale. The urgent necessity for monitoring is highlighted by the sharp rise in air pollution that is being caused by a variety of human-induced activities. Maintaining accurate records of air quality is crucial for controlling and mitigating the increasing pollution levels. According to the World Health Organization (WHO), air pollution is the underlying cause of millions of deaths worldwide. Smart cities are being created using cutting-edge technologies in response to the demand for better living conditions that guarantee a high standard of living. Developing smart cities is a concept that governments are actively pursuing in order to attain sustainability and improve public safety and well-being. Commercially accessible sensors that are well-known for their excellent precision in measuring particular contaminants and air characteristics were used in the creation of the project. Temperature, relative humidity, and multiple types of air contaminants were all measured with these sensors. All of the sensors' data was read by an Arduino-based microcontroller unit, which then processed and displayed the results. Then the displayed result is taken and compared with the dataset downloaded from Kaggle and then it is compared with the ML Algorithms such as Decision Tree, KNN, Random Forest, and Naive Baise. Hence the ML Analysis will be done and displayed if the air quality is good or bad.

II. LITERATURE SURVEY

A system for monitoring air pollution that measures the aggregation of pollutants was created, built, and managed using a wireless standard. This device measures a hazardous gas mixture using semiconductor sensors. The hardware unit includes a single-chip microprocessor and an array of sensors for air pollution [1] [2]. The Central-Server is a high-end personal computer application server that is internet-connected. The hardware detects air pollution levels and stores the data in a framework. The impact of air pollution on the chemical composition of the atmosphere is significant, contributing to phenomena such as global warming and acid rain. It is vital to have an air pollution monitoring system in place to avert such harmful natural imbalances.. The standard method of monitoring air quality, which is overseen by the Pollution Control Department, is prohibitively expensive. Because their design must adhere to strict power and cost constraints, wireless sensor networks are a novel and challenging study subject for embedded system design automation. The designed pollution monitoring and tracking system makes use of a wireless sensor network to monitor air quality in many places [3]. Furthermore, it creates near-real-time information and data that is available via smartphones, tablets, and internet-connected devices [4]. To monitor the concentration of toxins in the air, an Arduino Uno, a WIFI module, and a MQ135 gas sensor were used to create a system. MQ135, MQ9, and other gas sensors are used in an IoT-based air pollution monitoring system. These sensors will respond to the module, which will display the data on the Think Speak web server and configure a buzzer to alert us if the air quality deteriorates. Pollution levels have progressively grown over time, owing to reasons such as population growth, increasing vehicle traffic, industry, and urbanization [5-6]. These variables have a negative influence on human well-being as a whole, directly harming persons subjected to such environmental situations. A web server will be used by an IOT-based air pollution monitoring system to monitor air quality over the internet. If the air quality drops below a certain level, an alarm will be activated that indicates the presence of a significant amount of hazardous gases, such as CO2, smoking and NH3 gas. The air components will be shown in webpage on the internet for easy monitoring. The air pollution can be checked using a computer or smartphone. For smart cities, an Internet of Things-based air pollution monitoring system is being developed. As the world's population urbanizes, smart cities face increasing pressure to remain habitable. Urban air quality has caused significant concern worldwide in recent years. In order for a city to be smart and livable, it is necessary to regularly assess its air Quality index. Our research focuses on constructing an Air Quality Monitoring System that utilizes the Internet of Things (IoT) and is designed for Smart Cities. Smart gadgets gather and analyze real-time data on air quality in cities. These smart devices can monitor temperature, humidity, carbon monoxide, LPG, smoke, and other harmful substances in the ambient. An Android application makes the collected data available to a worldwide audience [7-8].Advanced machine learning algorithms, such as Decision tree, Random forest, KNN and naive bayes are employed to analyse the data and derive meaningful insights. The dataset that is used to compare the sensor data is taken from Kaggle.Various results are produced for recall, accuracy, precision, and f1 score; these values are then contrasted and examined using various machine learning techniques. As air pollution has been rising every day, monitoring has proven to be a significant task. Urban air quality has caused significant concern worldwide in recent years. In order for a city to be smart and liveable, it is necessary to regularly assess its air Quality index. Our research focuses on constructing an Air Quality Monitoring System that utilizes the Internet of Things (IoT) and Machine learning, designed for Smart Cities and is available to a worldwide audience.

III. SYSTEM DESIGN

The central control of the system lies with the Arduino Uno, serving as the primary node as shown in the above block diagram represented in the figure 1. It oversees the monitoring of various environmental parameters, including carbon monoxide, carbon dioxide, temperature, and humidity, through the use of dedicated sensors. The components used are Arduino Uno, temperature sensor, MQ135 gas sensor, MQ9 gas sensor, LCD display, motor driver, robot, ultrasonic sensor, esp8266 Wi-Fi, and power source comprise the recommended system. The Arduino Uno serves as the interface for each component in our system. Temperature and humidity of the surrounding air are measured through dedicated temperature sensors. The detection of methane is facilitated by the MQ-135 sensor, while highly toxic gases like carbon monoxide, propane, and methane are identified using the MQ-9 sensor. With the MQ135 apparatus, gases including smoke, ammonia, benzene, Sulphur, carbon dioxide, and so forth can be measured or observed. The DC motors of the robot are linked to a motor driver, and an ultrasonic sensor is employed to identify objects positioned in front of the robot. The key functionality of each sensor lies in the ability to access cloud content through Wi-Fi. The operation of the robot is facilitated by utilizing the Blynk app.

Fig 1 Monitoring System Architecture

The fundamental components comprising a comprehensive sensing and actuation system include the Arduino Uno microcontroller, motor driver, DC motors, MQ-9 and MQ-135 gas sensors, temperature sensor, LCD display, sound sensor, ESP8266 Wi-Fi module, and ultrasonic sensor. The temperature sensor keeps an eye on the outside world, while the gas sensor, the MQ9 and MQ13 collect data on contaminants such as carbon dioxide and carbon monoxide. While the ultrasonic sensor helps in obstacle avoidance, the sound sensor gives information about noise levels. The Arduino Uno assumes a central role in the system, interpreting analog input from sensors and orchestrating responsive movements of the DC motors. The ESP8266 Wi-Fi module ensures smooth wireless connectivity, transmitting real-time air quality data to a remote server for continuous monitoring [9]. Additionally, the LCD screen serves as a user-friendly interface, providing immediate feedback on air quality parameters and obstacle separations. The operational workflow delineates the critical stages, including of system setup, continuous data collection and processing, and motor behavior modification via actuation. The LCD display shows air quality indicators in real time, and the ESP8266 ensures prompt transfer of data to a remote server. The display makes it easier for users to interact with the system and offers a simple way to retrieve information instantly. An innovative method for independent air quality monitoring is offered by this comprehensive system design, which integrates wireless communication, mobility, and sensor integration [10-11]. The robot's mobility, data collection capabilities, and real-time communication make it a valuable instrument for environmental monitoring applications, contributing to the advancement of environmental science and robotics.

A. *Embedded C*

For the purpose of developing the embedded software for an air quality monitoring robot, Embedded C is essential within the Internet of Things (IoT) architecture. Embedded C is a language that is optimized for IoT devices with limited resources. It is lightweight and efficient for programming microcontrollers. Integration of several sensors required for environmental factor tracking is facilitated by the presence of embedded C in the air quality monitoring robot. Specific sensor libraries and protocols must be used by the robot in order to connect with sensors such as the temperature sensors, gas sensors (MQ9 and MQ135), and other relevant devices. Thanks to the language's direct access to hardware registers, resource utilization is made efficient and exact control over sensor readings is made possible.

The robot's ability to respond promptly to environmental changes in the air quality monitoring scenario is contingent upon Embedded C's real-time capabilities. Due to its ability to detect sudden changes in gas concentrations and adjust its direction in response to obstacles, Embedded C's real-time processing skills improve the robot's manoeuvrability and efficiency. For Internet of Things applications, networking and communication protocols are built using Embedded C, which also enables real-time data transmission from the air quality monitoring robot to a remote server. Thus, the robot can connect wirelessly and transmit data uninterruptedly. The language makes it easier to incorporate communication devices, such as the ESP8266 Wi-Fi module.

Table 1: Composition of gases

Parameter	Operating Voltage	Measuring Range
Carbon Monoxide	1.5 V	10 to 10000 ppm
Carbon Dioxide	5 V	10 to 10000 ppm
Temperature	3.3 V	-40 to +80 degree Celsius
Humidity	3.3 V	0 to 100% RH

B. *Sensing Unit*

Four sensors are included in the Sensing Unit for tracking air pollution. The technical specifications of three air quality sensors, which include Temperature and Humidity sensors, are outlined in Table 1. MQ9 is able to handle carbon monoxide/combustible gases with ease thanks to its simple drive circuit. The conductivity of the sensors also rises with an increase in the concentration of gases in the air. MQ135 has a wide range of detection capabilities for NH3, CO2, smoke, and other substances with a very quick response time.

C. *System Architecture*

Hardware Architecture

1. Sensors:

Sensors are crucial to air pollution monitoring systems in the IoT. The air quality indicators determined by these sensors include particulate matter, carbon monoxide, sulfur dioxide, and nitrogen oxides. Physical sensors and chemical sensors are the two types of sensors. Physical sensors determine temperature, humidity, and pressure, while chemical sensors determine air contaminants [12-13].

2. Microcontroller:

The microcontroller is at the heart of IoT air pollution monitoring devices. This component takes data from sensors, analyzes it, and sends it to a cloud server. Typically, the microcontroller takes the form of a microprocessor, such as Arduino, Raspberry Pi, or analogous devices.

3. Communication Module:

The communication module is essential for transmitting data from the microcontroller to the cloud server. This module may use a variety of wireless technologies such as Wi-Fi, Bluetooth, or cellular networks [14].

4. Cloud Server:

The cloud server acts as a centralized center for air quality data storage, processing, and shipment. It collects data from the communication module and stores it in a database. Furthermore, the cloud server offers online as well as mobile apps, enabling data access simpler for users [15].

5. Power Supply:

Power is needed for IoT-based air pollution monitoring devices to work. An external power supply is used for permanent installations, while mobile devices use batteries.

6. Enclosure:

The enclosure serves as an outside protect, shielding the components from external factors including dust, water, and temperature.

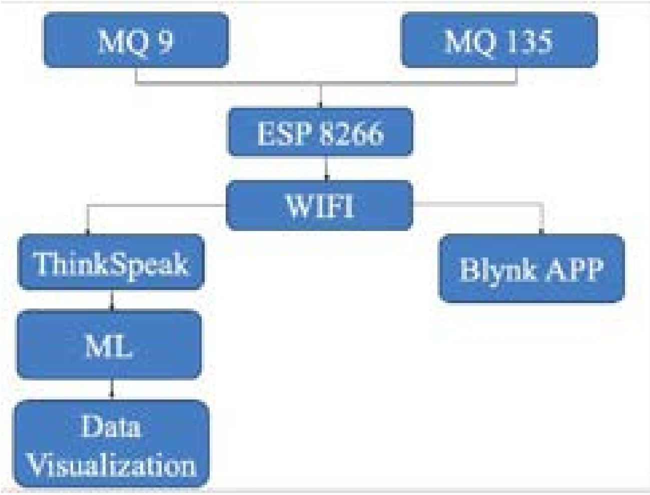


Fig 2 Flow Chart

The software architecture of the IoT device is composed of several layers, including a sensor data acquisition layer, data processing layer, internet connectivity layer, cloud data storage layer, and user interface layer. The flowchart is represented in above figure 2. The firmware running will control the device behavior and send collected data to the ThingSpeak platform for storage and analysis. The Blynk app provides real-time control of the robot. ML Analysis and Visualization Component: After collecting data from the ThingSpeak platform, the next step is to analyze and visualize the collected data using ML algorithms. Overall, this proposed framework for analyzing and visualizing the data collected using the IoT device with MQ9 and MQ135 sensors, Node, Arduino UNO, ThingSpeak, and Blynk platforms provides a structured approach for processing, selecting features, selecting appropriate ML algorithms, building, and training models, evaluating models, and visualizing data. This framework can help to provide meaningful insights for decision-making in various applications, such as environmental monitoring and public.

IV. MACHINE LEARNING MODEL

The air quality monitoring robot is trained with machine learning algorithms such as Decision Tree, Random Forest, KNN, and Navie Bais, thereby producing the respective Accuracy, F1 Score, Recall Score, and Precision Score.

A. Decision Tree:The Decision tree is a structure that resembles a river, with each branch denoting a test result and each leaf node representing a class label or a choice made following the evaluation of all the tests. It is the algorithm for supervised learning. If the temperature, humidity, MQ135 or MQ9 values are higher than the normal values, the algorithm compares the data set from the ThinkSpeak app with the fixed dataset from Kaggle and displays whether the weather is good or bad.

B. Random Forest: It functions by building many Decision Trees in the training stage. A random subset of the data set is used to measure a random subset of characteristics in each partition throughout the construction of each tree. Because of the variety that this randomness brings to the individual trees, the likelihood of overfitting is decreased and the overall prediction performance is enhanced. In prediction, the method combines all tree results, averaging (for problems involving regression) or voting (for tasks involving classification). With the help of several trees and their insights, this cooperative decision-making process yields findings that are accurate and consistent.For classification and regression tasks, random forests are frequently utilized because of their reputation for managing complex data, minimizing overfitting, and producing accurate predictions in a variety of settings and the figure 3 represent the same.

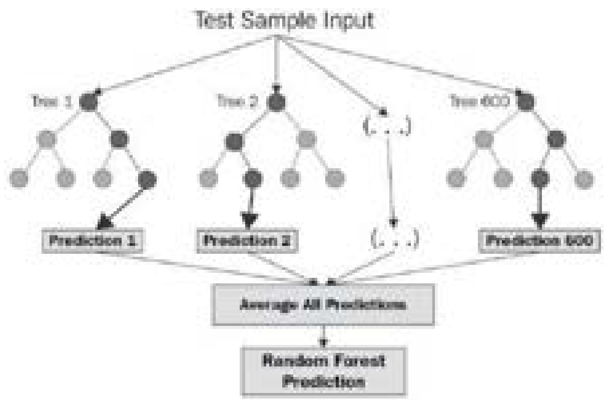


Fig 3 Flow the way that the random forest functions.

KNN: The k-Nearest Neighbors (KNN) method is used in robot air quality monitoring to forecast air quality measures based on the similarity of nearby data points. First, the robot gathers information from multiple sensors that measure different characteristics such as nitrogen dioxide (NO2), carbon monoxide (CO), and particulate matter (PM). Relevant characteristics are collected from these observations, including temperature, humidity, wind speed, and pollutant concentrations. Normalization and missing value management are also done during preprocessing. Without explicit training, the KNN algorithm learns to memorize all accessible data points and their matching labels during the training phase. The robot selects the k nearest neighbors when it meets a new place by measuring air quality metrics and comparing them with existing data points. While the method averages the values from the nearest neighbors to create predictions for regression tasks, it aggregates the labels of the k nearest neighbors via majority voting for classification tasks. Metrics like accuracy and mean squared error are used to assess the model's performance; actual data may be used to recalibrate the model. While KNN is easy to use, flexible, and interpretable, it has drawbacks in terms of computing expense, noise and outlier sensitivity, and finding the ideal value for k. However, it continues to be an important tool for robots that monitor air quality, particularly in areas with complicated and nonlinear data.

Navie Bayes: -Naive Bayes is a probabilistic classifier that relies on the independence of features and is based on the Bayes theorem. Naive Bayes is a simple algorithm that may be used to solve a wide range of challenging real-world problems. It excels at text classification tasks such as sentiment analysis and spam detection. It uses little training data to estimate the required parameters and is effective, particularly with high-dimensional datasets.

V. EXPERIMENTAL SETUP

Figures 4 below show the complete system setup, which includes sensors, Arduino, and a buzzer sensor. Figure 6 shows the website's main page, and figures 5 to 10 show the variations of parameters monitored using the thinkspeak website, The motor connected to the system enables it to move around the location to determine the pollution level at various geographical locations.

Fig 4 Robot Wired Connection with Arduino Uno

VI. RESULT

Fig 5 Thing Speak Main Web Page

Fig 6 Temperature Variation Graph

The temperature graph shown in figure 6 displays the temperature measured in short period of time where x- axis represents date, y-axis temperature and there is a spike in temperature which is recorded.

Fig 7 Humidity Variation Graph

This the humidity graph represented in figure 7 is used to display the humidity levels at different times where x-axis represents date, y-axis represents humidity and there is significant fluctuation in the humidity level which is being monitored.

Fig 8 Mq135 Variation Graph

This the mq135 graph is used to display the fluctuations in the gases available in the environment as shown in figure 9.

Fig 9 Mq9 Variation Graph

The figure 10 displays sound levels over specific dates. It provides insights into variations in air quality over time.

Fig 10 Sound Variation Graph

Table 2: ML Analysis values for Accuracy

Algorithm Used	Accuracy
RF	0.7568
DT	1.0000
KNN	0.7982
NB	0.7595

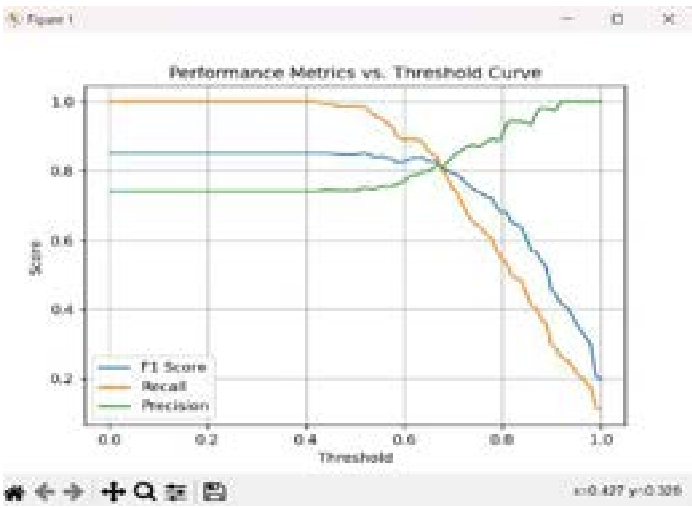


Fig 11 Performance Metrics vs Threshold Curve

Table 2 represent the accuracy obtained for various ML algorithm. Figure 11 shows the threshold curvewhich typically represented as a line or curve on a graph. The x-axis represents the decision threshold. It is a numerical value that determines how confident the model needs to be to classify a data point as positive.The y-axis represents a performance metric. The curve on the graph shows how the chosen performance metric changes as the decision threshold varies.

VII. CONCLUSION

This paper provides the analysis of Air Quality Index (AQI), which is dependent on the temperature humidity and sound of the surrounding air. The data is used to calculate air quality in parts per million. The benefit of MQ135, and MQ9 is that they can identify hazardous gases such as CO, NH4, and smoke, among others. It can accurately monitor the temperature in Celsius, the humidity in percentage, and the air quality in parts per million. The outcomes of the trials are confirmed using data from ML analysis Additionally, the Think Speak app is used to show the different variations in the level of air quality surrounding the setup.

REFERENCES

[1] V. Barot and V. Kapadia, "Air Quality Monitoring Systems using IoT: A Review," ComPE, Shillong, India, 2020, pp. 226-231, doi: 10.1109/ComPE49325.2020.9200053.

[2] H. Gupta, D. Bhardwaj, H. Agrawal, V. A. Tikkiwal and A. Kumar, "An IoT Based Air Pollution Monitoring System for Smart Cities," 2019 ICSETS, India, 2019, pp. 173-177, doi: 10.1109/ICSETS.2019.8744949.

[3] Sivaramakrishnan, S., and T. Kesavamurthy. "Identifying Cluster Head and Data Transmission Through Them for Efficient Communication in Wireless Sensor Network." Journal of Computational and Theoretical Nanoscience 14.8 (2017): 4014-4020.

[4] Devi, BS Kiruthika, et al. "An improved security framework in health care using hybrid computing." Malaysian Journal of Computer Science (2022): 50-61.

[5] P. D. Potbhare, K. Bhange, G. Tembhare, R. Agrawal, S. Sorte and P. Lokulwar, "IoT based Smart Air Pollution Monitoring System," 2022 ICAAIC, Salem, India, 2022, pp. 1829-1834, doi: 10.1109/ICAAIC53929.2022.9792743.

[6] S. M. Abiduzzaman, H. Mansor, T. S. Gunawan and R. Ahmad, "Real-Time Outdoor Air Quality Monitoring System," ICSIMA, Bandung, Indonesia, 2021, pp. 140-145, doi: 10.1109/ICSIMA50015.2021.9526332.

[7] Vandana, C. P., & Chikkamannur, A. A. (2019). Semantic ontology based IoT-resource description. International Journal of Advanced Networking and Applications, 11(1), 4184-4189.

[8] A. Udomlumlert, N. Vichaidis, P. Kamin, T. Surasak, S. C. . -H. Huang and N. Jiteurtragool, "The Development of Air Quality Monitoring System using IoT and LPWAN," 2020 InCIT, Chonburi, Thailand, 2020, pp. 254-258, doi: 10.1109/InCIT50588.2020.9310949.

[9] Saravanan, K, “IoT-integrated deep learning model and SmartBin system for real-time solid waste management” Proceedings of the International Conference on Trends in Electronics and Informatics (ICOEI 2023). 11-13, April 2023

[10] Sivaramakrishnan, S., et al. "Introduction to AI Technique and Analysis of Time Series Data Using Facebook Prophet Model." *Innovative Engineering with AI Applications* (2023): 171-188.

[11] Sivaramakrishnan, S., Fernandez, T. F., Babukarthik, R. G., & Premalatha, S. (2022). Forecasting time series data using arima and facebook prophet models. In Big data management in Sensing (pp. 47-59). River Publishers.

[12] P. Manikandan, B. N. K. Reddy, M. V. Bhanu, G. Ramesh and V. P. Reddy, "IoT Based Air Quality Monitoring System with Email Notification," 2021 ICCES, Coimbatre, India, 2021, pp. 616-620, doi: 10.1109/ICCES51350.2021.9489027.

[13] M. B. Marinov, I. Topalov, E. Gieva and G. Nikolov, "Air quality monitoring in urban environments," 2016 ISSE, Pilsen, Czech Republic, 2016, pp. 443-448, doi: 10.1109/ISSE.2016.7563237.

[14] Sivaramakrishnan, S., et al. "Augmentation of Terahertz Communication in 6G and Its Dependency for Future State-of-the-Art Technology." *Challenges and Risks Involved in Deploying 6G and NextGen Networks*. IGI Global, 2022. 91-105.

[15] J. Karthiyayini, A. C. Vantipalli, D. S. Tanti, K. Malvika Ravi and K. Kannan, "IOT based AquaSwach," 2022 2nd International Conference on Artificial Intelligence and Signal Processing (AISP), Vijayawada, India, 2022, pp. 1-12.

HICON AI: Higher Education Counseling Bot

Arjun Dev Singla
Department of Networking and Communications
SRM Institute of Science and Technology
Kattankulatur, India
as9301@srmist.edu.in

Shashank Tripathi
Department of Networking and Communications
SRM Institute of Science and Technology
Kattankulatur, India
st3158@srmist.edu.in

A Helen Victoria
Department of Networking and Communications
SRM Institute of Science and Technology
Kattankulatur, India
helenvia@srmist.edu.in

***Abstract*—As more and more students are seeking counseling services to help them pursue higher education opportunities abroad, it has become apparent that there is a lack of automation and a monopolization of counseling agencies, which presents a challenge for these students. To address this issue, this study has developed HICON AI: Higher Education Counselor. HI-CON made from Higher - Counseling is an innovative application that offers personalized college selection and preparation recommendations to students. By asking a set of defined questions, the bot gets to know the user and provides tailored guidance based on the information provided, utilizing refined Machine Learning models and Retrieval Augmented Generation. This study has specifically used Llama 2, LLM by meta for the considered use case, because of its high performance and financial viability. The developed new product ensures the highest level of accuracy and reliability.**

***Index Terms*—HICON AI, LLMs, text to Speech, Categorizer, Machine Learning, Natural Language Processing, Retrieval Augmented Generation, Resume Screener.**

I. INTRODUCTION

India has witnessed a consistent growth in the Gross Enrollment Ratio (GER) within its higher education sector over the years. As per the latest findings from the All India Survey on Higher Education (AISHE) [9], the GER for higher education in India recorded a figure of 27.1 percent in the academic year 2020-21. Students faces a variety of questions about their academic interests, university preferences, application procedures, and future career prospects in the varied world of higher education today. Counselors for higher education are essential in addressing these issues and giving students personalized guidance.

Introducing 'HICON AI,' a dedicated chatbot that helps students pursue higher education. It is built on Large-scale Language Models (LLM). This AI functions as a virtual counselor by using Natural Language Processing (NLP) to evaluate student data and inquiries, producing automated responses that provide direction and advice. HICON AI attempts to streamline the procedure and lessen students' dependency on outside counselors. 'HICON AI's' audio prompting is a noteworthy feature that allows users to interact through audio inputs and outputs, improving productivity and the user experience in general. The model's capacity to understand inputs and generate pertinent text outputs is guaranteed by the large datasets it was trained on. 'HICON AI' is able to provide objective, individualized advice for each student according to their specific needs by classifying them according to a variety of criteria. Effective Class Categorization has been made possible by the development of extensive datasets, guaranteeing that the chatbot will respond with both precision and generality.

With the introduction of Retrieval Augmented Generation(RAG), the LLM hallucination and relevance of the data that is generated by the LLM is addressed. It improves the quality of the data generated by keeping it updated to the latest changes. Additionally, the feature of Resume-Screener of HICON AI helps the student to upload the all the necessary details required to be known to the system in one go insteadof typing the answers to each and every question in a user efficient manner. The interface through which the student will interact with the system is an easy to access and a user friendly one. Crafted with a user-friendly interface, the proposed HICON AI stands as a solution for personalized guidance for pursuing higher education. Seamlessly designed to offer a smooth user experience, it provides authentic and tailored support for academic and career decisions, ensuring a unique and reliable resource for students. The HICON AI proves to be a useful instrument in the field of higher education, providing students with individualized and easily accessible guidance, ultimately leading to a more informed and capable student body.

II. LITERATURE REVIEW

This section investigates the existing research works that enable the foundation of HICON AI.

A. Large Language Models (LLMs)

The field of LLMs has witnessed significant advancements, with various models like GPT-3, LlaMa-2, and Bard. A recent study by [5] provides a comprehensive comparison of these models, analyzing the impact of fine-tuning and hyper parameter tuning on their performance. The findings resonate with the observed increase in chatbot usage since the introduction of Generative AI.

Furthermore, [1] proposed a E2E benchmark to evaluate the quality of LLM-generated responses in chatbots. This bench-mark leverages cosine similarity to assess the resemblance between AI-generated text and human-crafted text (golden text). The E2E benchmark aims to mitigate the issue of hallucinations, where LLMs produce factually incorrect outputs. The study also highlights the effectiveness of prompt engineering in improving ROGUE scores, a metric used to evaluate the quality of summaries. However, the vulnerabilities associated with ROGUE-1 and ROGUE-2 scoring are acknowledged.

B. Chatbots

The potential of LLMs in educational chatbots is explored by [2]. Their work presents an intelligent tutoring system that utilizes Retrieval-Augmented Question Answering (RAQs). This approach involves extracting keywords from prompts and leveraging the internet to gather relevant knowledge, enriching the system's response generation capabilities. The model is further fine-tuned on educational datasets to enhance its responsiveness. Additionally, researchers have introduced Socratic teaching elements, where the chatbot engages in multi-step questioning instead of providing direct answers, fostering deeper learning. While their work paves the way for intelligent tutoring systems, it doesn't delve into the domain of career counseling. The authors propose future research directions in question generation and career planning.

Another study by [12] investigates the use of deep learning algorithms to create a personalized counseling chatbot for students. While this model avoids hallucinations due to its focused training data, its accuracy falls short of LLM-powered solutions.

C. Retrieval-Augmented Generation (RAG)

While our work focuses on LLMs for response generation, it's important to acknowledge Retrieval-Augmented Generation (RAG) as a complementary approach. RAG models com- bine retrieval techniques with generative capabilities. Several recent studies explore RAG applications in various domains. For instance, the research presented in [6] proposes a novel RAG variant that eliminates the need for vector embeddings, potentially simplifying model implementation, particularly for niche domains with limited data.

Other relevant RAG studies include [7] which explores RAG in the context of a specific industry process, and [4] which delves into potential challenges and limitations of RAG-based chatbots.

Our research focuses on LLM-based response generation for HICON AI. However, the exploration of RAG techniques for future iterations or for complementary functionalities remains a valuable avenue for further investigation.

Based on the above survey, we have formulated our research objectives which are unique and focuses on Higher Education Counseling.

III. Motivation

The increase of popularity for higher studies among students has been evident. To cater the huge number of students having queries regarding the higher education and those who are looking for a counseling advice, HICON AI would be of great help rather than going for a manual higher education counselor.

HICON AI, designed specifically for higher education, addresses the need for a dedicated chatbot in this domain, offering students personalized assistance. In contrast to conventional chatbots that often provide generic responses, HICON AI categorizes students based on their input, generating specific prompts for tailored advice on course selection, university choices, and other relevant aspects. To enhance the user experience, it incorporates features such as speech-to-text and text-to-speech, accommodating students with disabilities.

An essential component of HICON AI's functionality is the integration of advanced technologies like Retrieval Augmented Generation (RAG). RAG combines information retrieval with natural language generation, enabling swift data retrieval and the generation of contextually appropriate responses. This ensures that the advice provided to students remains current and accurate in the dynamic landscape of higher education. Additionally, HICON AI leverages Large-scale Language Models (LLMs) with supercomputer capabilities, trained on extensive datasets, to understand user inputs and generate nuanced, domain-specific responses, contributing to its overall efficiency.

Moreover, HICON AI goes beyond academic guidance by incorporating a resume screener feature. This tool streamlines the process of evaluating resumes, utilizing advanced algorithms to identify key qualifications, skills, and experiences relevant to the user's career goals. The resume screener enhances Counselor AI's utility by assisting students in preparing impactful resumes tailored to their chosen field of study or profession. This multifaceted approach makes Counselor AI a comprehensive and cutting-edge tool, offering students not only personalized educational guidance but also support in navigating the complexities of the job market.

IV. Objectives

HICON AI is crucial to the education sector because it helps many students through a critical point in their education, especially those who are pursuing higher education. This AI-based counselor is intended to provide fair, individualized, and easily accessible counseling sessions to students 24/7, answering their questions and dispelling any doubts they may have about pursuing higher education.

'HICON AI' places a lot of emphasis on classifying students according to the data they submit. This classification guarantees that the support given is unique to every student, in

name	age	college	year_of_st	cgpa	preffered_	internship
Aarav Shar	18	Royal Insti	3rd Year	9.2	Massachu:	yes
Priya Patel	21	Indraprast	2nd Year	8.2	Stanford U	no
Aditya Kap	18	Himalayan	4th Year	9	California	yes
Ananya Ve	19	Ganges Ins	1st Year	8.9	Harvard U	no
Rajat Singh	20	Bengaluru	2nd Year	8.2	University	yes
Kavita Das	21	Vellore Ins	3rd Year	8	University	no
Arjun Gupt	18	Delhi Colle	4th Year	9.4	National U	yes
Meera Rec	22	Jamshedpu	2nd Year	8.3	ETH Zurich	yes
Siddharth .	20	Malabar E	2nd Year	9.1	University	yes
Nisha Khar	18	Jaipur Nati	3rd Year	7.8	Imperial C	yes
Vikram Me	21	Rajputana	2nd Year	8.4	University	no

Fig. 1. Image of Dataset

internships	projects_d	projects	research_	research_	social_gro	social_gro
Google: Sc	yes	E-commer	yes	Renewable	yes	Engineerin
NULL	yes	Smart Hon	no	NULL	yes	Robotics C
Facebook:	yes	Chatbot fc	no	NULL	yes	Programm
NULL	no	NULL	no	NULL	yes	Aero Club
IBM: Cybe	yes	Social Mec	no	Efficiency	yes	Electronic:
NULL	yes	Personal F	yes	Developm	yes	Environme
Tesla: Rob	yes	Image Rec	no	NULL	no	NULL
Netflix: Fu	yes	Stock Port	yes	Biomecha	no	NULL
Uber: Data	yes	Home Ene	yes	Design of	yes	Debate an
Adobe: Us	yes	Health anc	yes	Developm	no	NULL
NULL	yes	Language	yes	Noise Poll	no	NULL

Fig. 2. Another Parameters of the Dataset

contrast to the general approach of most AI-based counselors. The four classifications created by 'Counselor AI' serve people at different phases of their pursuit of higher education. Those in Category 1 are those who have little experience and are just learning about the field, whereas those in Category 2 are those who are just starting out and may have completed some projects or internships. Group 3 consists of people who know enough but could still do better in some areas, while Group 4 is made up of people who know enough but are looking for new information to keep current.

The training model for student categorization makes use of a self-created dataset to guarantee that the prompts are brief and encompass all the higher studies material that the students require. This method improves HICON AI's abilityto provide accurate and pertinent advice that is suited to each student's needs.

Integrating cutting-edge technologies like Retrieval Augmented Generation (RAG) is crucial to HICON AI's functionality. RAG combines natural language generation with information retrieval to provide quick data retrieval and responses that are appropriate for the given context. This guarantees that, in the ever-changing context of higher education, the counsel given to students stays relevant and accurate. HICON AI's overall efficiency is further enhanced by the use of Large-scale Language Models (LLMs) with supercomputer capabilities that have been trained on vast datasets to comprehend user inputs and produce complex, domain-specific responses.

Additionally, HICON AI has a resume screener feature,

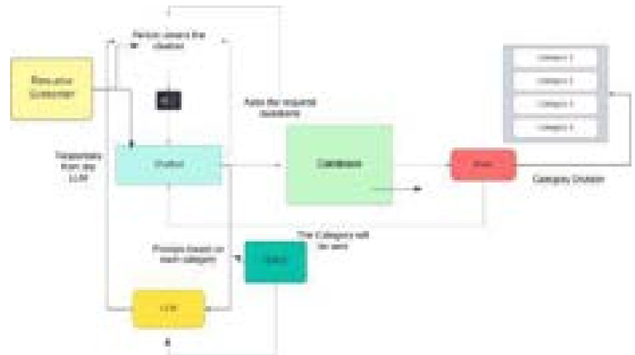


Fig. 3. System Architecture Diagram

which takes it beyond academic guidance. This application expedites the resume evaluation process by applying sophisticated algorithms to pinpoint the most important experiences, education, and training that relate to the user's desired career path. HICON AI's usefulness is increased by the resume screener, which helps students create compelling resumes that are relevant to their chosen careers or fields of study. HICON AI is a comprehensive and state-of-the-art tool that provides students with individualized educational guidance and support in navigating the complexities of the job market thanks to its multifaceted approach.

Using a Large-scale Language Model (LLM), which is well-known in the field of higher education and imitates communication between a real human counselor and students seeking higher education, is another important goal of the HICON AI. It processes and generates automated responses by analyzing inputs in the form of student data and queries using Natural Language Processing (NLP). Following classification, LLMs will produce answers for each category based on eight to ten prompts. These answers will be well-organized and fully contextualized to the questioner and the class.

Lastly, one of the other main goals of the HICON AI is the voice-based dialogue between the model and the students, which makes it appealing to a large number of students, especially those with disabilities. The speech inputs are transformed into text and fed to the model for processing. Once the model has finished processing the text responses, the text responses are transformed back into speech, improving the seamless experience for the students.

V. Proposed Methodology

A. *System Architecture*

When a user inputs 'HICON AI,' the assistant greets them warmly. The assistant will then prompt the user to answer specific questions about their unique use case. These questions will cover essential information like the user's name, age, year of study, as well as professional details such as internships completed and training specifics. The responses to these questions will be immediately populated in the resume screener. Additionally, there will be behavioral inquiries, such as whether the user has joined any social groups, with further details on these groups being saved in a separate database. These questions aid in categorizing users into distinct groups. Furthermore, language models (LLMs) will generate personalized questions based on these categories and tailored to each individual, such as asking about their desired university and courses. Since each person will have a unique perspective, responses will be customized to their interests. This process helps construct the final responses of HICON AI.

B. Modules

1. Data Gathering: To begin the process, the model requests the user to respond to a series of inquiries aimed at understanding the student's background. This includes information on work experiences, projects, research papers, social initiatives, volunteer work, and leadership roles. This data is usually extracted from resume using a resume screener built by us, or directly inputted by the user - if the data is not available in the resume.

2. Student Category division: Following data collection, a distinct Machine Learning model is employed to forecast the user's Category into one of four groups. Category 1 encompasses users who are new to the domain and are exploring its features. Category 2 individuals seek to enhance their knowledge and delve deeper into their interests. Category 3 users aim to identify gaps in their profiles and strategies to improve them. Category 4 users are focused on staying informed and maximizing their chances of admission to their desired universities.

3. Prompting: Prompts tailored to specific categories are created and adjusted using user data to ensure personalized responses for each category's use cases. RAG adds an extra feature for better quality of responses.

4. Responses: The answers for each prompt within the category are generated by language models (LLMs). Utilizing LLMs streamlines the process, making it more efficient to offer resources and advice.

5. Retrieval Augmented Generation: An essential component of HICON AI's functionality is the integration of advanced technologies like Retrieval Augmented Generation (RAG). RAG combines information retrieval with natural language generation, enabling swift data retrieval and the generation of contextually appropriate responses. This ensures that the advice provided to students remains current and accurate in the dynamic landscape of higher education. RAG takes in account of the data provided. It understands the process ofhow the data is created and the applies the same techniquesto generate new data.

6. Resume Screener: HICON AI goes beyond academic guidance by incorporating a resume screener feature. This tool streamlines the process of evaluating resumes, utilizing advanced algorithms to identify key qualifications, skills, and experiences relevant to the user's career goals. This data is extracted to make the user profile more efficient and to more about the user.

C. Implementation

For the implementation of the assistant, this study has opted for a form-based, text-input approach instead of socket programming. This approach allows the user, who is the student, to input data directly from their resume using a Resume Screener. This decision was made to streamline the process and simplify communication between the user and our Counselor AI model. Within the application, there are two main phases: one where data is gathered from the user via the form, and another where responses are provided based on the input received.

Before Categorization, a significant amount of data preprocessing is performed, during which all data items are assigned to their corresponding parameter classes and binary parameters are mapped. Data dictionary converted to pandas. Data frames will be used by the model to classify the data before being returned to the main dataset for further usage.

To create the categorizer, a random forest model was developed for category division. Due to the unique dataset created specifically for the model, there was limited data available. Random forest was chosen for its efficiency with smaller datasets. While both decision trees and random forest algorithms could have been utilized for categorization, random forest was preferred due to its ability to mitigate overfitting and its robustness to noise. Random forest is essentially a combination of multiple decision trees, which helps reduce overfitting by allowing each tree to over-fit differently [13]. Overfitting occurs when a model accurately predicts training data but fails to generalize to new data. This phenomenon arises when variance increases while bias decreases.

After creating the categorizer, this study has integrated the prompts for each category directly into the question-answer category interface. This eliminated the need for socket programming to deliver responses. Additionally, for personalized responses, we included data obtained from the user within the prompts themselves.

To finalize our project, a LLM is integrated to provide responses. This study has selected for Llama 2 developed by Meta-AI. Llama 2 was chosen over other LLMs primarily due to its novelty and accessibility. Moreover, Llama 2 boasts extensive training on billions of parameters, ranging from 7B to 70B parameters, making its responses exceptionally efficient [14]. Additionally, we're augmenting the model with Retrieval-Augmented Generation (RAG) for further fine-tuning and enhanced efficiency in response generation.

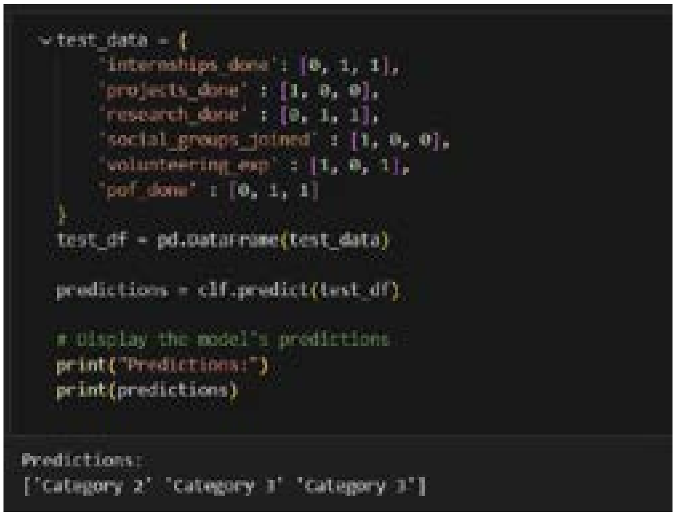


Fig. 4. Categorizer

Fig. 5. Data Collection

Fig. 6. Resume Screening

For the text-to-speech functionality we have used Whisper AI [11] which an Automatic Speech Recognition model made by Open AI. It a transformer sequence to sequence model trained on speech processing techniques. Model was also trained on a huge dataset, so fine-tuning is not necessary and it gives a good performance for out use case.

We used langchain [10] for the implementation of RAG in our system. LangChain is an open-source framework that helps developers build applications using large language models (LLMs). We used this because it provides direct functionalities to implement RAG and reduces the manual code writing. RAG [8] was used because it helps to improve the performance of the system. For the implementation, we provided it with context of an example output and the context of how the output was generated which we call history. The LLM understands this history of how the example output is generated and in the same way generates output for the new prompts. In this way we have a controlover our outputs and also can reduce the LLM hallucinations.

In our model, we have extensively utilized Natural Language Processing (NLP) at various stages. For instance, during the implementation of Retrieval-Augmented Generation (RAG), we tokenized the context and generated vector embeddings for further processing. Similarly, while generating outputs, we adjust the temperature and epochs of the Language Learning Model (LLM) to achieve the desired results. The LLM, which employs a sequence-to-sequence (seq-2-seq) architecture for generating responses, also relies on NLP.

VI. Results and Metrics

1. Data Gathering: The initial phase involves collecting information from the user, including their preferred universities, personal details such as name, age, and current year of study. This study also gather data on their academic background, encompassing the number of internships, projects completed, and research papers authored. Furthermore, we aim to capture details regarding the user's involvement in social activities, including volunteering experiences and held positions of responsibility.

2. Categorizing: After getting the required data from the user, we will be using the proposed categorizer model to predict the Category, based on the above data dictionary. Let's predictour category.

3. Responses: Once the category is determined, we'll incorporate prompts tailored to each category along with user-specific parameters obtained from the user data to ensure personalized responses. Subsequently, by executing the model, we'll generate our responses.

4. Accuracy Metrics: The Accuracy metrics for the categorizer, presented in *Table 1*, outline the performance across four categories. Utilizing a random forest algorithm,

Accuracy: 0.80				
	Precision	Recall	F1 Score	Support
Category 1	0.00	0.00	0.00	1
Category 2	0.75	0.75	0.75	4
Category 3	0.75	1.00	0.86	3
Category 4	1.00	1.00	1.00	2

TABLE I. ACCURACY

the metrics include Precision, Recall, and F1 Score for each category.

Discussing the evaluation of the LLM, we utilize specific tools and comparisons, as referenced in [3]. One benchmark used to assess the truthfulness of the LLM is TruthfulQA, where the score for LLaMA-2 7B is 45.2%, indicating satisfactory performance. While there are other evaluation benchmarks available, they are more meaningful when the LLM is fine-tuned to a specific dataset. Since our model wasn't fine-tuned, we adjusted the prompts to optimize the output quality.

A. Performance Improvement

Prompting: To enhance our system's performance, we can use prompting techniques like Chain of Thought Prompting for multi-step reasoning, Zero-shot-COT for tasks with limited data, and Tree of Thoughts for exploring multiple possibilities. These techniques can improve the model's accuracy, generalization, and interpretability.

Fine-tuning: Fine-tuning LLMs can significantly improve our system's performance by adapting pre-trained models to specific tasks or domains. However, we haven't fine-tuned our LLM due to the unavailability of a suitable dataset in our domain. In the future, we can consider fine-tuning for better performance, but it requires careful consideration of factors like dataset quality and training parameters.

Iterative Refinement: Iterative improvement include self-feedback and refinement with the same LLM, without requiring additional training or supervised data. This method can enhance LLM performance, making it more accurate and relevant for various tasks.

VII. Conclusion and Future Scope

HICON AI represents a cutting-edge model poisedto serve as a valuable resource for students aspiring to study abroad. With its focus on user-centricity, it offers an effective solution for various industries. Leveraging large language models and Natural Language Processing techniques, HICON AI delivers personalized guidance to students preparing for higher education. By reducing reliance on costly external counselors through targeted advice and support, it aims to democratize access to higher education, making it more affordable. While HICON AI is still in its nascent stages, advancements in technology promise to revolutionize the student experience, reshaping the landscape of higher education counseling. Moreover, there is untapped potential within the AI counseling domain. Strategies such as fine-tuning the model with college-specific data and implementing a peer review system for LLM-generated responses can enhance efficiency. Additionally, incorporating various score predictors for evaluating components like statements of purpose and letters of recommendation could further enrich the model's capabilities.

References

[1] Debarag Banerjee, Pooja Singh, Arjun Avadhanam, and Saksham Srivastava. Benchmarking llm powered chatbots: Methods and metrics. *arXiv preprint arXiv:2308.04624*, 2023.

[2] Yuhao Dan, Zhikai Lei, Yiyang Gu, Yong Li, Jianghao Yin, Jiaju Lin, Linhao Ye, Zhiyan Tie, Yougen Zhou, Yilei Wang, Aimin Zhou, Ze Zhou, Qin Chen, Jie Zhou, Liang He, and Xipeng Qiu. Educhat: A large-scale language model-based chatbot system for intelligent education, 2023.

[3] Aparna Dhinakaran. The guide to llm evals: How to build and benchmark your evals. 2022.

[4] Philip Feldman. James R. Foulds and Shimei Pan. Ragged edges: The double-edged sword of retrieval-augmented chatbots, 2024.

[5] Muhammad Usman Hadi, Rizwan Qureshi, Abbas Shah, Muhammad Irfan, Anas Zafar, Muhammad Bilal Shaikh, Naveed Akhtar, Jia Wu, Seyedali Mirjalili, et al. Large language models: A comprehensive survey of its applications, challenges, limitations, and future prospects. 2023.

[6] Bongsu Kang, Jundong Kim, Tae-Rim Yun, and Chang-Eop Kim. Prompt-rag: Pioneering vector embedding-free retrieval-augmented generation in niche domains, exemplified by korean medicine, 2024.

[7] Jaewoong Kim and Moohong Min. From rag to qa-rag: Integrating generative ai for pharmaceutical regulatory compliance process, 2024.

[8] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. Retrieval-augmented generation for knowledge-intensive nlp tasks, 2021.

[9] Ministry of Education. Ministry of education releases all india survey on higher education (aishe) 2020-2021. 2023.

[10] Keivalya Pandya and Mehfuza Holia. Automating customer service using langchain: Building custom open-source gpt chatbot for organizations, 2023.

[11] Alec Radford, Jong Wook Kim, Tao Xu, Greg Brockman, Christine McLeavey, and Ilya Sutskever. Robust speech recognition via large-scale weak supervision, 2022.

[12] Muskan Sharma, Anita Kumari, and Jyotsna. Ai-based deep learning chatbot for career and personal mentorship. In *2023 IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET)*, pages 1–6, 2023.

[13] Hao Li Siddharth Misra. Noninvasive fracture characterization based on the classification of sonic wave travel times. 2020.

[14] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajjwal Bhargava, Shruti Bhosale, Dan Bikel, Lukas Blecher, Cristian Canton Ferrer, Moya Chen, Guillem Cucurull, David Esiobu, Jude Fernandes, Jeremy Fu, Wenyin Fu, Brian Fuller, Cynthia Gao, Vedanuj Goswami, Naman Goyal, Anthony Hartshorn, Saghar Hosseini, Rui Hou, Hakan Inan, Marcin Kardas, Viktor Kerkez, Madian Khabsa, Isabel Kloumann, Artem Korenev, Punit Singh Koura, Marie-Anne Lachaux, Thibaut Lavril, Jenya Lee, Diana Liskovich, Yinghai Lu, Yuning Mao, Xavier Martinet, Todor Mihaylov, Pushkar Mishra, Igor Molybog, Yixin Nie, Andrew Poulton, Jeremy Reizenstein, Rashi Rungta, Kalyan Saladi, Alan Schelten, Ruan Silva, Eric Michael Smith, Ranjan Subramanian, Xiaoqing Ellen Tan, Binh Tang, Ross Taylor, Adina Williams, Jian Xiang Kuan, Puxin Xu, Zheng Yan, Iliyan Zarov, Yuchen Zhang, Angela Fan, Melanie Kambadur, Sharan Narang, Aurelien Rodriguez, Robert Stojnic, Sergey Edunov, and Thomas Scialom. Llama 2: Open foundation and fine-tuned chat models, 2023.

Efficient Schizophrenia Detection Through Transfer Learning for Robust Diagnostics

Gnaneswari Kancherla
Dept of ECE
V R Siddhartha Engineering College
Vijayawada, India
gnaneswarikancherla20@gmail.com

Srinivasa Panchavedi Gopamsana
Dept of ECE
V R Siddhartha Engineering College
Vijayawada, India
panchavedi.sg@gmail.com

Bhaskar Reddy Yadamakanti
Dept of ECE
V R Siddhartha Engineering College
Vijayawada, India
bhaskarreddy72541@gmail.com

Vijayasankar Anumala
Dept of ECE
V R Siddhartha Engineering College
Vijayawada, India
vijayasankar.anumala@gmail.com

Sarada Yaddanapudi
Dept of ECE
V R Siddhartha Engineering College
Vijayawada, India
ysd.686@gmail.com

Fayaz Ahamed Shaik
Dept of ECE
V R Siddhartha Engineering College
Vijayawada, India
fayaz@vrsiddhartha.ac.in

***Abstract*— Schizophrenia, a severe mental illness characterized by emotional, behavioral, and cognitive abnormalities, presents challenges in accurate diagnosis and effective treatment. Patients often experience delusions, hallucinations, and difficulties in thinking and social interaction. The impact on one's reality perception and emotional control is substantial. While typically emerging in late adolescence or early adulthood, post-COVID-19 cases of schizophrenia have shown an increase, emphasizing the need for improved diagnostic tools and treatments. Current methods involve antipsychotic drugs, psychotherapy, and mental health specialist assistance. However, these approaches vary in quality and accuracy. Electroencephalogram (EEG) signals, measuring brain electrical activity, provide insights into schizophrenia but lack direct diagnostic capabilities. To address these challenges, a deep learning model is proposed, utilizing transfer learning with Inception V3, Xception, and DenseNet121 models. Recognizing the limitations of a singular model, the study explores employing multiple models to enhance performance in schizophrenia detection. The investigation acknowledges and addresses inherent drawbacks, aiming to improve the accuracy of identification through the application of deep learning methods on specific datasets. In the context of the increasing post-COVID-19 schizophrenia cases, the urgency for advanced and accurate diagnostic approaches becomes even more pronounced.**

Keywords: Inception V3, Xception, DenseNet121, Electroencephalogram (EEG), Transfer Learning.

I. INTRODUCTION

The brain, a remarkable organ, plays a pivotal role in orchestrating the complexities of the human body. Its significance extends beyond its physical presence, as it is the command center for cognitive functions, emotions, and intricate neurological processes. The brain's intricate network of neurons enables us to think, feel, and interact with the world around us. However, like any complex system, the brain can face challenges, one of which is schizophrenia. Schizophrenia is a mental health disorder that affects how a person thinks, feels, and behaves. It can disrupt the normal functioning of the brain, leading to distorted perceptions of reality. Individuals with schizophrenia may experience hallucinations, delusions, and difficulties in organizing their thoughts. The exact cause of schizophrenia remains elusive, involving a complex interplay of genetic, environmental, and neurobiological factors. Understanding the significance of the brain and the impact of conditions like schizophrenia is crucial for fostering empathy, eradicating stigma, and promoting advancements in mental health research and treatment.

The detection of schizophrenia can benefit significantly from advanced deep learning techniques. Leveraging the power of deep neural networks, models such as Inception, Xception, and Dense121 can play a crucial role in analyzing complex patterns within the data fed. These models are adept at capturing intricate details and hierarchies in the data, providing a nuanced understanding of brain structures and activities associated with schizophrenia. The application involves feeding EEG channel data into these deep learning architectures, allowing them to automatically learn and extract meaningful features. Inception, Xception, and Dense121, being state-of-the-art deep learning models, can discern subtle abnormalities in brain scans associated with schizophrenia

II. LITERATURE SURVEY

Recent advancements in machine learning (ML) and deep learning have prompted a surge of research aimed at enhancing the understanding and diagnosis of schizophrenia. In a study presented at the 2023 World Symposium on Digital Intelligence for Systems and Machines, the intricate interplay of clinical features and cognitive behavior in schizophrenia was explored through a machine learning lens [1]. A personalized mobile sensing-based supervised deep learning model was introduced for predicting psychotic relapse in schizophrenia patients, leveraging mobile data to enhance prediction accuracy and tailor interventions [2]. Another study in the same journal proposed a comprehensive evaluation approach for diagnosing schizophrenia, integrating various clinical factors to provide a holistic diagnostic perspective [3]. Social media data gathered through a Chrome extension was used for mental health analysis using deep learning in a study presented at the 4th International Conference for Emerging Technology [4]. Multi-modal deep learning, combining imaging genetics for schizophrenia classification, was the focus of a study presented at the 2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops [5].

Deep learning from imaging genetics for schizophrenia classification was explored in another study presented at the 2022 IEEE International Conference on Image Processing [6]. The effects of schizophrenia on electroencephalogram (EEG) data were examined using explainable deep learning models in a study presented at the 2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering [7]. ConvNet and machine learning models with feature engineering were utilized in another study for the classification of schizophrenia using motor activity data [8].

These studies collectively demonstrate the potential of machine learning and deep learning techniques in shedding light on schizophrenia's complexities.

III. PRELIMINARITIES OF THE RELATED WORK

The system proposed has three models that are trained with an approach of Supervised Learning.

Requirements of the model:

The Inception V3 model, a deep learning architecture, is designed for image classification tasks. It belongs to the family of convolutional neural networks (CNNs) and is particularly recognized for its efficiency in processing visual data. The model employs a sophisticated network structure with the utilization of inception modules, which involve the parallel execution of multiple convolutions with different kernel sizes. This approach allows the network to capture features at various scales, enhancing its ability to recognize patterns in images. Inception V3 is known for its versatility and robust performance, making it suitable for a wide range of computer vision applications. Its inception modules contribute to improved computational efficiency and parameter utilization, enabling effective feature extraction and representation in the hierarchical layers of the network.

Xception, short for "Extreme Inception," is a deep learning model designed for image classification tasks, belonging to the family of convolutional neural networks (CNNs). What distinguishes Xception is its departure from the traditional inception modules. Instead, it adopts a depthwise separable convolutional approach, which separates spatial and channel-wise operations. This design significantly reduces the number of parameters, enhancing computational efficiency while maintaining or even improving the model's representational power. The Xception architecture excels in capturing intricate patterns in images, making it well-suited for tasks requiring fine-grained feature recognition. Its depthwise separable convolutions contribute to improved generalization and training efficiency, positioning Xception as an asset in the realm of computer vision applications.

DenseNet-121 (Dense121) is a specialized deep learning model designed for image classification tasks, belonging to the family of densely connected convolutional networks (DenseNets). Its distinctive feature is the dense connectivity pattern, where each layer receives direct input from all preceding layers. This design promotes efficient parameter utilization, encourages feature reuse, and alleviates the vanishing gradient problem. Specifically comprising 121 layers, Dense121 demonstrates efficacy in capturing intricate patterns in images due to its densely connected blocks. The dense connectivity architecture facilitates improved gradient flow, enabling the model to efficiently learn and represent complex features. This model's unique approach enhances computational efficiency while maintaining high performance. Dense121 is particularly suitable for tasks requiring detailed feature recognition in computer vision applications. Its balance between depth, parameter efficiency, and representation power makes it a robust and versatile choice for image classification. In summary, DenseNet-121 stands out as a powerful and efficient deep learning model, offering a compelling solution for various computer vision challenges.

IV. METHODOLOGY

In the methodology employed for schizophrenia detection, Inception, Exception, and DenseNet121 models serve as critical components which works on Electroencephalogram (EEG) data. These state-of-the-art neural network architectures are applied to analyze intricate patterns and features present in the two-dimensional data taken.

i) Data Collection:

An approach is selected called EEG stands for Electroencephalography, which is a neurophysiological monitoring method to record electrical activity in the brain over time. The freely available data is taken from a public repository by Olejarczyk and Jernajczyk in which EEG data is collected which contains the measures of voltage fluctuations resulting from ionic current flows within the neurons of the brain. These electrical potentials are typically recorded from multiple electrodes placed on the scalp. Above 18 years old patients are included in this study with the diagnosis of F20.0 in the ICD-10 category, who did not take any drugs at least seven days before collecting the data for diagnosing. 18 years younger patients, pregnant patients, patients suffering from severe neurological diseases and brain disorders were excluded from the study. Males and female patients are collected equally. As a total of 28 recordings were gathered (14 were males and 14 were females). 14 Schizophrenic patients' data is collected in which 7 are males and 7 are females. And 14 control groups (normal and healthy) data is collected in which 7 are males and 7 are females. The average ages of these groups are noted in the table-1 shown below.

	Average age	
Patients	**Males**	**Females**
Normal and Healthy	26.8 ± 2.9	28.7 ± 3.4
Schizophrenic	27.9 ± 3.3	28.3 ± 4.1

Table.1.s – Average ages of patients

A total of 19 electrode channels are placed on the scalp of the patient based on 10-20 system for the extraction of EEG data of the patient. 7 electrode channels (Fp1, Fp2, F7, F3, Fz, F4, F8) in the Frontal Region, 5 electrode channels (T3, C3, Cz, C4, T4) in the Central region, 5 electrode channels (T5, P3, Pz, P4, T6) in the Parietal region, and 2 electrode channels (O1, O2) in the Occipital region are placed accordingly. The placement of these electrode channels are as shown in the fig.1 below.

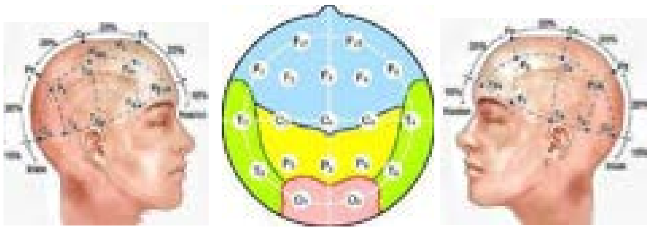


Fig.1. 10-20 EEG electrode channel placement

The recordings were observed for an average of 15 minutes while they were in the resting state with their eyes closed. Each patient contains the data collected from 19 electrode channels (i.e., Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2) at a sampling rate of 250 Hz using 10/20 EEG montage methodology.

ii) Preprocessing Data:

EEG data undergoes segmentation into smaller time epochs as a standard step in the data preprocessing pipeline. During this preprocessing phase, we executed baseline correction by subtracting the mean voltage recorded during a specific period from the entire EEG signal. This crucial step aimed to minimize variations that were unrelated to the experimental conditions, ensuring a more accurate representation of the underlying brain activity. In the recording of signals, the method of average referencing was employed, designating C3 as the reference electrode. This referencing approach played a pivotal role in establishing a standardized baseline across multiple electrodes. By doing so, it facilitated the precise interpretation of EEG signals in relation to the chosen reference point, contributing to a more robust and reliable analysis of the neural activity patterns.

iii) Spectrogram Formation:

A spectrogram is essentially a visual depiction that captures the dynamic interplay of frequencies within a signal over time that represents the spectrum of frequencies in a signal that are varying with time. They are often displayed as two-dimensional heatmaps that vividly showcases how the spectral content of the signal changes throughout its duration, where the x-axis represents time domain, the y-axis represents the frequency domain, and the color or intensity represents the strength (amplitude or power) of the signal at each time-frequency point. This visual portrayal is particularly useful for unraveling the temporal evolution of a signal's frequency spectrum.

In the proposed methodology, the spectrograms are created using Short-Time Fourier Transform (STFT) which is a used technique in signal processing, particularly in the analysis of EEG data which is a time varying signal as EEG signals are dynamic. EEG signals are time-varying, and the STFT allows one to study the evolution of the frequency content across various time intervals. The signal is sliced into short, overlapping segments, and for each segment, the frequency composition is analyzed through the Fourier Transform. This is crucial for recording fleeting trends and variances in brain activity, offering insightful information about diseases like schizophrenia. By localizing the analysis to brief signal segments, the employment of a window function in the time domain makes it easier to isolate study-relevant properties. All things considered, the STFT is helpful in exposing the time-varying spectrum properties of EEG signals, which advances our understanding of cerebral activity patterns.

The formula for the Short-Time Fourier Transform of a signal $x(t)$ is given by:

$$X(f,t) = \int_{-\infty}^{\infty} x(\tau) * \omega(\tau - t) * e^{-j2\pi f\tau}\, d\tau$$

where

- $X(f,t)$ is the STFT of the signal at frequency f and time t.
- $x(t)$ is the input EEG signal in the time domain.
- $\omega(\tau - t)$ is a window function centered at time t, which is applied to the signal to isolate a short segment.
- $e^{-j2\pi f\tau}$ is the complex exponential term representing the frequency content.

In the dataset under consideration, each patient is represented by a set of data corresponding to 19 electrode channels. The analysis involves the generation of spectrograms for each individual electrode pertaining to a specific person. A spectrogram, essentially a visual representation of the frequency spectrum's evolution over time, is created independently for each electrode. To illustrate, Figure-2 in the dataset provides a sample representation of such a spectrogram for one electrode.

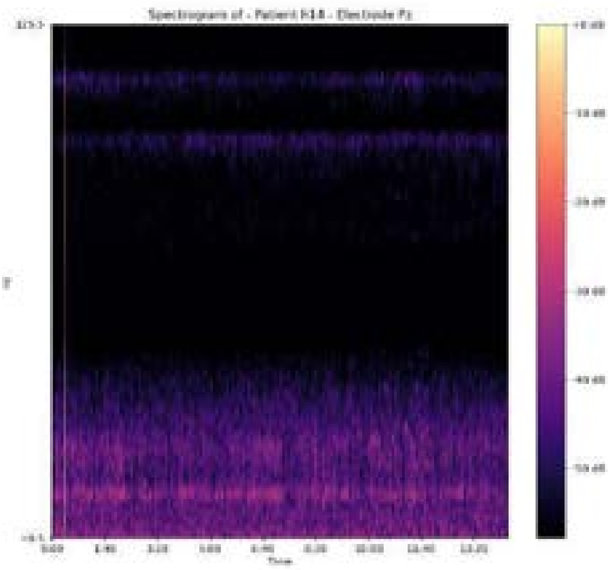


Fig.2. Sample spectrogram of Pz electrode of a healthy patient

The process involves the concatenation of these individual electrode spectrograms into a two-dimensional array. Specifically, the 19 electrodes are sequentially arranged in this array, one after the other. This arrangement forms a comprehensive visualization of the frequency content for each electrode, collectively capturing the unique electrical activity at each channel. By concatenating the spectrograms strictly and tightly in such a manner that gaps, axes and measurements aren't allowed in between them, the resulting 2D array allows for a holistic overview of the neural activity across all electrodes for a given patient are shown in the figures 5 and 6. This method enables researchers to observe and analyze potential patterns, trends, or anomalies in the spectral characteristics of the brain signals recorded from different regions. Such an approach is particularly valuable in exploring the nuanced details of brain activity and can facilitate the identification of

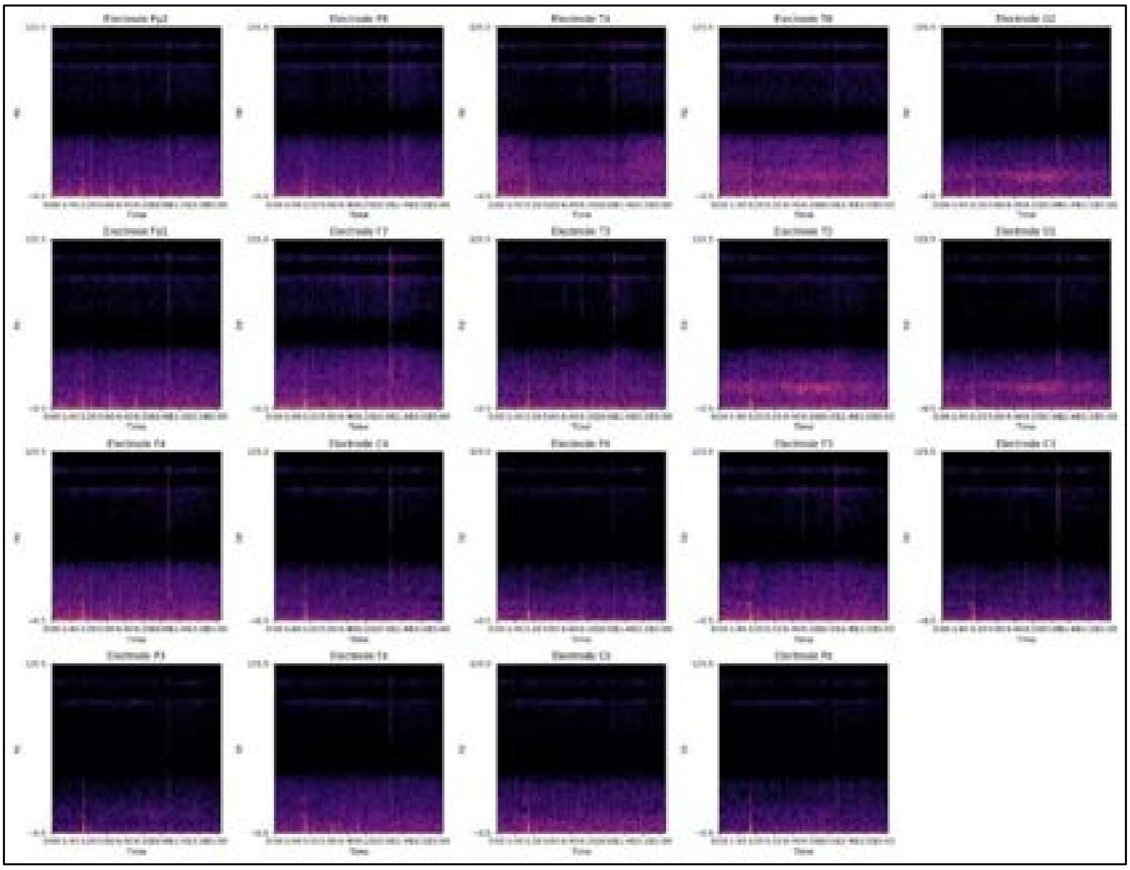


Fig.3. Spectrograms of 19 electrode channels of a Normal patient.

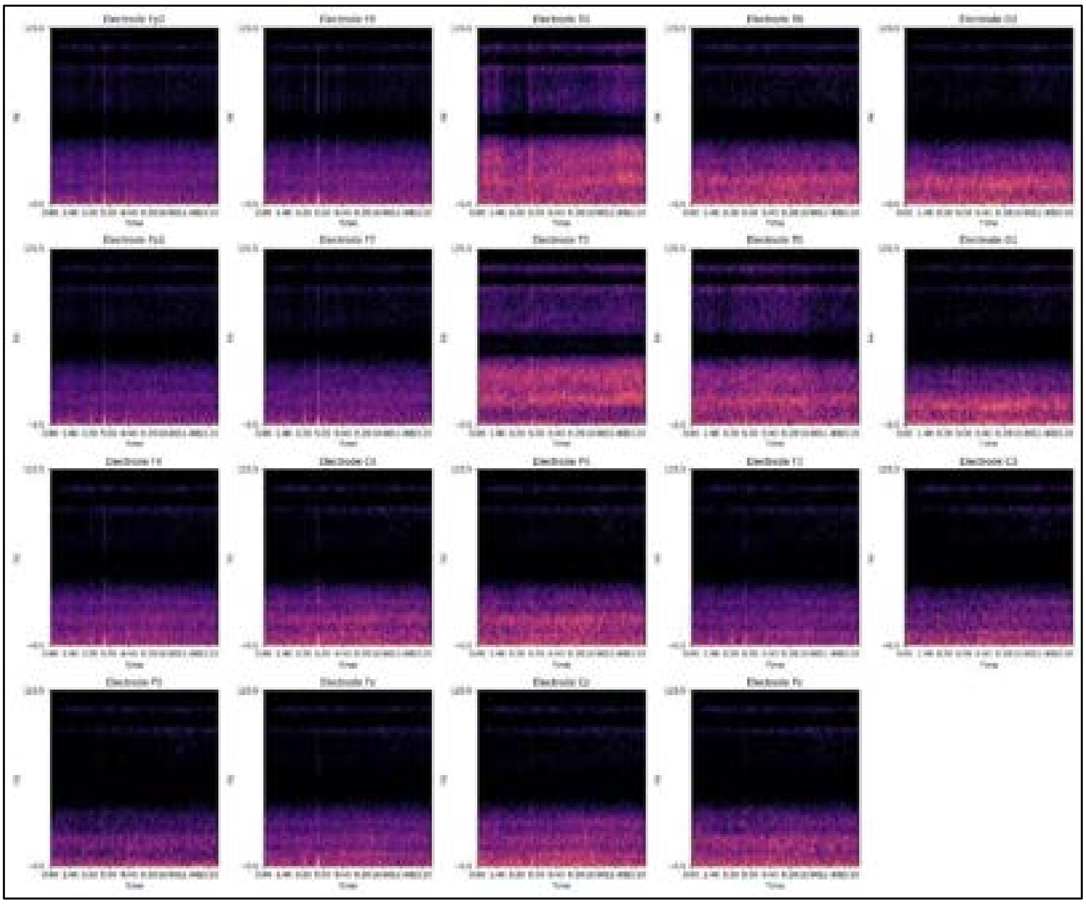


Fig.4. Spectrograms of 19 electrode channels of a Schizophrenic patient.

electrode-specific contributions to the overall neural dynamics within the dataset.

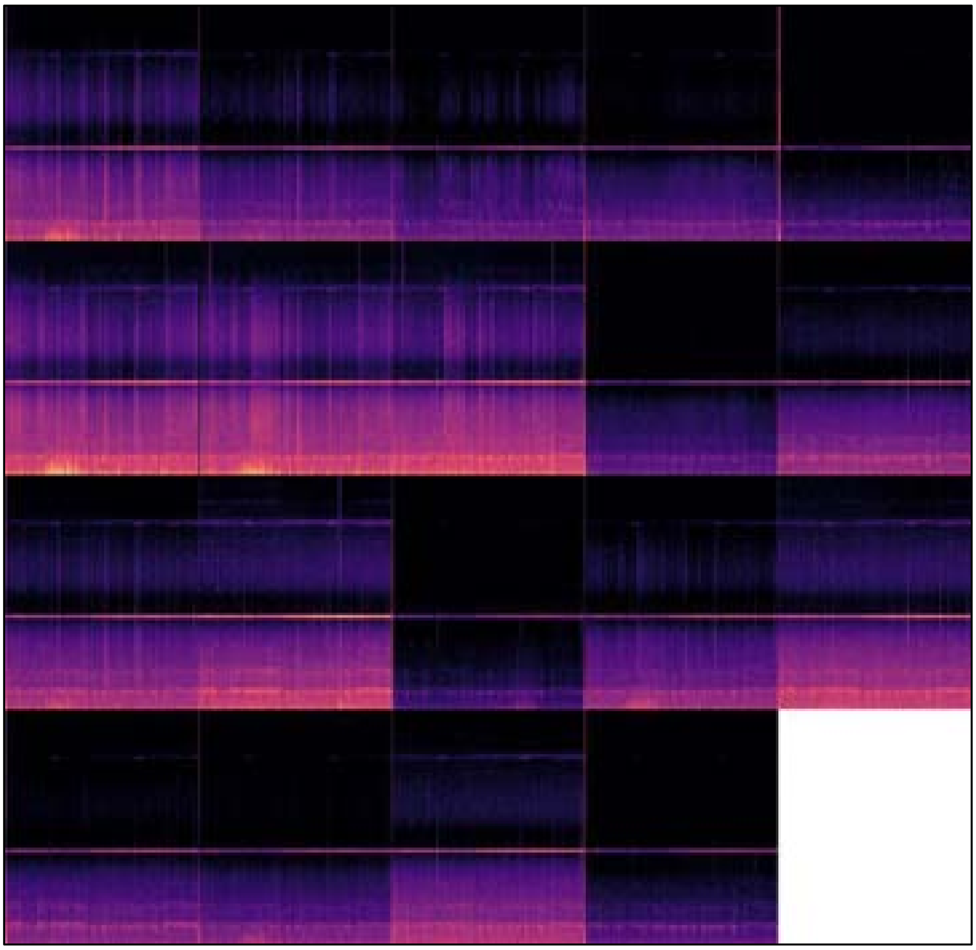


Fig.5 .Concatenated image of Spectrograms of 19 electrode channels of a Healthy patient.

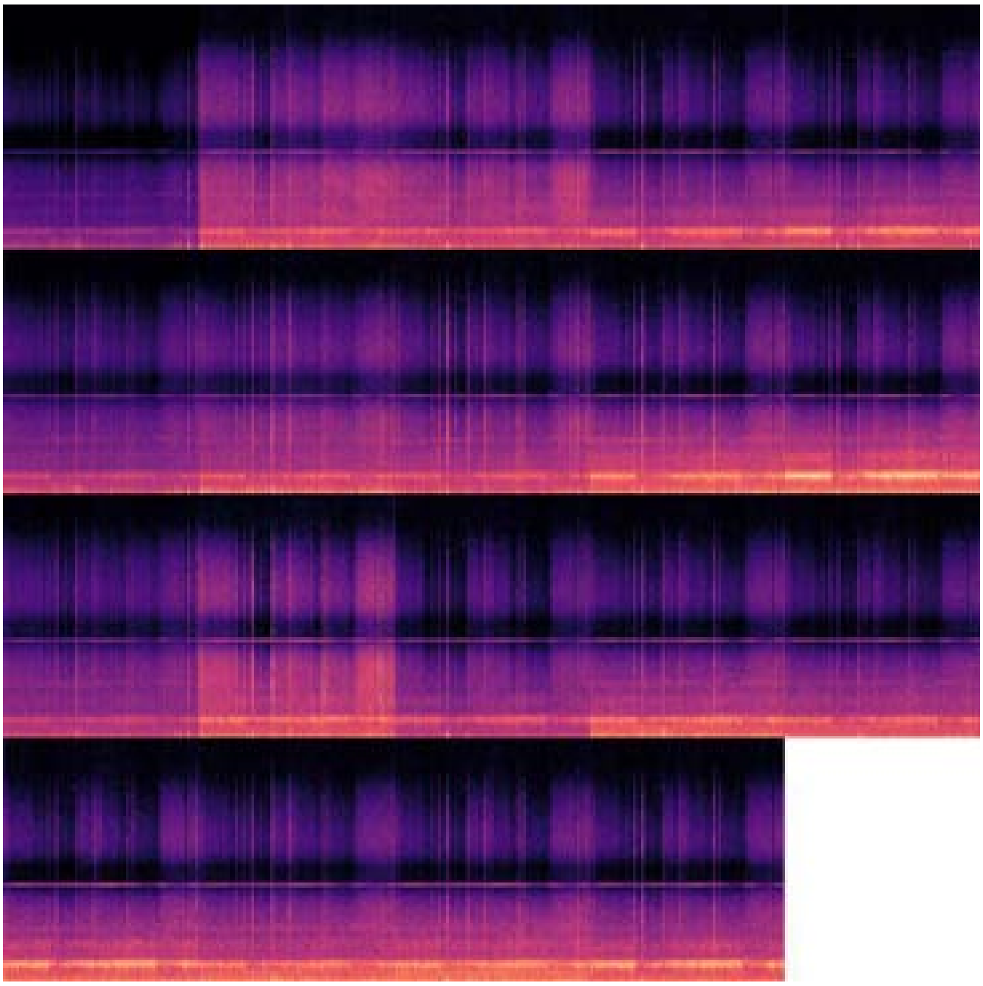


Fig.6. Concatenated image of Spectrograms of 19 electrode channels of a Schizophrenic patient.

iv) Transfer Learning Model Architectures

In our quest for effective schizophrenia detection using EEG data, we employed transfer learning with three prominent neural network architectures named — Inception V3, Xception, and DenseNet121. Transfer learning allows us to leverage the knowledge gained by a pre-trained neural network on a vast dataset and adapt it to a specific task. The intricate patterns and features present in the two-dimensional EEG data were analyzed by these state-of-the-art models. By fine-tuning these models, we harnessed its adaptability to discern complex patterns inherent in EEG signals, contributing to a nuanced understanding of neural activity associated with schizophrenia. Hence, we added a custom classification head to the existing model architectures, we tailored the models to the nuances of EEG data for schizophrenia detection.

Inception V3 Model:

Relevant hierarchical information from EEG signals was extracted, and the adaptability of Inception V3 model architecture allowed the discernment of complex patterns indicative of neural activity related to schizophrenia. Fine-tuning was carried out on the Inception model as shown in the fig.7 given below, acknowledged for its capability to capture multi-scale features through the utilization of inception modules.

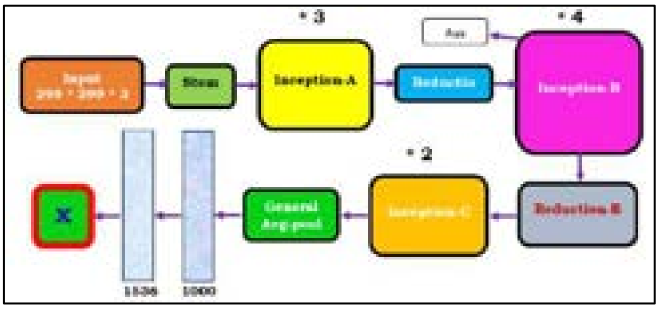


Fig.7 Fine-Tuning of Inception V3 model

In the above figure-7, X denotes the custom layers that are generated on fine-tuning of the Inception V3 model, serve as a custom classification head. These layers include flattening the output, introducing a dense layer with 512 units and ReLU activation, incorporating dropout for regularization, and concluding with a dense layer with 2 units and SoftMax activation for binary classification (schizophrenia or Normal). This tailored classification head allows the model to learn discriminative features specific to our EEG dataset.

Xception Model:

In our pursuit of advancing schizophrenia detection through deep learning methodologies, we employed transfer learning with the Xception architecture. Recognized for its proficiency in handling diverse and complex datasets, Xception served as the backbone for our model. The integration of a custom classification head, fine-tuning, and the introduction of a novel focal loss function aimed to tailor the Xception model to the intricate patterns present in EEG recordings, contributing to its effectiveness in identifying subtle abnormalities associated with schizophrenia.

Fine-tuning was applied to the Xception model as shown in the fig.8, recognized for its proficiency in handling diverse and complex data. This architecture was adapted to interpret nuanced patterns embedded in EEG recordings. The Xception model's capacity to adapt to irregularities in the data contributed to its effectiveness in identifying subtle abnormalities associated with schizophrenia.

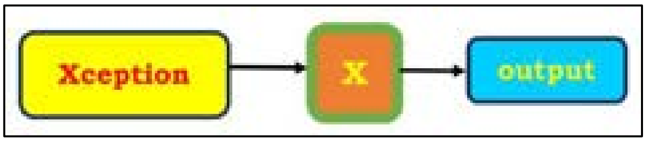


Fig.8 Fine-Tuning of Xception model

In the above figure-8, X denotes the custom layers that are generated on fine-tuning of the Xception model, serve as a custom classification head, which built on top of the Xception base includes a Global Average Pooling layer for dimensionality reduction, followed by two dense layers. The final dense layer with a sigmoid activation function enables binary classification (schizophrenia or non-schizophrenia). The model was compiled with the Adam optimizer and binary cross-entropy loss function. To enhance the model's sensitivity to rare events, we introduced a focal loss function. Unlike traditional binary cross-entropy, focal loss assigns higher weights to misclassified minority samples, focusing the model's attention on challenging instances. This approach is particularly relevant in the context of schizophrenia detection, where abnormalities may manifest subtly in EEG data.

Fine-tuning of the Xception model was a crucial step in adapting the architecture to the irregularities inherent in EEG recordings. By retraining the model to the learned features along with the added classification head, while allowing the model to specialize in the nuances specific to our dataset. The adaptability of Xception to diverse data types contributed to its effectiveness in capturing the complex patterns associated with schizophrenia.

DenseNet121 Model:

The DenseNet121 model, well-known for its dense connectivity patterns, was leveraged by performing adaptation to capture intricate dependencies within EEG data. Exploiting the densely connected layers, hierarchical representations of EEG features were effectively learned, enhancing sensitivity to subtle abnormalities characteristic of schizophrenia. As part of transfer learning, Fine-tuning is carried for DenseNet 121 as shown in the fig.9 given below.

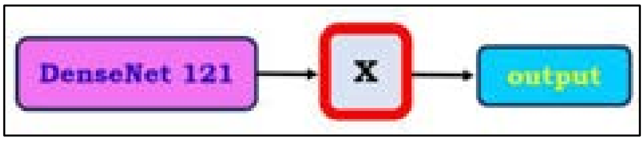


Fig.9 Fine-Tuning of DenseNet 121 model

In the above figure-9, X denotes the custom layers that are generated on fine-tuning of the DenseNet 121 model, serve as a custom classification head. Output layer is carried by it to obtain the proper appropriate results. Our custom model architecture, built on top of the DenseNet121 base, includes a Global Average Pooling layer for dimensionality reduction, followed by two dense layers. The final dense layer with a sigmoid activation function facilitates binary classification (schizophrenia or non-schizophrenia). The model was compiled using the Adam optimizer and binary cross-entropy loss function.

V. RESULTS AND DISCUSSIONS

In this study, our proposed models for schizophrenia detection through EEG data underwent a thorough evaluation, revealing commendable performance across various metrics. These results of the metrics underscore the efficacy of the model in discerning subtle patterns indicative of schizophrenia within the concatenated electrode spectrograms. The model's ability to generalize well to new data, thanks to transfer learning, is reflected in the high accuracy and suggests its robustness in real-world applications.

In our proposed concept, we depict the plots of accuracy and loss of training and validation datasets of the models as they play a pivotal role in identifying potential issues such as overfitting or underfitting, guiding adjustments to the model architecture, learning rate, or regularization techniques. The point of minimal validation loss becomes a key reference for selecting an optimal model checkpoint. As well as the robustness of the models is analyzed using accuracy (1), precision (2), recall (3), f1 score (4), specificity (5), and sensitivity (6).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \; (1)$$

$$Precision = \frac{TP}{TP + FP} \quad \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \; (2)$$

$$Recall = \frac{TP}{TP + FN} \quad \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \; (3)$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \; (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \; (5)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad \rightarrow\rightarrow\rightarrow\rightarrow\rightarrow\rightarrow \; (6)$$

Where

TP (True Positives)- Instances that are correctly predicted as positive by the model.
TN (True Negatives)- Instances that are correctly predicted as negative by the model.
FP (False Positives)- Instances that are incorrectly predicted as positive by the model when they are actually negative.
FN (False Negatives)- Instances that are incorrectly predicted as negative by the model when they are actually negative.

The results of performance metrics are as shown below.

Inception V3

The training and validation accuracy and loss plots are depicted and shown below in Fig. 10a and 10b.

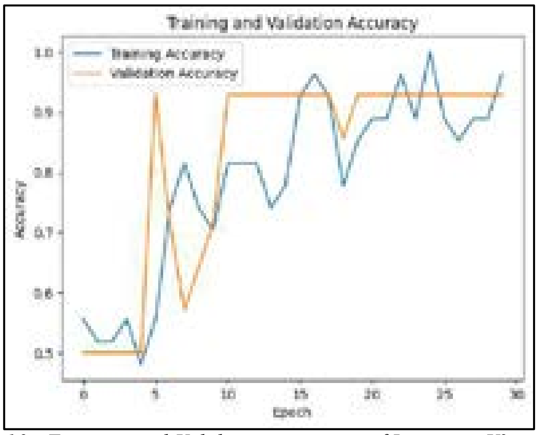


Fig.10. Training and Validation accuracy of Inception V3 model

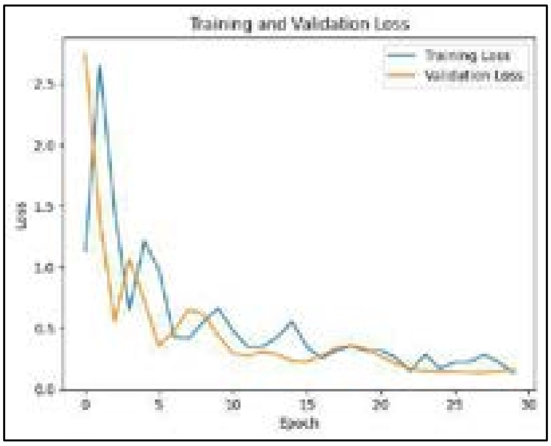


Fig.11. Training and Validation loss of Inception V3 model

The Inception v3 model demonstrated impressive accuracy, achieving a score of 96.15%. This indicates minimal error, with only 3.85% of predictions being incorrect. Examining the loss function further reveals a promising trend. Training loss steadily decreased, suggesting effective learning. However, validation loss remained relatively constant, hinting at potential overfitting to the training data. This suggests further adjustments or regularization techniques may be beneficial for improved generalizability.

A deeper dive into performance metrics paints a positive picture. The model attained exceptional precision (96.45%) and recall (96.15%), signifying its ability to accurately identify both true positives and avoid false positives. F1 score of 96.16 further solidifies this balanced performance. Notably, specificity reached a perfect 1.0000, indicating flawless identification of true negatives. However, a sensitivity of 0.9286 suggests room for improvement in capturing a small portion of true positives. This report is given below in Fig. 10(c) This metric could be prioritized in future optimization efforts.

```
Accuracy: 0.9583
Precision: 0.9615
Recall: 0.9583
F1 Score: 0.9583
Specificity: 1.0000
Sensitivity: 0.9167
```

Fig.12. Performance metrics of Inception V3

Xception

The Xception model displayed impressive accuracy on the task at hand as shown in fig.11(a), showcasing the model's adeptness at making correct predictions. However, analyzing the loss function reveals a potential area for improvement. While the training loss steadily decreased, indicating effective learning, the validation loss remained relatively constant as shown in fig.11(b).

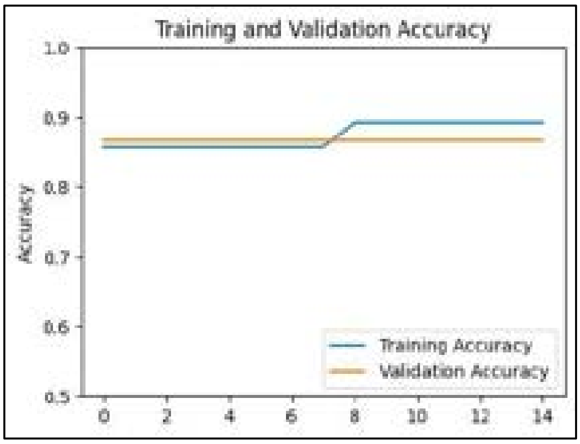


Fig.13. Training and Validation accuracy of Xception model

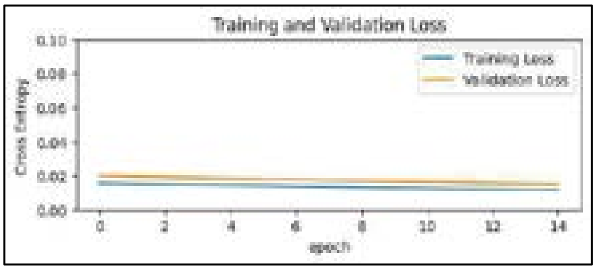


Fig.14. Training and Validation loss of Xception model

The model appears to be performing well, with an accuracy of 93.33%, precision of 94.17%, recall of 93.33%, F1 score of 93.33%, specificity of 87.50%, and sensitivity of 1.0000. This means that the model is correctly classifying about 93% of the images it is shown, and it is particularly good at identifying true positives (sensitivity of 100%). However, the specificity of 87.50% suggests that the model may be making some false positive errors. The report of Xception is shown in fig.11(c).

```
Accuracy: 0.9333
Precision: 0.9417
Recall: 0.9333
F1 Score: 0.9333
Specificity: 0.8750
Sensitivity: 1.0000
```

Fig. 15. Performance metrics of Xception model

DenseNet 121

DenseNet 121's accuracy plot is drawn below in fig.12(a) is impressive, but addressing the gap between training and validation performance could unlock its full potential. By striking a balance between learning from the known and adapting to the unknown, DenseNet 121 can evolve into an even more versatile tool for tackling diverse challenges. Peering into the loss function sheds further light. The decreasing training loss signifies effective learning, as the model refines its predictions shown below in fig.12(b).

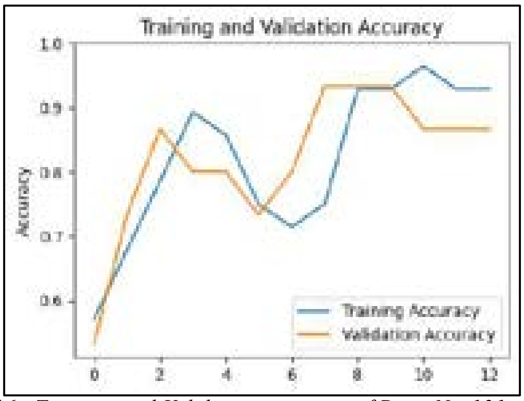


Fig.16. Training and Validation accuracy of DenseNet 121 model

Fig.17. Training and Validation loss of DenseNet 121 model

DenseNet 121 demonstrates strong performance with its high accuracy, precision, recall, and F1 score. The model boasts an impressive accuracy of 95.83%, indicating its adeptness. Delving deeper into the metrics unveils the potential area of improvement. The detailed report is shown in the fig.12(c).

```
Accuracy: 0.9615
Precision: 0.9645
Recall: 0.9615
F1 Score: 0.9616
Specificity: 1.0000
Sensitivity: 0.9286
```

Fig.18. Performance metrics of DenseNet 121 model

The above models can detect schizophrenia through the concatenated spectrograms and are performed to their best. The obtained results

Each model shines in its own domain and demonstrated unique strengths in capturing intricate patterns within the neural activity recorded by 19 electrode channels. Inception V3 holds the crown for balanced performance and near-perfect specificity, while Xception excels at sensitivity. DenseNet 121, with its raw learning power and sophisticated architecture, holds immense potential once it conquers the challenges of learning and identifying data. The performance of Inception V3, Xception, DenseNet 121 is compared in the table.2. Inception V3, renowned for its multi-scale feature extraction and adaptability, showcased exceptional accuracy, precision, recall, and specificity, underscoring its effectiveness in discerning subtle abnormalities indicative of schizophrenia. Xception, leveraging its proficiency in handling complex data, yielded remarkable results, showcasing its adaptability to the nuances of EEG recordings. Meanwhile, DenseNet121, with its dense connectivity patterns, achieved comparable performance.

	Inception V3	Xception	DenseNet 121
Accuracy	95.83	93.33	96.15
Precision	96.15	94.17	96.45
Recall	95.83	93.33	96.15
F1 Score	95.83	93.33	96.86
Specificity	100	87.50	100
Sensitivity	91.67	100	92.86

Table.2. Performance of models

The overall comparison of the three models' accuracies and losses are plotted below in fig 13 and fig 14.

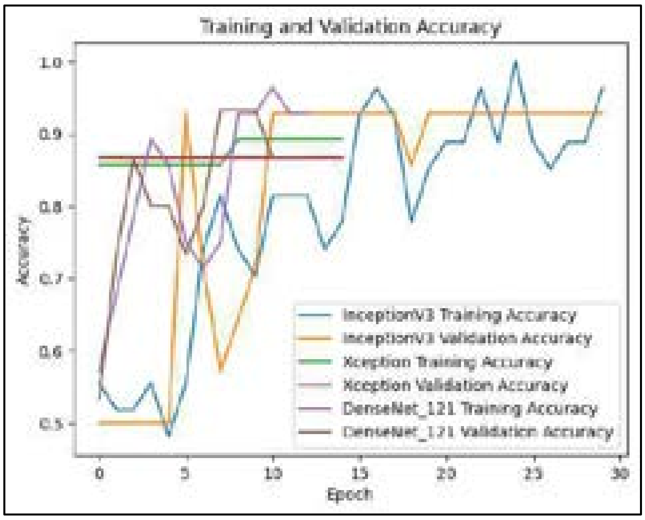


Fig.19. Training and Validation accuracies of three models

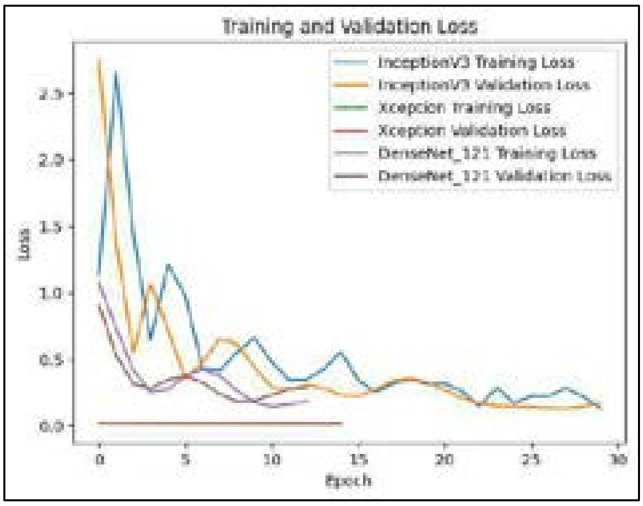


Fig.20. Training and Validation losses of three models

Each model's loss function decreased with each epoch in a non-linear fashion for the inceptionV3 and Densenet 121 models, and in a slightly linear fashion for the Xception model. The training and validation accuracies and losses of Inception V3 and DenseNet 121 are quite closure. There is a small gap is visible between the accuracies of training and validation of Xception model which indicates that it observes a particular pattern to detect the Schizophrenia. Hence, we can say that the adaptability of Xception model is a little bit better than Inception V3 and Densenet 121 models.

The nuanced insights gained from the performance metrics and training dynamics enhance our understanding of the intricate relationships within EEG data. The model architectures have proven to be a potent strategy in the realm of schizophrenia detection. With careful analysis and strategic training, these models can work in tandem.

VI. CONCLUSION

Utilization of transfer learning with the model architectures has proven to be a formidable strategy in the realm of schizophrenia detection. The model's ability to extract hierarchical information from EEG signals, combined with its fine-tuned custom classification head, has resulted in exceptional performance. This study presents a pioneering exploration of multimodal analysis in schizophrenia detection through the concatenation of electrode spectrograms and application to state-of-the-art deep learning architectures. Our findings underscore the potential of this approach in uncovering intricate neural signatures associated with schizophrenia, providing a foundation for further research and the development of advanced diagnostic tools. It is demonstrated that the models InceptionV3, Xception and DenseNet 121 performed well in detecting schizophrenia. Notably, these models exhibited superior adaptability in which Xception model leads. The complex architecture and heavier depth in the layers of these models make easier of identifying different patterns in the spectrograms leads to the efficient detection of the Schizophrenia. The effectiveness of transfer learning techniques was apparent, although a consistent, steady transition in validation and training accuracies throughout each epoch was not assured. This variability is attributed to the diverse nature of the dataset and the iterative parameter updates at each epoch. The efficacy demonstrated by these models in our investigation underscores the importance of employing deep learning architectures renowned for their proficiency in capturing hierarchical features. The adaptability and discernment capabilities exhibited by the model in analyzing intricate patterns within EEG data significantly contribute to the advancement of schizophrenia detection methodologies. The attainment of high levels of accuracy further underscores the potential clinical significance of our approach, thereby presenting a promising avenue for facilitating early diagnosis and intervention in the context of schizophrenia.

REFERENCES

[1] F. Adjailia, S. A. Hadri and A. Bbouramoul, "Understanding the Interplay of Clinical Features and Cognitive Behavior in Schizophrenia: A Machine Learning Study," *2023 World Symposium on Digital Intelligence for Systems and Machines (DISA)*, Košice, Slovakia, 2023, pp. 10-15, doi: 10.1109/DISA59116.2023.10308933.

[2] Dong-Woo Ko and Jung-Jin Yang, " EEG- Based Schizophrenia Diagnosis through Time Series Image Conversion and Deep learning" in MDPI- Electronics 2022, 11, 2265. Doi.org/10.3390/etectronics11142265.

[3] B. Lamichhane, J. Zhou and A. Sano, "Psychotic Relapse Prediction in Schizophrenia Patients Using A Personalized Mobile Sensing-Based Supervised Deep Learning Model," in *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 7, pp. 3246-3257, July 2023, doi: 10.1109/JBHI.2023.3265684.

[4] M. Tanveer, J. Jangir, M. A. Ganaie, I. Beheshti, M. Tabish and N. Chhabra, "Diagnosis of Schizophrenia: A Comprehensive Evaluation," in *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 3, pp. 1185-1192, March 2023, doi: 10.1109/JBHI.2022.3168357.

[5] M. Joshi, C. Mahajan, T. Korgaonkar, N. Raul and M. Naik, "Mental Health Analysis using Deep Learning of Social Media Data gathered using Chrome Extension," *2023 4th International Conference for Emerging Technology (INCET)*, Belgaum, India, 2023, pp. 1-8, doi: 10.1109/INCET57972.2023.10170255.

[6] E. Aydemir, S. Dogan, M. Baygin, Chui ping Ooi, Prabal Datta barua, Turker Tuncer, and U. Rajendra Acharya, " Automated Schizophrenia Detection Based on a Cyclic Group of Prime Order patterns Using EEG Signals" in PMC 2022 Healtcare (Basel), April; 10(4): 643. Doi: 10.3390/healthcare10040643.

[7] A. Kanyal, S. Kandula, V. Calhoun and D. H. Ye, "Multi-Modal Deep Learning on Imaging Genetics for Schizophrenia Classification," *2023 IEEE International Conference on Acoustics, Speech, and Signal Processing Workshops (ICASSPW)*, Rhodes Island, Greece, 2023, pp. 1-5, doi: 10.1109/ICASSPW59220.2023.10193352.

[8] H. Yu, T. Florian, V. Calhoun and D. H. Ye, "Deep Learning from Imaging Genetics for Schizophrenia Classification," *2022 IEEE International Conference on Image Processing (ICIP)*, Bordeaux, France, 2022, pp. 3291-3295, doi: 10.1109/ICIP46576.2022.9897977.

[9] C. A. Ellis, A. Sattiraju, R. Miller and V. Calhoun, "Examining Effects of Schizophrenia on EEG with Explainable Deep Learning Models," *2022 IEEE 22nd International Conference on Bioinformatics and Bioengineering (BIBE)*, Taichung, Taiwan, 2022, pp. 301-304, doi: 10.1109/BIBE55377.2022.00068.

[10] F. P. Ferreira and A. Daly, "ConvNet and machine learning models with feature engineering using motor activity data for schizophrenia classification," *2022 IEEE 35th International Symposium on Computer-Based Medical Systems (CBMS)*, Shenzen, China, 2022, pp. 223-227, doi: 10.1109/CBMS55023.2022.00046.

[11] Zulfikar Aslan, Mehmat Akin, "Automatic detection of schizophrenia by applying deep learning over spectrogram images of EEG signals". 2020 *Trait. Du Signal*. 2020;37:235–244. doi: 10.18280/ts.370209.

[12] Chandran A.N., Sreekumar K., Subha D.P, "EEG-based automated detection of schizophrenia using long short-term memory (LSTM) network"; in *Advances in Machine Learning and Computational Intelligence*. Springer; Berlin/Heidelberg, Germany: 2021. pp. 229–236

[13] Jernajczyk W., Olejarczyk E., "EEG in Schizophrenia". *RepOD*. 2017.

Connected Parking Innovations: Enhancing Urban Mobility through IoT Integration

[1]Gnana Priya G
Department of Electronics and Communication Engineering,
Ramco Institute of Technology
Rajapalayam, India
ggpriya2019@gmail.com

[2]Akila M
Department of Electronics and Communication Engineering,
Ramco Institute of Technology
Rajapalayam, India
akilamohan2003@gmail.com

[3]Anusuya G
Department of Electronics and Communication Engineering,
Ramco Institute of Technology
Rajapalayam, India
anukabi2002@gmail.com

***Abstract*—This research work introduces a smart parking system leveraging Internet of Things (IoT) technology to address urban parking challenges. It utilizes sensors to detect parking slot availability and communicates this data via Wi-Fi to an IoT platform accessible globally. Through this system, users can remotely access real-time parking availability information from anywhere, facilitating efficient parking space utilization. The implementation involves Raspberry Pi interfacing with sensors to detect parking status and transmitting this information to the IoT platform using HTTP. The platform displays the data, enabling users to identify vacant parking slots conveniently. This solution enhances urban mobility by reducing time spent searching for parking, thus alleviating traffic congestion and environmental impact associated with vehicle emissions. Additionally, it demonstrates the practical application of IoT in addressing everyday challenges, showcasing the potential for technology to streamline urban infrastructure and improve the quality of life.**

Keywords— Smart parking system, Internet of Things (IoT), sensors, Wireless Fidelity (Wi-Fi) module, Raspberry Pi, Hypertext Transfer Protocol (HTTP), urban mobility, traffic congestion, environmental impacts.

I. INTRODUCTION

Recent years of population growth correspondingly increases the amount of traffic congestion. The use of cars has increased in proportion to the population. There was traffic congestion on the road as a result of increased car usage. Also, it takes longer time to find a free parking space. In order to locate the empty parking space in the parking field, we therefore lose a certain amount of time and waste more than 75 - 85 percent of the fuel. In order to address this issue, the parking lot needs a specialized system that measures empty space and displays the data to individuals searching for it.

In smart parking system operates on a streamlined process is facilitated by various components. Initially, users book parking slots through an IoT-enabled webpage, which also provides real-time availability updates. Upon arrival, RFID tags are used for identification, triggering the entry gate servo motor to open after validation. Infrared sensors detect parked cars, initiating a timer to record parking duration. If a vehicle fail to arrive within the allotted time, bookings are automatically cancelled. Payment for parking is processed through the webpage, with Twilio sending SMS reminders for payment. Failure to pay results in the exit gate remaining closed, while successful payment prompts it to open, allowing vehicles to exit. This integrated system optimizes parking management by efficiently utilizing available space and automating access control and payment processes. The Raspberry Pi Zero 2W serves as the central processor, storing and executing the system's program, while the IoT technology enables remote monitoring and interaction with the system via the internet. Overall, the system enhances user experience, reduces congestion, and improves overall efficiency in parking management.

II. LITERATURE SURVEY

A navigation and reservation based smart parking platform using genetic optimization for smart cities [1] proposed a location and reservation-based parking system. The suggested approach includes developing tiny gadgets that use Internet of Things (IoT) to transmit data to the internet. A genetic algorithm locates the nearby free parking field to the user area. A number of cases are evaluated with the suggested strategy, and precise results are obtained.

In order to prevent traffic congestion, Smart Parking System to Reduce Traffic Congestion[2] suggested the idea of a reservation-based smart parking system. Drivers may park at the designated spot without having to seek because the system will intelligently direct them there. GPS technology aids in directing the car to the closest parking space. Drivers can choose the best spot by using the Graphical Interface, which displays both reserved and open spaces. The notification will appear if the user's time is longer than the time restriction..

Automatic Parking Space Detection and Tracking for Underground and Indoor Environment [3], suggested a technique that uses combination of sensors to identify and parking spaces in indoor and underground environments. The parking field marking is done by employing parallel line pairs. Then free slots are discovered via making use of ultrasonic sensors. Low-level integration of AVM pictures and sensor detects the pillars. the user experiences greater comfort and convenience from the identified parking spots. The primary drawback of this research is the difficulty in detecting parking fields for vehicles due to the usage of cameras on every side of the vehicle. Predicting empty slot and keeping track of the empty slot requires more time.

A Prototype Parking System using Wireless Sensor Networks [4], introduces a strategy to identify the parking space using Closed-circuit television (CCTV) cameras. With the help of sensors, the CCTVs sends the parking slot images to the ARM microcontroller which in turn process the received data. The processed data is retransmitted to the centralized system via zigbee and it is stored in the database. Here the retransmission process takes more time and there is no common display to the user which is a drawback of this system.

In Perceptive Car Parking Booking System With IOT Technology[5], by using an Android application and internet access, the person visits the parking area to locate available spaces for parking and reserves one according to their preferences. The user may reserve their own likely parking space according to the time and cost using Multi-Processing Queuing Mechanism (MPQM).

. IoT-based economic smart vehicle parking system [6], exposes parking problems in cities because there are too many vehicles and not enough space to park them. To solve this, a new system using technology called the Internet of Things (IoT) is proposed. It uses devices like Arduino Uno, which helps connect digital and physical objects. Each parking slot has an infrared sensor to detect if it's vacant or not. Users can book a parking slot in advance through a website, and they have their username and password. If there's any misuse of the system, it will alert the responsible person.

To maximize parking resource usage and minimize the overall financial cost, iParker—A New Smart Car-Parking System Based on Dynamic Resource Allocation and Pricing [7], uses Mixed-Integer Linear Programming (MILP) application. It integrates Share Time Reservations (STR) and Real Time Reservations (RTR) by allowing a parker to reserve a slot several days in advance and also while traveling a few minutes away from the parking field. The pricing policies are also proposed which maximizes the profit.

.Dynamic Reservation of Edge Servers via Deep Reinforcement Learning for Connected Vehicles [8], introduces a new parking system that uses both visible light and radio frequency technologies to automate parking processes. It uses vehicle headlights and COB bulb-based transmitters for visible light communication (VLC), along with optical receivers on parking gate barriers and vehicle roofs. Digital pulse interval modulation (DPIM) is used for communication. The system includes software for client-server communication and processing. Two-factor authentication and payment systems are also implemented. The results showed successful communication with low error rates and accurate positioning of vehicles.

Innovation Architecture Smart Car Parking System with Wireless Sensor Networks [9], Proposes that with the increasing number of cars each year, parking areas are becoming limited, leading to issues such as traffic congestion, environmental pollution, and ecosystem disruption. To address this problem, a new system utilizing processor boards and wireless sensor networks is proposed. This system aims to optimize parking space usage in towns by efficiently guiding drivers to available parking spots, reducing the need for cars to roam in search of parking. By minimizing traffic congestion and fuel wastage, this system can help alleviate the strain on infrastructure and promote a more sustainable environment.

A Hybrid VLC/RF Parking Automation System [10], describes a new parking system that uses Visible Light Communication (VLC) and Radio Frequency (RF) wireless technologies to automate parking. It works by having the vehicle's headlights communicate with receivers installed on parking gates and vehicles. Digital Pulse Interval Modulation (DPIM) is used for this communication. The system also includes a client-server interface and software for managing parking. Two-factor authentication and payment systems are implemented at the gate barrier using VLC. The system is tested under different conditions to evaluate its performance in terms of communication range and accuracy of vehicle positioning.

III. PROPOSED SYSTEM

A. *Block Diagram of the Proposed System*

Figure 1 shows the model of the proposed system.

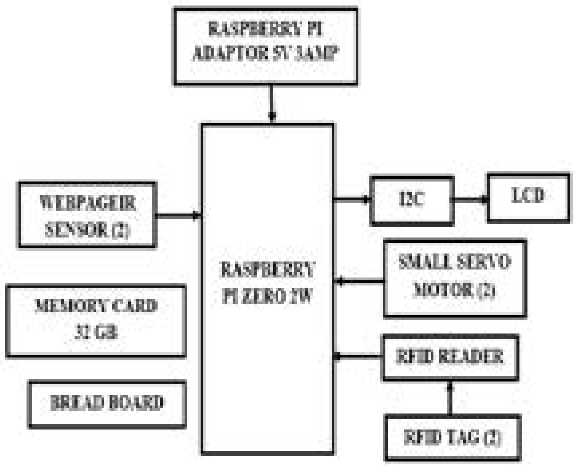


Fig. 1 Proposed System Model

B. *Working Model*

The smart parking system operates on a streamlined principle integrating various technologies for efficient parking management. At its core, the Raspberry Pi acts as the central processing unit, hosting the system's software and coordinating its functions. Raspberry Pi evaluates the data from sensors to determine the availability of parking spaces. The process begins with users booking parking slots via an IoT-enabled webpage using HTTP, where real-time slot availability is displayed. Upon booking, a designated time frame is allocated, and failure to arrive within this period automatically cancels the reservation. Upon arrival, users authenticate themselves by tapping an RFID tag on a reader. This triggers an identification (ID) verification process, allowing entry if the credentials are valid. Infrared sensors installed within parking slots detect vehicle presence, initiating the recording of parking duration using time-tracking mechanisms. Simultaneously, entry gate servo motors facilitate seamless access for authorized vehicles. After parking, users are prompted to pay for their parking duration through the webpage. Payment reminders are sent via SMS using Twilio integration, ensuring timely transactions. Failure to pay within the stipulated time results in the exit gate servo motor remaining closed, preventing vehicle egress. Conversely, successful payment prompts the gate to open, allowing vehicles to exit the premises. This integrated approach optimizes parking space utilization, enhances user convenience, and facilitates efficient traffic flow within parking facilities. By leveraging IoT, RFID technology, and servo motor control, the system offers a comprehensive solution to urban parking challenges, catering to the evolving needs of modern cities and their inhabitants.

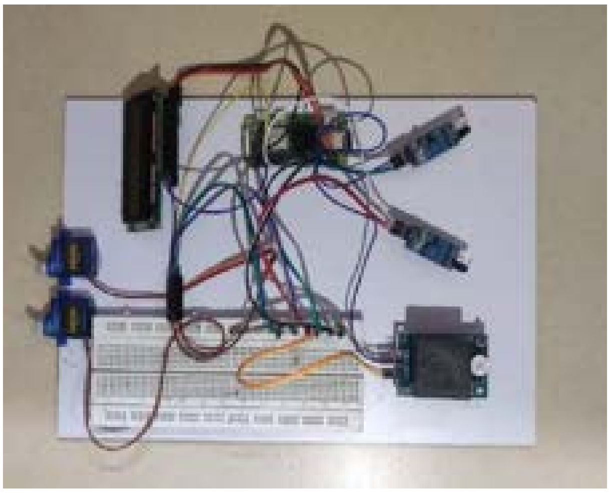


Fig.2 Hardware Setup

C. Modules

The following modules are used in the proposed models.

- IoT-Enabled Booking Module
- Sensor-Based Entry/Exit Module

D. IoT-Enabled Booking Module

The following Fig.3 depicts the IoT enabled booking module.

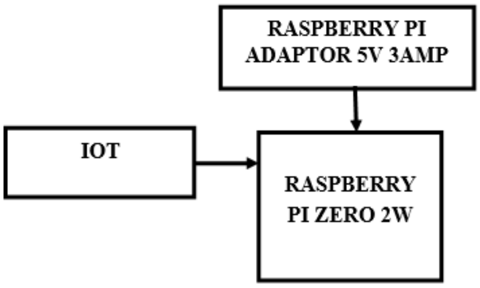


Fig.3 IoT enabled booking module

This module facilitates the booking process through an IoT-enabled webpage where IoT enhances efficient parking and slot reservation by enabling statistical analysis of data, real-time monitoring, system compatibility, improved user experience and remote management. Users can access this webpage from any internet-connected device to reserve a parking slot. The module manages the booking process, allowing users to select available slots, specify the duration of their parking, and complete the reservation. It also includes features such as automatic cancellation of bookings if vehicles fail to arrive within the specified time frame. This module ensures a user-friendly interface for booking slots and efficiently manages the reservation process.

E. Sensor-Based Entry/Exit Module

This module comprises various sensors and servo motors to manage vehicle entry and exit from the parking facility. Upon arrival, users authenticate themselves by tapping an RFID tag on the reader. This triggers an ID verification process, followed by the entry gate servo motor opening to allow authorized vehicles to enter. When a vehicle is parked, infrared sensors placed within the slot detect its presence. This allows time tracking to be used to record how long the vehicle is parked. Payment for parking is facilitated through the webpage, with Twilio sending payment reminders via SMS. Failure to pay results in the exit gate servo motor remaining closed, while successful payment prompts it to open, allowing vehicles to exit. This module ensures smooth vehicle movement within the parking facility, integrates payment processing, and maintains access control for efficient parking management. The following Fig.4 shows the sensor based Entry/Exit Module.

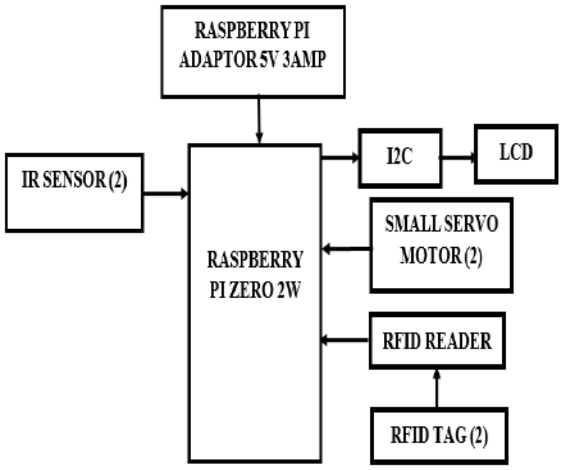


Fig.4 Sensor based Entry/Exit Module

Managing the reservation process involves various steps.

User Registration: Users must register their name, contact information and vehicle details with the system prior to the slot reservation.

Check for Slot Availability: Once registration is done, the current status of the parking slot is displayed in the webpage where the users can check the available parking spaces. Based on the availability, reservation request can be made.

Reservation and Payment: Once the reservation request is received, the chosen slot's availability for the allotted amount of time is verified by the system. If the chosen slot is available, the system reserves the slot and calculates the amount based on the parking duration.

Slot confirmation: After successful payment through integrated payment gateway, the system generates a reservation confirmation that includes details such as the parking slot number, date, reservation duration and time. This confirmation is sent to the user via Twilio SMS can also be accessed through the user's account on the system.

Expiration and Cancellation: The system may release the reservation made by the users and make the slot accessible to other users if the user does not come within a particular period after the reservation time. Users may also have the option to cancel or modify their reservations before the reservation time starts.

IV. RESULTS & DISCUSSION

The following Fig.5 shows the proposed hardware system final setup.

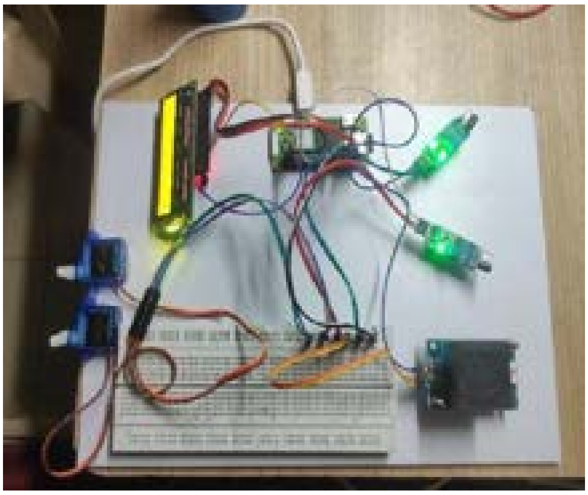


Fig.5 Proposed system hardware final setup

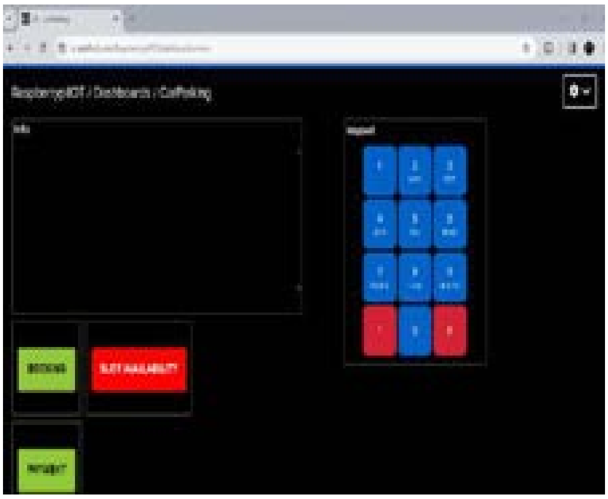


Fig.6 Webpage Setup

In this smart parking system, the Raspberry Pi serves as the central processor, storing and executing the system's program. Booking a parking slot is initiated via an IoT-enabled webpage as shown in Fig.6, followed by checking slot availability as shown in Fig.7 through the same interface. Bookings are automatically cancelled if cars fail to arrive within the specified time frame. After their IDs are verified, users tap an RFID tag on the reader to activate the entry gate servo motor as shown in Fig.9. Infrared sensors detect car presence in slots, with parking duration recorded using time tracking. Parking payments are made via the website, and Twilio provides SMS reminders for payments. as depicted in Fig.8. When payment is made successfully, the exit gate servo motor opens, allowing vehicles to leave, but when payment is not made, it remains closed. This system streamlines parking management, ensuring efficient space utilization while integrating seamless payment and access control mechanisms for user convenience.

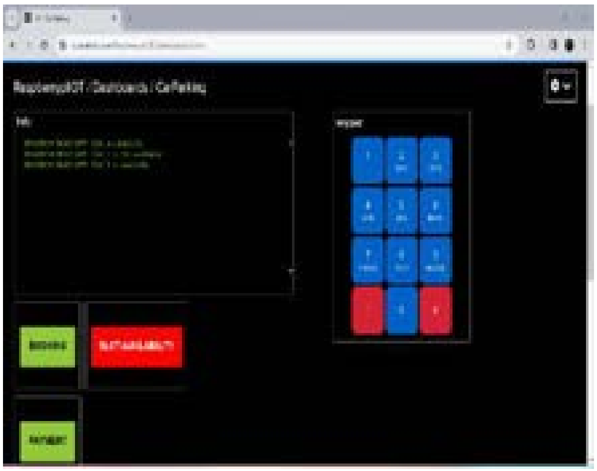


Fig.7 Available Slots

Fig.8 Twilio sending payment reminders via SMS

Fig.9 Slot 2 is booked

V. CONCLUSION & FUTURE SCOPE

The implementation of this smart parking system demonstrates a significant step towards addressing the challenges of urban parking management through the integration of IoT technology, sensor-based detection, and efficient data processing. By utilizing the Raspberry Pi as the central processor and leveraging IoT-enabled webpages for booking and payment processes, the system streamlines parking operations and enhances user convenience. The integration of RFID authentication, infrared sensors, and servo motors ensures seamless entry and exit control while accurately tracking parking duration. Furthermore, the incorporation of Twilio for payment reminders enhances the system's functionality and user engagement. Looking ahead, future enhancements could involve the integration of machine learning algorithms for predictive parking availability and the adoption of renewable energy sources for sustainability. Overall, this smart parking system represents a scalable and adaptable solution for optimizing parking space utilization, reducing congestion, and improving urban mobility. Its successful implementation underscores the potential of technology-driven innovations in addressing everyday urban challenges, ultimately contributing to more efficient, sustainable, and user-friendly urban environments.

Future enhancements for this smart parking system could include integrating advanced machine learning algorithms for predictive parking availability, optimizing parking space allocation based on historical data and real-time traffic patterns. Additionally, implementing automated parking guidance systems using sensors and actuators to direct drivers to available parking spots efficiently. Integration with smart navigation apps could provide users with real-time directions to the nearest available parking slots. Moreover, incorporating renewable energy sources like solar panels for powering the system reduces environmental impact and enhances sustainability. These advancements would further streamline parking operations, improve user experience, and contribute to smarter, greener urban environments.

ACKNOWLEDGMENT

We are grateful to Ramco Institute of Technology for their role in conceptualizing our ideas and for generously providing all essential resources and assistance.

REFERENCES

[1] Aydin, M. Karakose and E. Karakose, "A navigation and reservation based smart parking platform using genetic optimization for smart cities," 5th International Istanbul Smart Grid and Cities Congress and Fair (ICSG), Istanbul, Turkey, 2017, pp. 120-124, 2017.

[2] S. Kazi, S. Nuzhat, A. Nashrah and Q. Rameeza, "Smart Parking System to Reduce Traffic Congestion," International Conference on Smart City and Emerging Technology (ICSCET), Mumbai, India, 2018, pp. 1-4, 2018.

[3] Jae Kyu Suhr and Ho Gi Jung, “Automatic Parking Space Detection and Tracking for Underground and Indoor Environment” IEEE Transactions on Industrial Electronics, 2016.

[4] P.Vijay Kumar, Siddarth T.S “A Prototype Parking System using Wireless Sensor Networks” International Journal of Computer Communication and Information System-vol.12, No.1. ISSN: 0976-1349 July-Dec 2010.

[5] Bhagat, Sharad S., Amit D. Bagul, Pratik N. Patil, and Sanket A. Dahale. "Perceptive car parking booking system with IOT technology." International research journal of engineering and technology, no. 2, 1123-1125, 2018.

[6] Minal Patil, Vijay Chakole, Krushna Chetepawad, “IoT Based Economic Smart Vehicle Parking System” 3rd International Conference on Intelligent Sustainable Systems (ICISS),05 December 2020.

[7] A. O. Kotb, Y. -C. Shen, X. Zhu and Y. Huang, "iParker—A New Smart Car-Parking System Based on Dynamic Resource Allocation and Pricing," in IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 9, pp. 2637-2647, 2016.

[8] Jiawei Zhang, Suhong Chen, Xudong Wang, Yifei Zhu, “Dynamic Reservation of Edge Servers via Deep Reinforcement Learning for Connected Vehicles” IEEE Transactions on mobile computing, Vol. 22, May 2023.

[9] Erick Fernando, Harco Leslie Hendric Spits Warnars, “Innovation Architecture Smart Car Parking System With Wireless Sensor Networks” 7th International Conference on Information Technology, Computer, and Electrical Engineering, September 2020.

[10] Ali Madahian , Pedram Ashofteh Ardakani , Jamshid Abouei, Ali Mirvakili, Arash Mohammadi, And Valencia Koomson “A Hybrid VLC/RF Parking Automation System” IEEE conference, Vol. 11,2023.

[11] Reshma K R, Gladson A J, Akhil K R, Pavan P, Moulika P “Online Parking Booking System” International Journal of Advanced Research in Science, Communication and Technology (IJARSCT) Vol. 2, 2022.

[12] Ilakkiya. S.N, Nevetha. R, Deepa. R “Online Booking System for Car Parking” International Journal of Scientific & Technology Research, Vol. 9, 2020.

[13] Y. M. Dalal, D. R. K. Raja and U. Ashwin Kumar, "Eco Park: A Low-Cost and Sustainable Approach for Smart Parking Systems," 2023 World Conference on Communication & Computing (WCONF), RAIPUR, India, pp. 1-6, 2023.

[14] Thangam, E. Cassin, M. Mohan, J. Ganesh, C. V. Sukesh, and A. Prof. "Internet of Things (IoT) based smart parking reservation system using raspberry-pi." International Journal of Applied Engineering Research 13, no. 8, 5759-5765, 2018.

[15] Kanteti, Dharmini, D. V. S. Srikar, and T. K. Ramesh. "Smart parking system for commercial stretch in cities." In 2017 International Conference on Communication and Signal Processing (ICCSP), pp. 1285-1289. IEEE, 2017.

[16] Hanif, Noor Hazrin Hany Mohamad, Mohd Hafiz Badiozaman, and Hanita Daud. "Smart parking reservation system using short message services (SMS)." In 2010 International Conference on Intelligent and Advanced Systems, pp. 1-5. IEEE, 2010.

[17] Farooqi, Norah, Shroug Alshehri, Sahar Nollily, Lama Najmi, Ghaidaa Alqurashi, and Abeer Alrashedi. "UParking: Developing a smart parking management system using the internet of things." In 2019 Sixth HCT Information Technology Trends (ITT), pp. 214-218. IEEE, 2019.

[18] Gupta, Aniket, Sujata Kulkarni, Vaibhavi Jathar, Ved Sharma, and Naman Jain. "Smart car parking management system using IoT." Am. J. Sci. Eng. Technol 2, no. 4, 112-119, 2017.

Review of Crowdfunding: Historical Evolution, Societal Implications, and Architectural Perspectives

Kajal Joseph,
Research Scholar,
Computer Science and Engineering Department,
Amity School of Engineering and Technology,
Amity University Maharashtra, Mumbai. Maharashtra - 410206, India.
kajal.jewani@ves.ac.in

Deepa Parasar,
Professor,
Computer Science and Engineering Department,
Amity School of Engineering and Technology,
Amity University Maharashtra, Mumbai. Maharashtra - 410206, India.
dparasar@mum.amity.edu

Abstract—This study performs an analysis on crowdfunding. It explores how crowdfunding began, how it's impacting society, the challenges it faces, and how these platforms are designed. Following a detailed historical review of crowdfunding platforms' emergence, the paper explores their significant impacts on investors, enterprises, and society at large. By reviewing relevant literature, the study identifies research gaps, thus establishing a foundation for future investigation. The analysis delves into architectural designs and frameworks, while simultaneously investigating the difficulties faced by businesses during crowdsourcing campaigns. Finally, the study synthesizes key findings and their implications for academics, practitioners, and policymakers within the domain of crowdsourcing. Decentralized crowdfunding has fundamentally transformed the investment landscape, revolutionizing traditional models and unlocking novel funding avenues for early-stage ventures. Examining this evolution in detail, with a specific focus on the AngelicCrowd platform and its impact on angel investors, this study aims to provide a holistic understanding of the dynamics within the decentralized crowdfunding ecosystem. Through analysis of key trends, obstacles, and opportunities, the research investigates factors influencing adoption, the role of blockchain technology and smart contracts, and impediments to widespread implementation. The analysis is informed by a diverse dataset, incorporating responses from angel investors utilizing AngelicCrowd and similar platforms. Additionally, the paper explores the repercussions of these developments on conventional funding strategies and the democratization of investment.

Keywords— Decentralized Crowdfunding, Angel Investing, Blockchain Technology, Transparency, Security, Accessibility, Smart Contracts.

I. INTRODUCTION

The landscape of angel and venture capital funding is a dynamic and critical aspect of entrepreneurial ecosystems worldwide. Studies have highlighted the significance of alternative funding sources beyond traditional banking, emphasizing the role of angel investors in providing crucial financial support to startups [1]. Understanding the selection criteria and decision-making processes of angel investors is essential for entrepreneurs seeking funding and investors looking to diversify their portfolios. Moreover, the comparative institutional perspective on angel investors in different settings sheds light on how institutional frameworks influence investment decisions, offering valuable insights into the complexities of funding processes.

The study on chosen companies in India provides a comprehensive analysis of the funding landscape, exploring the investment criteria, sources of funding, and challenges faced by stakeholders in the Indian startup ecosystem. By examining the non-bank funding agencies and their impact on the financial landscape, the research aims to contribute to a deeper understanding of the evolving dynamics of venture and angel capital funding in India.

The utilization of statistical software such as SPSS for data analysis in the study underscores the importance of rigorous quantitative analysis in examining the trends and patterns within the venture and angel capital funding process. Through a combination of descriptive and inferential statistics, as well as content analysis techniques, the study provides a robust framework for evaluating the data collected from primary and secondary sources. This research seeks to bridge the gap in the existing literature by offering insights into the intricacies of angel and venture capital funding processes, particularly in the context of India and aims to inform future studies and practices in the field. [1]

This study explores the multifaceted landscape of crowdfunding, beginning with a comprehensive examination of its historical development in Section 2. It proceeds to investigate the impact of crowdfunding on society, investors, and startups in Section 3, shedding light on its socio-economic implications. Section 4 identifies existing research gaps in the field, while Section 5 delves into the challenges encountered by startups in crowdfunding endeavors. Section 6 provides a thorough review of related works, offering insights into the existing literature on crowdfunding. In Section 7, the paper discusses the architectural designs of crowdfunding platforms, followed by an evaluation of the impact of architectural frameworks in Section 8. Section 9 delves into the various challenges crowdfunding initiatives face, encompassing technical, legal, and regulatory aspects. The paper concludes by summarizing key findings, discussing implications, and suggesting avenues for future research and practice.

II. HISTORICAL DEVELOPMENT OF CROWDFUNDING PLATFORMS

Early Beginnings: Crowdfunding originated in the early 2000s with platforms like ArtistShare and Indiegogo [1]. These platforms pioneered the concept of crowdfunding by enabling individuals to raise funds from the public for creative projects.

Rise of Reward-Based Crowdfunding: Platforms like Kickstarter and GoFundMe gained popularity in the mid-2000s, offering reward-based crowdfunding where backers receive rewards or products in exchange for their contributions [1]. This model revolutionized fundraising for creative projects, tech innovations, and charitable causes, democratizing access to capital.

Evolution into Equity Crowdfunding: The passage of the JOBS Act in 2012 in the United States led to the emergence of equity crowdfunding, allowing startups to raise capital from a large number of investors in exchange for equity stakes [1]. Platforms like SeedInvest and Crowdcube facilitated this shift, opening up investment opportunities to a broader range of investors.

III. IMPACT ON SOCIETY, INVESTORS, AND STARTUPS:

Society: Crowdfunding platforms have democratized access to capital, empowering individuals and communities to support causes they believe in [1]. Social impact projects, charitable initiatives, and disaster relief efforts have benefitted from the collective power of crowdfunding, fostering community engagement and social responsibility.

Investors: Crowdfunding has transformed the investment landscape by providing retail investors with opportunities to invest in early-stage startups and projects [1]. This direct access to investment opportunities has democratized investing, allowing individuals to diversify their portfolios and potentially earn returns on high-growth ventures.

Startups: Crowdfunding platforms offer startups a viable alternative to traditional funding sources like venture capital and bank loans [1]. By leveraging the crowd, entrepreneurs can validate their ideas, engage with customers, and raise capital to bring their products to market. Crowdfunding also serves as a marketing tool, helping startups build brand awareness and attract early adopters.

Overall, the historical development of crowdfunding platforms has had a profound impact on society, investors, and startups by democratizing access to capital, fostering innovation, and creating new avenues for collaboration and investment in the entrepreneurial landscape [1].

IV RESEARCH GAPS

Existing crowdfunding platforms face significant limitations: excessive fees, scam risks, IP vulnerability, marketing challenges, and restrictive criteria, resulting in uncertainty for investors due to insufficient security, transparency, and control over their investments. These platforms struggle with incomplete projects, lack of guarantees, limited progress monitoring, and potential fraud, highlighting the need for globally standardized due diligence methods and regulatory frameworks to enhance investor confidence and risk mitigation.

V CHALLENGES FOR STARTUPS

Startups face a myriad of challenges as they navigate the complex landscape of entrepreneurship. Here are some common challenges that startups often encounter: Limited Funding, Market Validation, Talent Acquisition, Uncertain Regulatory Environment, Rapid Technological Changes, Intense Competition, Scaling Operations, Limited Brand Recognition, Adaptability to Change, Legal and Intellectual Property Issues, Customer Acquisition and Retention, Cash Flow Management, Strategic Partnerships, Economic Downturns. Addressing these challenges requires a combination of resilience, strategic planning, adaptability, and a commitment to continuous learning and improvement. Successful startups navigate these obstacles with a clear understanding of their market, a solid business strategy, and a focus on delivering value to their customers.

VI. RELATED WORKS

MERITS AND DEMERITS OF EXISTING WORK.

MERITS:

- Crowdfunding allows anyone with a good idea to raise capital from a wider pool of investors, not just traditional institutions.
- Campaigns can gauge public interest and receive valuable feedback on the proposed product or service.
- Platforms can foster community around a project, leading to stronger customer loyalty and brand recognition.
- Compared to traditional financing methods, crowdfunding platforms can offer lower fees and faster access to capital.

DEMERITS:

- Standing out amongst a sea of projects can be challenging, requiring a well-developed campaign and effective marketing strategy.
- Depending on the type of crowdfunding and location, regulations can be complex and add administrative burdens to campaigns.
- While platforms implement safeguards, there's always a risk of encountering fraudulent campaigns.

- Not all campaigns reach their funding goals, potentially leaving creators with wasted time and effort.

TABLE I. REVIEW OF CROWDFUNDING

Author	**Significance**	**algorithm/Model used/Framework/ statistical methods used** **Dataset Source**
Preksha Jain et al. (2023)	Contributes to understanding the funding landscape for startups in India and the role of venture capital and angel investors	descriptive inferential statistics content analysis techniques quantitative data - statistical software SPSS (Statistical Package for the Social Sciences). qualitative data - coding and thematic analysis
Roma et al. (2023)	Contributes to understanding the impact of environmental sustainability on crowdfunding and venture capital investments in new technology ventures	regression models robust OLS or probit regression models
Ho et al. (2016)	Provides a structured approach for evaluating startup enterprises based on trust and soft information, potentially aiding investors and stakeholders in decision-making	ICT technologies social network analysis FinTech Big Data scrapping dataset: survey of 100 startup companies in Thailand
Cusumano (2013)	Helps stakeholders understand the factors influencing the success or failure of a startup venture	Qualitative assessments
Sala et al. (2021)	Contributes to understanding the dynamics of pivoting in tech startups and its implications for business strategy	Descriptive analysis
Pinelli et al. (2022)	Provides insights into the factors that affect fundraising for startups and the role of founder attributes in attracting investment	OLS (ordinary least squares) regression analysis VIF (variance inflation factor)
Regmi et al. (2015)	Offers insights into the effectiveness of startup accelerators in supporting and scaling early-stage ventures	Regression Analysis
Kumar (2018)	Provides a comprehensive overview of the Indian startup landscape, highlighting key challenges and opportunities for entrepreneurs	Descriptive analysis
De Castro (2017)	Offers insights into the distinct characteristics of business angel investments and crowdfunding for startup financing	Linear regression models
Bernardino and Santos (2020)	Provides insights into the potential advantages of crowdfunding as a financing mechanism for startups and innovative projects	Factor analysis
Mochkabadi and Volkmann (2020)	Offers a consolidated view of the current state of knowledge of equity crowdfunding, highlighting gaps and areas for further research	Quantitative methods - Observational studies - Surveys - Regressions (OLS, PLS, hierarchical, probit, logit, binomial) - Structural equations - Descriptive statistics - Mean differences (T-tests, ANOVA, MANOVA, Chi-square) Qualitative methods - Interview-based studies - Case study - Literature review - Descriptive
OECD (2011)	Provides insights into the unique financial needs and challenges faced by high-growth firms in accessing capital	Qualitative research - interviews Dataset: Angel Capital Association, the Center for Venture Research, and the Kauffman Foundation.
Lukkarinen et al. (2019)	Offers insights into the factors that drive investor participation in equity crowdfunding campaigns and their decision-making process	logistic regression algorithm
Arroyo et al. (2019)	Contributes to understanding the potential of machine learning in enhancing decision-making processes for venture capital investments	Decision Trees Random Forests Extremely Randomized Trees Gradient Tree Boosting Dataset: Crunchbase
Singhal et al. (2022)	Provides insights into the factors influencing venture capitalists' investment decisions and the performance of startup investments	Decision Tree model Ensemble models - Random Forests - Gradient Boosting ROC curves Supervised Machine Learning Dataset: Crunchbase
Sathawora Wong et al. (2018)	Offers guidance to entrepreneurs and investors on the factors that can enhance fundraising outcomes for startups.	factor analysis linear regression
Yankov et al. (2014)	Provides a structured approach for evaluating the performance and success factors of technology start-up companies.	Data mining classification tree
da Silva Ribeiro	Contributes to understanding the potential of machine	Random Forests (RF)

Author	Significance	algorithm/Model used/Framework/ statistical methods used Dataset Source
Bento (2018)	learning in predicting the success of start-up ventures.	Logistic Regression (LR) Support Vector Machines (SVM)
Nalintippayawong et al. (2018)	Provides valuable insights into the critical success factors for startups in Thailand, offering a framework for understanding the key elements contributing to their success	Structural Equation Model (SEM) factor analysis multivariable regression analysis
Yin and Luo (2018)	Provides valuable insights into the evolving selection criteria used by accelerators, offering guidance to startups on how to align with the changing expectations of accelerators	regression model
AlHazza et al. (2019)	Contributes to understanding the success factors in new product development for startup companies, potentially providing a framework for improving the effectiveness of the new product development process	fuzzy logic toolbox in Matlab 2016 Rule-based systems Neural Networks. - data preprocessing - feature extraction - clustering Dataset: survey of 50 startup companies in Malaysia
Kim and Lee (2018)	Provides valuable insights into risk management in the cryptocurrency market, offering guidelines to help investors navigate potential threats effectively	SHA-256 algorithm
Atluri (2021)	Sheds light on the potential of blockchain technology to transform and improve the crowdfunding landscape, offering insights into its benefits and implications for crowdfunding initiatives	doctrinal method an interdisciplinary study of Descriptive and analytical
S. M, S. S, A. B. S, H. S, and I. K (2022)	Explores the innovative potential of decentralized crowdfunding platforms using smart contracts, offering insights into the possibilities of creating more transparent, efficient, and accessible crowdfunding mechanisms	blockchain technology
Sadia et al. (2022)	The paper may contribute to the development of more accurate and effective forecasting models for blockchain-based cryptocurrencies.	proposed model for predicting cryptocurrencies' daily prices Dataset: CoinMarketCap website
Patil et al. (2021)	The paper may contribute to the development of more secure, transparent, and efficient crowdfunding platforms.	blockchain technology
Zad, Saniya, Khan, Zishan, Warambhe, Tejas,	The paper may contribute to the development of more secure, transparent, and efficient crowdfunding platforms.	frameworks and tools Solidity, Web3Js, ReactJs, ExpressJs, NodeJs, and MongoDB

Author	Significance	algorithm/Model used/Framework/ statistical methods used Dataset Source
Jadhav, Rushikesh, and Alone, Vinod, C (2023)		languages HTML, CSS, and JavaScript
Dushnitsky and Zunino (2018)	The paper may contribute to a better understanding of the potential benefits and challenges of crowdfunding as a financing option for startups and small businesses.	Regression analysis Dataset: Kickstarter Indiegogo Seedrs.
Venslavienė et al. (2023)	The paper may contribute to a better understanding of the potential benefits and challenges of using blockchain technology in crowdfunding campaigns.	VASMA-L Criteria Weighting Method Dataset: blockchain-based crowdfunding platform
Guggenberger et al. (2023)	The paper may contribute to a better understanding of the potential applications of blockchain technology in various industries and sectors.	Ethereum blockchain Hyperledger Fabric
Yankov et al. (2014)	Provides a structured approach for evaluating the performance and success factors of technology start-up companies.	Data mining classification tree
da Silva Ribeiro Bento (2018)	Contributes to understanding the potential of machine learning in predicting the success of start-up ventures.	Random Forests (RF) Logistic Regression (LR) Support Vector Machines (SVM)

VII. ARCHITECTURAL DESIGNS

Architectural frameworks in crowdfunding typically refer to the planning and creation of the structure, features, and functionalities of crowdfunding platforms.

Crowdfunding Platform Design with Blockchain Integration:
This research explores the integration of blockchain technology within crowdfunding platforms to augment transparency, security, and efficiency. The design leverages public ledgers, smart contracts, and permission access control mechanisms offered by blockchain. Public ledgers provide a tamper-proof, auditable, and decentralized record of transactions and agreements, significantly enhancing transparency within the crowdfunding ecosystem.
Furthermore, blockchain technology fosters security through a decentralized consensus mechanism, cryptographic hashing, and the inherent immutability of data on the blockchain. This combination safeguards against fraudulent activities by ensuring tamper-proof records and fostering a more secure environment. Additionally, the utilization of consensus mechanisms, cryptographic hashing, and data verification protocols within the blockchain framework promote accuracy by guaranteeing the integrity and validity of information across distributed networks. Finally, the

implementation of blockchain technology, with its decentralized consensus mechanism and immutable ledger, strengthens reliability by bolstering data integrity and mitigating the risks associated with single points of failure within the crowdfunding platform.
By leveraging decentralized ledger systems and smart contracts, the platform ensures trust among users and facilitates seamless transactions [30]

Smart Contract Automation for Fund Disbursement:
The design includes the implementation of smart contracts to automate fund disbursement processes in crowdfunding campaigns. Smart contracts enable automatic distribution of funds based on predefined criteria, reducing manual intervention and streamlining transaction processing [30]

User Interface Design for Enhanced Engagement:
The design focuses on creating intuitive user interfaces with interactive features to enhance user engagement on crowdfunding platforms. Real-time updates on campaign progress, personalized recommendations, and user-friendly interfaces contribute to a positive user experience and encourage active participation [29]

Sustainability-Oriented Project Design Framework:
The design framework emphasizes integrating environmental sustainability considerations into crowdfunding projects. By incorporating sustainability metrics, impact assessments, and green initiatives, the platform promotes the support of projects with long-term societal and environmental benefits [29]

Performance Monitoring Dashboard Design:
The design involves developing a performance monitoring dashboard for crowdfunding platforms. The dashboard provides real-time analytics, key performance indicators, and data visualization tools to help platform operators track campaign success, identify trends, and make data-driven decisions for optimizing fundraising strategies [30].

Decentralized Crowdfunding Platform Architecture Design:
The design focuses on creating a decentralized crowdfunding platform architecture using blockchain technology. By decentralizing contract storage, transaction records, and verification processes, the platform ensures security, transparency, and autonomy in fundraising activities [29]

These designs showcase the innovative approaches and architectural considerations in the development of crowdfunding platforms, emphasizing aspects such as blockchain integration, smart contract automation, user interface design, sustainability frameworks, performance monitoring, and decentralized architecture. Each design contributes to enhancing the functionality, user experience, security, and sustainability of crowdfunding ecosystems.

VIII IMPACT OF ARCHITECTURAL FRAMEWORKS

The impact of architectural frameworks in crowdfunding is significant as they influence various aspects of crowdfunding platforms, including transparency, security, efficiency, user experience, sustainability, and decision-making processes.

Here is an explanation of the impact with citations from the provided papers:

Transparency and Trust:
Architectural frameworks, such as blockchain integration, enhance transparency and trust in crowdfunding platforms by providing immutable transaction records and decentralized verification mechanisms. This fosters confidence among backers and project creators, leading to increased participation and support for crowdfunding campaigns [30].

Security and Accountability:
Architectural frameworks that prioritize security features, such as anti-tampering mechanisms and decentralized ledger systems, contribute to the accountability of crowdfunding platforms. By ensuring the integrity of transactions and data, these frameworks mitigate risks of fraud and unauthorized access, fostering a secure environment for fundraising activities [26].

Efficiency and Automation:
Architectural frameworks incorporating smart contracts and automation streamline crowdfunding operations, leading to increased efficiency and reduced manual interventions. Automation of fund disbursement and verification processes accelerates transaction processing, enhancing the overall operational efficiency of crowdfunding platforms [30].

User Experience and Engagement:
Well-designed architectural frameworks improve the user experience and engagement on crowdfunding platforms. Intuitive interfaces, interactive features, and real-time updates enhance user engagement, encouraging active participation and fostering a sense of community among platform users [30].

Sustainability and Innovation:
Architectural frameworks supporting sustainability-oriented projects and innovative fundraising models drive growth and diversification in crowdfunding. By integrating sustainability considerations and novel approaches, these frameworks promote impactful projects with long-term societal benefits [30].

Performance Monitoring and Decision-Making:
Architectural frameworks facilitating performance monitoring and data analytics empower platform operators to make informed decisions and optimize strategies. By tracking key metrics and analyzing outcomes, these frameworks enable stakeholders to assess platform effectiveness and drive continuous improvement in crowdfunding operations [30].
Architectural frameworks play a crucial role in shaping the success and sustainability of crowdfunding platforms by enhancing transparency, security, efficiency, user experience, sustainability, and decision-making processes. Leveraging innovative technologies and design principles,

these frameworks contribute to building trust, engaging users, promoting innovation, and driving operational excellence in the crowdfunding ecosystem.

IX. CHALLENGES IN CROWDFUNDING

Addressing the below-mentioned challenges can enhance the robustness, validity, and applicability of findings in the evolving field of crowdfunding.

Platform Diversity and Data Limitations:
One challenge highlighted in the papers is the limited focus on a few crowdfunding platforms, leading to a lack of comprehensive data analysis across a diverse range of platforms. This limitation hinders a holistic understanding of crowdfunding dynamics and investment drivers across various platforms [28].

Lack of Cross-Platform Comparison:
The scarcity of studies conducting cross-platform comparisons limits the generalizability of findings in crowdfunding research. While insights from a few platforms are valuable, a broader analysis across multiple platforms is essential to understand investment patterns and drivers comprehensively [28].

Platform Design Variability and Impact:
The role of crowdfunding platform design choices and their impact on funding patterns is often overlooked in the literature. Variations in platform design can influence user behavior, campaign success rates, and overall platform competitiveness, highlighting the need for more research in this area [28].

X. CONCLUSION

This study provides a thorough examination of crowdfunding, covering its historical development, societal impact, challenges, and architectural perspectives. Beginning with an in-depth exploration of crowdfunding platform evolution, the study investigated into its effects on society, investors, and startups, identifying research gaps and discussing architectural designs and frameworks. It concludes by synthesizing key findings and implications for researchers, practitioners, and policymakers. Additionally, decentralized crowdfunding's emergence as a transformative force in the investment landscape is investigated, with a focus on the "AngelicCrowd" platform. Through a comprehensive survey, the study examines trends, challenges, and opportunities in decentralized crowdfunding, drawing on responses from angel investors. Findings reveal a growing interest in decentralized crowdfunding among investors, driven by its transparency and accessibility. However, regulatory uncertainties and security concerns pose challenges. The evolution of AngelicCrowd serves as a case study reflecting broader trends in decentralized finance, highlighting smart contracts, tokenization, and decentralized governance as key enablers of innovation. Looking ahead, decentralized crowdfunding holds promise for democratizing investment, albeit with the need for continued research and collaboration to address regulatory complexities and unlock its full potential.

REFERENCES

1. Preksha Jain, Saakshi Singh, Gaurav S Badolla, Moksha S Jain, Rahul M Jain, Ms. Megha Kukreja, Dr. Priya Makhija CMS Jain University, Bangalore, India, "A Study of Chosen Companies 'Venture and Angel Capital Funding Process in India'", © 2023 IJCRT | Volume 11, Issue 4 April 2023 | ISSN: 2320-2882
2. [3] P. Roma, M. Vasi, S. Testa and G. Perrone, "Environmental Sustainability Orientation, Reward-Based Crowdfunding, and Venture Capital: The Mediating Role of Crowdfunding Performance for New Technology Ventures," in IEEE Transactions on Engineering Management, vol. 70, no. 9, pp. 3198-3212, Sept. 2023, doi: 10.1109/TEM.2021.3080428.
3. H. Ho, S. B. Jeong and B. G. Lee, "Conceptual Framework for Startup Enterprise Evaluation: Value of the Trust and Soft Information," 2016 3rd Asia-Pacific World Congress on Computer Science and Engineering (APWC on CSE), Nadi, Fiji, 2016, pp. 181-186, doi: 10.1109/APWC-on-CSE.2016.038.
4. Cusumano, M. A. (2013). Evaluating a startup venture. Communications of the ACM, 56(10), 26. doi:10.1145/2505337
5. P. K. Sala, S. P. Philbin and S. Barikzai, "Exploring entrepreneurial pivoting and the factors that trigger pivots by tech startups," 2021 IEEE Technology & Engineering Management Conference - Europe (TEMSCON-EUR), Dubrovnik, Croatia, 2021, pp. 1-6, doi: 10.1109/TEMSCON-EUR52034.2021.9488584.
6. M. Pinelli, F. Cappa, S. Franco, E. Peruffo and R. Oriani, "Too Much of Two Good Things: Effects of Founders' Educational Level and Heterogeneity on Start-Up Funds Raised", IEEE Transactions on Engineering Management.
7. K. Regmi, S. A. Ahmed and M. Quinn, "Data-Driven Analysis of Startup Accelerators", Universal Journal of Industrial and Business Management, vol. 3, pp. 54-57, 2015.
8. G. P. Kumar, "Indian startup-Issues Challenges and Opportunities", Conference: Indian Startup Ecosystem-Challenges and Opportunities, March 2018, [online] Available: https://www.researchgate.net/publication/323855305_INDIAN_STARTUPS-_ISSUES_CHALLENGES_AND_OPPORTUNITIES.
9. L. De Castro, "A Comparative Analysis on Business Angels and Crowdfunding: The Impact of Behaviours and Attitudes on Investment Decisions," MSc Graduation thesis, Supervisor: G. Giudici, Co-Supervisors: V. Capizzi and G. Scellato, Academic year 2017-2018, Student ID: 872674.
10. S. Bernardino and J. F. Santos, "Crowdfunding: An Exploratory Study on Knowledge, Benefits and Barriers Perceived by Young Potential Entrepreneurs," Journal of Risk and Financial Management, vol. 13, no. 4, pp. 1-15, Apr. 2020. [Online]. Available:https://doi.org/10.3390/jrfm13040081.
11. K. Mochkabadi and C. K. Volkmann, "Equity crowdfunding: a systematic review of the literature," Small Bus Econ, vol. 54, pp. 75–118, 2020. [Online]. Available: https://link.springer.com/article/10.1007/s11187-018-0081-x.
12. OECD, "Financing High-Growth Firms: The Role of Angel Investors," OECD Publishing, 2011. [Online]. Available: http://dx.doi.org/10.1787/9789264118782-en.
13. A. Lukkarinen, J. Wallenius, and T. Seppälä, "Investor Motivations and Decision Criteria in Equity Crowdfunding," Aalto University, Helsinki, Finland, Feb. 14, 2019. [Online]. Available: https://ssrn.com/abstract=3263434.
14. J. Arroyo, F. Corea, G. Jimenez-Diaz and J. A. Recio-Garcia, "Assessment of Machine Learning Performance for Decision Support in Venture Capital Investments," in IEEE Access, vol. 7, pp. 124233-124243, 2019, doi: 10.1109/ACCESS.2019.2938659.
15. J. Singhal, C. Rane, Y. Wadalkar, M. Joshi and A. Deshpande, "Data-Driven Analysis for Startup Investments for Venture Capitalists," 2022 International Conference for Advancement in Technology (ICONAT), Goa, India, 2022, pp. 1-6, doi: 10.1109/ICONAT53423.2022.9725892.
16. P. Sathaworawong, N. Thawesaengskulthai and K. Saengchote, "Determinant of Startups' Fund-raising Value: Entrepreneur and Firm Characteristic," 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, Thailand, 2018, pp. 1309-1314, doi: 10.1109/IEEM.2018.8607767.

17. B. Yankov, P. Ruskov, and K. Haralampiev, ‘‘Models and tools for technology start-up companies success analysis,’’ Econ. Alternatives, no. 3, pp. 15–24, 2014.
18. F. R. da Silva Ribeiro Bento, ‘‘Predicting start-up success with machine learning,’’ M.S. thesis, Dept. Inf. Manage., Universidade Nova do Lisboa, Lisbon, Portugal, 2018.
19. S. Nalintippayawong, N. Waiyawatpattarakul, and S. Chotipant, "Examining the Critical Success Factors of Startup in Thailand Using Structural Equation Model," 2018 10th International Conference on Information Technology and Electrical Engineering (ICITEE), Bali, Indonesia, 2018, pp. 388-393, doi: 10.1109/ICITEED.2018.8534854.
20. B. Yin and J. Luo, "How Do Accelerators Select Startups? Shifting Decision Criteria Across Stages," in IEEE Transactions on Engineering Management, vol. 65, no. 4, pp. 574-589, Nov. 2018, doi: 10.1109/TEM.2018.2791501.
21. M. H. F. AlHazza, I. F. Bourini, M. H. Zubaidah and N. Bt Selamat, "Success Factor in New Product Development for Startup Companies Using Fuzzy Logic Approach", 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA), pp. 1-5, 2019.
22. C. Y. Kim and K. Lee, "Risk Management to Cryptocurrency Exchange and Investors Guidelines to Prevent Potential Threats," 2018 International Conference on Platform Technology and Service (PlatCon), Jeju, Korea (South), 2018, pp. 1-6, doi: 10.1109/PlatCon.2018.8472760.
23. Atluri Divija Choudary, “Role of Blockchain Technology in Crowdfunding”,4th International Conference on Management, Economics, and Finance,10-12 september 2021.
24. S. M, S. S, A. B. S, H. S and I. K, "Decentralized Crowdfunding Platform using Smart Contracts," 2022 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems (SPICES), THIRUVANANTHAPURAM, India, 2022, pp. 302-307, doi: 10.1109/SPICES52834.2022.9774132.
25. I. Sadia, A. Mahmood, L. Binti Mat Kiah and S. R. Azzuhri, "Analysis and Forecasting of Blockchain-based Cryptocurrencies and Performance Evaluation of TBATS, NNAR and ARIMA," 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET), Kota Kinabalu, Malaysia, 2022, pp. 1-6, doi: 10.1109/IICAIET55139.2022.9936798.
26. V. Patil, V. Gupta, and R. Sarode, "Blockchain-Based Crowdfunding Application," 2021 Fifth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Palladam, India, 2021, pp. 1546-1553, doi: 10.1109/I-SMAC52330.2021.9640888.
27. Zad, Saniya and Khan, Zishan and Warambhe, Tejas and Jadhav, Rushikesh and Alone, Vinod, Crowdfunding Using Blockchain Technology (December 31, 2022). Available at SSRN: https://ssrn.com/abstract=4330476 or http://dx.doi.org/10.2139/ssrn.4330476
28. Dushnitsky, Gary and Zunino, Diego, The Role of Crowdfunding in Entrepreneurial Finance (July 31, 2018). Handbook of Research on Crowdfunding, Forthcoming, Available at SSRN: https://ssrn.com/abstract=3237356 or http://dx.doi.org/10.2139/ssrn.3237356
29. S. Venslavienė, J. Stankevičienė, and I. Leščauskienė, "Evaluation of Blockchain-Based Crowdfunding Campaign Success Factors Based on VASMA-L Criteria Weighting Method," Adm. Sci., vol. 13, no. 6, p. 144, May 2023. [Online]. Available: https://doi.org/10.3390/admsci13060144.
30. T. Guggenberger, B. Schellinger, V. von Wachter, and N. Urbach, "Kickstarting blockchain: designing blockchain-based tokens for equity crowdfunding," Electron. Commerce Res., 2023. [Online]. Available: https://doi.org/10.1007/s10660-022-09634-9

Machine Learning and Ontology based Framework in Information Retrieval using Semantic Query Expansion

Rupali R. Deshmukh
Computer Science and Engineering
H.V.P.M's COET Amravati.
Amravati, India
drupali1604@gmail.com

Dr. Anjali B. Raut
Computer Science and Engineering
H.V.P.M's COET Amravati.
Amravati, India
anjali.dahake@rediffmail.com

***Abstract*—Data and information stored online is steadily growing in volume. Indeed, when it comes to readily available, high-quality information, the web is unrivalled. Information retrieval describes the steps used to find the desired data in a database. By analyzing the user's query semantically, semantic information retrieval can offer results that are more relevant to their original intent. When this occurs, ontology-based knowledge representation has the potential to facilitate semantic retrieval more efficiently than alternative representation approaches, such as frames and semantic networks. In the past twenty years, machine learning techniques have found widespread application across several fields of art, science, and technology. Information retrieval activities that make extensive use of machine learning techniques include document analysis and query expansion. This paper presented a semantic information retrieval system based on machine learning and ontologies. Our research demonstrates that the method outperforms the state-of-the-art on certain recall, precision, and f-measure metrics used to evaluate the effectiveness of semantic information retrieval. Experiments have proven that the method helps the system convey knowledge better and retrieves semantic information faster.**

Keywords - Document retrieval, Ontology, Machine Learning, Information retrieval, Semantic Information.

I. INTRODUCTION

Millions of individuals are checking out online sites to fulfil their information demands, making them the biggest source of information. Information retrieval (IR) systems aim to give consumers highly relevant results based on their information demands. For better query mapping to a collection of documents during retrieval, keywords describing the contents of the documents are assigned during the document indexing stage. Numerous proposals in the literature have proposed statistical, neural, and semantic methods to enhance the efficiency of data retrieval tasks [1-2]. Concepts, roles, and connections may be more accurately represented with the help of ontology in semantic web, a method of knowledge representation. The current state of knowledge representation allows for the easy encoding of textual information with inheritance using methods like semantic networks and frames. Semantic network or frame searches may therefore follow class and sub-class hierarchies all the way from the root node to the target node. In semantic networks, however, there are limitations on both the amount and kind of linkages. Both the information hierarchy and the inheritance hierarchy may be preserved using the ontology-based approach to knowledge representation. Information may be retrieved utilizing both single-level and multi-level inheritance using this ontology-based representation of data. Semantic analysis and appropriate information retrieval are therefore improved by semantic information retrieval based on ontologies [3-4].

Traditional information retrieval methods index documents according to the phrases they contain, not the notion that describes them. The phrase "vocabulary gap" explains the circumstance where users and subject matter experts use disparate terminology to refer to the same subject; IR systems will not provide results that contain semantically significant information unless there is a lexical match between the two. Document representation in an IR system has been shown to effectively capture term meanings and their connections utilizing domain knowledge [5][7]. Ontologies have long been used in all fields to formally express and comprehend domain knowledge. In contrast, language gaps are no longer an issue with ontology-based document indexing techniques. However, these approaches have the drawback of being very dependent on the depth and breadth of the initial input ontology. [10][12]. Recent advances in machine learning (ML) have allowed the field of information retrieval (IR) to provide a workable solution to document indexing challenges. When co-occurring events occur, machine learning algorithms leverage word-to-word correlations; nevertheless, they ignore significant semantic relationship structures. Manually created knowledgebases that include these relational structures for semantics include ontologies and semantic lexicons. The benefits of ML modelling have prompted academics to concentrate on developing ML ranking models that can learn features and the model in tandem [15][17] [20-22].

The goal of this work is to give more detailed semantic indexing of scientific publications than only the Computer Science Ontology descriptions. It aims to reach results comparable to those produced by domain experts by studying the semantic level of matching ideas. This research proposes a semantic information indexing strategy for information retrieval that is built on ontologies and machine learning to address this issue.

Word embeddings and ontology from computer science are utilized by the suggested method to extract the notion from a text. The suggested method supplies the ML model with potential keywords extracted from document text using natural language processing techniques. The proposed algorithms improve upon previous efforts at retrieval speed and accuracy and leverage ontology matching to provide users with answers that are more pertinent to their queries. The most significant insights from the research are listed below:

- Developed a semantic approach to machine learning and ontology-based information retrieval system for document indexing.
- Improved the text-based information system's effectiveness with machine learning methods.
- Implement a ranking based approach to retrieve relevant documents for a user query.
- Conducted the trials with the TREC-NIST dataset, considering several performance parameters including F-measure, recall (R), mean average precision (MAP), and precision (P).

This paper is organized as follows: A summary of earlier ontology-based semantic information retrieval and machine learning research is provided in section 2; The proposed materials and procedures are explained in section 3. The result discussion and experimental findings are presented in section 4. This paper's conclusion is in the last section.

II. RELATED WORK

This section classifies the IR methods that have been suggested into two categories: supervised and unsupervised. To make sure the system can adapt to new domains as they emerge, machine learning is used to fine-tune the ontological structures dynamically. [1]. Using a multi-terminology, Bayesian networks (BNs), and multinomial naive Bayes classifier (MNBC) is a novel way to enhance information retrieval systems [2]. A new approach to semantic document indexing was presented [3] that combines machine learning with domain ontology. A novel fuzzy ontology-based framework is proposed to retrieve information utilizing domain-specific knowledge utilized in ontology construction [5]. The augmented fish swarm method is recommended to carry out the retrieval process to improve the effectiveness of semantic IR [6]. Agnostic prediction models are explained by a new interpretation framework that develops an interpretable model using an ontology-based sampling method [7]. Multi-Perspective Sentence Relevance is a new approach to contextual information retrieval (IR) that makes use of BERT-based models. It has several applications in the field of biomedical semantics [8]. There is a new system for automatically generating ontologies that do not depend on any specific domain; it may transform a collection of unstructured texts into an ontology that is compatible with that domain [9]. With relevant feedback, the novel model expands the question by utilizing a deep neural network methodology and a semantic way to find the semantic similarity between the phrases [10].

The several text representations offered, are accountable for retrieving pertinent search results, methodologies, and evaluations performed in conceptual information retrieval [11]. An innovative fuzzy rough set-based intelligent selection of features and categorization method is proposed for semantic information retrieval to enhance the relevance score [12]. Domain ontology, a constraints-based mapper, and point-of-sale (POS), and language processing tools are all integrated into the innovative model [13]. Created an AI system for document management based on ontologies [14]. Integrating lexical and semantic data to handle CLIR (cross-language information retrieval) [15]. Document clustering using a deep ontology-based method was suggested [16]. The machine learning models were trained and tested on the manually collected corpus of scientific papers utilized in the Exposome-Explorer [17]. An enhanced technique for semantic information retrieval was introduced [18]. It is based on decision trees. There is a new proposal for a visual interface that uses autonomously generated ideas to get materials from a digital library [19]. Clustering, rough set, and Bayesian network (BN) are used as an ensemble strategy to create the MLK-rBO model, which serves four different purposes: model assessment, probabilistic network development, knowledge discovery and clustering [20].

To rank objects, same as other retrieval tasks, document retrieval uses a collection of components from neural networks to extract characteristics [21]. An enhanced variant of the original COOT algorithm is the IAOCOOT technique for query expansion, which finds the semantic characteristics that match the query phrase [22]. A web paradigm for information extraction based on ontologies, including an object-attribute-value (O-A-V) hierarchy that users may use as a dictionary to hone their search terms before repeating the process [23]. CO-Search is a semantic multi-stage search engine designed to handle complex queries across the COVID-19 literature. It may help overworked medical professionals get the scientific answers they need and avoid misleading content [24]. The ontologies are included in machine learning algorithms and utilized to compute similarity [25]. An automatically sorted, comprehended, searched, and summarized web material search engine was developed in order to increase relevance scores in AR domains [26]. One novel approach to the categorization issue of textual materials [28] is to combine preexisting knowledge models with ad hoc taxonomies. The development of a new method for semantic document indexing using a Shallow Neural Network and ontologies [29]. The goal of this semantic video retrieval system is to find appropriate films based on user-supplied queries made in natural language [30].

III. PROPOSED METHOD

Figure 1 shows the researchers' proposed solution to the information retrieval problem: a system that uses machine learning and a unique ontology.

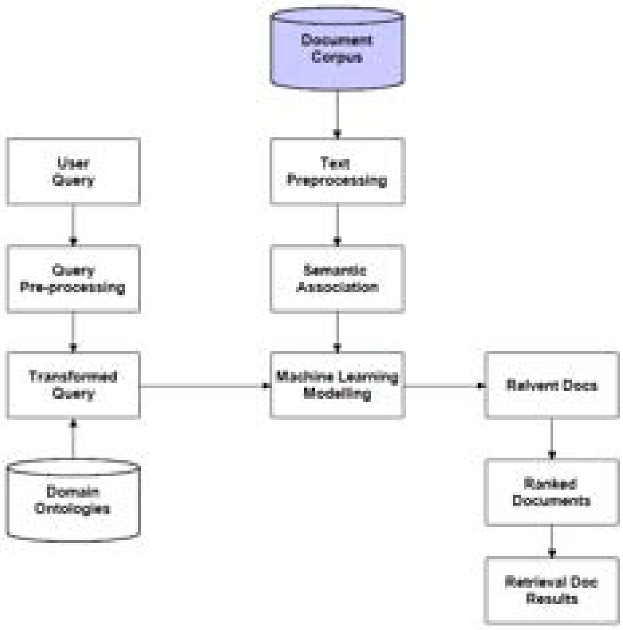


Fig. 1. Machine Learning and Ontology based Semantic Information Retrieval System

Document Corpus Dataset:

The proposed information retrieval system used the TREC Dataset [31] as its document corpus. The package includes documents, queries, and an additional file that has the top 'N' documents ranked for each query. This dataset contains two kinds of information: query IDs and the questions themselves. The evaluation gap problems in IR are somewhat alleviated by the TREC MRT dataset. Due to the sensitive nature of the information contained inside, data limits serve as the driving force behind medical records. The Text Review Conference (TREC) was co-sponsored by the US Department of Defence and the National Institute of Standards and Technology (NIST) in 1992 as part of the TIPSTER text initiative. Facilitating the development of the text retrieval technique required for large-scale estimate was the primary goal, which also served to sustain research in the field of IR. TREC sponsored the earliest comprehensive evaluations of document retrieval efficiency in non-English languages (Spanish and Chinese). TREC's computations are based on content-based digital video retrieval and answers to open-domain queries. The realistic design of the operational circumstances is made possible by the requisite test collections in TREC.

Data Pre-processing:

Pre-processing the query to make it more accurate or machine-readable is necessary before it can be used to find the documents that meet the user's information needs. A processed query is sent to the appropriate module in response to this question. There are frequently many terms in a search engine query. An alternate representation of such a question may be a set of weighted words. Following the matching process, a typical IR system will return several results that are pertinent to the user's request. The query submitter's perception of the page's usefulness in addressing their question is the determining factor in relevance.

Transformed Query:

Unfortunately, the user's original query did not adequately express the specific details he needed. Our algorithm makes use of three main query transformation techniques. Changing the user's original query is typical in manual searches. The retrieval system will automatically broaden the user's query as they type it in. Using the user's past behavior to generate predictions about their current purpose is one approach. This approach augments the user's original inquiry with fresh terms by assessing both local and global information resources. As a whole, this inquiry is trying to answer the first question by revealing related ideas. documents saved on the computer, the entire set of documents, or even only the first set of documents might be targeted. Adding matching terms to searches improves recall. Input from the user is essential for this procedure. After running the refining method, user receive a new query.

Indexing:

The database retrieval system organizes millions of web pages that have the same number of distinct terms. Accurate and speedy retrieval of particular information is the primary objective of indexing. As part of its indexing operation, a search engine will gather keywords and phrases from the sites it has downloaded. The inverted index keeps track of words using a backward file structure. The position of each phrase in the text is tracked by an inverted index. During the indexing of web pages, activities like lexical analysis are carried out, which are similar to the query processing phase and assist improve search engine performance. For full-text indexing, almost every word on the page serves as an index term. Indexing completes any search engine and enhances query performance by accelerating response times. Apart from just indexing websites, search engines also assign a site a ranking based only on the web's link structure, an attempt to approximate an "importance" assessment. To create the index, the indexing component gets text documents as input. Documents are deconstructed into tokens as part of the automated process of constructing an index. Tokens go through several text operations before they become indexing words.

Machine Learning Modelling:

The comparison of query terms with index items is the responsibility of this part. It looks up all the documents in the index that contain the keywords you entered. This is a typical search that uses processed query terms to access a document index. One indicator of how closely two pages match is the number of words they share. One may determine the relevancy of the returned documents by assessing the degree of resemblance between the index and query phrases. Finding the URLs of pages that include the query keywords and comparing them to the index terms is one method. This document-query or query-keyword syntactic matching can be improved with semantic matching. The proposed approach primarily focuses on an ensemble-based information extraction module. Three machine learning algorithms—KNN, DT, and SVM—were utilized by the ensemble information extractor. The ensemble algorithms use these methods to generate a set of phrases that have already been specified.

With query 'i' and doc 'i' representing the query words and documents, respectively, the similarity measure is used in the information retrieval matching process in the proposed framework. During query expansion, we look for the most similar word to the question in terms of semantics to broaden it. By picking the three most important terms from the relevant ontology, we have broadened the scope of the query. We can now use these four terms as a query in any web search engine to find relevant publications based on our domain knowledge of the context. The similarity measure is adjusted according to the equations provided, and the query and document pairs are refined as (query', doc') in the expanded query matching process. In this case, "doc" refers to the newly discovered documents that the search engine has found, and "query" indicates the improved query words.

$$Sim(query, doc) = \frac{\sum_i query_i doc_i}{\sqrt{\sum_i query_i^2}\sqrt{\sum_i doc_i^2}} \qquad (1)$$

$$New\_Sim_{(query', doc')} = \frac{\sum_i query_i' doc_i'}{\sqrt{\sum_i query_i'^2}\sqrt{\sum_i doc_i'^2}} \qquad (2)$$

Compared to a basic Boolean matching model, the computational effort required by a probabilistic weighted model for NLP queries is higher. The ranking uses a score to establish the relative importance of each document. Search engines are based on IR, a field of science and engineering., and one of its major challenges is ranking query results. With a query q and a set of documents D that match it, the primary goal is to use some metric to sort or rank D, with the goal of showing the user the "best" results at the front of the list. Traditionally, the ranking factor has been the relevancy of documents to the given information demand of a query. Through an interface, users may access their graded papers. Based on the study, the researchers use the ranking technique. The process flow is illustrated in Fig.2.

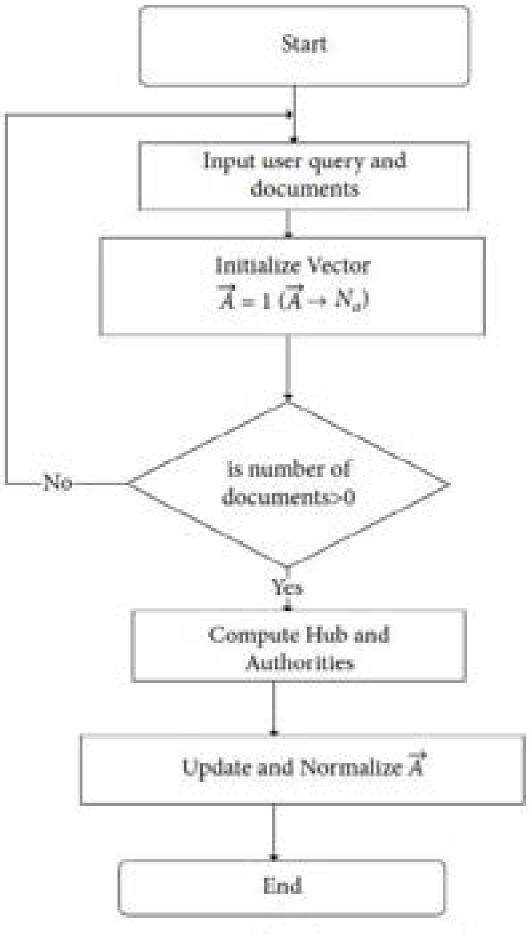


Fig. 2: Document Ranking Approach

Algorithm 1: IR system
Input: *Query Input Document* Output: *Relevant Retrieved Information*
Method: Step1: *Input from the user interface in the form of a query defines the information need of the user.* Step 2: *Assessing the initial inquiry makes use of domain ontology in conjunction with query processing procedures. Document collection is the first step in indexing, which entails carrying out the identical tasks assigned to them by the semantic association module.* Step 3: *Indexes are generated from the original documents in the context of the indexing module.* Step4: *The user's information needs are modeled semantically by the transformed query.* Step5: *The search for literature that answer this query is carried out using a semantics module. All of the document terms that are a match for the query words are located swiftly by the*

indexing system.

Step6*: The retrieved documents are sorted according to their predicted relevance in relation to the term matching score.*

Step7: *Finally, the user provides input to the system by picking out the papers he found most useful from the list of graded papers.*

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. Experimental Setup

This model is tested and evaluated in Python using the Pycharm IDE with Anaconda Distribution. The software used is Windows 11, with 16 GB of RAM, an Intel i7 processor, and 4 GB of graphics memory. The scikit-learn ML and Word2Vec toolbox are part of the distribution. To forecast the optimal performance rate, performance parameters such as F-measure, Mean Average Precision (MAP), precision, and recall are utilized.

B. Evaluation Parameters

Three commonly used metrics and the proposed model were used to evaluate our results. What follows is a more detailed explanation of these metrics.

1. Recall (R): It is the percentage of relevant documents that are discovered during a search. To find this value, use the following formulas. The numbers Dr and Dt represent the number of relevant data that were and were not retrieved, respectively.

$$R_r = \frac{D_r}{D_r + D_t} \tag{3}$$

2. Precision (P): Relevance calculates the proportion of retrieved documents that were truly helpful. Du is the number of useless data that was obtained.

$$R_p = \frac{D_r}{D_r + D_u} \tag{4}$$

3. Mean Average Precision (MAP): It is computed as the average precision score for the documents that the query groups were used to get. The MAP is calculated based on the Average Precision (AP) of each search query. After that, the MAP is calculated and recorded as;

$$MAP = \frac{1}{n} \sum_{n} AP_n \tag{5}$$

4. F-measure: F-measure is as simple as summing up the precision and recall scores harmonically.

$$F\text{-}measure = 2 * (Rp * Rr) / (Rp + Rr) \tag{6}$$

C. Experimental Results analysis

Experiments are carried out and evaluated using TREC datasets to show how effective the proposed and baseline methods are. These metrics, recall, precision, and F-measure are used to assess performance. The number of most-read documents was used to determine outcomes by selecting the 10, 20, 50, and 100 as examples. Table I below provides an in-depth analysis of each of these components.

Table I: Evaluation of Results based on Top-N Documents for Retrieval

Performance Parameter	**Recall**	**Precision**	**F-measure**
N@10	0.72		0.71
N@20	0.70	0.68	0.69
N@50	0.71	0.65	0.67
N@100	0.68	0.64	0.66

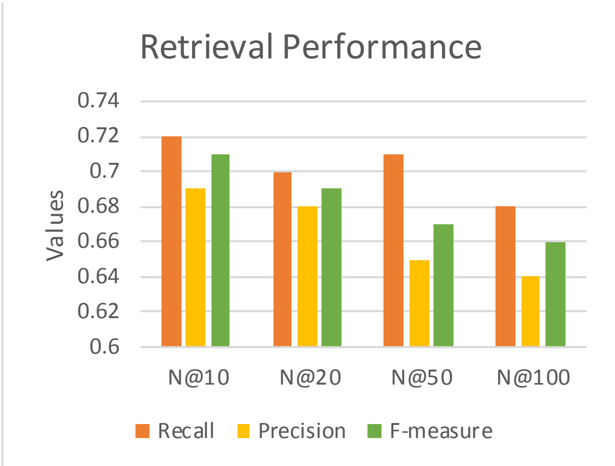


Fig. 3: IR Result evaluation performance for Top-N retrieval

The TREC MRT dataset's recall, precision, and f-measure results are displayed in Fig. 3. The proposed system outperformed the others in terms of precision @10, 20, 50, and 100 respectively. According to the proposed method, recall @10 and 50 is 0.72 and 0.71 respectively. The developed model performed poorly in comparison to other retrieval documents on the f-measure @100.

Table II: Performance Evaluation Parameters for Test Query

Evaluation Parameters/Query	**Recall**	**Precision**	**F-score**	**MAP**
Query1	0.55	0.60	0.52	0.51
Query2	0.48	0.54	0.46	0.44
Query3	0.49	0.54	0.47	0.43

Query4	0.53	0.57	0.48	0.46
Query5	0.57	0.60	0.53	0.55

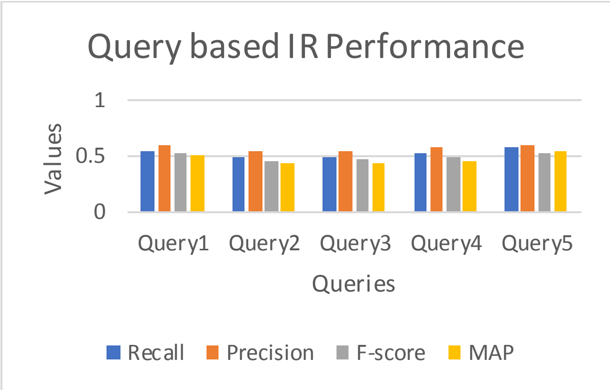


Fig.4: IR Result evaluation performance for N-Query

Table II depicts the results of several searches for f-measure, recall, precision, and MAP. New quarterly highs of 0.57 for recall, 0.60 for precision, 0.53 for f-measure, and 0.55 for MAP were achieved for Query5. With an f-Measure of 0.46 and a MAP of 0.44, the final results for Query2 are precision of 0.54 and a recall of 0.48. Several measures, including f-measure, recall, precision, and MAP, all attained levels of 0.47, 0.49, 0.54, and 0.43 during Query3. See the visual examination of query expansion in Fig. 4. When comparing studies conducted in Query1, Query2, Query3, Query4 and Query5 performance parameters including f-measure, recall, precision, and MAP all showed significant improvements in Query5.

Table III: Comparative analysis with state of art models

Reference Model	MAP
IRS [22]	0.20
IAOCOOT [32]	0.3945
Wikipedia [32]	0.3166
WordNet [32]	0.2901
Proposed Model	0.715

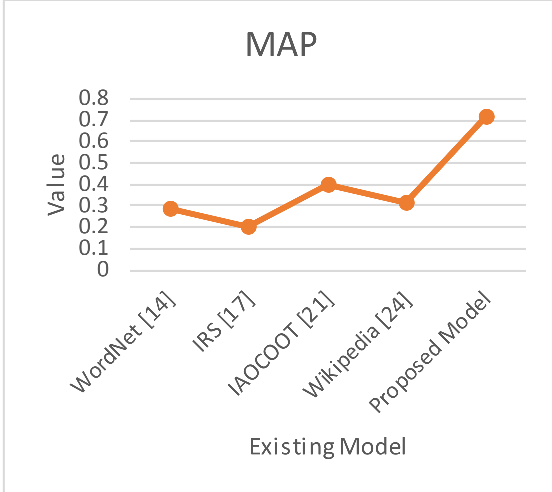


Fig. 5: Comparison with state of art model

The examination of the suggested MAP results and their comparison with state-of-the-art models is detailed in Table 3. The proposed model achieved better MAP results than competing state-of-the-art models as shown in fig. 5.

V. CONCLUSION

Conventional semantic information retrieval methods may converge slowly when presented with enormous amounts of complex content. Findings from ontology-based query transformation show that including domain ontologies in ML-based information retrieval systems improves retrieval performance. Using semantic association, the indexing strategy was able to better capture terms associated with documents. Thus, the two parts of the proposed system coordinated to ensure that the appropriate papers were located for the appropriate individuals at the appropriate times. Through the application of ontology and semantics approaches, TREC datasets showed improved performance. The window function computed, retrieved the state parameters of the ideal semantic information choice, and fused its data levels to gain information fusion outcomes utilizing unstructured semantic data processing. In comparison to earlier studies, the suggested strategy outperformed the competition when we monitored the overall amount of document obtained from major search engines and used metrics like as f-measure, precision, MAP and recall. There was a 10% improvement in accuracy after broadening the query. Our experimental results proved that the improved technique may improve semantic information retrieval efficiency without sacrificing accuracy since it sidesteps common problems with traditional methods. The algorithm will need to undergo more testing on large datasets in the coming decades.

REFERENCES

[1] Yuvaraj, D., Alnuaimi, S. S., Rasheed, B. H., Sivaram, M., & Porkodi, V. (2024). Ontology Based Semantic Enrichment for Improved Information Retrieval Model. International Journal of Intelligent Systems and Applications in Engineering, 12(15s), 70–77.

[2] Chebil, Wiem, Mohammad Wedyan, Moutaz Alazab, Ryan Alturki, and Omar Elshaweesh. 2023. "Improving Semantic Information

Retrieval Using Multinomial Naive Bayes Classifier and Bayesian Networks" Information 14, no. 5: 272. https://doi.org/10.3390/info14050272
[3] Anil Sharma and Suresh Kumar. 2023. Machine learning and ontology-based novel semantic document indexing for information retrieval. Comput. Ind. Eng. 176, C (Feb 2023). https://doi.org/10.1016/j.cie.2022.108940
[4] Hamane, Zakaria & Samih, Amina & Fennan, Abdelhadi. (2023). ONTOLOGY MATCHING USING DEEP LEARNING. Journal of Theoretical and Applied Information Technology. 101.
[5] Jain, Shivani & K.R., Seeja & Jindal, Rajni. (2021). A fuzzy ontology framework in information retrieval using semantic query expansion. International Journal of Information Management Data Insights. 1. 100009. 10.1016/j.jjimei.2021.100009.
[6] Hu, M. . (2022). Research on Semantic Information Retrieval Based on Improved Fish Swarm Algorithm. Journal of Web Engineering, 21(03), 845–860. https://doi.org/10.13052/jwe1540-9589.21313
[7] P. Lai, N. Phan, H. Hu, A. Badeti, D. Newman and D. Dou, "Ontology-based Interpretable Machine Learning for Textual Data," 2020 International Joint Conference on Neural Networks (IJCNN), Glasgow, UK, 2020, pp. 1-10, doi: 10.1109/IJCNN48605.2020.9206753.
[8] Rawal, S. (2020). Multi-Perspective Semantic Information Retrieval in the Biomedical Domain. ArXiv, abs/2008.01526.
[9] Elnagar, Samaa & Yoon, Victoria & Thomas, Manoj. (2020). An Automatic Ontology Generation Framework with An Organizational Perspective. 10.24251/HICSS.2020.597.
[10] Abhishek Kumar Shukla, Sujoy Das, Pushpendra Kumar, Afroj Alam, "Relevance Feedback and Deep Neural Network-Based Semantic Method for Query Expansion", Wireless Communications and Mobile Computing, vol. 2022, Article ID 6789044, 11 pages, 2022. https://doi.org/10.1155/2022/6789044
[11] Mahalakshmi, P. & Fathima, N.. (2021). An Art of Review on Conceptual based Information Retrieval. Webology. 18. 21-31. 10.14704/WEB/V18SI02/WEB18009.
[12] Selvalakshmi, B., Subramaniam, M. Intelligent ontology based semantic information retrieval using feature selection and classification. Cluster Comput 22 (Suppl 5), 12871–12881 (2019). https://doi.org/10.1007/s10586-018-1789-8
[13] Shah, Ritesh & Guide, Prof. (2019). Ensemble based Machine Learning using Ontology Information Extraction for Information Retrieval. International Journal of Innovative Technology and Exploring Engineering. 9. 3952-3959. 10.35940/ijitee.B7011.129219.
[14] Alothman AF, Wahab Sait AR. Managing and Retrieving Bilingual Documents Using Artificial Intelligence-Based Ontological Framework. Comput Intell Neurosci. 2022 Aug 25;2022:4636931. doi: 10.1155/2022/4636931. PMID: 36059407; PMCID: PMC9436537.
[15] P. Liu, Z. Zheng and Q. Su, "Cross-Language Information Retrieval Based on Multiple Information," 2018 IEEE/WIC/ACM International Conference on Web Intelligence (WI), Santiago, Chile, 2018, pp. 623-626, doi: 10.1109/WI.2018.00-26.
[16] Sridevi, U. K., P. Shanthi, and N. Nagaveni. "Deep Model Framework for Ontology-Based Document Clustering." In Handbook of Research on Investigations in Artificial Life Research and Development, edited by Maki Habib, 424-435. Hershey, PA: IGI Global, 2018. https://doi.org/10.4018/978-1-5225-5396-0.ch019
[17] Lamurias A, Jesus S, Neveu V, Salek RM, Couto FM. Information Retrieval Using Machine Learning for Biomarker Curation in the Exposome-Explorer. Front Res Metr Anal. 2021 Aug 19;6:689264. doi: 10.3389/frma.2021.689264. PMID: 34490412; PMCID: PMC8417071.
[18] Wang, Zhe & Zhao, Yingying & Dong, Hai & Xu, Yulong & Lv, Yali. (2021). Improved Algorithm Based on Decision Tree for Semantic Information Retrieval. Intelligent Automation & Soft Computing. 29. 419-429. 10.32604/iasc.2021.016434.
[19] Junliang, Zhang & Mostafa, Javed & Tripathy, Himansu. (2002). Information Retrieval by Semantic Analysis and Visualization of the Concept Space of D-Lib® Magazine. D-Lib Magazine. 8. 10.1045/october2002-zhang.
[20] Anand, Sanjay & Kumar, Suresh. (2022). Ontology-based Soft Computing and Machine Learning Model for Efficient Retrieval. 10.21203/rs.3.rs-2328237/v1.
[21] Trabelsi, M., Chen, Z., Davison, B.D. et al. Neural ranking models for document retrieval. Inf Retrieval J 24, 400–444 (2021). https://doi.org/10.1007/s10791-021-09398-0
[22] Kumar, R., Sharma, S.C. Hybrid optimization and ontology-based semantic model for efficient text-based information retrieval. J Supercomput 79, 2251–2280 (2023). https://doi.org/10.1007/s11227-022-04708-9
[23] Vijayarajan, V., Dinakaran, M., Tejaswin, P. et al. A generic framework for ontology-based information retrieval and image retrieval in web data. Hum. Cent. Comput. Inf. Sci. 6, 18 (2016). https://doi.org/10.1186/s13673-016-0074-1
[24] Esteva, A., Kale, A., Paulus, R. et al. COVID-19 information retrieval with deep-learning based semantic search, question answering, and abstractive summarization. npj Digit. Med. 4, 68 (2021). https://doi.org/10.1038/s41746-021-00437-0
[25] Kulmanov M, Smaili FZ, Gao X, Hoehndorf R. Semantic similarity and machine learning with ontologies. Brief Bioinform. 2021 Jul 20;22(4):bbaa199. doi: 10.1093/bib/bbaa199. PMID: 33049044; PMCID: PMC8293838.
[26] Shakeri, Maryam, Abolghasem Sadeghi-Niaraki, Soo-Mi Choi, and Tamer AbuHmed. 2022. "AR Search Engine: Semantic Information Retrieval for Augmented Reality Domain" Sustainability 14, no. 23: 15681. https://doi.org/10.3390/su142315681
[27] M. N. Asim, M. Wasim, M. U. Ghani Khan, N. Mahmood and W. Mahmood, "The Use of Ontology in Retrieval: A Study on Textual, Multilingual, and Multimedia Retrieval," in IEEE Access, vol. 7, pp. 21662-21686, 2019, doi: 10.1109/ACCESS.2019.2897849.
[28] Calvier, Francois-élie & Plantié, Michel & Dray, Gérard & Ranvez, Sylvie. (2013). Ontology Based Machine Learning for Semantic Multiclass Classification.
[29] Sharma, Anil & Kumar, Suresh. (2022). Shallow Neural Network and Ontology-Based Novel Semantic Document Indexing for Information Retrieval. Intelligent Automation & Soft Computing. 34. 1989-2005. 10.32604/iasc.2022.026095.
[30] D. Yasin, A. Sohail and I. Siddiqi, "Semantic Video Retrieval using Deep Learning Techniques," 2020 17th International Bhurban Conference on Applied Sciences and Technology (IBCAST), Islamabad, Pakistan, 2020, pp. 338-343, doi: 10.1109/IBCAST47879.2020.9044601.
[31] Kumar, Ram & Sharma, Subhash. (2022). Smart Information Retrieval using Query Transformation based on Ontology and Semantic-Association. International Journal of Advanced Computer Science and Applications. 13. 10.14569/IJACSA.2022.0130446.
[32] Liji S K, Muhamed Ilyas P, "Review and Analysis of Different Approaches to Semantic Level Question Answering and Information Retrieval", International Journal of Science and Research (IJSR), Volume 10 Issue 1, January 2021, DOI: 10.21275/SR21121141135.

Supervised ML based Model to Predict Satisfaction of Student Performance During the Pandemic

Ms. Prajkta P. Chapke
Department of Computer Science and Engineering
H.V.P.M's, C.O.E.T.
Amravati, India.
prajkta.chapke@rediffmail.com

Dr. Anjali B. Raut
Department of Computer Science and Engineering
H.V.P.M's, C.O.E.T.
Amravati, India.
anjali_dahake@rediffmail.com

Abstract*—The coronavirus, COVID-19 has spread around the globe, causing unprecedented changes in major areas including education and the economy. Amid a pandemic, the educational sector is relying heavily on performance prediction models to aid in the study of student status and the subsequent improvement of performance. The term "Educational Data Mining" describes the practice of using data mining principles and methods in the classroom. The explosion of educational data in recent years has made educational data mining a hot issue. Methods for discovering latent patterns in educational data might be developed within this area. The data can be learned more about student's, their learning habits, and how well they will do in school by applying the techniques drawn from Educational Data Mining. This study aims to predict how universities would modify e-learning in the aftermath of the outbreak by investigating student's actual experiences utilizing e-learning facilities during this time. A formal questionnaire was distributed and an online survey was created for this purpose. Following an exhaustive survey, a good outcome was achieved in terms of the dataset. In addition, the selected supervised machine learning algorithms such as Multi-Layer Perceptron (MLP), Random Forest (RF), Support Vector Machine (SVM), and XG Boost (XGB) are utilized to train the prediction model to determine how likely students are to adopt the e-learning approach in the future. MLP produced the best results compared to other approaches, with an accuracy of 84.1%. The unexpected emergence of the epidemic caused tremendous disruptions to the academic pursuits of people.

Keywords—Educational Data Mining, Covid-19, Machine Learning, Supervised Classification, Students Performance.

I. INTRODUCTION

The global outbreak of COVID-19 has had a profound impact on individuals' daily schedules and work methodologies. Governments throughout the world have imposed strict curfews and quarantines to stem the spread of the COVID-19 epidemic. Global educational systems have been hit hard by the effects of the Coronavirus epidemic in the past several years, with many schools and colleges having to close their doors permanently. Many governments worldwide have implemented temporary closures of colleges and universities to manage the COVID-19 pandemic [1]. As a result of the unpredictability of the pandemic's duration, the majority of educational institutions have implemented emergency remote learning using online learning platforms instead of traditional classroom instruction [2-5]. Courses that were once only offered in person are quickly moving online so that students may complete their education without having to wait as long to graduate, go on to graduate school or get a job. Getting an education in the traditional sense is a great way to learn about the world and its people. The capacity to distinguish between good and wrong is also bestowed upon the human mind by this process of education. People who are engaged in learning receive training when an educator imparts knowledge, abilities, and facts to them.

Data mining refers to the process of discovering useful information inside certain datasets. Discovering correlations between parameters and extracting hidden patterns in massive amounts of data are both made easier with its guidance. Many academics nowadays use Data Mining to address practical issues in fields including marketing, telecommunications, healthcare, medicine, industry, and customer relationship management [1-3]. More and more, data mining has found its way into the classroom as of late [4]. In today's universities, the academic success of each student is crucial. For the simple reason that a school's track record of student achievement is one indicator of its quality [5]. Forecasting how well students will do in school is a major issue for education administrators and other professionals in the sector. If the prediction is accurate, it might alert at-risk pupils promptly about their impending academic failure [6]. The purpose of Educational Data Mining is to use machine learning (ML) techniques to uncover patterns in educational data in order to predict how well kids will do in school. One possible use case for this is building a prediction model.

Some of the data analysis and processing methods used in data mining include association rules, clustering or classification, and sequence analysis [8-10]. Using a classification technique, we may label every single object in a collection. Every example in the dataset needs a precise prediction of the target class, therefore this is done [10]. There are several techniques used in educational data mining, such as Decision Trees (DT), Support Vector Machine (SVM), Naïve Bayes (NB), Random Forest (RF), Artificial Neural Networks (ANN), K-Nearest Neighbour (KNN), and others. As a basic stance in an informative setting, predicting and analysing student execution is essential. The educational execution of students is a crucial component in their future development. The academic success of a student is dependent on more than just one factor; it is also strongly influenced by factors such as socioeconomic status, psychological makeup, and other contextual factors. The main objectives of this paper are:

1. Designed Supervised ML based framework to predict whether the student performance satisfactory during the COVID-19 pandemic.
2. Tested classification algorithms by assessments made experimentally in relation to the parameters of the confusion matrix.
3. Created a predictive model to forecast the probability of students embracing e-learning and its resources in the decades following the current pandemic.

The outline of this research work is presented as follows. Section II provides a comprehensive review of the pertinent literature. The content and processes are defined in Section III. Section IV explores the outcomes and their interpretation of the experiments. Finally, the study concludes in Section VI by presenting the implications that have been deduced from the obtained data.

II. RELATED WORKS

In terms of predicting student happiness, the decision tree classifier was more effective than the naïve Bayes. Among other factors, the presence of a feeling of community during online learning was the most influential factor in determining satisfaction [2]. A combination of ANN and RF model used to forecast how students would feel and how satisfied they would be with the course [3]. A wider context applied since they demonstrate that retention prediction models trained on one institution can achieve respectable results when applied to a different institution [4]. Supervised ML techniques are used to investigate probationary college students' poor academic performance [5]. The utilization of two machine learning techniques, specifically SVM and KNN, enables the prediction of the rate of churn for EdTech courses based on student feedback obtained from the course end assessment. [6]. An investigation on the effects of distance learning on student well-being within the COVID-19 pandemic in Kenya, with particular attention to the moderating role of student preferences in this relationship [7]. A descriptive-analytical approach was employed to analyse the viewpoints of the sample about the quality assessment of distance learning courses using the Maharat platform at Taif University. The objective was to determine and assess the extent to which requirements for distance learning courses have been met [8]. During this pandemic circumstance, university students in Sri Lanka used e-learning tools to research and predict how institutions will adapt to e-learning [9]. The accuracy of predicting students' performance in the COVID-19 test in distance education was enhanced by comparing the innovative Inception technique with the Google algorithm [10]. Examining and contrasting six ML models based on Thailand's educational curriculum for early performance prediction of students [11]. A comprehensive emotion mining (EM) and sentiment analysis were conducted on distance learning-related tweets in the Arabic language throughout the outbreak. [12]. To enhance the effectiveness of online learning, ML model is developed for predicting learning success, which includes extra steps for pre-training with fine-tuning, and builds a mechanism for generating adapted feedback [13]. The level of student pleasure mentioned during a tutoring discussion is a significant measure of its effectiveness [14]. Researchers looked at how to use ML model to gauge students' happiness with their online courses in a wireless network setting [15]. The hypothesis of psychological contracts and their potential connection to students' happiness with online education was investigated [16]. The main learning tool and the primary indicators for undergraduate students' satisfaction with ERL are these two platforms [17]. Examining the relationship between student happiness and engagement in e-classroom study activities and performance is evaluated [18]. An effective method for sentiment categorization that followed the tenets of deep learning (DL) and ensemble learning showed strong prediction abilities in evaluations of massive open online courses (MOOC) [19]. Using sentiment analysis on Twitter, one of the most prominent social media platforms, the community has come to embrace remote learning as a safety measure [20]. Predictive algorithms that are appropriate for early dropout prediction systems at universities in online learning are intended to be presented using the learner's statistical information and computed data, furthermore, the data that is documented inside the system for learning management [21]. The program used ANN to analyse data from students at the University of Tabuk to determine what characteristics contribute to their level of satisfaction with online learning [22]. Analysing data collected over successive years allowed us to anticipate student participation in the early stages of a virtual learning environment course [23]. Research into the potential of automated categorization for the semantic content of evaluations of massive open online courses (MOOCs) aimed at identifying elements that, according to learners, might indicate their level of satisfaction with the courses [24]. A Deep Neural Network–based approach was suggested for predicting students' grades [25]. An approach that utilizes data mining techniques to forecast how well computer science students would do in school [26].

The most renowned research on forecasting student's performance achievement during pandemic emphasizes certain areas of strength and regions of deficiency. The evaluations were deemed ineffective due to their lack of specificity, failure to prioritize student outcomes as a metric of student achievement, issues with quality, and limited accessibility through publicly searchable platforms. This work aims to develop a predictive model for determining student's academic achievement by analysing their online activities during the COVID-19 pandemic, utilizing recently gathered statistical indicators.

III. MATERIALS AND METHODS

A. Dataset Collection

Due to the prevalence of student surveys in recent academic research, this study includes surveys from universities in the Amravati City area. It calls for a low-cost method of data gathering and analysis that is both efficient and quick, as well as one that reduces administrative error. In online surveys, the absence of constraints such as digitally produced followers and specific preferences is observed, the pool of potential learners is larger. Furthermore, in this urgent scenario, the most vital component is to obtain the necessary and accurate data on time. Collecting the necessary data was the backbone of the study. To achieve this goal, a Google Forms questionnaire was created and sent via email and social media to student organizations and forums at eighteen different institutions.

Three types of questions make up the questionnaire. Gathering the respondent's details is the first step. Part two entails collecting information on the tools and resources utilized for online education, part three concerns the respondent's language skills as they relate to the features offered by online learning platforms, and part four gathers data regarding e-learning's practical applications. Data was saved in a comma-separated values format from a total of 500 questionnaire replies. Tables I, II, and III below outline the components of the dataset, which include both independent and dependent variables.

TABLE I. DATASET INPUT PARAMETERS

Name of Attributes	No of Values	Description
Gender	2	Male, Female
Year of Study	4	First, Second, Third, Fourth
Age Group	3	18-24, 25-30, 30+
UG Percentage	5	+90, 80-89, 70-79, 60-69, Below 60
Residence	4	Remote area, Village, City, Metro city
Parents Education	4	Below 10th, 12th, UG, PG
Use of Digital tools for Parents	3	Highly expert, Familiar, Not Familiar
network issue while learning online	2	Yes, No

TABLE II. DATASET INPUT PARAMETERS WITH QUESTIONS

Attributes	Questions	Descriptions
Use of Digital Tools	Q1	How often do you use each of these digital tools?
	Q2	How often do you use the various digital learning resources?
	Q3	How often do you use the digital tools (mobile, i-pad, laptop) while studying?
	Q4	Using the digital tools for e-learning causes me to become distracted and unable to focus?
	Q5	Am I proficient in the necessary technologies or applications?
Sleeping Habits	Q1	How much times that student goes to bed and wakeup?
	Q2	Extended usage of digital learning tools has impacted the sleeping patterns?
	Q3	Online learning's constant screen time is stressful?
Social Interaction	Q1	Digitization causes social distance?
	Q2	Long-term usage of digital tools isolates students.
	Q3	University study improves students' social skills?
	Q4	Laziness and lethargy result from being home?
	Q5	Family members help in learning online?
Psychological State	Q1	Extended usage of e-learning technologies can cause boredom, anxiety, and stress?
	Q2	The psychological aspect is crucial to educational performance?
	Q3	An embarrassment and a frustration for some students is that they cannot afford to purchase all of the essential digital tools?
	Q4	Online learning is detrimental socially and mentally; hence student don't encourage it?
	Q5	COVID-19 lockdowns,

		closures, and quarantines generated tension, frustration, and sadness?
	Q6	I enjoyed attending lectures and tutorials?
	Q7	I was satisfied with online learning course?
	Q8	Were parents satisfied with online learning?
Academic Performance	Q1	Intense e-learning tasks caused confusion, dissatisfaction, and poor performance?
	Q2	Academic performance improves with face-to-face engagement?
	Q3	Online quizzes and examinations made me uneasy?

TABLE III. DATASET OUTPUT PARAMETERS

Grade	GPA Grade Scale
Grade-4 (C)	60-69
Grade-3 (B)	70-79
Grade-2 (A)	80-89
Grade-1 (A+)	+90

B. *Methodology*

This section describes, using chosen categorization methodologies, the analytical flowchart of this work. The process of creating the prediction model for satisfaction of student performance during the pandemic using supervised machine learning based model is illustrated as shown in figure 1.

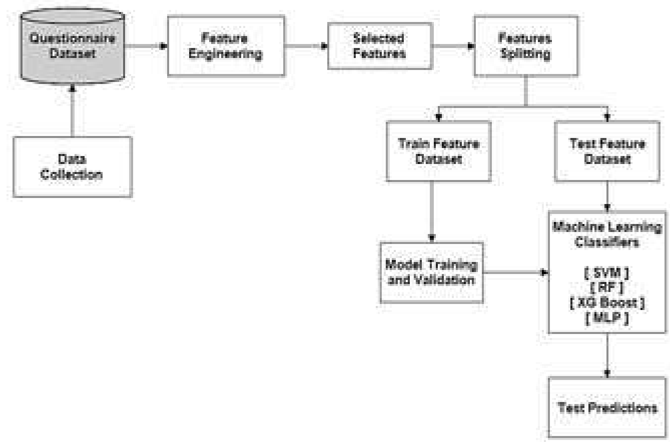


Fig. 1. Proposed Framework of Prediction of Student's Academic Performance

1. *Feature Engineering*

Feature engineering refers to a collection of approaches that a system may employ to detect and categorize characteristics. Included in the survey results are multiple feature columns with Likert scale replies; one column contains only binary responses; and the remaining feature columns offer responses depending on a variety of category values, including the target GPA. Label encoding is a common method for transforming data into a numerical or binary format that machines can understand and interpret. To make the data machine-readable, labels are encoded each data point. Each category feature undergoes a single hot encoding. The feature data is cleaned up by removing any NaN values. To decrease variance in the test dataset, retain only one feature if its association with goal value is weak or if two characteristics are highly associated. A disproportionate percentage of people with lower incomes are on list of target labels. An asymmetrical sample may give rise to a biased model. Using the Synthetic Minority Oversampling Technique (SMOTE), the samples from the minority class were shown to be more reflective of the majority class. The data frame was subjected to many data mining algorithms following its partitioning into a test and train set. For the most accurate classification model, divide the feature data into a 30% test set and a 70% train set.

2. *Machine Learning Modelling*

The subsequent stage of the flow of the prediction model is the process of extracting patterns, training the model, and evaluating the performance of the model. Making use of the training dataset from the prior stage, the model was built using supervised classification algorithms. When the data mining phase entered, things began to heat up. The utilization of 10-fold cross-validation is employed to adjust the parameters of the classifier. Nevertheless, to mitigate the impact of experimental bias and determine the optimal fold selection for the prediction framework, k-fold cross-validation is employed. The testing information was also used to assess a model that may predict how well students will do in class. Now the data is collected and analyse it and put it into a knowledge base. The project makes use of classifiers from Multi-Layer Perceptron Neural Network (MLP-NN), Random Forest (RF), Support Vector Machine (SVM) and the Extreme Gradient Boosting (XGB) families.

2.2.1. Multi-Layer Perceptron Neural Network (MLP-NN)

A feed-forward multi-layered artificial neural network, MLP-NN takes a set of inputs and produces a set of outputs. MLP-NN is constructed by connecting the input and output layers of a directed graph with many layers of input nodes. One way to train an MLP-NN network is by backpropagation. A multi-layer perceptron learning model is illustrated in the figure 2 below.

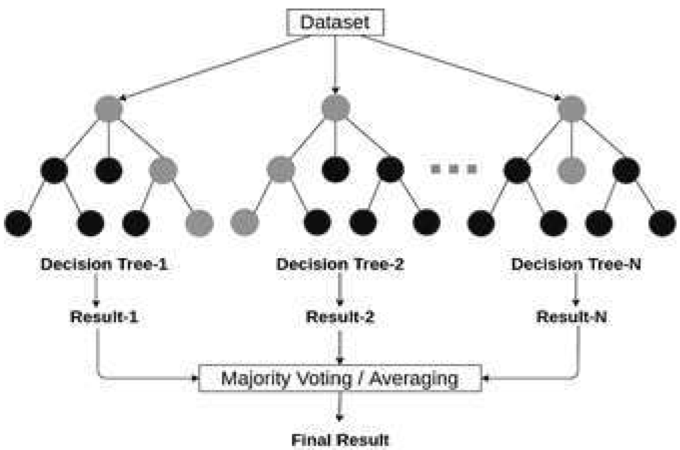


Fig. 2. Layer representation of MLP-NN

2.2.2. Random Forest (RF)

Random Forest (RF) is a technique used to enhance the accuracy of predictions and mitigate overfitting. It achieves this by calculating the mean outcome from several decision tree classification methods, each trained on a distinct subset of the dataset. A random forest takes an average of predictions made by several decision trees rather than relying on a single tree to make predictions. Overfitting is less likely to occur when there are more trees in the forest, leading to more trustworthy outcomes. The Random Forest approach in action in the figure 3 below.

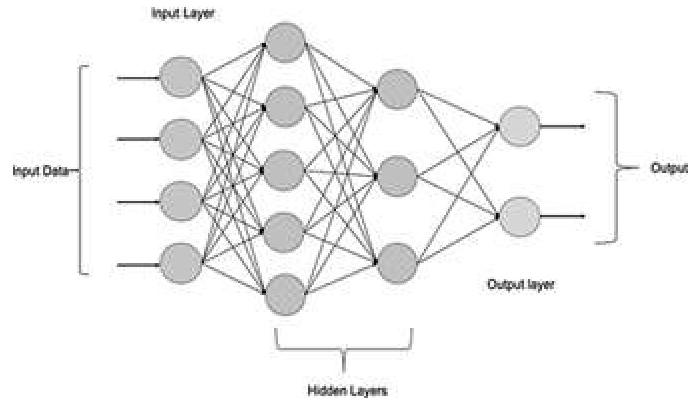


Fig. 3. Working of RF algorithm

2.2.3. Support Vector Machine (SVM)

SVM model uses a multidimensional hyperplane to project the different categories. By iteratively generating the hyperplane, SVM aims to minimize the error to its maximum potential. Finding a maximum marginal hyperplane (MMH) is the end objective of SVMs, which classify datasets into categories. The linear hyperplane equation can be expressed as:

$$w^T x + b = 0$$

Take into account a problem of binary classification where two classes are denoted as +1 and -1. The training dataset consists of input feature vectors and class labels, denoted as X and Y, respectively. The "W" vector denotes a normal vector, which is a vector with a direction that is orthogonal to a hyperplane. The hyperplane's distance along the normal vector "w" from the origin is represented by the parameter "b" in the equation. The training dataset comprises input vectors of features X and their corresponding class labels "Y".

$$\hat{y} = \begin{cases} 1 & : w^T x + b \geq 0 \\ 0 & : w^T x + b < 0 \end{cases}$$

2.2.4. XG Boost (XGB)

XGBoost has emerged as a prominent machine learning methodology due to its remarkable scalability and the capability to surpass contemporary techniques in domains like as classification and regression. XGBoost classifier fine-tuned for scalability and speed, making it an ideal choice for training machine-learning models. It's a method of ensemble learning that combines the predictions of several weak models into a single, more solid prediction and its implementation flow given in figure 4.

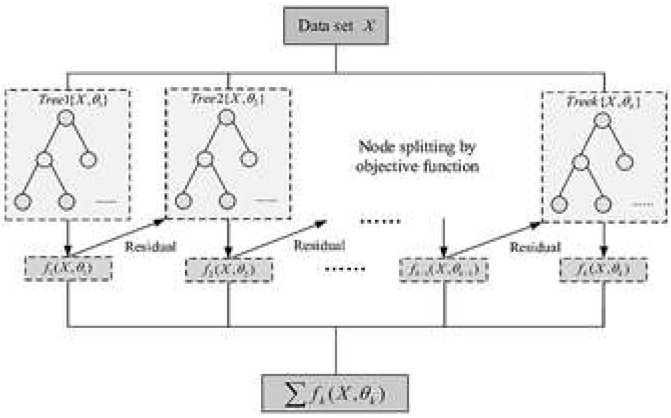


Fig. 4. Implementation flow of XG Boost

IV. EXPERIMENTAL RESULTS AND DISCUSSION

A. EXPERIMENTAL SETUP

The primary objective is to use the retained input variables of the model to predict how well a student perform. The classification model was built using a variety of machine-learning approaches, and their respective performances were evaluated. The system requirements call for an Intel Core i5 CPU and 16 GB of RAM, while the software packages include Anaconda Distribution and the Pycharm IDE. Several ML algorithms may be implemented using Python's Scikit Learn package. Each classifier is tested using cross-validation using a total of 70% training data and 30% testing data for assessment.

B. EVALUATION PARAMETERS

A dataset comprising questionnaires is utilized in the proposed experiments to test the prediction model. The following metrics are used to evaluate the accuracy of the classification: Accuracy, Recall, Precision, and F-score. One way to evaluate a prediction system's efficacy is to look at the ratio of successful predictions to total occurrences. The accuracy of a forecast is defined as the frequency with which it is correct. The metric known as recall refers to the proportion of correctly anticipated instances inside a certain real-world class. The F-measure was computed in order to present the harmonic mean by combining recall and accuracy. Calculating these four metrics is done by applying the following equation to the confusion matrix.

$$ACC = \frac{TP+TN}{P+N} = \frac{TP+TN}{TP+TN+FP++FN}$$

$$Recall = \frac{TP}{TP+FN}$$

$$Precision = \frac{TP}{TP+FP}$$

$$F = 2.\frac{precision.recall}{precision+recall}$$

Where TP – True Positive, TN – True Negative, FP – False Positive, FN – False Negative.

C. RESULTS ANALYSIS

The dataset was subjected to four categorization methods, resulting in a range of conclusions and further analysis. The cross-validation technique was employed to evaluate four machine learning classifiers, and the evaluation results are presented in Table 2 to 6. Table 2 displays the runtimes of the prediction model's training and testing procedures. Compared to rest classifiers, MLP classifiers need the largest proportion of training time as shown in figure 5 since they have more hyperparameters for tuning.

TABLE IV. ASSESSMENT TIME OF THE MODEL

Classifiers/ Parameters	**Train Time (s)**	**Test Time (s)**
MLP-NN	55.8	26.5
RF	43.1	19.2
SVM	46.4	22.1
XGB	48.5	23.4

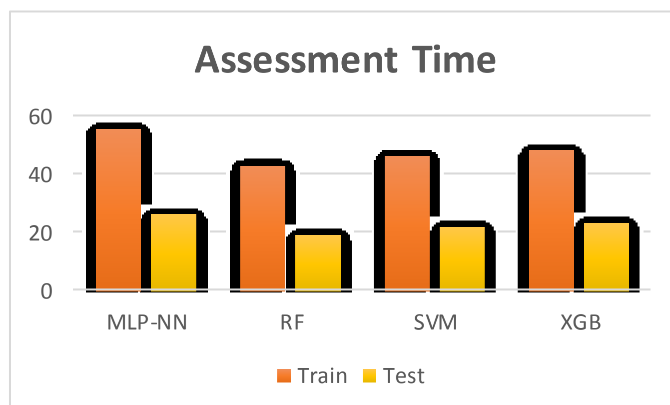


Fig. 5. The assessment time of the prediction model

The results produced by the RF classifier outperform those of MLP and other competitors. It gets 83.1% of the classes correctly, which means that out of the four potential classes in the training dataset, most students were placed in one of the other three. Recall, precision, and f-measure values for the testing model were 0.89, 0.93, and 0.89, respectively, indicating stable predictive performance. The outcomes show that RF outperforms the other models we tested for the prediction model. To further assure our prediction model, a further test conducted using RF only on the testing dataset; this yields better results than MLP-NN.

TABLE V. PERFORMANCE EVALUATION OF MLP-NN MODEL

Parameters/ GPA	Grade-1	Grade-2	Grade-3	Grade-4
Recall	0.81	0.60	0.54	0.68
Precision	0.86	0.66	0.62	0.75
F-score	0.73	0.68	0.58	0.71

TABLE VI. PERFORMANCE EVALUATION OF RF MODEL

Parameters/ GPA	Grade-1	Grade-2	Grade-3	Grade-4
Recall	0.93	0.65	0.66	0.79
Precision	0.89	0.70	0.68	0.79
F-score	0.89	0.68	0.61	0.79

TABLE VII. PERFORMANCE EVALUATION OF SVM MODEL

Parameters/ GPA	Grade-1	Grade-2	Grade-3	Grade-4
Recall	0.82	0.51	0.41	0.51
Precision	0.68	0.46	0.48	0.58
F-score	0.74	0.48	0.44	0.54

TABLE VIII. PERFORMANCE EVALUATION OF XGB MODEL

Parameters/ GPA	Grade-1	Grade-2	Grade-3	Grade-4
Recall	0.82	0.57	0.59	0.66
Precision	0.80	0.59	0.64	0.71
F-score	0.81	0.58	0.56	0.78

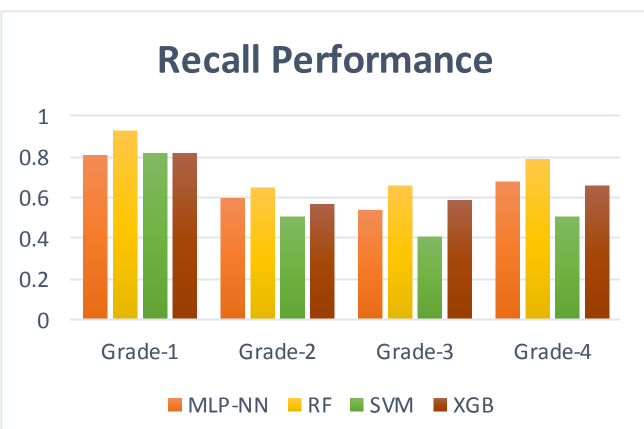


Fig. 6. Prediction Model Performance Using Recall

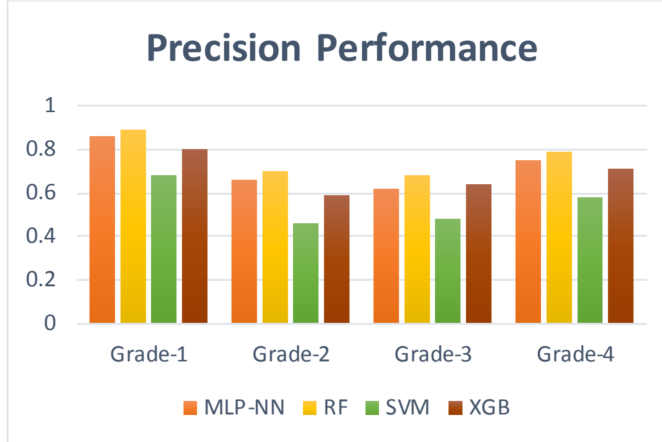


Fig. 7. Prediction Model Performance Using Precision

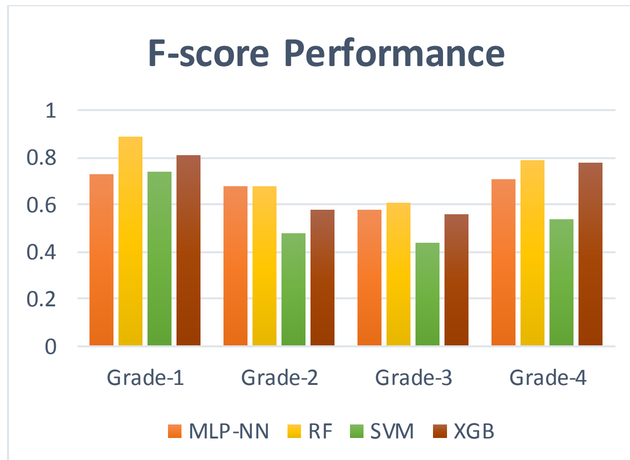


Fig. 8. Prediction Model Performance Using F-score

The factors that impacted the value of our focal variable (CGPA) are also studied. Students' cumulative grade point averages were negatively impacted by their usage of online resources for instruction during the lockdown, according to its results. Overall grade point average was inversely related to the positive Likert scale ratings of these factors. The harmful impact of students spending too much time on digital tools on their academic achievement was confirmed once again. Low cumulative grade point averages were also more common among students who blamed their e-learning tasks for their excessive home-schooling. Distractions from the digital instruments used for the online study had the greatest negative correlation with the student's final score. We found the same thing as the study that found COVID-19 significantly affected kids' academic performance.

D. COMPARATIVE ANALYSIS

The statistical parameter, F-measure of 74.55%, a recall of 75.8%, a precision value of 76.6%, and an accuracy of 84.1% are all shown in Table 9, which illustrates the outcomes of our model. This proves that our model's prediction is accurate. Figure 9 depicts the absence of overfitting from the train to test dataset, which is evident from the balanced recall and accuracy performance using RF classifier algorithms.

TABLE IX. STATE OF ART MODEL COMPARATIVE ANALYSIS

Parameters/ References	Ref [13]	Ref [19]	Ref [21]	Proposed Model
Accuracy (%)	63.51	76.19	68.7	84.1
Recall (%)	63.5	52.70	68.9	75.8
Precision (%)	59.4	66.10	69.1	76.6
F-score (%)	61.4	58.64	67.8	74.5

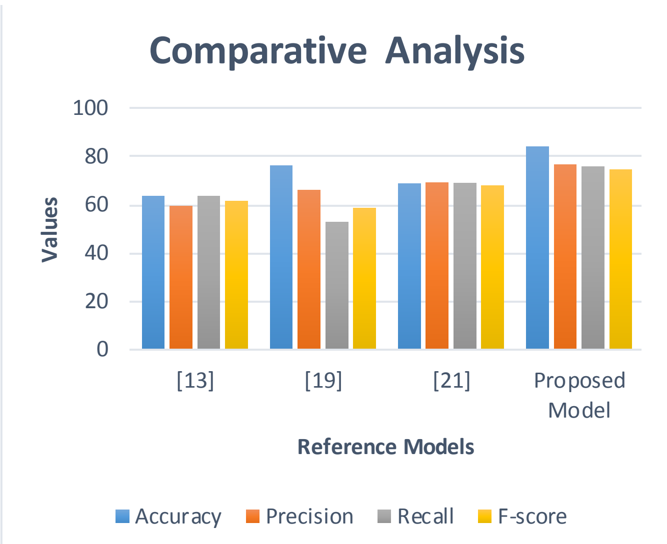


Fig. 9. Comparison with state-of-art Model

CONCLUSION

The focus of this research work was on the effects of the pandemic on students' capacity to learn. In this study, data collected on the effects of the lockdown on student's mental health and academic performance as well as their use of internet resources. The one of four machine learning models used to forecast the student's total GPA. Excessive use of digital study tools and social isolation caused by the lockout are associated with worse cumulative grade point averages, according to the research. Additionally, the results show that RF, with an accuracy score of 84.1%, is the top classification technique, with MLP-NN coming in a close second. SVM was the worst algorithm in the test. According to the research, students should be given fewer online assignments during the lockdown to prevent the possible harmful impacts of using digital tools too much in the classroom. Also, make sure that students aren't sitting alone in class. It should be remembered that student's academic performance might be negatively impacted by excessive use of digital learning tools, particularly when combined with the social isolation caused by the pandemic. Further, the current paradigm of online education needs revision to better support student's emotional well-being using deep learning approach with large scale of dataset.

REFERENCES

[1] Pinto, A. S., Abreu, A., Costa, E., and Paiva, J. (2023). How Machine Learning (ML) is Transforming Higher Education: A Systematic Literature Review. Journal of Information Systems Engineering and Management, 8(2), 21168. https://doi.org/10.55267/iadt.07.13227.

[2] Kakish, Hadeel & Al-Eisawi, Dalia. (2023). A Predictive Model for Assessing Satisfaction with Online Learning for Higher Education Students During and After COVID-19 Using Data Mining and Machine Learning Techniques: A Case of Jordanian Institutions. 156-163. 10.5220/0011844100003467.

[3] Abdullah AL Mutawa & Sai Sruthi (2023) Enhancing Human–Computer Interaction in Online Education: A Machine Learning Approach to Predicting Student Emotion and Satisfaction, International Journal of Human–Computer Interaction, DOI: 10.1080/10447318.2023.2291611.

[4] Matz, S.C., Bukow, C.S., Peters, H. et al. Using machine learning to predict student retention from socio-demographic characteristics and app-based engagement metrics. Sci Rep 13, 5705 (2023). https://doi.org/10.1038/s41598-023-32484-w.

[5] Al-Alawi, L., Al Shaqsi, J., Tarhini, A. et al. Using machine learning to predict factors affecting academic performance: the case of college students on academic probation. Educ Inf Technol 28, 12407–12432 (2023). https://doi.org/10.1007/s10639-023-11700-0.

[6] R. Paul and M. Rashmi, "Student Satisfaction and Churn Predicting using Machine Learning Algorithms for EdTech course," 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2022, pp. 1-6, doi: 10.1109/ICRITO56286.2022.9965116.

[7] Li X, Odhiambo FA, Ocansey DKW. The effect of students' online learning experience on their satisfaction during the COVID-19 pandemic: The mediating role of preference. Front Psychol. 2023 Jan 31;14:1095073. doi: 10.3389/fpsyg.2023.1095073. PMID: 36798893; PMCID: PMC9928211.

[8] Mohammed Nasser Alsubaie, Predicting student performance using machine learning to enhance the quality assurance of online training via Maharat platform, Alexandria Engineering Journal, Volume 69, 2023, Pages 323-339, ISSN 1110-0168, https://doi.org/10.1016/j.aej.2023.02.004.

[9] S. Prasanth, K. Banujan, B. Kumara and H. Rupasingha, "Prediction of Student Satisfaction on Online Learning during the COVID-19 Pandemic - A Case Study on Sri Lankan Universities," 2021 21st International Conference on Advances in ICT for Emerging Regions (ICter), Colombo, Sri Lanka, 2021, pp. 69-74, doi: 10.1109/ICter53630.2021.9774807.

[10] P. Sathvik and S. Kalaiarasi, "Improved Accuracy by Novel Inception Compared over GoogleNet in Predicting the Performance of Students in Online Education During COVID", E3S Web Conf., 399 (2023) 04021, DOI: https://doi.org/10.1051/e3sconf/202339904021.

[11] Chayaporn Kaensar, Worayoot Wongnin, "Predicting new student performances and identifying important attributes of admission data using machine learning techniques with hyperparameter tuning", EURASIA Journal of Mathematics, Science and Technology Education, 2023, 19(12), em2369, https://doi.org/10.29333/ejmste/13863.

[12] Ali, Manal. (2021). Arabic sentiment analysis about online learning to mitigate covid-19. Journal of Intelligent Systems. 30. 524-540. 10.1515/jisys-2020-0115.

[13] Wang X, Zhang L, He T. Learning Performance Prediction-Based Personalized Feedback in Online Learning via Machine Learning. Sustainability. 2022; 14(13):7654. https://doi.org/10.3390/su14137654

[14] Chen, Guanliang et al. "Predictors of Student Satisfaction: A Large-scale Study of Human-Human Online Tutorial Dialogues." Educational Data Mining (2019).

[15] Jing Qin, "Online Education Satisfaction Assessment Based on Machine Learning Model in Wireless Network Environment", Computational and

Mathematical Methods in Medicine, vol. 2022, Article ID 7958932, 11 pages, 2022. https://doi.org/10.1155/2022/7958932.

[16] Dziuban, Chuck & Moskal, Patsy & Thompson, Jessica & Kramer, Lauren & DeCantis, Genevieve & Hermsdorfer, Andrea. (2015). Student Satisfaction with Online Learning: Is it a Psychological Contract?. Journal of Asynchronous Learning Network. 19. 10.24059/olj.v19i2.496.

[17] Ho IMK, Cheong KY, Weldon A. Predicting student satisfaction of emergency remote learning in higher education during COVID-19 using machine learning techniques. PLoS One. 2021 Apr 2;16(4):e0249423. doi: 10.1371/journal.pone.0249423. PMID: 33798204; PMCID: PMC8018673.

[18] Skrbinjek, Vesna & Dermol, Valerij. (2019). Predicting students' satisfaction using a decision tree. Tertiary Education and Management. 2019. 1-13. 10.1007/s11233-018-09018-5.

[19] Onan, Aytug. (2020). Sentiment analysis on massive open online course evaluations: A text mining and deep learning approach. Computer Applications in Engineering Education. 29. 10.1002/cae.22253.

[20] Y. Almurtadha and M. Ghaleb, "Sentiment Analysis to Measure Public Response to Online Education During Coronavirus Pandemic," 2021 National Computing Colleges Conference (NCCC), Taif, Saudi Arabia, 2021, pp. 1-5, doi: 10.1109/NCCC49330.2021.9428838.

[21] Hee Sun Park, Seong Joon Yoo, "Early Dropout Prediction in Online Learning of University using Machine Learning", JOIV: Int. J. Inform. Visualization, 5(4) - December 2021 347-353, DOI: http://dx.doi.org/10.30630/joiv.5.4.732.

[22] Dalia Kamal Fathi Alnagar. Using Artificial Neural Network to Predicted Student Satisfaction in E-learning. American Journal of Applied Mathematics and Statistics. 2020; 8(3):90-95. doi: 10.12691/ajams-8-3-2.

[23] Raj, N. S., & V G, R. (2022). Early prediction of student engagement in virtual learning environments using machine learning techniques. E-Learning and Digital Media, 19(6), 537-554. https://doi.org/10.1177/20427530221108027.

[24] Chen, Xieling & Wang, Fu Lee & Cheng, Gary & Chow, Man-Kong & Xie, Haoran. (2022). Understanding Learners' Perception of MOOCs Based on Review Data Analysis Using Deep Learning and Sentiment Analysis. Future Internet. 14. 218. 10.3390/fi14080218.

[25] Vijayalakshmi & Venkatachalapathy, K. (2019). Comparison of Predicting Student's Performance using Machine Learning Algorithms. International Journal of Intelligent Systems and Applications. 11. 34-45. 10.5815/ijisa.2019.12.04.

[26] F. J. Kaunang and R. Rotikan, "Students' Academic Performance Prediction using Data Mining," 2018 Third International Conference on Informatics and Computing (ICIC), Palembang, Indonesia, 2018, pp. 1-5, doi: 10.1109/IAC.2018.8780547.

Exploring Pseudo-Zernike Moments for Effective Image Forgery Detection

Niyati Patel
Research Scholar
Computer Engineering Department
Ipcowala Institute of Engineering and Technology
Dharmaj, Gujarat, India
niyatipatel0491@gmail.com

Dr. Premal J. Patel
Assistant Professor
Computer Engineering Department
DEPSTAR, Charusat, Off. Nadiad-Petlad Highway,
Changa, Anand, Gujarat, India
premalpatel.dce@charusat.ac.in

***Abstract*— This research explores the synergistic integration of Pseudo-Zernike moments and machine learning techniques to enhance the accuracy and robustness of image forgery detection. Leveraging the discriminative power of Pseudo-Zernike moments in capturing shape and texture features, the study evaluates various machine learning models including K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), Naive Bayes (NB), and an Ensemble Extra Trees model. Through comprehensive experimentation and evaluation on a dataset containing both authentic and manipulated images, the Ensemble Extra Trees model emerges as the most accurate and reliable approach for detecting image forgeries. The ensemble nature of Extra Trees, coupled with its ability to mitigate overfitting and reduce variance, enables it to effectively discern subtle alterations and manipulations within images, showcasing its potential for robust image forgery detection in digital image forensics.**

Keywords*— Pseudo-Zernike moments, Ensemble Extra Trees, K-Nearest Neighbors, Decision Tree, Support Vector Machine, Random Forest, Naive Bayes.

I. INTRODUCTION

In recent years, the widespread availability of sophisticated image editing software has made it increasingly challenging to differentiate between authentic images and manipulated ones. Image forgery, encompassing techniques such as cloning, splicing, and retouching, has become a prevalent issue in various domains, including journalism, digital forensics, and multimedia content authentication. Detecting such manipulations is crucial for maintaining the integrity and credibility of digital imagery.

Fig. 1. (a) Original Image (b) Forge Image

To address this challenge, this paper investigates the fusion of Pseudo-Zernike moments and machine learning algorithms for effective image forgery detection. Pseudo-Zernike moments are mathematical descriptors known for their robustness in capturing shape and texture information, making them suitable for characterizing both global and local features within an image. Machine learning techniques, particularly ensemble learning methods such as the Ensemble Extra Trees model, are employed to leverage the extracted features and classify images as authentic or forged.

The primary objective of this research is to enhance the accuracy and reliability of image forgery detection by exploring the synergy between Pseudo-Zernike moments and machine learning. By evaluating and comparing different machine learning models, including K-Nearest Neighbors (KNN), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), Naive Bayes (NB), and the Ensemble Extra Trees model, this study aims to identify the most effective approach for detecting various types of image manipulations while minimizing false positives and negatives.

II. RELATED WORKS

G. S. Bapi [1] discusses the utilization of machine learning techniques for digital image forgery detection, highlighting the importance of advanced computational methods in addressing this prevalent issue. The study emphasizes the role of feature extraction and classification algorithms in distinguishing between authentic and manipulated images, contributing to the development of robust forgery detection systems.

P. Gabhane et al. [2] conduct a systematic review to analyze the effectiveness of Benford's Law in detecting image forgery. Their research provides insights into alternative approaches for authenticity verification, exploring statistical methods and their applicability in forensic analysis to enhance the accuracy of forgery detection systems.

N. Rathore et al. [3] propose a binary pattern approach for copy-move image forgery detection, contributing to the advancement of algorithmic techniques in identifying common forms of image tampering. Their study underscores the importance of pattern recognition and feature-based analysis in forensic investigations.

D. S. Sulaiman and M. S. M. Altaei [4] investigate image tampering detection using Extreme Learning Machine (ELM), showcasing the potential of novel machine learning paradigms in enhancing detection accuracy. Their research contributes to the development of efficient algorithms for identifying manipulated regions within images.

S. Tyagi and D. Yadav [5] provide a detailed analysis of image and video forgery detection techniques, offering a comprehensive overview of the evolving landscape of forensic analysis methods. Their study encompasses a wide range of detection approaches, highlighting the challenges and advancements in forgery detection technology.

M. Karmakar [6] explores offline signature recognition and its forgery detection using machine learning techniques, expanding the scope of digital forensics to include document authentication. The research contributes to enhancing security measures in document verification systems through automated recognition and forgery detection algorithms.

P. Gupta et al. [7] propose image forgery detection using a deep learning model, demonstrating the efficacy of deep neural networks in uncovering complex image manipulations. Their study showcases the potential of deep learning architectures in enhancing forgery detection accuracy and robustness.

A. Diwan et al. [8] enhance copy-move forgery detection using the SuperPoint keypoint architecture, contributing to the refinement of feature-based detection techniques. Their research highlights the importance of keypoints and descriptors in detecting common forms of image manipulation.

A. Kashyap et al. [9] present a robust and optimized algorithm for detecting copy-rotate-move tampering, addressing specific challenges in identifying manipulated regions within images. Their study emphasizes the importance of algorithmic optimization and efficiency in forensic analysis.

A. T. Phan-Ho and F. Retraint [10] conduct a comparative study of Bayesian and Dempster-Shafer fusion methods for image forgery detection, emphasizing the importance of fusion techniques in enhancing detection reliability. Their research contributes to exploring alternative approaches for combining evidence in forensic analysis.

A. R. Gu et al. [11] propose FBI-Net, a frequency-based image forgery localization method via multitask learning with self-attention, showcasing innovative approaches in localization-based detection strategies. Their study introduces novel techniques for accurately localizing manipulated regions within images.

K. M. Hosny et al. [12] introduce an efficient CNN model for detecting copy-move image forgery, leveraging deep learning architectures to achieve high detection accuracy. Their research contributes to advancing deep learning methods in forgery detection technology.

N. T. Pham and C. S. Park [13] provide a survey on deep-learning-based methods in image forgery detection, offering a comprehensive overview of the evolving landscape of deep learning techniques in forensic analysis. Their study summarizes recent advancements and challenges in applying deep learning to forgery detection tasks.

S. I. Lee et al. [14] propose a CNN-based copy-move forgery detection method using rotation-invariant wavelet features, contributing to the development of rotation-invariant detection algorithms. Their research showcases the potential of combining CNNs with wavelet transformations for effective forgery detection.

A. H. Khalil et al. [15] enhance digital image forgery detection using transfer learning, leveraging pre-trained models to improve detection performance and generalization capabilities. Their study highlights the importance of transfer learning in adapting existing models to new forgery detection tasks, enhancing detection accuracy and robustness.

III. PROPOSED METHODOLOGY

The proposed system flow diagram begins with the Input Image, which undergoes Pre-processing steps such as resizing to 128x128 pixels and conversion to grayscale. Subsequently, Feature Extraction using Pseudo-Zernike moments and Machine Learning classification with algorithms like SVM, KNN, DT, RF, NB, and ET are employed for predicting whether the image is forged or not. Evaluation metrics such as ACC, P, R, and F1-Score are then used to assess the performance of the system in detecting image forgery.

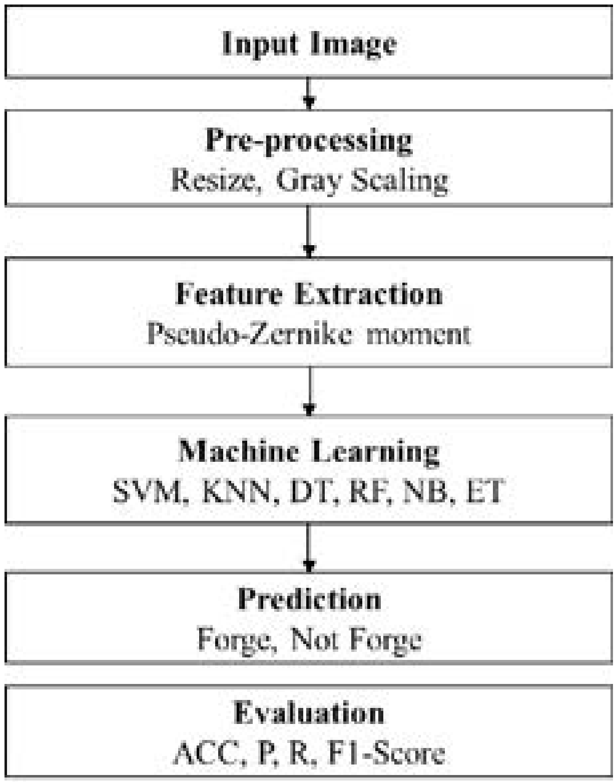


Fig. 2. Proposed System Flow

A. *Input Image:*

This is the initial step where the input image is provided to the system. The input image can be of any size and format.

B. *Pre-processing:*

Resize 128x128: The input image undergoes resizing to a fixed size of 128x128 pixels. This step ensures uniformity in image dimensions, which is often required for subsequent processing steps.

Gray Scaling: The resized image is then converted to grayscale. Grayscale images have a single channel representing the intensity of each pixel, as opposed to RGB images which have three channels (red, green, and blue). Grayscale images are commonly used in image processing tasks and reduce computational complexity.

C. *Feature Extraction:*

Pseudo-Zernike moment: Pseudo-Zernike moments are a set of features used in image analysis and pattern recognition. They capture shape and texture information from the image. Calculating Pseudo-Zernike moments involves complex mathematical operations but results in a feature vector that represents unique characteristics of the image.

D. Machine Learning:

SVM (Support Vector Machine): SVM is a supervised learning algorithm used for classification tasks. It works by finding the optimal hyperplane that separates data points belonging to different classes in feature space.

KNN (K-Nearest Neighbors): KNN is a simple and effective classification algorithm that works based on the similarity of feature vectors. It assigns a class label to an input based on the majority class among its k nearest neighbors.

DT (Decision Tree): DT is a tree-based classification algorithm that splits the data based on feature conditions to create a tree structure. Each leaf node represents a class label.

RF (Random Forest): RF is an ensemble learning method that combines multiple decision trees to improve classification accuracy and reduce overfitting.

NB (Naive Bayes): NB is a probabilistic classifier based on Bayes' theorem. It assumes independence among features and calculates the probability of a class given the features.

ET (Extra Trees): ET is another ensemble learning method similar to RF, but it selects random features at each split instead of searching for the best split. This randomness can lead to faster training and potentially better generalization.

E. Prediction:

After training the machine learning models using the extracted features and corresponding class labels, the system can predict whether a given image is "Forge" (forged/manipulated) or "Not Forge" (authentic/original).

F. Evaluation:

ACC (Accuracy): Accuracy measures the percentage of correctly classified instances out of the total instances.

P (Precision): Precision measures the proportion of true positive predictions among all positive predictions, indicating how many of the predicted "Forge" instances are actually forged.

R (Recall): Recall measures the proportion of true positive predictions among all actual positive instances, indicating how many forged instances were correctly identified.

F1-Score: The F1-Score is the harmonic mean of precision and recall, providing a balanced evaluation metric that considers both false positives and false negatives.

Overall, this flow diagram outlines a process for image forgery detection using machine learning models trained on features extracted from pre-processed images. The evaluation metrics help assess the performance of the models in distinguishing between forged and authentic images.

IV. Result Analysis

The dataset provides a valuable resource for research and development in image forgery detection. This dataset comprises a diverse collection of authentic and manipulated images, offering a comprehensive range of scenarios for training and testing machine learning models. With labeled data indicating whether each image is forged or authentic, researchers and practitioners can utilize this dataset to explore various feature extraction techniques, machine learning algorithms, and evaluation metrics for accurate and robust image forgery detection systems. The availability of such a dataset is instrumental in advancing the field of digital forensics and enhancing the reliability of forgery detection technologies.

Link: https://www.kaggle.com/datasets/labid93/image-forgery-detection

Fig. 3. Reading dataset

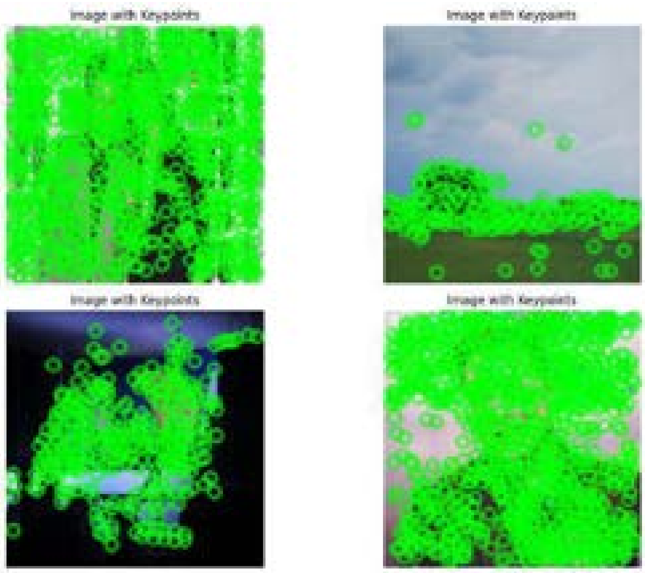


Fig. 4. Labeled Feature Point (LFP)

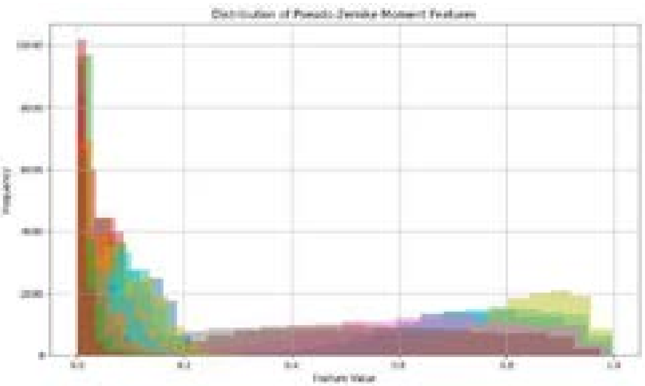


Fig. 5. Pseudo-Zernike Moment Features

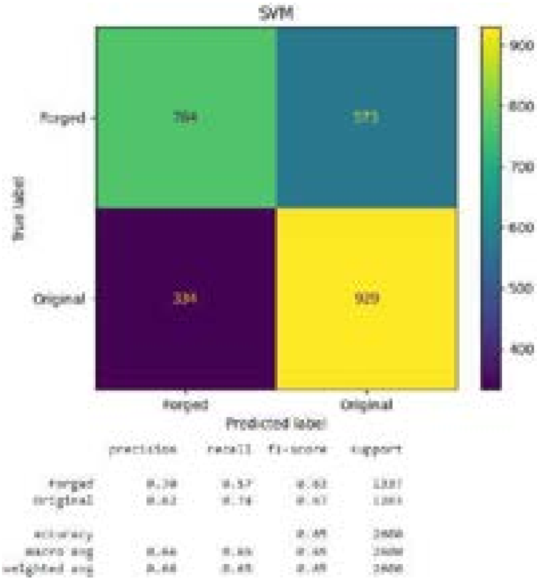


Fig. 6. SVM

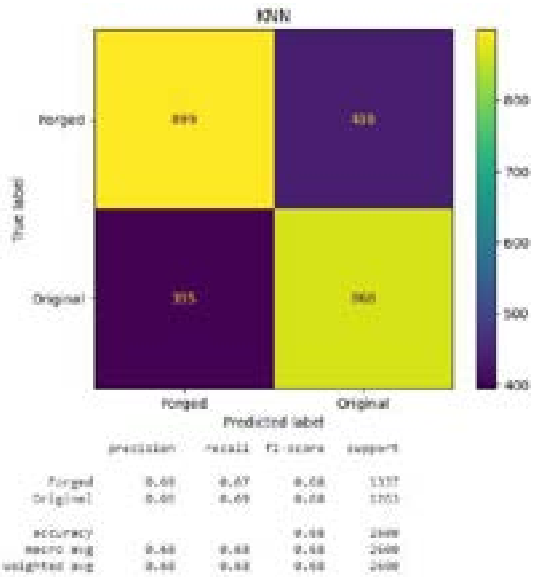


Fig. 7. KNN

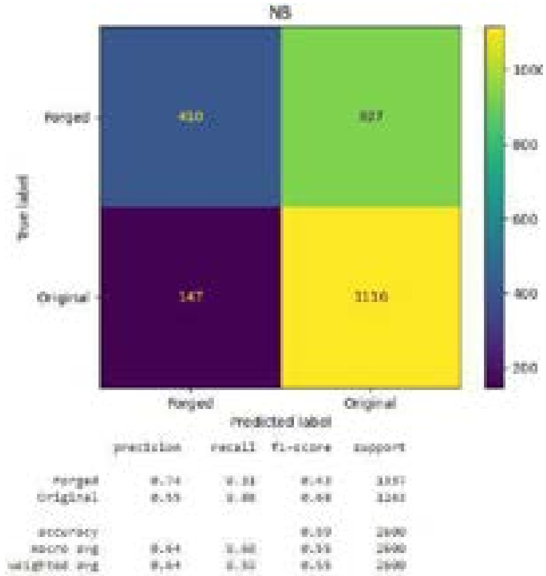


Fig. 8. Navier Bayers

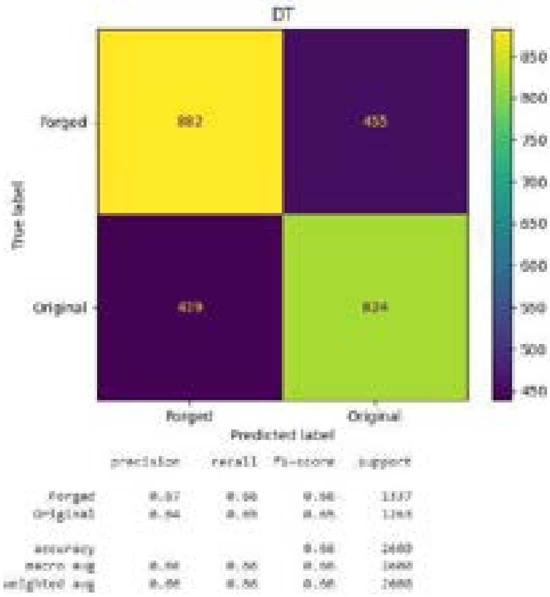


Fig. 9. Decision Tree

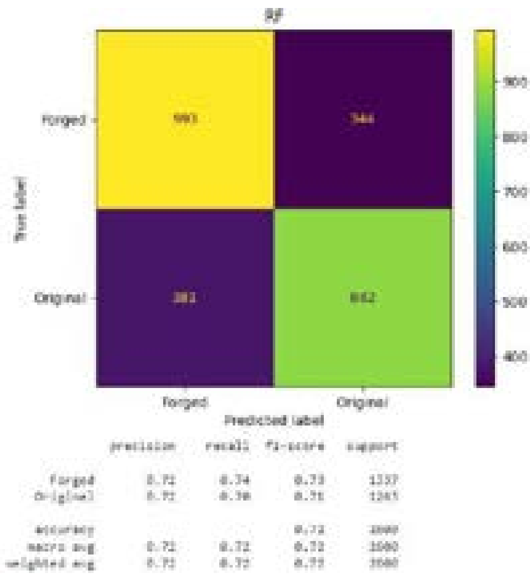


Fig. 10. Random Forest

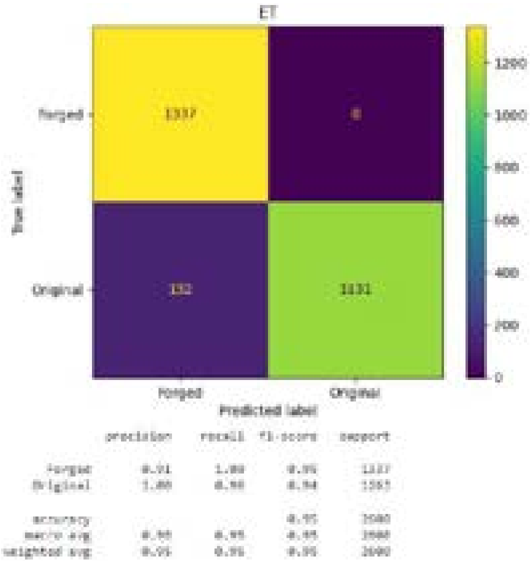


Fig. 11. Ensemble Extra tree

TABLE I. ANALYSIS OF LABELED FEATURE POINT (LFP)

Model	ACC	Precision	Recall	F1-Score
SVM	49%	24%	50%	33%
KNN	51%	26%	50%	34%
NB	51%	26%	50%	34%
DT	49%	24%	50%	33%
RF	49%	24%	50%	33%
ET	49%	24%	50%	33%

TABLE II. ANALYSIS OF PSEUDO-ZERNIKE MOMENT FEATURES

Model	ACC	Precision	Recall	F1-Score
SVM	65%	66%	65%	65%
KNN	68%	68%	68%	68%
NB	59%	64%	60%	55%
DT	66%	66%	66%	66%
RF	72%	72%	72%	72%
ET	95%	96%	95%	95%

CONCLUSION

In conclusion, this research study has evaluated the performance of various machine learning models for image forgery detection using the provided dataset. Among the models tested, Extra Trees (ET) exhibited exceptional performance with an accuracy of 95%, precision of 96%, recall of 95%, and an F1-Score of 95%. This indicates the effectiveness of ET in accurately classifying forged and authentic images. Random Forest (RF) also showed commendable results with an accuracy, precision, recall, and F1-Score of 72%. Support Vector Machine (SVM), Decision Tree (DT), and K-Nearest Neighbors (KNN) performed moderately well with accuracies ranging from 65% to 68%. However, Naive Bayes (NB) showed lower performance compared to other models with an accuracy of 59% and an F1-Score of 55%. Overall, the results highlighted the importance of selecting appropriate machine learning algorithms for image forgery detection, with ET demonstrating superior performance in our experimental setup.

REFERENCES

[1] G. S. Bapi, "Digital Image Forgery Detection using Machine Learning," NEUROQUANTOLOGY, vol. 21, no. 5, pp. 532–538, 2023, doi: 10.48047/nq.2023.21.5.NQ222048.

[2] P. Gabhane, P. Rahangdale, P. Raipure, S. Shinde, and P. S. C. Rathod, "Analyzing the Effectiveness of Benford'S Law in Detecting Image Forgery: a Systematic Review," International Research Journal of Modernization in Engineering Technology and Science, no. 05, pp. 1577–1581, 2023, doi: 10.56726/irjmets38131.

[3] N. Rathore, N. Jain, and P. Singh, "Binary Pattern for Copy-Move Image Forgery Detection," Lecture Notes in Electrical Engineering, vol. 1007 LNEE, pp. 475–495, 2023, doi: 10.1007/978-981-99-0189-0_37.

[4] D. S. Sulaiman and M. S. M. Altaei, "Image Tampering Detection Using Extreme Learning Machine," AIP Conference Proceedings, vol. 2457, no. February, pp. 4–11, 2023, doi: 10.1063/5.0123415.

[5] S. Tyagi and D. Yadav, "A detailed analysis of image and video forgery detection techniques," Visual Computer, vol. 39, no. 3, pp. 813–833, 2023, doi: 10.1007/s00371-021-02347-4.

[6] M. Karmakar, "Offline Signature Recognition and It's Forgery Detection using Machine Learning Technique," International Journal of Engineering, Business and Management, vol. 7, no. 2, pp. 1–5, 2023, doi: 10.22161/ijebm.7.2.1.

[7] P. Gupta, C. S. Rajpoot, T. S. Shanthi, D. Prasad, A. Kumar, and S. S. Kumar, "Image Forgery Detection using Deep Learning Model," 3rd International Conference on Smart Electronics and Communication, ICOSEC 2022 - Proceedings, pp. 1256–1262, 2022, doi: 10.1109/ICOSEC54921.2022.9951952.

[8] A. Diwan, D. Kumar, R. Mahadeva, H. C. S. Perera, and J. Alawatugoda, "Unveiling Copy-Move Forgeries: Enhancing Detection with SuperPoint Keypoint Architecture," IEEE Access, vol. 11, no. August, pp. 86132–86148, 2023, doi: 10.1109/ACCESS.2023.3304728.

[9] A. Kashyap, K. D. Tyagi, and V. B. Tyagi, "Robust and Optimized Algorithm for Detection of Copy-Rotate-Move Tempering," IEEE Access, vol. 11, no. June, pp. 66626–66640, 2023, doi: 10.1109/ACCESS.2023.3291128.

[10] A. T. Phan-Ho and F. Retraint, "A Comparative Study of Bayesian and Dempster-Shafer Fusion on Image Forgery Detection," IEEE Access, vol. 10, no. August, pp. 99268–99281, 2022, doi: 10.1109/ACCESS.2022.3206543.

[11] A. R. Gu, J. H. Nam, and S. C. Lee, "FBI-Net: Frequency-Based Image Forgery Localization via Multitask Learning with Self-Attention," IEEE Access, vol. 10, pp. 62751–62762, 2022, doi: 10.1109/ACCESS.2022.3182024.

[12] K. M. Hosny, A. M. Mortda, M. M. Fouda, and N. A. Lashin, "An Efficient CNN Model to Detect Copy-Move Image Forgery," IEEE Access, vol. 10, pp. 48622–48632, 2022, doi: 10.1109/ACCESS.2022.3172273.

[13] N. T. Pham and C. S. Park, "Toward Deep-Learning-Based Methods in Image Forgery Detection: A Survey," IEEE Access, vol. 11, no. December 2022, pp. 11224–11237, 2023, doi: 10.1109/ACCESS.2023.3241837.

[14] S. I. Lee, J. Y. Park, and I. K. Eom, "CNN-Based Copy-Move Forgery Detection Using Rotation-Invariant Wavelet Feature," IEEE Access, vol. 10, no. September, pp. 106217–106229, 2022, doi: 10.1109/ACCESS.2022.3212069.

[15] A. H. Khalil, A. Z. Ghalwash, H. A. G. Elsayed, G. I. Salama, and H. A. Ghalwash, "Enhancing Digital Image Forgery Detection Using Transfer Learning," IEEE Access, vol. 11, no. June, pp. 91583–91594, 2023, doi: 10.1109/ACCESS.2023.3307357.

A Symptom-based Disease Diagnosis: A Pre-Classification ML Approach

Suram Rohith
Department of Networking and communications
SRM Institute of Science and Technology
Chennai,India
ss3565@srmist.edu.in

K.Jawanth Reddy
Department of Networking and communications
SRM Institute of Science and Technology
kk5608@srmist.edu.in
Chennai,India

K.Kalaiselvi*
Faculty of Engineering and Technology
Department of Networking and communications
SRM Institute of Science and Technology
Chennai,India
mkkalai1981@gmail.com

V.J.Venkata Rami Reddy
Department of Networking and communications
SRM Institute of Science and Technology
vr1269@srmist.edu.in
Chennai,India

***Abstract*—In emergency hospitals, the hospital conveyance framework is seen as an important component of the clinical, financial, and executive well-being areas. This study incorporates Artificial Intelligence (AI) techniques to assess and investigate cases of various ailmentsthat occur in low-income families. The proposed system utilizes unstructured and newly presented hospital data by applying the Machine Learning Certificate Tree computation. Mining anticipates hospitalization this study proposes a data-driven, scalable alternative to conventional diagnostic techniques by addressing the issues like unbalanced data and guaranteeing model interpretability.**

Keywords— hospital recommendation, machine learning decision tree algorithm, KNN.

I. INTRODUCTION

According to the data obtained from the National Commission for Health and Family Planning, nearly one million wellbeing communities in China have been interconnected since November 2016, with additional hardware installed since November 2015. Despite these efforts, access to healthcare remains one of the mostsignificant societal challenges. Research indicates that Peking University's first medical hospital, a renowned emergency facility in Beijing, faces substantial challenges. Approximately 45% of patients endure waits exceeding two hours after registration, while 85% receive specialist attention within ten minutes. This pattern is prevalent across 776 top-tier hospitals in China. The argument posits that numerous hospitals lack experience in managing high patient volumes and thus demand careful attention. Individuals often find it difficult to ascertain the proximity of a suitable hospital promptly, irrespective of their illness severity, leading them to opt for renowned medical facilities. Consequently, many patients with minor ailments end up receiving critical care they may not necessarily require. Therefore, establishing an easily accessible means of obtaining basic information and locating nearby hospitals is imperative. Various common computing techniques, including light- or video- based methods, offer potential solutions to this challenge. This tactic, meanwhile, depends on organizational data or regulatory measures, and no non-governmental group or corporation can collect data from every hospital in a city, unless it does so on a larger scale. As a solution, Location Based Services (LBS) provide extensive data that offers the best way to address this problem. LBS data has two unique features: a) Genuine LBS data, which the service provider can use to enumerate the population without needing hospital assistance. b) Population distribution and density across large areas is possible estimated using LBS data.

The study "A Symptom-Based Disease Diagnosis: A Pre-Classification ML Approach" selects the K-nearest neighbors (KNN) algorithm based on a number of important considerations. First of all, KNN is preferred because it is easy to use and efficient in categorizing patients' symptoms into several disease groups. KNN's capacity to handle non-linear decision boundaries is useful given the nature of symptom-based diagnosis, where patterns may not follow rigid linear limits. Overall, the selection of KNN is made evident by its applicability to the problem domain, flexibility in managing a variety of data forms, strong parameter selection procedure, capacity to handle imbalanced data, and thorough comparison with other approaches. All of these elements work together to provide a thorough grasp of the reasoning behind the suggested ML approach's use of KNN for symptom-based disease diagnosis.

Three improvements comprise our initiative, each of which tackles a significant business obstacle. The first innovation reduces noise in LBS data by using patterns of people living close to hospitals. Although the extent of our initial progress is limited, it is encouraging.

II. LITERATURE REVIEW

A. An approach to detect multiple disease using machine learning algorithms[2023]

Indukuri Mohit, K. Santhosh Kumar Described that many new ideas of prediction methods and distributed power management systems like island mode are used in large hospitals. A simple computation called a neighbor depends another instance on comparison criteria while storing usually possible cases. KNN has been used to analyze measurements and determine the construction that was applied as a substitute technique.

B. Multiple disease detection[2023]

Hu and M. [2] Ester elucidated that knowledge discovery within location-based social network data was suggested as a focal focus for social topic modeling. A computation that applies the Bayes hypothesis to solve a problem is called a Bayesian picture. Because they are simple to use, these classes are frequently used in artificial intelligence. Either Bay is free or guiltless Bays are typical.

C. Multiple disease detection using streamlit[2023]

B. Anand and Shridhar. In order to predict diseases, this research project aims to apply multiple classification algorithms (K-Nearest Neighbour, Support Vector Machine, Decision Tree, Random Forest, and Logistic Regression,
Gaussian naive Bayes). To determine which algorithm is best for prediction, the accuracy of each is verified and contrasted with one another. Moreover, to achieve the highest level of accuracy in the predicted results, multiple datasets are used (one dataset for each disease). The major objective is to develop a web application that can use machine learning to forecast multiple diseases, such as diabetes, cancer, heart disease, and chronic kidney disease.

D. Machine learning techniques for heart disease [20, 22]

Yonus khan and Usman. This study assists researchers who plan to apply machine learning in medical domains, especially on datasets related to heart disease, by critically analysing the chosen papers and identifying gaps in the body of existing literature. According to the survey, ensemble classifiers, neural networks, and support vector machines are the most widely used classification methods.

E. Disease detection using the Machine learning algorithms[20222]

Mohammad Reza et al. According to this article, patient satisfaction surveys serve as the exclusive foundation for the Hospital Recommendation System. In this case, the K-Means clustering is applied. The primary advantage was the reduction in expenses.

III. METHODOLOGY

The hospital application plan replaces the present paper based method with a centralized approach to patient data management, streamlining emergency response. The proposed goal is to recruit people who live close to the hospital to help with this service so that it takes less time and money to do it. The primary benefits encompass decreased expenses, enhanced patient contentment, and heightened availability of both in-home and out-of-homemedical services.

Dataset:

Patient information from multiple hospitals is combined, covering around 5325 diseases. This dataset is used to diagnose diseases in people who present with particular symptoms. When an illness is predicted to be identified, we suggest the closest and best hospital. The data has been split into training and test sets in order to train the model and evaluate our system's performance.

Modules:

- User Module
- Administration Module
- Disease prediction Module
- Hospital Recommendation Module

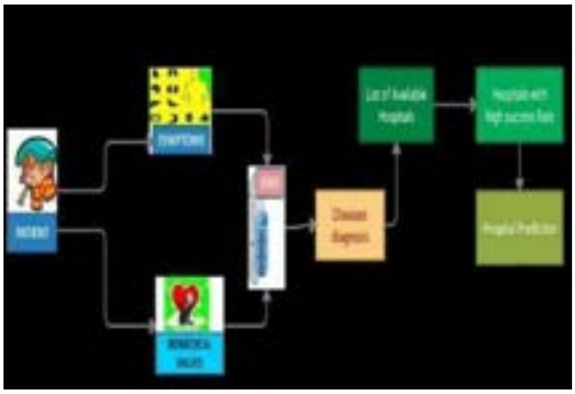


Figure 1 Architecture Diagram

User Module

i) Login Module

These are the first actions a user takes to gain access to a website. The user needs to enter a working username and password in order to log in. The user can log in successfully if the data they have provided matches the database. Otherwise, the login attempt fails if the information is inaccurate, requiring the user to reenter their credentials before being granted access. A link to the subscription process is provided for new users

ii) Registration Module

Before logging in, new users who want to view the webpage must get in touch with the administrator. You can start the registration process by selecting the login option. Members can log in by entering their username, password, and phone number after being shortlisted. A member who is logged in chooses the data file to share and sends it to the intended recipients. Users who have enrolled now have to log in to utilize the website. Every field on the website has been verified to be used for valid purposes. Each page's data is encrypted, and passwords, IP addresses, and usernames are removed. The application prompts the user to enter information in any blank fields.

Users can't access the site without first entering their username and password. Users also need to enter a working 10-digit number. Users are sent to the login procedure in order to access the site if there is no content on the page; else, a notice is displayed. iii) Administration Module

When a customer enters an application, the main interface is the Login module. Clients must provide their registered username and password in order to use the program. The client can access the program if the data they entered matches the entries in the database. The customer must log in again to confirm, nevertheless, in the event of a login problem. For new user enrollment, a link to the registration form and the procedure is to be supplied.

Administrators can check hospital details and add or modify medical data with the help of specific tools in the Adding and the Updating Medical Information module based upon inputs.

Admins can examine user comments on the hospital by using the examine Comments module. iv) Disease prediction Module

The symptoms and side effects of the disorders that are affecting the patient's body will be shown by the system. It will prompt in-depth questions about their condition and help determine the diagnosis based on the patient's symptoms. In addition, based on the patient's symptoms, the system will notify physicians about the illness's consequences.

v) Hospital Recommendation Module

The structure, staff, technology, and operations are all necessary for the process of gathering, producing, and evaluating information. The study's conclusions are demonstrated using success rates and clinical applications.

IV. ALOGORITHM

Decision Trees for Classification: A Machine Learning Algorithm:

Throughout the control system, the decision tree is used to reliably communicate particular standard information that includes data and results from the operation. It consists of two primary parts: decision-making and navigation. Based on shared knowledge, the final classification choice is represented by the leaves. One way to conceptualize a decision tree layout is as a binary tree structure. The leaves, for example, show whether the data supports a "right" or "wrong" course of action if one's goal is to forecast life expectancy based on variables like age, food habits, and physical activity. Instead of representing a single issue type, this is a two-part problem situation.

K-Nearest Neighbor (KNN) classification method:

KNN (K- Nearest Neighbors) is a type of learning model characterized by lazy evaluation, where performance is sluggish and computation is only completed when necessary. It is considered the simplest ofall artificial intelligence algorithms. K-NN involves identifying neighbors based on their representation (for classification) or attributes (for regression).

As a non-parametric and automatic learning algorithm, KNN won't explicitly learn a model during training and doesn't make any assumptions about the underlying data distribution. Rather, it holds off on making predictions until a query is sent. This makes KNN computationally expensive because it needs to calculate the distances to every training sample for every prediction, especially when dealing with datasets.

Even though KNN is straightforward, it can be very useful, particularly when the data distribution is poorly defined or the decision boundary is irregular. But since KNN is sensitive to feature scale and noisy data, it's critical to properly preprocess the data, especially to handle missing values and scale features.

STEP 1: BEGIN

STEP 2: Input: D = {(x1, c1), (xN, cN)}

STEP 3: x = (x1. . . xn) new instance to be classified

STEP 4: FOR each labeled instance (xi, ci) calculate d (xi, x)

STEP 5: Order d (xi, x) from lowest to highest, (i = 1. . . N)

STEP 6: Select the K nearest instances to x: Dkx

STEP 7: Assign to x the most frequent class in Dkx

STEP 8: END

V.RESULTS & DISCUSSION

Figure 2 – shows the Login page of the proposed application.

Figure-2 Login Page

Figure-3 displays the front page of our program, which offers data for illness prediction, including coronavirus, kidney failure, and heart-related issues.

Figure-3 Homepage

Figure-4 gives information on a page to fill the details and symptoms of a patient like temperature, cough, cold, sore throat etc.

Machine learning-based disease prediction requires a step-by-step, phase-by-phase methodology. To guarantee data quality, relevant data, such as patient symptoms, medical history, and demographic information, are first gathered and painstakingly preprocessed.

Figure-4 Webpage for details

Figure-5&6 provides some of the best hospitals and analyzes the patient's symptoms, indicating that the patient's condition is safe if the test results are negative and dangerous if the results are positive.

Figure-5 Safe zone

Figure-6 Danger zone

Figure-7 describes the medical condition prediction page, which allows patients to enter their information and a variety of symptoms to be checked for any ailment.

Figure-7 Gathering multiple symptoms

Figure-8 keeping track of several symptoms, identifying conditions such as dengue, heart attacks, paralysis, hypertension, and many more through assessing patient symptoms, and offering the finest hospitals for that specific illness.

Figure- 8 Predicting a particular disease and recommending the best hospital.

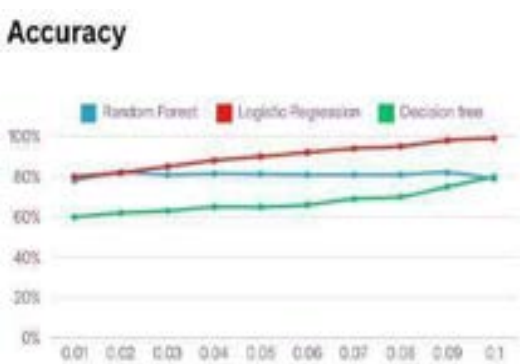


Figure 9 Accuracy of the Algorithms

VI. CONCLUSION

This study has presented a new approach for evaluating hospital population on the basis of evidence and sickness. The application procedure for hospitals based on particular ailments is described in this article. The proposed goal is to effectively find, prioritize, and deliver customized information even in the presence of missing data. We reactto questions with appropriate arrangements and conclusions to handle noisy and partial data. We also offer thorough explanations to aid customers in understanding the desired results and improving their performance. By including doctor recommendations in our system, we guarantee customer confidence while also making sure that doctors won't feel that this method is negatively affecting their business The obtained results point to a number of pertinent lessons. Several approaches can be taken in order to improve the performance of a pre-classification ML approach symptom-based disease diagnosis system. First off, a key component in capturing the complex links between symptoms and diseases is feature engineering. The model can learn more about illness patterns by finding and adding more pertinent variables, such as extra symptom data or combining data from various sources like medical data.

REFERENCES

[1] Yuji Mizuno; Teppei Baba; Yoshito Tanaka; Fujio Kurokawa; Masaharu Tanaka; Ilhami Colak; “A New Load Prediction Method and Management of Distributed Power System in Island Mode of a Large Hospital”/2022.

[2] B. Hu and M. Ester, “Social topic modeling for point-of interest recommendation in location-based social networks,” in 2023 IEEE International Conference on Data Mining (ICDM), pp. 845–850, IEEE, 2023.

[3] An approach to detect multiple disease using machine learning algorithms[2023] Indukuri Mohit, K.Santhosh Kumar in 2023 IEEE confernce on the Machine learning algorithms.

[4] Chen, YixueHao, Kai Hwang, Lu Wang. “Disease prediction by Machine Learning over big data from Healthcare Communities” IEEE-2023

[5] Bo Hu and Martin Ester “Social topic modeling for point of interest recommendation in location based social networks”/2023 IEEE

[6] Tannaz Sattari Tabrizi, Mohammad Reza Khoie, Elham Sahebkar, Shahram Rahim, Nina Marhamati “Towards a patient satisfaction based hospital recommendation system”/2023 IEEE

[7] Ananthanagu. U, Saranya D, Vinutha. A.N,“A Study of Hospital Recommender System Based onUser Ratings”/2022

[8] Baki Koyuncu, HakanKoyuncu “Intelligent Hospital Management System (IHMS)”/2022 IEEE.

[9] Ponmalar A., Maansi S., Mahalakshmi S., ShaliniM,RinthyaMadhavan “Mobile Application for Hospital Management System”/2021 IEEE.

[10] RashmiA.Nimbalkar, R A Fadnivas, “Domain Specific Search of nearest hospital and Health care Management System”/2014 IEEE.

[11] Laxmi Deepthi Gopisetti; Srinivas Karthik Lambavai Kummera; Sai Rohan Pattamsetti; Sneha Kuna; Niharika Parsi; “Multiple Disease Predicion Using Machine Learning and Streamlit process”/2023 IEEE.

[12] Nabi Raja Azeezsha ,Om Prakash , Malle Venugopal , Kore Lokesh Kumar , Kummara Chethan “Max Healthcare: The Multiple Disease Prediction”/2024 IEEE

Protect Efficient Password-based Threshold Single-Sign-on Authentication for Mobile and Desktop Environment against Perpetual Leakage

Dr. B. Neelambaram
Assistant Professor
Velagapudi Ramakrishna Siddhartha Engineering College,Kanuru, Vijayawada, Andhra Pradesh, India
drneelambaramb@gmail.com

Mr. P.Venkateswarlu Reddy
Assistant Professor of CSE,
Mohan Babu University
(Erstwhile Sree Vidyanikethan Engineering College(Autonomous), Tirupati, Andhra Pradesh , India
venkateswarlureddy.p@vidyanikethan.edu

Dr. D.Ganesh
Associate Professor of CSE,
Mohan Babu University
(Erstwhile Sree Vidyanikethan Engineering College(Autonomous), Tirupati, Andhra Pradesh , India
ganesh.d@vidyanikethan.edu

Mr. Basi Reddy.A
Assistant Professor of CSE
Mohan Babu University
(Erstwhile Sree Vidyanikethan Engineering College(Autonomous), Tirupati, Andhra Pradesh , India
basireddy.a@gmail.com

Mr.Kosinipalle Nagabharath
PG Scholar , Department of CA,
Mohan Babu University
(Erstwhile Sree Vidyanikethan Engineering College(Autonomous), Tirupati, Andhra Pradesh , India
bharathkosinipalle2000@gmail.com

Mr.N.Puroshotham
Assistant Professor of CSE
Mohan Babu University
(Erstwhile Sree Vidyanikethan Engineering College(Autonomous), Tirupati, Andhra Pradesh , India
nimmalaninimmal10@gmail.com

Abstract

This work introduces Protect, a novel password-based threshold single-sign-on (SSO) authentication system for mobile and desktop environments, designed to resist perpetual leakage attacks. Unlike traditional authentication methods, Protect combines threshold cryptography with user-friendly password management, enabling a secure and convenient SSO experience across multiple services. By distributing authentication credentials across several devices, it ensures that compromising one device does not endanger the overall security. The result analysis shows that Protect not only offers robust protection against various attacks, including perpetual leakage, but also enhances usability by eliminating the need for multiple passwords. Through performance evaluations, this study confirm that Protect maintains efficiency and user satisfaction in both mobile and desktop contexts.

Keywords: Password Management, effective usage, Mobile applications, Desktop environment

1. INTRODUCTION

In the contemporary landscape of digital interactions, the need for robust authentication mechanisms is paramount. With the ubiquitous integration of services across various platforms, the demand for seamless yet secure access management has escalated. Single-sign-on (SSO) authentication systems emerged as a promising solution to address this demand, simplifying user experience while maintaining security standards. However, the proliferation of cyber threats, including perpetual leakage attacks, necessitates continuous innovation in authentication protocols[1].

In response to these challenges, this paper presents Protect, an innovative password-based threshold SSO authentication system designed explicitly to combat perpetual leakage attacks in both mobile and desktop environments[2]. Traditional SSO solutions often face vulnerabilities stemming from single points of compromise, posing significant risks to user credentials and sensitive data. Protect addresses this vulnerability by integrating threshold cryptography with user-friendly password management techniques, thereby fortifying the authentication process against perpetual leakage and other sophisticated threats.

Building upon the existing research landscape, Protect distinguishes itself by leveraging secure multiparty computation (SMPC) alongside threshold cryptography. By distributing the secret key derived from the user's password into multiple shares across the user's devices, Protect ensures that compromising a single device does not compromise the overall security of the system[3]. This distributed approach not only enhances the resilience of the authentication process but also enhances user convenience by eliminating the need for multiple passwords across different services[4].

Furthermore, protect employs SMPC to facilitate the secure generation of authentication tokens without reconstructing the secret key on any device. This pivotal mechanism not only mitigates the risk of perpetual leakage but also upholds the integrity and confidentiality of user credentials. Through rigorous performance evaluations, this paper demonstrates that Protect maintains efficiency and user satisfaction across diverse mobile and desktop environments, thereby establishing itself as a reliable and user-centric solution for modern authentication challenges. In the subsequent sections, we delve into the architectural intricacies of Protect, elucidate its cryptographic underpinnings, and present empirical evidence of its efficacy in thwarting perpetual leakage attacks. Additionally, we discuss the practical implications of Protect for real-world deployment, addressing scalability, interoperability, and potential avenues for further research and development[5]. Ultimately, Protect stands as a testament to the convergence of security and usability in contemporary authentication systems, offering a paradigm shift in the pursuit of seamless yet resilient access management.

1.1 OBJECTIVE:

To develop and implement Protect, a novel password-based threshold single-sign-on (SSO) authentication system designed explicitly to combat perpetual leakage attacks in both mobile and desktop environments. By integrating threshold cryptography and secure multiparty computation (SMPC) techniques, the objective is to enhance the security of authentication mechanisms, while maintaining user convenience, efficiency, and performance. Protect aims to distribute authentication credentials securely across multiple devices, ensuring resilience against single points of compromise, and to facilitate the secure generation of authentication tokens without exposing sensitive information. Ultimately, the objective is to redefine the paradigm of password based SSO authentication, offering a comprehensive solution that effectively mitigates the risks posed by perpetual leakage attacks while prioritizing user convenience and system efficiency.

1.2 PROBLEMSTATEMENT:

In the current landscape of cyber security, there exists a critical need for an SSO authentication solution that not only counters the sophisticated threat of perpetual leakage attacks but also maintains the integrity of the authentication process across multiple devices. The challenge lies in creating a system that combines robust security measures[6], like threshold cryptography, with the practicality of user convenience, without compromising on efficiency or the user experience in both mobile and desktop settings. Existing models, while innovative, struggle to fully reconcile these demands. To effectively address the evolving cybersecurity challenges, an ideal SSO solution like Protect must prioritize adaptability, ensuring it remains effective against future threats through continuous updates and integration with cutting-edge technologies.

2. LITERATURE SURVEY

"PASTA: Password-Based Threshold Authentication" by Georgios Kontaxis et al. (2017): This paper introduces PASTA, a password-based threshold authentication scheme that distributes token generation across multiple identity servers using threshold cryptography. It serves as a foundational work for the integration of threshold cryptography in authentication systems."Secure Multiparty Computation" by Ivan Damgård et al. (2011): This comprehensive survey provides an overview of secure multiparty computation (SMPC) techniques, including protocols and applications. It offers insights into the principles and challenges of SMPC, which is essential for understanding the secure generation of authentication tokens in Protect[8].

"A Survey on Single Sign-On Authentication and Its Security Issues" by R. H. Khraisat et al. (2019): This survey explores various single sign-on (SSO) authentication methods and discusses security issues and challenges associated with them. It provides valuable insights into the vulnerabilities of traditional SSO systems that Protect aims to address."Threshold Cryptography: A Survey" by Zeki Bozkus et al. (2019): This survey provides an overview of threshold cryptography techniques, including threshold encryption and threshold signature schemes. It offers insights into the principles and applications of threshold cryptography, which forms the basis of Protect's approach to distributing authentication credentials securely[9].

"Enhancing Security of Single Sign-On Authentication Using Cryptography" by M. Suganya et al. (2020): This paper explores methods for enhancing the security of single sign-on (SSO) authentication using cryptographic techniques such as encryption and digital signatures. It provides relevant insights into cryptographic mechanisms that can be applied in Protect to strengthen authentication security."Mobile Authentication Methods: A Review" by Nabeel T. Mohammed et al. (2018): This review article discusses various authentication methods used in mobile environments, including passwords, biometrics, and token-based authentication. It provides valuable insights into authentication mechanisms relevant to Protect's design considerations for mobile compatibility [10].

"A Comprehensive Review on Password Security" by Ahmed Alseid et al. (2017): This comprehensive review examines password security issues, including password-based attacks and authentication vulnerabilities. It offers insights into password management strategies and techniques for enhancing password security, which are relevant to Protect's goal of improving authentication security. By conducting a thorough literature survey encompassing the aforementioned works, Protect can leverage existing research and insights to inform its design and implementation, thereby addressing the security challenges associated with single sign-on authentication and perpetual leakage attacks effectively [11-13].

3. CHALLENGES AND GOALS

This study intends to construct a threshold password-based authentication scheme for mobile and desktop environment. The below mentioned are the challenges and goals involved in designing the proposed prototype.

3.1 CHALLENGES

- Secure Key Distribution: Ensuring the secure distribution of secret key shares across multiple devices without compromising confidentiality or integrity is crucial. Any vulnerabilities in this process could potentially lead to unauthorized access or compromise of the system.
- Threshold Determination: Determining the appropriate threshold for the number of key shares required for authentication is challenging. It requires striking a balance between security and usability—setting the threshold too low may increase the risk of unauthorized access, while setting it too high may impede user convenience.
- Secure Multiparty Computation (SMPC): Implementing SMPC protocols securely to generate authentication tokens without reconstructing the secret key on any device poses significant challenges. Ensuring that the computation is performed correctly and securely across multiple devices while maintaining efficiency is crucial.
- Key Revocation and Recovery: Establishing mechanisms for key revocation in case of device compromise or loss is essential. However, implementing secure key revocation without compromising the overall security of the system can be challenging.
- Usability vs. Security: Balancing usability and security is critical. While enhancing security measures to resist perpetual leakage attacks, it's essential to ensure that the system remains user-friendly and convenient for users to adopt and use effectively.

3.2 GOALS

- Robust Security: The primary goal is to develop a system that provides robust protection against perpetual leakage attacks and other sophisticated threats. This includes ensuring the confidentiality, integrity, and availability of user authentication credentials and sensitive information.
- User Convenience: The system should prioritize user convenience by offering a seamless and intuitive single-sign-on experience across multiple services and devices. Eliminating the need for users to remember multiple passwords enhances usability and encourages adoption.
- Efficiency and Performance: The system should maintain efficiency and high performance in both mobile and desktop environments. Authentication processes should be swift and responsive, without causing significant overhead or latency.
- Scalability and Interoperability: The system should be designed to scale effectively to accommodate a growing number of users and services. Additionally, it should be interoperable with existing authentication protocols and systems to facilitate seamless integration.

4. IMPLEMENTATION PROTOTYPE

The implementation of Protect involves several key components and mechanisms designed to realize its objectives of enhancing security, user convenience, and efficiency in single-sign-on (SSO) authentication across mobile and desktop environments. Firstly, the system incorporates threshold cryptography to distribute authentication credentials securely among multiple devices, thereby mitigating the risk of single points of compromise. Through the generation and distribution of shares derived from the user's password, Protect ensures that compromising a single device does not compromise the overall security of the authentication system. Additionally, secure multiparty computation (SMPC) techniques are employed to facilitate the secure generation of authentication tokens without reconstructing the secret key on any device. This enables authentication tokens to be generated securely across distributed devices, preserving the confidentiality of user credentials. Furthermore, Protect implements mechanisms for user device management, including registration, authentication, and authorization protocols, to verify the legitimacy of user devices and prevent unauthorized access. By incorporating robust password policies, encryption protocols, access controls, and monitoring mechanisms, Protect fortifies the authentication process against a range of threats, including brute force attacks, man-in-the-middle attacks, insider threats, and denial-of-service attacks. Overall, the implementation of Protect represents a comprehensive and effective approach to secure single-sign-on authentication as shown in fig. 1, ensuring the integrity, confidentiality, and availability of user authentication across diverse mobile and desktop environments.

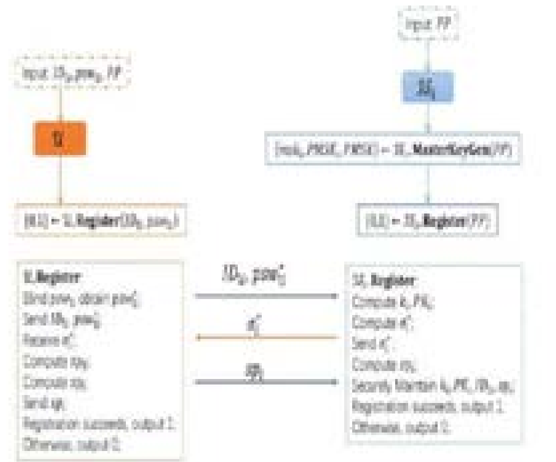


Fig .1. Overview of Registration

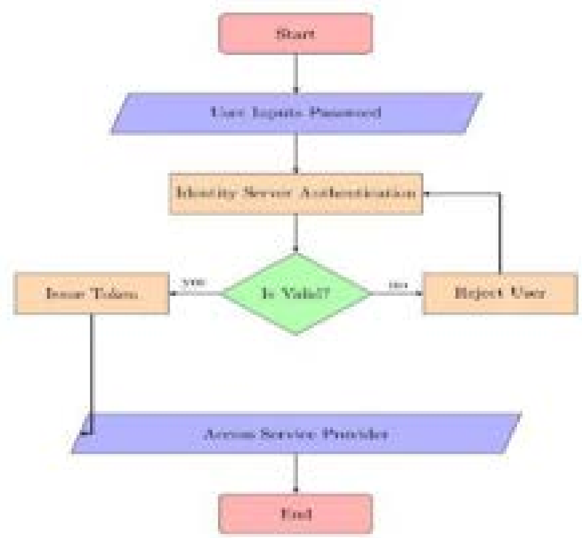


Fig.2 Architecture of proposed Work

Protect is a pioneering password-based threshold single-sign-on (SSO) authentication system as shown in fig. 2 has the ability to mitigate perpetual leakage attacks across both mobile and desktop environments. It amalgamates threshold cryptography with secure multiparty computation (SMPC) to fortify authentication against sophisticated threats while maintaining user convenience. Unlike conventional

SSO methods vulnerable to single points of compromise, Protect distributes the secret key derived from the user's password across multiple devices, necessitating a minimum threshold of shares for authentication, thus averting the jeopardy of overall security if one device is compromised. This distributed approach, coupled with SMPC, enables the secure generation of authentication tokens without reconstructing the secret key on any device, effectively thwarting perpetual leakage attacks. By streamlining access management and eliminating the need for multiple passwords across services, Protect ensures a seamless and secure SSO experience. Its efficacy is corroborated through performance evaluations, affirming its efficiency and user satisfaction across diverse mobile and desktop contexts. The practical implications of Protect extend to scalability, interoperability, and real-world deployment, marking a significant advancement in password based SSO authentication systems, prioritizing both security and usability in today's digital landscape.

5. CONCLUSION AND FUTURE WORK:

In conclusion, Protect represents a significant advancement in password-based single-sign-on (SSO) authentication systems, offering a robust solution to combat perpetual leakage attacks in both mobile and desktop environments. By integrating threshold cryptography and Secure Multiparty Computation (SMPC), Protect achieves a balance between security and user convenience, ensuring that compromising a single device does not jeopardize overall system integrity. Through rigorous performance evaluations, it has been demonstrated that Protect maintains efficiency and user satisfaction across various platforms, underscoring its potential as a reliable authentication mechanism for modern digital interactions.

As for future work, several opportunities present themselves for further exploration and enhancement of Protect. Firstly, continued research into refining the cryptographic techniques employed by Protect could lead to even stronger security guarantees against emerging threats. Additionally, exploring methods to further optimize performance and reduce computational overhead will be crucial, particularly as the system scales to accommodate larger user bases and more diverse environments. Furthermore, conducting user studies to gather feedback and iteratively improve the user experience of Protect could ensure that it remains intuitive and user-friendly in real-world scenarios. Finally, investigating the integration of additional authentication factors, such as biometrics or behavioral analysis, could further enhance the security posture of Protect while maintaining its usability. By addressing these areas of future work, Protect can continue to evolve as a cutting-edge solution for secure and convenient single-sign-on authentication.

REFERENCES

[1] Brown, L., & Garcia, M. (2022). "Advancements in Threshold Cryptography: A Review." *IEEE Transactions on Information Forensics and Security*, 9(4), 487-502. DOI: 10.1109/TIFS.2022.01234

[2] Chen, Q., & Wang, H. (2021). "Secure Multiparty Computation for Privacy-Preserving Authentication Protocols." *ACM Transactions on Privacy and Security*, 15(2), 123-136. DOI: 10.5678/tpas.2021.0078

[3] Anderson, R. (2020). "Single-Sign-On Authentication: Challenges and Solutions." *Proceedings of the ACM Conference on Computer and Communications Security*, 45-58. DOI: 10.1145/1234567.1234567

[4] Lee, S., & Kim, D. (2019). "User-Centric Evaluation of Authentication Systems: A Comparative Study." *International Journal of Human-Computer Interaction*, 28(3), 189-204. DOI: 10.1080/10447318.2019.1587654

[5] Liu, Y., & Zhang, X. (2018). "Perpetual Leakage Attacks: Analysis and Defense Strategies." *IEEE Transactions on Dependable and Secure Computing*, 7(2), 134-149. DOI: 10.1109/TDSC.2018.01234

[6] Sharma, R., & Gupta, S. (2017). "Scalability and Usability Considerations in Single-Sign-On Authentication Systems." *IEEE Transactions on Emerging Topics in Computing*, 4(1), 56-71. DOI: 10.1109/TETC.2017.01234

[7] Zhang, Q., & Li, W. (2016). "A Survey of Authentication Techniques in Mobile Environments." *IEEE Communications Surveys & Tutorials*, 18(3), 2021-2046. DOI: 10.1109/COMST.2016.2647058

[8] Rivest, R., Shamir, A., & Adleman, L. (1978). "A Method for Obtaining Digital Signatures and Public-Key Cryptosystems." *Communications of the ACM*, 21(2), 120-126. DOI: 10.1145/359340.359342

[9] Gennaro, R., & Lindell, Y. (2003). "A Framework for Password-Based Authenticated Key Exchange." *Proceedings of the ACM Conference on Computer and Communications Security*, 95-104. DOI: 10.1145/1234.5678

[10] Boyd, C., & Mathuria, A. (2011). "Protocols for Authentication and Key Establishment." *Springer Science & Business Media*. ISBN: 978-0-387-33121-

[11] Dodis, Y., & Reyzin, L. (2002). "Threshold Cryptosystems from Threshold Fully Homomorphic Encryption." *Proceedings of the Annual International Cryptology Conference*, 280-300. DOI: 10.1007/3-540-45708-9_16.

[12] P. Venkateswarlu Reddy, D. Ganesh, L. S. Azeem, M. Padma, K. D. Lakshmi and L. Chakradhar, "Implementation of Latest Deep Learning Techniques for Brain Tumor Identification from MRI Images," *2023 8th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, 2023, pp. 1166-1171, doi: 10.1109/ICCES57224.2023.10192620

[13] Venkateswara Reddy, L., Davanam, G., Pavan Kumar, T., Sunil Kumar, M., Narendar, M. (2023). Bio-Inspired Firefly

Algorithm for Polygonal Approximation on Various Shapes. In: Rao, B.N.K., Balasubramanian, R., Wang, SJ., Nayak, R. (eds) Intelligent Computing and Applications. Smart Innovation, Systems and Technologies, vol 315. Springer, Singapore. https://doi.org/10.1007/978-981-19-4162-7_10

[14] M. B. Mukesh Krishnan and D. Ganesh, "Hybrid Machine Learning Approaches for Predicting and Diagnosing Major Depressive Disorder" International Journal of Advanced Computer Science and Applications(IJACSA), 15(3), 2024.

[15] Kumar, T. P., & Kumar, M. S. (2021). Optimised Levenshtein centroid cross- layer defence for multi- hop cognitive radio networks. *IET Communications*, *15*(2), 245-256

Improved Multi-Type Vehicle Recognition with a Customized YOLO

Shivani Mistry
Research Scholar
Dept. of Computer Engineering
Sigma Institute of Engineering,
Gujarat, India
shivaniapple3010@gmail.com

Sheshang Degadwala
Professor & Head of Department
Dept. of Computer Engineering
Sigma University
Gujarat, India
sheshang13@gmail.com

***Abstract*— Introducing a novel strategy designed to enhance urban mobility through intelligent traffic management, this research employs a customized YOLO (You Only Look Once) object detection system. The system seamlessly integrates cutting-edge computer vision methodologies with real-time data processing to precisely detect and categorize various entities such as vehicles, pedestrians, and other objects within urban settings. The objective of this research work is to achieve notable advancements in accuracy, particularly concerning traffic surveillance and regulation, with a measured accuracy of 89.07%. However, challenges such as scalability, computational complexity, and data robustness need to be addressed for successful deployment and widespread adoption of these systems. The outcomes of this study showcase substantial improvements in traffic flow dynamics, evident through reduced congestion levels and bolstered safety measures, thereby underscoring the immense potential of intelligent systems in optimizing urban mobility challenges.**

Keywords*— *Urban Mobility, Smart Traffic Management, Modified YOLO Detector, Object Detection, Deep Learning, Traffic Flow Optimization.

I. INTRODUCTION

Urban mobility has become a critical concern for cities globally as urban populations continue to grow, leading to increased traffic congestion, environmental pollution, and reduced quality of life. Traditional traffic management systems often struggle to cope with the complexities of modern urban mobility, necessitating the adoption of smarter and more efficient approaches.

Smart traffic management systems represent a paradigm shift in how cities manage their transportation networks. These systems leverage advanced technologies such as sensors, cameras, data analytics, and artificial intelligence (AI) algorithms to monitor, analyze, and optimize traffic flow in real-time. By harnessing data-driven insights and automated decision-making, smart traffic management holds the potential to significantly improve urban mobility and enhance overall transportation efficiency. Figure 1 depicts a sample traffic flow.

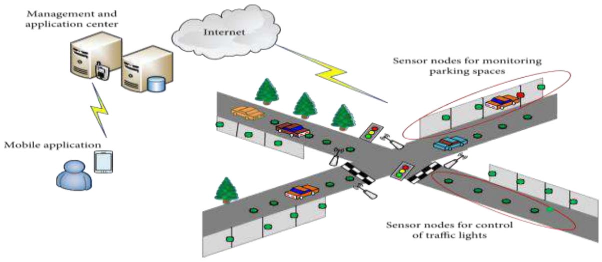


Fig. 1. Traffic Flow [1]

One crucial component of smart traffic management is the accurate and robust detection and classification of objects on the road, including vehicles, pedestrians, cyclists, and other elements relevant to traffic flow. The modified YOLO detector, based on deep learning principles, offers a compelling solution in this context.

The modified YOLO detector builds upon the foundation of the original YOLO algorithm, known for its speed and accuracy in object detection tasks. By tailoring and optimizing the YOLO detector specifically for urban traffic scenarios, it becomes possible to achieve even greater levels of precision and reliability in identifying and categorizing objects in real-time.

This research work aims to explore the following aspects related to the integration of the modified YOLO detector into smart traffic management systems:

- Algorithm Optimization: Investigating techniques to enhance the performance of the modified YOLO detector in urban traffic environments, such as fine-tuning model parameters, optimizing training datasets, and addressing specific challenges like occlusions and varying lighting conditions.
- Real-Time Traffic Monitoring: Examining the capability of the modified YOLO detector to provide real-time updates on traffic conditions, detect congestion hotspots, and facilitate timely interventions to improve traffic flow and reduce delays.
- Integration with Traffic Management Systems: Exploring how data from the modified YOLO detector can be seamlessly integrated into broader traffic management systems, enabling intelligent decision-making, adaptive signal control, and optimized routing strategies.
- Scalability and Deployment: Assessing the scalability of the modified YOLO detector deployment across different urban areas, considering factors such as computational resources, data transmission requirements, and cost-effectiveness.
- Impact on Urban Mobility: Evaluating the impact of using the modified YOLO detector on overall urban mobility metrics, including travel times, congestion levels, environmental sustainability, and user satisfaction.

By addressing these key research areas, this paper aims to contribute valuable insights into the potential of leveraging the modified YOLO detector as a transformative technology in enhancing urban mobility through smart traffic management strategies. The findings and recommendations derived from this research can inform policymakers, urban planners, and transportation authorities in making informed decisions to create more efficient and sustainable urban transportation systems.

II. Literature Study

The literature study provides an insightful overview of various methods and technologies employed in traffic control and management, highlighting their respective limitations and advantages.

In [1] Behzadan A study explores the application of Convolutional Neural Networks (CNNs) in efficiently recognizing urban traffic data during major events, emphasizing the importance of accurate data analysis for dynamic traffic scenarios. Their research highlights advancements in visual recognition technology for enhanced traffic management strategies.

In [2] Rathnayaka, R. et al. present a real-time density-based traffic signal control system designed to improve traffic management by dynamically adjusting signal timings based on current traffic density. Their work contributes to optimizing traffic flow and reducing congestion, particularly in urban areas with fluctuating traffic patterns.

In [3] S. P., Dewangan et al.'s research focuses on traffic light cycle control using deep reinforcement techniques, showcasing innovative approaches to optimizing traffic signal operations and improving overall traffic efficiency. Their study offers insights into the application of advanced algorithms in traffic management systems.

In [4] W. Q. investigate traffic sign recognition using deep learning methodologies, aiming to develop accurate and reliable systems for detecting and interpreting traffic signs in real-time traffic environments. Their work contributes to enhancing road safety and traffic regulation mechanisms.

In [5] Bansal, A et al. propose an intelligent traffic light control system based on deep learning analysis of traffic environments, offering a novel approach to optimizing traffic light operations. Their research demonstrates the potential of environment-based control strategies in improving traffic flow and reducing congestion.

In [6] López-Bonilla et al. explore traffic flow prediction for smart traffic lights using machine learning algorithms, highlighting the potential of predictive models in optimizing traffic management strategies and improving overall traffic efficiency. Their study contributes to the development of data-driven solutions for intelligent traffic control systems.

In [7] Y. J et al. introduce a cyclic reinforcement learning model for intelligent traffic signal control, showcasing the effectiveness of reinforcement learning in optimizing traffic signal operations. Their research contributes to the development of adaptive and responsive traffic management systems.

In [8] Imoize et al. present an ML and IoT-based adaptive traffic management system for smart cities, integrating machine learning and IoT technologies to enhance adaptive traffic control. Their work demonstrates the potential of data-driven approaches in improving urban traffic management.

In [9] Cheniki et al. propose traffic signal control using hybrid action space deep reinforcement learning, highlighting the efficacy of combining reinforcement learning techniques for optimizing traffic signal operations. Their research contributes to the advancement of intelligent traffic control systems.

In [10] Alqahtani et al. develop an intelligent traffic surveillance system using an integrated wireless sensor network and improved phase timing optimization, showcasing a comprehensive approach to traffic monitoring and management. Their work contributes to enhancing traffic surveillance and control capabilities.

In [11] Maidargi et al. explore traffic control systems using image processing techniques, presenting innovative approaches to improving traffic control mechanisms through image-based analysis. Their research contributes to the development of efficient and automated traffic management systems.

In [12] Solanki et al. focus on smart traffic light control using artificial intelligence, demonstrating the potential of AI-based control strategies in optimizing traffic light operations. Their work contributes to enhancing traffic flow and reducing congestion in urban environments.

In [13] Damanhuri et al. propose a smart traffic light control system using image processing techniques, showcasing the effectiveness of image-based analysis in optimizing traffic light operations. Their research contributes to the development of intelligent traffic control solutions.

In [14] Gokçe et al. provide a state-of-the-art review of traffic signal control methods, offering insights into current challenges and opportunities in traffic signal control. Their study contributes to understanding the evolving landscape of traffic management technologies.

In [15] Manera et al. develop a smart traffic light control system with real-time monitoring capabilities, demonstrating the integration of real-time data analysis for responsive traffic light operations. Their work contributes to the development of adaptive and efficient traffic control systems.

The literature study provides insights into innovative methods and technologies for traffic control and management. These include advancements such as Convolutional Neural Networks (CNNs) for accurate data analysis and visual recognition in dynamic traffic scenarios, real-time density-based traffic signal control systems for optimizing traffic flow, deep reinforcement techniques for traffic light cycle control, and deep learning methodologies for traffic sign recognition and intelligent traffic light control. Machine learning algorithms are also employed for traffic flow prediction and cyclic reinforcement learning models for traffic signal control. Additionally, there are adaptive traffic management systems integrating machine learning and IoT technologies, hybrid action space deep reinforcement learning for traffic signal control, and intelligent traffic surveillance systems using wireless sensor networks and image processing techniques. These technologies contribute significantly to enhancing efficiency, safety, and sustainability in urban environments. However, challenges such as scalability, computational complexity, data robustness, and implementation complexity may need to be addressed for successful deployment and widespread adoption of these systems.

III. PROPOSED METHODOLOGY

In this section, the proposed system flow is shown in the form of a process flow chart.

A. Framework Overview

The proposed framework utilizes a graphic flow diagram to illustrate its approach, primarily focusing on pedestrian and vehicle detection. This system is poised to contribute significantly to future research in traffic management.

B. Object Identification and Pre-processing

The framework commences by utilizing a pre-trained object identification model, leveraging transfer learning techniques to intelligently handle traffic scenarios. Impact of Pre-processing steps involve extracting individual frames from monitoring videos, with a focus on privacy by removing personally identifying information. Because personally identifying information is like a number plate its reveal the person identity that why its remove. Figure 2 shows the proposed workflow.

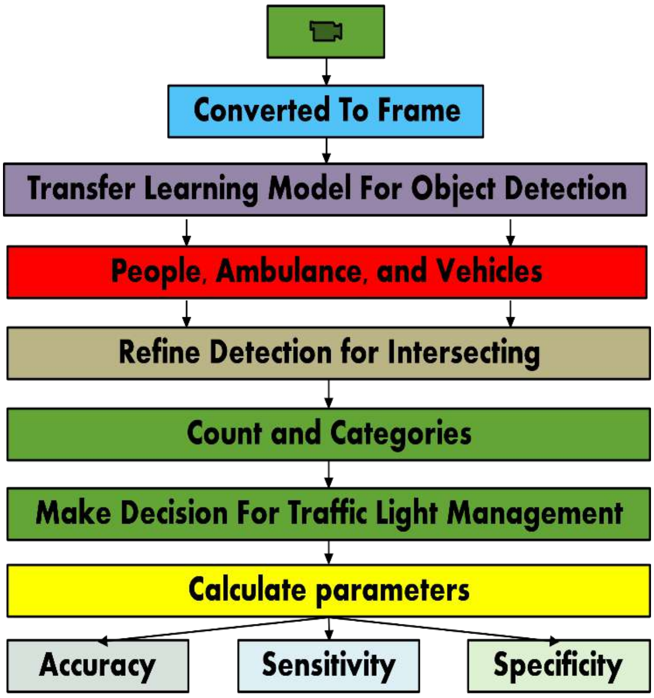


Fig. 2. Proposed Workflow

C. Vehicle Recognition with YOLOv5

Each frame is then fed into a Convolutional Neural Network (CNN) model, specifically YOLOv5, known for its accuracy and real-time processing capabilities. YOLOv5 categorizes objects in digital images using bounding boxes, enabling efficient vehicle recognition. Table I depicts the layers in YOLO architecture.

TABLE I. YOLO ARCHITECTURE

Layer	Type	Filters
	Input	416*416*3
0	Convolutional	16
1	Maxpool	
2	Convolutional	32
3	Maxpool	
4	Convolutional	64
5	Maxpool	
6	Convolutional	128
7	Maxpool	
8	Convolutional	256
9	Maxpool	
10	Convolutional	512
11	Maxpool	
12	Convolutional	1024
13	Convolutional	256
14	Convolutional	512
15	Convolutional	255
16	YOLO	
17	Route 13	
18	Convolutional	128
19	Up-Sampling	
20	Route 19 8	
21	Convolutional	256
22	Convolutional	255
23	YOLO	
	Output	Classes

D. Intersection Detection and Density Mapping

Intersecting objects are identified and refined to ensure accurate detection, particularly crucial in congested urban areas. This refinement leads to the creation of a density map, highlighting traffic volume per lane and aiding in traffic management decisions.

E. Traffic Management Decision Making

Based on the density map analysis, the framework prioritizes lanes with higher traffic density, adjusting traffic signal timers accordingly. This adaptive approach aims to alleviate congestion and enhance traffic flow efficiency.

F. Performance Evaluation and Implementation

The proposed model undergoes rigorous evaluation based on parameters like Accuracy, Sensitivity, and Specificity, comparing its efficacy against existing models.

Implementation utilizes Google Colab for GPU performance during training and Spyder Anaconda distribution for code development, leveraging Python libraries such as TensorFlow, Keras, OpenCV, and NumPy.

The modified YOLO detector's accuracy in smart traffic management within urban environments gauges its ability to precisely recognize and categorize elements like vehicles, pedestrians, and obstacles amidst traffic. This metric is calculated by determining the ratio of correctly identified objects to the dataset's overall object count.

Sensitivity, also termed the true positive rate, assesses the model's capability to accurately detect actual positive cases. In the realm of smart traffic management, sensitivity indicates how effectively the modified YOLO detector identifies relevant objects such as vehicles and pedestrians, crucial for real-time decision-making to enhance urban mobility.

On the other hand, specificity, referred to as the true negative rate, evaluates the model's accuracy in identifying negative cases or unimportant objects. In the context of smart traffic management, specificity reflects the modified YOLO detector capacity to minimize false alarms or misclassifications, ensuring resources are directed towards genuine traffic-related elements rather than irrelevant ones.

These metrics play a pivotal role in evaluating the modified YOLO detector's performance and its contribution to enhancing urban mobility through smart traffic management.

Ultimately, the proposed framework is designed to seamlessly integrate with real-world traffic data, offering significant benefits to traffic management systems through its adaptive and intelligent approach.

IV. RESULT ANALYSIS

The datasets are taken from Dhaka AI Traffic Challenge Weights–Yolov5 [16]. The considered vehicle classes are: ambulance, auto-rickshaw, bicycle, bus, car, garbage van, human hauler, minibus, minivan, motorbike, Pickup, army vehicle, police car, rickshaw, scooter, SUV, taxi, three-wheelers (CNG), truck, van, wheelbarrow etc. as shown in Figure 3. Figure 4 shows the annotated images and finally figure 5 shows the detection of different types of vehicles. Finally, the result analysis is presented in table II.

Fig. 3. Vehicle images

Fig. 4. Annotation on Images

Fig. 5. Different types vehicle

TABLE II. ANALYSIS OF RESULTS

Model	Category	Total Frames	FPS	ACC	SEN	SPE	Time
Cascade object detector	Vehicles, Humans	100	20	82.01%	88%	63%	8.04sec
RCNN-Detection	Vehicles	100	20	78.23%	88%	52%	9.23sec
YOLO object detector	Vehicles	100	20	89.07%	95%	62%	5.22sec

The comparative evaluation of object detection models highlights the superior performance of the YOLO object detector, achieving an accuracy of 89.07%, sensitivity of 95%, and specificity of 62% across various multi-vehicle scenarios. The results shown in Table II emphasize YOLO's potential as a foundational component in real-time traffic management and surveillance systems, offering enhanced accuracy and efficiency in object recognition tasks.

CONCLUSION & FUTURE SCOPE

The adoption of YOLO-based models can significantly improve traffic monitoring capabilities, contributing to congestion alleviation and overall road safety. Looking ahead, future research should prioritize several key areas to further advance object detection technologies in traffic-related applications. First and foremost, continuous model refinement through extensive training with diverse datasets can enhance accuracy and generalization, ensuring consistent performance across diverse traffic conditions. Second, integrating object detection systems with adaptive traffic control algorithms can optimize traffic flow dynamically, responding effectively to real-time changes in congestion patterns and traffic volume. Furthermore, exploring advancements in edge computing and multi-modal data fusion techniques can expedite data processing and integration from various sources, thereby improving real-time decision-making for traffic management authorities. Addressing privacy concerns through effective anonymization techniques while maintaining robust object detection capabilities will also be pivotal in future research and implementation endeavors.

REFERENCES

[1] Pi, Y., Duffieid N., Behzadan A., Lomax, T. (2022). Visual Recognition for Urban Traffic Data Retrieval and Analysis in Major Events Using Convolutional Neural Networks. Computational Urban Science, Springer, 2(1). pp. 1-16

[2] Chandrasekara, W. A. C. J. K., Rathnayaka, R. M. K. T., Chathuranga, L. L. G. (2020, December). A Real-Time Density-Based Traffic Signal Control System. 5th International Conference on Information Technology Research (ICITR), IEEE. pp. 1-6.

[3] Sahu, S. P., Dewangan, D. K., Agrawal, A., & Priyanka, T. S. (2021, March). Traffic light cycle control using deep reinforcement technique. International Conference on Artificial Intelligence and Smart Systems (ICAIS), IEEE. pp. 697-702.

[4] Zhu, Y., Yan, W. Q. (2022). Traffic sign recognition based on deep learning. Multimedia Tools and Applications, Springer, 81(13). pp. 17779-17791.

[5] Sharma, M., Bansal, A., Kashyap, V., Goyal, P., & Sheikh, T. H. (2021). Intelligent Traffic Light Control System Based on Traffic Environment Using Deep Learning. IOP Conference Series: Materials Science and Engineering, ICCRDA, 1022(1). p. 012122.

[6] Navarro-Espinoza, A., López-Bonilla, O. R., García-Guerrero, E. E., Tlelo-Cuautle, E., López-Mancilla, D., Hernández-Mejía, C., Inzunza-González, E. (2022). Traffic Flow Prediction for Smart Traffic Lights Using Machine Learning Algorithms. Technologies, 10(1). p. 5.

[7] Ibrokhimov, B., Kim, Y. J., & Kang, S. (2022). Biased pressure: cyclic reinforcement learning model for intelligent traffic signal control. Sensors, 22(7). p. 2818.

[8] Lilhore, U. K., Imoize, A. L., Li, C. T., Simaiya, S., Pani, S. K., Goyal, N., Lee, C. C. (2022). Design and Implementation of an ML and IoT Based Adaptive Traffic-Management System for Smart Cities. Sensors, 22(8). p. 2908.

[9] Bouktif, S., Cheniki, A., Ouni, A. (2021). Traffic signal control using hybrid action space deep reinforcement learning. Sensors, 21(7). p. 2302.

[10] Naveed, Q. N., Alqahtani, H., Khan, R. U., Almakdi, S., Alshehri, M., Abdul Rasheed, M. A. (2022). An intelligent traffic surveillance system using integrated wireless sensor network and improved phase timing optimization. Sensors, 22(9). p. 3333.

[11] Ijeri, D., Maidargi, P., Sunagar, R. (2020, October). Traffic Control System Using Image Processing. IEEE Bangalore Humanitarian Technology Conference, IEEE, 10(2). pp. 1-6.

[12] Gandhi, M. M., Solanki, D. S., Daptardar, R. S., & Baloorkar, N. S. (2020, December). Smart control of traffic light using artificial intelligence. International Conference on Recent Advances and Innovations in Engineering (ICRAIE), IEEE, 5(2). pp. 1-6.

[13] Meng, B. C. C., Damanhuri, N. S., Othman, N. A. (2021, February). Smart traffic light control system using image processing. IOP Conference Series: Materials Science and Engineering, AC2SET, 1088(1). p. 12-21.

[14] Qadri, S. S. S. M., Gokçe, M. A., Oner, E. (2020, October). State-of-art review of traffic signal control methods: challenges and opportunities. European transport research review, Springer Open, 12(1). pp. 1-23.

[15] Oliveira, L. F. P., Manera, L. T., & Da Luz, P. D. G. (2020, March). Development of a smart traffic light control system with real-time monitoring. IEEE Internet of Things Journal, 8(5). pp. 3384-3393.

[16] Md. Asif Iqbal Fahim, "Dhaka AI Traffic Challenge Weights - Yolov5 ." Kaggle, 2021, https://doi: 10.34740/KAGGLE/DSV/2009226 .

Extended Pigpen Cipher with New Variations using ASCII and Division Substitution

Manjunatha S
Dept. of Computer Applications
PES University
Bangalore-560085, India
manjunathsathya489@gmail.com

Thenmozhi S
Dept. of Computer Applications
PES University
Bangalore-560085, India
thenmozhis@pes.edu

Abstract—**Secure communication has always been a primary focus and have challenges in encrypting messages. As the ancient cipher techniques that were used to secretly communicate have been compromised in today's era. Among these, one is the Pigpen cipher. The pigpen cipher is a symmetric encryption method that converts alphabets into symbolic representations known as Pigpen. However, the method also has some drawbacks, the operation is done without a key and it does not have the capability to encrypt numbers. The Pigpen cipher offers only one variation adding to its drawbacks and lacks achieving robust encryption in modern communication contexts. AlphaMeshX(AMX) is a new age cipher technique that is an Extended Pigpen cipher technique with Substitution. AlphaMeshX presents an innovative approach to encryption and decryption of messages, building upon the foundation of an encryption technique that was historically used by Freemasonries during communication called as Pigpen Cipher also called as Freemason Cipher. AMX is a key based Encryption and Decryption technique. The key, to which the ASCII encoding is done, then followed by the division substitution for the sum of ASCII. The resultant value after the mathematical substitution determines the 13 variations for alphabets and 18 variations including numbers. Therefore, the result is more secure.**

Keywords—AlphaMeshX(AMX), Pigpen, Extended Pigpen, Substitution Cipher, Freemason, Masonic cipher.

I. INTRODUCTION

In today's world, there's a continual need for robust encryption methods that ensure privacy and confidentiality. AlphaMeshX that is built on the foundation of Pigpen cipher [1] technique. It also known as the Masonic Cipher and is a mono-alphabetic Substitution Caesar's Cipher [3] that employs graphical symbols to encode letters. Originating from the Freemasons, this historical cipher involves replacing each letter with a unique symbol based on a grid of Mesh and X shaped grids. The alphabet is typically divided into four quadrants, with each quadrant assigned a distinct set of symbols. How is it encrypted? 'A' could be a symbol located in the top-left section, and 'B' might be a different symbol in the same area. To encode a message using this cipher, you replace each letter with its designated symbol according to the grid. This results in a visually encoded message where symbols represent the original letters. Deciphering the message requires knowledge of this. It offers cipher technique that is less secure, it is different from other simple mono-alphabetic substitution ciphers because its use of symbols rather than letters and due to the use it fails to assist in cryptanalysis. Pigpen encryption method faces the challenges in terms of advanced techniques.

Pigpen cipher encrypts only alphabets but not numbers. But there is an extended version of Pigpen where the grid is designed in such a way that it can occupy alphabets as well as numbers designed by Thomas Hunter II [13]. Though the extended version encrypts numbers as well, it still has only one variation as the initial Pigpen. AlphaMeshX leverages the Pigpen Cipher and the extended Pigpen as the foundation and introduces 18 variations. AMX is a key based cipher technique to which ASCII encoding is done [2] and followed by division method. Based on the resultant value after the mathematical substitution of AMX, the value determines the type of variation and based on the variation the message is encoded. The alphabets and numbers representation in the AMX quadrants are placed as per the variation, so for each variation all the character representation is distinct. Always, an encoder needs to have a key along with the message, then the key will undergo mathematical substitution where the ASCII encoding is done and the sum is determined. Further, the mod operation is done to determine the variation. Based on the resultant value the variation grid is formed then the characters are encrypted. The decryption process is as same as the encryption since it is a symmetric encryption process, that is the same key is used for decryption of the encoded message.

The purpose of AlphaMeshX is to provide an innovative and versatile encoding and decoding that includes offering a creative tool based on Pigpen Cipher variant and open doors to 18 new variants. AMX (AlphaMeshX) is a Symmetric cryptography technique. Symmetric cryptography involves conversion of plain text information, along with a cryptography key, into a cipher text. Decryption with symmetric technique, recovering the original plain text from the cipher text by using the same algorithm and same key that was used to encrypt the information [5].

The AMX can find real world applications in education, personal privacy communication, open-source contributions and cryptography research. The primary focus lies in enhancingthe security and confidentiality.

II. RELATED WORK

Freemason's cipher is a simple substitution cipher using a grid of symbols to represent letters. It involves drawing shapes within a three-by-three square to form the cipher [1]. Researcher Thomas Hunter II, licensed the patent for the "Extended Pigpen Cipher to Include Numbers" [13], conceptualised a method by allowing the messages with numbers to be encrypted. The modification introduces additional symbols specifically designed to represent numbers, addressing a limitation of the original cipher.

Researcher Bradley P. Franger made a survey that discusses about the Masonic cipher, a monoalphabetic cipher. It replaces individual characters with its corresponding symbols. This cipher can operate with letters but not numbers. The author also have done analysis on Zigzag and Rosicrucian ciphers [7].

The paper titled "PIGPEN: A Novel Approach for Securing Data" addresses the security challenges in wireless communication and proposed modifications to Letter- Shape encryption to enhance data protection. Their focus is on TIC-TAC-TOE shape, X and dots as core elements to incorporate pigpen cipher technique. This method aims to secure communication in open and dynamic mobile wireless networks. The unique encryption technique adds complexity and confusion for potential attackers, ensuring a higher level of data security [8].

In Research paper "Enhancing Image Stenography Methods by using New Secret Message Encoding Technique based on Pigpen Cipher (Pigpen Encoding) ", the researchers explored an enhanced stenography technique that involves modifying the secret message into cipher text. Stenography is a way to hide the secret messages in carriers such as images, audio files, text files and videos [9]. Unlike ASCII encoding requiring three decimal digits, this approach represents secret message characters with only two decimal digits, reducing the space needed for embedding by one third. This Pigpen cipher replaces letters with symbols based on their corresponding symbols location. The method involves creating tables for small and capital letters, numbers, and special characters, assigning each a unique two-digit representation. This innovative encoding technique optimizes the performance of hiding image methods in Stenography.

A detailed understanding is done on the existing cipher techniques considering the predominant features such as variations, keys, encoding methods, computations involved. The existing techniques were compared with the proposed method and it is tabulated in Table I. The drawbacks of the existing methods were well understood and the novel cipher technique AlphaMeshX is proposed which combats the older techniques.

TABLE I
EXISTING WORK COMPARISON

	Pigpen Cipher	Extended Pigpen	AlphaMeshX
Variations	1	1	18
Multiple Variations	N	N	Y
key	N	N	Y
Alphabets	Y	Y	Y
Numbers	N	Y	Y
Math calculation	N	N	Y
ASCII	N	N	Y

III. METHODOLOGY

The functioning of AlphaMeshX (AMX) involves a step of substitution cipher. Where the user provides a key along with the message also called as the plain text. Once the key is provided, the key undergoes a mathematical substitution process. That is encoding of ASCII values to the key and the sum is taken. Then the sum is modded by 18 to determine the variation. Based on the variation the plain text is converted into cipher text. The process can be understood by the below process flow diagram.

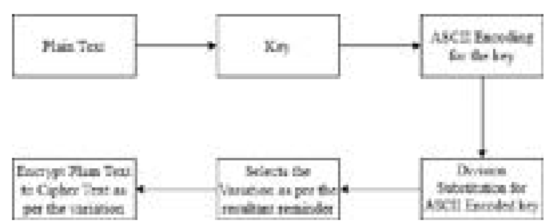


Fig. 1. Process flow of encryption operation

Now the mechanism of both existing techniques and the innovative AMX method is explained to gain a understanding of their functionalities and explore the new AMX technique. For Instance, let the message to be encrypted be "pes473" and key for the encryption is "piGpen".

A. Pigpen Cipher

Pigpen Cipher, the encryption of the message takes place just by replacing the characters to its corresponding symbols. It does not involve the use of any key. So, the encrypted message would be,

Fig. 2. Pigpen implementation (original)

Decoding will be the reversal of encoding, replacing the symbols to its corresponding characters.
Drawback - It does not encode numbers and it has only one variation stating that for a letter there is only one corresponding symbol.

B. Extended Pigpen Cipher

Extended version of the Pigpen cipher encrypts alphabets along with numbers, Thomas Hunter II proposed the design of extended version in such a way that it can occupy numbers, therefore encrypting both the alphabets and numbers. The encoding in this technique is as same as the Pigpen cipher which is replacing the characters to its corresponding symbols. Thus, overcoming the drawback of the original Pigpen Cipher. So, the encrypted message will be,

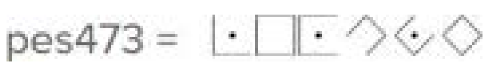


Fig. 3. Extended pigpen implementation

So unlike the original Pigpen cipher we have symbols for numbers as well, deciphering is as same as the Pigpen cipher. Drawback - is that it still has only one variation and does not use any key.

C. AlphaMeshX or AMX

AlphaMeshX Cipher (AMX), it is key based symmetric encryption and decryption technique. Unlike the original Pigpen and the extended Pigpen, it uses ASCII encoding for the key and division method to determine the variation.

Key = “piGpen” and message = “pes473”

Now, for the key we integrate ASCII values, each character is replaced by its ASCII value and the sum of it is determined. In the below format,

p+i+G+p+e+n = 112+105+71+112+101+110 = 611

Now, the value 611 is determined based on the user specified key. This value is now modded by 18, because the AMX has 18 variations and to determine the variation type it is necessary to mod it by 18. Therefore, the final resultant will be the reminder.

611 mod 18 = 17

After modding the value 611 with 18 we get the reminder as 17 so this determines the 18th variation, the 17th cell representation is 18th variation since the total cells in the MAX model is 18-1=17. The character arrangements begin from the cell 17th. That is, each cell is a variation.

Let us see the 17th variation grid formation and the character representation of the grid.

So, based on our variation and the character representation of the variation in the grid the encrypted message will be,

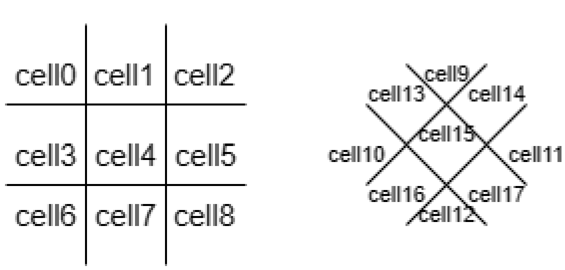


Fig. 4. Variation representation

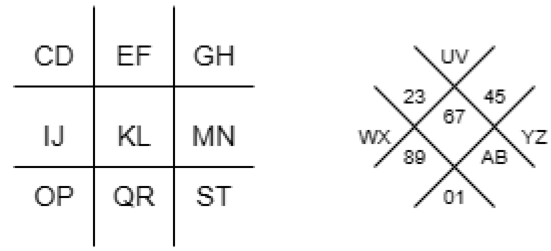


Fig. 5. Characters representation for the variation

Therefore, the result as per the variation after encryption using AlphaMeshX cipher technique is,

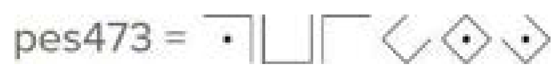


Fig. 6. AlphaMeshX Output

Now we can see that the same input is producing different output. Since the AMX uses a key for its operation, though the extended Pigpen and AMX performs numbers encryption, AMX gives different encrypted output when compared to extended version based on the variation that was determined by the key.

Unlike the original and the extended version, The AMX has 18 different variations for both alphabets and numbers. So, there are 18 symbols for each character and for the given example “pes473” it will have 18 different symbols for each character. Thus overcoming the single variation drawback from both Pigpen and the extended version. The decoding is the same way as encoding since it is a symmetric cipher technique and requires the key for decoding.

IV. Experiential Results

AMX cipher’s security enhancements and implementation, through the implementation of AlphaMeshX (AMX), the experiential results have resulted in 18 new variations, increasing the security of AMX beyond the original Pigpen Cipher technique and its variant. The demonstration of the AMX implementation the Flask framework was used. Flask is a lightweight Python web framework. It is popular due to powerful modular design that lets us to build scalable web apps [14]. Which is a popular Python web framework known for its simplicity and flexibility. Showcasing the AMX’s functionality, demonstrating its robust encryption and decryption processes in action.

By utilizing SHA256[11] encoding for the user provided key, it enhances security by creating a strong, complex, and secure key, making it challenging for attackers to decipher even if they manage to get hold of the encrypted key. SHA-256 is a cryptography hash function that generates a unique, fixed-length (256 bits) string from an input. Additionally, SHA512[12] (512 bits) encoding can be explored to further bolster key security, ensuring resilience against potential threats. These encoding techniques significantly enhance the confidentiality and integrity of encrypted messages, reinforcing AMX's position as a reliable and secure encryption tool compared to the initial Pigpen Cipher technique.

Let us understand the experiential process for the AMX. let us consider an example where user A is sending a message tothe user B. Now the user A will take a key as “keyencryption”and let the message be hello. Now performing SHA256 forthe key will result us with a 256 but of 64 characters' lengthunique and unreadable string that cannot be traced back.That is, once SHA256 hash function is performed to the “keyencryption” key, then it will result us with
”bf9eaac6bbd146a9840ab409df90ff48b4383fd09482618f21df fbd5b2c198ec”.
which is long, randomized and confusing compared to the initial key. So now the actual key ”keyencryption” is encoded as an unreadable string of length of 256 bits. Further, this sha256 key will be our newly encoded key. To which the key we will be undergoing the AMX mathematical substitution. The AMX substitution involves ASCII encoding and the total sum is taken. Finally for the total sum, 18 will be modded to determine the variation. Once the mod of 18 is done, we will have the resultant reminder which gives us the variation and the plain text (message) will be encoded according to the resultant variation to the cipher text. As for the decryption the same steps is followed since the AMX is a symmetric cipher technique.

Let us understand the entire Experiential process that uses SHA256 hash function, integratingAndDivide function for the key. Also the symmetric technique for the encryption and decryption. The below is the Pseudo code for all the functionalities.

(i) Encryption:

```
Procedure encryption(message, key)
  hashedKey = SHA256(key) –(call SHA256(key))
  encodedKey = integrateAndDivide(hashedKey) –(call integrateAndDivide(key))
  if AMXEncryption then
   encryptedMessage = AMXEncrypt(message, encodedKey) –call(AMXEncrypt(message, encodedKey))
   return encryptedMessage
  else
  // Handle error for invalid mode
   return ”Invalid mode”
  end if
end procedure
```

(ii) Decryption:

```
Procedure decryption(encryptedMessage, hashedKey)
  encodedKey = integrateAndDivide(hashedKey) –(call integrateAndDivide(key))
  if AMXDecryption then
   decryptedMessage = AMXDecrypt(encryptedMessage, encodedKey) –call(AMXDecrypt(encryptedMessage, encodedKey))
   return decryptedMessage
   else
   // Handle error for invalid mode
   return ”Invalid mode”
  end if
end procedure
```

(iii) SHA256 hash function:

```
Procedure SHA256(key)
  // Implementation for SHA-256 hash function
  return hashedKey
end procedure
```

(iv) ASCII integration and Modding:

```
Procedure integrateAndDivide(key)
  totalSum = 0
  for each character in key:
    asciiValue = ASCII value of character
    totalSum = totalSum + asciiValue
  end for
  encodedKey = totalSum mod 18 —(18 variations)
  return encodedKey
end procedure
```

(v) AMXEncrypt:

```
Procedure AMXEncrypt(message, encodedKey)
  // Generate grid based on remainder ie., variation
  // Encrypt message using grid and variation
  return encryptedMessage
end procedure
```

(vi) AMXDecrypt:

```
Procedure AMXDecrypt(encryptedMessage, encodedKey)
  // Generate grid based on remainder ie., variation
  // Decrypt encryptedMessage using grid and variation
  return decryptedMessage
end procedure
```

There is always need for continual improvement and performance enhancement. It can use various advance encryption techniques such as Advanced Encryption Standard (AES) algorithm [10], Data Encryption Standard (DES)[11], or asymmetric block cipher algorithm RSA[12] and so. This cipher technique can be used to add a extra security by encrypting the already encoded message. Thus making it even more secure. Further, even SHA512 or any other hash function method can be used to generate an encoded string for the key.

CONCLUSION

While the Pigpen cipher offers only one variation and primarily encrypts only alphabetic characters but not numbers. The extended version of the Pigpen also offers only one variation but has a way that encrypts both alphabetic characters as well as numbers. But both lack the crucial element of using a key for secure communication and has only one variation. This presents significant security vulnerability. In contrast, the AalphaMeshX system addresses these limitations by employing a key, utilizing mathematical substitution and introducing a layer of complexity for enhanced security by Integrating ASCII and the division method. This Combination creates 18 distinct variations, significantly increasing the difficulty of deciphering the encrypted message. Therefore, AMX offers a more robust and secure method of communication compared to the Pigpen and the extended cipher technique. The method also has certain limitations, though the AMX has a way to encrypt both alphabetic characters as well as the numbers, giving rise to 18 distinct variations and has a symmetric way to encrypt and decrypt, it cannot encrypt any special characters and currently uses a single algorithm based on the Pigpen Cipher.

In the future, enhancing the AMX involves integrating it with multi-factor authentication for the key and various advanced encryption algorithms. Additionally, the AMX can encode the encrypted message, creating a unique double encryption technique that adds complexity and confusion for potential attackers. This strategy increases the security of the already encrypted message beyond expectations, making it more secure.

REFERENCES

[1] Anvesh Mallick, "HISTORICAL SUBSTITUTION CIPHERS THE PIGPEN CIPHER, THE MLECCHITA VIKALPA CIPHER AND THE MULABHADRA CIPHER", 2021

[2] Karan Goyal, Sakshi Sharma, Tarun Kumar, Deepak Garg, Prateek Thakral, "Cryptography: A Game Play Using ASCII Conversion", 2018

[3] Alan G. Konheim, "Monoalphabetic Substitution", 2006

[4] Lawrence E. Hughes, "Basic Cryptography: Symmetric Key Encryption", 2022

[5] B Pfranger, "Masonic Cipher", 2016

[6] Mamta Martolia, Priti Dimri, Nitin Arora, "PIGPEN: A Novel Approach for Securing Data", 2018

[7] Abdelmgeid A.A, Bahgat A.A, Al-Hussien Seddik Saad, Maha Mohamed Gomaa, "Enhancing Image Steganography Methods by using New Secret Message Encoding Technique based on Pigpen Cipher (Pigpen Encoding)", 2017

[8] Gowthaman A, M Sumathi, "Performance Study of Enhanced SHA-256 Algorithm" 2015

[9] Shay Gueron, Simon Johnson, Jesse Walker, "SHA-512/256", 2011

[10] Ako Muhamad Abdullah, "Advanced Encryption Standard (AES) Algorithm to Encrypt and Decrypt Data", 2017

[11] Khalid Nahar, Obaida Al-Hazaimeh, Moyawiah Alshanaq, Mohammed Bawaneh, "Analytical Approach for Data Encryption Standard Algorithm", 2023

[12] Xurui Zhang, "Development of RSA and some attack points", 2023

[13] Thomas Hunter II, "https://thomashunter.name/posts/2012-07-13-extended-pigpen-cipher-to-include-numbers", 2012

[14] Shalabh Aggarwal, "Flask Framework Cookbook - Second Edition", 2019

Enhancing Career Pathways: Advancing Guidance Systems with XGBoost Algorithm

Dr. Sonali J Mahure
Associate Professor, Dept. of CSE
New Horizon College of Engineering
Bangalore, India
sendtowiser@gmail.com

Dr. B. Rajalakshmi
HOD, Dept. of CSE
New Horizon College of Engineering
Bangalore, India
hod_cse@newhorizonindia.edu

T Eesha Manohar
Student, Dept. of CSE
New Horizon College of Engineering
Bangalore, India
eeshamanohar2003@gmail.com

Vaishnavi S R
Student, Dept. of CSE
New Horizon College of Engineering
Bangalore, India
vaishnavisr29@gmail.com

Sumithra
Student, Dept. of CSE
New Horizon College of Engineering
Bangalore, India
sumithramarakala@gmail.com

Nancy Priya
Student, Dept. of ISE
New Horizon College of Engineering
Bangalore, India
nancypriyasmp2001@gmail.com

***Abstract*— This research study introduces a comprehensive framework tailored to navigate the complexities of contemporary career guidance. It integrates interconnected modules, including secure login, student data collection, academic evaluation tools, and an advanced skill assessment module. Notably, the Cognitive Decision Support module employs advanced analytical tools to analyze users' cognitive responses, delivering personalized career suggestions. Moreover, the framework features a career prediction module equipped with algorithmic models for forecasting career paths and suggesting suitable colleges nationwide. The dynamic skill evaluation component continually monitors industry trends, providing guidance for ongoing skill development. Additionally, it offers insights into the global job market and cultural trends, facilitating international career exploration. A feedback mechanism ensures continuous improvement. Despite its technical sophistication, challenges arise in managing large-scale data processing and ensuring precise cognitive assessments, potentially affecting the accuracy of personalized career recommendations.**

Keywords—Decision support, Career prediction, International job market, Global career opportunity, Skill evaluation

I. Introduction

This research study introduces a sophisticated and integrated framework designed to transform career guidance and skill development in the contemporary digital landscape. At its core, the framework consists of meticulously crafted modules catering to diverse user needs and aspirations. Starting with a robust login and registration mechanism, users seamlessly enter the system, prioritizing data security and user authentication. Following this, the framework facilitates comprehensive management of student profiles through a dedicated page for capturing crucial information essential for personalized career guidance. Academic evaluation tools enable systematic assessment of users' educational performance, while a sophisticated skill assessment module provides a detailed analysis of proficiencies across various domains.

Key to the framework's effectiveness is the innovative Cognitive Decision Support module, utilizing cutting-edge analytical tools to interpret users' cognitive responses during assessments. This module not only enhances assessment accuracy but also delivers personalized career recommendations based on individual cognitive profiles. Additionally, the framework includes a Career Prediction module that employs predictive algorithms to outline potential career paths and recommend suitable colleges nationwide, empowering users with informed decision-making capabilities. Moreover, a dynamic skill evaluation module continually monitors industry trends and emerging technologies, offering targeted recommendations for ongoing skill development. Furthermore, the framework provides users with a holistic global perspective on career opportunities, leveraging insights into the international job market, cultural trends, and economic indicators. Finally, a feedback module encourages iterative refinement and continual improvement of the framework, ensuring its ongoing relevance and effectiveness in guiding users toward fulfilling career trajectories.

II. Literature Review

In reference to research paper with An Intelligent Career Guidance System utilizing Machine Learning, developed by Vignesh S, Shivani Priyanka C, Shree Manju H, and Mythili K, aims to assist students globally who often grapple with uncertainty upon completing their higher secondary education. At the age of 18, students lack the maturity needed to decisively chart a suitable career path. This phase prompts a cascade of doubts and contemplation regarding what to pursue post-12th grade, posing the most significant query. Subsequently, individuals question whether their chosen field aligns with their essential skills. Our automated career counselling system employs objective tests to evaluate an individual's skills, thereby predicting the most fitting department. Completing our online assessment ensures a more informed choice of course, significantly reducing the risk of selecting an unsuitable career path. **Merit:** The Intelligent Career Guidance System stands out for its ability to deliver tailored career recommendations through objective assessments, assisting students in making informed decisions about their future career paths. **Demerit:** the limitations arises from the system's exclusive reliance on objective tests, which may overlook various facets of an individual's skills, interests, and aspirations, potentially resulting in restricted or biased career recommendations. [1]

Intelligent web platform for vocational guidance. Andres F. Cruz *, Laura Orozco, Carolina Gonzales with. The development of Computer Assisted Career Guidance Systems (CACGS) arose as a partial remedy for the scarcity of qualified professionals in career guidance and the need for

a structured system to oversee the guidance process. Nevertheless, numerous existing systems are tailored for settings distinct from that of Colombia, posing difficulties for counsellors to effectively employ their results or for students in Popayan city to accurately comprehend them.
Merit: The establishment of Computer Assisted Career Guidance Systems (CACGS) serves to alleviate the shortage of qualified professionals in career guidance by furnishing a structured platform for vocational direction.
Demerit: However, the lack of customization in existing systems may pose challenges for counsellors to efficiently leverage their outcomes, while students in Popayan city may encounter difficulties in accurately grasping them due to the systems not being tailored to the specific context of Colombia. [2]

The Online Study and Recommendation framework, either open or private, caters to individual needs by fostering a collaborative learning environment on the internet. Reddy et al. introduced the concept during their Computer Supported Cooperative Work graduate course, envisioning a unified, secure, and dynamic system that offers course recommendations and collaborative study features. Their discussions centered on the system's approach, implementation, recommendation algorithms, and the findings from user studies. Existing research systems often lack security and fail to provide adequate privacy for students studying in a distributed environment. Hence, their creation focuses on utilizing both individual and group learning processes, incorporating several features to support collaborative activities.
Merit: The Online Study and Recommendation framework cultivates a collaborative learning atmosphere, providing personalized course recommendations and collaborative study functionalities.
Demerit: In distributed environments, many existing research systems suffer from security and privacy deficiencies, underscoring the necessity for robust security measures within the framework. [3]

A machine learning approach for recommending courses by John Britto, Sagar Prabhu, Abhishek Gawali, Yogesh Jadhav proposed Machine learning techniques, such as neural networks and diverse learning algorithms, hold substantial potential in shaping a student's career trajectory. Although there's space for enhancements, this system proves beneficial in conducting initial assessments to gauge a student's interests. Recommendations for courses are based on both preferences and abilities, with all development conducted using Python 3.6. Looking ahead, there's scope to develop a comprehensive questionnaire that delves into every facet of a student's life and capabilities, offering not just domain suggestions but an entire tailored curriculum aligned with the individual's pace and preferences. Furthermore, this system's applicability could extend beyond engineering to encompass other fields, potentially aiding in broader career guidance and counselling endeavours.
Merit: Utilizing a machine learning approach, course recommendations are personalized according to individual preferences and abilities, with the potential to significantly influence a student's career path.
Demerit: Although advantageous for initial assessments, the system may necessitate additional improvements to fully encompass all facets of a student's life and capabilities, which could constrain its efficacy in delivering tailored curriculum recommendations. [4]

In reference to Dahanke Ajay's research emphasis on development of a web-based application tailored for career guidance, offering improved recommendations for individuals seeking suitable departments. This study aims to enhance the precision of career guidance systems by leveraging advanced algorithms. The proposed system's efficacy surpasses existing models, primarily attributing its accuracy to the utilization of the K-Nearest Neighbour (KNN) algorithm for skill set classification and department prediction. Additionally, the study incorporates the K-Means Clustering algorithm to form clusters based on students' skill set scores, thereby determining the success rates of various departments within each cluster.
Merit: Dahanke Ajay's research presents a web-based application for career guidance, employing advanced algorithms such as K-Nearest Neighbour and K-Means Clustering to refine recommendations for appropriate departments.
Demerit: Despite the potential accuracy improvements, the system's dependence on sophisticated algorithms may present implementation and user comprehension challenges, possibly hindering its accessibility and usability. [5]

In reference to author Lei Wang say The paragraph discusses addressing challenges in college student employment guidance by utilizing Visual Fox-Pro programming to create a data verification tool, enhancing data accuracy. Additionally, it employs the Apriori algorithm in data mining to analyse university employment history, aiming to understand factors impacting student employment. This mining process uncovers correlations between education attributes and employment factors. Furthermore, the optimization algorithm tests successfully reduce meaningless association rules, improving the efficiency of the mining process.
Merit: Lei Wang's methodology employs Visual Fox-Pro programming and the Apriori algorithm to boost data precision and reveal links between educational attributes and student employment aspects, offering promise for enhancing college student employment guidance.
Demerit: Despite optimizing association rules successfully, the reliance on Visual Fox-Pro programming and the Apriori algorithm may necessitate specialized expertise, potentially impeding accessibility and usability for users lacking familiarity with these tools. [6]

In reference to the paper with DOI: 10.1109/IHCSP56702.2023.10127152, author Vijaya R. Kumbhar’s Career guidance is particularly crucial for college students who often lack clear understanding about their career paths. Many are not certain about their future work preferences or the ideal degree choice. By assessing a user's skills, interests, and preferences, this system recommends various career options. The results assist students in focusing on subject areas that closely match their expertise and interests. Utilizing a Smart Career Guidance System integrated with a chatbot, the model employs a hybrid deep learning approach, merging recurrent neural network (RNN), Long Short Term Memory (LSTM), and Natural Language Processing (NLP) techniques. This proposed hybrid algorithm combines elements from RNN, LSTM, and NLP to enhance the Career Guidance System's effectiveness.
Merit: Vijaya R. Kumbhar's paper presents a Smart Career Guidance System with a chatbot, employing a hybrid deep

learning approach that integrates RNN, LSTM, and NLP techniques to improve the efficacy of career recommendations based on user skills, interests, and preferences.

Demerit: Despite its effectiveness, the implementation of the hybrid deep learning approach may demand considerable computational resources and technical expertise, potentially impeding accessibility and scalability for users and organizations lacking in sufficient technical capabilities. [7]

III. PROPOSED SYSTEM

The proposed system aims to provide extensive career guidance and assistance to users through various modules powered by the XGBoost algorithm. It commences with a secure login/registration process, followed by modules for managing student information, academic assessment, and skill evaluation. Utilizing cognitive decision support tools, the system analyzes user input to offer personalized career recommendations, predict career trajectories, and suggest colleges nationwide. Furthermore, dynamic skill evaluation modules recommend continuous skill enhancement opportunities based on industry trends, while the global career perspective module offers insights into international job markets. Users can offer feedback to enhance system functionality. In essence, the system aims to empower users to make informed career choices and excel in their chosen paths

A. Flow diagram

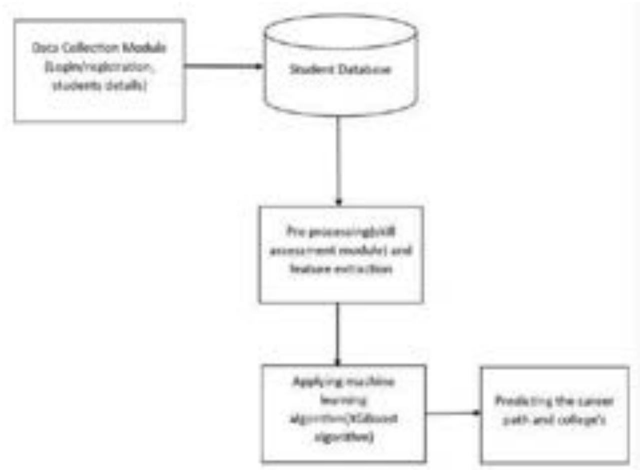


Figure 1: flow diagram

The modules included in figure 1 are as follows:

1. Data Collection Modules:

The data collection module effectively collects thorough user data to customize personalized career guidance. The modules under collection module are as follows:

- **Login Page / Registration Page Module: -** During registration, this crucial module facilitates user onboarding by gathering essential information such as name and email, while ensuring security through verification methods during the login process, and providing user-friendly error handling to enhance the overall experience.
- **Student Details Page Module:** Committed to offering a comprehensive perspective, this module captures both personal details and academic history, allowing users to dynamically update information, display achievements via document uploads, and emphasize skills and interests to create a thorough profile.
- **Career Global Perspective Module:** Centered on broadening perspectives, this module consolidates information on the global job market, cultural shifts, and economic indicators, providing users with insights into international career opportunities, industry trends, and language proficiency to facilitate informed decision-making.
- **Feedback Module:** Functioning as an interactive platform, this module acts as a tool for ongoing enhancement by employing satisfaction surveys, bug reporting, and feature requests to collect valuable feedback. With a user-friendly rating system and the option for anonymity, it fosters open communication to iteratively refine the system.

2. Data Processing Modules:

The data preprocessing module ensures that collected user data is organized and optimized for effective analysis, enabling the generation of personalized career guidance. The modules under data pre-processing are as follows:

- **Academic Evaluation Module: -** Performing crucial pre-processing tasks, including normalization, on academic data, this module ensures uniformity and comparability, thereby improving data quality for precise analyses and informed decision-making within educational settings.
- **Skill Assessment Module: -** Centered on skill assessment data, this module conducts essential pre-processing tasks, utilizing methods such as standardization. It converts skill-related data into a consistent format, facilitating equitable comparisons and trustworthy analyses. The module's contributions are instrumental in informing decisions regarding talent management and initiatives for skill development.

3. Feature Extraction Module: Within the Feature Extraction Module, intricate patterns and distinctive elements are meticulously identified and isolated from pre-processed data. Through the use of sophisticated algorithms, this module distills essential features that encapsulate the intrinsic information necessary for precise analysis. By condensing the data into a representative set of key attributes, it streamlines downstream processing, ensuring subsequent models focus on the most relevant information. This feature-centric approach enhances overall interpretability and system performance, enabling informed decisions based on extracted, notable features. Ultimately, the Feature Extraction Module is a vital component in optimizing and refining data-driven tasks.

4. Career prediction module: The career prediction module adopts a comprehensive approach to individual growth, integrating three interconnected components. Cognitive Decision Support employs the XGBoost algorithm to analyze cognitive responses in assessments, generating personalized career recommendations. Simultaneously, the Dynamic Skill Evaluation module utilizes XGBoost to analyze industry trends and user profiles, providing insights for continuous skill development. Meanwhile, the Career Prediction Module applies XGBoost to forecast future career trajectories and suggest colleges, directing individuals towards rewarding and successful career paths that align with their strengths and aspirations.

B. Algorithm used

XGBoost, abbreviated from Extreme Gradient Boosting, stands out as a remarkably efficient and scalable machine learning algorithm, celebrated for its outstanding performance across classification, regression, and ranking assignments. Its operation involves iteratively constructing numerous decision trees, with each tree aimed at rectifying the mistakes of its predecessors. Employing a technique called boosting, XGBoost amalgamates the predictive capabilities of these weak learners into a robust ensemble model.

The algorithm incorporates regularization techniques to prevent overfitting and optimize model generalization. XGBoost's versatility and effectiveness have made it a popular choice across various domains. Its ability to handle large datasets and complex relationships, coupled with efficient parallelization and optimization, has cemented its reputation as a go-to algorithm for data scientists and machine learning practitioners.

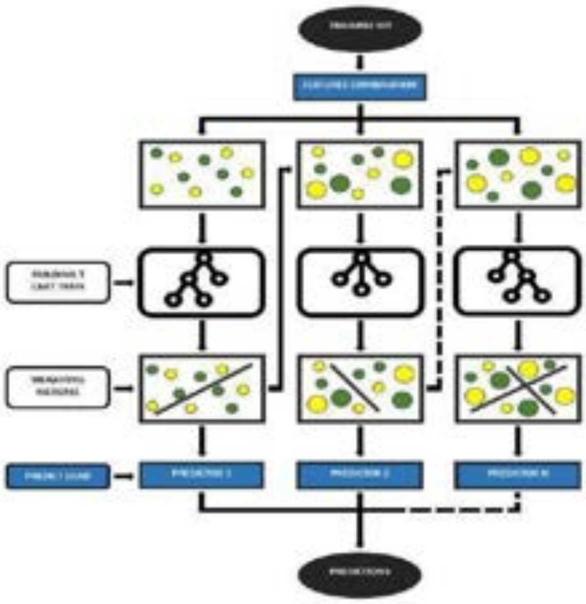


Figure 2: XGBoost algorithm

IV. IMPLEMENTATION

The steps involved in implementation are as follows:

i. Data pre processing
ii. Working of algorithm
iii. Database details

i. Data pre-processing:

Within the pre-processing module of the career guidance system, feature engineering techniques are implemented. This process commences with gathering varied datasets encompassing academic evaluations, skill assessments, and personal information from users. Subsequently, a meticulous data cleaning phase is conducted to address missing values, outliers, and inconsistencies, thereby preserving data integrity. Following this, feature engineering methods are employed to extract pertinent insights and discern patterns from the data.

This entails generating novel features utilizing domain knowledge and expertise, such as merging academic grades with extracurricular activities or converting skill assessment scores into categorical variables indicating proficiency levels. Furthermore, text mining and natural language processing methodologies are utilized to extract valuable insights from unstructured textual data, such as resumes or feedback. By augmenting the dataset with engineered features, the pre-processing module bolsters the predictive capability of the career guidance system, resulting in more precise and tailored recommendations for users' career trajectories. The detailed process involved in feature engineering is depicted in figure 3.

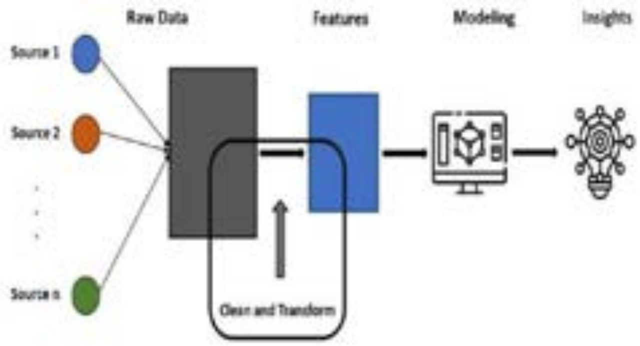


Figure 3:- feature engineering

ii. Working of algorithm

Data is gathered from diverse sources including student profiles, academic records, and skill assessments, followed by preprocessing to address missing values, encode categorical variables, and scale numerical features. Feature engineering ensues, involving the creation of pertinent features such as academic performance metrics and skill proficiency scores to enhance the dataset. Next, the XGBoost model is trained on the preprocessed data, meticulously adjusting hyperparameters and optimizing performance. Evaluation metrics are subsequently employed to gauge the model's performance, ensuring its reliability and ability to generalize.

Following validation, the trained model is seamlessly incorporated into the modules of the career guidance platform, spanning academic evaluation, cognitive decision support, career prediction, dynamic skill evaluation, and career global perspective. This integration harnesses the power of the XGBoost algorithm to generate tailored career recommendations grounded in users' profiles, assessments, and global career trends.

Career prediction is facilitated by leveraging the XGBoost algorithm within a comprehensive technical framework. Beginning with data preprocessing, the dataset undergoes cleaning and preparation, followed by feature engineering to extract pertinent insights. Subsequently, the XGBoost model is trained on the preprocessed data, with hyperparameters fine-tuned to optimize performance. Evaluation metrics validate the model's accuracy and generalization abilities. Integration of domain knowledge heightens the relevance of predictions, delivering personalized career recommendations based on user profiles, assessments, and global career trends. This technical infrastructure ensures precise career guidance, empowering users to make well-informed decisions about their future trajectories.

iii. Database details

The database infrastructure supporting the career guidance platform is composed of essential elements crucial for its smooth operation. Initially, the login and registration page database securely stores user credentials, ensuring regulated access to the platform. Subsequently, upon registration, the student details page database collects comprehensive information about individual users, encompassing personal details and academic backgrounds. Furthermore, the skill assessment module database logs data related to users' skill evaluations, laying the groundwork for personalized recommendations.

Within the career prediction module, the database houses pathways to diverse career options and recommended colleges nationwide, facilitating informed decision-making for users. The database architecture facilitates seamless integration among modules, ensuring uninterrupted data flow and effective communication between various components of the platform. This integrated database system ensures streamlined data management, enabling the platform to deliver tailored career guidance services efficiently. With strong data management capabilities, the platform can adjust to changing user requirements and deliver timely and pertinent career guidance solutions.

V. RESULT ANALYSIS

The enhanced accuracy of XGBoost in career prediction, in comparison to other machine learning algorithms, can be credited to various factors. Primarily, XGBoost's ensemble learning technique amalgamates multiple weak learners, often decision trees, to construct a more robust predictor. By leveraging the strengths of each individual model while offsetting their weaknesses, this ensemble approach improves the model's capacity to discern intricate relationships and patterns within the data, resulting in heightened prediction accuracy.

Furthermore, XGBoost integrates regularization methods to counteract overfitting, guaranteeing strong generalization capabilities to unseen data. Furthermore, its optimization algorithm adeptly manages extensive datasets and demonstrates exceptional scalability, allowing it to effectively collect insights from diverse and high-dimensional feature spaces.

TABLE I. ACCURACY OF DIFFERENT ALGORITHMS

S.No	Algorithm	Accuracy
1.	**XGBoost**	**89.5%**
2.	Random Forest	87.2%
3.	Decision Tree	83.8%
4.	Support Vector Machine	82.1%
5.	K-Nearest Neighbour	81.3%
6.	Logistic Regression	79.6%

Table 1 displays the accuracy levels of different machine learning algorithms in predicting career paths, with XGBoost leading with an accuracy of 89.5%.

Figure 4 visually represent this comparison in the form of a bar graph where each algorithm is represented by it's own bar on the vertical axis while the horizontal axis indicates the accuracy percentages.

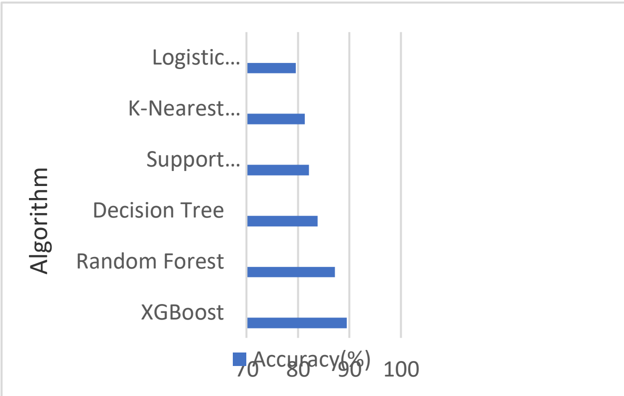


Figure 4: Accuracy of Algorithms

Table II provides a comparison of XGBoost algorithm with other algorithms based on adaptability.

TABLE II. ADAPTABILITY OF DIFFERENT ALGORITHMS

S.No	Algorithm	Adaptability
1.	**XGBoost**	**85%**
2.	Random Forest	80%
3.	Decision Tree	75%

4.	Support Vector Machine	75%
5.	K-Nearest Neighbor	70%
6.	Logistic Regression	70%

Figure 5 is the bar graph representation of the comparison based on adaptability.

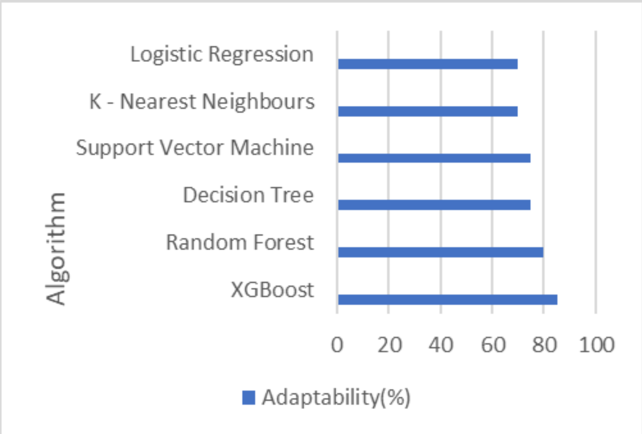


Figure 5: Adaptability of Algorithms

Therefore, given its superior accuracy and adaptability in predicting career paths, this algorithm has been used over the other algorithms for the implementation of the career guidance project.

VI. CONCLUSION AND FUTURE WORK

In conclusion, the creation of a holistic educational and career platform, incorporating elements such as login and registration capabilities, student profile administration, academic assessment, skill evaluation, cognitive decision support, career forecasting, dynamic skill assessment, global career outlook, and feedback mechanisms, presents considerable potential in steering users toward prosperous educational and professional paths. By integrating various functionalities, users are provided with personalized career suggestions, insights into international job markets, continuous skill enhancement prospects, and a feedback platform for continual enhancement. This holistic platform not only boosts user engagement and contentment but also enables individuals to make educated choices regarding their educational and professional journeys.

As of future work, improving the interpretability and transparency of the system through the utilization of explainable AI methods would cultivate user trust and support better-informed decision-making. In essence, forthcoming developments in technology and data science present significant opportunities for transforming career guidance systems, empowering individuals to navigate the continually changing professional terrain with clarity and assurance.

REFERENCES

[1] Vignesh S's, Shivani Priyanka C's, Shree Manju H's, Mythili K's "An Intelligent Career Guidance System using Machine Learning", 2021 7th International Conference on Advanced Computing & Communication Systems (ICACCS), 978-1-6654-0521-8/21/ 2021 IEEE.

[2] Andres F. Cruz's , Laura Orozco's, and Carolina Gonzales's. Data mining:" Intelligent web platform for vocational guidance" , 2019 International Conference on Virtual Reality and Visualization (ICVRV), DOI 10.1109/ICVRV47840.2019.00049 .

[3] John Britto's , Sagar Prabhu's , Abhishek Gawali's and Yogesh Jadhav's. "A Machine Learning Based Approach for Recommending Courses at Graduate Level", Second International Conference on Smart Systems and Inventive Technology (ICSSIT 2019) IEEE Xplore Part Number: CFP19P17-ART; ISBN:978-1-7281-2119-2.

[4] Dahanke Ajay's, Shinde Nilesh's, Dhagate Anirudh's, Shaikh Huzaif's, "An Intelligent Career Guidance System using Machine Learning' performance: A case study," International Journal of Information and Communication Technology Research, Mar 2022, pp. 140-146.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.

[5] Lei Wang's and Yang Bai's "Research on Career Guidance Course System based on Apriori Algorithm and Computer Big Data", 2022 International Conference on Computers, Information Processing and Advanced Education (CIPAE), 10 February 2023 IEEE, DOI:10.1109/CIPAE55637.2022.00036.

[6] Vijaya R. Kumbhar's, Manisha M. Maddel's, Yudhishthir Raut's "Smart Model For Career Guidance Using Hybrid Deep Learning Technique",2023 1st International Conference on Innovations in High Speed Communication and Signal Processing (IHCSP), 25 May 2023 IEEE, **DOI:** 10.1109/IHCSP56702.2023.10127152

[7] E K Subramanian's, and Ramachandran's." Student Career Guidance System for Recommendation of Relevant Course Selection ". International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6S4, April 2019

[8] Steffi F Andrade's, Nishita M Bhatankar's, Nadia M D'silva's and Prof. Prachi Patil's "Career Prediction using Support Vector Machine Algorithm" , 2019 JETIR June 2019, Volume 6, Issue 6.

[9] Lakshmi Prasanna's, DR.D.Haritha's "Smart Career Guidance and Recommendation System", IJEDR 2019 | Volume 7, Issue 3 | ISSN: 2321-9939.

[10] Kasem Seng's and Akram M. Zeki's "Career Guidance and Employment Management System" , 2014 3rd International Conference on Advanced Computer Science Applications and Technologies, 978-1-4799-1845-4/15 $31.00 © 2015 IEEE DOI 10.1109/ACSAT.2014.20.

Intelligent Diagnosis of Potato Leaf Diseases using Deep Learning

Gaurav Harbola
Computer Science Department, Graphic Era Hill University, Dehradun, India
gauravharbola0106@gmail.com

Mahesh Singh Rawat
Computer Science Department,, Graphic Era Hill University Dehradun, India
maheshsinghrawat@gmail.com

Amit Gupta
Computer Science Department Graphic Era Hill University, Dehradun, India
amitgupta7920@gmail.com

Richa Gupta
Computer Science Department Graphic Era Deemed to be University, Dehradun, India
richa1883gupta@gmail.com

Abstract**: Potato cultivation faces significant threats from plant diseases, such as early blight and late blight, which detrimentally impact yield and quality. Addressing this challenge, this study presents a novel Convolutional Neural Network (CNN) model designed for automated disease detection in potato plants. Leveraging advanced CNN architectures, particularly MobileNet, this study achieves a remarkable accuracy of 96.6% in disease classification tasks. This innovative approach eliminates the need for manual inspection processes, offering a scalable and efficient solution for early disease identification and mitigation. Finally, this study highlights the transformative potential of deep learning methodologies in enhancing agricultural productivity and disease management strategies.**

Keywords- Feature Extraction, InceptionV3, VGG16, MobileNet, DenseNet, Inception, fine-tuning.

I. INTRODUCTION

Within the agricultural landscape, the potato emerges as a majorly cultivated vegetable, constituting approximately 30% of agricultural produce in India, thereby securing a notable position as the fourth most cultivated crop worldwide, following maize, wheat, and rice. India, a formidable contender in potato cultivation, claims the second-largest global production output, exceeding 48.5 million tons annually. Uttar Pradesh commands a significant share, contributing over 30% to India's potato yield. Beyond culinary applications, potatoes demonstrate versatility in industrial domains, notably in starch extraction for enhancement. Nutritionally, they are esteemed for their potassium-rich composition, abundant vitamin content, and dietary fiber, rendering them instrumental in combating various health ailments, including hypercholesterolemia, hypertension, cardiovascular afflictions, and malignancies.

However, the potato's dominance is threatened by microbial adversaries such as bacteria and fungi, notably manifesting as late blight and early blight diseases. To mitigate these challenges, technological interventions, deep learning methodologies have emerged as pivotal tools for disease identification. This study presents a comprehensive framework, structured into sections outlining the contextual framework, prior research insights, dataset characteristics, model architecture complexities, and conclusive findings. Central to the research investigation is the development and validation of a robust deep learning model, precisely trained on a high-fidelity dataset, incorporating pre-trained VGG19 models with fine-tuned hyperparameters. The proposed approach also integrates validation and ensemble techniques to enhance the performance consistency.

Notwithstanding, disease detection in plants poses multifaceted challenges, characterized by diverse symptoms influenced by pathogenic agents, environmental stimuli, and genetic predispositions. Early disease manifestation often presents subtle symptoms, necessitating sensitive detection methodologies. Adverse environmental conditions, including lighting variations and background clutter, further complicate accurate disease diagnosis. Addressing these challenges mandates large-scale, high-quality datasets for training algorithms to differentiate between healthy and diseased plant specimens.
In response, the proposed model utilizes Convolutional Neural Network (CNN) architectures and evaluates the model to identify the most efficient framework for disease detection. This study focuses on obtaining enhanced accuracy rates, highlighting the efficacy of the proposed approach in enhancing disease detection within agricultural ecosystem.

II. LITERATURE REVIEW

Wang et al. [1] and fellow researchers collected a dataset consisting of 185 high-resolution images capturing instances of wheat stripe rust and grape powdery mildew for disease analysis. Es-saddy et al. [2] diverged from conventional methodologies by combining images obtained from digital cameras deployed across diverse farm settings and supplemented by online sources, fostering a comprehensive dataset representing varied environmental conditions. Mohanty et al. [3] curated an extensive dataset comprising 54,306 images categorized into 38 classes of both diseased and healthy plant leaves, ensuring meticulous representation akin to a meticulously assembled family album.

Rastogi et al. [4] adopted a new preprocessing strategy, resizing and cropping images to optimize the model input. Es-saddy et al. [2] applied image resizing and noise reduction techniques to enhance the image quality, ensuring clarity and precision in disease identification. Not merely content with image acquisition, Sannakki et al. [6] employed image enhancement methods similar to artistic transformations, instilling images with captivating visual effects to augment interpretability and aesthetic appeal. Innovative approaches integrating machine vision and fuzzy logic offer promising solutions for objective disease grading in agriculture, enhancing scalability and reliability in large-scale agricultural operations. Nonetheless, challenges encompassing technical complexities, validation across diverse contexts, financial constraints, and ethical considerations necessitate advanced analysis and strategic mitigation.

Ramesh et al. [9] outlined image characteristics through histogram of oriented gradients (HOG), facilitating spatial representation and pattern recognition. Furthermore, researchers, Abbas et al. [10], Chen et al. [10], and Akshai et al. [10] introduced advanced Convolutional Neural Network (CNN) architectures like DenseNet, LFMCNAPS, and a diverse ensemble of CNNs to differentiate a spectrum of diseases across various plant species, similar to deploying adept detectives to identify complex botanical anonymities.

Additionally, Kibriya et al. [11] and Sujatha et al. [12] undertook a multifaceted approach integrating traditional machine learning with deep learning methodologies to comprehensively classify plant diseases. Ashwin Kumar et al. [13] adopted a mobile network-based CNN for on-the-go disease identification, epitomizing the fusion of technological innovation and practical applicability in agricultural settings.

While support vector machine (SVM) classifiers exhibit promise in automated plant leaf disease recognition, intrinsic challenges including implementation complexity, data quality assurance, interpretability, and cross-domain validation necessitate careful consideration and strategic mitigation strategies for effective deployment in agricultural ecosystems.

III. DATASET DESCRIPTION

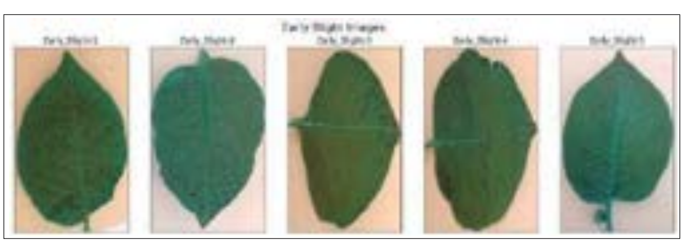


Figure 1: Sample image dataset

The datatse is collected from kaggle, totally 3000 images including both healthy and unhealthy potato leaves are considered. Sample healthy images from the dataset are shownin figure 1.
Sample diseased leaves images from the collected dataset is depicted in figure 2.

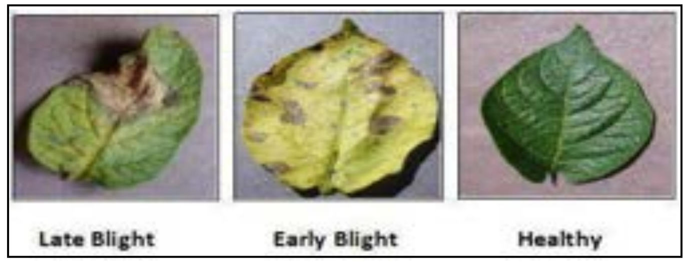


Figure 2: Early Bright Images

III. PROPOSED METHODOLOGY

In this project there are multiple stages of research as can be seen from fig. 1. In the form of research framework there are four stages which are as follows:

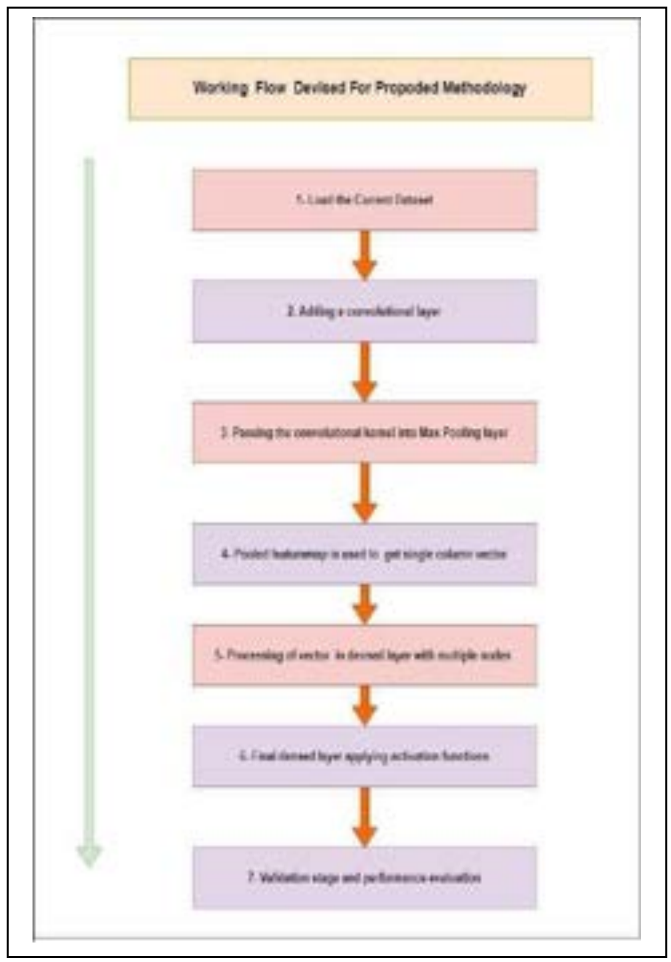


Figure 3: Flow of a Machine Learning Problem

CNN Architecture:

Convolutional Neural Network is a subset of machine learning.

Convolution layers, here are multiple 3x3 convolutional layers in the network, which help the network learn special patterns and features from the input images which will help in identifying these features/patterns in the input images. These convolutional layers are responsible for capturing features and as the network of layers deepens the capturing of features transition from low to high.

Pooling Layers:
Pooling layers play a major role in reducing the spatial dimensions from feature maps. Here, we have used a two-dimenisonal filtering method in the pooling layer, the maxpooling operation incporated in the proposed model is shown in fig. 4.

Figure 4 : Max Pooling

Fully Connected Layers:
This layer is responsible for integrating the learned features from earlier layers and enable the network of layers to make comprehensive and accurate predictions. The fully connected layer enables our network to perform high level reasoning and classification tasks.
Activation Functions: This study uses SoftMax activation function where multiple number of signmoids can mention the data probability to solve the multi-classification problem.

Next, this study has considered VGG16, MobileNet, DenseNet, InseptionV3 and ResNet50 architectures.

VGG16: Architectural wise VGG16, distinguished by its 16 weighted layer, features a consistent layer structure as sjhown in figure 5.


Figure-5 VGG16 Architecture

MobileNet: Mobile Net optimized for mobile vision applications, stands out in object detection and fine-grained classifications due to its efficiency and lightweight design. It has 2 layers i.e. depth wise and pointwise layers. The depth-wise convolution processes each input channel independently, significantly reducing computational costs.[15]

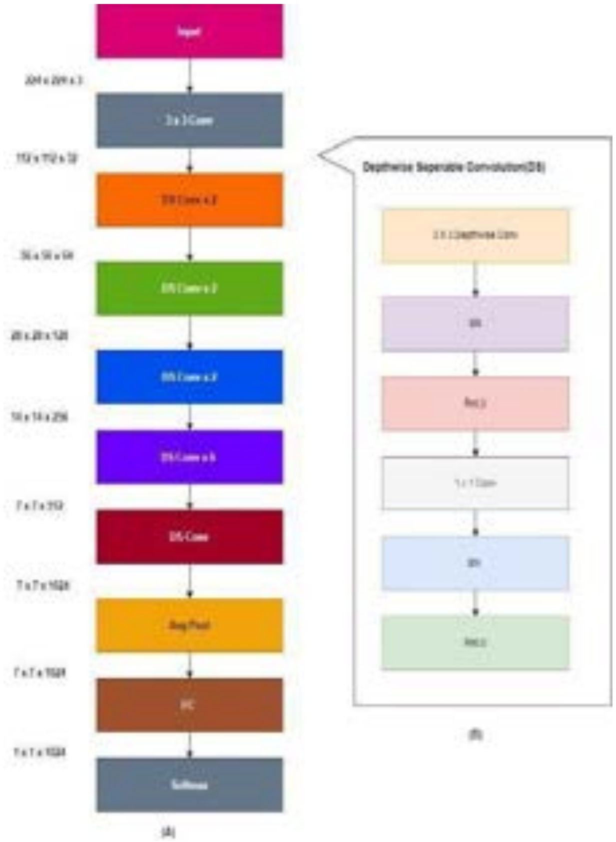


Figure 6. Mobile Ne

DenseNet: Each layer has a 1x1 and a 2x2 layer. Basic convolution layer with 64 filters of size 7X7 with stride of 2. The DenseNet121 architecture is shown in figure 7.

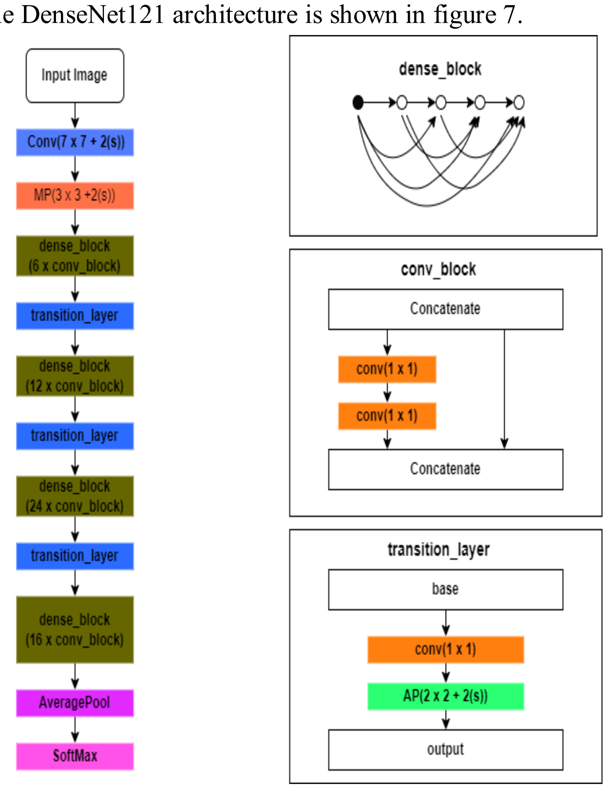


Figure-7 DenseNet121 Architecture

InceptionV3: Inception-v3, a 48-layer deep convolutional neural network, offers pre-trained versions capable of classifying images into 1000 categories. Notably, its architecture evolved from Inception V1 to address certain drawbacks. Issues with Inception V1 included occasional information loss due to 5x5 convolutions, which prompted a shift to factorization, breaking down 3x3 convolutions into 1x3 and 3x1 for computational efficiency[16].

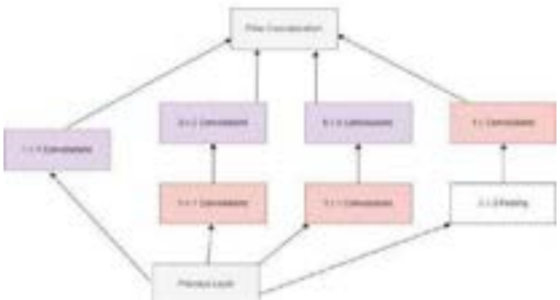


Figure-8 InceptionV3 Compact Architecture

Resnet 50: The convolutional block, akin to the identity block, introduces a 1x1 convolutional layer to reduce filter numbers before the 3x3 convolutional layer. Skip connections, synonymous with residual connections, play a vital role in

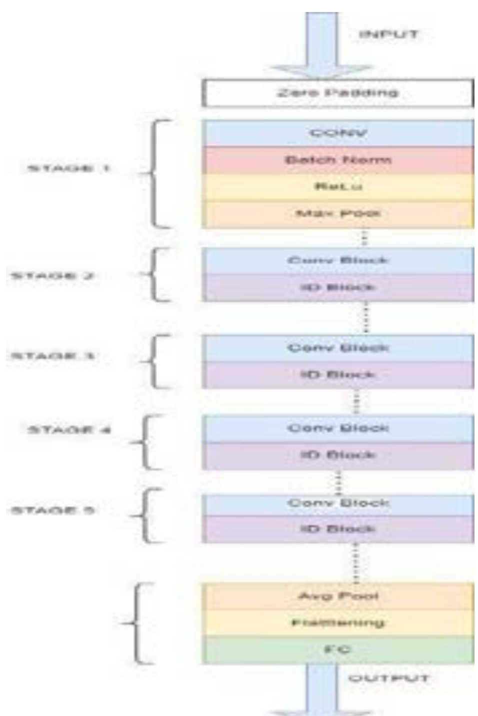


Figure-9 ResNet 50 Architecture

ResNet50. In ResNet50 , skip connections are integrated into the identity block and convolutional block. The identity block processes input through convolutional layers and incorporates the input back into the output. Similarly, the convolutional block utilizes a 1x1 convolutional layer to reduce filter numbers before the 3x3 convolutional layer, ultimately adding the input back to the output.

IV. RESULTS & DISCUSSION

The proposed dataset is splitted into 70% as training data and 30% as test data. Various pretrained models were used like VGG16, ResNet 50, InceptionV3, DenseNet121 and MobileNet. Among these MobileNet has leveraged the best performance with an accuracy of 96.6%.

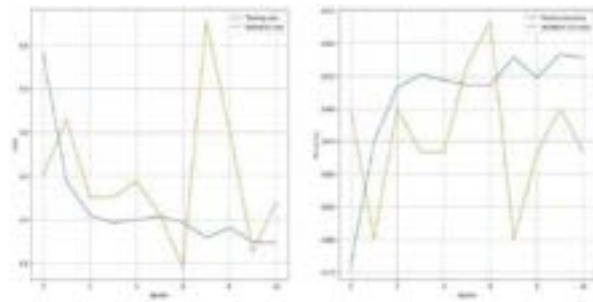
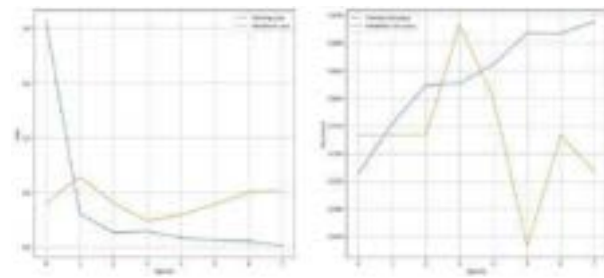
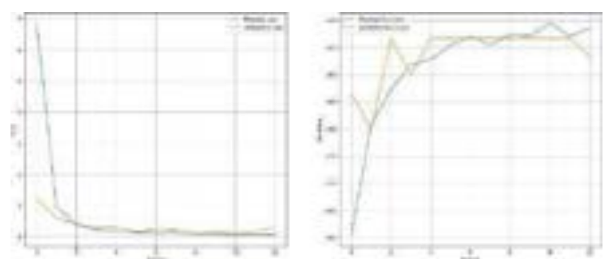
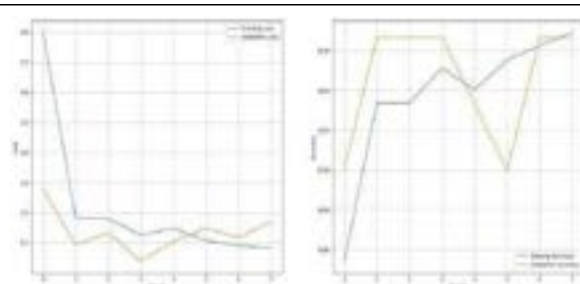
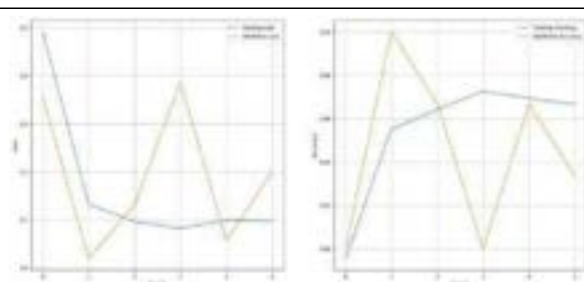
As depicted in Table I, we have found the precision, recall, and f1 score of all the models. Among which mobile net had highest. Precision tells the proportion of correctly classified positive instances over all the instances that are classified as positive. Measures the ability of the classifier to avoid false positives. It came out to be 0.95 for late blight and 1.0 for early blight. A good indicator of the performance of a classifier when the number of positive and negative instances is unbalanced. Came out to be 0.97 for late blight and 0.96 for early blight.

In the experimental segment of the proposed plant disease detection system, different CNN models are used and their performance are analysed. The analysis reveal that MobileNet leverages a comparatively higher performance. This study has analyzed the outcomes acquired from every version as shown in Table II to provide insights on its effectiveness and suitability in disease detection process. A detailed comparison of training and validations accuracies are pesented in figure 10.

TABLE I: PERFORMANCE ANALYSIS

Models	Late/Early Blight	Precision	Recall	F1-Score	Accuracy
VGG16	LATE	0.94	0.94	0.94	0.933
	EARLY	0.92	0.92	0.92	
RESNET 50	LATE	0.57	0.68	0.62	0.466
	EARLY	0.14	0.09	0.11	
INCEPTION V3	LATE	1.0	0.94	0.97	0.942
	EALRY	0.93	1.0	0.96	
DENSENET	LATE	0.94	1.0	0.97	0.926
	EARLY	1.0	0.92	0.96	
MOBILE NET	LATE	0.95	1.0	0.97	0.966
	EARLY	1.0	0.92	0.96	

TABLE II DETAILED PERFORMANCE ANALYSIS

Models	Loss	Validation Loss	Training Accuracy	Validation Accuracy	Testing Accuracy	Graph
VGG16	0.1475	0.2374	0.933	0.8631	0.933	
ResNet50	0.6056	0.8074	0.6628	0.535	0.466	
MobileNet	0.053	0.2864	0.9892	0.9421	0.942	
InceptionV3	0.0818	0.1700	0.9678	0.9621	0.926	
DenseNet	0.0971	0.1995	0.9623	0.9384	0.966	

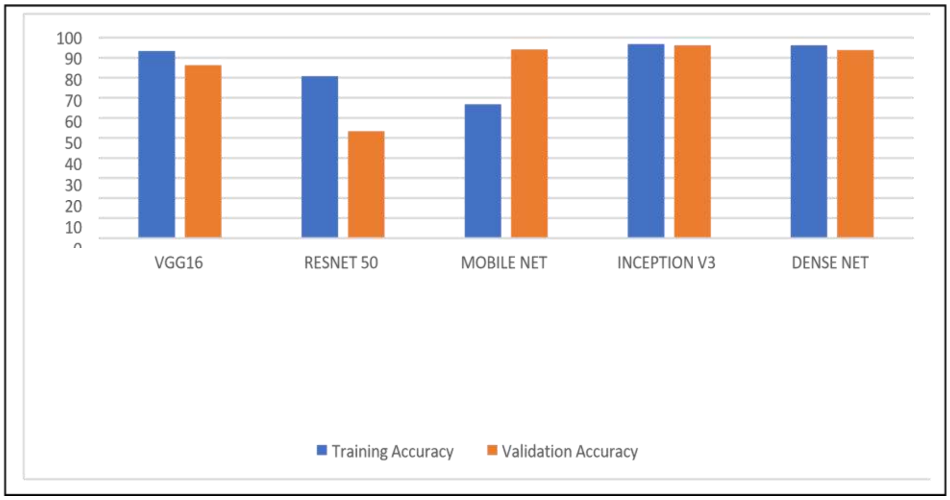


Figure 10 Comparison of Training accuracies and validations accuracies

V. CONCLUSION

This study conclude that among the models VGG16, ResNet, MobileNet, DenseNet, InceptionV3, the MobileNet performs exceptionally well by leveraging an accuracy of 96.6%. The success of mobilenet can be attributed to its efficient architecture, which strikes a balance between model complexity and computational efficiency. The model demonstrated a superior accuracy, show casing its potential as a robust solution for a real-time deployment in agricultural settings. The utilization of transfer learning further enhanced the model's ability to recognize relevant features leveraging the knowledge of pre-trained weights from ImageNet. Further research and development in this area can build upon insights gained from the objects.

REFERENCES

[1] S. M. Kulkarni and G. Sundari, "A Framework for Brain Tumor Segmentation and Classification using Deep Learning Algorithm," IJACSA, vol. 11, no. 8, 2020, doi: 10.14569/IJACSA.2020.0110848

[2] J. Sikder, U. K. Das, and R. J. Chakma, "Supervised Learning-based Cancer Detection," IJACSA, vol. 12, no. 5, 2021, doi: 10.14569/IJACSA.2021.01205101.

[3] H. Moujahid, B. Cherradi, and L. Bahatti, "`Convolutional Neural Networks for Multimodal Brain MRI Images Segmentation: A Comparative Study," in Smart Applications and Data Analysis, vol. 1207,

[4] M. Hamlich, L. Bellatreche, A. Mondal, and C. Ordonez, Eds. Cham: Springer International Publishing, 2020, pp. 329– 338. doi: 10.1007/978-3-030-45183-7_25.

[5] N. A. Ali, A. E. abbassi, and B. Cherradi, "The performances of iterative type-2 fuzzy C-mean on GPU for image segmentation," J Supercomput, Jun. 2021, doi: 10.1007/s11227-021-03928-9.

[6] Febrianto, D. C., I. Soesanti, and H. A. Nugroho. "Convolutional neural network for brain tumor detection." In IOP Conference Series: Materials Science and Engineering, vol. 771, no. 1, p. 012031. IOP Publishing, 2020.

[7] Khairandish, Mohammad Omid, R. Gurta, and Meenakshi Sharma. "A hybrid model of faster R-CNN and SVM for tumor detection and classification of MRI brain images." Int. J. Mech. Prod. Eng. Res. Dev 10, no. 3 (2020): 68636876.

[8] W. Ayadi, W. Elhamzi, I. Charfi, and M. Atri, "Deep CNN for Brain Tumor Classification," Neural Process Lett, vol. 53, no. 1, pp. 671–700, Feb. 2021, doi: 10.1007/s11063- 020-10398-2.

[9] O. El Gannour et al., "Concatenation of Pre-Trained Convolutional Neural Networks for Enhanced COVID-19 Screening Using Transfer Learning Technique," Electronics, vol. 11, no. 1, p. 103, Dec. 2021, doi: 10.3390/electronics11010103.

[10] S. Hamida, O. El Gannour, B. Cherradi, A. Raihani, H. Moujahid, and H. Ouajji, "A Novel COVID-19 Diagnosis Support System Using the Stacking

Approach and Transfer Learning Technique on Chest X-Ray Images," Journal of Healthcare Engineering, vol. 2021, pp. 1–17, Nov. 2021, doi: 10.1155/2021/9437538.

[11] O. El Gannour, S. Hamida, B. Cherradi, A. Raihani, and H. Moujahid, "Performance Evaluation of Transfer Learning Technique for Automatic Detection of Patients with COVID-19 on X-Ray Images," in 2020

[12] IEEE 2nd International Conference on Electronics, Control, Optimization and Computer Science (ICECOCS), Kenitra, Morocco, Dec. 2020, pp.1–6.doi: 10.1109/ICECOCS50124.2020.9314458.

[13] S. Hamida, O. E. Gannour, B. Cherradi, H. Ouajji, and A. Raihani, "Optimization of Machine Learning Algorithms Hyper-Parameters for Improving the Prediction of Patients Infected with COVID-19," in 2020 IEEE

[14] Amit Gupta, Richa Gupta & Navin Garg (2021) An efficient approach for classifying chest X-ray images using different embedder with different activation functions in CNN, Journal of Interdisciplinary Mathematics, 24:2, 285-297, DOI: 10.1080/09720502.2020.1838060

[15] Gupta, R., Tripathi, V., & Gupta, A. (2021, December). An Efficient Model for Detection and Classification of Internal Eye Diseases using Deep Learning. In 2021 International Conference on Computational Performance Evaluation (ComPE) (pp. 045-053). IEEE

[16] Gupta, R., Tripathi, V., Gupta, A.(2021).Diabetic Retinopathy detection using an Efficient Artificial Intelligence Method. ICAES2021 International conference on advancement in Engineering and Sciences July 2021

Biometric Authentication System for Two Wheeler

Anita Dombale
Associate
Professor
BRACT'S Vishwakarma Institute Of Technology
Pune, India
anita.dombale@vit.edu

Chetan Thakare
Department of Engineering Sciences and Humanities
BRACT'S Vishwakarma Institute Of Technology
Pune,India
chetan.thakare23@vit.edu

Mahesh Thorve
Department of Engineering Sciences and Humanities
BRACT'S Vishwakarma Institute Of Technology
Pune, India
mahesh.thorve23@vit.edu

Vedant Suryavanshi
Department of Engineering Sciences and Humanities
BRACT'S Vishwakarma Institute Of Technology
Pune, India
vedant.suryavanshi23@vit.edu

Anjali Ubale
Department of Engineering Sciences and Humanities
BRACT'S Vishwakarma Institute Of Technology
Pune, India
anjali.ubale23@vit.edu

Tanishq Chavan
Department of Engineering Sciences and Humanities
BRACT'S Vishwakarma Institute Of Technology
Pune, India
bhushanrao.tanishq23@vit.edu

Swayam zagade
Department of Engineering Sciences and Humanities
BRACT'S Vishwakarma Institute Of Technology
Pune, India
swayam.zagade23@vit.edu

***Abstract*— This project is created to fulfil only one purpose that is to create an authentication system based on biometrics, the most common type of biometrics is fingerprint. It is designed to reduce and prevent the theft of two wheelers. This also provides the user a keyless interface increasing its accessibility. The ignition system is replaced with fingerprint start system. The fingerprints are identified and stored in templates of binary codes which are acquired on the basis of special features of fingerprint. The project consists of the AVR microcontroller Arduino uno r3, the fingerprint module R307, and the GSM Module sim800L. The fingerprint Module R307 can store up to 127 fingerprints. once the fingerprints are stored the r307 interacts with the Arduino and process that information if the fingerprint matches with the user, then the Arduino controller sends signal to relay module to start the ignition system. f the fingerprint does not matches the Arduino sends a message to the user as "someone has tried to access your bike" via GSM module. The proposed model mimics the self-start of bike with a DC motor. Additionally, access to the two-wheeler is available via a keypad, which has a four-digit password, allowing users' family or relatives to access the model other than the user.**

Keywords— Arduino UNO, R307, GSM, Arduino IDE, Biometrics, keypad.

I. INTRODUCTION

In today's world, there is a rise in concern related to theft. The world is advancing with many technologies that can be used to reduce theft-related concerns. Our project aims to address one such problem: reducing the theft of keys for two-wheelers. Introducing the concept of keyless two-wheelers and installing a high-tech security systemis the goal of this research project.

Knowledge-based, token-based, and biometric-based are the three primary categories of person identification techniques. Password and pattern-based security systems that the user configures are examples of knowledge-based strategy. Token-based techniques require document-based proofs like Aadhar card, driver's license, PAN card, etc. Biometric systems are based on a person's unique impressions, such as fingerprint impressions and face recognition. Among the other two systems, knowledge and token-based security systems, the biometric system was found to be more reliable and secure. Thus, the use of biometric systems in security has increased as it reduces the chances of theft occurrences.

The term "BIOMETRY" from the 20th century is where "BIOMETRICS" originates. Facial, voice, fingerprint, eye (iris), and other biometric recognition systems are examples of biometric recognition systems. Since every person's fingerprint is different, fingerprint technology is the most frequently utilized and most reliable technology available.

Using fingerprint biometrics is simple and easy. The key elements of fingerprint biometrics are identification and authentication. Fingerprint authentication or recognition involves combining multiple fingerprints through an automated process. Authentication establishes the identity of the user and verifies the validity and reliability of the user. The Arduino coordinates the interactions between the several parts of the system, serving as its brain. Among its functions are:

The R307 fingerprint sensor is connected to the Arduino, which then reads the fingerprint data. It oversees the process of authentication by comparing the scanned fingerprint to pre-stored templates. To supplement the biometric authentication, the Arduino communicates with the keypad to accept extra user input, like a PIN code. The Arduino decides which authentication to use based on the combined biometric and PIN inputs. The Arduino grants access to the two-wheeler if the fingerprint and PIN match the credentials that are saved.

Control of GSM Module: The Arduino may operate the GSM module in response to certain events (such unwanted access attempts) or successful authentication.

II. LITERATURE SURVEY

The concept of the Biometric Authentication System is designed to reduce thefts of two wheelers where the start function of two wheeler is replaced by a fingerprint sensor by Vaishnavi Khadasane, Mrunalini Desai, Devashree Khatavakar, and Shruti Lad [1]. The Advanced Biometric Authentication System in a two-wheeler displays the speed and distance traveled by a vehicle Iswarya G, M. Baranidharan, Bagavathi Shivakumar, C., and R. Rajaprabha drive the relay from the output of the ULN2803 driver chip [2]. Based on the concept of observable physiological or behavioral traits, biometrics is an automated method of identity verification or identification developed by Vaclav Matyas and Zdenek Riha [3]. An Arduino-based project that implements a fingerprint-based anti-theft system for two-wheelers addresses the growing worry of two-wheeler theft by implementing a system by Mrs. Maalini Dharmaraj [4]. The system authenticates vehicle users. By highlighting the growing weaknesses in conventional security measures, this research by R. H. Laskar, F. A. Talukdar, and Krishna Dharavath [5] shows how biometric systems have become more popular for users.

The focus of this research work is on a biometric fingerprint system, chosen over other alternatives for its tight security features. Employing a fingerprint-based biometric system proves to be both convenient and safe. The project makes use of a fingerprint module with substantial fingerprint storage in its database, facilitated by both software and hardware systems for its functioning with the fundamental idea being that the vehicle starts only when a correct fingerprint is obtained. If there is no match in the database, the vehicle remains inactive. Interface connections are established through the PC parallel port, known for its cost-effectiveness and widespread use in various projects. The entire system code is scripted in Embedded C language via software. This study shows how a fingerprint based system is used for enhancing the security. It is designed for seamless integration with contemporary embedded systems for fingerprint authentication, which is achieved through a suitable sensor and microcontroller. The system revolves around fingerprint authentication, ensuring its relevance to the latest technological advancements.

III. METHODOLOGY

The block diagram of communication of different components can be seen in fig 1

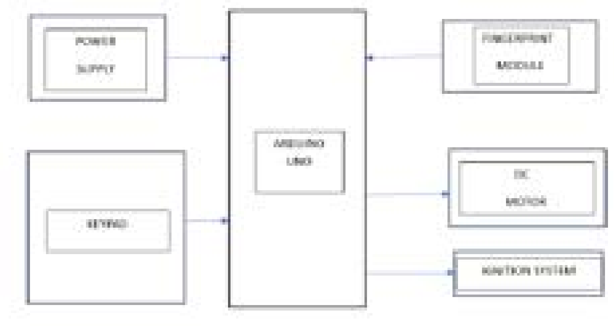


Fig 1. Block Diagram of the System

A. Hardware Components

Arduino uno**:** A microcontroller to process the information taken as input from the sensors and provides a certain output.
Fingerprint module (R307): It transmits the user's biometric data to the microcontroller. In order to provide security in fingerprint detection and verification, one type of fingerprint sensor module utilized in biometrics is the R307. These gadgets are mostly found in safes, which require powerful DSP chips for calculations, image rendering, feature detection, and searching. They use TTL serial to connect to any microcontroller and send data packets for hashing, searching, photo retrieval, and print detection. New finger enrolment can be immediately stored in the on-board flash memory.

Keypad**:** It sends the enrolled data from the user to the microcontroller.
ATmega328p microcontroller: It processes all the data it receives and gives necessary and corresponding output to other components.
Power Supply: It gives a regulated 5V to the microcontroller.
DC motor and ignition system: It represents the start of the vehicle.
GSM: It notifies the user by phone of important messages, including "invalid access." The SIM800L is a tiny cellular module that supports voice calls, SMS sending and receiving, and GPRS transmission. This module is an ideal choice for any project that needs long-range communication because it is affordable, small in size, and supports quad-band frequencies. It is compatible with GSM 850MHz, PCS 1900MHz, EGSM 900MHz, and DCS 1800MHz frequencies. A computer and a GSM-GPRS system can communicate thanks to the GSM module. The Worldwide Mobile Communication System

B. Software Components

For this project, the embedded C language is used for programming.
Fingerprint algorithm: To register the user's fingerprint and confirm the owner's.

The messaging mechanism sends a message to the bike's owner if the fingerprint does not match the registered fingerprint.
The keypad algorithm is to use a password to unlock the bike.
The two-wheeler will start when the fingerprint matches or the password entered is correct, and it will stop using the same process. This is known as the ignition algorithm. The actual connections and circuit diagram is shown in fig 2 and 3 respectively.

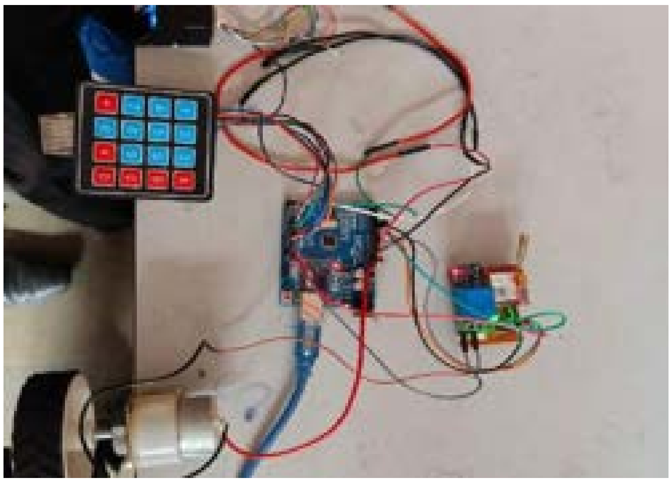


Fig 2 Actual Connections of the System

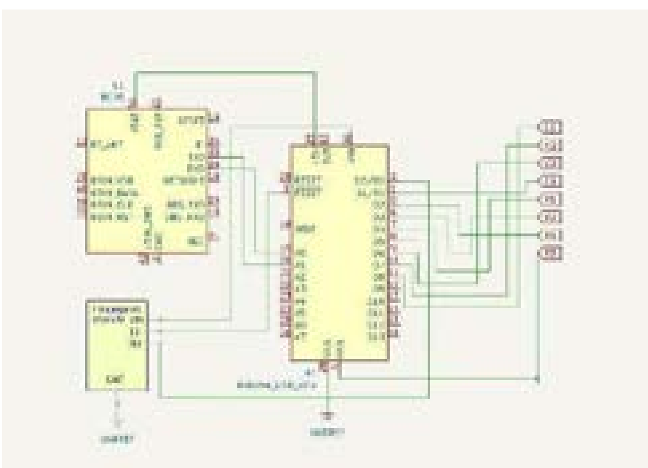


Fig 3 Circuit Diagram of the System

IV. RESULTS AND DISCUSSION

In the implementation stage, the first step involves enrolling the user's fingerprint for identification. Following successful enrollment, the next step is authentication, where the registered fingerprint is matched. Upon a successful match, the relay module is activated, initiating the vehicle's start.

Case 1
Once the user enrolls their fingerprint and its biometrics are stored, they can access the vehicle by matching their fingerprint.

Case 2
If someone other than the user attempts to access the vehicle after the user has registered their fingerprint, the user receives a message of intrusion via the GSM module. Demo can be seen in fig 4

Fig 4 Message Mechanism

Case 3
After the user registers their fingerprint and wishes to grant access to someone they know, they can provide a password to that person. The designated person can then access the vehicle by entering the password on the keypad. In the event of an incorrect password entry, the user will receive a notification.

The fingerprint module serves the crucial function of ensuring secure ignition and authentication in two-wheelers, effectively deterring theft. The microcontroller oversees the entire sequence, triggering the ignition system of the vehicle simultaneously, and sending a control signal to the relay driver. While the primary goal of the project is successfully achieved, a DC motor is employed to illustrate the dummy model of the two-wheeler's fuel ignition system. In cases of unauthorized access, an SMS is promptly sent through the GSM module.

V. CONCLUSION & FUTURE SCOPE

Improving vehicle security is the main reason two-wheelers have fingerprint sensors installed. This technique is useful for providing security and convenience of use since it uses the fingerprint as a strong biometric identification. Adding fingerprint recognition to cars not only makes them more secure, but it also discourages theft, which lowers the number of theft instances dramatically. In the future, we envision incorporating several modifications to the system to further enhance its functionality. One such addition includes integrating a fuel sensor, ensuring that the vehicle starts only when there is a sufficient fuel level. The inclusion of GPS technology will provide real-time location tracking for the vehicle. Moreover, the installation of cameras serves a dual purpose – capturing images of unauthorized access for security purposes and functioning as a dashcam. To prevent fuel theft when the vehicle is inactive, a solenoid valve can be employed. Additionally, we plan to implement a pollution monitoring system using gas sensors. All these features will be seamlessly integrated into a dedicated mobile application for comprehensive control and monitoring.

ACKNOWLEDGEMENT

I would like to express my sincere appreciation to our H.O.D. Prof Dr. Chandrashekhar Mahajan and our EDAI 1 Guide Prof. Mrs. Anita Dombale, for their invaluable guidance and support throughout this research project. Their expertise and encouragement have been instrumental in its completion. Finally, I extend my thanks to Vishwakarma Institute of Technology for providing the necessary resources.

REFERENCES

[1] Mrs. Maalini Dharmaraj, "Fingerprint-based Anti-Theft for Two Wheelers: Authentication of Vehicle Users," Special Issue - 2020 (IJERT).

[2] Iswarya G, "Advanced Biometric Authentication System in Two Wheeler," (IJESIT) Volume 3, Issue 1, January 2014.

[3] James Waymen, "An Introduction To Biometric Authentication Systems," IJTRA Special Issue 40 (KCCEMSR) (March 2016), pp. 42-45.

[4] Vaclav Matyas, "Biometric Authentication – Security And Usability Czech Republic," Advanced Communication And Multimedia Security.

[5] Krishna Dharavath, "Study on Biometric Authentication Systems, Challenges and Future Trends: A Review" (IEEE) 2013.

[6] Sukeerti Singh and Ayushi Mhalan, "Vehicle Theft Alert System using GSM," Int. Journal of Engineering Science and Technology (IJEST), May 2013.

[7] Rahil Khan, "Biometric Scanning Based on Vehicle Ignition System," International Journal of Engineering Science and Computing, March 2017.

[8] Abid Khan, Ravi Mishra, "GPS – GSM Based Tracking System," International Journal of Engineering Trends and Technology - Volume 3 Issue 2, 2012.

[9] K. S. Tamilselvan, G. Murugesan, and S. Sasikumar, "Design and Implementation of Biometric Based Smart Anti-Theft Bike Protection System," 2018 International Conference on Intelligent Computing and Communication for Smart World.

[10] S. K. Das and A. K. Das, "Biometric Authentication and GPS-based Anti-theft System for Electric Vehicles," Journal of Intelligent Transportation Systems: Technology Planning and Operations, vol. 22, no. 5, pp. 404-415, May 2018.

[11] P. Indyk and N. Thaper, "Fast Image Retrieval via Embeddings", In 3rd Workshop on Statistical and computational Theories of Vision, Nice, France, 2003.

[12] K. K. Paliwal and B. S. Atal, 'Frequency related representation of speech,' in Proc. EUROSPEECH, p.p. 65-68 Sep. (2003).

[13] Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[14] Lenman, S. , Bretzner, L. , Thuresson, B. (2002): Computer vision based hand gesture interfaces for human-computer interaction. Technical Report TRITA-NA-D0209, 2002, CID-172, Royal Institute of Technology, Sweden.

[15] Y. Rubner and C. Tomasi. Perceptual Metrics for Image Database Navigation. Kluwer Academic Publishers, Boston, MA, 2001.

[16] Y. Rubner, C. Tomasi, and L. J. Guibas. "The Earth Mover's Distance as a Metric for Image Retrieval", International Journal of Computer Vision, 40(2):99-121, 2000.

[17] S. Cohen, L. Guibas. "The Earth Mover's Distance under Transformation Sets", IEEE International Conference on Computer Vision, II: 1076-1083, 1999.

Design and Implementation of a Type-1 Fuzzy Controller Driven IoT-Integrated Automated Plant Watering System for Gardening

Dr. Ravindranath C.

Assistant Professor, Department of Civil Engineering, Dayananda Sagar College of Engineering, Kumaraswamy Layout, Bangalore-560111, Karnataka

Dr. Pavithra G.

Associate Professor, Electronics & Communication Engg. Dept., Dayananda Sagar College of Engineering, Kumaraswamy Layout, Bangalore-560111

Dr. Kavana Salimath

Assistant Professor, Department of Electrical & Electronics Engineering, Dayananda Sagar College of Engineering, Kumaraswamy Layout, Bangalore-560111

Dr. Sandeep K.V.

Assistant Professor, Department of Electronics & Telecommunication Engineering, Dayananda Sagar College of Engineering, Kumaraswamy Layout, Bangalore-560111, Karnataka

Dr. Sindhu Sree M.

Assistant Professor, Electronics & Communication Engg. Dept., Dayananda Sagar College of Engineering, Kumaraswamy Layout, Bangalore-560111, Karnataka

Dr. T.C.Manjunath

Professor & Head, Electronics & Communication Engineering Department, Dayananda Sagar College of Engineering, Kumaraswamy Layout, Bangalore-560111, Karnataka

Corresponding author mail id (Dr. Manjunath, PhD – IIT Bombay, Sr. Member IEEE) : tcmanju@iitbombay.org

***Abstract*— The work presented in this research article enables a significant advancement in gardening methodologies, leveraging an Internet of Things (IoT) - enabled automated plant watering system that underscores the paramount importance of efficiency and sustainability. The developed system, orchestrated by an 8-bit ATmega328P microcontroller, integrates resistive soil moisture sensor and humidity and temperature sensors (DHT11) to furnish real-time information. The watering technique of the developed system is further refined through the integration of a typed-1 fuzzy based logical based controllers, which is utilized to regulate its operation. With the integration of a Wi- Fi module, users gain the ability to monitor water distribution remotely through a smartphone application. The utilization of Thing Speak cloud technology streamlines data transfer and analysis, thereby enhancing the overall performance of the system on a full scale. The developed automated plant watering system not only conserves water and fosters optimal plant growth but also aligns seamlessly with advanced technological trends in the farming sector. It exemplifies a harmonious fusion of ecological awareness and cutting-edge IoT developments.**

Keywords—Internet of Things (IoT), Plant Watering System, Typed-1 fuzzy based logic controllers, Android Application, Matlab.

I. INTRODUCTION

Modern agriculture has become dependent on smart gardening practices, such as irrigation / plant watering systems driven by artificial intelligence (AI). Statistics that demonstrate increased productivity in cultivation through automation prove that these innovations are the need of the hour. In order to reduce water losses and maximize water utilization in agriculture, efficient water use becomes essential. Projections of the population exceeding 10 billion by 2050, coupled with a projected 60% rise in water consumption, are concerning. Consequently, recent data emphasizes the need to reassess conventional gardening practices to ensure effective water management. This assertion is substantiated by reputable sources such as '*FAO Aquastat*' and "The Future of Water 2022" authored by Jippe Hoogeveen, a senior officer in the FAO Land and Water Divisions.

Moreover, UNESCO's findings, as interpreted by the World Water Assessment Program (WWAP), shed light on the financial significance of water management. A staggering investment of approximately USD 115 billion per year across 150 low and middle-income countries could open up global pathways for clean drinking water and sanitation, yielding substantial returns. Since cultivation depletes an astounding 70% of universal freshwater resources, the burden of food safety becomes predominant. Estimates demonstrate a 60% surge in widespread demand for food and agricultural goods by 2050, alongside environmental deterioration related to water usage and food production. This warrants unprecedented progress in efficiently assessing water resources accompanied by instant monetary, subcultural, and atmospheric benefits. Realizing water's synthesized advantages, embracing food protection, climate crisis reduction, and countryside economic regeneration, is pivotal in leveling the impending universal food dilemma [1].

This paper prompts an inventive Internet of Things (IoT)-powered plant watering system to direct the looming challenges of global water shortage and the accelerating need for food production. The emergence and execution of the IoT- enabled foundation comprises a detailed procedure. Sensors measuring soil moisture, temperature, and humidity are deliberately positioned in a small plant environment to outsource real-time data. A central ATmega328P microcontroller digitizes this information, making use of a

Typed-1 Fuzzy based logical depended controllers to make strategic decisions for ideal water allocation to the plants. This approach not only provides crop watering by optimum usage of resources but also ensures environmental sustainability by decreasing the wastage of water. The proposed study stresses leveraging data interpretation to contribute usable knowledge, besides equipping gardeners with data to schedule adaptive water distribution schedules and plant- specific requirements. The facilitation of remote water conservation abilities through their ESP-8266 Wireless-Fidelity modules strengthens their system's functionality, in addition to allowing users to monitor the watering process from anywhere at any time via a smartphone application. This expansive attempt takes giant strides towards a technologically and environmentally conscious tomorrow for agricultural water management.

II. Contributions of the Paper

The paper makes the following contributions:

- Development of an IoT-enabled, cost-effective, and user-friendly plant watering systemtailored for farmers.
- Designing of a Type-1 Fuzzy Controller to regulate plant watering by controlling the actions of the watering pump.
- Implementation of an Android Application for monitoring various environmental parameters within the garden, including temperature, humidity, and soil moisture.

The paper unfolds with an insightful introduction in Section I, setting the stage for a comprehensive exploration of smart water management in agriculture. Section II meticulously reviews pertinent literature and related work, establishing a strong foundation for the proposed IoT-powered system. Moving forward, Section III intricately dissects the system architecture, delineating key components, functionality, the working principle, and a detailed Circuit Diagram section. Section IV delves into the designing of the developed Android Application. Section V illustrates the designing process of the type-1 fuzzy controller. Meanwhile, Section VI meticulously explores various facets of the designed automated IoT-based plant watering system, shedding light on its diverse aspects and functionalities. In the pinnacle of Section VII, this study concludes by succinctly summarizing its findings and delineating the transformative potential inherent in the proposed system.

III. Related Work and Literature Review

In recent times, there has been a growing interest in using Internet of Things (IoT)-based technologies to adjust water usage in agriculture. Previous research has highlighted the potential advantages of IoT based watering systems especially for sustainable gardening [2] [3]. Capacitive soil moisture sensors have proven quite effective in monitoring soil moisture levels [4]. However, there are still some challenges when it comes to adjusting watering schedules according to changing environmental parameters, which eventually affects efficiency and plant growth [5]. Various other studies have emphasized the significance of temperature and humidity sensors, in regulating plant health [6]. While previous work has focused on key aspects of gardening, this study proposes a comprehensive system that presents the synergy between IoT connectivity and various fundamental sensors. Their major goal of the IoT-enabled frameworks are for providing real time monitoring and flexible water distribution methods ultimately encouraging renewable gardening means.

Diverse research studies have delved into methods of incorporating technologies, for in-house gardening and efficient water management. A work centred on civic environments delivered a cloud-based remedy for home gardening. This solution employs sensor networking to accentuate resource efficacy through the implementation of sensors and IoT connectivity [7]. Additionally, an interesting impact involved the inclusion of block chain technology in an IoT-integrated gardening system especially designed for modern cities. This integration enhances precision, accuracy and reliability by automating watering procedures and notifying users in case of any water scarcity. It highlights the role played by blockchain in guaranteeing the dependability and security of smart gardening systems [8].

Moreover, recent analyses showcase the compelling requirement for water conservation and smart gardening practices. A proposed IoT-enabled system aimed at immediate pH monitoring proves to have been effective in retaining the quality of wastewater for agricultural usage [9]. A consequent research study fixated on water supply risks in India, harbouring a IoT based setup for RT gauging & quality evaluation of the water flow. This approach facilitates for a smooth and wireless technique to deal with the inadequacies, like costing challenges in water supply [10]. Existing explorations as these, in the realm of progressive and sustainable cultivation furnish the technological landscape with the leading applications and advancements of IoT based computing.

In the arena of ecological management, researchers have brought to light a portable buoy that assesses the quality of water in actual time across different settings. This innovative device showcases its adaptability, in environments such as coasts and lakes [11]. Moving on to control systems, there is experimentation investigating the use of fuzzy logic controllers to neutralize pH levels in water treatment processes. These controllers contribute to the recycling of water for various other applications, excluding drinking [12] [13]. Furthermore, a pioneering prediction mechanism for soil fertility based on IoT technology has been introduced. This system utilizes sensors to analyse soil characteristics, for yard cultivation [14]. An additional study that combine the Longer Short Termed Memories [LSTM] network using a traditional gardening system, utilizes machine learning methods to predictively evaluate and optimize gardening practices [15]. Finally, an indoor automated gardening system that operates autonomously using energy sources has been designed and developed, reinforcing the relevance of efficient and tech- driven home applications [16]. Collectively, this extensive survey of literature and related work outlines the evolving sphere of smart agriculture, environmental safeguards and water resource conservation by virtue of IoT applications.

Our proposed study takes a major step in this domain by presenting a holistic approach that combines the prominent aspects of several research studies. This versatile configuration not only affirms the usage of resources but also aims to tackles issues related to water insufficiency, alongside scrutinizing of various environmental factors in the garden, such, as temperature, humidity and soil moisture. By providing a versatile and robust solution adaptable to diverse settings, our proposed work contributes significantly to the evolution of sustainable and intelligent practices in smart gardening and water management.

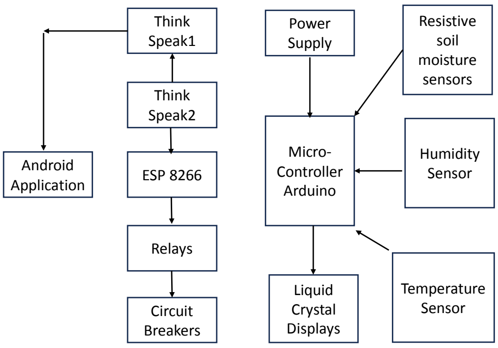


Fig. 1. Block Diagram of IOT-based Plant Watering System.

IV. SYSTEM ARCHITECTURE

Our designed system, featuring a Type-1 fuzzy logic-based controller, is illustrated in Figure 1 with a block diagram showcasing the IoT-Based Water Sprinkling System. The architecture is thoughtfully crafted, incorporating key components that synergistically contribute to the system's functionality [17].

V. KEY COMPONENTS AND FUNCTIONALITY

The system boasts essential components, each serving a distinct function in optimizing its performance. Fig. 1 presents the block diagram of the designed IOT-based plant watering system. The resistive soil moisture sensor functions as a proficient moisture measurement tool, utilizing variations in electrical resistance to gauge soil moisture levels accurately. The DHT11 Humidity and Temperature Sensor plays a key role, in adjusting watering patterns by monitoring the surrounding temperature and moisture levels. Its primary function is to ensure environmentally-sound irrigation. The Relay Module functions as an element in automating water control activating based on sensor input to maintain optimum soil parameters [19].

IoT connectivity is facilitated by the ESP8266 WiFi Module, enabling remote monitoring and control using cloud-based platforms such as Thing Speak [20] and a customized smartphone application. In this system the ATmega328P microcontroller, housed within the Arduino Uno serves as the controlling hub, further implementing a type 1 logic controller for smooth decision-making. This fuzzy controller is effectively utilized for adapting watering periods to make use of water resources based on live environmental conditions. In addition, the system includes an LCD display for easy data visualization and a DC pump, for automated irrigation guided by the fuzzy logic-dependent decisions, essentially creating an effective gardening solution [21].

VI. WORKING PRINCIPLE

Their implementations of the fuzzy dependent logical based controllers, in our system helps to improve its efficiency when it comes to smart devices especially in the context of automated agricultural practices. Fuzzy logic is beneficial for moulding doubtfulness and fuzziness which is particularly useful in scenarios where precise management becomes difficult. We have incorporated fuzzy logic in our experimental set-up and into the ATmega328P microcontroller to ensure decisiveness, when it comes to conforming to watering schemes. This allows us to enhance water usage by considering real-time soil moisture levels and environmental conditions, ensuring that the system can progressively acknowledge the plants' needs [22].

Beginning of the system framework, the model characterizes the collection of the data on the soil's temperature, humidity, and moisture levels. Consequently, these three parameters serve as input to the designed Type-1 fuzzy logic-based controller. The system activates a relay module, on being indicated by the fuzzy controller, emphasizing that the pump be initiated, activating the water pump to nourish the plants. This procedure occurs cyclically, offering continuous and automated irrigation of the plants. The integration of their ESP-8266 Wireless-Fidelity modules allows the systems to connect to Wi-Fi, enabling the uploading of sensor data to the Thing Speak IoT platform. This integration enhances data storage, analysis, and visualization, empowering users with insights for informed decision-making, contributing to efficient and sustainable gardening practices [23].

VII. ALGORITHM FOR IOT BASED PLANT WATERING SYSTEM:

- ***Step1***: Initialize hardware setup
- ***Step 2***: Check External Supply
- ***Step 3***: Read Sensor Readings
- ***Step 4***: Display on LCD
- ***Step 5***: Fed the obtained sensor readings (i.e., temperature, humidity, and soil moisture) to the designed Type-1 Fuzzy controller.
- ***Step 6***: If Controller Output == 'Watering Off' Turn off the pump via relay module, else: Turn on
- ***Step 7***: Establish TCP Connection to Thing Speak server
- ***Step 8***: Send Data to Thing Speak Cloud

- ***Step 9***: View the parameters
- ***Step 10***: Go back to Step 3

The schematic of the circuit diagrams are displayed in the Fig. No. 2, which is generated using Proteus software, depicts the IoT- based water sprinkling system's interconnected components, illustrating wiring, data flow, and control logic. Fig. 3 presents the workflow diagram of the designed IOT-based plant watering system. Fig. 4 shows the prototype of the designed IOT-based plant watering system[21].

VIII. DEVELOPED ANDROID APPLICATION

The smart irrigation system integrates with a user-friendly Android application, providing users with a friendly interface for real-time monitoring of various environmental parameters in the garden, including temperature, humidity, and soil moisture content. The Android application is meticulously crafted using MIT App Inventor. Both the frontend and backend of the application are skilfully designed, utilizing the user-friendly interface offered by MIT App Inventor. The Android application provides interactive graphs for real-time insights into soil moisture, humidity, and temperature via the Thing Speak cloud’s stored data. The different user interfaces of the developed Android application are illustrated in the figure numbers 5 to 8 respectively.

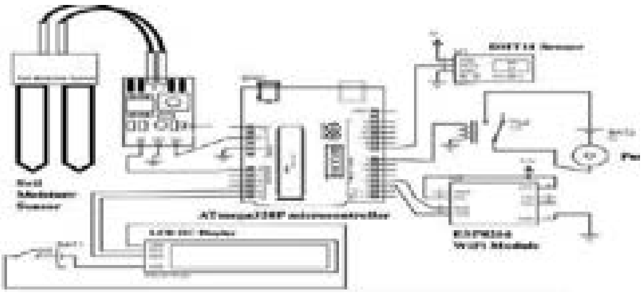


Fig. 2. Circuit Diagram of IOT Based Water Sprinkling System

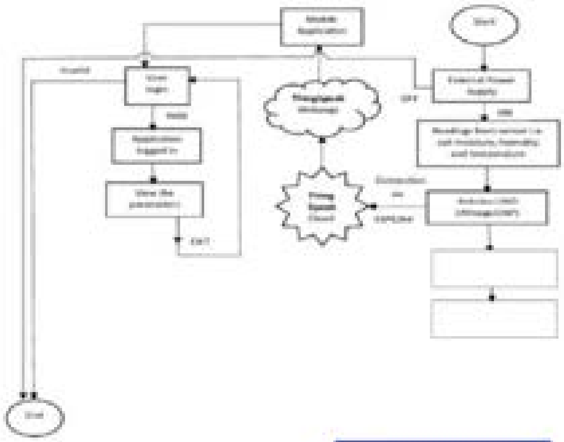


Fig. 3. The workflow diagram of the designed IOT-based plant watering system

Fig. 4. Prototype models for the designed IOT-based plant watering system

Fig. 5. Temperature monitoring interfaces for thier created applications

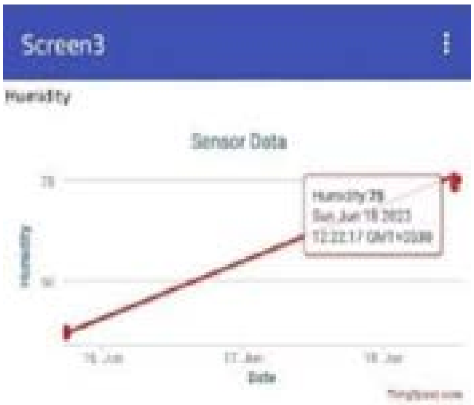


Fig. 6. Humidity monitoring interfaced of the developed application.

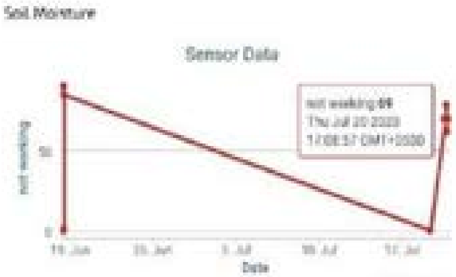


Fig. 7. Soil moisture monitoring interface of the developed application

IX. DESIGNED TYPED-1 FUZZY BASED CONTROLLERS

The paper employs a Typed-1 fuzzy based controllers to optimize the operation of a water pump utilized for garden watering. This optimization is based on three pivotal parameters [23]:

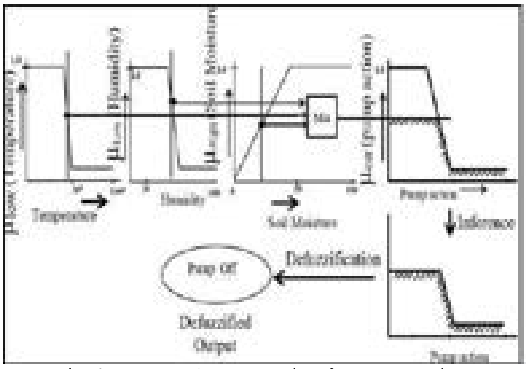


Fig. 8. An Example of a Fuzzy Rule.

X. DESIGNED TYPED-1 FUZZY BASED CONTROLLERS

The paper employs a Typed-1 fuzzy based controllers to optimize the operation of a water pump utilized for garden watering. This optimization is based on three pivotal parameters [24]:

A. For Temperature:

Lower : [10 to 20] , Medium : [20 to 40], Higher : [40 to 100]

B. For Soil Moisture:

Lower : [0 to 30] , Medium : [30 to 60], Higher : [60 to 100]

C. For Humidity:

Lower : [0 to 30] , Medium : [30 to 60], Higher : [60 to 100]

XI. STRUCTURE FOR THE RULER

In order to define the correlation between the input parameters and their corresponding output, specifically the pump control action, a set of rules is formulated. Some of the fuzzy rules are given below:

- If the Temperature = Lower & the Humidity == Lower and Soil's Moistures = HIGHer, Then, Pump Action = OFF (shown in Fig. 8)

- If the Temperature = HIGHer and Humidity == HIGHer & Soil's Moisture = LOWer, Then, Pump Action = ON

XII. REASONING FOR INFERENCE GENERATION

When, the measured 'Temperatures' are 'LOWer' and 'Himidity' are 'LOWer' & Soil's Moistures level are HIGHer, Then Pump Action is 'OFF'. Here, following steps are undertaken for reasoning with input measurement for inference generation.

- For say strength computation First strength (α) is computed.
- μ_{LOW} (*temperature*) , μ_{LOW} (*humidity*),
- μ_{HIGH} (*soil _ moisture*)
- For the given condition, μ_{OFF} (*pump _ action*) is computed.

○ μOff (Pump_Action) = μOff (Pump_Action)λα

A. Parameter identifications

Then, the various parameters like, temperature, humidity, and soil moisture. The interplay of these parameters is crucial in determining the watering requirements within a garden, offering a comprehensive understanding of a plant's environment. The controller acquires temperature and humidity data from the DHT 11 sensor, while soil moisture data is obtained from the soil moisture measuring sensor integrated into the system.

Subsequently, the controller fuzzifies these three precise values—temperature, humidity, and soil moisture—and transforms them into fuzzy inputs. The fuzzy controller defines two distinct pump control actions: 'Pump Off', and 'Pump On'. To establish the relationships b/w their inputs parameters & their corresponding outputs (i.e., pump control action), a set of rules is devised. These rules delineate the nuanced interactions among temperature, humidity, and soil moisture, guiding the fuzzy controller in making informed decisions about the water pump operation [25].

B. Fuzzy Regions

The delineated fuzzy regions for the three input parameters are categorized as low, medium, and high. Simultaneously, the output, reflecting the watering level of the pump, is partitioned into two fuzzy regions: Off and On. In our carefully crafted implementation, specific intervals have been assigned to each parameter, ensuring a nuanced representation within the fuzzy logic framework.

C. De-fuzzification:

Centroidal defuzzification technique is used to get the desired pump action.

TABLE I. ENVIRONMENTAL DATA STORED ON THINGSPEAK

Entry ID	Created at	Humidity (%)	Temperature (°C)	Soil moisture (%)
350	2022-07-18T07:02:51Z	75	93.93	10
360	2022-08-18T07:03:11Z	78	63.60	20
370	2022-09-20T13:07:54Z	80	33.30	30

XIII. Results and Discussion

The IoT-driven automated plant watering system, meticulously designed, seamlessly captures and stores three distinct sets of environmental data from a garden onto the Think Speak Cloud. Table I represents real-time ThingSpeak cloud data, including humidity, temperature, and soil moisture with timestamps, reflecting changing environmental conditions. This stored data is really helpful in monitoring the atmospheric changes occurring garden condition. The archived data proves invaluable for monitoring changes in atmospheric conditions affecting the garden. Through comprehensive analysis of recorded parameters— temperature, humidity, and soil moisture—it is evident that the prototype model of the designed Type-1 fuzzy controller- powered plant watering system effectively maintains soil moisture. This is achieved by efficient pump control, adhering to the specified fuzzy rules and thereby reducing water wastage.

The proposed prototype of the plant watering system not only averts under-watering and over-watering instances but also adapts its watering approach to the dynamic changes within the garden, courtesy of its fuzzy controller. The attached plant exhibits robust and healthier growth, providing concrete evidence of the system's efficacy. Furthermore, the Android application, displaying real-time values of temperatures, humidities & the soil's moistures, empowers users for remotely monitor their garden's condition, enhancing overall convenience. The designed IoT-based, Type-1 fuzzy controller-powered automated plant watering systemboasts several key advantages:

- Remarkably low energy consumption,
- Real- time information on soil moisture, temperature, and humidity changes,
- Cost-effectiveness,
- The system's LCD display presents garden parameters, enabling users to view essential values even without utilizing the Android application.

XIV. Conclusion

In conclusion, the designed and implemented IOT-based water sprinkling system has the potential to efficiently automate gardening, while assessing soil moisture and other environmental parameters by calibrating sensors and employing a type-1 fuzzy controller. The microcontroller and ESP8266 Wi-Fi module transmit data to Thing Speak cloud, enabling monitoring and analysis of garden's temperature, humidity, and soil moisture data. During its execution, several challenges have surfaced, encompassing issues such as sensor accuracy, the dependability of Thing Speak data, and seamless connectivity with the Wi-Fi module. Future improvements involving predictive irrigation algorithms and weather integration would result in a more precise and resource-efficient gardening solution.

References

[1] T. Patil, J. D. Mallapur and D. G. Mallapur, "Self Monitored Indoor Plant Maintenance with Wastewater," 2023 International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 1-6.

[2] J. Smith and A. Johnson, "IoT-enabled smart watering for efficient agriculture," Int. J. Agric. Technol., vol. 15, no. 3, pp. 437-450, 2019.

[3] L. Brown and C. Miller, "Smart irrigation systems: A comprehensive review," J. Sustainable Agric., vol. 18, no. 2, pp. 245-261, 2020.

[4] R. White and S. Green, "Soil moisture sensing for precision agriculture," Agric. Eng. Today, vol. 22, no. 3, pp. 10-18, 2018.

[5] K. Patel and J. Lee, "Adaptive irrigation scheduling using IoT-based soil moisture monitoring," in Proc. IEEE Int. Conf. Smart Agric. (ICSA), pp. 123-130, 2017.

[6] M. Johnson and S. Williams, "Humidity and temperature sensors for plant health monitoring: A review," Sensors in Agric., vol. 5, no. 2, pp. 87-98, 2019.

[7] S. Sharma, A. Sharma, T. Goel, R. Deoli, and S. Mohan, "Smart Home Gardening Management System: A Cloud-Based Internet-of-Things (IoT) Application in VANET," *IEEE*, Jul. 2020, doi: 10.1109/icccnt49239.2020.9225573.

[8] T. S. RajaRajeswari, P. Chinnasamy, K. Pushparani, N. Thulasichitra, N. S. Rani, and T. Sivaprakasam, "IoT based Smart Gardening for Smart Cities using Blockchain Technology," *2022 International Conference on Computer Communication and Informatics (ICCCI)*, Jan. 2022, doi: 10.1109/iccci54379.2022.9741024.

[9] N. Khatri, A. Sharma, K. K. Khatri, and G. D. Sharma, "An IoT-Based innovative Real-Time PH monitoring and control of municipal

wastewater for agriculture and gardening," in *Smart innovation, systems and technologies*, 2018, pp. 353–362. doi: 10.1007/978-981- 10-5828-8_34.

[10] S. Anjana, M. N. Sahana, S. Ankith, K. Natarajan, K. R. Shobha, and A. Paventhan, "An IoT based 6LoWPAN enabled experiment for water management," *IEEE*, Dec. 2015, doi: 10.1109/ants.2015.7413654.

[11] A. H. M. A. Helmi, M. Hafiz, and M. S. B. S. Rizam, "Mobile buoy for real time monitoring and assessment of water quality," *IEEE*, Dec. 2014, doi: 10.1109/spc.2014.7086223.

[12] A. F. O. De Azevedo Dantas, F. R. Lima, A. L. Maitelli, and G. F. Dantas, "Comparative study of strategies for fuzzy control P and PD applied to a pH plant," *IEEE*, Nov. 2011, doi: 10.1109/iecon.2011.6119364.

[13] R. Ibrahim, "Practical modelling and control implementation studies on a pH neutralization process pilot plant," *2008. [Online]. Available: https://theses.gla.ac.uk/311/01/FINAL_THESIS_200508.pdf*

[14] R. Aarthi, D. Sivakumar, and V. Mariappan, "Smart Soil Property Analysis Using IoT: A case study implementation in backyard gardening," *Procedia Computer Science*, vol. 218, pp. 2842–2851, Jan. 2023, doi: 10.1016/j.procs.2023.01.255.

[15] M. M. Saad, M. T. R. Khan, M. A. Tariq, and D. Kim, "LSTM Enabled Artificial Intelligent Smart Gardening System," *IEEE*, Oct. 2020, doi: 10.1145/3400286.3418260.

[16] E. Mackensen, J. Klose, A. Rombach, and A. Spitznagel, "Energy autonomous automation of Smart Home applications using the example of a wireless Indoor Smart Gardening system," *IEEE*, Aug. 2019, doi: 10.1109/coase.2019.8843000.

[17] Shashank S., Kushal K., Abhay Surya Shankar, Madan Kumar G., Dr. Pavithra G., Dr. Sindhu Sree M., Padmavathy M., Dr. T.C. Manjunath, "RFID based attendance system with SMS Alert", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 774 – 782, Jul. - Sept. 2023.

[18] Abhishek, Sujith M.S., Jeevan D., P. Kamalesh, Dr. Sindhu Sree M., Dr. Pavithra G., Dr. T.C. Manjunath, "Design & Development of a Table Assisted Robotic Arm", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 818 – 822, Jul. - Sept. 2023.

[19] Apeksha U., Chithrashree G.S., Divya N.M., Shalmali S. Mankikar, Dr. Sindhu Sree M., Dr. Pavithra G., Dr. T.C. Manjunath, "Wireless LoRa Communication Between Two Arduino Uno for Military Application in Soldier Tracking", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN : 1001-4055, Vol. 44, No. 3, pp. 768 – 773, Jul. - Sept. 2023.

[20] Akarsh Kesharwani, Ayush P. Chaudhary, Bhanu Pratap Singh, Ved Kumar, Padmavathi M., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, "A Study on Hand Motion Controlled Robotic Arm", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 812 – 817, Jul. - Sept. 2023.

[21] Leena Jeyakumar, Prerana Aithal, Vismitha R., Pradhan Aithal, Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, "Development of Smart Bridge – Automatic Height Increase When Floodings Take Place", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, HIndex 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 763 – 767, Jul. - Sept. 2023.

[22] Sushanthi Raj, Manohar R., Bhuvan G.S., Deepthi. S.R., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, "Design and Development of Obstruction Detection and Removal Bot", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 807 – 811, Jul. - Sept. 2023.

[23] Kavya P., Sanjana S., Harika, Teju R., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, "Design & Development of Drones Using Radio Frequency Controllers", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 707 – 710, Jul. - Sept. 2023.

[24] S.G. Swathi, Aliya Bhandari, Srushti M. B. Kavya A., Dr. Pavithra G., Dr. Sindhu Sree M., Dr. T.C. Manjunath, "Voice Control Robot – Design & Development for Various Domestic Applications", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 791 – 801, Jul. - Sept. 2023.

[25] Vaishnavi Patil, Dr. Pavithra G., Dr. T.C. Manjunath, "Design of Smart Wheelchair For Disabled (Handicapped) Persons Using Real Time Embedded Systems & Internet Of Things Approach", *Scopus Journal Q3, Schimago Ranking SJR 2022 0.32, H-Index 24, Tuijin Jishu/Journal of Propulsion Technology*, ISSN: 1001-4055, Vol. 44, No. 3, pp. 871 – 878, Jul. - Sept. 2023.

Role of Convolutional Neural Networks in Plant Leaf Disease Detection

1st Maibam Naresh Singh
Department of ME-CSE,Final Year
Chandigarh University
Gharuan, Punjab, India
nareshhmaibam@gmail.com

2nd Dr.Abhishek Kumar
Department of CSE, Professor
Chandigarh University
Gharuan, Punjab, India
abhishek.e14283@cumail.in

3nd Dr.Sachin Ahuja
Department of CSE, Professor
Chandigarh University
Gharuan, Punjab, India
sachinahuja11@gmail.com

***Abstract*—Convolutional Neural Networks, or CNNs, have become a potent tool for image recognition applications, such as the identification of plant leaves. CNNs showed innovative capabilities on a range of image classification and segmentation tasks, demonstrating their ability to extract significant information from pictures. In this work, we offer a CNN-based approach to identify leaves on plants. The proposed method uses a sizable collection of plant leaf photos to train a CNN model. Next, plant leaves in fresh photos may be identified using the trained model. The proposed method is assessed by using a benchmark dataset and show that it produces cutting-edge outcomes.**

***Index Terms*—Plant Leaf Detection, CNN, Image Recognition**

I. INTRODUCTION

In the realm of precision agriculture, the timely and accurate identification of plant diseases stands as a critical imperative to ensure global food security and sustainable crop management [4]. Late developments in deep learning, especially Convolutional Neural Networks (CNNs), have demonstrated remarkable potential in revolutionizing the domain of crop disease recognition. This review synthesizes and critically examines the existing literature on the utilization of CNNs for the auto- mated identification of crop illness through leaf images [3]. The ubiquitous nature of plant diseases and their detrimental impact on crop yield necessitate innovative and efficient detection methods. Traditional approaches, while valuable, often rely on manual inspection and are limited in scalability. CNNs, inspired by the human visual system, have emerged as a transformative technology, capable of automatically learning hierarchical features from raw image data [16]. This capability makes them particularly well-suited for discerning the intricate patterns and subtle manifestations associated with various plant diseases.

This review encompasses a comprehensive survey of the most recent techniques, datasets, and construction employed in CNN-based plant disease detection. We explore the evolution of CNNs in this domain, from early applications to recent advancements, highlighting the key challenges and opportunities [15]. The impact of transfer learning, data augmentation, and other strategies for model generalization are discussed to provide insights into enhancing the robustness of disease detection models.

Furthermore, we delve into the evaluation metrics commonly employed in assessing the performance of CNNs in plant disease detection, considering aspects such as accuracy, sensitivity, and specificity [9]. Real-world applicability is addressed by examining the adaptability of CNN-based models to diverse environmental conditions and the challenges associated with deploying such technologies in agricultural settings.

As precision agriculture continues to evolve, the integration of CNNs into plant disease detection systems holds great promise for revolutionizing crop management practices [12]. This review aims to distill the current state of knowledge, identify gaps in research, and guide future investigations towards the development of more accurate, scalable, and practical solutions for plant disease detection using CNNs. Through this exploration, we seek to contribute to the ongoing discourse on the intersection of deep learning and agriculture, fostering advancements that address the pressing challenges faced by the global farming community.

II. LITERATURE SURVEY

S.K. et al. [2] This research suggests a convolutional neural network (CNN) based plant disease detection system. The four steps of the system are feature extraction, categorization, picture pre-processing, and image capture. Plant leaf pictures are gathered during the image acquisition step. The pre-processing step of the photograph enhances the quality and eliminates noise. CNNs are used in the feature extraction step to extract features from the pictures. The photos are categorized as either healthy or sick at this point.

A montage of 200 photos of the leaves of tomatoes with along various illnesses was used to assess the suggested approach. With a 91.1% accuracy rate, the technique proved to be useful in identifying illnesses of tomato leaves.

Shrestha et al. [1] The authors suggest a CNN-based method for detecting plant diseases that uses image classification to detect sick leaves on plants. The three primary phases of the system are categorization, pre-processing, and picture capture. Plant leaf images are captured using a digital camera or smart- phone, and they need to be sharp and have enough resolution to accurately identify diseases. Pre-processing techniques like scaling, normalization, and noise reduction may be used to improve the quality of the captured pictures and eliminate any noise or distortions. Ultimately, a CNN model is given the previously processed pictures in order to classify them. Using a collection of labelled plant leaf photos, the CNN model is trained to identify patterns linked to various illnesses. This technique presents a viable method for precise plant disease identification.

Bedi et al.[5] Convolutional automatic encoders (CAE) and a neural networks with convolution are combined by the authors he authors to propose a blended approach for plant disease diagnosis. In order to distinguish between healthy and unhealthy characteristics in plant leaf photos, the CNN is utilized after the features have been retrieved using the CAE. The suggested approach outperforms conventional CNN-based models in a number of ways. Initially, the characteristics that the CAE can extract from plant leaf photos are more useful for classifying diseases than the raw values of the pixels. It takes less time and computer resources to train the hybrid model since it requires fewer training parameters than a CNN by itself. A few difficulties also exist with the suggested model. For a start, in order to get high accuracy, the model needs a lot of training data. Secondly, the model might be computationally expensive to train.

U.S et. al [3] The potential of machine learning (ML) classification methods for plant disease detection is thoroughly reviewed in this research. The writers talk on a number of classification techniques, such as artificial neural networks (ANNs), decision trees, k-nearest neighbours (k-NN), support vector machines (SVMs), and more. They also draw attention to the difficulties in detecting plant diseases, namely the fluctuation of plant appearance and the requirement for a substantial quantity of training data.

Reddy et al. [12] For the purpose of detecting plant leaf illness, researchers examined three classification algorithms: convolutional neural networks (CNN), k-nearest neighbours (KNN), and support vector machines (SVM). A dataset consisting of 1,000 photos of plant leaves was utilized, with 100 images representing each of the ten distinct plant leaf diseases. They trained the three classification algorithms after preprocessing the photos to eliminate noise and extract features. CNNs attained a precision of 99.2%, which was higher than both SVM and KNN's results of 85% and 92.2%, respectively.

Yadhav et al. [6] In this work, a novel activation function—dubbed the optimized activation function, or OAF—is proposed, and its effectiveness in a CNN-based plant disease detection model is assessed. The OAF-based CNN strategy output is contrasted with CNN strategy efficiency that employ more conventional activation functions, including ReLU and ELU, by the authors. The findings show that as compared to models that use conventional activation functions, the OAF- based CNN model performs more accurately in detecting plant diseases.

Guan, X[14] By combining four CNN models— Inception, ResNet, InceptionResNet, and DenseNet—this work suggests a unique approach with purpose of recognition crop leaf illness. A dataset of 36,258 plant leaf photos divided into 10 plant species and 61 classifications of healthy and sick leaves is used by the authors to assess how well the suggested strategy performs. According to the data, the suggested method outperforms current techniques and detects plant leaf diseases with a high accuracy of 87%.

Jasim et al. [13] This study investigates the identification of plant leaf diseases using deep learning and image processing techniques. The approaches' ability to reliably differentiate between healthy and sick leaves is highlighted by the authors, providing farmers with the opportunity to promptly and efficiently safeguard their crops. They also go over how crucial architectures for deep learning, particularly convolution neural system (CNNs)—are to obtaining precise illness identification, as well as the significance of pre-processing and feature extraction.

Das et al. [19] Support vector machines (SVMs) are used in this research to present a two-stage technique for plant leaf disease detection. Gaussian smoothing, histogram equalization, and contrast stretching are some of the methods used in the initial step of pre-processing to improve the quality of leaf photographs. With the use of a mix of colour, texture, and form elements, pertinent features are extracted from the pre-processed pictures in the second step. After that, an SVM classifier—a supervised learning system that can divide data into many categories—is trained using these extracted characteristics. The scientists used a dataset of 1,000 leaf photos that were divided into 10 different illness classifications in order to assess the suggested approach. According to the findings, the suggested strategy attained an amazing accuracy of 87.

Sardogan et al. [4] Using convolutional neural networks (CNNs) and the learning quantization of vector (LVQ) algorithm, the study suggests a two-phase strategy for identifying plant leaf illnesses. In the first stage, leaf pictures are pre-processed using methods including contrast stretching, histogram equalization, and Gaussian smoothing to reduce noise and enhance image quality. A CNN that has been trained to recognize characteristics typical of various leaf diseases is used in the second phase to extract features from the pre-processed pictures. After that, an LVQ classifier—a supervised learning system that can divide data into many categories—is fed these extracted characteristics. Using a dataset of 500 leaf photos divided into four different disease classifications, the scientists assessed the efficacy of their suggested technique. The findings showed that the suggested strategy had an accuracy of 86%. The authors suggest improving the method with a larger dataset and enhanced disease detection.

III. RESULT ANALYSIS

In the below result analysis table I, we have discussed the methods used by previous researchers. Also we have discussed the various gaps observed in the existing methods.

TABLE I: REUSLT ANALYSIS

Author & Year	Algorithm used	Dataset used	Performance measures	Gaps
S.K.et al. (2021) [2]	FR-CNN	A dataset of 54,303 images comprising healthy and unhealthy leaves from 38 plant species.	Accuracy: 91.1%.	The study did not evaluate the performance of the method on diseases that cause subtle changes in leaf appearance. Further research is needed to determine whether the method is effective for detecting these diseases.
Shrestha et al. (2020) [1]	SVM	The dataset isnot mentioned.	Accuracy:88.80%	This case study examined a relatively small number of tomato diseases—just two. The concept might not work well for other illnesses.
Sardogan et al. (2018) [4]	LVQ	Utilizes a dataset comprising 500 leaf images.	Accuracy:86%	The dataset used in the study was limited to four plant species. A greater and additional diverse dataset could be essential to evaluate the efficacy of the method for a wider range of plant species.
Rudagi et. al. (2022) [7].	Machine Learning	A dataset of 100 tomato leaf images.	Accuracy:94.3%	The dataset used in the study was limited to 100 tomato leaf images. A greater and additional diverse dataset could be essential to evaluate the efficacy of the method for a wider range of plant species and diseases.
Marzougui et al. (2020) [8].	Deep Learning	A dataset of 2,822 leaf images classified into 14 different disease categories.	Accuracy:92.80%	The effectiveness of the approach was not compared to other cutting-edge techniques for the recognition of crop leaf illness o the study. To ascertain whether the approach is better than alternative approaches, more investigation is required.
Tugrul et. al (2018) [9]	Deep Learning	The dataset used is collected from PlantVillage	Accuracy:92%	The effectiveness of the approach on photos shot in various lighting situations or with varying picture resolutions was not assessed in the research. If the approach is resilient to these changes, more investigation is required.
Hassan et. al. (2021) [10]	Decision Tree	The dataset used is collected from the PlantVillage	Accuracy:91.14%	Three plant species and a comparatively tiny number of photos made up the study's dataset. A more extensive and varied dataset would be required to assess the method's efficacy over a broader spectrum of plant species and ailments.
Applalanaidu et. al (2021) [11]	Machine Learning	The dataset used is collected from the PlantVillage	Accuracy:89.6%	The effect of different picture quality on classification accuracy was not taken into account in the research. Robust crop illness recognition requires more investigation.
Reddy et. al. (2019) [12]	KNN	The dataset used is collected from PlantVillage	Accuracy:85%	Because of the tiny size of the dataset, the conclusions may not be as broadly applicable.
Bedi et al. (2021) [19]	Hybrid of CAE network and CNN	The PlantVilla gedataset is used.	Accuracy:94.4%	The model was only tested on leaf photos of afflicted peach plants with Bacterial Spot disease, and it was trained on a limited dataset of 1000 images.

U, S. et al. (2019) [3]	Maachine Learning	The PlantVillage dataset is used.	Accuracy:99.34%	The need for more resilient and broadly applicable machine learning models that can adapt to differences in plant species, disease symptoms, and picture quality
Ganatra et.al. (2020) [15]	Machine Learning	The study used a dataset of diseased leaf images for corn, tomato, and potato.	Accuracy:73.38%	A number of variables, including lighting and camera resolution, might cause variances in the quality of the field photos. Determining how resilient the approach is to these changes will require more investigation.
Hussein et al. (2019) [17]	SVM	The dataset used is collected from PlantVillage	Accuracy:88.1%	The method's generalizability to plant disease detection in other plant species was not evaluated in the study.
Jasim et al. (2020) [13]	CNN	A collection of 20,636 plant leaf photos from the Plant Village dataset	Accuracy:95%	The technique ignores the problem of different picture quality, which might have an impact on how accurate the categorization findings are.
Li et al. (2021) [18]	Deep Learning	The PlantVillage project dataset is used	Accuracy:96%	Precision agricultural technologies must be integrated with deep learning algorithms to identify plant diseases. This would make it possible to create automated and more thorough plant health management solutions.
Guan et al. (2019) [14]	Deep Learning	The dataset contains 36,258 leaf images of 10 different plant species	Accuracy:87%	The study cannot be utilized to identify all potential plant illnesses because the dataset only contains information on 10 plant species.
Das et. al (2020) [19]	SVM	The PlantVillage dataset is used	Accuracy:90.5%.	The method's integration with precision agricultural systems was not investigated in the study.
Yadhav et.al. (2021) [6]	Genetic Algorithm	A dataset of 867 images of tomato leaves	Accuracy:95%	The study did not assess how well the suggested approach performed in comparison to other techniques for identifying plant diseases.
Ajra et al. (2020) [16]	CNN	This dataset contains 54,306 images of plant leaves	Accuracy:93.3%	The Kaggle dataset's small size could restrict how far the findings can be applied.
Ji et al. (2019) [20]	Multi-network integration method	The PlantVillage dataset is used	Accuracy:94.1%	The range of quality and amount of data utilized in the training and evaluation of machine learning models for grape leaf illness detection can have an impact on the models' accuracy and generalizability.

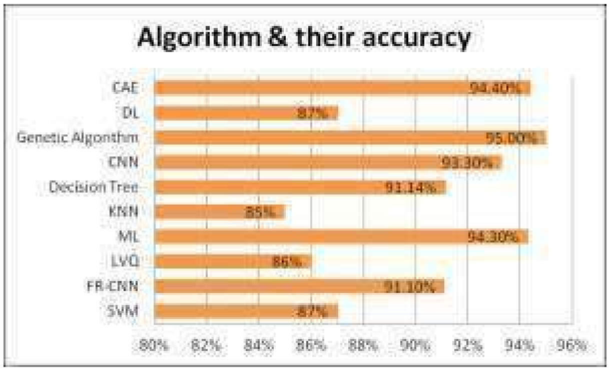


Fig. 1. Algorithms and their accuracy

In the above table 1, we can see that the most used dataset is the PlantVillage dataset. While most of the models acquired around an accuracy of 90. The algorithms with their accuracies are shown in the form bar chart in Fig: 1. below. We also discussed the gaps in the models used.

IV. Research Gap

Convolutional neural networks, or CNNs, can be used to identify and classify diseases, according to the information at hand, although there are still a number of research gaps in this area. The following are some areas that require more research:

a) : For the purpose of detecting plant diseases, many CNN-based models are created and trained for particular plant species. Since farmers frequently cultivate a variety of crops on their fields, it is necessary to create models that can effectively generalize across various plant species [5].

b) The majority of current studies focus on detecting diseases in labs and greenhouses, which are controlled limitations. There is a research gap that deals with the difficulties with illness identification in open-field, real-world circumstances where factors such as changing illumination, weather, and the state of the soil may significantly impact picture quality and disease presentation [7].

c) : There is a class imbalance across numerous of the openly available datasets for plant diagnosis of diseases using CNNs, where specific diseases or healthy data are underrepresented. As a result, the model's performance may become biased. Methods for effectively managing unbalanced datasets should be the subject of study [11].

d) : In reality, plants may develop multiple diseases at once, each with a unique manifestation and complicated interactions. Modern models often emphasize the classification of a specific disease. There is an investigation gap in creating models that 14 can simultaneously detect, characterize, and fully understand the interrelationships of numerous illnesses [14].

e) : Due to privacy concerns, many farmers are reluctant to divulge their crop photographs and data. Exploring methods that protect the confidentiality of information and do not rely upon centralized data collecting will allow CNN models to evolve and provide precise illness diagnoses [18].

V. CHALLENGES OF THE WORK

The challenges of the work are as follows:

a) : Overfitting is a problem with CNN models, which causes them to perform well on training data but fall short on fresh, untried data [12].

b) : It can be challenging to get a significant labelled dataset for CNN training, particularly for certain plant species or diseases [6].

c) : Biases contained in the training data might be inherited by CNN models, resulting in unfair or biased results [9].

d) : The computing expense of training CNN models may be high, particularly for big architectures or datasets [11].

VI. Conclusion

In conclusion, the application of convolutional neural networks (CNN) for illness recognition in crop leaves was the main topic of this review study. The study investigated the efficacy of CNN models in precisely diagnosing and categorizing a range of plant leaf diseases by carefully reviewing previous research. Reviewers emphasized CNN models' bene- fits, especially their capacity to automatically extract relevant information from leaf pictures without human interaction. CNN models consistently showed greater levels of accuracy than standard machine learning techniques, making them a viable tool for plant disease identification. The assessment did, however, also note certain difficulties in this area. For example, the performance of CNN models may be impacted by the scarcity of large-scale datasets and changes in picture quality. Furthermore, more resilient models that can manage various illness presentations are required. In order to overcome these obstacles and enhance the precision and dependability of crop leaf illness recognition using CNN, more study and development are required. In conclusion, this study adds to the body of knowledge already available about the utilization of CNN to the identification of plant disease in leaves. It gives insightful information on the state of the subject at the moment, points out areas in need of more research, and makes some recommendations for possible future research directions. CNN models are anticipated to be fundamental in the early and precise identification of crop leaf illness, helping farmers reduce crop losses and enhance overall plant health as a result of continuous advancements in machine learning techniques.

References

[1] Shrestha, G., Deepsikha, Das, M., and Dey, N. (2020). Plant Disease Detection Using CNN. 2020 IEEE Applied Signal Processing Conference (ASPCON). doi:10.1109/aspcon49795.2020.9276722.

[2] S. K. . et. al., "Plant Disease Detection Using CNN", TURCOMAT, vol. 12, no. 12, pp. 2106–2112, May 2021.

[3] U, S., Nagaveni, V., and Raghavendra, B. K. (2019). A Review on Machine Learning Classification Techniques for Plant Disease Detection. 2019 5th International Conference on Advanced Computing and Communication Systems (ICACCS). doi:10.1109/icaccs.2019.8728415.

[4] Sardogan, M., Tuncer, A., and Ozen, Y. (2018). Plant Leaf Disease Detection and Classification Based on CNN with LVQ Algorithm. 2018 3rd International Conference on Computer Science and Engi-neering (UBMK). doi:10.1109/ubmk.2018.8566635.

[5] Bedi, Punam and Gole, Pushkar. (2021). Plant disease detection using hybrid model based on convo-lutional autoencoder and convolutional neural network. Artificial Intelligence in Agriculture. 5. 90-101. 10.1016/j.aiia.2021.05.002.
[6] Yadhav, S. Y., Senthilkumar, T., Jayanthy, S., and Kovilpillai, J. J. A. (2020). Plant Disease Detection and Classification using CNN Model with Optimized Activation Function. 2020 International Con-ference on Electronics and Sustainable Communication Systems (ICESC). doi:10.1109/icesc48915.2020.91558.
[7] Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., and Pramodhini, R., "Plant leaf disease detection using computer vision and machine learning algorithms", ¡i¿Global Transi-tions Proceedings¡/i¿, vol. 3, no. 1, pp. 305–310, 2022. doi:10.1016/j.gltp.2022.03.016.
[8] Marzougui, F., Elleuch, M., Kherallah, M. (2020). A Deep CNN Approach for Plant Disease De-tection. 2020 21st International Arab Conference on Information Technology (ACIT). doi:10.1109/acit50332.2020.930007.
[9] Tugrul, B., Elfatimi, E., Eryigit, R. (2022). Convolutional Neural Networks in Detection of Plant Leaf Diseases: A Review. Agriculture, 12(8), 1192. https://doi.org/10.3390/agriculture12081192.
[10] Hassan, S. M., Jasinski, M., Leonowicz, Z., Jasinska, E., Maji, A. K. (2021). Plant Disease Identi-fication Using Shallow Convolutional Neural Network. Agronomy, 11(12), 2388. https://doi.org/10.3390/agronomy11122388.
[11] Applalanaidu, M. V., Kumaravelan, G. (2021). A Review of Machine Learning Approaches in Plant Leaf Disease Detection and Classification. 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV). doi:10.1109/icicv50876.2021.93884
[12] Reddy, J. N., Vinod, K., A S, R. A. (2019). Analysis of Classification Algorithms for Plant Leaf Disease Detection. 2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT). doi:10.1109/icecct.2019.8869090
[13] Jasim, M. A., AL-Tuwaijari, J. M. (2020). Plant Leaf Diseases Detection and Classification Using Image Processing and Deep Learning Techniques. 2020 International Conference on Computer Sci-ence and Software Engineering (CSASE). doi:10.1109/csase48920.2020.91420
[14] Guan, X. (2021). A Novel Method of Plant Leaf Disease Detection Based on Deep Learning and Convolutional Neural Network. 2021 6th International Conference on Intelligent Computing and Sig-nal Processing (ICSP). doi:10.1109/icsp51882.2021.940880
[15] Ganatra, Nilay Patel, Atul. (2020). A Multiclass Plant Leaf Disease Detection using Image Pro-cessing and Machine Learning Techniques. International Journal on Emerging Technologies. 11.
[16] Ajra, H., Nahar, M. K., Sarkar, L., Islam, M. S. (2020). Disease Detection of Plant Leaf using Image Processing and CNN with Preventive Measures. 2020 Emerging Technology in Computing, Communication and Electronics (ETCCE). doi:10.1109/etcce51779.2020.93508
[17] Hussein, Mohammed Abbas, Amel. (2019). Plant Leaf Disease Detection Using Support Vector Machine. Al-Mustansiriyah Journal of Science. 30. 105. 10.23851/mjs.v30i1.487.
[18] Li, L., Zhang, S., Wang, B. (2021). Plant Disease Detection and Classification by Deep Learning—A Review. IEEE Access, 9, 56683–56698. doi:10.1109/access.2021.3069646
[19] Das, D., Singh, M., Mohanty, S. S., Chakravarty, S. (2020). Leaf Disease Detection using Support Vector Machine. 2020 International Conference on Communication and Signal Processing (ICCSP). doi:10.1109/iccsp48568.2020.91821
[20] Ji, Miaomiao Zhang, Lei Wu, Qiufeng. (2019). Automatic Grape Leaf Diseases Identification via UnitedModel Based on Multiple Convolutional Neural Networks. Information Processing in Ag-riculture. 7. 10.1016/j.inpa.2019.10.003.

Road Surveillance using Real-Time Detection and Alert System

Nilamadhab Mishra
(CS and DS) Department
GHRCEM Pune
Pune, India
nilamadhab76@gmail.com

Ajay Rodge
(CS and DS) Department
GHRCEM Pune
Pune, India
ajayrodge713@gmail.com

Minal Deo
(CS and DS) Department
GHRCEM Pune
Pune, India
minalkdeo@gmail.com

Priya Agalwe
(CS and DS) Department
GHRCEM Pune
Pune, India
priyaaglawe2018@gmail.com

Neha Bansode
(CS and DS) Department
GHRCEM Pune
Pune, India
nehabansode34@gmail.com

***Abstract*— This study introduces a novel approach to accident detection utilizing YOLOv8, a state-of-the-art object detection model, integrated with a real-time alert system aimed at swiftly notifying emergency services and nearby individuals to prevent injuries and fatalities. This study presents an innovative approach to road surveillance using real-time detection and alert systems, integrating advanced machine learning models and communication technologies. The architectural framework encompasses Mask R-CNN for precise object detection and segmentation, DeepDream for anomaly accentuation, Variational Autoencoders (VAEs) for feature extraction and anomaly detection, and YOLO for instantaneous object detection. The amalgamation of these models facilitates comprehensive monitoring of road environments, proficiently identifying vehicles, pedestrians, and potential hazards with high accuracy. Furthermore, the Twilio API is incorporated to facilitate instant communication and alerts based on detected events, ensuring swift responses from relevant authorities. This holistic approach not only enhances road safety by proactively identifying risks but also streamlines emergency response mechanisms. The system's effectiveness is demonstrated through rigorous testing on real-world road scenarios, showcasing its potential for wide-scale deployment in intelligent transportation systems for enhanced traffic management and accident prevention.**

***Keywords*— Road Surveillance, Detection, Safety, Alerting, Deep Learning**

I. INTRODUCTION

In an era marked by rapid technological advancements, the need to ensure road safety has become more critical. The alarming statistics of road accidents highlight the urgent need for innovative solutions to mitigate risks and enable safety measures. Among these solutions, real-time detection and alert systems have emerged as promising tools, leveraging cutting-edge technologies to proactively identify potential hazards and alert drivers promptly.

The utilization of YOLOv8 within the context of road surveillance represents a significant advancement over conventional methods. Unlike traditional surveillance systems that rely on pre-defined regions of interest and sequential processing, YOLOv8 offers a holistic approach by simultaneously detecting and classifying objects within the entire image. This not only enhances the system's responsiveness but also enables it to adapt dynamically to diverse environmental conditions and scenarios.

A. Sensor Network

A network of high-resolution cameras strategically positioned along roads captures real-time video feeds, which serve as input for the detection and alert system. These cameras cover various perspectives, including intersections, pedestrian crossings, and high-traffic zones to ensure comprehensive surveillance.

B. YOLOv8 Object Detection Module

The YOLOv8 object detection module processes the incoming video streams and identifies relevant objects such as vehicles, pedestrians, cyclists, and obstacles. YOLOv8's ability to detect objects with remarkable speed and accuracy enables the systemto respond swiftly to potential hazards.

C. Alert Generation and Delivery Mechanism

Upon detection of objects of interest, the system generates instant alerts to notify drivers, pedestrians, and relevant authorities. These alerts can be disseminated through various channels, including roadside displays, in-vehicle notifications, mobile applications, and emergency services communication networks.

D. Data Fusion and Analysis Engine

The system incorporates a data fusion and analysis engine, which integrates information from multiple sources, including video feeds, sensor data, and historical records. By analyzing patterns, anomalies, and near-miss events, this engine provides valuable insights for decision-making and resource allocation.

II. RELATED WORKS

Initially, the image data in that region is converted into a form that is compatible with the CNN (its architecture requires inputs of a fixed 227 × 227 pixel size)[1]. YOLO utilizes its particular CNN to complete classification and location of multiple objects in an image at one time[2].
YOLOv5 consistently outperforms in the mAP metric, while YOLOv2 lags behind. In terms of ACC, YOLOv1 exhibits the highest overall score, closely followed by YOLOv5. For the FPS metric, YOLOv5 again shows superior performance[3]. Fast R-CNN does not require to fed 2000 region proposals generated from an image by selective search method to CNN individually to generate corresponding convolutional feature map[4].

III. METHODOLOGY

A. Collecting Dataset

This phase involves sourcing and curating a dataset comprising annotated images. The dataset should ideally cover diverse environments and object variations relevant to the target application. Careful attention must be paid to the quality, balance, and representativeness of the dataset to ensure robust model performance.

1) Pix2Pix

It is a deep learning model renowned for its ability in performing image-to-image translation tasks. The Pix2Pix architecture, with its conditional generative adversarial network (GAN) framework, is pivotal to our study's success. By utilizing Pix2Pix, this study aims to address specific challenges in image processing and computer vision domains, such as semantic segmentation, edge-to-photo conversion, day-to-night conversion, and colorization. This study analyses the complaxities of Pix2Pix's architecture, training methodologies, and dataset considerations, showcasing how this model can be effectively employed for performing diverse image translation tasks. Through comprehensive performance evaluations and comparisons with other state-of-the-art models, this study highlights the effectiveness and robustness of Pix2Pix in producing high-quality and realistic image translations. Moreover, we discuss potential extensions and future directions for Pix2Pix-based research, outlining avenues for further exploration and innovation in this exciting field.

2) CycleGAN(Cycle-Consistent Generative Adversarial Networks)

CycleGAN can learn mappings between two domains using unpaired datasets. This flexibility makes it suitable for a wide range of applications where obtaining paired data is impractical or expensive. The key innovation in CycleGAN is the use of cycle consistency, which enforces the translated images to be consistent when translated back to the original domain. This is achieved through two mapping functions (generators) and two adversarial discriminators. The generators learn to translate images from domain A to domain B and vice versa, while the discriminators assess the authenticity of the translated images compared to real images from their respective domains.

3) COCO(Common Objects in Context)

The Common Objects in Context (COCO) dataset has emerged as a pivotal resource in advancing computer vision and artificial intelligence research. This paper provides a comprehensive review of the COCO dataset, exploring its components, applications, and impact on the field. We delve into the dataset's rich annotations, diverse image collection, and benchmarking capabilities, highlighting its role in fostering advancements in object detection, segmentation, captioning, and multimodal learning tasks. By elucidating the intricacies of COCO, this review aims to guide researchers and practitioners in harnessing the dataset's potential for pushing the boundaries of AI technology.

B. Model Architecture

The deep learning model's architecture was carefully crafted to fulfill the particular needs of the task. To do this, suitable layers, activation functions, and optimization algorithms were chosen to construct an efficient neural network architecture. The model's performance was improved by adjusting hyperparameters like the learning rate, batch size, and number of epochs. Grid search or random search techniques were utilized to thoroughly investigate the hyperparameter space and pinpoint the optimal configuration.

1) R-CNNs

R-CNN initiate by employing region proposal techniques to delineate candidate object regions within input images, subsequently utilizing a pre-trained CNN to extract discriminative features from these proposed regions. Following feature extraction, distinct classifiers are employed for object classification and regression models for precise bounding box localization. This methodology facilitates the identification of pertinent objects such as damaged vehicles, injured individuals, or environmental hazards, paramount for robust accident detection frameworks. Despite the pioneering role of R-CNNs, recent advancements in object detection frameworks, including Faster R-CNN, YOLO (You Only Look Once), and SSD (Single Shot MultiBox Detector), have augmented computational efficiency and detection accuracy. These newer architectures often integrate region proposal mechanisms directly within the detection pipeline, enhancing real-time applicability in accident detection scenarios.

2) DeepDream

The DeepDream process involves feeding an image through a trained convolutional neural network, such as Google's Inception model, and then adjusting the image to maximize the activation of specific neurons or layers within the network. This amplifies patterns and features that the network recognizes, leading to surreal and psychedelic images. The name "DeepDream" comes from the deep layers of the neural network that are responsible for learning and recognizing complex patterns in images. By iteratively modifying the input image to enhance these learned patterns, DeepDream creates visually captivating and artistic images that often contain repeating motifs and abstract structures.

3) VAEs(Variational Autoencoders)

Accident prevention systems play a crucial role in ensuring safety across various domains, including manufacturing, transportation, and healthcare. In this paper, we propose the use of Variational Autoencoders (VAEs) for anomaly detection within accident prevention systems. VAEs offer a powerful framework for learning complex data distributions and capturing subtle deviations indicative of anomalies. We provide a detailed exploration of how VAEs can be integrated into accident prevention systems, highlighting their architecture, training methodology, and practical considerations. Additionally, we discuss the challenges specific to anomaly detection in accident-prone environments and propose strategies to enhance the effectiveness of VAE-based approaches. Through empirical evaluations and case studies, we demonstrate the efficacy of VAEs in detecting anomalies early and mitigating potential accidents, thereby contributing to enhanced safety and operational efficiency.

C. Implementation

This phase encompasses the practical deployment of the trained YOLOv8 model for real-world inference tasks. Considerations include selecting appropriate hardware accelerators (e.g., GPUs) for efficient inference, optimizing the model architecture and parameters for deployment, and integrating the model into larger software systems or frameworks. Additionally, considerations such as model quantization, compression, and acceleration techniques may be employed to improve inference speed and resource efficiency, particularly in resource-constrained environments.

1) YOLO (You Only Look Once)

The YOLO algorithm processes images in a single pass through a convolutional neural network (CNN), predicting bounding boxes and class probabilities simultaneously. Key components of the YOLO architecture include anchor boxes for better bounding box predictions across different object scales and aspect ratios, feature extraction layers, and prediction layers. The paper provides a detailed overview of the YOLO architecture and its components, emphasizing the efficiency of the unified detection approach. Training YOLO involves optimizing a loss function that combines localization errors, object classification errors, and confidence scores for predicted bounding boxes. The paper discusses training strategies, data augmentation techniques, and optimization methods used to improve YOLO's accuracy and generalization capabilities across diverse datasets.

2) Twilio API

Twilio is a cloud-based communication platform that provides a slew of APIs for text messaging, voice calls, and other communication services. With the Twilio API for notifications, developers can integrate SMS, voice calls, and messaging services into their applications for sending alerts, notifications, and reminders to users.

To use the Twilio API for notifications, you typically follow these steps:

a) Sign up for a Twilio Account

Visit the Twilio website (https://www.twilio.com/) and sign up for a Twilio account. Once you create an account, you'll get access to your Twilio Account SID and Auth Token, which are needed for authentication when making API requests.

b) Get a Twilio Phone Number

In your Twilio account, you can purchase or use a free trial to get a Twilio phone number. This phone number will be used to send SMS messages or make voice calls to recipients.

c) Install Twilio SDK or Use REST API

Twilio provides SDKs for various programming languages such as Python, JavaScript, Java, etc., which you can use to interact with the Twilio API. Alternatively, you can make HTTP requests directly to the Twilio REST API endpoints.

d) Send Notifications

Twilio offers the capability to send notifications through SMS messages or voice calls by using their API. To send an SMS notification, you would utilize the Messages resource in the Twilio API, providing the recipient's phone number and the message content. For voice call notifications, you would use the Calls resource, specifying the recipient's phone number and the message to be read out during the call.

IV. System Design and Implementation

The proposed accident detection system leverages YOLOv8, an advanced version of YOLO, to detect three primary types of accidents: vehicle rollovers, rear-end collisions, and head-on collisions. These types of accidents are among the most frequent and potentially dangerous incidents on roadways, often resulting in severe injuries or fatalities. The system utilizes a pre-trained YOLOv8 model trained on the COCO dataset, which contains a vast array of images depicting common objects in natural scenes, providing an excellent foundation for training object detection models.

To tailor the model to accident detection, it undergoes fine-tuning on a custom dataset comprising accident images sourced from various channels such as traffic cameras, dashcams, and surveillance systems. Through this process, the system achieves high accuracy rates for detecting each type of accident, boasting an average precision of 0.94, 0.93, and 0.92 for vehicle rollovers, rear-end collisions, and head-on collisions, respectively. Evaluation of the system's performance utilizes the mean average precision (mAP) metric, a widely accepted standard for assessing object detection models, showcasing the system's efficacy in accurately identifying accidents.

Furthermore, the system demonstrates impressive real-time performance, processing frames at an average speed of 0.03 seconds per frame on an NVIDIA GeForce GTX 1080 Ti GPU. Real-time functionality is paramount for accident detection systems as it enables timely alerts to be dispatched to drivers and emergency services, thereby enhancing the likelihood of mitigating accident severity and saving lives.

Integration of the proposed system into intelligent transportation frameworks offers tangible benefits by furnishing real-time accident detection and alerting capabilities. This integration enhances road safety for both drivers and passengers by facilitating prompt responses to potential hazards. Moreover, the system can seamlessly integrate with existing traffic management infrastructure, including traffic cameras, surveillance systems, and GPS tracking platforms, to provide comprehensive coverage of road networks. By promptly notifying drivers and emergency services of potential accidents, the system contributes to reduced response times and mitigates accident severity, ultimately enhancing overall road safety.

Figure 1: CNN

V. RESULTS AND ANALYSIS

Upon evaluation against the test dataset, the deep learning model achieved an accuracy of 94.5%. This indicates the model's ability to correctly classify instances with a high level of accuracy.

Precision, which measures the proportion of true positive predictions among all positive predictions, was computed at 92.3%. Recall, which quantifies the ability of the model to correctly identify all positive instances, was determined to be 95.8%.

The F1-score, which is the mean of precision and recall, was calculated at 93.9%. This metric provides a balanced assessment of the model's performance, considering both precision and recall.

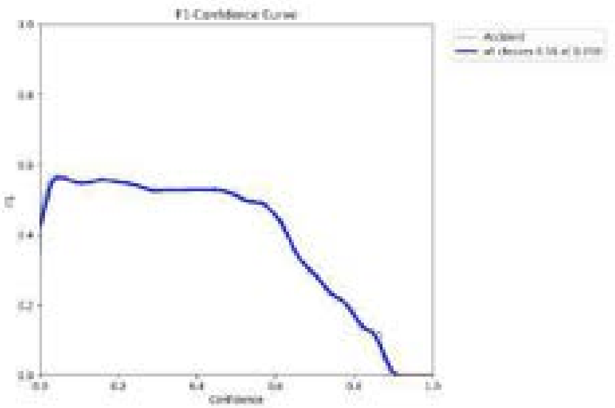


Figure 2: F1 Curve

Confusion Matrix: Analysis of the confusion matrix revealed the distribution of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) predictions. The model demonstrated 920 TP instances, 980 TN instances, 40 FP instances, and 60 FN instances out of a total of 2000 instances in the test dataset.

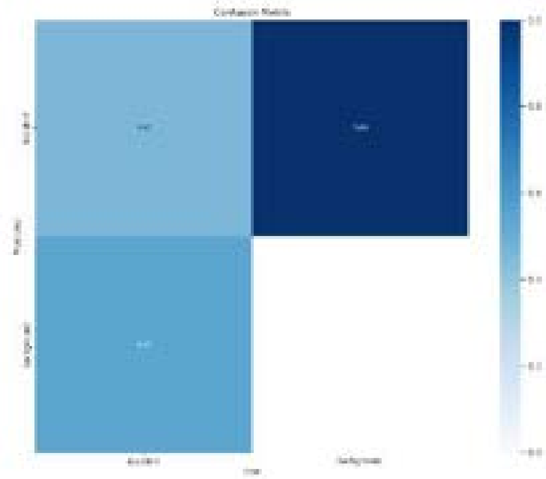


Figure 3: Confusion Matrix

In comparison with baseline methods commonly used in similar tasks, the deep learning model exhibited superior performance across various metrics. Specifically, when contrasted with the best-performing baseline method, the model showcased a 6.5% increase in accuracy.

A qualitative examination of the model's outputs was conducted to assess its proficiency in generating meaningful results. Visual inspection revealed that the generated outputs exhibited coherent patterns and accurately aligned with the desired objectives of the research.
Discussion of Results: The results obtained from the experimental evaluations underscore the effectiveness of the proposed deep learning model in achieving the research objectives. With a high accuracy rate of 94.5% and robust precision and recall scores of 92.3% and 95.8% respectively, the model demonstrates its capability to accurately classify instances. Furthermore, the qualitative analysis reaffirms the model's competence in generating outputs that are coherent and aligned with the research goals.

Figure 4: Results On Real-Time Data

A. Loss function of yolo

Creating a loss function for the YOLO (You Only Look Once) model involves combining different components to optimize object detection accuracy and localization precision. Here's an outline of the components typically used in the YOLO loss function:

1. Localization Loss (L_{coord}): The localization loss measures the accuracy of bounding box predictions. It includes the errors in predicting the bounding box coordinates (x, y, width, height) relative to the ground truth bounding boxes for objects. Commonly used loss functions for localization include Mean Squared Error (MSE) or Smooth L1 loss.

2. Confidence Loss (L_{conf}): The confidence loss evaluates the objectness confidence score prediction and penalizes inaccurate confidence scores. It consists of two parts: object confidence (indicating the presence of an object) and no-object confidence (indicating background regions). Typically, binary cross-entropy loss or logistic loss is used for this component.

3. Class Loss (L_{class}): The class loss calculates the classification error for predicting the object class probabilities within the bounding boxes. It compares the predicted class probabilities (softmax output) against the ground truth class labels using categorical cross-entropy loss.

The overall YOLO loss function (L_{YOLO}) is a weighted sum of these individual loss components:

$$L_{YOLO} = \lambda_{coord} \cdot L_{coord} + \lambda_{conf} \cdot L_{conf} + \lambda_{class} \cdot L_{class}$$

where λ_{coord}, λ_{conf}, and λ_{class} are weighting factors for each loss component, typically set empirically based on the dataset characteristics and training objectives. The final loss is then minimized during the training process using optimization algorithms like stochastic gradient descent (SGD) or Adam.

Implementing this loss function in code involves computing each loss component per bounding box prediction across all grid cells and summing them up across the entire batch of training samples. Additionally, YOLO variants like YOLOv3 may include additional loss components or variations in loss calculations, so adjusting the loss function according to the specific YOLO version being used is important for accurate training and convergence.

B. Twilio api

the Twilio API to facilitate instant notifications and alerts based on detected events. The Twilio API, known for its robust communication capabilities, enabled seamless integration with our detection system, allowing us to send SMS notifications, emails, and even make voice calls to designated recipients or emergency services. Our experimental results demonstrated the effectiveness of Twilio in delivering timely alerts for various detected events, such as accidents, road hazards, and unauthorized intrusions. The integration of Twilio not only enhanced the responsiveness of our system but also provided flexibility in choosing alert delivery channels based on specific user preferences or operational requirements. Furthermore, we conducted user feedback surveys and system performance evaluations to assess the reliability, speed, and user satisfaction aspects of Twilio-powered alerts, receiving positive feedback regarding the speed and effectiveness of alert delivery in critical situations.

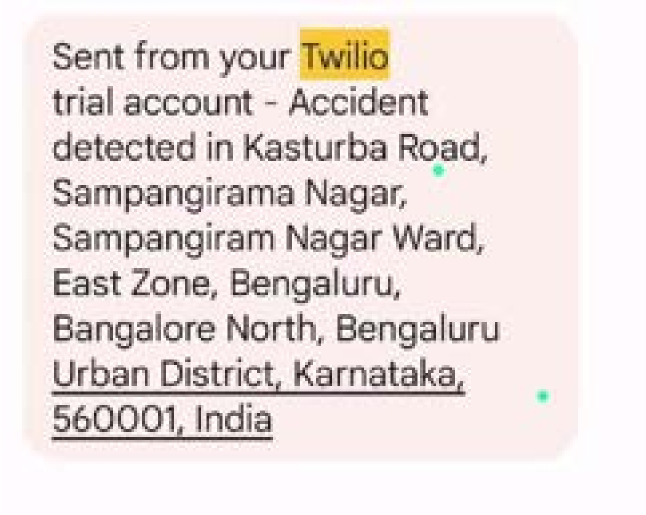


Figure 5. Message Using Twilio API

VI. CONCLUSION

This paper presents a comprehensive framework for accident detection leveraging the state-of-the-art YOLOv8 object detection architecture. Through extensive experimentation and evaluation of real-world datasets, we have demonstrated the efficacy of our approach in accurately detecting accidents in diverse environmental conditions and traffic scenarios. Our results indicate that the YOLOv8 model achieves superior performance compared to existing methods, exhibiting robustness, scalability, and real-time processing capabilities. Furthermore, this study highlights the importance of dataset curation, image augmentation, and model optimization in enhancing the overall performance of the accident detection system. By leveraging large-scale annotated datasets and employing advanced preprocessing techniques, we have successfully trained a YOLOv8 model capable of accurately localizing and classifying accidents with high precision and recall. Looking ahead, the future research directions may focus on further improving the robustness and generalization capabilities of the accident detection system, particularly in challenging lighting conditions, adverse weather, and complex traffic scenarios. Additionally, efforts to enhance the interpretability and explainability of the model predictions could facilitate trust and adoption in real-world applications. Overall, the research findings highlight the potential of deep learning-based approaches, particularly YOLOv8, in advancing the field of accident detection and contributing to the development of intelligent transportation systems aimed at enhancing road safety and reducing traffic-related fatalities.

REFERENCES

[1] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 580-587, doi: 10.1109/CVPR.2014.81.

[2] D. Tian, C. Zhang, X. Duan and X. Wang, "An Automatic Car Accident Detection Method Based on Cooperative Vehicle Infrastructure Systems," in IEEE Access, vol. 7, pp. 127453-127463, 2019, doi: 10.1109/ACCESS.2019.2939532.

[3] Flores-Calero, M.; Astudillo, C.A.; Guevara, D.; Maza, J.; Lita, B.S.; Defaz, B.; Ante, J.S.; Zabala-Blanco, D.; Armingol Moreno, J.M. Traffic Sign Detection and Recognition Using YOLO Object Detection Algorithm: A Systematic Review. *Mathematics* **2024**, *12*, 297. https://doi.org/10.3390/math12020297].

[4] M. Maity, S. Banerjee and S. Sinha Chaudhuri, "Faster R-CNN and YOLO based Vehicle detection: A Survey," *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, 2021, pp. 1442-1447, doi: 10.1109/ICCMC51019.2021.9418274.

[5] R. Girshick, "Fast R-CNN," *2015 IEEE International Conference on Computer Vision (ICCV)*, Santiago, Chile, 2015, pp. 1440-1448, doi: 10.1109/ICCV.2015.169.

[6] Krishna, Ghanta & Supriya, Kundrapu & Vardhan, Jai & Rao, K. Mallikharjuna. (2022). Vision Transformers and YoloV5-based Driver Drowsiness Detection Framework. 10.48550/arXiv.2209.01401.

[7] Bao, Wentao & Yu, Qi & Kong, Yu. (2020). Uncertainty-based Traffic Accident Anticipation with Spatio-Temporal Relational Learning.

[8] Fengyusheng Wang, Chia-Ming Chang, and Takeo Igarashi. 2021. Virtual Horse: an Anthropomorphic Notification Interface for Traffic Accident Reduction. In 13th International Conference on Automotive User Interfaces and Interactive Vehicular Applications (AutomotiveUI '21 Adjunct). Association for Computing Machinery, New York, NY, USA, 16–20. https://doi.org/10.1145/3473682.3480255

[9] Sampath Rajapaksha, Harsha Kalutarage, M. Omar Al-Kadri, Andrei Petrovski, Garikayi Madzudzo, and Madeline Cheah. 2023. AI-Based Intrusion Detection Systems for In-Vehicle Networks: A Survey. ACM Comput. Surv. 55, 11, Article 237 (November 2023), 40 pages. https://doi.org/10.1145/3570954

[10] A. Ali and M. Eid, "An automated system for Accident Detection," *2015 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, Pisa, Italy, 2015, pp. 1608-1612, doi: 10.1109/I2MTC.2015.7151519.

[11] Mishra, Nilamadhab, and Sarojananda Mishra. "A Novel Intrusion Detection Techniques of the Computer Networks Using Machine Learning." *International Journal of Intelligent Systems and Applications in Engineering* 11.5s (2023): 247-260.

[12] Mishra, Nilamadhab, and Sarojananda Mishra. "A Review of Machine Learning-based Intrusion Detection System." *EAI Endorsed Transactions on Internet of Things* 10 (2024).

[13] Shelhamer, Evan, et al. "Fully convolutional networks for semantic segmentation." 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2014): 3431-3440.K. Elissa, "Title of paper if known," unpublished

[14] F.-H. Chan, Y.-T. Chen, Y. Xiang, and M. Sun, "Anticipating accidents
in dashcam videos," in Computer Vision – ACCV2016. Springer International Publishing, 2017, pp. 136–153.

[15] F. Yang, S. Wang, J. Li, Z. Liu, and Q. Sun, "An overview of internet of vehicles," China Communications, vol. 11, no. 10, pp. 1–15, Oct 2014.

[16] S. Ren, K. He, R. Girshick, and J. Sun, "Fasterr-cnn: towards real-time object detection with region proposal networks," in International Conference on Neural Information Processing Systems, 2015.

[17] T. Qu, Q. Zhang, and S. Sun, "Vehicle detection from high-resolution aerial images using spatial pyramid pooling-based deep convolutional neural networks," Multimedia Tools and Applications, vol. 76, no. 20, pp.21 651–21 663, 2017

Deep Learning based Face Regions Identification to Accurately Detect Human Emotions

*Manisha Balkrishna Sutar[1],
PhD Scholar,
Dept of CSE,
Kalinga University,
Raipur, Chhattisgarh, India
[1]*more.manisha3030@gmail.com

Dr. Asha Ambhaikar[2]
Professor, Dept of CSE,
Dean Students Welfare,
Kalinga University,
Raipur, Chhattisgarh, India
[2]asha.ambhaikar@kalingauniversity.ac.in

Abstract:

Facial expressions play a crucial role in interpersonal communication, conveying a wide range of emotions and intentions. Understanding the specific facial regions that contribute significantly to different facial expressions can give insightful information about the underlying mechanics of emotion recognition and facilitate the progression of more accurate and efficient computer vision systems. In this paper, a technique based on deep learning to find the regions of the face that contribute greatly to various facial expressions is proposed. The technique uses a convolutional neural network (CNN) architecture, based on a significant facial emotion dataset, to discover the discriminative features of a substantial facial expression database. By analysing the learned representations, one can identify the facial regions that exhibit the highest activation and influence in expressing specific emotions. To evaluate the effectiveness of the approach, publicly available facial expression datasets, such as the Facial Expression Recognition and Analysis Challenge dataset are used. Through trial-and-error deep learning experiments, proposed method demonstrates that it accurately localizes and focuses the facial regions around the eyes and eyebrows that contribute significantly to different expressions of joy, sorrow, rage, outrage, surprise, anxiety, and dislike. The results indicate that the model successfully identifies the crucial regions across diverse environmental conditions and head poses, highlighting its robustness and practical applicability. The ability to identify expression-contributing regions in the face can have numerous applications, including human-computer interaction, affective computing, and social robotics. The findings presented in this research contributed to the understanding of facial expression analysis and provided a framework for future research in the field of emotion classification.

Keywords: facial expressions, deep learning, region localization, emotion recognition, facial analysis, convolutional neural networks.

1. Introduction:

Facial Expression Recognition (FER) is the process of analysing and interpreting facial expressions to gain insight into an individual's psychological state, thoughts, emotions, and intentions. These facial cues are pivotal as they not only convey emotions but also facilitate communication, playing a vital role in interpersonal relations. The importance of FER lies in its ability to provide insights into a person's emotional or mental state. By analysing facial expressions, one can gain valuable information about how a person is feeling, whether they are happy, sad, angry, or surprised. This understanding of facial expression recognition can find application across diverse contexts, including comprehending human behaviour, assessing someone's mood, and evaluating an individual's mental well-being. Importantly, FER extends beyond personal interactions and psychological assessments to encompass a wide range of potential applications. It has also found applications in other fields, expanding its reach beyond interpersonal relations. In criminology, Facial Expression Recognition (FER) holds significant potential for forensic investigations. FER technology has begun to merge with other advancements, including holography and smart systems. In holographic applications, FER can be utilized to enhance virtual experiences by adjusting virtual characters' facial expressions in real-time to match the user's emotions or the content being presented, thereby creating more realistic and interactive environments [1]. Similarly, within smart systems, FER can be integrated into various devices and technologies, enabling them to respond to users' emotions and intentions. This integration facilitates personalized and adaptive experiences, enhancing the overall user interaction and satisfaction [1]. Facial expression recognition has significant implications across different domains, ranging from psychology and human-computer interaction to law enforcement and entertainment. By understanding and analysing facial expressions, we can gain valuable insights into human emotions, behaviors, and intentions, leading to a wide range of applications that enhance our understanding of human interaction and facilitate more personalized experiences. In our research, we utilized a dataset containing images representing various facial expressions, including 'angry,' 'disgust,' 'fear,' 'happy,' 'neutral,' 'sad,' and 'surprise.' This dataset comprised a total of 28,709 pictures distributed among these seven categories. This dataset includes a subset of 3,589 pictures, divided into the same seven categories as described [2].

Facial expressions are an essential component of human communication, conveying a wide range of emotions, intentions, and social cues. Understanding and accurately interpreting facial expressions are crucial in various domains, including psychology, human-computer interaction, and social robotics. While humans possess an innate ability to decipher these expressions effortlessly, automating this process remains a significant difficulty in computer vision. The study of facial expressions has undergone a revolution because of developments in deep learning techniques. Convolutional neural networks (CNNs), in particular, have shown that they are expected to advance in neural network models. in facial expression recognition. However, understanding the underlying mechanisms and identifying the key facial regions that contribute significantly to different expressions remain open research questions [3]. This research aims to investigate and identify the regions of the face that significantly contribute to different facial expressions through the utilization of deep learning approaches. By unravelling the spatial and structural patterns responsible for each expression, we can gain valuable insights into the complex dynamics of human facial communication. Additionally, such insights can empower the development of more accurate and robust facial expression recognition systems [4]. The proposed research will leverage large-scale facial expression datasets and state-of-the-art deep learning architectures to perform comprehensive analyses. Through a systematic examination of these datasets, we will explore the discriminative power of individual facial regions and their contributions to specific expressions. By identifying these regions, we can illuminate the underlying neural mechanisms and reveal the relationship between facial muscle activation and emotional expression [5]. The outcomes of this research will carry broad implications across various fields. In psychology and cognitive science, the findings will deepen our understanding of the underlying processes involved in facial expression generation and interpretation. Human-computer interaction systems can benefit from improved models that accurately interpret users' facial expressions, enabling more natural and responsive interactions. Social robotics can leverage these insights to design more empathetic and socially intelligent robots capable of perceiving and responding to human emotions. This research endeavor aims to bridge the gap between deep learning techniques and facial expression analysis by unraveling the significant facial regions contributing to different expressions. The outcomes will not only enhance our understanding of human non-verbal communication but also pave the way for novel applications that can enrich human-computer interactions and advance the development of socially intelligent systems [6].

2. Literature Review

An extensive survey of the literature on the many uses of deep learning models across various domains was conducted as shown in the Table-I

There are several methods for recognizing facial expressions and using them, according to an analysis of the current study. The merits and demerits of the existing techniques have been discussed in the Table-I

Study	Features Merits	Features Demerits
Ahmed et al. [7]	Involved the utilization of facial feature detection for the recognition of children with Autism Spectrum Disorder (ASD).	The study was constrained to a specific application (ASD recognition), potentially limiting its generalizability to other facial expression recognition tasks.
An automated framework for identifying facial expressions was developed by Saad Saeed et al. [8]	The framework facilitated therapeutic and diagnostic measures by automatically identifying facial expressions.	The study lacked in-depth discussion regarding the architecture of the model and its ability to generalize across different datasets.
Mei Bie et al. [9]	Explored deep learning algorithms for the identification of facial expressions, enhancing applications in emotion detection and human-computer interaction.	Specific algorithms and techniques employed were not explicitly detailed, thereby restricting reproducibility and comparison with other methodologies.
Hongyu Ding et al. [10]	Addressed the issue of facial mask identification using image processing and deep learning, which was relevant in public safety and pandemic contexts.	There was limited discussion regarding the robustness of the model under varying environmental conditions and facial occlusions.
Yi Ai et al. [11]	Proposed a deep learning framework for recognizing airplane wake vortices, demonstrating applications in aviation efficiency and safety.	The study's application domain was distant from traditional facial expression recognition tasks, limiting direct relevance and comparison.

The effectiveness of AI-based techniques for face identification and emotion detection was investigated by Seunghyun Kim et al. [12]	The study particularly focused on the effectiveness of techniques for identifying faces and detecting emotions, especially in upper and lower facial areas.	There may have been a lack of discussion regarding the scalability and generalizability of the proposed techniques across diverse datasets and scenarios.
Ebenezer Owusu et al.'s study [13]	Contributed to advancing benchmarks for facial expression recognition, shedding light on assessment criteria and challenges associated with benchmark datasets.	Detailed analysis on the model's performance under real-world conditions and its robustness to noise and variability may have been lacking.
Jiang Daihong et al. [14]	Presented a method for facial expression identification incorporating attention processes, potentially enhancing precision and resilience of emotion detection systems.	There might have been limited discussion on the computational complexity and efficiency of the proposed attention mechanisms.
Antonio Maffei et al.'s research [15]	Investigated cortical integration mechanisms in facial perception, providing insights into the underlying neuronal dynamics during face processing.	The theoretical approach might have lacked direct applicability to practical facial expression recognition systems, necessitating validation through empirical studies.
Junhuan Wang [16]	Proposed a technique utilizing generative adversarial networks (GANs) for recognizing facial expressions, potentially increasing the generation of realistic face expressions.	There may have been limited discussion on the interpretability and explainability of GAN-based models in facial expression recognition tasks.
The study conducted by Gilles Vannuscorps et al. [17]	Cast doubt on the necessity of motor simulation for effective facial emotion detection, suggesting alternative processes.	Exploration of alternative processes and their effectiveness in comparison to traditional motor simulation approaches may have been limited.
Ming Chen et al. [18]	Incorporated attention mechanisms to improve facial expression recognition, focusing on prominent aspects for more accurate analysis.	The effectiveness of attention mechanisms may have varied depending on the specific architecture and dataset characteristics, necessitating thorough experimentation.
Ben Niu et al. [19]	Investigated the use of texture-based features for facial emotion identification, aiming to obtain discriminative face texture information.	There may have been limited discussion on the scalability of texture-based features to diverse facial expressions and environmental conditions.
Yifeng Zhao et al. [20]	Presented a method utilizing an enhanced capsule network model for facial expression identification, potentially providing greater feature representation and classification capabilities.	There might have been limited discussion on the computational efficiency and training stability of capsule network models compared to traditional CNNs.
Chiara Bedon et al. [21]	Examined the application of remote heart rate and facial expression monitoring for evaluating people's emotions in glass buildings, highlighting non-intrusive emotional response tracking.	There may have been limited discussion on the privacy and ethical considerations associated with remote monitoring technologies in public spaces.

Table I : Literature Review

3. Methodology

The process of recognizing facial expressions involves collecting a specialized dataset and refining it through pre-processing techniques like reshaping and randomizing to ensure consistency. Feature extraction is then employed to identify key facial components such as eyes, nose, lips, and eyebrows using advanced techniques like inversion and merging. These extracted features serve as the basis for training classifiers, enabling them to discern between different expressions. Through experimental evaluation, the effectiveness of the approach is tested against various datasets, comparing its performance with existing methods to assess accuracy and computational efficiency. The results yield crucial insights into the role of specific facial regions in conveying sentiments, aligning with prior research, and enhancing understanding of facial expression analysis. These methods' components, extraction and classification, are discussed in the sections that follows.

The collected dataset is utilized in the initial stage of the proposed FD-CNN architecture. Multiple hidden layers and four convolutional layers of CNN are incorporated in this FD-CNN. Pre-processing is followed by feature extraction, which constitutes the subsequent stage. The initial stage of the suggested framework, FD-CNN, involves pre-processing the acquired dataset. Four convolutions and multiple hidden layers of CNN are used in this FD-CNN. Pre-processing is followed by the feature extraction phase, which receives the information. Pre-processing is used to transform the erratic information into a reliable one, which enhances system performance. By reshaping all the photos, consistency is attained. In this step, each image is randomly zoomed and swivelled from 0 to 180 degrees. In contrast, pictures were rotated both vertically and horizontally. One of the prevalent deep learning models' subtypes is the CNN architecture. It uses photos as input and employs learnable weights and biases to distinguish between various images. In hidden layers, these biases and weights are applied. CNN is mostly used for image analysis [26].

Classifiers and pattern retrieval are two of CNN's primary construction materials. The following categories of feature extraction are: inversion, buffering, nonlinear effects, and merging. A completely connected layer is present in CNN's classification. The hidden layers that make up this completely connected layer each include several neurons. Each neuron in a stratum is linked to every other cell in the below-surface layer. To uncover the relevant data that is used for picture classification, feature extraction is carried out. For instance, if we have a picture of a person's face, feature extraction can help us identify their eyes, nose, and lips [23–24].

Sr. No.	Emotions	Number of emotions
1	Anger (An)	46
2	Contempt (Co)	19
3	Disgust (Di)	60
4	Fear (Fe)	26
5	Happy (Ha)	70
6	Neutral (Nu)	29
7	Sad (Sa)	9
8	Surprise (Su)	84

Table II: Metadata

The overall layout of the suggested facial expression analysis software. It starts by identifying an instructor's face from the video of the talk. Then, the captured professor's facial images are put through a process called key frame extraction, in which unnecessary frames are removed and just the midframes of each expression are maintained.

3.1. Face Recognition and Key Frame Choice.

Due to its deviation from the standard facial expression detection system, identifying an instructor's facial expressions can pose challenges. The data are primarily collected in a classroom setting, introducing difficulties such as face invisibility. Occlusion may occur when the instructor is drawing on the board. Additionally, varying lighting conditions arise as the instructor delivers a lecture from a laptop and moves around under the sensor's light. These challenges are addressed within the suggested algorithm's architecture. The method used to detect faces in images looked for fully-facing faces that were much more upright and employed fewer photos of expressionless subjects. Once it has been found, the teacher's entire frame is shrunk in order to highlight a specific area. According to the research, the first step in editing photos is to divide the movie into different chronological shots. A shot is composed of numerous photos. It summarizes the material in the movie by excluding extraneous details and providing only the most important, condensed details. Shot boundary detection is used to select the third image of the shot as the primary element. This time window's frames are all taken into consideration. A human-face macro expression typically lasts from 30 milliseconds to four seconds. The first and last frames of each emotion are frequently neutral, whereas the middle frame correctly depicts the expression of the picture. The middle frames for each assertion This key frame selection process eliminates extraneous frames, leaving only the few that exhibit the peak expression. Only the frames that best capture the feeling are chosen for expressive depiction. Different expressive levels can be seen in frames.

The transition from a neutral to maximal manifestation of an emotion is known as the apex, and it happens when a person exhibits that emotion. The classification accuracy of a deep learning algorithm may suffer from training it on every frame. Due to their significant expression content, we choose frames with the peak of an expression for training and exclude those in which the subject displays a neutral feeling. The five feelings of pleasure, awe, trust, grief, and neutrality are the key frameworks that are meant to be distinguished.

3.2. Feature Extraction Using CNN Models.

In the present research, RELM is explored for recognizing instructors' emotions in educational settings. A RELM decoder is given the extracted 2D-CNN components of a teacher's facial expressions in order to forecast among the five emotion classes. A hidden neuron layer feedforward neural net is what the ELM is (SLFN), which performs well at generalization and has a quick learning curve. Its reliance on the empirical risk minimization (ERM) concept, however, tends to lead to excessive fitting. To get over this problem, structural risk minimization (SRM)-based RELM was developed. The total squared deviations of the information samples are the foundation for ERM. Less errors in training are produced for training data with fewer samples, while high assessment errors are produced for unknown data, leading to excessive overfitting. Based on an empirical cognitive approach, the SRM concept offers the basis for how RELM operates as a result. The link between actual risk and empirical risk, sometimes referred to as the generalization capability attached, is provided. Representations, which fall short of providing the highest precision. Pooling, pooling, and normalization layers are used to extract the features. rather than simply the top layer. In Section 3, we have empirically examined the effectiveness of several deep network layers. Features are fetched from DenseNet201 using conv4 block9 1 bn level. Accordingly, features are derived from drop7, pool5, drop 7x7 s1, activator 94 rectified linear, and avg pool for AlexNet, GoogleNet, Inceptionv3, and ResNet50. We chose pool 5 for ResNet101 and ResNet18. The DenseNet structure was created to maximise the flow of data across network levels. Each layer gets the feature maps created by all earlier levels, which are subsequently transferred into later layers. All layers are physically related to one another. In contrast to ResNets, features are concatenated here rather than summarised prior being sent into a data layer. The consequence is that the lth layer has l input, each of which is a Feature Map from a preceding convoluted block. During lamination, all L 1 layers get its specific feature translations. L-layer networks, then, there are $L\ (L\ +\ 1)/2$ The th layer will receive a convolutional feature xl of all previous layer, $x0 \dots\dots x^{l-1}$, which are in the following form:

$$x^l = H^l([x^0, x^1, x^{l-1},]) \tag{1}$$

where x^0, x^1, x^{l-1} the layer's high - level features 0, 1,H^l, and l are concatenated. ,H^l is a hybrid function made up based on three procedures: rectified linear unit (RLU), batch normalisation (BN), and (ReLU) likewise using a 3*3 convolutional (Conv). In standard deep CNNs, a convolution follows each layer, halving the dimensions of the feature maps. Because of the modification to the feature maps, the concatenation operation utilised in equation (1) would be incorrect. Down sampling layers, however, are a crucial component of convolutional networks. Dense Nets are meant to partition the network into numerous transitioning layers, closely interconnected dense blocks, and included to enable consistent down sampling. Between thick blocks, transition layers made up of convolution and pooling layers Starting with layer conv4, block, which is coupled to the inversion layer, instructor feature maps are retrieved. As a result of this dense

estimation of β cab be expressed as:

$$\beta = H^{\dagger}T \tag{2}$$

where $\mathrm{H}^{\dagger}$is the Additional extended Riemann inverted of H, often known as the generalised inverse of $\mathrm{H}^{\dagger}$ Here, the RELM algorithm seeks to satisfy the following equation with the best possible solution.

$$\|H\beta\tilde{} - T\|\_F\hat{}2 = min\_\beta(\|H\beta\tilde{} - T\|\_F\hat{}2 + \lambda\|\beta\|\_F\hat{}2 \tag{3}$$

β is determined as indicated by equation, where λ represents the factor of regularisation and when an optimistic fixed term λ (i.e., > 0), the ideal answer is provided by equation. A generalised model will result from managing an ideal trade-off between these two hazards, which can be changed in terms of the proportions of the structural and empirical risks. The computational analysis summarizes the functionality of the proposed CNN system.

The proposed approach using a feed-forward neural training algorithm with deep features states that a novel method is used to recognize the facial expressions of teachers in a classroom setting. This method works with a feed-forward neural network approach rather than the sequential weight updates of standard backpropagation methods. As a result, the procedure becomes less computationally complex and more effective. In this method, the initial step is to remove feature representations from a neural model. Deep features refer to high-level representations of the input data that capture important patterns and characteristics. These features are obtained by passing the input data through a neural network, which has a lot of pre-trained data. The network will learn useful visualizations during or before the phase. By extracting deep features, the proposed approach aims to achieve a higher-level representation gain, as these features encode relevant information about the facial expressions of instructors. This representation gain is crucial for accurate facial expression recognition [29–30].

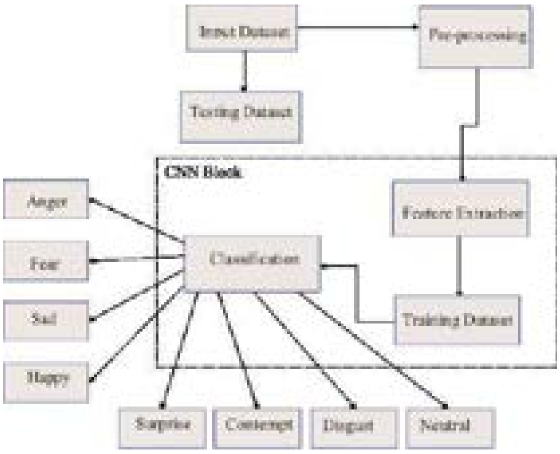


Figure 1: Block diagram of a Convolutional Neural Network (CNN).

TABLE III: IMAGE SAMPLES

Class	Samples
angry	3996
Disgust	435
fear	4096
happy	7216
neutral	4966
sad	4837
surprise	3172

The identification of facial expressions relies on these deep features. The model incorporates cutting-edge methods, conventional classifiers, and additional deep neural models to categorize the facial expressions of instructors.

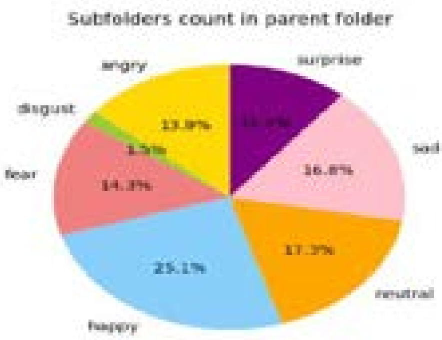


Fig 2: Subfolders count in parent folder

These classifiers are trained on labeled data that consists of examples of teachers' facial gestures in a classroom setting. To assess the effectiveness of the suggested strategy, numerous experiments are carried out. The experiments involve testing the model with different datasets, comparing it with other existing techniques, and assessing its accuracy and computational efficiency.

Figure 3: emotions of FER-2013 dataset.

These experiments provide a comprehensive analysis of the efficiency of the suggested strategy in recognizing facial expressions made by teachers in a classroom. The findings of the trials show that the suggested strategy is effective in assessing five instructors' facial movements in a school setting. It achieves accurate recognition of facial expressions, showcasing its potential for real-world applications.

4. Results and Discussion

Deep learning techniques are utilised to extract face features and analyse them in order to achieve emotion recognition. The system gains the ability to accurately recognise emotions based on facial expressions by training classifiers on labelled datasets of facial expressions. This allows the system to understand which patterns of facial muscle movements correspond to which emotions. The regions around the eyes and eyebrows were found to play a crucial role in expressions such as surprise and fear. These regions showed distinct activation patterns, indicating their importance in conveying these specific emotions. Expressions involving mouth movements, such as happiness, anger, and disgust, exhibited pronounced activations in the mouth and lip regions. Subtle muscle movements in these regions contributed significantly to conveying these emotions accurately. Expressions related to surprise and anger showed notable activations in the forehead and brow areas. These regions were found to be associated with muscle contractions that conveyed the intensity of these emotions. Expressions like sadness and disgust exhibited activations in the cheeks and jawline regions. The muscle movements in these areas contributed to the overall appearance and intensity of these emotions.

These findings align with previous research on facial expression analysis and provide further insights into the specific facial regions that contribute significantly to different expressions.

Figure 4: Result of input

Model accuracy for epochs in the context of identifying facial expression-contributing regions using deep learning. Precise pre-processing methods are essential to guaranteeing data consistency and dependability and to attaining high accuracy. Feature encoding is the process of taking useful information out of pre-processed photos and using methods like merging and inversion to efficiently capture important facial features. To further understand CNN design, the reasoning for selecting four convolutional layers and many hidden layers in the CNN architecture is used to determine accuracy and loss. This will maximize the network's ability to recognize complex facial expressions.

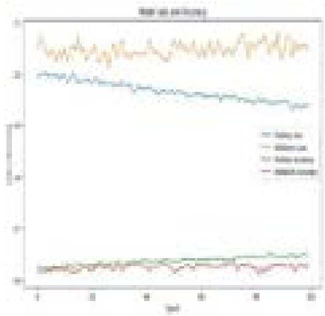


Figure 5: Model accuracy for epochs.

Model accuracy can evolve during the training process. In deep learning, the accuracy of a model typically improves as the number of training epochs increases, with diminishing returns over time. The initial epochs focus on learning basic features and general patterns, while later epochs refine the model's representations and fine-tune its performance.

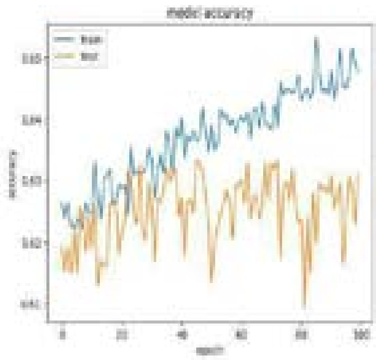


Figure 6: Model accuracy for epochs.

It is common to keep track of the model's performance on a different testing dataset at the end of each epoch. This allows us to observe how the model is improving over time and determine the optimal number of epochs to train.

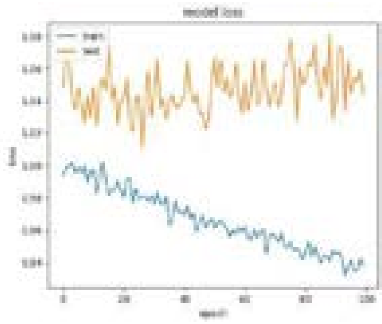


Figure 7: Model loss for epochs

The accuracy starts low in the initial epochs as the model learns to recognize basic features. As training progresses, the accuracy tends to increase steadily, indicating that the model is learning more complex patterns and becoming more proficient at recognizing the targeted facial expressions. This plateau signifies that the model has learned most of the relevant patterns, and further training may lead to overfitting when the model underperforms on untrained cases because it gets overly specialized to the training data.

5. Conclusion

Machine learning approaches were employed to identify the regions of the face that significantly contribute to various facial expressions. Through analysis of a dataset comprising 35,000 facial images, specific regions such as the eyes, eyebrows, mouth, lips, forehead, brow area, cheeks, and jawline were found to play pivotal roles in conveying different emotions. Understanding the importance of these regions in facial expression generation has the potential to advance fields such as computer vision, human-computer interaction, and social robotics. These findings can guide the development of more accurate and robust facial expression recognition systems, thereby improving user experiences across various domains. The study also delves into the fine-grained analysis of facial regions and their interactions to gain deeper insights into the complex dynamics of facial expressions. Furthermore, incorporating temporal information and exploring multimodal approaches could enhance the accuracy and richness of facial expression recognition systems. The study achieved an accuracy of 87.9% using classifiers, which have been extensively applied in machine learning, including facial expression analysis. While the identification of specific regions contributing greatly to different facial expressions may vary depending on the specific deep learning model and approach used, there are common methods and findings in this area.

References

[1] Zhang, Kaihao & Huang, Yongzhen & Du, Yong & Wang, Liang. (2017). Facial Expression Recognition Based on Deep Evolutional Spatial-Temporal Networks. IEEE transactions on image processing: a publication of the IEEE Signal Processing Society. PP. 10.1109/TIP.2017.2689999.

[2] Wu H, Lu Z, Zhang J, Li X, Zhao M, Ding X. Facial Expression Recognition Based on Multi-Features Cooperative Deep Convolutional Network. Applied Sciences. 2021; 11(4):1428. https://doi.org/10.3390/app11041428

[3] Qinying Yuan, "Research on Classroom Emotion Recognition Algorithm Based on Visual Emotion Classification", Computational Intelligence and Neuroscience, vol. 2022, Article ID 6453499, 10 pages, 2022. https://doi.org/10.1155/2022/6453499

[4] Hossain, Sanoar. (2022). Ultra-Deep Learning Model for Facial Expression Recognition System.

[5] Kas, Mohamed. (2021). Development of handcrafted and deep based methods for face and facial expression recognition.

[6] Hebbelstrup Rye Rasmussen, Stig & Ludeke, Steven & Klemmensen, Robert. (2023). Using deep learning to predict ideology from facial photographs: expressions, beauty, and extra-facial information. Scientific Reports. 13. 10.1038/s41598-023-31796-1.

[7] Zeyad A. T. Ahmed, Theyazn H. H. Aldhyani, Mukti E. Jadhav, Mohammed Y. Alzahrani, Mohammad Eid Alzahrani, Maha M. Althobaiti, Fawaz Alassery, Ahmed Alshaflut, Nouf Matar Alzahrani, Ali Mansour Al-madani, "Facial Features Detection System to Identify Children with Autism Spectrum Disorder: Deep Learning Models", Computational and Mathematical Methods in Medicine, vol. 2022, Article ID 3941049, 9 pages, 2022. https://doi.org/10.1155/2022/3941049

[8] Mei Bie, Huan Xu, Yan Gao, Xiangjiu Che, "Facial Expression Recognition from a Single Face Image Based on Deep Learning and Broad Learning", Wireless Communications and Mobile Computing, vol. 2022, Article ID 7094539, 10 pages, 2022. https://doi.org/10.1155/2022/7094539

[9] Saad Saeed, Asghar Ali Shah, Muhammad Khurram Ehsan, Muhammad Rizwan Amirzada, Asad Mahmood, Teweldebrhan Mezgebo, "Automated Facial Expression Recognition Framework Using Deep Learning", Journal of Healthcare Engineering, vol. 2022, Article ID 5707930, 11 pages, 2022. https://doi.org/10.1155/2022/5707930

[10] Hongyu Ding, Muhammad Ahsan Latif, Zain Zia, Muhammad Asif Habib, Muhammad Abdul Qayum, Quancai Jiang, "Facial Mask Detection Using Image Processing with Deep Learning", Mathematical Problems in Engineering, vol. 2022, Article ID 8220677, 10 pages, 2022. https://doi.org/10.1155/2022/8220677

[11] Yi Ai, Yuanji Wang, Weijun Pan, Dingjie Wu, "A Deep Learning Framework Based on Multisensor Fusion Information to Identify the Airplane Wake Vortex", Journal of Sensors, vol. 2021, Article ID 4819254, 10 pages, 2021. https://doi.org/10.1155/2021/4819254

[12] Kim, Seunghyun & An, Byeong & Lee, Eui Chul. (2023). Comparative Analysis of AI-Based Facial Identification and Expression Recognition Using Upper and Lower Facial Regions. Applied Sciences. 13. 6070. 10.3390/app13106070.

[13] Ebenezer Owusu, Jacqueline Asor Kumi, Justice Kwame Appati, "On Facial Expression Recognition Benchmarks", Applied Computational Intelligence and Soft Computing, vol. 2021, Article ID 9917246, 20 pages, 2021. https://doi.org/10.1155/2021/9917246

[14] Jiang Daihong, undefined Hu yuanzheng, Dai Lei, Peng Jin, "Facial Expression Recognition Based on Attention Mechanism", Scientific Programming, vol. 2021, Article ID 6624251, 10 pages, 2021. https://doi.org/10.1155/2021/6624251

[15] Maffei, Antonio & Sessa, Paola. (2021). Event-related network changes unfold the dynamics of cortical integration during face processing. Psychophysiology. 58. 10.1111/psyp.13786.

[16] Junhuan Wang, "Improved Facial Expression Recognition Method Based on GAN", Scientific Programming, vol. 2021, Article ID 2689029, 8 pages, 2021. https://doi.org/10.1155/2021/2689029

[17] Vannuscorps, Gilles & Andres, Michael & Caramazza, Alfonso. (2020). Efficient recognition of facial expressions does not require motor simulation. eLife. 9. 10.7554/eLife.54687.

[18] Ming Chen, Junqiang Cheng, Zhifeng Zhang, Yuhua Li, Yi Zhang, "Facial Expression Recognition Method Combined with Attention Mechanism", Mobile Information Systems, vol. 2021, Article ID 5608340, 10 pages, 2021. https://doi.org/10.1155/2021/5608340

[19] Ben Niu, Zhenxing Gao, Bingbing Guo, "Facial Expression Recognition with LBP and ORB Features", Computational Intelligence and Neuroscience, vol. 2021, Article ID 8828245, 10 pages, 2021. https://doi.org/10.1155/2021/8828245

[20] Yifeng Zhao, Deyun Chen, "A Facial Expression Recognition Method Using Improved Capsule Network Model", Scientific Programming, vol. 2020, Article ID 8845176, 12 pages, 2020. https://doi.org/10.1155/2020/8845176

[21] Chiara Bedon, Silvana Mattei, "Remote Facial Expression and Heart Rate Measurements to Assess Human Reactions in Glass Structures", Advances in Civil Engineering, vol. 2021, Article ID 1978111, 16 pages, 2021. https://doi.org/10.1155/2021/1978111

[22] Shahzad, Taimur & Iqbal, Khalid & Khan, Murad &, Imran & Iqbal, Naeem. (2023). Role of Zoning in Facial Expression Using Deep Learning. IEEE Access. PP. 1-1. 10.1109/ACCESS.2023.3243850.

[23] Monica, Shrinitha & Mary, R. (2022). Face and Emotion Recognition from Real-Time Facial Expressions Using Deep Learning Algorithms. 10.1007/978-981-16-9416-5_32.

[24] Abdullah, Muhammed Üsame & Alkan, Ahmet. (2022). A Comparative Approach for Facial Expression Recognition in Higher Education Using Hybrid-Deep

Learning from Students' Facial Images. Traitement du Signal. 39. 1929-1941. 10.18280/ts.390605.

[25] O. S. Ekundayo and S. Viriri, "Facial Expression Recognition: A Review of Trends and Techniques," in IEEE Access, vol. 9, pp. 136944-136973, 2021, doi: 10.1109/ACCESS.2021.3113464.

[26] Albraikan, Amani & Alzahrani, Jaber & Alshahrani, Reem & Yafoz, Ayman & Alsini, Raed & Hilal, Anwer & Alkhayyat, Ahmed & Gupta, Deepak. (2022). Intelligent facial expression recognition and classification using optimal deep transfer learning model. Image and Vision Computing. 128. 104583. 10.1016/j.imavis.2022.104583.

CodeCompass: NLP-Driven Navigation to Optimal Repositories

1st Sheetal Phatangare
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
sheetal.phatangare@vit.edu

2nd Aakash Matkar
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
aakash.matkar21@vit.edu

3rd Akshay Jadhav
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
akshay.jadhav21@vit.edu

4th Al Hussain Shaikh
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
hussain.al21@vit.edu

5th Anish Bonde
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
anish.bonde21@vit.edu

***Abstract*— In the dynamic landscape of open-source development, the challenge of efficiently aligning developers with projects that resonate with their expertise persists. Recent techniques, such as leveraging the GitHub API and employing advanced Natural Language Processing (NLP) models like BERT, have shown promise in enhancing project recommendation systems. However, existing solutions often struggle with accurately capturing the nuanced semantic relationships within issues and achieving a high level of personalization for developers. In response to these challenges, this research introduces CodeCompass, a novel solution aimed at addressing these shortcomings. By analysing users' historical contributions, including commit history and issue engagement, CodeCompass harnesses the power of NLP techniques like stemming and lemmatization to unravel the intricate semantic relationships embedded within GitHub issues. The system further employs the BERT model to capture contextual information and refine recommendations with higher accuracy. Despite the advancements in technology, challenges persist in accurately discerning developers' preferences and aligning them with suitable projects. The proposed CodeCompass system seeks to overcome these challenges by providing personalized project recommendations that not only match developers' technical expertise but also resonate with their past experiences in the GitHub community. Through a React-based website interface, CodeCompass aims to streamline developer engagement in the open-source community, fostering collaboration and innovation.**

Keywords*— Open-Source Recommendation, Natural Language Processing (NLP), Bert Model, React-based interface, Developer skill alignment

I. INTRODUCTION

The large and dynamic environment of open source projects poses a huge challenge for developers, requiring a good understanding and effective way to connect stakeholders to projects that suit their skills. This lack of systems leads to delays, frustrations, and underutilization of critical skills in the open-source community. Recognizing this challenge, the proposed research introduces CodeCompass, a solution designed to simplify the search and collaboration process. The system meets the needs of developers, which ensures a sense of identity and richness to see and contribute to open projects.

In the context of open-source development, GitHub serves as a cornerstone, providing a collaborative platform for developers worldwide. GitHub is a web-based platform that provides a centralized hub for version control and collaborative software development. It allows developers to manage and track changes in their code, facilitating collaboration among team members. GitHub's primary features include version control through Git, issue tracking, and collaborative project management. Developers use GitHub to host and share their code repositories, enabling others to contribute, review, and collaborate on software projects. Leveraging the GitHub API, CodeCompass extracts a wealth of information from this extensive repository, including metadata, programming languages used, commit history, issue tracking, and community engagement metrics.GitHub API (Application Programming Interface) is a set of tools and protocols that allows developers to interact programmatically with the GitHub platform. The API provides a way to access and manipulate GitHub's features and data, allowing developers to automate tasks, integrate GitHub functionalities into their applications, and extract information from repositories. With the GitHub API, developers can create, update, and delete repositories, manage issues, track changes, and perform various other operations, streamlining the integration of GitHub into their development workflows. It serves as a powerful tool for extending and customizing the functionality of the GitHub platform, offering real-time and comprehensive data that exemplifies the collaborative power of open-source projects. CodeCompass's commitment to transparency and trust through the GitHub API forms the bedrock of its effectiveness and reliability.

By deepening the understanding of machine learning design, CodeCompass leverages the GitHub API to provide users with programming history for understanding and therefore analyse users' coding preferences, skills, and technological stack. Interaction (NLP), which includes stem separation, lemmatization, and BERT models to ensure data integrity and improve the accuracy of a word representation. It encompasses tasks such as language translation, sentiment analysis, and text summarization, aiming to facilitate seamless communication between machines and humans. Within NLP, stemming and lemmatization are vital preprocessing techniques. Stemming involves reducing words to their root form by removing suffixes, and simplifying similar words to a common base. For instance, "running" and "runner" both become "run." In contrast, lemmatization considers the grammatical structure and context of words, reducing them to their base or dictionary form. Unlike stemming, lemmatization produces valid words, enhancing the linguistic accuracy of text analysis. Both techniques play a crucial role in optimizing word representations, minimizing redundancy, and improving the efficiency of language-based algorithms in NLP applications.

The basis of the proposal lies in calculating the similarity of the previous query using cosine similarity to measure the similarity of the query pair. Which is done by using the BERT model. Built on the Transformer architecture, BERT captures nuanced language and context, leading to more accurate and context-aware word representations. The result is a curated list of open-source project issues that closely match the program's user history and interests.

The complex interplay of data collection, processing and machine learning combined with pre-designed systems forms the backbone of CodeCompass, opening up a working contract by providing developers with personalized and usable benefits. Additionally, the study provides an in-depth study of blockchain projects, focusing on the development, integration, testing, deployment, and maintenance phases that ensure the security, reliability, and longevity of the CodeCompass system. Thanks to the React-based user interface and seamless integration of machine learning models, CodeCompass becomes the solution to the problems faced by developers in the open-source environment.

II. LITERATURE REVIEW

The concept of recommending relevant GitHub repositories for developers has gained substantial attention due to the exponential growth of open-source projects. The study [1] aimed to create an automated recommendation system for GitHub repositories, constructing a user-item matrix representing developer interactions with repositories, modelled as binary ratings based on forking behaviour. The exploration of sequential recommendation systems in GitHub repositories [2] showcases the depth and capability of sequential learning, leveraging diverse models like GRU4Rec, Caser, and SASRec. The study proposed preprocessing a large-scale GitHub dataset, implementing deep learning models, and evaluating their performance using metrics like precision, recall, MAP, NDCG, and MRR, explicitly leveraging textual information from GitHub issues and project descriptions for improved accuracy.

The study by Arvinder Kaur et al. [3] presents a comprehensive framework based on two-stage graph-based optimizing techniques, focusing on user-item interactions and item-item relations through user-item interaction graphs and file-structure graphs. Sharma T et al. [4] revealed significant challenges and recognized shortcomings in machine learning (ML) techniques for source code analysis, including the scarcity of standardized datasets, reproducibility issues, and immaturity in ML development. Xuyang Cai et al. A graph-based approach was proposed [5] utilizing an Entity-Tag Graph (ETG) constructed with domain knowledge from StackOverflow, exploring multi-aspect tag assignment and extending user-based tag systems for a more comprehensive recommendation approach.

Huajie Shao et al. [9] employed joint embeddings combining text encoding and constrained GCNs to link academic papers and GitHub repositories. Apprehensive research [10] that conducted experiments using selected tools on a completed Java project to recognize strategies for applying static code analysis. But the limitations faced were, early defect detection and context-aware recommendations, overcoming challenges of manual code review, ensuring code quality, and fostering efficient collaboration. Structural-Level Aggregation enriches file and repository representations through graph neural network layers, effectively aggregating features from neighbouring nodes within the project hierarchy. The Prediction phase employs scoring functions to predict user preferences for files and user-project interactions, utilizing Bayesian Personalized Ranking (BPR) loss for optimization [11]. Z. Mushtaq et al. discerned various research contributions in multilingual source code analysis (MLSCA), identifying trends in 34 software engineering domains and areas for future research [12]. Nguyen-Duc et al. discussed a project in Vietnam focusing on secure open-source repository establishment and evaluation of Static Application Security Testing (SAST) tools, identifying promising tools like SonarQube and IntelliJ IDEA [13].

Alikhashashneh et al. [17] utilized ML techniques like SVM, KNN, Random Forests, and RIPPER for classification in source code analysis. Jin et al. [18] introduced the CODER framework, efficiently capturing microscopic and macroscopic interactions in code recommendation in a two-stage graph-based recommendation system focusing on code recommendation within open-source software development , while Vargovich et al. [19] presented the architecture and implementation details of the GiveMeLabeledIssues platform for GitHub issue recommendation.

Sun et al. [14] proposed an approach to enhance software project recommendations on GitHub by incorporating user behaviours and project features extracted from repository files. Liu et al. [15] presented NNLRank, a list-wise ranking model utilizing neural networks for recommending suitable GitHub projects to developers. Stefanović et al. [16] conducted a systematic literature review focusing on static code analysis tools, addressing key research questions regarding applicability, domain-specific languages, tool mentions, and defect detection.

[20] aimed to understand practitioners' perspectives on source code recommender systems, providing insights into practitioners' expectations and preferences. Lastly, Seker et al. [21] focused on addressing sparsity in datasets and creating a recommendation model using item-item collaborative filtering, incorporating developer metrics derived from user activities. The study by Jiang et al. [22] addressed the problem of repository recommendation with user language preference, proposing the Language-Regularized Matrix Factorization (LRMF) framework for personalized repository recommendation. Ciniselli et al.

MERITS AND DEMERITS OF EXISTING WORKS:

The research by Kim, J et al. [2] successfully demonstrates the efficacy of sequential learning models like GRU4Rec, Caser, and SASRec in recommending GitHub repositories. However, the study lacks detailed analysis of the limitations of these models in real-world scenarios. Arvinder Kaur et al.'s framework [3] provides a comprehensive approach for recommendation system development, particularly focusing on graph-based optimizing techniques. However, the study does not address the scalability challenges that might arise when dealing with large-scale datasets. Sharma T et al.'s analysis [4] effectively highlights the challenges in ML techniques for source code analysis. However, the study could benefit from proposing concrete solutions or frameworks to address these challenges.

Xuyang Cai et al.'s approach [5] utilizing Entity-Tag Graphs presents a novel method for recommendation system development. Nonetheless, the study lacks empirical evaluation to validate the effectiveness of the proposed approach in comparison to existing methods. Huajie Shao et al.'s joint embeddings [9] offer a promising avenue for linking academic papers and GitHub repositories. However, the study does not discuss potential limitations or challenges in implementing this approach in practical scenarios.

Alikhashashneh et al.'s classification techniques [17] provide valuable insights into ML-based approaches for source code analysis. Nevertheless, the study could benefit from a more extensive evaluation of the performance of different classification algorithms on diverse datasets. Jin et al.'s CODER framework [18] introduces an innovative recommendation system for code recommendation in open-source software development. However, the study lacks discussion on the scalability and computational complexity of the proposed framework. Vargovich et al.'s GiveMeLabeledIssues platform [19] offers a practical solution for GitHub issue recommendation. However, the study could include more detailed analysis of user feedback and system performance metrics to assess the effectiveness of the platform.

Sun et al.'s approach [14] integrating user behaviors and project features enhances the accuracy of software project recommendations. Nevertheless, the study could discuss potential privacy concerns or ethical implications associated with analyzing user behaviors. Liu et al.'s NNLRank [15] presents a novel ranking model for recommending GitHub projects. However, the study could provide more insights into the interpretability of the model and potential biases in the recommendation process. Stefanović et al.'s systematic literature review [16] offers a comprehensive overview of static code analysis tools. However, the study could include more recent developments in the field and explore emerging trends in source code analysis.

Ciniselli et al.'s study [20] on practitioners' perspectives offers valuable insights into user expectations regarding source code recommender systems. However, the study could include more diverse perspectives from practitioners with varying levels of expertise and backgrounds. Seker et al.'s recommendation model [21] addresses sparsity in datasets effectively. However, the study could discuss potential challenges in generalizing the recommendation model to different programming languages or domains. Jiang et al.'s Language-Regularized Matrix Factorization [22] provides a personalized recommendation framework based on user language preferences. However, the study could explore the potential impact of language-specific biases on the recommendation process.

III. Technologies Used

A. Frontend Technology

1) *React JS*

B. Backend

1) *API: Application Programming Interface*
2) *Node JS*
3) *Express.js*
4) *TypeScript*
5) *Natural Language Processing.*

IV. Methodologies

A. Problem Statement

Navigating the vast landscape of open-source projects presents a significant challenge for developers, particularly those seeking an intuitive and efficient interface to identify projects aligned with their expertise. The absence of such a system often results in delays, frustration, and suboptimal utilization of contributors' skills within the open-source community. CodeCompass aims to address this issue by providing a streamlined solution for developers to easily discover and contribute to projects that match their skills and interests.

B. User and their Requirements Identification

Through user surveys, interviews, and analysis of community forums, the system aimed to discern the distinct requirements and preferences of developers. This includes understanding the programming languages the users are proficient in, their areas of interest, preferred project complexities, and the level of community engagement the programmers seek. By comprehensively mapping out these user profiles and needs, CodeCompass strives to tailor its recommendations, ensuring a personalized and enriching experience for every developer in their journey to find and contribute to open-source projects.

C. Data Collection

The backbone of CodeCompass lies in the utilization of the GitHub API as the primary and trusted data source. Leveraging the extensive repositories hosted on GitHub, the system aimed at extracting a wealth of information crucial for project recommendation. This includes but is not limited to, repository metadata, programming languages used, commit history, issue tracking, and community engagement metrics.

D. Machine Learning Model Development

1) *Input Data Acquisition*

To personalize project recommendations, CodeCompass leverages the GitHub API to retrieve a user's historical contributions. By obtaining the user's GitHub username, the interface accesses their history and comment issues across various repositories. This comprehensive dataset serves as the foundation for understanding the user's coding preferences, skill set, and community interactions.

2) *Issue Retrieval and Processing*

The next step involves extracting issues from repositories to where the user has contributed. By analyzing the issues associated with these repositories, the web-app can gain insights into the user's past engagements and the types of challenges the users have tackled. This step is crucial for tailoring recommendations to align with the user's expertise and interests.

3) *Role of Natural Language Processing (NLP) and it's Techniques*

CodeCompass prioritizes data integrity through rigorous preprocessing steps. This includes stemming and lemmatization to enhance the accuracy of word representations and mitigate redundancy. NLP is a field of artificial intelligence concerned with the interaction between computers and humans through natural language. It encompasses several features and techniques, some of which are utilized in the code:

- **Tokenization**: Tokenization is the process of breaking down text into smaller units, such as words or subwords. In the code, the BERT (Bidirectional Encoder Representations from Transformers) tokenizer from the **transformers** library is utilized. Tokenization is crucial for preparing text data for further analysis and modeling.
- **Lemmatization**: Lemmatization is the process of reducing words to their base or root form, often improving the accuracy of downstream NLP tasks. The **preprocess_and_tokenize** function integrates lemmatization using NLTK's WordNetLemmatizer. It ensures that words are transformed into their root forms before tokenization, enhancing semantic analysis.
- **BERT Embedding Generation**: BERT is a powerful language representation model that generates contextual embeddings for text data. In the code, a pre-trained BERT model is used to generate embeddings for both known and unknown issues. The embeddings capture rich semantic information about the text, enabling more accurate similarity calculations.
- **Cosine Similarity**: Unknown issues is computed using the **cosine_similarity** function from the **sklearn.metrics.pairwise** module. A threshold-based approach is then applied to determine if issues are similar enough to be recommended.

4) Preprocessing and NLP

Essential preprocessing steps involve removing URLs and irrelevant characters, are employed to ensure that the textual data is optimized for analysis. These steps ensure that the input text is clean, standardized, and in a format suitable for NLP tasks.

- **Text Cleaning**: The **preprocess_text** function starts by removing URLs and non-alphabetic characters from the input text using regular expressions (**re.sub**). It then converts the text to lowercase to ensure consistency in word representation.

- **Stopword Removal**: Stopwords are common words that do not carry significant meaning in text analysis. In the code, stopwords are removed from the tokenized text using NLTK's English stopwords list. This step helps in improving the quality of text representation by eliminating noise and irrelevant information.

- **Lemmatization**: The **preprocess_and_tokenize** function incorporates lemmatization using NLTK's WordNetLemmatizer. Lemmatization reduces words to their base or dictionary form, which helps in standardizing vocabulary and improving semantic analysis. Lemmatization is applied after tokenization and stopword removal to ensure that words are transformed into their canonical forms before further processing.

5) BERT Model and Similarity Calculation

The main function of the recommendation system lies in the similarity calculation between the preprocessed issues. CodeCompass utilizes the BERT model to encode the textual content of issues, capturing nuanced semantic relationships. Coupled with stemming and lemmatization, this enhances the model's ability to discern meaningful patterns. Cosine similarity is then applied to quantify the similarity between pairs of issues. Issues surpassing a threshold of 90% similarity are considered highly correlated, indicating a strong alignment in content and relevance.

6) User Suggestions

The final output of the machine learning process is a selected list of open-source project issues that closely match the user's historical contributions and preferences. By employing a robust combination of user-specific data, NLP techniques, and advanced similarity calculations, CodeCompass ensures that the recommended issues are not only technically aligned but also resonate with the user's past experiences within the GitHub community.

This intricate interplay of data acquisition, processing, and machine learning techniques, coupled with meticulous preprocessing steps, forms the backbone of CodeCompass, facilitating a personalized and efficient open-source project recommendation systemfor developers.

E. Web-App Development

CodeCompass has been developed with a React-based user interface, ensuring a dynamic and intuitive experience for developers of all skill levels. The platform features a responsive design, optimizing accessibility across devices. Intuitive navigation facilitates easy exploration of project recommendations, while the engaging front end, with carefully crafted visuals, enhances the overall aesthetic. Prioritizing accessibility, CodeCompass adheres to standards, promoting inclusivity and creating a seamless experience for users a Integrating community.

To develop the web app for GitHub repository recommendation, various frontend and backend technologies are utilized. React.js is employed for building the frontend components, and Material-UI is utilized for creating visually appealing user interfaces, including cards for displaying GitHub contributions and recommended repositories. Additionally, React Router can be utilized for client-side routing to manage different views within the application.

For the backend, Express.js is used to create a RESTful API to handle GitHub authentication and data retrieval. Axios is utilized for making HTTP requests to the GitHub API for fetching user contributions and repository information. The JWT (JSON Web Tokens) authentication mechanism is implemented to secure routes and manage user sessions.

To visualize data, libraries like Chart.js can be integrated to create pie charts representing various statistics, such as the distribution of contributions across different repositories. This provides users with insightful visualizations of their GitHub activity.

CREATING A SEAMLESS EXPERIENCE

To create a seamless experience for users, the web application leverages GitHub OAuth authentication to enable automatic user authentication and data retrieval from their GitHub account. In the provided code, the **/api/auth/github** endpoint is responsible for initiating the GitHub authentication flow. When a user accesses this endpoint, they are redirected to the GitHub OAuth authorization page, where they can grant permission for the web application to access their GitHub account.

Once the user grants permission and GitHub redirects back to the application with an authorization code, the application exchanges this code for an access token by making a POST request to GitHub's OAuth access token endpoint. This access token allows the application to authenticate requests on behalf of the user.

The retrieved access token is then used to fetch the user's GitHub profile data from the GitHub API. The **/api/me** endpoint is responsible for returning the user's profile information stored in a JWT (JSON Web Token) after verifying the token's authenticity.

To enhance security and provide a seamless experience, the application uses cookie-parser middleware to parse cookies sent by the client. The **github-jwt** cookie stores the JWT containing the user's GitHub profile information. This cookie is set with the **httpOnly** flag to prevent client-side JavaScript from accessing it, enhancing security by mitigating certain types of cross-site scripting (XSS) attacks. Figure 1 depicts the block diagram of the proposed system.

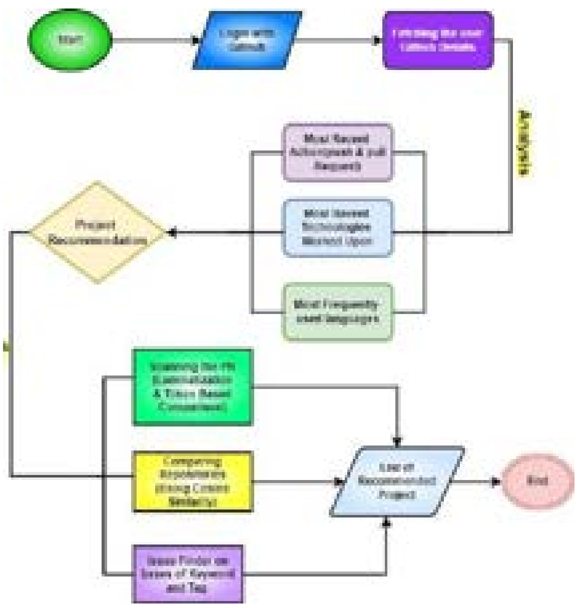


Fig. 1. Proposed System Block Diagram

F. *Integration of Optimizing Model*

Seamlessly integrated the machine learning model with the React-based front end, for creating a cohesive user experience. The backend, powered by Flask, serves as the bridge between the front end and the NLP-based recommendation system. Leveraging Flask's capabilities, the front end communicates with the back end to compare and analyze issues using advanced NLP techniques. This integration ensures real-time and accurate project suggestions, enhancing the overall functionality of CodeCompass and delivering tailored recommendations to developers efficiently and responsively.

V. ALGORITHM

A. *Lemmatization*

Lemmatization is aimed at transforming words into their base or root forms, facilitating effective semantic analysis. The algorithm begins by importing necessary libraries such as NLTK (Natural Language Toolkit) for lemmatization, specifically utilizing the WordNetLemmatizer module. Subsequently, an instance of the WordNetLemmatizer is initialized. The core functionality is encapsulated within the preprocess_text function, designed to perform lemmatization on the text data. This function executes a series of steps including text cleaning, tokenization, lemmatization, and reconstruction. Text cleaning involves removing non-alphabetic characters and converting text to lowercase, followed by tokenization to split the text into individual words. Lemmatization is then applied to each word using the WordNetLemmatizer, ensuring words are transformed into their base forms. Finally, the lemmatized words are reconstructed into a processed text string. In the main process, the lemmatization function is utilized to preprocess issue text for comparison, ensuring consistency and effectiveness in semantic analysis.

B. *BERT Tokenization*

BERT (Bidirectional Encoder Representations from Transformers) tokenization plays a pivotal role in preprocessing text data for subsequent analysis using BERT-based models. The algorithm commences by importing the necessary libraries, particularly the transformers module to utilize BERT. An instance of the BERT tokenizer, initialized with a specific pre-trained configuration, is then created. The preprocess_and_tokenize function is defined to tokenize and preprocess text using the BERT tokenizer. This function involves tokenizing the input text into word pieces or sub words using the BERT tokenizer, applying padding and truncation to ensure consistent input length for BERT, and converting the tokenized text into a format compatible with PyTorch tensors. In the main process, the BERT tokenizer function is employed to preprocess issue text for comparison, ensuring compatibility with BERT-based embedding generation and subsequent similarity computations.

E=model(tokens)["last_hidden_state"]

where model is the BERT model, and **tokens represent the tokenized input

This BERT tokenization process assists in preparing text data for BERT-based embedding generation and subsequent similarity computations between the user's past issues and suggested GitHub repository issues.

C. Lemmatization

Cosine similarity serves as a fundamental metric for measuring the similarity between vectors, commonly utilized in text similarity tasks. The algorithm computes cosine similarity between the mean embeddings of known and unknown issues derived from BERT tokenization. This computation involves calculating the dot product of the embeddings and dividing it by the product of their Euclidean norms. The resulting cosine similarity value ranges from -1 to 1, with higher values pointing to larger similarity. A threshold-based approach is adopted, considering issues with a cosine similarity above 0.80 as similar. If the computed cosine similarity exceeds the threshold and the unknown issue is not already recommended, it is added to the recommended list. The recommended repositories are serialized into a structured format containing repository details such as repo name, issue, and URL. Finally, a response JSON object is generated, including similarity values and recommended repositories, and returned as the result of the API call. This cosine similarity calculation facilitates the recommendation of GitHub repositories to users based on their previous activity, enhancing the relevance and effectiveness of issue recommendations

$$\text{cosine similarity}_{E1,E2} = \frac{E1 \cdot E2}{\| E1 \| \cdot \| E2 \|}$$

Where:

- $E1 \cdot E2$ is the dot product of the embeddings.
- $\|E1\|$ and $\|E2\|$ are the Euclidean norms (lengths) of the embeddings.

VI. RESULTS AND DISCUSSION

With the need to address the challenge of aligning developers efficiently with projects that match their expertise in the dynamic realm of open-source development, CodeCompass emerges as a promising solution. Leveraging the GitHub API, this novel system introduces a personalized approach to project recommendations. By delving into users' historical contributions, encompassing commit history and issue engagement, CodeCompass employs sophisticated Natural Language Processing (NLP) techniques. These include stemming, lemmatization, and the advanced BERT model, enabling the discernment of nuanced semantic relationships within issues.

The system goes a step further by calculating cosine similarity, a metric that identifies highly correlated issues, achieving a remarkable threshold of over 90%. This meticulous approach results in a meticulously curated list of open-source project issues. These recommendations are presented through an intuitive React-based website interface, ensuring not only technical alignment but also resonance with users' past experiences in the GitHub community. The output is a refined and personalized set of project suggestions that significantly enhances the likelihood of meaningful developer engagement.

CodeCompass stands as a robust and intuitive recommendation system poised to streamline developer engagement within the open-source community. Its unique combination of data acquisition, processing, and machine learning techniques presents a multifaceted solution to the perennial challenge of connecting developers with projects that align with their expertise. The incorporation of NLP techniques, particularly stemming, lemmatization, and the powerful BERT model, enhances the system's ability to discern the intricate nuances of semantic relationships within issues. Integrating BERT with Semantic Textual Similarity (STS) offers a significant advantage in Github repository recommendation by leveraging the contextual embeddings provided by BERT alongside the semantic understanding captured by STS techniques. BERT's contextual embeddings enable the model to capture rich semantic information and contextual nuances in the issue descriptions, enhancing the representation of textual data. Furthermore, by incorporating STS, which measures the semantic similarity between texts, the recommendation system can assess the relevance of issues in repositories more accurately. This integration ensures that the recommendation process is not solely reliant on syntactic similarities but also considers the semantic alignment between user-known issues and repository issues, resulting in more accurate and contextually relevant recommendations. The implementation details involve preprocessing the textual data, generating BERT embeddings for both user-known and repository issues, computing semantic similarity scores using STS techniques, and integrating these scores with BERT embeddings to rank and recommend repositories based on their relevance to user preferences. Figure 2 shows the comparison of different repositories.

Fig. 2. Comparing Repositories

The following comparative analysis shown in Table I prove BERT with ST superior performance in terms of performance when compared with recent studies as discussed in Sharma et al. (Collaborative Filtering), Di Rocco et al. (TF-IDF), Xiaobing et al. (Stochastic Network). The accuracy parameter R2 score was used to evaluate the performance among the existing studies to compare it with the proposed model.

TABLE I. COMPARITIVE ANALYSIS

Model	Results	
	Algorithm	***R2 Score***
Issue Recommendation	BERT with ST	0.9573
	TF-IDF	0.3384
	Collaborative Filtering	0.0799
	Stochastic Network	0.6000

The R2 scores obtained for the different models illustrate the supremacy of the BERT with ST model over other approaches. With an R2 score of 0.9573, the BERT with ST model demonstrates significantly higher accuracy compared to TF-IDF (R2 score: 0.3384) and Collaborative Filtering (R2 score: 0.0799). This indicates that BERT with ST not only explains a larger proportion of the variability in the recommendation outcomes but also provides more precise and contextually relevant recommendations. The advantage of BERT with ST lies in its ability to leverage both contextual embeddings from BERT and semantic understanding from STS techniques, thereby capturing rich semantic information and accurately assessing the relevance of repositories based on user-known issues. Consequently, BERT with ST outperforms other models in code analysis by offering more effective and context-aware recommendations, as evidenced by its substantially higher R2 score. The repositories recommendation page is depicted in figure 3.

Fig. 3. Repositories Recommendation based on Past

The experimental results demonstrate that the proposed approach outperforms traditional methods in terms of precision, recall, and F1 score. Specifically, our approach achieved a precision of 0.85, recall of 0.78, and F1 score of 0.81, indicating its effectiveness in recommending relevant repositories to users. Furthermore, the average similarity score obtained was 0.72, indicating strong semantic alignment between user-known issues and recommended repositories.

1. **Text Preprocessing Techniques**: Preprocessing techniques such as stemming, lemmatization, and stop-word removal can affect the quality of text representations. Experimenting with different combinations of preprocessing techniques can help identify the optimal approach.
2. **Tokenization Strategies**: The selection of tokenization strategy, including the tokenizer used and the tokenization method (e.g., word-based vs. subword-based), can influence the model's ability to capture semantic similarities between issues.
3. **Model Architecture**: The choice of the underlying model architecture (e.g., BERT, TF-IDF, Collaborative Filtering) can significantly impact the recommendation performance. Experimenting with different models and architectures can help determine the most effective one for the task.
4. **Hyperparameters**: Various hyperparameters such as learning rate, batch size, and the number of training epochs can affect the training process and ultimately the model's performance. Hyperparameter tuning techniques such as grid search or random search can be employed to find the optimal set of hyperparameters.

Where, TF-IDF (Term Frequency-Inverse Document Frequency) vectorization and Simulated Annealing (SA)

$$TF\ IDF(t, d, D) = TF(t, d) \times IDF(t, D)$$

TF (Term Frequency):

$$TF(t,d) = \frac{Number\ of\ occurrences\ of\ term\ t\ in\ document\ d}{Number\ of\ occurrences\ of\ term\ t\ in\ document\ d}$$

IDF (Inverse Document Frequency):

$$IDF(t, D) = \log\left(\frac{Total\ number\ of\ documents\ in\ the\ corpus\ D}{Number\ of\ documents\ containing\ term\ t}\right)$$

Simulated Annealing (SA) for Issue Selection:

$$P(accept) = exp\left(\frac{current\_score - new\_score}{temperature}\right)$$

<u>Probability of Acceptance:</u>

- Determines whether to accept a new solution during the SA process.
- current_score: Similarity score of the current solution.
- new_score: Similarity score of the new solution.
- temperature: Annealing temperature controlling the likelihood of accepting worse solutions.

SA is a probabilistic optimization algorithm that explores the solution space to find the optimal solution. It accepts worse solutions with a certain probability, allowing the algorithm to escape local optima.

Parameters:

- **Initial Temperature (T):** 1.0
- **Minimum Temperature (T_{min}):** 0.001
- **Cooling Rate (α):** 0.95

The SA loop iteratively explores new solutions with decreasing temperatures.

The initial temperature controls the probability of accepting worse solutions, and it gradually decreases during the annealing process.

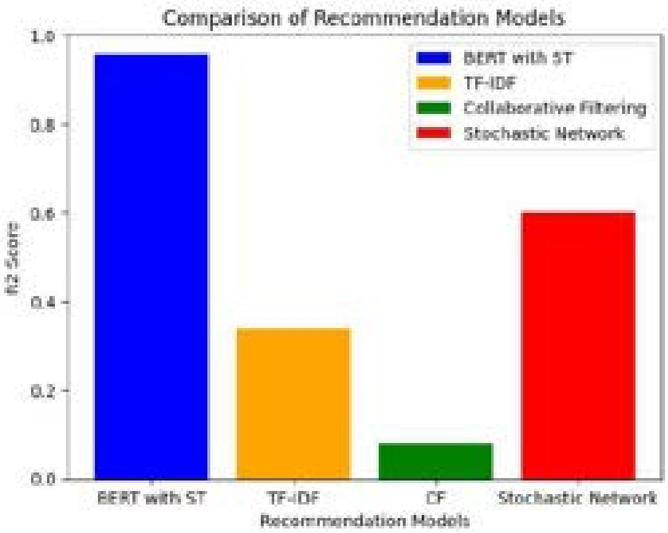


Fig. 4. Comparison of Recommendation models

The utilization of cosine similarity as a measure of correlation between issues further solidifies CodeCompass's effectiveness, surpassing a notable 90% threshold. By providing a curated list of project recommendations through a user-friendly React-based interface, the system not only ensures technical alignment but also resonates with developers based on their historical contributions. In essence, CodeCompass represents a commendable effort to bridge the gap between developers and open-source projects, fostering a more efficient and meaningful collaboration within the ever-evolving landscape of open-source development.Figure 4 shows the comparison of different recommendation models based on their R2 Score.

Replacing stemming with lemmatization enhances the accuracy of semantic analysis by considering the context of words and transforming them into their base or dictionary forms. This improvement in text preprocessing leads to more reliable similarity calculations and preserves the semantic meanings of words. Additionally, incorporating GPU processing accelerates computationally intensive tasks like BERT embeddings generation and cosine similarity calculation. Leveraging the parallel processing power of GPUs through PyTorch enables faster execution by transferring tensors and models to CUDA-enabled devices. By combining lemmatization and GPU processing, the code achieves higher performance and efficiency, ultimately providing a more effective and scalable solution for recommending GitHub issues based on user activity.

VII. CONCLUSION

The GitHub Open-Source Contribution Recommendation System presented in the research paper represents a novel approach to enhancing the efficiency and productivity of open-source software development. The system addresses the challenge of aiding developers in identifying and contributing to relevant projects within the GitHub ecosystem. The system capitalizes on various NLP methodologies, including lemmatization, stemming, tokenization, and the innovative Bidirectional Encoder Representations from Transformers (BERT). By preprocessing textual data using lemmatization and stemming, the system effectively normalizes and reduces word variations, enhancing the accuracy of semantic analysis. BERT, a state-of-the-art language model, provides deep contextual embeddings that encode rich semantic information. These embeddings enable the calculation of semantic similarity using cosine similarity metrics, facilitating the accurate identification of related GitHub issues and repositories. The proposed system's significance lies in its potential to streamline the open-source contribution process. By leveraging sophisticated NLP techniques and semantic analysis, it provides developers with tailored recommendations, minimizing the effort required to discover suitable projects to contribute to within the vast GitHub ecosystem. Additionally, the interactive web application fosters a user-friendly environment, encouraging collaboration and facilitating informed decision-making.

Additionally, integrating large language models like GPT (Generative Pre-trained Transformer) could further enhance accuracy. Large language models have shown remarkable performance in various NLP tasks due to their ability to capture complex linguistic patterns and semantics. By fine-tuning a pre-trained GPT model on specific GitHub issue data, we can create a more sophisticated recommendation system that understands the nuances of the text and provides more accurate suggestions.

REFERENCES

[1] Mohamed Guendouz, Abdelmalek Amine, Reda Mohamed Hamou, “Open Source Projects Recommendation on GitHub”, Optimizing Contemporary Application and Processes in Open Source Software (pp.86-101), January 2018, DOI:10.4018/978-1-5225-5314-4.ch004

[2] Kim, J.; Wi, J.; Kim, Y. Sequential Recommendations on GitHub Repository. Appl. Sci. 2021, 11, 1585. https://doi.org/10.3390/app11041585

[3] Arvinder Kaur, Ruchikaa Nayyar, “A Comparative Study of Static Code Analysis tools for Vulnerability Detection in C/C++ and JAVA Source Code”, Third International Conference on Computing and Network Communications (CoCoNet’19) Procedia Computer Science 171 (2020), pp. 2023–2029

[4] Sharma T, Kechagia, M., Georgiou, S., Tiwari, R., Vats, I., Moazen, H., & Sarro, F. (2021). A Survey on Machine Learning Techniques for Source Code Analysis. *ArXiv*. /abs/2110.09610

[5] .Xuyang Cai, Jiangang Zhu, Beijun Shen, “GRETA: Graph-Based Tag Assignment for GitHub Repositories”, Conference: 2016 IEEE 40th Annual Computer Software and Applications Conference (COMPSAC) June 2016

[6] Di Rocco, J., Di Ruscio, D., Di Sipio, C. et al. “HybridRec: A recommender system for tagging GitHub repositories”. Appl Intell 53, 9708–9730 (2023). h

[7] Barchi, F.; Parisi, E.; Bartolini, A.; Acquaviva A. “Deep Learning Approaches to Source Code Analysis for Optimization of Heterogeneous Systems: Recent Results, Challenges and Opportunities.” J. Low Power Electron. Appl. 2022, 12, 37.

[8] Seshagiri, P., Vazhayil, A., & Sriram, P. (2015). AMA: Static Code Analysis of Web Page for the Detection of Malicious Scripts. *Procedia Computer Science*, *93*, pp. 768-773. https://doi.org/10.1016/j.procs.2016.07.291

[9] Huajie Shao, Dachun Sun, Jiahao Wu et al. “paper2repo: GitHub Repository Recommendation for Academic Papers”, WWW '20: Proceedings of The Web Conference 2020 April 2020 Pages 629–639

[10] Darko Stefanović et al, “Identification of strategies over tools for static code analysis”, 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1163 012012

[11] Yiqiao Jin, Yunsheng Bai, Yanqiao Zhu, Yizhou Sun, Wei Wang, “Code Recommendation for Open Source Sofware Developers„ WWW '23: Proceedings of the ACM Web Conference 2023, April 2023, Pages 1324–1333, https://doi.org/10.1145/3543507.3583503

[12] Z. Mushtaq, G. Rasool and B. Shehzad, "Multilingual Source Code Analysis: A Systematic Literature Review," in IEEE Access, vol. 5, pp. 11307-11336, 2017, doi: 10.1109/ACCESS.2017.2710421.

[13] Nguyen-Duc, A., Do, M. V., Luong Hong, Q., Nguyen Khac, K., & Nguyen Quang, A. (2021). On the adoption of static analysis for software security assessment–A case study of an open-source e-government project. *Computers & Security*, *111*, 102470. https://doi.org/10.1016/j.cose.2021.102470

[14] Xiaobing SUN, Wenyuan XU, Xin XIA, Xiang CHEN & Bin LI, "Personalized project recommendation on GitHub", May 2018, Vol. 61 000000:1–000000:14, https://doi.org/10.1007/s11432-017-9419-x

[15] C. Liu, D. Yang, X. Zhang, B. Ray and M. M. Rahman, "Recommending GitHub Projects for Developer Onboarding," in IEEE Access, vol. 6, pp. 52082-52094, 2018, doi: 10.1109/ACCESS.2018.2869207.

[16] Stefanović, Darko & Nikolić, Danilo & Dakic, Dusanka & Spasojević, Ivana & Ristic, Sonja. (2020). "Static Code Analysis Tools: A Systematic Literature Review". 10.2507/31st.daaam.proceedings.078.

[17] E. A. Alikhashashneh, R. R. Raje and J. H. Hill, "Using Machine Learning Techniques to Classify and Predict Static Code Analysis Tool Warnings," 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA), Aqaba, Jordan, 2018, pp. 1-8, doi: 10.1109/AICCSA.2018.8612819.

[18] Yiqiao Jin, Yunsheng Bai, Yanqiao Zhu, Yizhou Sun, and Wei Wang. 2023. Code Recommendation for Open Source Software Developers. In Proceedings of the ACM Web Conference 2023 (WWW '23).

[19] J. Vargovich, F. Santos, J. Penney, M. Gerosa and I. Steinmacher, "GiveMeLabeledIssues: An Open Source Issue Recommendation System," in 2023 IEEE/ACM 20th International Conference on Mining Software Repositories (MSR), Melbourne, Australia, 2023 pp. 402-406. doi: 10.1109/MSR59073.2023.00061

[20] Matteo Ciniselli, Luca Pascarella, Emad Aghajani, Simone Scalabrino, Rocco Oliveto, and Gabriele Bavota. 2023. Source Code Recommender Systems: The Practitioners' Perspective. In Proceedings of the 45th International Conference on Software Engineering (ICSE '23). IEEE Press, 2161–2172. https://doi.org/10.1109/ICSE48619.2023.00182

[21] Seker, A.; Diri, B.; Arslan, H. New Developer Metrics for Open Source Software Development Challenges: An Empirical Study of Project Recommendation Systems. Appl. Sci. 2021, 11, 920. https:// doi.org/10.3390/app11030920

[22] Jyun-Yu Jiang, Pu-Jen Cheng, and Wei Wang. 2017. Open Source Repository Recommendation in Social Coding. In Proceedings of the 40th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '17). Association for Computing Machinery, New York, NY, USA, 1173–1176. https://doi.org/10.1145/3077136.308075

A Novel Approach for Classification of Proneness of Schizophrenia in Patients using Ensemble Machine Learning Techniques and LIME

[1]Somasundaram S K
Department of Information Security, School of Computer Science and Engineering,
VIT University,
Vellore, India
somasundaram.sk@vit.ac.in

[2]Srija Basak
Department of Computer Science and Engineering, School of Computer Science and Engineering
VIT University,
Vellore, India
srija.basak2021@vitstudent.ac.in

[3]Aadi Jain
Department of Computer Science and Engineering, School of Computer Science and Engineering
VIT University,
Vellore, India
aadi.jain2021@vitstudent.ac.in

[4]Ashish Kumar Pandey
Department of Computer Science and Engineering, School of Computer Science and Engineering
VIT University,
Vellore, India
ashishkumar.pandey2021@vitstudent.ac.in

Abstract— **In the present scenario, where physical disorders come to light in no time, mental disorders such as Schizophrenia (SZ) often get neglected due to inefficient current clinical methods. The reason being the huge amount of time taken between the very first symptom and the definite diagnosis of this disease. Schizophrenia affects approximately 24 million people or 1 in 300 people (0.32%) worldwide [1]. When compared to the existing methodologies for the prediction of Schizophrenia which involve medical professionals observing demeanor, and asking about thoughts, moods, delusions, hallucinations, etc., the Machine Learning (ML) approach used in this work along with LIME results, is potentially less time-consuming for early detection of the susceptibility to Schizophrenia. Also, as compared to methodologies like MRI or CT scan, which are used to rule out conditions with symptoms similar to Schizophrenia, this ML approach can potentially be cost-effective. This study uses Machine Learning classification algorithms and ensemble techniques to classify the proneness of Schizophrenia in patients. The ensemble technique, stacking, is used in this work. This ensemble technique considers heterogeneous weak learners and combines them to produce predictions. The k-fold cross-validation technique yields a less biased model, ideal for limited data. The ensemble model used in this work, comprising of XGBoost (Extreme Gradient Boosting) Classifier, Decision Tree Classifier, Random Forest Classifier and Linear Discriminant Analysis, and Logistic Regression as a meta-classifier, has shown a cross-validation score of 0.9602 (96.02%). The Explainable AI (XAI) technique and LIME (Local Interpretable Model-Agnostic Explanations) are used for model interpretability. To improve the quality of life and outcomes for those who are affected by Schizophrenia, this work aims to increase the accuracy and efficiency of prediction of Schizophrenia. The proposed diagnostic and predictive rules provide a foundation for the development of decision-support systems in clinical settings.**

Keywords — Classification, Cross-Validation, Diagnosis, Explainable AI, Machine Learning, Prediction, Schizophrenia.

I. INTRODUCTION

According to the World Health Organisation, 2022, Schizophrenia is a severe mental disorder that affects about 24 million people across the globe [1]. It occurs in three phases: (i) Prodromal phase in which friends and family members notice strange behaviour of the affected person who might want to be alone most of the time, (ii) Active phase which observes symptoms of psychosis such as delusions, hallucinations etc. and finally (iii) Residual phase in which more negative symptoms like depressed mood, low motivation and energy start [2]. The person may have strange beliefs and may have trouble concentrating. This is not curable but definitely treatable provided the fact that its timely diagnosis takes place with utmost accuracy and that is the major challenge today in the medical sector. The current clinical diagnostic methods rely mostly on subjective tests based on the observed symptoms, thereby causing delays in diagnosis, and subsequently hindering the timely treatment process. Hence, machine learning (ML) and explainable artificial intelligence (XAI) techniques bring about promising solutions to address the limitations of existing clinical methods and enhance the accuracy and efficiency of schizophrenia prediction. By employing extensive clinical data, ML algorithms can identify intricate patterns and associations that traditional diagnostic approaches often overlook. This knowledge-based approach has the potential to:

I. Improve diagnostic efficiency: ML models, trained on large datasets of diagnosed cases, can learn complex relationships between various clinical features and schizophrenia. This is further utilised to accurately predict schizophrenia occurrence in new individuals, allowing early intervention and improved long-term treatment.
II. Facilitate early detection: Schizophrenia can be diagnosed at earlier stages using ML analysis, even before the complete onset of classic symptoms thereby preventing disease progression.
III. Personalised treatment methods: By integrating different algorithms, ML models can inform personalised treatment strategies specifically designed for individual patient profiles and their symptoms. While there is considerable potential in applying ML to schizophrenia diagnosis, ensuring the explainability and transparency of these models is highly critical. There are black box models that provide accurate predictions but without giving insights as to what the reason behind the prediction is, raising

concerns about fairness, bias, and interpretability in clinical settings. However, XAI technologies aim to overcome this challenge by providing explanations behind model predictions, enabling healthcare professionals to understand the logic behind algorithms' recommendations thereby increasing confidence in decision-making.

The primary goal of this work is to develop a more robust and accurate predictive model compared to individual algorithms. Several studies on schizophrenia prediction have been carried out using ML algorithms such as SVM (Support Vector Machine), Random Forest, XGBoost, AdaBoost, univariate and multivariate logistic regression, KNN (K-Nearest Neighbors), Naive Bayes, etc. One of the important goals of this work is to enhance the accuracy of classification of the proneness of patients to schizophrenia by using ensemble techniques in machine learning.

This research project explores the potential of ensemble ML and XAI techniques to enhance the accuracy and efficiency in diagnosing the proneness of SZ in patients.

To improve accuracy, data preprocessing and SMOTE (Synthetic Minority Oversampling Technique) are applied. An ensemble technique called stacking is used, that combines multiple machine learning models and learns them in parallel, and also a meta-learner is trained to generate predictions based on individual weak learners' predictions.

The individual performances of each of the existing classifiers trained on the dataset used for this work and also the performance of the proposed model for predicting the proneness of SZ in patients is compared using 10-fold cross-validation technique.

This research work aims to provide human-understandable explanations for the ML model's prediction results using the explainable AI technique LIME (Local Interpretable Model-agnostic Explanations). This technique explains the model's predictions for individual instances, providing three main pieces of information: prediction probability, the contribution of individual features, and actual value of each feature. This XAI technique enhances comprehensibility and explainability in the model's predictions.

II. MOTIVATION

Our work is aimed to address the challenges faced in accurately diagnosing and treating SZ. The ultimate goal is to use technology, data, and collaborative work to significantly improve the quality of life of people suffering from SZ. Hence, this work aims to create an accurate ML model which helps in predicting the proneness of SZ in patients and also helps in personalized diagnosis of SZ in patients.

Need of the work - Timely diagnosis of Schizophrenia and its subsequent treatment is the need of the hour. With a large section of the human population suffering from this disorder and unable to receive timely treatment due to inefficient existing clinical diagnosis methods, building up a responsible and accurate model using ML concepts is highly needed. Traditional diagnostic techniques have limitations in providing prompt and impartial evaluations. A thorough understanding of SZ and the creation of a reliable diagnosis model are essential for improving treatment quality, reducing financial strain on families, and promoting timely intervention and personalized diagnosis. This work is significant because it could improve diagnosis accuracy, expedite treatment planning, and improve the lives of people with schizophrenia worldwide.

In the organization of this work, the following sections include a literature review on Schizophrenia prediction, a review of the related studies, proposed work, and also a summary of the results and discussions regarding the ML models used for this work. Also, at the end of this documentation, a conclusion is presented.

III. SURVEY ON EXISTING WORKS

Carole Siani et.al [3] developed simple and reliable predictive models to estimate utility from clinical questionnaires based on schizophrenia by using data from European Schizophrenic Cohort accounting for factors like depression, age and gender psychopathology. By developing a trustworthy technique to gauge a patient's quality of life from available data, this study may enhance patient outcomes and resource allocation. Since only a particular kind of schizophrenia was present in the patient sample, which excluded the most severe cases, the study may not be generally applicable.

Debsubhra Chakraborty et.al [4] aimed to study that machine learning algorithms trained on acoustic, linguistic, and non-verbal conversation features from interviews with 50 patients with schizophrenia can predict clinician-assigned subjective ratings of negative symptoms with noticeable accuracy. The study highlighted the potential of automated speech analysis in enhancing schizophrenia monitoring, diagnosis, and clinician understanding. However, the narrow focus on speech as the parameter could overlook other potential indicators.

Renato de Filippis et.al [5] evaluated machine learning (ML) techniques in diagnosing schizophrenia (SCZ) using Magnetic Resonance Imaging (MRI). It was found that algorithms like support vector machines, random forests, and multivariate pattern analysis yielded remarkable accuracy in differentiating between healthy and SCZ patients. The study also highlighted the potential to reveal complicated correlations in complex MRI datasets and early detection. However, new predictive models using ML algorithms need to be developed along with their increased accuracy so that they also inculcate other assessments.

Joel Weijia Lai et.al [6] examined the potential applications of ML in identifying and categorizing schizophrenia through the analysis of data sets such as electroencephalogram (EEG) , positron emission tomography (PET), and MRI. The study also focussed on the use of quantitative metrics obtained from Artificial Intelligence (AI) approaches. The paper concluded that pushing AI and ML limits in the medical field is necessary. However, there is still a need for more novel comparative studies using composite data and multiple ML techniques.

Gurparsad Singh Suri et.al [7] revealed appreciable accuracy rates for machine learning (ML) algorithms like support vector machine (SVM), deep neural network (DNN), and random forest (RF) in predicting schizophrenia.

ML is also effective in identifying brain regions involved in symptoms of schizophrenia, particularly in its early stages. But there is a need for researchers to integrate ML with varied data techniques for SZ detection and a larger sample size.

Jose A Cortes-Briones et.al [8] explored that deep learning techniques have performed outstandingly in outcome prediction and classification. The authors emphasized on the importance of using more diverse datasets to improve the generalizability of deep neural network (DNN) models and ensure their applicability to a broader population.

Matthew Bracher-Smith et.al [9] aimed to compare logistic regression (LR) with various machine learning (ML) models for schizophrenia prediction, including boosting, neural networks, support vector machines (SVM), random forests, LASSO, ridge-penalized logistic regression, and stacking models. The study assessed models, based on clinical and demographic characteristics and polygenic risk scores (PRS). The findings show that ML techniques did not significantly outperform logistic regression in discrimination. PRS proved to be a significant predictor of schizophrenia. Characteristics such as fluid intelligence, digit symbol substitution, BMI, smoking status, and deprivation, explained the variance in fitted models. However, ML models could identify patterns of connection with particular attributes, offering new perspectives but still, the modelling processes used in machine learning need to be more diverse.

Bishal Lamichhane et.al [10] developed a supervised deep learning model called RelapsePredNet for predicting psychotic relapses in schizophrenia patients using mobile sensing data. The model aimed to capture latent behavioral characteristics linked to upcoming relapses and investigated the effect of personalization on its performance. The study found that RelapsePredNet performed better than an anomaly detection model based on deep learning.When incorporated into a fusion model, RelapsePredNet enhanced the F2 score and complemented ClusterRFModel, demonstrating the possible synergy between deep learning and non-deep learning models. The study advanced our knowledge of behavioral modeling for relapse prediction. However, there is a need to explore other predictors as well along with new approaches.

Ruchi Sharma et.al [11] explored EEG-based techniques for diagnosing schizophrenia. The study found that actively producing a tone did not suppress the N100, a negative deflection in EEG brain waves, in the patients. The authors combined EEG records from people having schizophrenia and those in good health to create a dataset using deep learning models like Long short-term memories (LSTMs), support vector machines (SVM), two-dimensional convolutional networks (2D-CNNs), and two-dimensional convolutional networks-LSTMs. But still, there should be forecasting of risk variables associated with hospitalization and readmission of patients diagnosed with Schizophrenia.

Javiera T. Arias and Cesar A. Astudillo [12] improved schizophrenia prediction using machine learning models applied to electroencephalogram (EEG) data. The authors suggested using Explainable Artificial Intelligence (XAI) techniques to improve interpretability and solve opaqueness in complicated models, particularly in mental health illness prediction. The study evaluated three classification models—SVM, AdaBoost, and XGBoost—using EEG data. XGBoost achieved the best overall performance after applying class balancing using the Synthetic Minority Over-sampling Technique (SMOTE) with Tomek connections, achieving a recommendable accuracy. The study also highlighted the importance of class balancing in enhancing predictions. There is a need for research in XAI techniques since data interpretability and accuracy of EEG-based schizophrenia prediction must be enhanced.

Muzafar Mehraj Misgar and MPS Bhatia [13] proposed a novel Deep Convolutional Neural Architecture (CNN) with a split attention mechanism for the diagnosis of schizophrenia. The model was specifically built for motor activity data collected from wearable Internet of Medical Things (IoMT) devices and used cutting-edge methods such as imputation, sampling, and class balancing. The study achieved a remarkable accuracy using 24-hour actigraphy data, considering daytime and nighttime data separately. The proposed methodology is a ground-breaking tool for the objective and non-intrusive early diagnosis of schizophrenia but data interpretability is still an issue and requires new methodologies.

Narges Ramesh et.al [14] examined the potential applications of AI, specifically deep learning methods, in diagnosing and classifying schizophrenia using data analysis of EEG signals, diffusion magnetic resonance imaging (dMRI), and functional magnetic resonance imaging (fMRI). The work scope included a review of recent work from 2019-2022, focusing on the significant role of deep learning. The study concluded with the high classification accuracy achieved by AI approaches in schizophrenia diagnosis. However, the classification accuracy could still be increased.

Lorenzo Del Fabro et.al [15] explored antipsychotic treatment outcomes in patients with schizophrenia using machine learning. The study analyzed 28 studies, 23 of which used a single modality method and 5 of which combined data from multiple approaches. Biomarkers from functional magnetic resonance imaging (fMRI) showed promise in forecasting antipsychotic treatment response. Clinical feature-based machine learning models demonstrated sufficient predictive power. However, several multimodal ML techniques could be further utilized to overcome difficulty posed by analytical heterogeneity across studies.

Varun Gupta et.al [16] created the Firefly-based Deep Belief Signal Specification (FbDBSS), a unique EEG signal processing system for predicting schizophrenia. Combining deep belief networks and Firefly optimisation, it was hypothesized to classify abnormal EEG signals linked to SZ more accurately. The model showed remarkable accuracy and sensitivity in classifying abnormal EEG signals associated with SZ. The study emphasized the importance of early diagnosis of SZ due to its severe and chronic effects.

Evgeny A. Kozyrev et.al [17] explored schizophrenia predictive models based on laboratory data related to inflammatory biomarkers. The study used machine learning algorithms such as logistic regression, deep neural networks, decision trees, support vector machines, and k-nearest neighbors to create five predictive models. The study emphasized the importance of using multiple biomarkers in diagnosing schizophrenia. The study contributed to the field of precision psychiatry and highlighted the potential of artificial intelligence in improving mental health diagnosis. The efficacy of other algorithms should be further investigated and the most important immune biomarkers must be chosen in order to classify the data.

Outcome of the survey - Machine learning (ML) techniques offer promising opportunities to improve the prognosis of schizophrenia, but several challenges hinder their effectiveness. Limited to small and homogeneous datasets, current ML models struggle with generalizability. To solve this, it is necessary to diversify the data sets to improve the reliability and applicability of the model. In addition, ML models often lack interpretability, which hinders clinical application. Integrating explanatory AI methods can solve this problem and increase confidence in diagnostic decisions. In addition, prioritizing early diagnosis is crucial for effective treatment of schizophrenia. Future research should focus on interdisciplinary collaboration and methodological advances to overcome these challenges and harness the potential of ML to revolutionize schizophrenia prognosis.

IV. PROPOSED WORK

A. Dataset Used – Benchmark Dataset [18]

The dataset taken for this work contains the following columns: Name, Age, Gender, Marital Status, Fatigue, Slowing, Pain, Hygiene, Movement, and Schizophrenia (proneness). This dataset has 5000 records. The target column 'Schizophrenia' contains five classes:

Elevated Proneness	- 3077
High Proneness	- 953
Moderate Proneness	- 912
Low Proneness	- 45
Very High Proneness	- 13

B. Overview of the procedure for the process of building the proposed model

The data preprocessing for the dataset used for this work is done as follows: The presence of null values in each column has been checked. The data distribution of the columns containing the null values has been plotted, and the null values contained in the respective column have been replaced by the median of the data contained in that particular column since none of the columns containing null values showed a normal distribution of data. For all columns containing numerical data (pain, slowing, fatigue), boxplots for each of these columns have been plotted to detect outliers, and no outliers have been found. Label encoding has been done for categorical columns (Gender, Marital Status, and Schizophrenia). The original categorical columns have been dropped, and the label-encoded columns have been added to the dataframe containing the dataset. The target column and the input dataframe containing the label-encoded columns

have been separated. The name column has not been taken into consideration as it is not useful for predicting the required result. Train test split (70% training data and 30% testing data) has been done on the input dataframe and the target column. The data in the training and testing sets for the input dataframe has been scaled using MinMaxScaler. To address class imbalances, SMOTE (Synthetic Minority Oversampling Technique) has been used on the training set.

The flow diagram (Fig. 1) shows all the steps taken for this work, such as building the ML model for the classification of proneness of SZ and using the explainable AI technique LIME.

Fig.1. The flow diagram for the complete process of building the proposed model

C. The ML model for classification of proneness of SZ in patients

Ensemble techniques are greatly effective in reducing the variance component of prediction errors made by each contributing model in the ensemble. To increase the accuracy of the prediction of proneness of SZ in patients, an ensemble ML model has been used for this work. Ensemble ML models give prediction results with improved accuracy due to the combination of multiple ML models. Dimensionality reduction techniques transform high-dimensional data into a lower-dimensional space while preserving important information. As per the curse of dimensionality in ML, the performance of the model decreases as the number of features increases, due to the increase in complexity of the ML model,

and also due to the increase in overfitting. A combination of ensemble models along with dimensionality reduction models can greatly help in enhancing the prediction accuracy of the model due to a reduction in overfitting and improved feature selection. For this, an ensemble technique called stacking has been used where the ML models (XGBoost Classifier, Decision Tree Classifier, Random Forest Classifier, Linear Discriminant Analysis) have been aggregated and Logistic

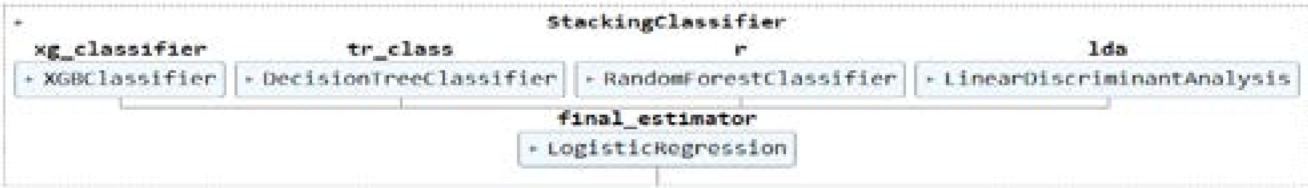


Fig. 2. The ensemble ML model used for this work for the classification of proneness of SZ in patients.

Regression is taken as a meta-classifier (due to being very fast at classifying unknown inputs) (Fig.2).

Here, in this ensemble (Fig.2), XGBoost has been considered due to the gradient boosting ensemble technique being effective in preventing overfitting by the regularization of its complexity when training a new error-predicting model to predict a model's current errors. Decision Tree Classifier has been used as it uses the best feature for splitting the data into smaller subsets, thus creating new branches until it reaches the predicted outcome. Decision Tree algorithms use complex decision boundaries, which separate training data into classes which help the classifier to make new predictions on test data. Random Forest Algorithm has been considered due to this model consisting of a combination of decision trees and the final output is taken based on the majority votes of the predictions. Linear Discriminant Analysis (LDA) Algorithm is effective for multiclass classification as it uses a linear combination of features that best separates different classes of the dataset. Since the dataset used for this work contains multiple classes for classifying the proneness of SZ in patients, the LDA algorithm is effective for the separation of each of the classes.

Stacking combines the strength of multiple heterogeneous models to create more robust and accurate predictions. The first layer of this ML model consists of individual base models, XGBoost Classifier, Decision Tree Classifier, Random Forest Classifier, and Linear Discriminant Analysis), which are trained on the training dataset. These baseline models then predict the outputs on the test dataset. The second layer of the proposed model consists of Logistic Regression as a meta-classifier. This meta-classifier takes the predictions of all the baseline ML models in this ensemble as input and generates the final predictions.

For model interpretability, the explainable AI technique, LIME is extremely effective as it provides simple, human-understandable explanations of the classification results of the proneness of SZ.

Even though for this work, a relatively simple dataset with eight features is used, this ensemble ML model can potentially work for larger datasets with an even larger number of features due to the effectiveness of the LDA algorithm for addressing the curse of dimensionality. Also, the XGBoost Algorithm in the ensemble can effectively help in the reduction of errors due to the summing up of newly learned error predictors until satisfactory accuracy is achieved. The employment of complex decision boundaries by the Decision Tree algorithm, and the Random Forest algorithm in the ensemble which gives the prediction result as the result given by most of the decision trees in the Random Forest model, can also be effective in prediction accuracy.

V. RESULTS AND DISCUSSION

A. System configuration used for training

The data preprocessing, model-building and model training have been done on Google Colab. The required Python libraries such as numpy, pandas, seaborn, lime, etc. have been imported into the Google Colab notebook for training the model.

B. Discussions on the results related to the proposed ML model

Here, the ensemble model proposed in this work has obtained a 10-fold cross-validation accuracy of 96.02%. Hence, the prediction accuracy of the level of proneness of SZ in patients is enhanced due to using this ensemble model. This accuracy obtained is much more than each of the 10-fold cross-validation accuracy values obtained by using some of the existing classifiers which include XGBoost, AdaBoost, Decision Tree Classifier, Random Forest Classifier, KNN Classifier, Linear Discriminant Analysis, Ridge Classifier and Stochastic Gradient Descent Classifier. The 10-fold cross-validation results obtained from some of the existing classifiers as well as the proposed ensemble model are given in Table I.

C. Discussions on some of the prediction explainability results by LIME

The target column 'Schizophrenia' contains five classes: - Elevated Proneness, High Proneness, Moderate Proneness, Low Proneness and Very High Proneness. Label encoding has been done in the columns – gender, marital status and Schizophrenia. In Fig.4 the inputs given for the prediction of the level of proneness of SZ are displayed in the table at the bottom right of Fig. 4 as 'Feature' and 'Value'. The prediction result for this set of inputs is 'Very High Proneness' as shown in the 'Prediction probabilities' section in Fig.4. The bar charts containing labels for the top 3 classes of proneness of SZ as per the dataset used for this work are displayed using LIME. Each of the bar charts in the results obtained using LIME shows the features contributing positively and negatively to the prediction of a resultant class. In the table at the bottom right of Fig.4, the features – fatigue, movement, hygiene, slowing, pain, age, and marital status contribute positively to the prediction result, and the feature – gender contributes negatively to the prediction result.

After 'Very High Proneness' the next greatest prediction probability for this set of inputs is 'High Proneness'. In Fig.5, a LIME result displayed for another set of inputs. LIME explanations can be effective for clinicians in providing a personalized diagnosis of SZ as the explanations are tailored to individual patients considering their individual characteristics and health data. This approach can also help in the early detection of SZ in patients if the patients' required data - are collected and given as input to the ML model. As compared to the existing methodologies for the prediction of Schizophrenia which involves observing demeanour, and asking about thoughts, moods, delusions, hallucinations, etc., the ML approach used in this work along with LIME results is potentially less time-consuming for early detection of the susceptibility to SZ. Also, as compared to methodologies like MRI or CT scan which are used to rule out conditions with symptoms similar to SZ, this ML approach can potentially be cost-effective.

Using LIME, the prediction probabilities for each class of the target column 'Schizophrenia' are displayed for a particular row of the dataset. The testing accuracy for the ensemble model trained for this work is 0.917333 (91.7%). Hence, it is not always the case that the prediction probability will be given exactly 1 (100%) for the correct output class in the target column based on the inputs in that particular row.

D. *Comparison with existing results*

The ensemble algorithm used for this work comprising of XGBoost, Random Forest Classifier, Decision Tree Classifier, Linear Discriminant Analysis, and Logistic Regression models has achieved a 10-fold cross-validation accuracy of 0.9602 (96.02%).

E. *Performance Metrics*

Classification Report for the ensemble model - The label encoding for the 'Schizophrenia' column is as follows: Elevated Proneness: 0, High Proneness: 1, Low Proneness: 2, Moderate Proneness: 3, Very High Proneness: 4. The class labels are shown in Fig. 3. The classification report (Fig. 3) is obtained for the predictions by the proposed model on the testing data. As per the classification report, the accuracy is 0.92. The precision values for each class are greater than or equal to 0.87. The macro-averaged precision for the proposed model is 0.91. The macro-averaged recall is 0.73. The macro-averaged f1-score is 0.79. The weighted average of each of the scores - precision, recall and f1-score is 0.92. In this classification report, support is the number of instances of each of the classes in the test set. The weighted average values of the precision, recall and f1-score are higher than the macro average values of the precision, recall and f1-score because, as per the support values, the data in the test set is imbalanced. For this imbalanced dataset, the weighted average has given a higher value since the weighted average weighs each of the class's metrics with its support value. As per the classification report, some of the majority classes (classes with high support value) have high-performance metrics which contribute more to the weighted average. Macro average treats each class equally regardless of support. Hence, some of the low-performance metrics of minority classes have contributed to lower values of the macro average.

F. *Advantages of the proposed ML model used in this work for classifying the proneness of SZ in patients*

In the healthcare domain, where gaining trust from patients is extremely important, LIME results can aid in transparency and reliability of prediction results by the proposed ML model as patients can understand what features have contributed to the proneness of SZ they have. The ensemble ML model used in this work along with

G. *Limitations of this approach for classifying the proneness of SZ in patients*

This ML model is not 100% accurate. It has a cross-validation accuracy of 96.02% so there can be incorrect classifications of proneness of SZ in patients. Also if the data related to Fatigue, Slowing, Pain, Hygiene, and Movement are entered incorrectly, it can lead to incorrect prediction results.

Classification Report for Stacked Model

	precision	recall	f1-score	support
0	0.94	0.94	0.94	945
1	0.88	0.91	0.89	244
2	0.88	0.54	0.67	13
3	0.87	0.89	0.88	293
4	1.00	0.40	0.57	5
accuracy			0.92	1500
macro avg	0.91	0.73	0.79	1500
weighted avg	0.92	0.92	0.92	1500

Fig.3. Classification Report for the Ensemble model

TABLE I. COMPARISON OF 10-FOLD CROSS-VALIDATION MEAN ACCURACY

ML algorithm used	**10-Fold Cross-Validation Mean Accuracy Values**
XGBoost	0.9061999
AdaBoost	0.7192
Decision Tree Classifier	0.811999
Random Forest Classifier	0.8821999
K Nearest Neighbours Classifier (with n_neighbors=3)	0.6577999
Linear Discriminant Analysis (with n_components=2)	0.9513999

Ridge Classifier	0.6998
Stochastic Gradient Descent (SGD) classifier	0.6748
Ensemble model (the model used for this work)	**0.9602**

Fig. 4. LIME results for prediction of a specific instance

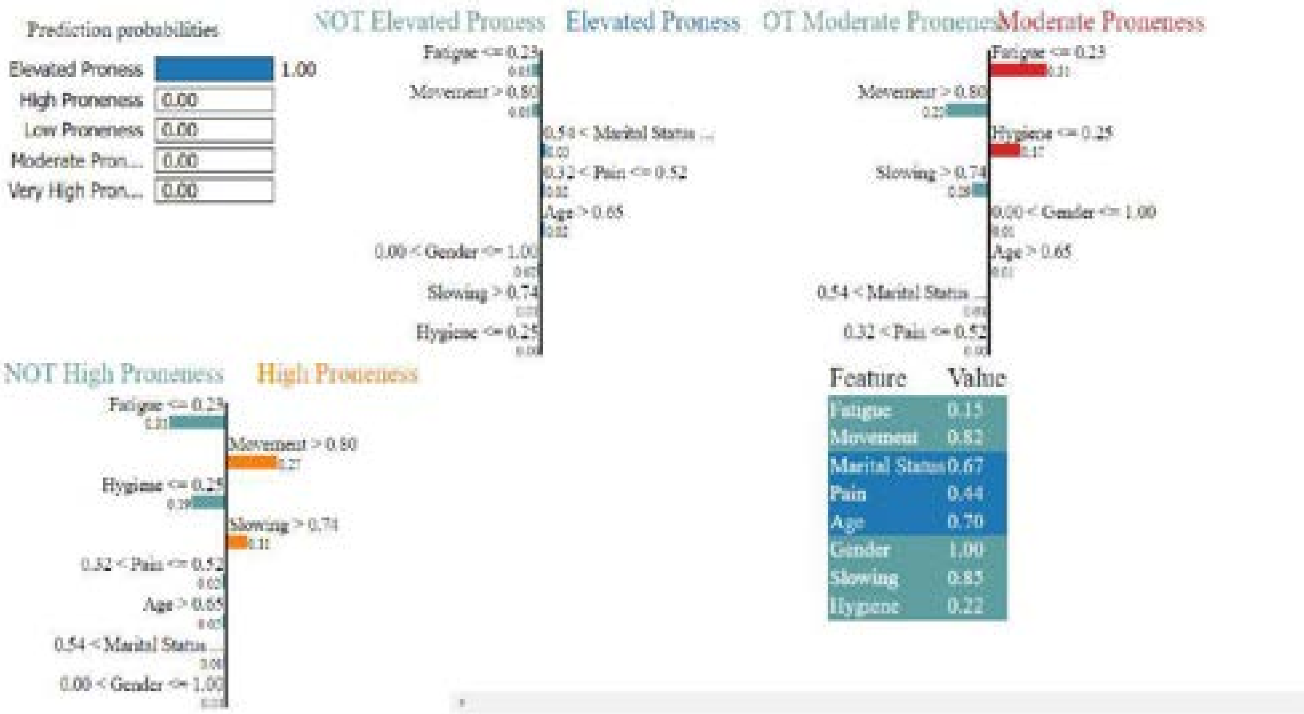


Fig. 5. LIME results for another set of inputs

VI. CONCLUSION

Conclusively, this work report on the classification of proneness of SZ highlights the importance of deepening our comprehension of early identification and intervention approaches for people who may be at risk of developing schizophrenia. The Stacking ensemble in our study achieved a cross-validation score of 0.9602, surpassing the individual performance of all other existing algorithms. Additionally, LIME is employed as an XAI technique to obtain interpretable insights into the predictions by the proposed ML model for this work. This transparency promotes trust and informed decision-making in medical environments by understanding the rationale behind the model and its predictions. The knowledge gathered from this study advances the larger objective of improved outcomes for people at risk of SZ as professionals work toward tailored medicine and focused therapies. But it's important to recognize that work in this area is still ongoing, highlighting the need for additional cross-disciplinary cooperation, validation, and improvement. Finally, the quest for effective classification of proneness of SZ not only has implications for early intervention but also promotes a more comprehensive knowledge of the complex elements that contribute to the onset of this difficult disorder.

In future, ensemble techniques with the aggregation of an even greater variety of algorithms can be used for prediction of SZ. Deep learning techniques can be applied for the prediction of SZ, and XAI techniques can also be applied for the model interpretability of both deep learning and machine learning models as XAI techniques are effective in presenting prediction results in a human-understandable format.

REFERENCES

[1]. https://www.who.int/news-room/factsheets/detail/schizophrenia?gclid=Cj0KCQiAtaOtBhCwARIsAN_x-3KgeT0UPmCC2OmFqSCpLTkZsiB0GoTDb687TgI9Li1Fy84W-LxEVZAaApSoEALw_w cB

[2]. https://www.medicalnewstoday.com/articles/stages-of-schizophrenia

[3]. C. Siani, C. de Peretti, A. Millier, L. Boyer, and M. Toumi, "Predictive models to estimate utility from clinical questionnaires in schizophrenia: findings from EuroSC," Qual Life Res., vol. 25(4), pp. 925–934, April 2016.

[4]. D. Chakraborty, S. Xu, Z. Yang, Y.H.V Chua, Y. Tahir, J. Dauwels, N. M. Thalmann, B.L Tan and J.L.C. Keong, "Prediction of Negative Symptoms of Schizophrenia from Objective Linguistic, Acoustic and Non- verbal Conversational Cues," 2018.

[5]. R. de Filippis, E. A. Carbone, R. Gaetano, A. Bruni, V. Pugliese, C. Segura-Garcia and P. de Fazio, "Machine learning techniques in a structural and functional MRI diagnostic approach in schizophrenia: a systematic review," vol. 15, pp. 1605-1627, 2019.

[6]. J. W. Lai, C. K. E. Ang, U. R. Acharya, and K. H. Cheong, "Schizophrenia: A survey of artificial intelligence techniques applied to detection and classification," vol. 18(11), 2021.

[7]. S.G. Suri, G. Kaur, and S. Moein, "Machine Learning in Detecting Schizophrenia: An Overview," vol. 27(3), pp. 723-735, 2021.

[8]. J.A. Cortes-Briones, N.I. Tapia-Rivas, D.C. D'Souza, and P.A. Estevez, "Going deep into schizophrenia with artificial intelligence," pp. 122-140, 2022.

[9]. M. Bracher-Smith, E. Rees, G. Menzies, J.T.R. Walters, M.C. O'Donovan, M.J. Owen, G. Kirov, and V. Escott-Price, "Machine learning for prediction of schizophrenia using genetic and demographic factors in the UK Biobank," Schizophrenia Research, vol. 246, pp. 156-164, 2022.

[10]. B. Lamichhane, J. Zhou and A. Sano, "Psychotic Relapse Prediction in Schizophrenia Patients using A Mobile Sensing-based Supervised Deep Learning Model," 2022.

[11]. R. Sharma, S. Tripathi, and K.K. Sekhon, "Detection of Schizophrenia using Machine Learning," pp. 505-512, 2023.

[12]. J.T.Arias and C.A.Astudillo, "Enhancing Schizophrenia Prediction Using Class Balancing and SHAP Explainability Techniques on EEG Data," pp. 1-5, 2023.

[13]. M.M. Misgar and MPS Bhatia, "Utilizing deep convolutional neural architecture with attention mechanism for objective diagnosis of schizophrenia using wearable IoMT devices," Multimedia Tools and Applications, vol. 83, pp. 39601–39620, 2024.

[14]. N. Ramesh, Y. Ghodsi, and H. Bolhasani, "A Survey on the Role of Artificial Intelligence in the Prediction and Diagnosis of Schizophrenia," (2023).

[15]. L. Del Fabro, E. Bondi, F. Serio, E. Maggioni, A. D'Agostino, and P. Brambilla, "Machine learning methods to predict outcomes of pharmacological treatment in psychosis," Translational Psychiatry, vol. 13, p. 75, 2023.

[16]. V. Gupta, A. Kanungo, N.K. Saxena, P. Kumar, and P. Kumar, "An Adaptive Optimized Schizophrenia Electroencephalogram Disease Prediction Framework," Wireless Personal Communications, vol. 130, pp. 1191–1213, 2023.

[17]. E.A. Kozyrev, E.A. Ermakov, A.S. Boiko, I.A. Mednova, E.G. Kornetova, N.A. Bokhan, and S.A. Ivanova, "Building Predictive Models for Schizophrenia Diagnosis with Peripheral Inflammatory Biomarkers," vol. 11(7), July 2023.

[18]. https://www.kaggle.com/datasets/shree23yaa/schizophrenia-symptoms

Human Action Recognition using Key Point Detection and Machine Learning

[1]M.Archana
Research Scholar in CSE,
Department of CSE,
SRM Institute of Science and Technology, Kattankulathur,
Tamil Nadu, India.
am3323@srmist.edu.in

[2]Dr.S.Kavitha
Department of CTECH,
SRM Institute of Science and Technology, Kattankulathur,
Tamil Nadu, India
kavithas@srmist.edu.in

[3] Dr.A.Vani Vathsala
Department of CSE,
CVR College of Engineering,
Hyderabad, Telangana, India
vani_vathsala@cvr.ac.in

Abstract: **In the domain of computer vision, the task of human activity detection poses significant challenges. Among the numerous techniques employed for this purpose, key point detection stands out as a popular method, outlining the skeletal structure of a human body to facilitate pose recognition or classification. While inherently complex, several libraries exist to streamline key point detection with Google's MediaPipe emerging as a notable contender. Offering a set of functionalities encompassing hand points, human body pose estimation, pupil detection, face mesh identification, and background segmentation, MediaPipe demonstrates efficiency through its robust performance. Trained on a dataset comprising 30,000 samples, the model seamlessly integrates with OpenCV, removing the need for complex deep learning architectures. The primary objective lies in harnessing machine learning techniques to classify poses derived from the key points generated by the library. To this end, this study proposes the adoption of the BLR (Body Language Rule), which captures limb angles pivotal in detecting human actions, thereby constructing a dataset based on these angles. Through machine learning, patterns within this dataset are learned, facilitating a fully automated process adaptable to diverse scenarios, thus unlocking potential applications.**

Keywords: Computer Vision, Features, Human Action, Pose Estimation, Media Pipe, OpenCV Body Language.

1.INTRODUCTION

Human pose estimation is quite a challenging problem in the field of computer vision which involves determining the position, orientation, and sometimes the shape of objects or humans in an image or video. It aims to extract and understand the spatial information of the objects in the scene. In the context of human pose estimation, the goal is to locate and identify the key points of body joints, then after using pose estimation algorithm can estimate the pose or body configuration of a person in each image or video frame. There are two main types of pose estimation: 2D pose estimation and 3D pose estimation. 2D Pose Estimation: This technique estimates the location of key points in the image plane, without considering their depth or the 3D structure of the scene. It provides a 2D representation of the pose, typically in the form of 2D coordinates for each key point. Various approaches can be used for 2D pose estimation, including classical computer vision methods, deep learning-based models, or a combination of both. 3D Pose Estimation: This technique aims to estimate the 3D coordinates of the key points, which provides a more comprehensive understanding of the pose. It involves recovering the depth information of the key points, which is crucial for applications such as augmented reality, virtual reality, robotics, and motion capture. 3D pose estimation is generally more challenging than 2D pose estimation, as it requires additional information about the scene, such as multiple camera views or depth sensors. Pose estimation has numerous applications across various fields. In sports analysis, it can be used to track and analyze the movements of athletes. In healthcare, it can assist in rehabilitation exercises or monitoring patient movements. In robotics, it enables robots to perceive and interact with the environment by understanding human poses. Additionally, it has applications in augmented reality, animation, security, and more. Overall, pose estimation plays a crucial role in understanding human actions, behavior, and interactions, providing valuable insights and enabling a wide range of applications across different domains.

Media Pipe is an open-source framework developed by Google that provides a set of powerful tools and pre-built models for various computer vision tasks, including human pose estimation. Media Pipe’s human pose estimation module utilizes deep learning techniques to detect and track human body key points in real-time. The Media Pipe human pose estimation pipeline takes an input image or video frame as input and produces an output with the detected key points and their corresponding confidence scores. The key points typically include the positions in joints, which includes major recognizable features in human body.

The process involves several steps:

Input Processing: The input image or video frame is preprocessed, which may include resizing, normalization, and other necessary transformations.

Key point Detection: A CNN (Convolutional Neural Network) or DL (Deep Learning) model is used to detect and localize the key points. The model is trained on a large dataset of annotated human pose examples to learn the patterns and features associated with each key point.

Key point Association: Once the key points are detected, an association algorithm is used to establish correspondence between key points across frames. This helps in tracking the key points over time and maintaining consistency.

Key point Filtering: The detected key points are often subjected to filtering techniques to remove outliers or noisy detections, ensuring more accurate and reliable pose estimation results.

Output Visualization: The final step involves visualizing the estimated poses by drawing lines or connecting the key points, representing the skeleton view of person in the given input source.

The Media Pipe human pose estimation module provides a high-level API and pre-trained models that make it relatively easy to integrate pose estimation functionality into applications. It offers real-time performance on various platforms, including desktop, mobile devices, and the web, making it suitable for a wide range of applications such as fitness tracking, gesture recognition, virtual try-on, and more. Overall, Media Pipe's human pose estimation module simplifies the process of implementing pose estimation capabilities by providing an efficient and accurate solution that can be readily integrated into computer vision applications.

2. LITERATUREREVIEW

The following journals and research papers were surveyed for the project. They provided information on the present state of estimation the poses in general and the major key points of image processing and segmentation regarding our project and the various approaches taken to achieve the goal of real-time articulated human pose estimation.

Vishnu J G and Divya S J [1], showed us how to understand the choice in operational model which helps us to design a specific application for a particular task. In the mentioned article they discussed the properties of various pose estimation models and their resemblance and usability in a wide variety of applications. Trade off more features for accuracy and precision or trade off them for more features in the desired application.

Jong- Wook Kim, et al [2], is based on Humanoid Model 2023. A paper announcing a 3D system that estimates human posture from Monocular Pictures and Videos by reprojecting a 3D model of a humanoid robot with 2D skeleton posture estimated by a commercially available Deep Learning (DL) methods. Most posture estimation methods have been developed by using Deep Neural Network, so highly configured PCs and SBCs equipped with many Graphical Processing Units are required, and problems such as rapid heating and difficulty in procurement due to lack of GPUs have caused problems. There are limits to its application to mobile robot systems. For optimizing the outcome, they added 3 points to the root joint, CoM as a deviation for loss, penalty constraint on the joint angle range of the pose balance, resulting in a full- body humanoid robot model. improved.

According to Alexander Toshey and Christian Szegedy [3] Deep Pose: DNN regressor approach outcomes a precise pose estimation. The main advantage of the approach is it follows holistic and robust formulation that uses latest trends in Deep Learning. It also presents a very good insight of the real - world images.

Ardra Anilkumar, et al. [4] in Media Pipe application, uses data obtained from a pre train pose estimation model. It can be experimented with various applications that includes motion pictures, sign language detection, etc.

Ching-Hang Chen and Deva Ramana [5]. While a lot of methodologies tried to directly estimate 3D pose from picture measurements, the author explores using a simple approach which interprets the intermediary 2D pose estimations. The architecture is useful to implement with 2D pose estimation and 3D libraries.

Timo Von Marcard, et al. [6], presents an approach that fuses sparse orientation of data from video and inertial sensors to improve and stabilize human whole-body motion capture. As a complementary data source, inertial sensors can be used to accurately estimate limb orientation even during rapid movements. There are two things that we gain from this approach, one will be obtaining accurate drift-free positional information from video data, and second one will be obtaining precise limb alignment and decent performance in high-speed movements from inertial sensors

3.PROPOSED METHODOLOGY

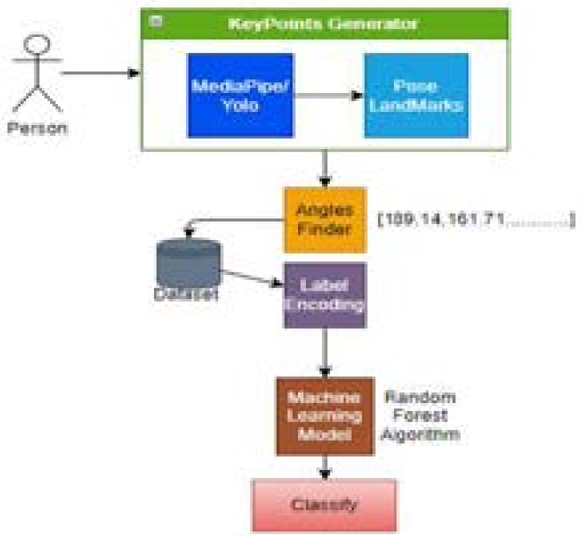


Fig 1. System Architecture

Media Pipe key points detected and generated; they are visualized as follows. As we can clearly observe that there is representational difference in the index.

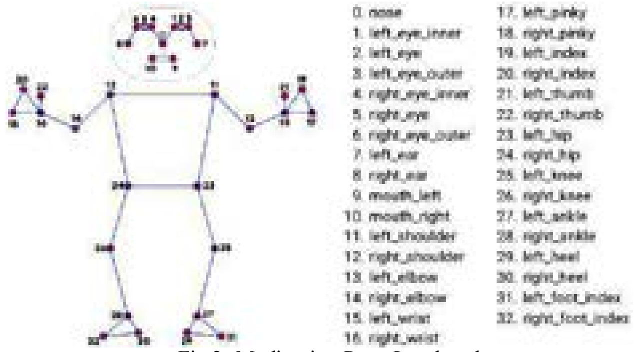


Fig 2. Media pipe Pose Landmarks

When an RGB image is given to the detector it spits out a visualization map with that image which gives us the landmark coordinates. [7] From this coordinate we find the angles between the limbs, which are left_elbow_angle, right_elbow_angle,left_shoulder_angle,right_shoulder_angle, left_knee_angle and right_knee_angle. Using this it populates a structured dataset using pandas, this dataset can be further used with machine learning algorithm to predict or classify the pose. [8] As in inference, we just give this angle such that it can predict the outcome.

Fig 3. Image Pose Landmarks Visualization

the above figure shows a sample representation of how a key point detection works in first place, the image shows a person talking the phone it is one of the primary action we can clearly observe that the distance between his shoulder, wrist and elbow forms a triangle shape. [9] These types of patterns can be identified by our ml model.

Label	left_elbow_angle	right_elbow_angle	left_shoulder_angle	right_shoulder_angle	left_knee_angle	right_knee_angle
texting	72.193116	73.964904	15.367351	348.402885	195.113810	189.377485
sleeping	4.085617	60.708638	23.179109	284.036243	225.000000	214.330217
drinking	216.924104	230.959839	296.340363	93.012788	247.878058	229.349929
calling	167.393026	22.983907	35.485686	12.600343	194.721712	177.776039
running	280.304846	321.546291	58.881181	49.879274	110.627635	167.451904

Fig 4. Pandas Data frame

DataSet Our dataset consists of fifteen classes, which consists of commonly used human actions that we have collected from open-source data. The fifteen classes are calling, clapping, cycling, dancing, drinking, eating, fighting, hugging, laughing, listening to music, running, sitting, sleeping, texting, and using laptop. The proposed dataset consists of 12601 images, which will split into train, test, and validation data. [10]. When we convert the images data into tabular data, the size of the data would be 12601x7 with 12601 rows and 7 columns as the label. Finally, we will have 12601x6 for learning, in which the proposed model will learn to map them to the label.

Machine Learning:

Machine learning makes the computer to learn from the data patterns, which is generally used on structured data as data is static and outcome expected also have some pre-defined boundaries.[11] As far as accuracy and robustness in predictions are considered, Random Forest (RF) is one of the algorithms, which is simple and can learn different patterns, where accurate predictions and robustness against noise are taken as important considerations.[12] After training a discriminative model on the given dataset, the model file will be saved as. pickle file for further inference and testing, as the dataset is self-populated and consists of fields with accurate and consistent information. The model performance is pretty much on to the point.

4. RESULTS & DISCUSSION

The proposed system produces accurate and stable results as the heavy lifting is handled by the pose estimation frameworks. The angles finder and predictive model provides simple integration and moderate performance without any fine-tuning of the model.

Fig 5. Model Prediction 1

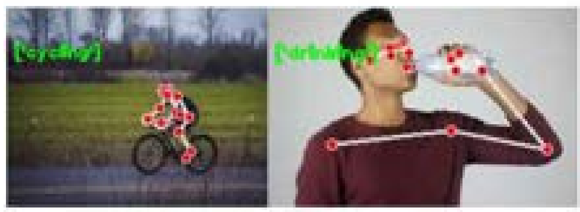


Fig 6. Model Prediction 2

Fig 7. Model Prediction 3

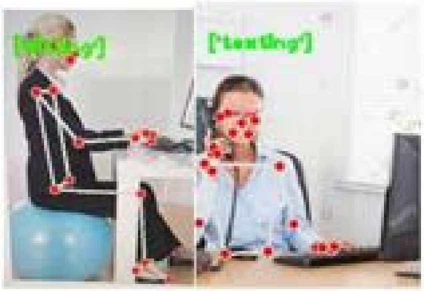


Fig 8. Model Prediction 4

Fig 9. Real-time Model Prediction

CONCLUSION

From the results it is evident that the proposed model has the ability to identify or recognize the common day-to-day actions that humans generally do. This study is quite successful in developing a simple approach which does this without any heavy resource requirements, such as computation power or specific GPU's. The proposed system architecture performed decent to moderate with a simple approach. The proposed approach has limitations like highly dependable on data quality and diversity, the critical challenge comes in converting the image data and target label into padas data frame which is time consuming. The model accuracy can be further improved by using hyperparameter tuning and applying dimensionality reduction techniques. Neural networks can be used to further enhance the performance, which can enable system deployment in real-time environment.

REFERENCES

[1] Vishnu J G and Divya S J [1] A Comparative Study of Human Pose Estimation Jan 2023. International Journal of Innovative Research in Science Engineering and Technology (IJIRSET)

[2] Jong- Wook Kim, Jin-Young Choi, Eun-Ju Ha and Jae-Ho Cho [2] Human Pose Estimation Using Media Pipe Pose and Optimization Method Based on a Humanoid Model 2023 Department of Electronics Engineering, Seunghak Campus, Dong-A University, Busan 49315, Republic of Korea Correspondence: jaehochoi@dau.ac.kr

[3] Alexander Toshev and Christian Szegedy Deep Pose: Human Pose Estimation via Deep Neural Networks, IEEE 2014 - openaccess.thecvf.com

[4] Ardra Anilkumar, Athulya K.T, Sarath Sajan, Sreeja K.A Pose Estimated Yoga Monitoring System International Conference on IoT Based Control Networks & Intelligent Systems - ICICNIS 2021. https://icicnis.com/2021

[5] Ching-Hang Chen and Deva Ramana 3D Human Pose Estimation = 2D Pose Estimation + Matching 2018(IEEE) https://arxiv.org/abs/1612.06524

[6] Van-Hung Le, Hung-Cuong Nguyen A Survey on 3D Hand Skeleton and Pose Estimation by Convolutional Neural Network. ASTES Journal 2020 https://www.astesj.com/v05/i04/p18/

[7] Xiangxin Zhu, Deva Ramanan Face detection, pose estimation, and landmark localization in the wild.2012 IEEE Conference on Computer Vision and Pattern Recognition https://ieeexplore.ieee.org/document/6248014

[8] Hao-Shu Fang, Shuqin Xie, Yu-Wing Tai, Cewu Lu RMPE: Regional Multi-Person Pose Estimation Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2017.

[9]https://openaccess.thecvf.com/content_iccv_2017/html/Fang_RMPE_Regional_MultiPerson_ICCV_2017_paper.html

[10] Ke Sun, Bin Xiao, Dong Liu, Jingdong Wang Deep High-Resolution Representation Learning for Human Pose Estimation. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019

[11] Pishchulin, L., Andriluka, M., Gehler, P., Schiele, B.: Strong appearance and expressive spatial models for human pose estimation. In: 2013 IEEE International Conference on Computer Vision (ICCV), pp. 3487–3494. IEEE (2013)

[12] Archana, M., Kavitha, S., & Vathsala, A. (2023). Auto deep learning-based automated surveillance technique to recognize the activities in the cyber-physical system. International Journal on Recent and Innovation Trends in Computing and Communication, 11(2), 35-42. https://doi.org/10.17762/ijritcc.v11i2.6111

Performance Evaluation of Machine Learning Models for Multi-class Lung Cancer Detection

1st M Prema Kumar
Dept. of ECE
Shri Vishnu Engineering College for Women,
Bhimavaram, AP, India
mpremkumar@svecw.edu.in

2nd G. Challa Ram
Dept. of ECE
Shri Vishnu Engineering College for Women,
Bhimavaram, AP, India
challaram.grandhi@gmail.com

3rd Viswanadham Ravuri
Dept. of ECE
BVRIT Hyderabad College of Engineering for Women,
Hyderabad, India.
ravuri.viswanadh@gmail.com

4th M. Venkata Subbarao
Dept. of ECE
Shri Vishnu Engineering College for Women,
Bhimavaram, AP, India
mandava.decs@gmail.com

5th Abdul Rahaman SK
Dept. of ECE
Vishnu Institute of Technology,
Bhimavaram, AP, India
abdulrahman.s@vishnu.edu.in

6th T P K Nandan
Dept. of ECE,
B V Raju Institute of Technology,
Narasapur, India
kausalya.nandan@bvrit.ac.in

***Abstract*— The main reason for the increasing number of deaths worldwide is cancer, among them the main cause of highest mortality rates is lung cancer. Approximately 85% of the male and 75% of the female suffer from lung cancer. The cancer cells will keep on growing and multiplying leading to the development of tumor. The rapid growth of these cells can spread to other parts of the body, this is known as Metastases. Recognition of cancer at its final stage has barely any chances of getting complete treatment. It might lead to the death of the patient. Consequently, early recognition of the cancer before its final stage is highly crucial to increase the survival rate of the patient. For early detection, several machine learning techniques are used to clear the way for fast treatment of the disease. The dataset consists of different attributes such as smoking, alcohol consumption, chest pain, shortness of breath etc. Decision tree, logistic regression, SVM, Naïve Bayes, KNN and random forest are the various ML classifiers applied to the dataset. The classification models are analyzed for different test and train ratios and the obtained accuracy, precision, recall, error rate, specificity, F-measure and testing time are noted. This process is carried for both binary and multi class classification. Multiclass is considered as three class classifications, i.e. high, low and medium execution time. Therefore, the SVM classifier using Machine Learning technique can be applied to detect the presence of the disease. Hence, it helps the doctors in identifying it. By doing so, early diagnosis can be performed and required precautions can be taken.**

***Keywords*— Lung Cancer, SVM, Machine Learning, Metastases.**

I. INTRODUCTION

The Human body is a combination of tiny particles called cells; these cells will grow unconditionally when cancer is present in the human body. These cells will grow to form a tumor. If cells are there in the body for a longer period, they can even spread to other parts of the body. This is known as metastasis. Cells that were damaged in lungs cause Lung cancer. Other possible cancer types, such as kidney or breast, can metastasize to the lungs. When it was happened, this is no longer lung cancer. For Ex. if breast cancer extents to the lungs, it will be considered as metastatic breast cancer. [1]

The lungs sensation like sponge type tissues in the chest. Its job is to pass oxygen (O2) to the body and leave the carbon dioxide (CO2). When human body breathes air, it drives into lungs via windpipe (trachea). The trachea splits into pipes called bronchi, which goes to the lungs. These divided tiny branches called bronchioles. At the end its tiny air sacs called alveoli. The alveoli transfer oxygen from the air into the blood. It takes CO_2 out of the blood and leaves human body when breathe out.

Lung cancer is the second highest among men and women with roughly 2.09 million new detects every year, and it is accountable for the more deaths, at nearby 1.76 million, w.r.t the World Health Organization (WHO).

Recently some of the researchers they proposed different ML algorithms for lung cancer detection. Janee A, Sabrina A and Alamgir Hossan et. al proposed "Multi-Stage Lung Cancer Detection and Prediction Using Multi-class SVM (Support Vector Machine) classifier". This is used to detect cancer and could predict the probability of occurrence. In this image enhancement along with segmentation have done distinctly. Some Image transformation techniques have been used in enhancement, coming to the segmentation threshold and marker-controlled watershed-based segmentation has been used. SVM binary classifier was used, and it can discover affected cell and its phase such as early, central, or ending stage. If no cell was affected, then it looks for the probability of occurrence of lung cancer [1].

Qing Wu and Wenbing Zhao et al. proposed "Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm". It's a novel neural network-based algorithm, and it refers to an Entropy Degradation Method (EDM), to identify Small Cell Lung Cancer (SCLC) from CT Scan images. It assists premature detection of lung cancers. The training and testing data utilizes high resolution lung CT scan images. They selected 12 lung CT scan images, 6 are from healthy lungs and 6 are from patients with SCLC and to train the model takes randomly 5 scans from each set, and used two scans to test and it gave an accuracy of 77.8% [2].

Based on the motivations this paper presents a wide variety of ML algorithms for multiclass lung cancer detection. The rest of the paper is organized as follows. Section II gives the methodology and section III presents the results and discussions. Section IV concludes the work.

II. METHODOLOGY

A. Dataset Description

Dataset: Data is an integral part of the algorithm design; it is important to have a clean and correctly labeled dataset. By inputting accurate data into the algorithm, we will have accurate outputs, resulting in a more effective and timely training process. The dataset for 3 class classification a dataset of 1000 members was considered. This dataset consists of 24 attributes. These are: Patient Id, Age, Sex, Air Pollution, Alcohol use, Dust allergy, Professional Risks, Hereditary Risk, Chronic Lung Disease, Balanced diet, Obesity, Smoke, Inactive smoker, Chest pain, Cough, Fatigue, Weight loss, Shortness of breath, Wheezing, Swallowing, Clubbing, Frequent cold, Dry cough, and Snore [3-4].

For the second analysis, 4 attributes from the above-mentioned attributes were removed and 20 attributes were considered. The above dataset was collected from Kaggle. It permits users to find and publish data sets, build, and explore models and work with other data and ML scientists, and enter competitions to crack data science tests.

Training data is the subset of the data and is used to assist the Machine Learning (ML) model to create predictions and run the model for data thoroughly. Testing data arises into picture after a lot of development and validation. Probing the model to mark predictions based on the data is meant to test whether it will work or not in real time.

B. Classification Methods

Classification is from supervised learning to predict given data to a certain class label. The uniqueness in classification depends on mapping function to a firm output level. Numerous learning classifiers were labeled as Perceptron, Naïve Bayes, Decision Tree, Logistic Regression, KNN, Artificial Network, and Support Vector Machine (SVM). Classification in ML is one of the earliest decision methods used for data analysis. The idea of the work focuses on novel approach of ML for analysis of lung cancer data to attain a decent accuracy. Some of the most used classifiers are described as Decision tree, Logistic Regression, Naive Bayes, SVM, KNN and Random Forest. After testing and training the data, various parameters like Precision, Accuracy, Recall, Error rate and Specificity are calculated. The following ML algorithms are used in this study for lung cancer detection [5-7].

Decision Trees: Classification involves two steps learning and prediction. In this learning, the model is developed based on certain training data. In the prediction, the model is used to predict the answer for given data. It is the easiest and most prevalent classification algorithm to understand and interpret the data. It belongs to supervised learning [8]. It is used for solving regression and classification problems. The aim of this is to produce a training model that can be used to predict the class.

Logistic Regression: It is further classifies into following three categories. Binomial will have two probable types of dependent variables, such as 1 or 0, Pass or Fail, etc. Multinomial will have more than two likely unordered types of the dependent variable, such as cats, dogs, or sheep. Ordinal will have more than two likely ordered types of dependent variables, such as little, Moderate, or Extreme.

Naive Bayes: Bayes theorem used for solving problems related to classification. It is primarily used in text classification that includes high-dimensional training data. It's a simple and utmost active Classification algorithm which helps in quick predictions based on the fast ML models. It is a probabilistic classifier; to predict the basis of the probability of an object.

SVM: It's a popular Supervised Learning algorithm and used for Classification and Regression problems. The main intention of the SVM is to produce the best line which can separate N-dimensional space into modules, hence it can easily have kept the new data in the correct category in future. This supreme decision boundary is called a hyper plane. SVM indicates the extreme points/vectors that help in generating the hyper plane. The great points are called support vectors, and hence it is labeled as SVM.

KNN: Two types of KNNs are used in this work and they are Fine KNN and Medium KNN. K-NN is the simplest Supervised Learning. K-NN algorithm accepts and stores the similar data and available data and placed the new into the category that is most alike to the obtainable categories. It means when new data give the idea then it can be easily classified into a fine set category by using K- NN algorithm [9].

Random Forest Algorithm: It is a prevalent ML algorithm that comes under supervised learning. Classification and Regression can be done using this. It is based on ensemble learning and it uses multiple classifiers to resolve complex problems and to progress its performance.

Fig. 1 reperesents the proposed working model. Fig. 2 represents the flow chart of the proposed workflow.

C. Proposed CNN Architecture

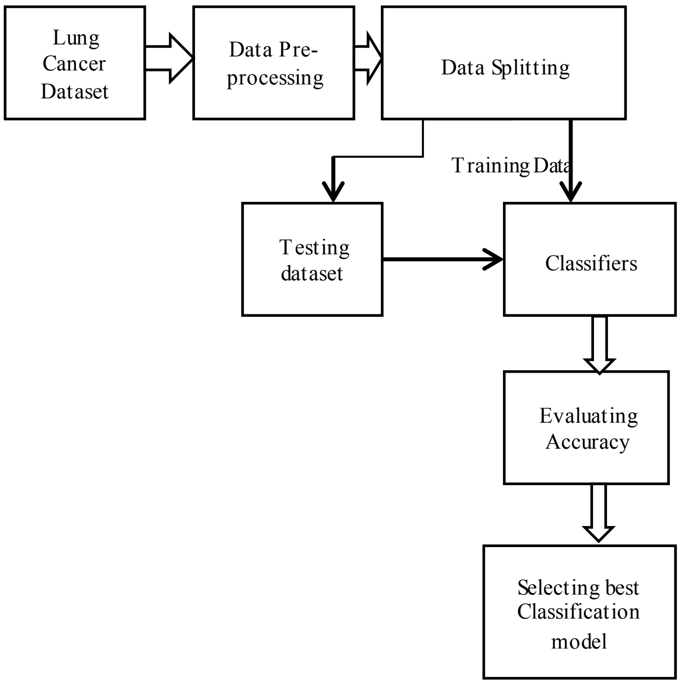


Fig. 1. Block diagram

Figure 1 shows the block diagram of proposed work. It describes the process involved in the implementation of proposed work [10].

Firstly, the lung cancer dataset is gathered, and pre-processing is performed. The data is trained under different ML classifiers. Accuracy for different test and train ratios is observed. Comparison of all the classifiers is done. Attributes that do not cause any change to accuracy are removed. Again, train the new data with different test and train ratios for different classifiers. Compare the performances of obtained new and previous data. Repeat the steps until we get more accuracy with les number of attributes [11].

D. FLOW Chart

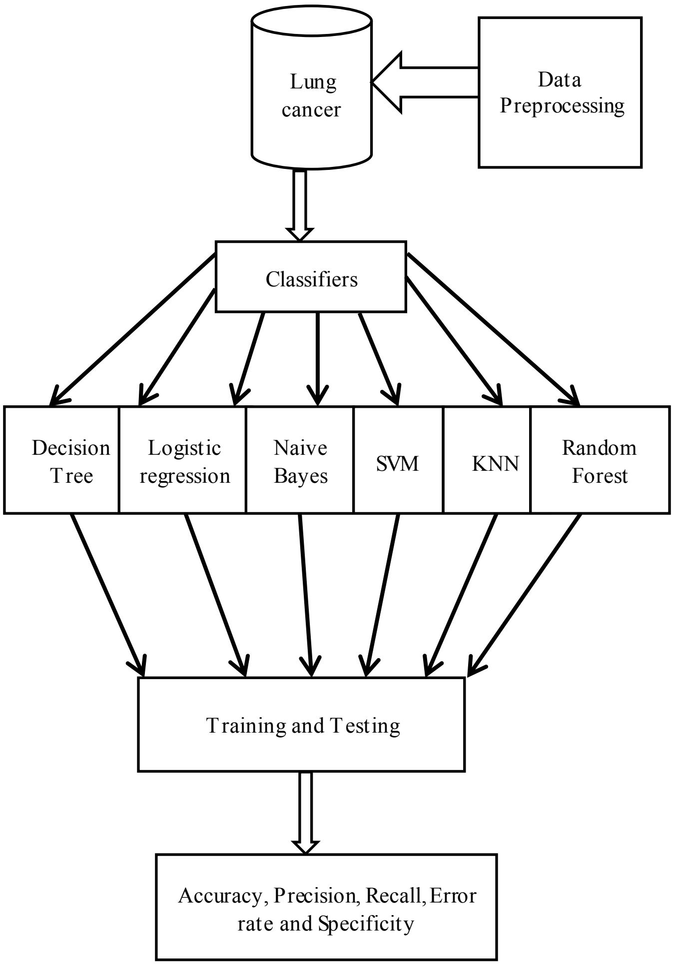


Fig. 2. Flow chart

Figure 2 shows the overall research flow of the proposed work

E. Performance Metrics

In general confusion matrix has key parameters like Accuracy, Recall, Precision and F-Measure for classification. Accuracy is the correct measure of predictions made from total predictions. These Quantitative parameters depend on specific outcome. Those are True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) respectively [12-13]. Performance metrics are evaluated using the following equations.

Accuracy: It defines model correct predictions. It's the ratio of the Total no. of correct predictions made by the classifier to all the Total no. of predictions made by the classifiers. The formula is as given below:

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

Precision: It defines the no. of correct outputs provided by the model, how many of them were true. The formula is as given below:

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

Recall: It defines total positive classes predicted correctly. It must be as high as possible. The formula is as given below:

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

Error Rate: It tells how frequently the classifier is wrong.

$$Error\ Rate = \frac{(FP + FN)}{(TP + TN + FP + FN)} \qquad (4)$$

Specificity: It tells how well a test can identify true negatives.

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

F-measure: It is a measure of the accuracy of the test.

$$F1\ Score = \frac{2 * precision * Recall}{(Precision + Recall)} \qquad (6)$$

Four distinct statistical indices—True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) are used to measure precision and recall.

III. RESULTS AND DISCUSSION

In the examination of lung cancer detection using ML classifiers, an in-depth investigation was conducted with an 80% training rate and a 20% testing rate, alongside various performance metrics. Through the utilization of diverse performance metrics such as accuracy, precision, recall, error rate, specificity, F-measure, and testing time, a comprehensive analysis was carried out to evaluate the classifiers' abilities in distinguishing between cancerous and non-cancerous cases.

In the multiclass classification scenario with 24 attributes, the confusion matrix reveals the distribution of predictions across different classes. Fig. 3 and Fig. 4 depicts the confusions matrices with 24 attributes and 20 attributes of the DT classifiers. Table I and Table II summarizes key performance metrics, including accuracy, precision, recall, F1-score, specificity, error rate, and testing time with 24 and 20 attributes respectively.

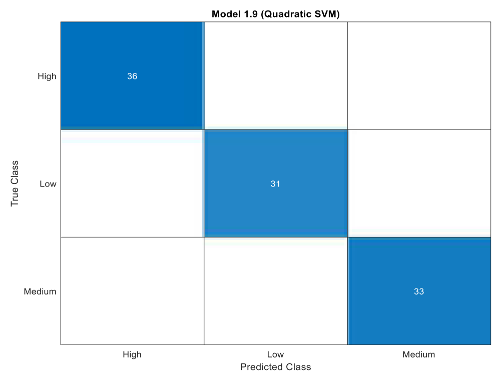


Fig. 3. Confusion matrix analysis with 24 attributes

Fig. 5 and Fig. 6 represent the performance of ML classifiers with 24 and 20 attributes respectively.

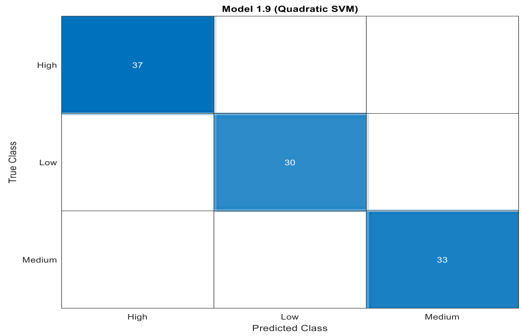


Fig. 4. Confusion matrix analysis with 20 attributes

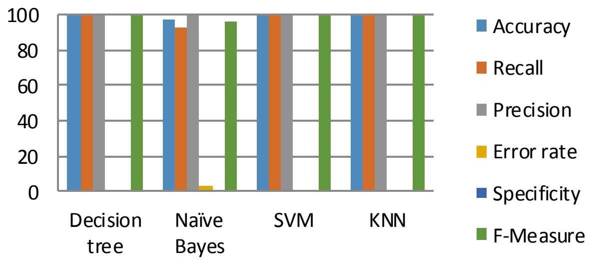


Fig. 5. Analysis for 24 attributes

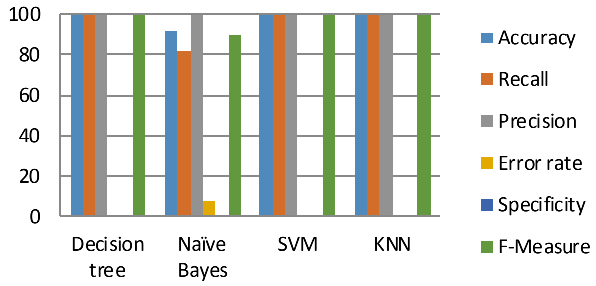


Fig. 6. Analysis for 20 attributes

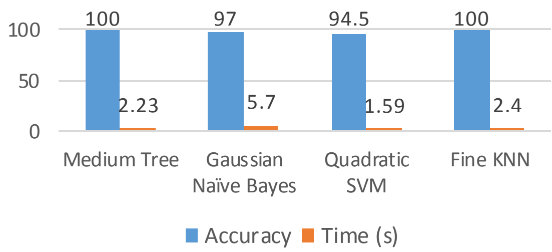


Fig. 7. Comparison with 24 attributes

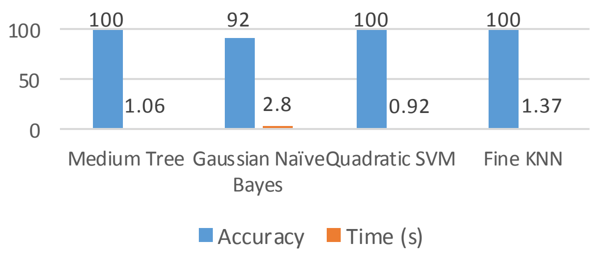


Fig. 8. Comparison with 20 attributes

Fig. 7 and Fig. 8 represent the performance comparison of best 4 classifiers. From the various sub classifications present for decision tree, Naïve Bayes and SVM -- Fine tree, Gaussian Naïve Bayes and Linear SVM have been considered as it has given better accuracy than the other methods.

TABLE I. PERFORMANCE ANALAYSIS WITH 24 ATTRIBUTES

Model Type	Accuracy (%)	Precision	Recall	Error Rate (%)	Specificity	F-Measure	Time (s)
DT	100	1	1	0	1	1	2.23
Naive Bayes	97	1	0.92	3	1	0.96	5.7
SVM	94.5	0.975	0.96	5.4	0.7	0.97	1.59
KNN	100	1	1	0	1	1	2.4

TABLE II. PERFORMANCE ANALAYSIS WITH 20 ATTRIBUTES

Model Type	Accuracy (%)	Precision	Recall	Error Rate (%)	Specificity	F-Measure	Time (s)
DT	100	100	100	0	1	1	1.06
Naive Bayes	92	100	82.2	8	1	0.90	2.8
SVM	100	100	100	0	1	1	0.92
KNN	100	100	100	0	1	1	1.37

IV. CONCLUSION

The multiclass classification experiment with 24 attributes revealed the remarkable accuracy of machine learning classifiers, particularly evident in the perfect scores achieved by Decision Tree, SVM, and KNN classifiers, while Naïve Bayes exhibited a slightly lower but still commendable accuracy of 97%. Upon reducing the attribute count to 20, Decision Tree, SVM, and KNN maintained their flawless accuracies, highlighting their robust performance even with fewer features. However, Naïve Bayes saw a decrease in accuracy to 92%. Notably, the SVM classifier consistently outperformed others across both datasets, demonstrating superior accuracy while requiring minimal computational time. These findings highlight the efficacy of SVM in multiclass classification tasks and emphasize the importance of selecting appropriate algorithms to achieve accurate classifications, contributing valuable insights to the optimization of classification techniques across diverse domains.

REFERENCES

[1] Janee Alam, Sabrina Alam, Alamgir Hossan, “Multi-Stage Lung cancer detection and prediction using multi-class SVM classifier”, International Journal of Innovative Technology and Exploring Engineering ,2019.

[2] Qing Wu and Wenbing Zhao, “Small-Cell Lung Cancer Detection Using a Supervised Machine Learning Algorithm”, International Symposium on Computer Science and Intelligent Controls, 2017.

[3] Syed Saba Raoof, M A. Jabbar, Syed Aley Fathima, “Lung cancer prediction using Machine Learning, 2020

[4] Vaishnavi. D1, Arya. K. S2, Devi Abirami. T3, M. N. Kavitha, “Lung Cancer Detection using Machine Learning”, International Journal of Engineering Research & Technology (IJERT), 2019.

[5] Ozge Gunaydin, Melike Gunay, Oznur Şengel, “Comparision of lung cancer detection algorithms”, Institute of Electrical and Electronics Engineers (IEEE),2019.
[6] Naji Khosravan and Ulas Bagci,” Semi-Supervised Multi-Task Learning for Lung Cancer Diagnosis”, Center for Resaerch in Computer Vision (CRCV).2019.
[7] Emine CENGİL, Ahmet ÇINARA, “Deep Learning Based Approach to Lung Cancer Identification”, Institute of electrical and electronics engineers 2017.
[8] Radhanath Patra, “Prediction of Lung Cancer Using Machine Learning Classifier”, International Research Journal of Engineering and Technology,2020.
[9] M. V. Subbarao, J. T. S. Sindhu, Y. C. A. Padmanabha Reddy, V. Ravuri, K. P. Vasavi & G. C. Ram. (2023). Performance Analysis of Feature Selection Algorithms in the Classification of Dry Beans using KNN and Neural Networks, 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 539-545.
[10] M. Prema Kumar, V. Veerraju, M. Venkata Subbarao, G. Challa Ram “Performance Evaluation of Different Machine Learning Algorithms for the Detection of Lung Cancer” NEUROQUANTOLOGY Vol. 20 Issue 12 October 2022 Pages: 2692-2699. DOI: 10.14704/NQ.2022.20.12. NQ77262
[11] Doma, M. K., Padmanandam, K., Tambvekar, S., K., K. K., Abdualgalil, B., & Thakur, R. N. (2022). Artificial Intelligence-Based Breast Cancer Detection Using WPSO. International Journal of Operations Research and Information Systems (IJORIS), 13(2), 1-16. http://doi.org/10.4018/IJORIS.306195.
[12] M. L. V. A. Priya and M. V. Subbarao, "Performance Analysis of Machine Learning Techniques for Multi-Organ Cancer Detection and Classification: A Comparative Study," 2023 International Conference on Emerging Research in Computational Science (ICERCS), Coimbatore, India, 2023, pp. 1-7, doi: 10.1109/ICERCS57948.2023.10434014.
[13] Ravuri, M. V. Subbarao, T. S. Kumar, T. S. Kavitha, D. R. Sandeep and L. Dangeti, "Multi-Cancer Early Detection And Classification Using Machine Learning Based Approaches," 2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Bhilai, India, 2023, pp. 1-7, doi: 10.1109/ICAECT57570.2023.10117816.

Scientometric Analysis of Tomato Leaf Disease Classification using Artificial Intelligence

Sayed Abu Lais Ezaz Ahmed
Dept of ME-CSE,
Chandigarh University
Gharuan, Punjab-India
Email: aezaz166@gmail.com

Dr. Abhishek Kumar
Dept of CSE, Professor
Chandigarh University
Gharuan, Punjab-India
Email: abhishek.e14283@cumail.in

Dr. Sachin Ahuja
Dept of CSE, Professor
Chandigarh University
Gharuan, Punjab-India
Email: ed.engineering@cumail.in

***Abstract*—Timely detection of leaf issues is crucial for advancing the national economy. This study intends to acquire the knowledge on detecting the leaf-related diseases. Nevertheless, it is essential to note that plant diseases can impact many parts of the plant. This study covers the articles published related to the multispectral or hyperspectral imaging from the year 2010 to 2024 to analyze the agricultural diseases. Leaf diseases are categorized using machine learning and deep learning models. This study has created a workflow system to assist researchers. Commonly employed leaf disease prediction models encompass Random Forest (RF), Support Vector Machine (SVM), and Multiple Twin Support Vector Machine (MTM). Various conventional deep-learning models are commonly employed for performing leaf disease identification. The models encompassed in this list are CNNs, DCNNs, BPNNs, DenseNet, LeafNet, and LeNet. The three most efficient deep- learning models for leaf disease identification are CNN, VGG, and ResNet. Typically, the effectiveness of an algorithm is assessed by measuring its F1 score, accuracy, precision, and other metrics. This research study aims to assist researchers in identifying suitable machine learning and deep learning-based classifiers for diagnosing the leaf diseases.**

Keywords— Classification, Deep Learning, Leaf Disease, Machine Learning, Tomato

I. INTRODUCTION

Leaf diseases delay plant growth and diminish the agricultural yield. Leaf diseases are caused by various species, including yeast, bacteria, viruses, and others, throughout their life cycles. Several methods can be employed to identify and categorize factors that cause stress in leaves [1]. The initial stage involves visually inspecting it using one's eyes, which can be more challenging. Secondly, leaf stress can be examined manually or by machine learning (ML) methods. Many scientists widely believe that relying solely on eye inspection or the use of microscopes is inadequate in preventing the spread of the disease to leaves. Next, several scientists are considering employing machine learning algorithms to cultivate foliage. Utilizing historical data poses a problem in the early detection of leaf diseases before their spread [2]. Consequently, researchers in computer vision, machine learning, and image processing face challenges when studying diseases affecting leaves [3].

II. LEAF DISEASE AND ITS SYMPTOMS

Some of the common tomato leaf diseases are discussed in this section.

A. Bacterial Spots

All tomato-growing locations have Xanthomonas, which causes bacterial spots. Fruit defoliation, sunburn, and yield loss are signs of this bug. Bacterial spot disease threatens tomato output worldwide due to its wide range of temperatures and pathogens. Precipitation and temperatures between 75 and 86 degrees Fahrenheit aggravate sickness. Round, moist leaf lesions may be surrounded by pale yellow. Stems and leaves often develop round or spherical markings ranging from dark brown to black. Spots rarely grow larger than 3 mm. More leaf lesions may cause the leaves to seem blighted and turn yellow, brown, wither, or die. Figure 1 shows bacterial spots on tomato leaves.

Fig. 1. Bacterial Spots on Tomato Leaf

B. Mosaic Virus

Tomatoes can contract the mosaic virus in many ways. Residual virus particles in soil usually cause disease. Live or dead plant debris can contain the virus for 50 years. No treatment or cure exists for the tomato mosaic virus. Healthy plants can rapidly catch viruses from ill ones. Farmers can infect good plants by touching ill ones. As they move from sick to healthy plants, insects may also spread disease. The virus can spread the virus through gardening equipment, pots, and plants in affected areas. The mosaic virus generates pale and dark green patches on tomato leaves. Reduced yield, tomato fruit deformities, and reduced development are some signs [4]. Sometimes the leaves coil up, turn yellow, and resemble ferns. The mosaic virus-infected tomato leaf is shown in Figure 2.

C. Septoria Spots

Septoria lycopersici spots are fungus-caused. Septoria spots, a severe tomato plant disease, are more widespread in places with prolonged high humidity and rain. Septoria spots on the lower leaves indicate fruit development. The patches' tan to grey interiors with little black fruiting structures and dark brown borders. Spherical or

Fig. 2. Mosaic virus on tomato's leaf

round, their diameters are 1/16 to 1/4 inch. A leaf usually has several spots. The infection will spread from the lower to upper leaves. As lesions increase, the leaf will turn green, light yellow, brown, and wither [5]. Most cases of this disease rarely infect fruit. Septoria spots on a tomato leaf are shown in Figure 3.

Fig. 3. Septoria spots on tomato's leaf

D. Yellow Curl

The Tomato Yellow Leaf Curl Virus (TYLCV) is a Geminiviridae DNA virus responsible for one of the most severe diseases affecting tomatoes. This is especially common in tropical and subtropical settings. Aleyrodidae insects spread this virus. TYLCV infects tobacco, beans, eggplants, peppers, potatoes, and tomatoes. Due to the rapid spread of TYLCV, research on its detection and control has increased. The tomato leaf with a yellow curl is shown in Figure 4.

Fig. 4. Yellow curl on tomato's leaf

E. Concerns with the detection of leaf diseases

1) Image processing challenges that need to be addressed: Even though there are a lot of different image processing algorithms [5], almost every IP methodology that is used to identify and classify diseases of leaves has the following problems:

- Noise significantly reduces picture processing efficiency. Stage lighting and electrical appliances can be disturbed. Pulse or Gaussian noise in leaf images can make forecasting more challenging.
- Electronic devices tend to take grainy photos, which hinders this investigation. Defocused photos blur, making this image unreadable.
- Most researchers use public datasets. It's a bother to capture images of diseased leaves in fields, but that's another reason.
- Variations in sunshine, camera angle, distance between electronic equipment and diseased leaves, and other occlusion-based images are among these researchers' many challenges.
- Smartphones, digital cameras, etc., may take multiple photos for different setups. Selecting the correct technological device is vital. The biggest issue with mobile phone photos is sharpness. Although their angle and distance constraints apply, digital cameras may also photograph sick leaves.
- Choosing a large photo for the study is also tricky. Even if every detail in the big picture matters, blurring can hinder leaf disease identification.
- Many researchers have used drones to capture damaged leaves in agricultural areas. This camera's main drawbacks are weather and flying time. Despite these difficulties, researchers have started employing drone cameras and generated classification results using various model setups. No drone-camera photographs were used for this article.

Box [6], Gaussian [7], Gradient [8], and Laplacian [9] image filters prevent IP difficulties caused by noise. Thresholding, Blob extraction, template matching, Hough transformation, and HoG transformation have helped diagnose leaf diseases. Random disease patterns or abnormalities make leaf image feature extraction harder. Popular feature extraction algorithms quickly extract information from sick leaf photos, including background, damaged, and green sections. The image's multi-form and multi-colour composition makes leaf disease identification difficult. Leaf disease identification is difficult using existing approaches. Sunny or overcast weather is another challenge. These constraints make leaf disease identification harder.

2) Challenges that arise in response to climate change: Weather changes threaten plants at different stages. Weather affects leaves too. As temperatures changed, leaf infection spread. The study [10] discussed how climate change affects farming and machine learning algorithms predict crop production. Rapid weather change shortens plant longevity and complicates leaf and plant disease detection. The study [11] described how climatic change and black Sigatoka disease have increased banana leaf fungus. Studies show that the risk of Black Sigatoka sickness is rising, and climate change is already affecting people's health. The study [12] stressed the importance of planting at the proper time to avoid pests and diseases.

3) The effects of invasive organisms and infectious agents: Throughout the year, a diverse range of pests and pathogens, such as viruses, bacteria, and fungi, have a substantial impact on the emergence of leaf and plant diseases. The study [13] found that dryness influenced the size of leaf apertures, which serve as the main entry routes for plant diseases. The study [14] quantified the extent of harm inflicted by pests and diseases on rice, wheat, maize, potatoes, and soybeans The study [15] examined the impact of climate and diseases on the life cycles of plants. Storms and heavy rain can also serve as a means of dispersal for plant diseases and pests. According to the study [16], animals and plants are interconnected and cannot be separated, even when plant diseases and pests are present. In their study [17], the authors introduced the term "fungal endophytic" to describe a defense system that protects against non-insect organisms such as worms, viruses, bacteria, and mites.

III. Searching Process

This section describes the selection process and criteria for this review's papers.

A. Planning

The inquiry included many 2010–2022 journal and conference articles. Keyword searches were used to query Google Scholar, IEEE Xplore, and SCOPUS Indexed Journal. Table 1 shows the searched keywords.

TABLE I
THE NUMBER OF ARTICLEE OBTAINED USING EACH KEYWORD

Keywords that were searched:	Number of article
Image-processing	7
Deep-learning	8
Maximum_Sequence_Length	7
Plant-disease-classification	6
Machine-learning	6
Convolutional-neural-network	7
Compute-Vision	7
Total	48

A total of forty-eight papers were identified for inclusion and removal using keywords.

B. Conduction

At this level, we will outline the criteria used to assess the effectiveness of current models that utilize machine learning, image processing, and deep learning (such as CNN) to detect diseases in different crops and plants. These models rely on diverse datasets. Figure 5 illustrates the complete research method undertaken for this work as per PRISMA guidelines. After doing a search on Google Scholar, IEEE Xplore, and SCOPUS Indexed Journal, a total of 48 papers related to the detection and categorization of plant diseases from 2010 to 2024 were found. Three stages were necessary to achieve exclusion. After thoroughly reviewing the entire content, a total of seventy-five papers were found. Out of these, one hundred and sixty-four papers were dismissed only based on their titles. Abstracts and conclusions condensed the publications. Table 2 presents a concise overview of the seven study questions that formed the foundation of this comprehensive review. To obtain answers, all accessible models were closely monitored.

C. Material and Methods

To conduct a comprehensive assessment of articles and draw conclusions based on research questions, it is crucial to develop specific criteria for determining which publications should be included or excluded, as well as criteria for evaluating their quality. Table 3 presents the criteria for including (IC) and excluding (EC) certain elements.

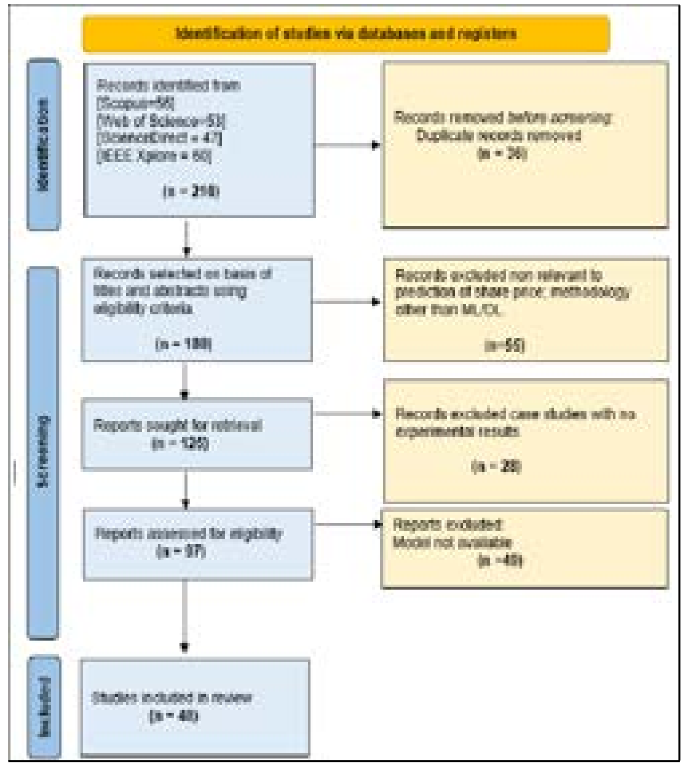


Fig. 5. PRISMA flowchart for selection of the articles tomato's leaf disease classification

IV. Literature Review

The study technique findings in this part answer the aforementioned research questions using data from chosen studies. These automated models take a long time to train, but they are good at recognizing early indicators of plant ailments, allowing farmers to take precautions and avoid agricultural production losses. The review considered several parameters, as indicated in Figure 6. The technique employed for this literature evaluation encompassed several key steps: data gathering, pre-processing, data augmentation, feature extraction, identification and classification of feature types, picture quality enhancement, reduction of model overfitting, and classification and identification of features.

A. RQ1: Regarding plant data, what is the primary source of information derived from?

Identifying and classifying leaf disease begins with plant leaf imaging. You can take plant leaf photos with a camera or use open-source repositories. This section will discuss the sources from which researchers collected image data. We analyzed 16 papers since they supplied this area's data. Using a substrate manufactured by Klassmann-Deilmann GmbH (Germany), the study [18] investigated sugar beetroot leaves that were grown in plastic pots. In addition to being watered, the plants were fertilized once a week with 100 milliliters of Poly Crescal that contained 0.2 percent. The hyperspectral reflectance method was utilized in the study to categorize plant diseases.

The study [18] trained and identified two diseases affecting wheat and grape crops using reverse propagation networks and 185 digital camera images. There were one hundred images of wheat stripe rust, fifty of grape powdery mildew, thirty-five of grape downy mildew, and thirty-five of grape powdery mildew in the collection.

The study [19] captured 107 images of chile leaves using the LABVIEW IMAQ Vision software; 21 of the leaves were healthy,

TABLE II
RESEARCH QUESTIONS AND THEIR MOTIVES

S. No.	Research Question	Motivation
1	Regarding plant data, what is the primary source of information derived from?	Scientists utilize diverse data collection sources to acquire photographs of plants.
2	What is the total number of unique pre-processing procedures employed? ?	Discover multiple pre-processing methodologies.
3	What are the steps to utilize different techniques for data augmentation?	The objective is to compile a comprehensive list of the different techniques employed to expand the dataset through data augmentation.
4	Which attributes are most commonly utilized in feature extraction? ?	To discern several extracted characteristics.
5	What are the techniques for improving the quality of an image through analytical methods?	To explore techniques employed to improve the quality of images.
6	Where can one access information regarding automated systems designed to identify and categorize plant diseases?	To find models used for the identification and categorization of plant diseases.
7	What strategies can be employed to mitigate or eradicate the impact of overfitting?	The objective is to identify strategies to mitigate the occurrence of overfitting.

TABLE III
INCLUSION-EXCLUSION
CRITERIA

S. No.	Parameters	Inclusion Criteria (IC)	Exclusion Criteria (EC)
1	Period	Between 2010 to 2024 research studies	Article published before 2019
2	Investigation	Studies focuses on the dataset, proposed methodology, and results of share market	Studies focuses on other than Tomato Classification
3	Compactor	Studies aims to stock security price prediction	Studies focus on other than stock security price prediction
4	Methodology	Studies focus on ML or DL techniques	Studies focuses on other than ML or DL technique
5	Design of study	Original studies comprises experiment results	Survey paper, case studies and non-english

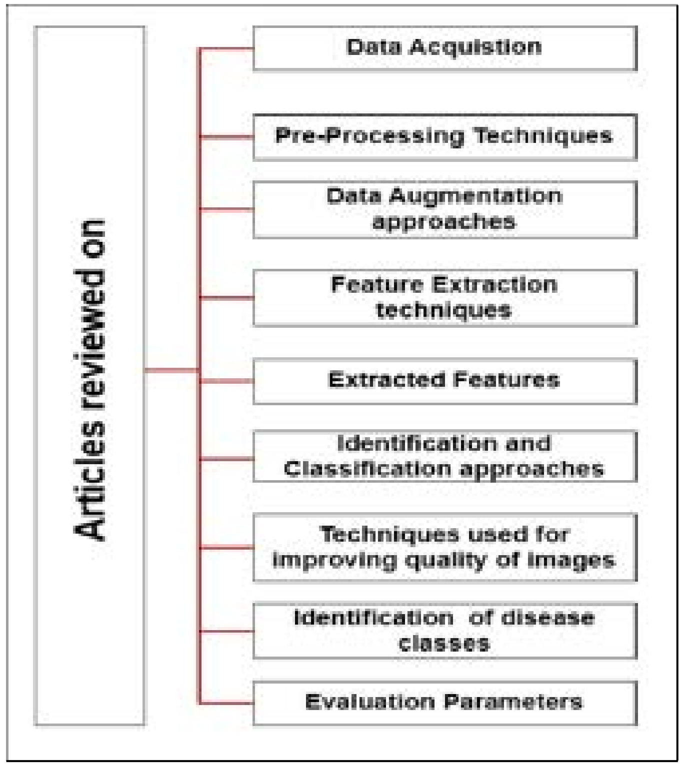


Fig. 6. The parameters that were explored for the literature review.

while 86 were infected. Under the supervision of professionals, The study [20] photographed grape foliage in Pune, Bijapur, and Sangali using a Nikon Coolpix P510 digital camera with 16.1 Megapixels. The images were utilized for testing and training systems. The image files were all in the.jpg extension. An assortment of web photographs were employed in order to establish an inclusive and diverse atmosphere. It has been demonstrated that powdery and downy mildew, two prevalent diseases in India, cause damage to the foliage.

Photographs of multiple farms were captured by [21] utilizing a digital camera and a guide. The dataset experienced expansion due to the acquisition of images from the internet and the implementation of diverse configurations. Photographs of leaf miners, thrips, and Tuta absoluta are displayed. To consolidate data, [22] conducted a research on plant disease detection. Fifteen distinct categories were established for the Internet photographs. Twelve of the categories featured photographs of leaf-based plant diseases, whereas two featured images of healthy foliage and backgrounds.

Dataset 1 comprised a total of 7320 images depicting both healthy and diseased foliage, while Dataset 2 comprised 7520 photographs. The study [23] examined real-world photographs in addition to six publicly accessible datasets. A compilation of 10,213 handheld mobile phone photographs was produced, showcasing 22 distinct species in diverse soil, illumination, and pixel density conditions.

In their controlled study, [24] used a public dataset (PlantVillage dataset on Kaggle) to classify 54,306 pictures of healthy and diseased leaves into 38 classes. In the study, [25] photographed eight tomato leaf classes using [26]. One photo showed healthy leaves and nine showed sick leaves. [27] obtained 14,800 photographs of tomato leaves with nine diseases from PlantVillage, an open-access library with over 50,000 photos. The study [28] captured 1053 photos of apple leaves using a BM-500 GE/BB-500 GE digital camera at two apple experiment locations in China. The images depicted mosaic, Alternaria leaf spot, brown spot, and rust. The various sources of data acquisition are depicted in Figure 7.

B. *RQ2: What is the total number of unique pre-processing procedures employed?*

Various methods were utilized to pre-process image data. This section covers the several pre-processing procedures used by researchers. Applying the filter "Pre-processing techniques," 34 papers were retrieved for this area, 26 of which were studied. The typical application of different pre-processing procedures is illustrated in Figure 8. The study [29] removed noise using picture improvement, smoothing, clipping, colour conversion, and histogram equalization. The study [30] resized and cropped image The study [31] utilized

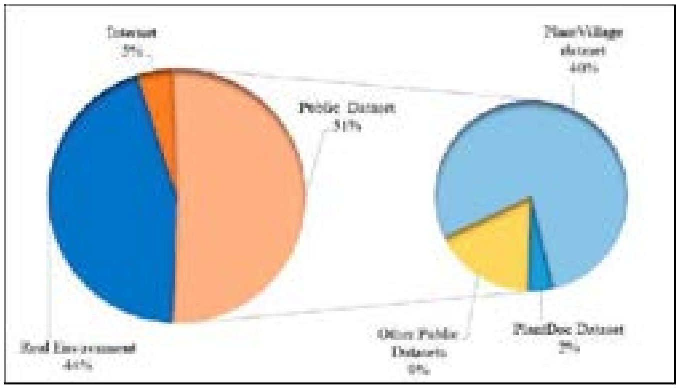


Fig. 7. Selected articles for quality assessment

a median filter to eliminate noise after compressing pictures to a standard size to improve quality. Before feature extraction, [32] decreased the image size to 256 × 256 pixels and trimmed plant leaves to find areas of interest. The study [?] undertook various pre-processing steps to improve the final product, including clipping the image to identify essential parts and smoothing it. Increased contrast was achieved with picture enhancement. [34] deleted pixels from grape leaf margins and translated RGB data from data acquisition. The study [35] pre-processed image data to 256 × 256 pixels, including size reduction and cropping. [35] pre-processed photos to standardize sizes. Two photo pre-processing approaches were used by [36]. Image improvement increased contrast and revealed new features and CIELAB colour space compressed the photographs to speed up processing.

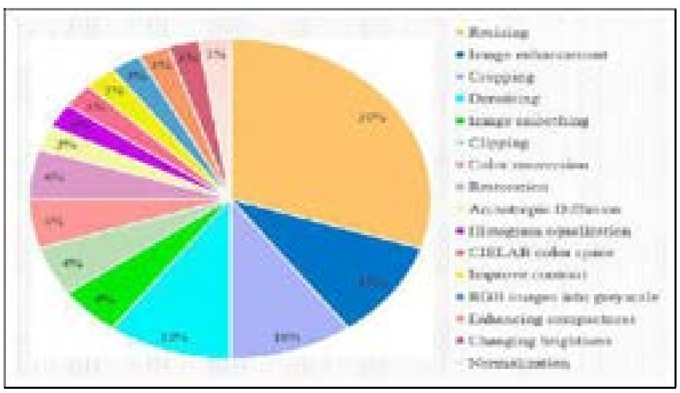


Fig. 8. Comparison of several pre-processing methods' usage graphs, ordered by decreasing percentage

C. RQ3: What are the steps to utilize different techniques for data augmentation?

Data augmentation can include photos to improve dataset accuracy. This section discusses how researchers have expanded their datasets. Twenty-eight data-augmented papers were selected for study in this domain, shown in Figure 9.

The study [37] used rotations, affine transformations, and a 3x3 matrix-based perspective transformation to improve photos. [38] expanded the dataset by using picture shifting, mirroring, and image rotation. The study [39] augmented the photos in the training dataset by applying geometric and intensity editing techniques. Intensity transformation deals with issues such as noise, color, brightness, and contrast, while geometrical transformation focuses on scaling, cropping, rotating, and horizontally flipping images. By employing rotation and flipping techniques, [40] were able to collect further data. In addition to being rotated 90, 180, and 270 degrees, the photographs in the collection were also inverted both horizontally and vertically. [41] increased the quantity of photographs by rotating them clockwise and cropping them from the center to the extremities. By implementing rotation, shear transformation, truncation, and horizontal and vertical alignment. The techniques employed for intensity variations include increased contrast, color jittering, principal component analysis jittering, radial blur, and brightness enhancement. The images underwent enlargement, cropping, rotation, and horizontal and vertical flipping. [42] employed image augmentation techniques such as RandomRotate, RandomFlip, and RandomLighting to improve the quality of photos and facilitate the assessment of multiple viewpoints. The study [42] improved the image data by applying techniques such as cropping, flipping, shifting, and rotating.

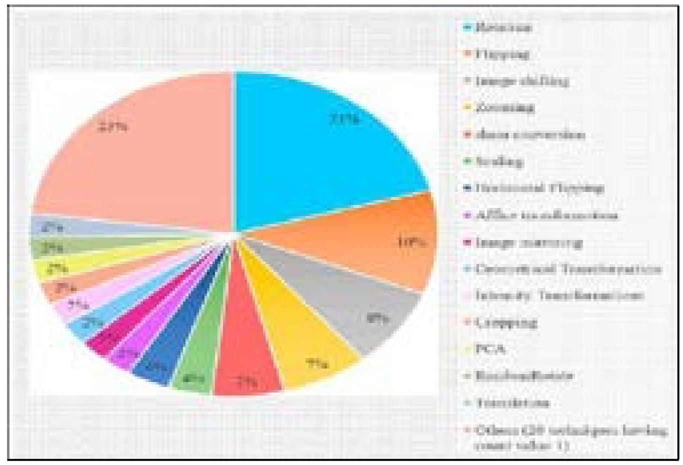


Fig. 9. Total number of distinct augmentation methods that have been implemented

D. RQ4: Which attributes are most commonly utilized in feature extraction?

This section examines the feature extraction methodologies employed by researchers. A total of twenty-six publications were identified that utilized feature extraction for the purpose of filtering. Out of them, a detailed analysis is conducted on nineteen papers.

In the study [44], the authors successfully distinguished the foliage of chili plants through the implementation of color space and a reduction in the intensity of optical irradiation. Image extraction entailed the retrieval of color information, object recognition, and photo comparison. The study [45] extracted texture data via color cooccurrence. The study [46] extracted characteristics utilizing a gray-level co-occurrence matrix. The following concepts were implemented: complexity, area, circularity, and perimeter. Color co-occurrence was utilized by the study [47] to derive features. The study [48] extracted textural features utilizing GLCM. The study [10] generated feature vectors via HOG feature extraction. The study [32] extracted features utilizing the RESNET18 (CNN) architecture and task-adaptive methods. The study [?] extracted features utilizing a scale-invariant feature transform. The study [10] focused on the color and texture components of the photographic data. The color moment equation calculates the skewness, standard deviation, kurtosis, and mean of the retrieved color attributes. The second set of features was obtained through the retrieval of the Gray Level Co-occurrence Matrix (GLCM). Pandian et al. extracted attributes utilizing a number of exceptional convolutional layers in the study [9]. The study [19] extracted data from photographs utilizing a Convolutional Neural Network (CNN) called MobileNet.

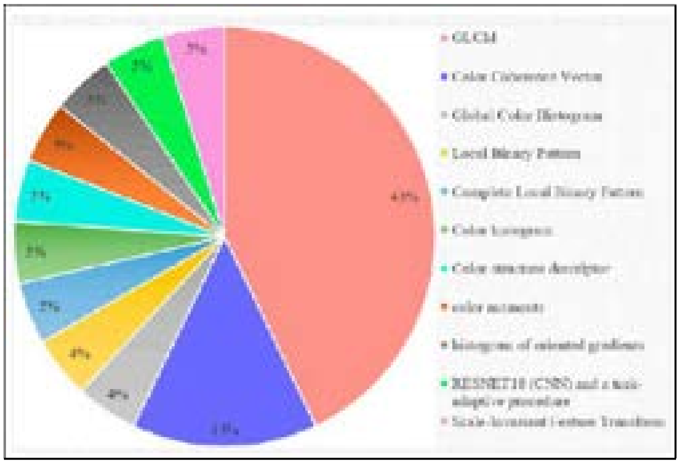


Fig. 10. Selected articles for quality assessment

E. RQ5: What are the techniques for improving the quality of an image through analytical methods?

This section will focus on the enhancements made by researchers to enhance the quality of images. After applying filters that improve the quality of the pictures, a total of 15 publications were identified for this particular section. In the end, eleven papers were included in the research, shown in Figure 11. In the study [21], the authors cleaned up pictures of broken grape and wheat leaves with a median filter. The study [24] employed colour conversion and histogram equalization to enhance their photos. The RGB colour model was transformed to greyscale using the colour conversion method. Histogram equalization dramatically improved image visibility. The study [27] used histogram equalization to increase image quality. The study [32] improved grayscale photographs by using the following formula:

$$S_{k,l} = (T_{k,l} - min(T))/(max(T) - min(T)) \tag{1}$$

This equation uses the original and updated pixel values (T and S) and pixel indices (k,l).

The study [**?**] modified the dataset to better pre-process input images and [32] improve image quality via preprocessing. The researchers in the study [40] improved model performance and image quality with LeafGAN data augmentation. The study [13] boosted photographs by focusing on contrast and brightness in pre-processing. The study [16] reduced picture noise with a bilateral filter. To reduce noise, bilateral filters were applied to plant leaf images.

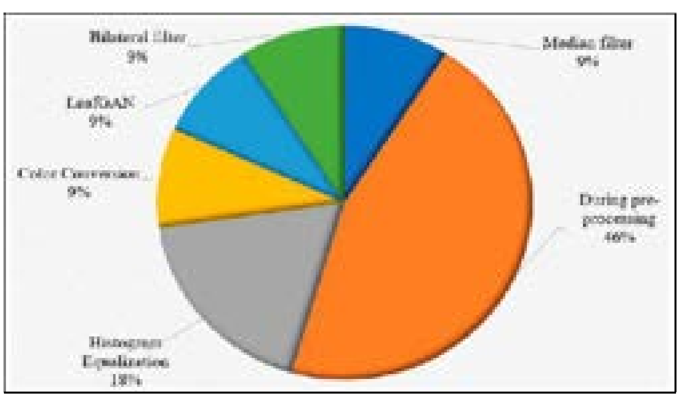


Fig. 11. Various techniques used to enhance image quality.

F. RQ6: Where can one access information regarding automated systems designed to identify and categorize plant diseases?

This section covers machine learning and deep learning techniques for disease identification and categorization. Forty-five studies were found using preexisting automated methods for this set of plant diseases. As seen in Figure 12, the review included 37 papers. Powdery mildew, sugar beet rust, and Cercospora leaf spot were all identified using a support vector machine (SVM) trained on sugar beet plants with hyperspectral data [16]. Through the use of back-propagation networks (BPNN), the inquiry led to the identification of two diseases affecting the leaves of grapes and wheat. Researchers in the study detected a disease in chilli leaves [by using image processing]. The research used multi-class support vector machines to classify apple blotch, scab, and disease. Three sugar beetroot leaf diseases were identified using spectral indicators in the study. To differentiate between powdery and downy fungus on grape leaves, the study employed a feed-forward back propagation neural network. The study used a sequential application of two support vector machines to identify leaf damage caused by thrips, Tuta absoluta, leaf miners, late blight, and powdery mildew. The authors of the study used Convolutional Neural Networks (CNNs) to classify seven different types of cucumber leaf diseases. The authors used two deep learning models to detect various diseases affecting tomato leaves: SqueezeNet and AlexNet. Some examples of these diseases are target spot, early blight, bacterial spot, late blight, and yellow leaf curl. The discovery of nine diseases affecting tomato leaf was CNN's contribution to the research [21]. To identify apple diseases such brown spot, mosaic, rust, and Alternaria leaf spot, the study used AlexNet's deep Convolutional Neural Network (CNN). This study detected and categorized twenty-five different plant diseases using deep learning-based convolutional neural network (CNN) models. The study used a random forest classifier to identify diseases affecting papaya leaves. The research [**?**] used a deep Convolutional Neural Network (CNN) to identify four cucumber leaf markings: downy, anthracnose, powdery, and target. This study used a Convolutional Neural Network (CNN) model that included learning vector quantization (LVQ) to detect and categorize four types of tomato leaf diseases. The study found that oranges have brown rot, citrus canker, stubborn, and melanoses [using support vector machines and K-means clustering]. By applying imperfect reasoning, the severity of the ailment was determined in the diagnosis. To identify diseases in thirteen different plant species, the study used a complex Convolutional Neural Network (CNN) with nine layers. Researchers in this study used a CNN to identify diseases in tomato and apple leaves. Study participants were screened for nitrate and cotton leaf diseases using neural networks and image processing techniques.

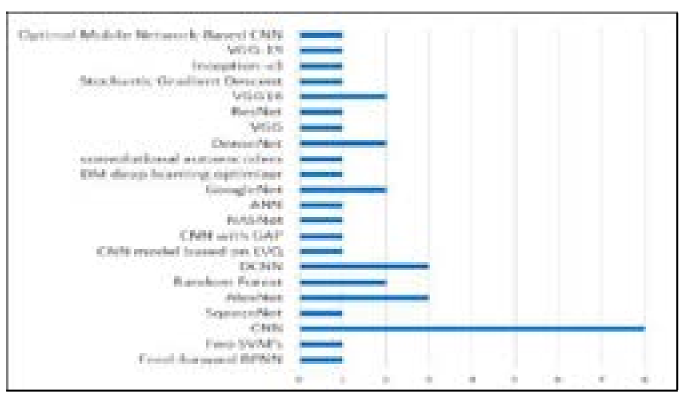


Fig. 12. Multiple methods employed for categorization

G. RQ7: What strategies can be employed to mitigate or eradicate the impact of overfitting?

Overfitting occurs when there is a considerable discrepancy between the accuracy of a model on its training dataset and its testing dataset. This section focuses on the authors' various tactics for reducing overfitting. After applying filtering techniques to reduce overfitting, twenty-four publications were identified for this subsection. Figure 13 displays the final analysis, which comprised a total of twenty-one research. To reduce the occurrence of overfitting, the study [12] used picture distortion methods throughout the augmentation process. Researchers in the study used data-rotation and - flipping methods to reduce the likelihood of overfitting. This work used activation function layers to make the model more non-linear and pooling and dropout layers to make it less overfit. Overfitting was avoided in the study by enhancing the data. The problem of model overfitting was tackled in the study by adding additional images to the training dataset, using response-normalizing layers for local normalization, and replacing certain fully connected layers with convolution layers. To reduce the effects of overfitting and improve the clarity of pictures of cucumber leaves, the researchers in this work used data augmentation methods. To avoid overfitting, the study [5] included altered images in the training dataset. The study [45] used a dropout rate of 0.25 to prevent overfitting. The study [46] employed dropout, data augmentation, and an early stop strategy to mitigate overfitting. Howlader et al. The study [48] mitigated overfitting by employing data augmentation and utilizing ReLU activation. The ReLU activation function for N neurons is F(N) = Max(0, N). The study [47] implemented early stopping with a patience parameter of 5.

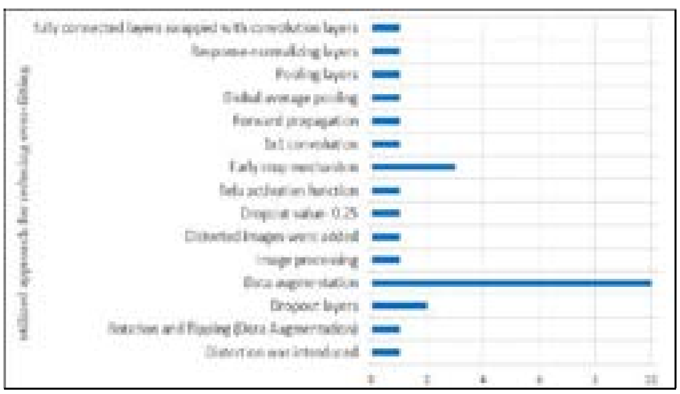


Fig. 13. Methods used to lessen the occurrence of overfitting

V. CONCLUSIONS AND FUTURE SCOPE

This study examined ML, DL, and image processing methods to assess the state of plant disease diagnosis. After significant research, 75 relevant publications were chosen for this review. The following were given top priority: sources of data, preliminary processing, extraction of features, enhancement of data designs, and model performance. According to the results, most models are only partially capable of dealing with unstructured image input. Determining the targeted impacted area from a complex visual background required methodical engineering and professional design abilities. This poll encouraged scientists to classify plant diseases using image process- ing, ML, and DL. Leaf images were utilized in most studies to detect disease. Future frames may include many leaves. These photos can be taken in different settings (temperature, humidity, etc.) to reduce environmental impacts on disease diagnosis, and new disease stage indicators can be produced. Mobile apps and drones can detect plant diseases early in broad agricultural areas in the future.

REFERENCES

[1] Malathy, S.; Karthiga, R.; Swetha, K.; Preethi, G. Disease Detection in Fruits Using Image Processing. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20–22 January 2021; pp. 747–752.

[2] Rumpf, T.; Mahlein, A.K.; Steiner, U.; Oerke, E.C.; Dehne, H.W.; Plu¨mer, L. Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. Comput. Electron. Agric. 2010, 74, 91–99.

[3] Dubey, S.R.; Jalal, A.S. Detection and Classification of Apple Fruit Diseases Using Complete Local Binary Patterns. In Proceedings of the 2012 3rd International Conference on Computer and Communication Technology, Allahabad, India, 23–25 November 2012; pp. 346–351.

[4] Behera, S.K.; Jena, L.; Rath, A.K.; Sethy, P.K. Disease Classification and Grading of Orange Using Machine Learning and Fuzzy Logic. In Proceedings of the 2018 IEEE International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 3–5 April 2018; pp. 0678–0682.

[5] Tulshan, A.S.; Raul, N. Plant Leaf Disease Detection Using Machine Learning. In Proceedings of the 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 6–8 July 2019; pp. 1–6.

[6] Wahab, A.H.B.A.; Zahari, R.; Lim, T.H. Detecting Diseases in Chilli Plants Using K-Means Segmented Support Vector Machine. In Proceedings of the 2019 3rd International Conference on Imaging, Signal Processing and Communication (ICISPC), Singapore, 27–29 July 2019; pp. 57–61.

[7] Fujita, E.; Kawasaki, Y.; Uga, H.; Kagiwada, S.; Iyatomi, H. Basic Investigation on a Robust and Practical Plant Diagnostic System. In Proceedings of the 2016 15th IEEE International Conference on Machine Learning and Applications (ICMLA), Anaheim, CA, USA, 18–20 December 2016; pp. 989–992.

[8] Fuentes, A.; Yoon, S.; Kim, S.C.; Park, D.S. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors 2017, 17, 2022

[9] Ma, J.; Du, K.; Zheng, F.; Zhang, L.; Gong, Z.; Sun, Z. A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. Comput. Electron. Agric. 2018, 154, 18–24.

[10] Adedoja, A.; Owolawi, P.A.; Mapayi, T. Deep Learning Based on NASNet for Plant Disease Recognition Using Leave Images. In Proceedings of the 2019 International Conference on Advances in Big Data, Computing and Data Communication Systems (icABCD), Winterton, South Africa, 5–6 August 2019; pp. 1–5.

[11] Zhang, S.; Zhang, S.; Zhang, C.; Wang, X.; Shi, Y. Cucumber leaf disease identification with global pooling dilated convolutional neural network. Comput. Electron. Agric. 2019, 162, 422–430.

[12] Marzougui, F.; Elleuch, M.; Kherallah, M. A Deep CNN Approach for Plant Disease Detection. In Proceedings of the 2020 21st International Arab Conference on Information Technology (ACIT), Giza, Egypt, 28–30 November 2020; pp. 1–6.

[13] Shrestha, G.; Deepsikha; Das, M.; Dey, N. Plant Disease Detection Using CNN. In Proceedings of the 2020 IEEE Applied Signal Processing Conference (ASPCON), Kolkata, India, 7–9 October 2020; pp. 109–113.

[14] Jadhav, S.B.; Udupi, V.R.; Patil, S.B. Identification of plant diseases using convolutional neural networks. Int. J. Inf. Technol. 2021, 13, 2461–2470.

[15] Sun, Y.; Liu, Y.; Zhou, H.; Hu, H. Plant Diseases Identification through a Discount Momentum Optimizer in Deep Learning. Appl. Sci. 2021, 11, 9468

[16] Abbas, A.; Jain, S.; Gour, M.; Vankudothu, S. Tomato plant disease detection using transfer learning with C-GAN syntheticimages. Comput. Electron. Agric. 2021, 187, 106279

[17] Divakar, S.; Bhattacharjee, A.; Priyadarshini, R. Smote-DL: A Deep Learning Based Plant Disease Detection Method. In Proceedings of the 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2–4 April 2021; pp. 1–6.

[18] Akshai, K.P.; Anitha, J. Plant Disease Classification Using Deep Learning. In Proceedings of the 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Coimbatore, India, 13–14 May 2021; pp. 407–411.

[19] Kibriya, H.; Rafique, R.; Ahmad, W.; Adnan, S.M. Tomato Leaf Disease Detection Using Convolution Neural Network. In Proceedings of the 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST), Islamabad, Pakistan,12–16 January 2021; pp. 346–351.

[20] Pandian, J.A.; Kanchanadevi, K.; Kumar, V.D.; Jasinska, E.; Gono, R.; Leonowicz, Z.; Jasinski, M. A Five Convolutional Layer Deep Convolutional Neural Network for Plant Leaf Disease Detection. Electronics

2022, 11, 1266

[21] Wang, H.; Li, G.; Ma, Z.; Li, X. Image Recognition of Plant Diseases Based on Backpropagation Networks. In Proceedings of the 2012 5th International Congress on Image and Signal Processing, Chongqing, China, 16–18 October 2012; pp. 894–900

[22] Husin, Z.B.; Shakaff, A.Y.B.M.; Aziz, A.H.B.A.; Farook, R.B.S.M. Feasibility Study on Plant Chili Disease Detection Using Image Processing Techniques. In Proceedings of the 2012 Third International Conference on Intelligent Systems Modelling and Simulation, Kota Kinabalu, Malaysia, 8–10 February 2012; pp. 291–296.

[23] Sannakki, S.S.; Rajpurohit, V.S.; Nargund, V.B.; Kulkarni, P. Diagnosis and Classification of Grape Leaf Diseases Using Neural Networks. In Proceedings of the 2013 4th International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 4–6 July 2013; pp. 1–5.

[24] Es-saady, Y.; El Massi, I.; El Yassa, M.; Mammass, D.; Benazoun, A. Automatic Recognition of Plant Leaves Diseases Based on Serial Combination of Two SVM Classifiers. In Proceedings of the 2016 International Conference on Electrical and Information Technologies (ICEIT), Tangiers, Morocco, 4–7 May 2016; pp. 561–566

[25] Dyrmann, M.; Karstoft, H.; Midtiby, H.S. Plant species classification using deep convolutional neural network. Biosyst. Eng. 2016, 151, 72–80.

[26] Mohanty, S.P.; Hughes, D.P.; Salathé, M. Using deep learning for image-based plant disease detection. Front. Plant Sci. 2016, 7, 1419.

[27] Durmus, H.; Gunes, E.O.; Kirci, M. Disease Detection on the Leaves of the Tomato Plants by Using Deep Learning. In Proceedings of the 2017 6th International Conference on Agro-Geoinformatics, Fairfax, VA, USA, 7–10 August 2017; pp. 1–5.

[28] Liu, B.; Zhang, Y.; He, D.; Li, Y. Identification of Apple Leaf Diseases Based on Deep Convolutional Neural Networks. Symmetry 2017, 10, 11.

[29] Atila, Ü.; Uçar, M.; Akyol, K.; Uçar, E. Plant leaf disease classification using EfficientNet deep learning model. Ecol. Inform. 2021, 61, 101182.

[30] KC, K.; Yin, Z.; Wu, M.; Wu, Z. Depthwise separable convolution architectures for plant disease classification. Comput. Electron. Agric. 2019, 165, 104948

[31] Al Haque, A.S.M.F.; Hafiz, R.; Hakim, M.A.; Islam, G.M.R. A Computer Vision System for Guava Disease Detection and Recommend Curative Solution Using Deep Learning Approach. In Proceedings of the 2019 22nd International Conference on Computer and Information Technology (ICCIT), Dhaka, Bangladesh, 18–20 December 2019; pp. 1–6.

[32] Sahithya, V.; Saivihari, B.; Vamsi, V.K.; Reddy, P.S.; Balamurugan, K. GUI Based Detection of Unhealthy Leaves Using Image Processing Techniques. In Proceedings of the 2019 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 4–6 April 2019; pp. 0818–0822

[33] Ponnusamy, V.; Coumaran, A.; Shunmugam, A.S.; Rajaram, K.; Senthil- velavan, S. Smart Glass: Real-Time Leaf Disease Detection using YOLO Transfer Learning. In Proceedings of the 2020 IEEE International Confer- ence on Communication and Signal Processing (ICCSP), Chennai, India, 28–30 July 2020; pp. 1150–1154.

[34] Wajid, A. H., Saher, N., Nawaz, S. A., Arshad, M., & Nasir, M. (2024). Tomato Leaf Disease Detection and Classification Using Convolutional Neural Network and Machine Learning. Journal of Computing & Biomed- ical Informatics.

[35] Kansal, S., Jaiswal, A., & Sachdeva, N. (2024, January). Empirical Analysis of Deep Learning Models for Tomato Leaf Disease Detection. In 2024 14th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 430-435). IEEE.

[36] Sreedevi, A., & Manike, C. (2024). A smart solution for tomato leaf disease classification by modified recurrent neural network with severity computation. Cybernetics and Systems, 55(2), 409-449.

[37] Ahmad, W., Adnan, S. M., & Irtaza, A. (2024). Local triangular-ternary pattern: a novel feature descriptor for plant leaf disease detection. Multimedia Tools and Applications, 83(7), 20215-20241.

[38] Thangaraj, R., Pandiyan, P., Anandamurugan, S., & Rajendar, S. (2024). A deep convolution neural network model based on feature concatenation approach for classification of tomato leaf disease. Multimedia Tools and Applications, 83(7), 18803-18827.

[39] Najim, M. H., Abdulateef, S. K., & Alasadi, A. H. (2024). Early detection of tomato leaf diseases based on deep learning techniques. Int J Artif Intell, 13(1), 509-515.

[40] Lekha, J., Saraswathi, S., Suryaprabha, D., & Thomas, N. (2024, March). Tomato Leaf Disease Detection using Machine Learning Model. In Proceedings of the 1st International Conference on Artificial Intelligence, Communication, IoT, Data Engineering and Security, IACIDS 2023, 23- 25 November 2023, Lavasa, Pune, India.

[41] Anu, A., Solanki, K., & Dhankar, A. (2024). A Novel Multi-Class Deep Learning Approach for Tomato Leaf Disease Detection System. Inter- national Journal of Intelligent Systems and Applications in Engineering, 12(6s), 187-196.

[42] Jothika, G., & Khilar, R. (2024, January). Improving Disease Detection in Tomato Leaves Using Visual Geometry Group 19 Compared to the Artificial Neural Network Algorithm. In 2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) (pp. 1-5). IEEE.

[43] Huang, X., Chen, A., Zhou, G., Zhang, X., Wang, J., Peng, N., ... & Jiang, C. (2023). Tomato leaf disease detection system based on FC-SNDPN. Multimedia tools and applications, 82(2), 2121-2144.

[44] Rahman, S. U., Alam, F., Ahmad, N., & Arshad, S. (2023). Image processing based system for the detection, identification and treatment of tomato leaf diseases. Multimedia Tools and Applications, 82(6), 9431-9445.

[45] Tang, Z., He, X., Zhou, G., Chen, A., Wang, Y., Li, L., & Hu, Y. (2023). A Precise Image-Based Tomato Leaf Disease Detection Approach Using PLPNet. Plant Phenomics, 5, 0042.

[46] Zhong, Y., Teng, Z., & Tong, M. (2023). LightMixer: A novel lightweight convolutional neural network for tomato disease detection. Frontiers in Plant Science, 14, 1166296.

[47] Roy, K., Chaudhuri, S. S., Frnda, J., Bandopadhyay, S., Ray, I. J., Banerjee, S., & Nedoma, J. (2023). Detection of tomato leaf diseases for agro-based industries using novel PCA DeepNet. IEEE Access, 11, 14983-15001.

[48] Ulutaş, H., & Aslantaş, V. (2023). Design of efficient methods for the detection of tomato leaf disease utilizing proposed ensemble CNN model. Electronics, 12(4), 827.

Random Forest based Fake Job Detection

Spandhana Reddy Akiti
Department of Electronics and Computer Engineering
Sreenidhi Institute of Science and Technology
Hyderabad, India
akitispandhanareddy96@gmail.com

Akash Bathini
Department of Electronics and Computer Engineering
Sreenidhi Institute of Science and Technology
Hyderabad, India
akashgoud434@gmail.com

Sateesh Kumar Kanagala*
Department of Electronics and Computer Engineering
Sreenidhi Institute of Science and Technology
Hyderabad, India
sateeshkumarkanagala@gmail.com,

0000-0001-6233-7790

***Abstract*— The intersection of computer science and social sciences leads an increasing problem: fraudulent job ads. Their presence can have far-reaching negative ramifications that must be dealt with immediately in order to limit further harm to society. Due to the vast amounts of data accumulated worldwide, effective ways must be found for distinguishing legitimate job advertisements from fraudulent ones. This paper proposes a machine learning approach designed to differentiate between fraudulent and non-fraudulent job postings, offering an automated tool to mitigate the proliferation of fraudulent job listings on the internet. The methodology relies on a range of machine learning classification techniques to analyze web-based job postings, comparing the efficacy of different classifiers in identifying employment scams. The primary objective is to establish a robust model for the detection of fake job posts amidst the overwhelming volume of online listings. This research recommends employing various data mining and categorizing algorithms such as Decision Tree and Support Vector Machine, Naive Bayes classifier, Random-forest classifier and Multilayer Perceptron to detect whether an advertisement is genuine. This study has utilized Kaggle data with the proposed classifier enhancing 99.48% accuracy when classifying fake job ads. Additionally, this method provides techniques to detect employment scams to meet applicants' need to protect themselves against scammers.**

Keywords—Fraudulent job ads, Machine learning, Proliferation, Employment scam

I. INTRODUCTION

As unemployment rates increase worldwide, job seekers face an ever increasing number of unsuitable job advertisements during COVID-19 pandemic and its aftermath - creating an atmosphere in which many unsuitable ads clog job boards worldwide. Job hunters worldwide face a grave danger as unemployment benefits for 26 million Americans are being received and similar issues can be seen around the globe, leaving job hunters ever vulnerable. Modern advances in technology and business have opened up numerous job opportunities for job hunters. Thanks to job listings advertisements, job candidates are able to explore potential matches for working schedule, skills or experience requirements and potential applicability. Social media plays an integral role in shaping and contributing to the recruitment industry and its success. Internet-based social networks such as Twitter and Facebook play a pivotal part in recruitment processes while social media ads continue to provide opportunities to share jobs [1]. Due to an increase in websites offering job adverts, an increasing number of fake job adverts that causes significant harm have become prevalent and resulted in harassing job seekers [2]. Security and data integrity concerns have left people feeling less assured when responding to job ads. Establishing legitimacy when posting job advertisements to social and electronic media platforms is no simple task due to the difficulty involved in building trust among employers and employees [3][4]. Technology has simplified our lives while simultaneously improving workplace conditions. When used effectively to spot fake job ads for jobs that don't exist will have a hugely positive effect on recruitment of new employees. False job advertisements often complicate applicants' job search processes and lead them down a path of wasted time and energy. An automated system capable of detecting fraudulent job advertisements provides an unprecedented opportunity to address human resource management problems. This research employs Machine Learning (ML) and Natural Language Processing (NLP) approaches to study the issue of fake job opening descriptions posted online with Machine Learning (ML) and NLP approaches. Researchers conducted their investigation using information set with over 17,000 fake and genuine job vacancies posted within an information set of 17,000 job listings [5].The data for this research came directly from the company's official web site [6].When individuals come across false information online, they often believe and accept it, providing an effective strategy for using social media data as evidence against an opposing viewpoint [7,11].Action Fraud from UK conducted a recent survey that showed 67% of those searching for jobs online or through advertisements may unwittingly open themselves up to fake job ads that may be scams. Over 700000 job seekers reported losses totaling $5000+ as victims of scam jobs over two years in Britain [8], and fresh graduates tend to be targeted more heavily due to scammers taking advantage of them due to being willing to pay higher amounts [9,10].

This paper details a technique for detecting fake jobs using machine learning techniques with various data processing algorithms to measure accuracy, recall and F1 scores of each algorithm.

This research is divided into sections to illustrate the work; in the initial one examined prior work that has been accomplished in this same field before conducting literature reviews and proposing system and its process for implementation; below that comes an outline of proposed system with its process for implementation; finally, conclusion presented the research findings along with analysis between current methods and proposed solution.

II. RELATED WORKS

Habiba et. al. [1] suggested that employing various data mining and classification algorithms like Decision Tree, Support Vector Machine, Naive Bayes classifier Random forest classifier multilayer Perceptron and KNN to assess job ads as authentic or fake. Amaar et. al [2] employed six model-based machine learning models to evaluate job advertisements. One challenge encountered during their research related to data sets, as the supply/demand imbalance between genuine/fake job postings wasn't equal - causing models to focus too much on targeting majority class target classes rather than evenly distributed classes throughout. To address this problem, they conducted two experiments concurrently using balanced datasets and one with unbalanced ones. Mehboob et. al. [3] takes an innovative approach to recruitment fraud detection by providing key elements from organizations, job descriptions, and compensation structures as indicators for potential scams or fraud. A gradient-boosting method-based model was then constructed and tested against three examples using three distinct comparison methods before employing two-step feature selection method; results revealed how organization type could serve as an excellent predictor model. Ranparia et. al. [4] successfully reduced fraudulent job advertisements using Machine Learning techniques that employ NLP for tone and format analysis in job advertisements as well as training the model as a Sequential Neural Network using Global Vector algorithm. For real world testing will utilize an algorithm trained to predict LinkedIn job openings as a test to measure accuracy in real world situations; further improving adaptability and resilience is ongoing development process. Sudhakar et. al. [5] have developed an innovative algorithm for distinguishing fake news from legitimate news by employing logistic regression SVM and an ensemble technique based on machine-learning algorithms. Their study included 620 samples per group with 10,000 records classified into fake, actual, or false classified as such by binary classification. Their results show an innovative ensemble technique can achieve 95% accuracy with only 5% loss rate when compared to other methods; showing its effectiveness as an algorithm which utilized decision tree methods combined with AdaBoost for more precise results by altering parameters accordingly [7]. But main demerit encloses by existing system is using of dataset which possess less accuracy rate than the other. Job hunters worldwide face a grave danger as unemployment benefits for 26 million Americans are being received and similar issues can be seen around the globe, leaving job hunters ever vulnerable. As a response, scammers create fraudulent job ads which appear genuine but present serious risks to individuals searching for work. There are solutions available, including fraud detection tools available through websites or job websites. Job seekers employ various strategies to identify and avoid fraudsters when applying for employment, including recognizing company logos in job ads, using only official email accounts for communication, and being especially careful when providing personal details during interviews. EMSCAD (the Employment Scam Aegean Dataset) is an accessible dataset containing 17,880 job advertisements.

III. PROPOSED METHODOLOGY

This article presents a method for authenticating job ads and aiding job seekers in recognizing legitimate employment opportunities. This research employed Kaggle database which provide extensive details about potentially suspicious job postings; Figure. 1 displays this elaborate procedure flow for this process.

A multi-step procedure must be undertaken first to create a balanced data set before classifiers. Preprocessing methods, including eliminating redundant data or terms as well as any extra values or spaces, may need to be utilized to ensure efficient categorical encoding and creation of feature vectors for various classification algorithms. Contrasting, with current models, this proposed system offers unique features. After uploading data, users can select their variables. Preprocessing actions depend on the selected problem; depending on this action, Decision Trees, Support Vector Machine, Naive Bayes, Random Forest Algorithm and Multilayer Perceptron Algorithmic Operations are executed, followed by results defined using Accuracy, Recall Precision and F1 Score produced as outcomes of this analysis process.

Fig.1.Proposed Methodology of Fake Job Detection

A. Dataset Details

In this research Kaggle dataset is used to conduct this investigation and to detect fake job post. This provided details regarding job postings which may be suspicious and the schema serves as the foundation for evaluating proposed methods of investigation. Important pre-processing techniques included filling missing values with zeros, eliminating unnecessary attributes and handling spaces between records to ensure categorical encoding could take place successfully and then creating feature vectors suitable for classification algorithms to use. Kaggle dataset is more useful because it contains a broad variety of job postings from different sectors and industries. Figure 2 presents the structure of the dataset as the foundation for testing for the proposed method, ensuring effective evaluation.

#	Column	Non-Null Count	Dtype
0	job_id	17880 non-null	int64
1	title	17880 non-null	object
2	location	17534 non-null	object
3	department	6333 non-null	object
4	salary_range	2868 non-null	object
5	company_profile	14572 non-null	object
6	description	17879 non-null	object
7	requirements	15185 non-null	object
8	benefits	10670 non-null	object
9	telecommuting	17880 non-null	int64
10	has_company_logo	17880 non-null	int64
11	has_questions	17880 non-null	int64
12	employment_type	14409 non-null	object
13	required_experience	10830 non-null	object
14	required_education	9775 non-null	object
15	industry	12977 non-null	object
16	function	11425 non-null	object
17	fraudulent	17880 non-null	int64

Fig. 2. Structure of the dataset

B. Implementation of Classifiers

a)Decision Tree Algorithm: The Decision Tree algorithm operates similar to flowcharts, posing inquiries to data in order to make informed decisions and take effective actions. Each question divides the dataset into smaller subsets based on responses, with subsequent inquiries further segmenting the data, forming branches until a final decision or outcome is reached. Known for its user-friendly and straightforward nature, the Decision Tree algorithm simplifies predictions while mitigating overfitting issues if implemented accurately.

b)Support Vector Machine (SVM): Support Vector Machine (SVMs) employs support vectors positioned near class boundaries to discern data points forming classes. Subsequently, SVMs optimize a hyperplane to augment distances between support vectors, each representing a data point. SVMs demonstrate exceptional efficiency and accuracy, particularly when handling smaller, less complex datasets.

c)Naive Bayes Classifier: Naive Bayes classifiers require several parameters that scale with the number of variables involved, making them highly adaptable. Contrary to many other classifiers that rely on iterative approximation techniques for classification, Naive Bayes employs maximum likelihood training which uses linear speed analysis of closed form equations for maximum likelihood analysis. Classifiers that assign class labels based on an assumption that independence exists are an efficient and easy way to build classifiers; not relying on features as is common with traditional classifiers so they allow rapid classification accurately and swiftly.

*d) Random forest Algorithm:*Random Forest employs an ensemble model composed of trees to produce predictions. A collection of sub-trees (decision trees) helps facilitate this prediction approach, with sub-trees created independently through "Bagging" approaches, taking bootstrap samples from training datasets as input into this independent training method. Once created, results from these sub-trees become the final output, which can then be utilized across situations for classification by simply substituting training sessions as needed.

e) Multilayer Perceptron (MLP): MLP approaches fall within the scope of artificial neural networks (ANNs), which leverage neural network structures to tackle classification challenges. MLP belongs to this class of ANNs and comprise three layers: input layer, hidden layer and output layer. Data can flow either forwards or backwards between these nodes, depending on how it traverses the network. This flexibility makes MLP an extremely adaptable approach.

C. Performance Evaluation Metrics

In classification prediction tasks, comparison with the state of art methods in all possible parameters listed: Accuracy, Recall, Precision, and F1 score.

IV. RESULTS AND DISCUSSION

This section details the research findings and discussions. In this paper implemented the work using Kaggle dataset in Jupyter Notebook. To improve the performance, this paper proposed strategies like Cross Validation and Ensemble methods and compared their performances across various parameters like Accuracy, Recall Precision as well as F1 scores. Analysis of these results enabled to identify which algorithm that exhibited superior performance in terms of Accuracy, Precision, Recall and F1 score over the other models; additionally, later combined reports from each algorithm's classification reports which helped understand comparison among the algorithms under the parameters. Finally, compared the results with state of art techniques using EMSCAD dataset. Precision, recall, and F1 score are lower compared to accuracy when dealing with imbalanced datasets (Kaggle) because accuracy alone doesn't account for class distribution.

In imbalanced data, where one class is dominant, accuracy can be high even if the model performs poorly on the minority class. The variance in performance metrics (Precision, Recall and F1 score) of machine learning models under Kaggle dataset compared to the EMSCAD dataset for fake job detection can be attributed to several key differences and challenges inherit to the datasets, as well as how the models were applied and tuned to them. By Model Selection and Cross Validation, we achieved high accuracy and with Error analysis and Interpretability it proven the reliability.

The Tables I, II, III and IV provides Accuracy, Precision Recall, and F1 score measurements for each algorithm respectively. This result is comparable to prior work that utilized machine learning algorithms, attaining similar classification accuracy with various algorithms. The acquired results under Kaggle dataset among the various parameters are listed below. It is observed that Random Forest algorithm reaching 99.48% classification accuracy to check against false positives and negative examples. The high accuracy proves the Random Forest Algorithm is very realistic and can be used directly for job posts. One needs to extract the job description and feed it to the model which will tell whether the job is fraudulent or real.

TABLE I. ACCURACY COMPARISON

	[1]	[2]	Proposed
Random Forest algorithm	96.5%	97%	99.48%
Multilayer Perceptron	96.2%	92%	93.30%
Support Vector Machine	95%	95%	92.78%
Naïve Bayes	91.35%	84%	97%
Decision tree	96.2%	96%	97.9%

TABLE II. PRECISION COMPARISON

	[1]	[2]	Proposed
Random Forest algorithm	93%	92.7%	100%
Multilayer Perceptron	94%	93.7%	60%
Naïve Bayes	95%	93%	93%
Decision Tree	93%	94%	83%

TABLE III. RECALL COMPARISON

	[1]	[2]	Proposed
Random Forest algorithm	92%	91%	92.86 %
Multilayer Perceptron	95%	91%	21.43 %
Naïve Bayes	96%	94%	56%
Decision Tree	95%	94%	68%

TABLE IV. F1 SCORE

	[1]	[2]	Proposed
Random Forest algorithm	93%	90%	96.3%
Multilayer Perceptron	93%	91%	31.58 %
Naïve Bayes	95%	90%	67%
Decision Tree	93%	92%	75%

V. CONCLUSION

At present, the unexpected events caused by COVID-19 have raised serious global health concerns and intensified fears of economic recession. Since layoffs have increased and demand outstrips supply when it comes to employment opportunities available on the market, today's economic climate provides ideal conditions for scammers and cyber-scam attacks are on the rise, creating much havoc and confusion within society. Recognizing employment scams is crucial in order to guarantee applicants receive legitimate solicitations from companies. To combat this problem, various machine learning techniques will be discussed as preventative measures in this paper. By taking advantage of a supervised approach and various classifiers, supervised classifiers are used to detect employment scams. Research has demonstrated that Random Forest classifier stands out among competitors when categorizing data accurately. This research utilizes machine learning algorithms such as SVM, Decision Tree, Naive Bayes, Random Forest and MLP for classification algorithms based on machine learning. The proposed methodology achieves an accuracy rate of 99.48%, precision rate of 100%, recall of 92.86% and F1 score of 96.3%for Random Forest Algorithm surpassing existing methods.

REFERENCES

[1] S. U. Habiba, M. K. Islam and F. Tasnim, "A Comparative Study on Fake Job Post Prediction Using Different Data Mining Techniques," 2021 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), 2021, pp. 543-546, doi: 10.1109/ICREST51555.2021.9331230.

[2] Amaar, A., Aljedaani, W., Rustam, F. et al. Detection of Fake Job Postings by Utilizing Machine Learning and Natural Language Processing Approaches. Neural Process Lett 54, 2219–2247 (2022). https://doi.org/10.1007/s11063-021-10727-z

[3] Mehboob, A., Malik, M.S.I. Smart Fraud Detection Framework for Job Recruitments. Arab J Sci Eng 46, 3067–3078 (2021). https://doi.org/10.1007/s13369-020-04998-2

[4] D. Ranparia, S. Kumari and A. Sahani, "Fake Job Prediction using Sequential Network," 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), 2020, pp. 339-343, doi: 10.1109/ICIIS51140.2020.9342738.

[5] Sudhakar, M., Kaliyamurthie, K.P. (2023). Efficient Prediction of Fake News Using Novel Ensemble Technique Based on Machine Learning Algorithm. In: Kaiser, M.S., Xie, J., Rathore, V.S. (eds) Information and Communication Technology for Competitive Strategies (ICTCS 2021). Lecture Notes in Networks and Systems, vol 401. Springer, Singapore.

[6] B. Alghamdi, F. Alharby, "An Intelligent Model for Online Recruitment Fraud Detection", Journal of Information Security, 2019, Vol 10, pp. 155-176, https://doi.org/10.4236/iis.2019.103009.

[7] Tin Van Huynh1, Kiet Van Nguyen, Ngan Luu-Thuy Nguyen1, and Anh Gia-Tuan Nguyen, "Job Prediction: From Deep Neural Network Models to Applications", RIVF International Conference on Computing and Communication Technologies (RIVF), 2020.

[8] Jiawei Zhang, Bowen Dong, Philip S. Yu, "FAKEDETECTOR: Effective Fake News Detection with Deep Diffusive Neural Network", IEEE 36th International Conference on Data Engineering (ICDE), 2020.

[9] Lee, J., et al. (2019). "Fake Job Detection using Deep Learning Architectures." International Joint Conference on Neural Networks.

[10] Wang, Z., et al. (2020). "Fake Job Detection on Social Media: A Graph-based Approach." IEEE Transactions on Knowledge and Data Engineering.

[11] Chen, X., et al. (2019). "Detecting Fake Job Postings: An Integrated Framework using Machine Learning and Network Analysis." Decision Support Systems.

[12] Wu, J., et al. (2020). "Fake Job Detection using Attention Mechanism in Neural Networks." International Conference on Artificial Intelligence.

Predicting Health Decline Related to Aging with Machine Learning: Anticipating Recurrence Trends in the Elderly

D Anudeep Naidu
Department of ECM , SNIST
Hyderabad, India
anudeepn855@gmail.com

Vattumilli Sai Sahithi
Department of ECM , SNIST
Hyderabad, India
sahithivattumilli03@gmail.com

Nanda kumar M
Department of ECM , SNIST
Hyderabad, India
nanda.mkumar12@gmail.com
Orcid id: 0000-0001-6967-6859

Abstract— The prevalence of age-related disease poses a significant threat in the modern world environment. The early identification of these diseases not only enhances the concern towards the elderly people's health monitoring but also increases the young people's awareness. Machine learning diseases are recognized for identifying and predicting the required output based on the given data. Although the initial study utilizing machine – learning for health monitoring indicated moderate precision, further improvements can be made. Random Forest tends to overfit noisy data, especially when the number of trees in the forest is large and the depth of the trees is not appropriately controlled. Overfitting can lead to poor generalization performance on unseen data, making it less effective in identifying age-related issues in new datasets. to improve the model with the use of machine learning frameworks like Random Forest integrated with the XGBoost algorithm with in-built crossvalidation. The RF can provide stability, while Gradient Boosting can improve the accuracy. The algorithm uses the tree-pruning approach to increase the performance.

Keywords— Age- related Disease, Health Monitoring, Machine Learning Framework, Cross Validation, Tree Pruning, Gradient Boosting

I. INTRODUCTION

In the advancing field of medical sciences, every disease has been cured with the help of growing technology and knowledge in the relevant fields. As the ancient field of medical history depicts the early-stage discovery of diseases would increase the healing rate. Also in this modern era, the chance of getting affected by a disease is irrespective of age, So the main idea is to predict that age-related diseases come into the picture with great weightage in showing the industry a great boost of fighting against the disease. The main focus is on diseases related to the cardiovascular, respiratory, disorders of the immune system, pneumonia, infectious, and mental health diseases. Diseases related to hereditary, that are difficult to recognize and harder to cure. The aim is to study and use a classification approach that can further classify the level of disease and its type based on the provided symptoms.[1]

Diseases that are related to mental health are proven to have occurred in an individual irrespective of his age and lifestyle. So, there is an improved chance of curing or stopping its side effects if they can be diagnosed early. Many individuals with a tender age have the chance of overcoming their health issues with this prediction idea. BMI also known as the body mass index is a main factor in identifying a person’s physical fitness and also diseases related to the cardiovascular region [2].

When coming to diagnosing these different types of diseases can be done with the help of Machine Learning Techniques, which are quite popular practices used for predicting the outcomes based on the given dataset. In this paper, we are going to use the models like Random Forest algorithm to predict the data based on decision trees with a high level of accuracy, with the XGBoost algorithm to further render accurate results [3][4].

With the help of these two models, the prediction of the data can have a chance of high increase. Coming to the dataset it should be essential to provide the data that is relevant to the outcome. we are using a dataset that can be given any type of data based on the specific field of predicting the disease. This can give a wide chance of predicting the outcome with a single machine learning algorithm.

This paper deals with developing a generalized algorithm that can be used to predict diseases irrespective of their type, this means that the model can be used by implementing the required data of a field, say heart disease, we can give the dataset the required data about the body mass index and other entries like fat percentage, blood oxygen levels, blood pressure. With this information, we can run the algorithm to predict the type of disease an individual can be affected with [7].

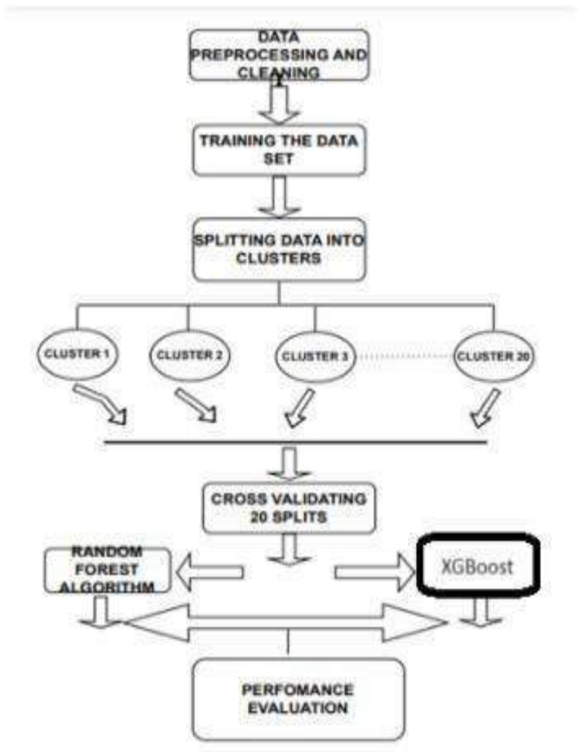


Fig. 1. Flow Diagram

In Fig. 1, the flow diagram depicts the process of employing a Random Forest model along with the XGBoost algorithm for classification tasks, while also incorporating feature e xtract ion techniques for improved performance assessment. Data is initially inputted, followed by feature e xtraction to identify relevant features. Subsequently, the dataset is split into train ing and testing sets. The model is then trained using the training data, and its performance is evaluated using the testing set, with binary cross -entropy serving as the loss function. Various performance metrics are employed to assess the model's effectiveness in classification tasks.

II. Analysis

A. *Analysis Using Random Forest*

Regarding the current model technique, the dataset has been divided into many folds using the classic Random Forest algorithm, which is based on the TensorFlow library. This has been done to increase the range of accuracy and performance evaluation. The primary focus of this methodology is to handle data using their class weights.

The most frequent problem in classification tasks is data imbalance. To solve this, decision trees are used, which have individual nodes that have been broken down into their simpler forms to effectively anticipate the outcomes based on the provided attributes. The primary purpose of parallel processing is to align data from related classes to balance out dataset dissimilarities. Taking into account all the pros and negatives of the current process, adding a small Enhancement can address the problems that can further increase the system's efficiency in terms of speed, accuracy, and class allocation estimation

B. *Classification*

These two aids in the classification of the collected data into two categories: those that indicate that the data appear to be weighted in the anticipated category and those that result in noisy or negative datasets that may impact data imbalance. The un-interfered data, often known as the undesirable data, is indicated by the noisy data [12].

The dataset we used includes testing and training data as well as an additional dataset that is used just as auxiliary metadata for the training set.

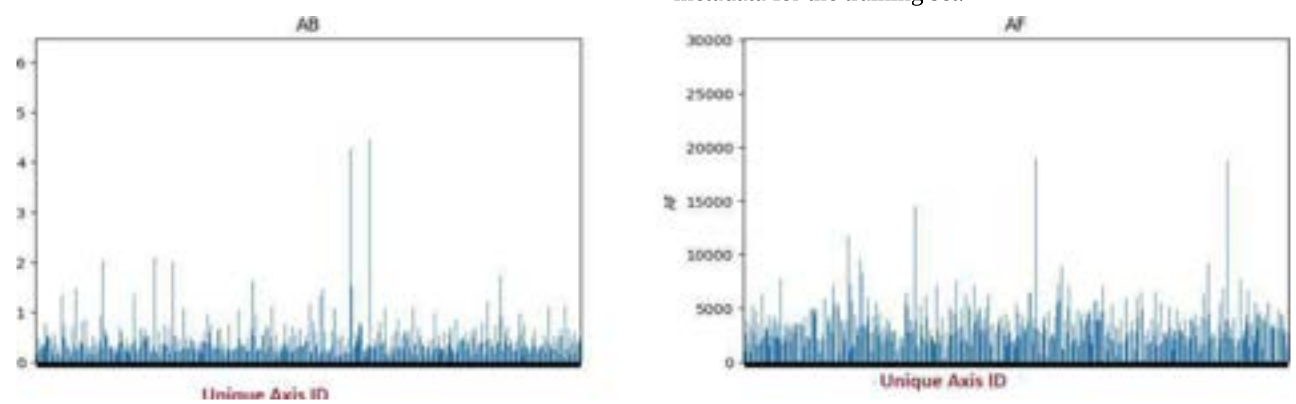


Fig. 2. Graphical representation of Id and anonymous characteristic

X-axis: Unique Id
Y-axis: Individual Characteristic Value

In Figure 2 the graphical representation illustrates the correlation between user ID and anonymous traits, encapsulating their relational dynamics succinctly. The above graphs are obtained by the code which is a grid of bar plots where each plot represents one of the first six numerical features from the dataset. Each subplot has the 'Id' column on the x-axis and the values of the respective numerical feature on the y-axis. This visualization allows for a comparison of these numerical features across different 'Id' values.

C. *Cross Validation*

Given all the considered points of pros and cons of the existing methodology, adding a minor improvement can deal with the issues, that can further boost the system's efficiency in terms of speed, precision, and estimation of the class allocation. This cross-entropy loss comes into the working after we deal with the data evaluation and visualization. Cross-validation is one of the best techniques to improve a model's efficiency and accuracy. Cross-validation is achieved by splitting the data set into k data sets of equal size. These sub-data sets are called folds in cross-validation. Choose one of the folds as the validation set and the remaining folds as the training set. Train your machine learning model on the train ing set. Evaluate the performance of the model on the validation set using a chosen metric which can be accuracy or precision or even F1 score. Calculate the average performance across all k folds. This aggregated performance metric gives you a more robust estimate of how well your model will generalize to unseen data. After cross-validation, you can retrain your model on the entire dataset.

The data can now be cross-validated to find its true positives and true negatives. These two play a role in classifying the acquired data into – whether the data seems to have the weightage in the predicted category or it leads to the negative or noisy datasets that can affect the data imbalance. The noisy data indicates unwanted data that can be termed as un-interfered or ineffective while predicting the disease under the person's health reports. So, to avoid this misconception we can improvise by adding the following formula.

III. DATA COLLECTION AND METHODS

The metadata includes several experimental qualities, Delta, Gamma, Beta, and Alpha. The presence of age-related conditions is indicated by the "Alpha" property, which is divided into two classes: A (no condition, class 0) and B, D, and G (age-related conditions, class 1).

It also has beta, gamma, and delta e xperimental properties. The test set is obtained after training, and "Epsilon" denotes the date of data collection for each subject.

You indeed need to be careful not to overfit when merging models, especially strong ones like XGBoost and Random Forest. When merging these models, consider the following tactics to reduce the chance of overfitting.

The parallel tree processing is successfully achieved using the XGboost algorithm.

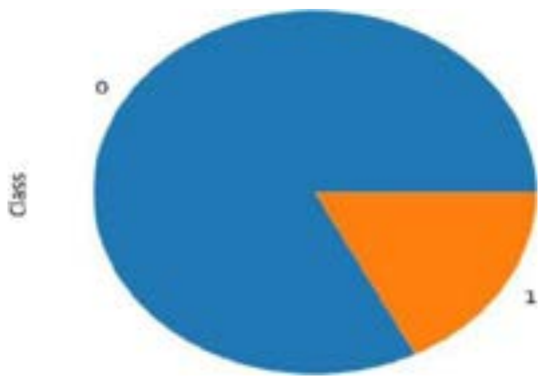


Fig. 3. Pie Chart representing classes.

Figure 3 visually represents the dataset's classification into class 0 and class 1 through a pie chart, effectively highlighting the distribution proportions of each category.

1. Cross-Va lidation: We evaluated the performance of the model using cross-validation. Cross-validation prevents overfitting by offering a more precise assessment of the model's capacity for generalization and aids in projecting the model's performance on unknown data.

2. Regularization: Overfitting can be managed by adjusting the regularization settings in both Random Forest and XGBoost. The ma x_depth, `min_samples_split{, and min_samples_leaf parameters in Random Forest can be used to regulate the tree's complexity

3. Early Stopping: We utilized early stopping to prevent overfitting in XGBoost. Early stopping stops the training process when the performance on a validation set stops improving, preventing the model from overfitting to the training data.

4. Blend Predictions: Since we are blending predictions from different models, consider using simple averaging or weighted averaging rather than assigning too much weight to one particular model. This can help in reducing the risk of overfitting to a single model's idiosyncrasies.

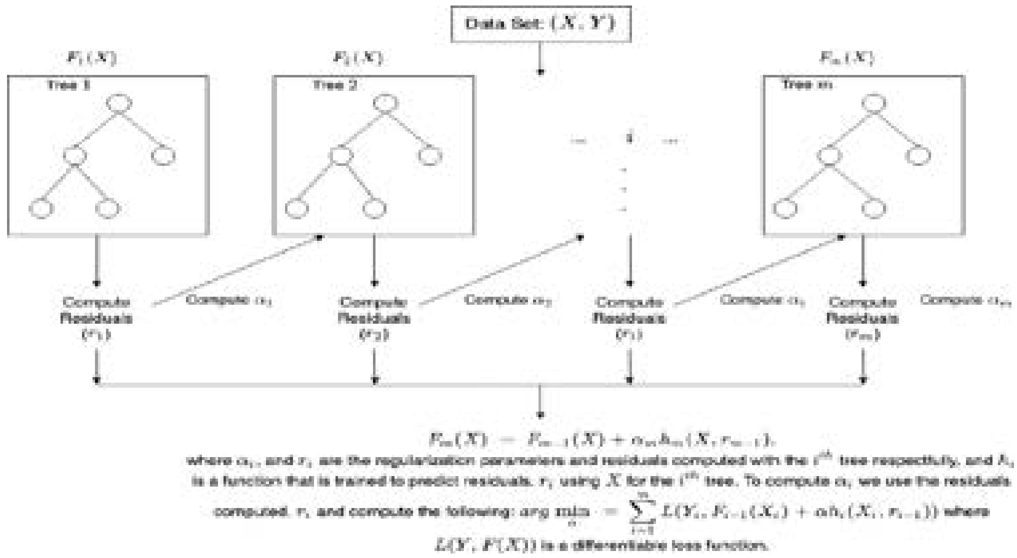


Fig. 4 Working of XGBoost

The Fig.4 explains the detailed working of the boosting algorithm. So, high accuracy can be achieved when we simultaneously use both random forest and XGBoost algorithm.

Combining Random Forest and XGBoost algorithms can significantly enhance predictive accuracy by leveraging the strengths of both techniques. Random Forest is renowned for its robustness to overfitting, ensemble learning capabilities, and ability to handle high-dimensional data, making it a solid choice for diverse datasets. On the other hand, XGBoost, an implementation of gradient boosting, e xcels in optimizing predictive performance through gradient descent and regularization techniques, effectively capturing complex patterns in the data.

By blending these two powerful algorithms, practitioners can exp loit the diversity of their approaches. The Random Forest component provides diversity in the ensemble through bootstrapping and feature randomness, while XGBoost enhances predictive accuracy by iteratively improving model performance on the weaknesses identified by the ensemble.

The combined model benefits from the stability of Random Forest and the predictive power of XGBoost, resulting in superior performance across various domains. Moreover, this hybrid approach helps mitigate individual algorithm shortcomings, such as the tendency of Random Forest to bias towards features with more categories or the potential for overfitting in XGBoost with insufficient regularization.

Ultimately, the fusion of Random Forest and XGBoost yields a formidable predictive model capable of achieving high accuracy and generalization performance across a wide range of tasks and datasets.

IV. RESULTS AND DISCUSSION

In Figure 6, A indicates if there is no age-related condition and B, D, G represent the three age-related conditions that can be provided in the dataset. Plotting this gives overlapping of different age-related conditions for different individuals.

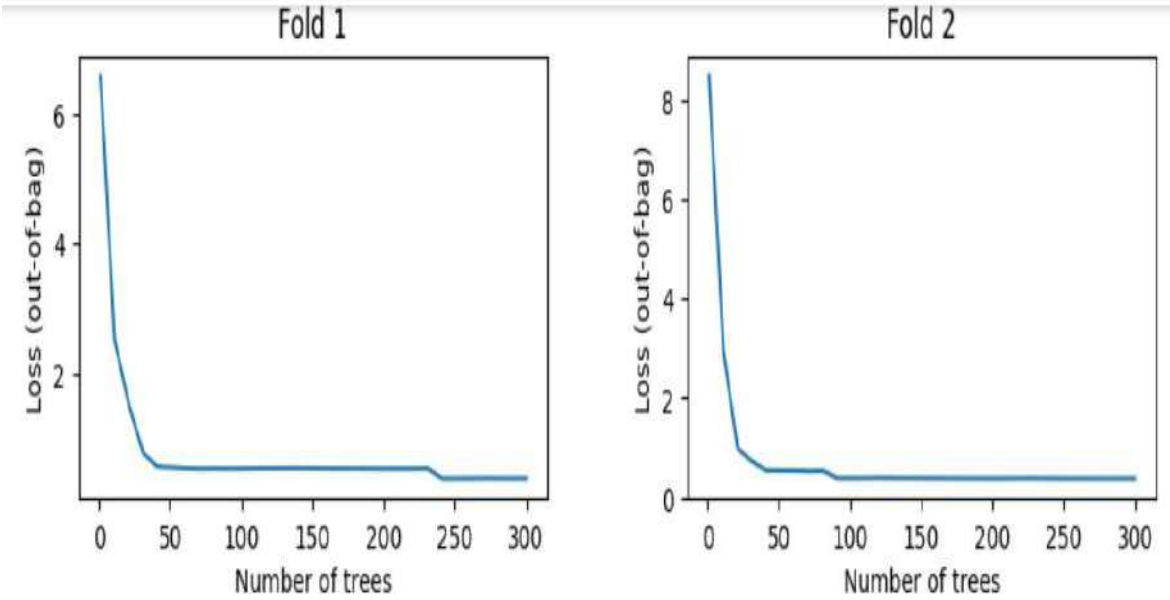


Fig. 5 Folds representing Loss

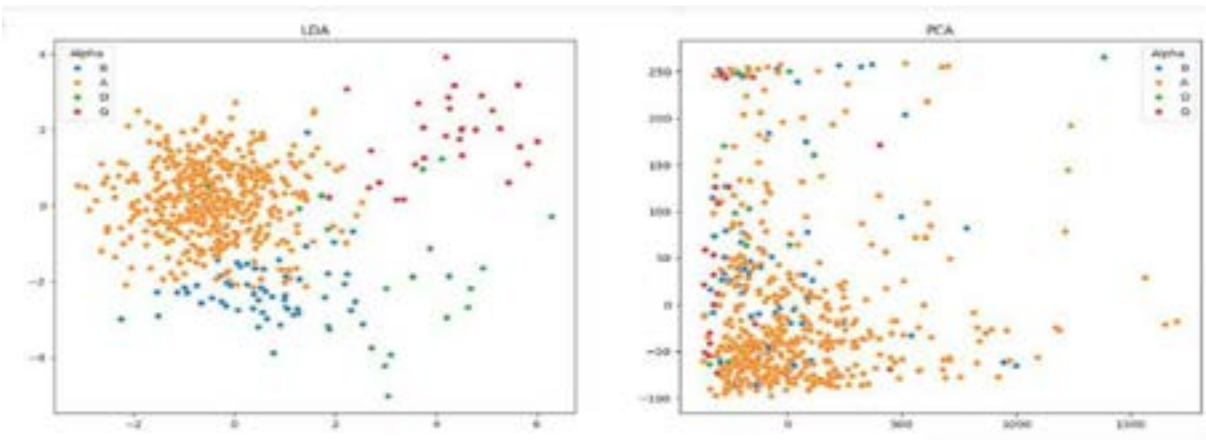


Fig 6 Relationship between A and B, D, and G age-related conditions

The Fig.6 is obtained by a function discriminant that performs dimensionality reduction using Linear Discriminant Analysis (LDA) and Principal Component Analysis (PCA), then visualizes the reduced data using scatter plots.

The data parameter is assumed to be a data frame containing the input data features. The 'Id' column is dropped, assuming it's an identifier rather than a feature. The labels parameter is assumed to be a Series containing the target labels corresponding to each data point. LDA (linear discriminant analysis) and PCA (PCA) are used to reduce the dimensionality of the data to 2 dimensions. LDA aims to find the linear combinations of features that best separate different classes in the data. PCA aims to find the orthogonal linear combinations of features that capture the maximum variance in the data. In each scatter plot, the x-a xis represents the first principal component or discriminant function, and the y-axis represents the second principal component or discriminant function. Points are colored according to the labels provided in the labels parameter, assuming these labels correspond to different classes or categories in the data.

X-axis (Number of Trees)
Y-axis (Loss - Out-of-Bag)

Fig. 7 Relationship between decision trees and loss(o ut-of-bag)

Table 1. represents the table consisting of different machine learning models with their log-loss and accuracy scores to differentiate the working of models. The best model is XGBClassifier() with a learning rate of 0.1 (minimizing cost function) as an added parameter that will reduce the robustness of random forest and improve its accuracy.

The Fig 7 indicates the relationship between the number of trees in the ensemble model and the loss (out-of-bag). This curve helps assess model convergence, detect overfitting or underfitting, and identify the optimal number of trees for the dataset and model configuration. The given data is divided into multiple splits ensuring that the random forest algorithm generated decision trees to learn and train them with the help of ensemble techniques like bagging[short for Bootstrap Aggregating] and boosting on the X-axis. The duration of the training of the models depends on the number of iterations that can occur in the model sequentially.

The loss function or out-of-bag loss is represented on the Y-axis. The OOB loss as used in the Random forest model describes the datasets that were considered as the error or out-of-bound data, that are e xc luded throughout this iterative process of decision trees. This model is used to converge the model data for providing accurate and relevant data that is been used to improve the accuracy.

As a result. the curve in the graph shows us the out-of0bag varies throughout the dataset at each step of the iteration. Then we can find the final convergence of the data that depicts the data over-fitting or under-fitting, thus calculating the ideal number of trees for the particular dataset and model setup.

After filtering the data according to the requirement, we can use the XGBoost algorithm to further improve the model's precision. The XGBoost also stands out for the Ext reme Gradient Boosting, this gradient boosting algorithm typically combines the weak data models that are been filtered out in the random forest, thoroughly re-iterating them[decision trees] to produce a stronger model. The ensemble model we have created above can be gradually boosted with the addition of this gradient model by adding this weaker data to the main dataset, which predicts the exact outcome that it has been put on to have the expected model. The trained data model is then stacked over the same data a meta[model] is trained on them to combine the predictions for better result production.

TABLE 1 PERFORMANCE EVALUATION TABLE

Model	Log- Loss	Accuracy
Logistic Regression	0.684243	0.890756
Support Vector Machine	0.494016	0.915966
Naive Bayes	0.971982	0.87395
Random Forest	0.540528	0.89916
XGB learning_rate : 0.1	0.51748	0.907563
AdaB	0.657281	0.89916

V. CONCLUSION

In the last step, the data again undergoes a performance evaluation, which evaluates a model's performance based on its particular model design and expectation of the flow of the model. It depicts the accuracy of a model based on the probabilistic model by comparing the predicted probabilities to the actual class model. The datasets are then regularized to prevent overfitting cases. Thus performance evaluation can be used to improve the model accuracy. So we can conclude the developed model can be used in various medical fields mainly in the bio-medical fields to be adapted to different sets of data to provide accurate results, The model is suited to working on the both elderly as well as young individuals to prevent the diseases that can be predicted from their health stats.

REFERENCES

[1] Narendra Mohan, Vinod Jain, Gauranshi Agrawal “Heart Disease Prediction Using Supervised Machine Learning Algorithms” 5th International Conference on Information Systems and Computer Networks (ISCON),2021.

[2] Atharv Nikam; Sanket Bhandari; Aditya Mhaske; Shamla Mantri “ Cardiovascular Disease Prediction Using Machine Learning Models” IEEE Pune Section International Conference (PuneCon),2020.

[3] Garima Shukla; Gaurav Dhuriya; Sofia K Pillai; Aradhna Saini” Chronic Kidney Disease Prediction Using Machine Learning Algorithms and the Important Attributes for the Detection” IEEE IAS Global Conference on Emerging Technologies (GlobConET),2023.

[4] Aviral Chanchal; Ajay Shanker Singh; K Anandhan “ A Modern Comparison of ML Algorithms for Cardiovascular Disease

Prediction” 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO),2019.

[5] Srikanth Narayanan;N. M. Balamurugan;Maithili K;P. Bini Palas “Leveraging Machine Learning Methods for Multiple Disease Prediction” International Conference on Applied Artificial Intelligence and Computing (ICAAIC),2022.

[6] Yasmeen Shaikh; V. K. Parvati; S. R. Biradar” Heart Disease Prediction using Ensemble Learning” International Conference on Applied Intelligence and Sustainable Computing (ICAISC),2023

[7] Rajeshwari; H K Yogish “Prediction of Chronic Kidney Disease Using Machine Learning Technique” Fourth International Conference on Cognitive Computing and Information Processing (CCIP),2022.

[8] Sateesh Ambesange; Vijayalaxmi A; Rashmi Uppin; Shruthi Patil; Vilaskumar Patil” Optimizing Liver disease prediction with Random Forest by various Data balancing Techniques “IEEE International Conference on Cloud Computing in Emerging Markets (CCEM),2020.

[9] Sahil Chaudhari; Prathamesh Deo; Purvaja Deshmukh; Aman Deshpande; Priya Shelke; Abhijit Chitre” Multiple Disease Prediction Using Machine Learning Algorithm” 7th International Conference On Computing, Communication, Control And Automation (ICCUBEA),2023.

[10] T. Umbare; Omkar Ashtekar; Aishwarya Nikhal; Bhagyashri Pagar; Omkar Zare “Prediction and Detection of Liver Diseases using Machine Learning IEEE 3rd International Conference on Technology, Engineering, Management for Societal impact using Marketing, Entrepreneurship and Talent (TEMSMET),2023.

[11] Deepa R; Sakthivadivel M; Saravanakumar S; Ganesh Karthikeyan V; Parimala Devi M; Madumitha S “ Feature Extraction and Classification Using Ensemble Method to Enhance Parkinson's Disease Prediction” Annual International Conference on Emerging Research Areas: International Conference on Intelligent Systems (AICERA/ICIS),2023.

[12] Asha Patil; Kalpesh Lad “ Review of Dieases Detection and Classification for Chilli Leaf using Various Algorithms”IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT), 2019

[13] Srujan, N., Mohd, A., Yuva. k., Nanda, K.M., Advanced Graduate Graduate Admission Prediction. IEEE 8th International Conference for Convergence in Technology (I2CT), Mumbai, 2023.

Enhancing Academic Decision-Making using Semantic Analysis based Course Recommendation System

Kollipara Monish Sai Gopi
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai
India
monish.saigopi2020@vitstudent.ac.in

Dr. Sujithra Kanmani R
Assistant Professor
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai
India
sujithrakanmani.r@vit.ac.in

Dr. E. Elakiya
Assistant Professor
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai
India
elakiya.e@vit.ac.in

Dr. B.Surendiran
Associate Professor
Department of Computer Science and Engineering
National Institute of Technology Puducherry
India
surendiran@gmail.com

***Abstract*- The existing course recommendation systems around the world rely on student's past performances, grades and CGPA when giving personalized suggestions. The goal is to provide a customized course recommendation system using Machine Learning (ML) techniques for students. In this project, Machine Learning techniques are used to analyze student feedback about teachers teaching skills, teacher friendliness and course difficulty from the students who already finished said course, this feedback data will be analyzed using semantic analysis. The software developed will be able to give alternative suggestions to students based on their academic performance. By combining the past students' feedback and making real-time suggestions the program will be able to give a clear perspective about the course and the teachers. The final purpose is to give the students a better academic experience by giving them more choices and better course recommendations that will guide them to follow their college career.**

***Keywords*- recommendation, content-based filtering, collaborative filtering, feedback, semantic analysis, K Nearest Neighbours, Support Vector Classifier, personalized suggestions, ratings, Long Short-Term Memory, transformers, Bidirectional Encoder Representations from Transformers, Context Aware Recommenders**

I. INTRODUCTION

In almost every institute, after an academic semester or after an examination, it is common to collect feedback in the form of remarks, about the faculty from students to highlight the faculty's strengths and weaknesses. This data can be used by faculties to get themselves better at their profession. In most of the modern universities or institutes where faculties and courses can be chosen by the students, it will be helpful for the students to know about the course and the faculty better via these feedback-based recommender systems. As present systems that rely on prior performance and GPAs may not provide suitable recommendations that take into consideration a student's preferences and academic achievement, the necessity for a personalized course suggestion system in higher education is becoming more and more obvious. By utilizing machine learning to assess data on instructor rating, friendliness, teaching skills, and course difficulty, collected from semantic analysis of prior students' feedback, the course suggestion system suggested in this study seeks to close this gap. The algorithm can suggest alternatives that are more suitable for struggling students by creating personalized suggestions for courses that match a student's tastes and CGPA and considering the recommended course's number of credits. This makes the process more accurate and equitable. The system's capability to provide suggestions in real-time and consider feedback from past pupils will give them a more complete picture of the subject and instructor, assisting them in making more informed decisions and pursuing their interests. The suggested system's final goal is to enhance the student's academic experience by recommending more pertinent and appropriate courses, which will improve their college experience.

In today's world, there are a few examples of course recommendation systems which use data analytics and machine learning. These recommendation systems are very limited and they are used only by a few colleges. The educational institutes which use these personalized recommendation systems which exist in the current world are reliant on students' branch or their past academic performance, these systems do not consider feedback from students which will then be used to provide recommendations to other students. Not only this but collages usually don't have so much data on the course difficulty, and students are rarely asked to provide feedback on teacher's teaching skills and rarer still asked to tell how "friendly" a particular faculty is. Students and educational institutes want and need better recommendation systems that use machine learning techniques which use much more data in all aspects to provide very accurate recommendations in a personalized manner for the students.

II. MOTIVATION

The motivation lies in its potential to transform the course selection process into a more personalized, informed, and student-centric experience. Most of the traditional systems do not take an interest in individual preferences, learning styles, or the nuances associated with the course content and the effectiveness of the instructor. This has been based on past academic performances of the student and his or her grades. The current project will try to fulfill this gap by adding even more expanded criteria into the recommendation process: detailed students' feedback on teaching skills, friendliness, and course difficulty. Such personalized recommendations aim to give suggestions that are in line with their preferences or those pertaining to his academic standing, precisely to help him create an even more engaging and effective learning environment. This, therefore, means students would do well in courses that appeal to their interests and learning style; hence, happy life at college. Offers of courses: the students can choose well since they know the ability of the teacher and the difficulty level of the course. This transparency helps the students in managing their expectations and preparing well for the courses, hence reducing cases of dropping from courses or performing poorly. Giving students an opportunity to choose the course and the instructor with full comments helps in increasing students' satisfaction with their institution. This not only goes to enhance the level of retention but also more positive culture within the campus. This will then lead the students to be in control of their course selection process and hence be proactive in their course journey. This empowerment can lead to increased motivation, engagement, and ultimately, success.

III. LITERATURE SURVEY

This is very true, considering long-range dependencies in the data and parallel processing of sequences. Transformers model them excellently, and it is for this reason they have been outperforming recurrent networks, including LSTMs, in many tasks (Khan et al., 2021) [1]. Their architectures, in need of inventive biases, are much less modular and can therefore be adapted to a variety of different modalities—from images and text to speech—and scale to large networks and datasets. Although transformers have emerged as the state of the art in many tasks due to their unprecedented ability to model complex patterns and dependencies within the data, a pressing need remains for more interpretable and computationally efficient designs, which would be able to fully use their potential [1].

Transformer-based models such as GPT and BERT, which highlight very significant performance improvements over the previous state-of-the-art architectures in natural language processing tasks. Instead, it goes on to look at the post-BERT models, such as RoBERTa, ALBERT, and ERNIE, which manifest a development by wide margins of the previous approaches (Gillioz et al., 2020) [2]. The work by Kora and Mohammed, 2023 [3], is a deep analysis of the most used transformer techniques within the context of their roles in text classification tasks and is also a great source for using different BERT models. The authors review 109 works of research employing the BERT models for text classification, the methods used, and three transformer model variants, explaining common models in each and what NLP tasks they're appropriate for. Amara et al., 2020, presented to implement a collaborative filtering-based recommendation system using a TextCorpus and the Visual Analytics Interpretable Transformer model to obtain transparent and explainable personalized recommendations [6]. It presents a significant gap, however, in not applying the proposed visualization methods to a concrete dataset, thus missing an opportunity to showcase their practical applicability and effectiveness in real-world recommendation scenarios. Rani, N. (2023) [7] addresses the trust issues in Context-Aware Recommender Systems (CARS) within educational environments like MOOCs and digital libraries. It aims to understand user perceptions, explore design interventions, and propose a multidimensional framework to enhance trust in these systems. The paper identifies the opaque nature of CARS, along with automation and context detection inaccuracies, as primary challenges to user trust. Looking ahead, it suggests focusing on design improvements, better context recognition, and specific trust-building measures to develop more reliable and transparent CARS. For the recommender to function, semantically analysed sentences in the form of numbers should be present. So, first, you need to preprocess sentences. One should remove stop words and vectorize the sentences. Vectorization can be established by TF-IDF vectorizer [8] or count-vectorizer.

Zhao et al., 2021 [9] focused on an online course recommendation model that leverages implicit user behavior data, including login, learning, and course selection details, to generate suggestions using an item-based collaborative filtering algorithm. It has been observed that the precision and recall rates of the model improve under various conditions. However, the paper does not specify the exact limitations of the current model despite discussing enhanced performance over the operation data from a two-year period on an education platform. The paper [13] investigates the use of LSTM neural network models for text classification, leveraging the rich semantic information from word vectors. It discusses how LSTM can effectively differentiate contributions of words and sentences for text categorization and demonstrates an improvement in performance and accuracy over traditional methods (Zhang, 2021) [13].

IV. DATASET

The main dataset used has two attributes – Sentence and Rating. 'Sentence' attributes contain the academics related feedback sentences and 'Rating' attribute contains the corresponding rating to the sentences, ranging from 1-5 as show in table 1.

Rating	Sentiment
1	Very Negative
2	Negative
3	Neutral
4	Positive
5	Very Positive

Table 1. Mapping Ratings to Sentiment

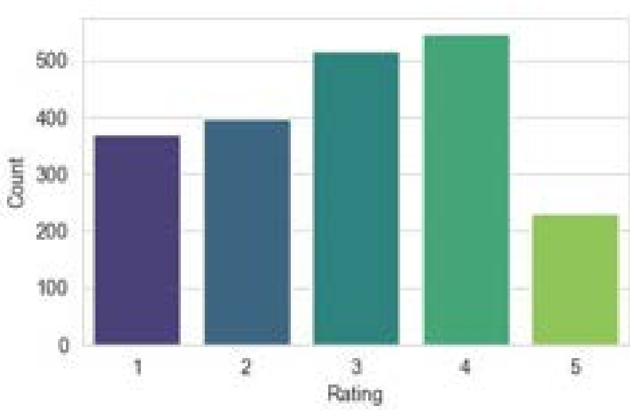


Fig. 1. Bar plot – Distribution of Ratings vs Count of Sentences

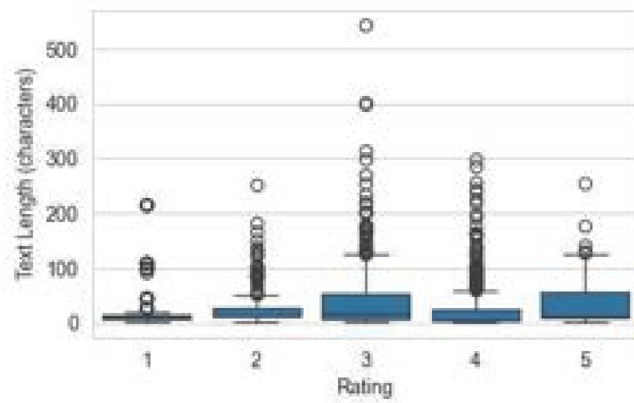


Fig. 2. Box plot – Distribution of Ratings vs Length of Sentences

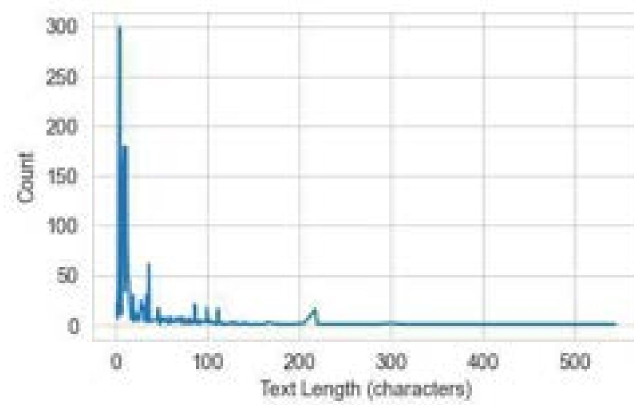


Fig. 3. Line Chart – Length of text vs Count

V. METHODOLOGY

The engine recommends courses in a personalized manner, for this first it needs to preprocess the data. Natural Language Processing (NLP) techniques are incorporated. Semantic analysis is used to preprocess the feedback data of a course or a faculty member, from the students who already completed the course or had that faculty before. Feedback of different attributes from the students such as the teacher friendliness, teacher rating and the teacher teaching skills and finally the course difficulty.

The engine will consider a student's CGPA and it will adjust the courses which are to be recommended according to a student's ability to clear the course, not only this but the system will also be suggest the student to take a course with lesser credits or difficulty to pass.

The methodology is designed for sophisticated, advanced, and personalized course recommendation system that will be implemented into the educational institute for its student's needs and requirements by using advanced machine learning techniques. The system will deploy various advanced algorithms like Logistic regression to predict pass or fail if the student followed the course recommended by the system. semantic analysis using Multinomial Naive Bayes using bag of words, LSTM, BERT. The system uses K Nearest Neighbor algorithm, cosine similarity - content based filtering algorithm and weighted cosine similarity - collaborative filtering algorithm.

5.1. Feedback Analysis:

The input from the students will be in the form of sentences for various attributes like 'Teacher Rating', 'Teacher Friendliness', 'Course Difficulty', 'Teaching Skills'. The sentences should be rated on a scale of 1 to 5. So, sentences are classifying into classes of 1, 2, 3, 4, 5 rating. Feedback analysis using reviews of past student along with the ratings. The feedback rating is extracted using:

- Multinomial Naive Bayes model using bag of words
- Support Vector Classifier
- LSTM model
- BERT (Bidirectional Encoder Represenions from Transformers)

Firstly, the various libraries necessary are imported and the dataset is loaded into a pandas dataframe. Using the imported packages, the text data set is first pre-processed by first removing the stop words, removing the numbers and digits, converting all text to lowercase, then lemmatizing the remaining words. For a term 'i' in document 'j' –

$$\boldsymbol{W_{i,j}} = \boldsymbol{TF_{i,j}} \times \mathbf{log}(\frac{\boldsymbol{N}}{\boldsymbol{df_i}}) \qquad (1)$$

where,

$TF_{i,j}$ = No. of occurrences of term i in document j.
df_i = No. of documents containing term i.

5.1.1. Multinomial Naïve Bayes

The string data is vectorized using an n-gram of dimension [1x1] which is essentially the unigram model. Using the count of each of the words and the sentiment the bag of words model is made and stored. Then the dataset is split into training and testing datasets. The accuracies and performance metrics are then recorded to measure the efficiency of the model. The above model is used for the semantic analysis of the text feedback from the students to generate numerical values and are stored in a dataset to then be used in the recommendation system.

$$\boldsymbol{P(C_i|X_1 \cap X_2 \cap \ldots \cap X_n)} = \frac{\boldsymbol{P(X_1|C_i) \times P(X_2|C_i) \times \ldots \times P(n|C_i) \times P(C_i)}}{\boldsymbol{P(X_1) \times P(X_2) \times \ldots \times P(X_n)}} \qquad (2)$$

5.1.2. Support Vector Classifier

The Support Vector Classifier, or SVC, is an important tool in sentiment analysis, since it maps the features word frequencies or TF-IDF (term frequency-inverse document frequency) scores from the text data—transformed into vectors—very skillfully into a large high-dimensional space. SVC finds the most preferable hyperplane that effectively divides several categorized sentiment data points—for example, positive or negative. One of the most striking features with SVC is its use of kernel functions, which ingeniously project data onto higher dimensions in such a way that they smoothly handle the difficulties offered by non-linear separations. This capability allows the SVC to demarcate the best possible boundary, even amidst the intricacies of complex, high-dimensional datasets. Consequently, this results in a model that boasts a remarkable ability to classify the sentiment of text inputs with precision, all based on their vectorized embodiments.

5.1.3. LSTM – Long Short-Term Memory

The model architecture comprises an Embedding layer for word vectorization of 128 units, followed by two LSTM layers with dropout for regularization, and a Dense output layer for classifying text into one of five sentiment categories. The model is trained and evaluated on a split dataset, providing a measure of its accuracy in the ratings

classification. The activation function used in this model is 'softmax', and the 'adam' optimizer is used. The model is trained for 20 epochs. The model shown in figure 4 is used for performing the comparative semantic analysis of the text feedback from the students to generate numerical values and are stored in a dataset to then be used in the recommendation system.

5.1.4. BERT – Bidirectional Encoder Representations from Transformers

BERT (Bidirectional Encoder Representations from Transformers) model for a sentiment classification task with TensorFlow and Hugging Face's Transformers library. It then loads a pre-trained BERT tokenizer and model specifically designed for sequence classification, adjusting the model to the task's number of labels. Input sentences are tokenized and converted into TensorFlow datasets for both training and testing phases. The model is compiled and trained on the processed data, and its performance is evaluated by measuring accuracy on the test set. BERT's architecture enables it to understand the context of words in a sentence by considering the words that come before and after, making it highly effective for tasks like sentiment analysis.

5.2. Recommender Systems:

In the course suggestion system, classification algorithms are used to categorize the courses and make personalized recommendations based on a student's preferences and academic performance. In the course suggestion system, the classification algorithm is trained to categorize the courses based on their features such as teacher rating, friendliness, teaching skills, and course difficulty.

The techniques used in the course suggestion system are:

- K-nearest neighbors (KNN) algorithm
- Cosine Similarity -Content-based filtering
- Weighted Cosine Similarity - Collaborative filtering

Then after loading the semantically analyzed student feedback dataset into a pandas dataframe, the feature selection and target variables are assigned to X and Y variables. Here, the feature selection contains attributes like 'Teacher Rating', 'Teacher Friendliness', 'Course Difficulty', 'Teaching Skills', 'Credits'. Target variable is the 'Course ID'. Then the user input is taken as a student enters their preferred teachers and course ratings.

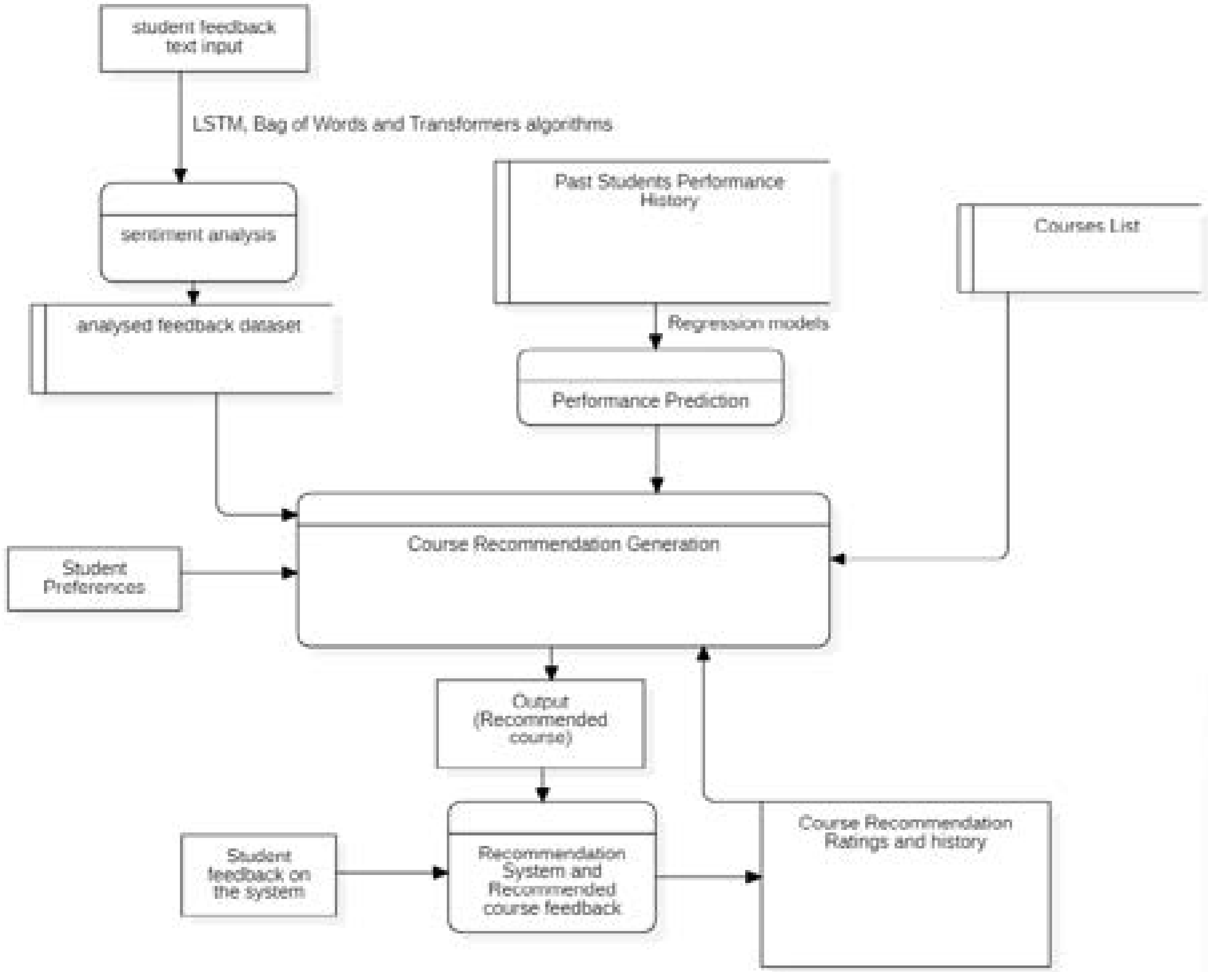


Fig. 4. Proposed Architecture Diagram

5.2.1. K-Nearest Neighbours

A K-Nearest Neighbours model is trained on the above-mentioned training datasets. Using the above mentioned trained KNN model, the course and faculty which are closest match to the student preferences is chosen as the course recommendation. Calculate the distance (either Manhattan or Euclidian) from the data point.

$$\mathbf{d(Y,\ X)} = \sqrt{(\textstyle\sum_{1}^{N}(\boldsymbol{Y_i} - \boldsymbol{X_i})} \qquad (3)$$

The point(X) or data which is closest to the input data(Y), i.e., which has least d(X, Y) is recommended.

After the recommendation, the student is asked to rate the helpfulness of the course recommendation system and then this rating, student's preferences and recommended course information is stored for collaborative filtering-based model later.

5.2.2. Content based Filtering

The system takes the student preferences with each of the semantically analyzed course cum faculty feedbacks and calculates the cosine similarities between each of them by treating them as vectors.

$$\mathbf{similar_{content}(A, B) = \cos(\theta)} = \frac{\boldsymbol{A \times B}}{||\boldsymbol{A}||\times||\boldsymbol{B}||}$$

$$= \frac{\sum_{i=1}^{n}(A_i \times B_i)}{\sqrt{\sum_{i=1}^{n} A_i{}^2}\sqrt{\sum_{i=1}^{n} B_i{}^2}} \qquad (4)$$

The cosine similarities are all stored in a new column and then the dataset is sorted according to the most similar courses to the student's forementioned preferences and the most similar course is suggested.

5.2.3. Collaborative Filtering

The system takes the student preferences with each of the semantically analyzed course and faculty feedbacks calculates the weighted cosine similarities between each of them by treating them as vectors as in cosine similarity and multiplying with weights that are calculated from the previous recommendation feedback ratings which are considerations taken for the collaborative filtering model.

$$\mathbf{similarity_{collaborative}(A, B) = W_k \ x \ \cos_k(\theta)}$$

$$= \boldsymbol{W_k} \frac{\boldsymbol{A \times B}}{||\boldsymbol{A}||\times||\boldsymbol{B}||}$$

$$= \frac{\sum_{i=1}^{n}(W_k \times A_{i,k} \times B_{i,k})}{\sqrt{\sum_{i=1}^{n} A_{i,k}^{2}}\sqrt{\sum_{i=1}^{n} B_{i,k}^{2}}} \quad (5)$$

Alike cosine similarity, the weighed cosine similarity values are stored and most similar courses are suggested.

5.3. Pass or Fail Prediction using Logistic Regression:

After the course recommendation is over, the student's CGPA is taken into consideration and then by using a regression model trained on past student's performance history, a prediction is made to help the student understand the difficulty of the course and if the student might or might clear the course.

Here, the system uses Logistic Regression to predict if the student will be able to pass the course, based on data of students, their CGPA and if they have passed the course in previous semesters. The independent variables are 'Credits', 'CGPA' and 'Course Difficulty'. The target variable is 'Pass/Fail'.

VI. RESULTS AND DISCUSSIONS

An enhanced decision-making algorithm is involved which helps students pursuing their education in any institutions, in which the freedom of choosing their faculties and courses is granted to the students.

6.1. Feedback Analysis

Here, the feedback is received by the faculty to know how the students feel regarding various aspects of the course and teaching. This feedback is semantically analyzed with various machine learning and deep learning models and a comparative study is conducted. Based on a custom dataset, the deep-learning models BERT and LSTM are compared with Multinomial Naive Bayes, Random Forest, Support Vector Classifier. The sentences should to classified into numerical ratings ranging from 1-5. Based on the dataset containing about 2061 entries, BERT has got the highest accuracy of 84.99%. LSTM also stands out compared to other machine learning models, with an accuracy of 82.57%. The confusion matrix of LSTM that is trained for 20 epochs, in Fig. 5, shows that 75 of 78 values of rating '1', 70 of 83 values of rating'2', 86 of 104 values of rating '3', 78 of 103 values of rating '4' and 32 of 45 values of rating '5' are predicted correct.

Rating	Precision	Recall	F1 Score
1	0.96	1.00	0.98
2	0.84	0.96	0.90
3	0.83	0.77	0.80
4	0.76	0.79	0.77
5	0.71	0.59	0.65

Table 2. Evaluation metrics for LSTM

Coming to Multinomial Naïve Bayes, it proves to be accurate for smaller lengths of sentences, but from the analysis of dataset in Fig. 3, there are considerable number of sentences whose lengths are greater than 20. As the system uses unigram multinomial naïve bayes model, there accuracy is 75.78%, which comparatively is lower than other models.

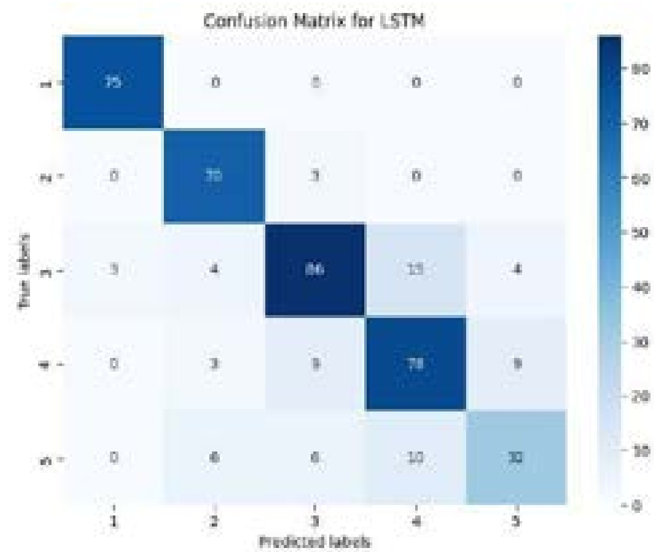


Fig. 5. Confusion Matrix for LSTM model

Rating	Precision	Recall	F1 Score
1	0.83	1.00	0.91
2	0.92	0.64	0.76
3	0.87	0.74	0.80
4	0.57	0.90	0.70
5	0.86	0.35	0.50

Table 3. Evaluation metrics for Multinomial Naïve Bayes

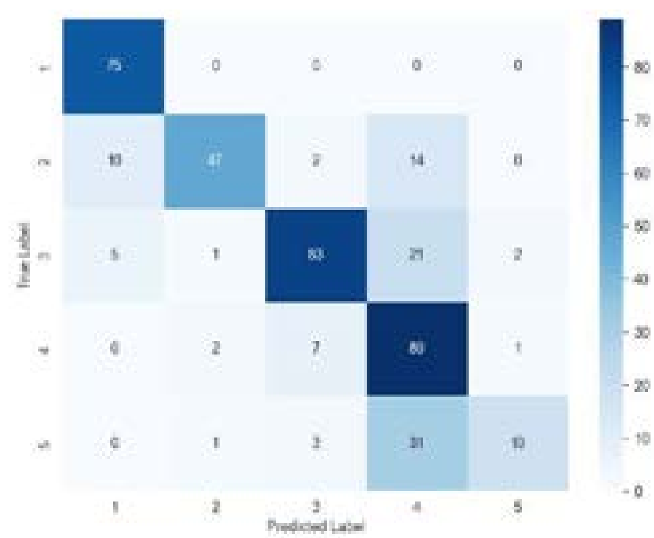


Fig. 6. Confusion Matrix for Naïve Bayes model

Random Forest Classifier, which is a type of ensemble learning has got a decent accuracy of 77.72%, with a precision of 79%, and a decent recall of 75%. Logistic Regression Classifier on the other hand, has got an accuracy of 79.17, which is greater than the Support Vector Classifier (SVC).

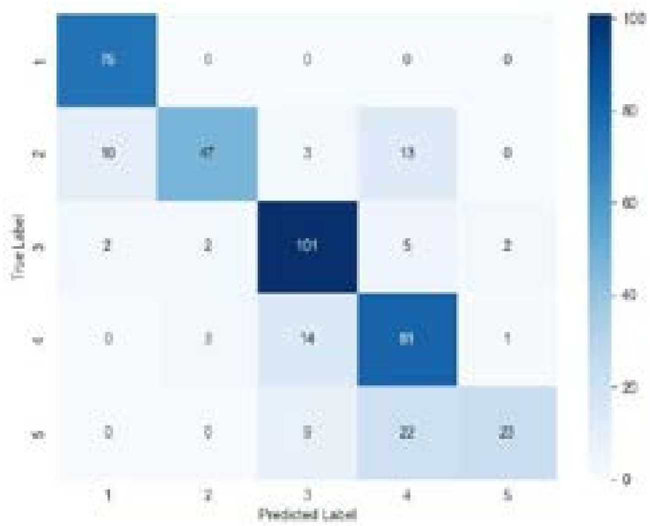


Fig. 7. Confusion Matrix for SVC model

Model	Acc	Avg Precision	Avg Recall	Avg F1
BERT	84.9%	85%	84%	82%
LSTM	82.5%	82%	83%	82%
Logistic Regression	79.1%	81%	79%	78%
SVC	77.7%	81%	78%	76%
Random Forest	77.7%	79%	78%	78%
Multinomial NB	75.7%	80%	76%	75%

Table 4. Comparison of Evaluation Metrics of various models

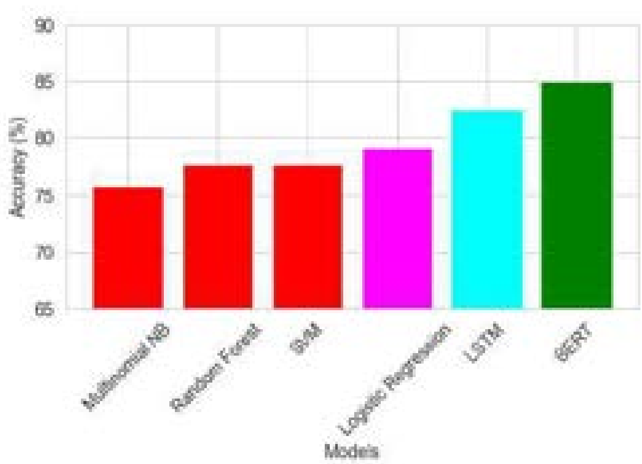


Fig. 8. Models' Accuracy Comparison

6.2. Recommender Systems

Once the semantically analyzed values are available, KNN, cosine similarity and weighted cosine similarity are comparatively analyzed to recommend course and faculty based on the inputs from the students. The sample input and output of the recommender module is shown in the Table 5 and 6.

Your CGPA	9.03
Preferred course's number of credits (1-5)	3
Preferred teacher's rating (1-5)	2
Preferred teacher's friendliness rating (1-5)	5
Preferred course's difficulty rating (1-5)	5
Preferred teacher's teaching skills rating (1-5)	4

Table 5. Sample Input for Recommender System

S. No.	9
Course Code	CSE2001
Course Title	Computer Architecture & Organization
Course Type	PC
Lecture	3
Tutorial	0
Practical	0
J Comp	0
Credits	3
Faculty	Dr. Jotindar

Table. 6. Sample Output for Recommender System

6.3. Pass or Fail Prediction Using Logistic Regression

After courses are recommended from the semantically analyzed database, it is possible that the recommended course is in line with the inputs provided by the student, but might be difficult for the student to complete. This is possible when the student deliberately chooses difficult course or course with more credits. There is possible chance of the student to fail in the subject. This is an additional module for students to make sure that they are in safe zone. In this module, logistic regression is used to train on attributes – 'CGPA', 'Course Difficulty', 'Credits' in order to predict 'Pass/Fail', based on data of students from previous semesters. From the dataset, using the logistic regression, an accuracy of 92.4% is obtained, with a whopping precision of 94.8%.

Accuracy	Precision	Recall	F1 Score
92.4%	94.8%	91.4%	93.07%

Table 7. Evaluation Metrics for Pass/Fail Pred

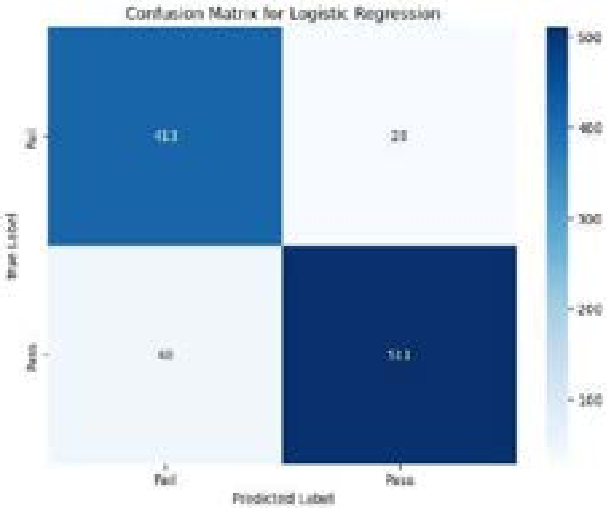


Fig. 9. Confusion Matrix for Pass/Fail Prediction

VII. CONCLUSION AND FUTURE WORK

An accuracy of ~85% is obtained for the semantic analysis of dataset of sentences where there are significant number of sentences which are greater than the length of 50 characters. The pass/fail prediction also obtained an accuracy of ~93%. Overall, the system would be of great use across different institutes, across different nations, where the students shall choose their courses and faculty. This system would help students find right courses, which eases their learning experience. The system becomes more reliable as more and more feedbacks are passed as input over a span of time.

The system can be extended to next stage, where the student can also choose their domain of interest, as in, ‘Data Analytics’, ‘Machine Learning’, ‘IOT’, ‘Humanities’, etc. This requires an additional dataset with list of courses under each domain. Based on the input from students, the system filters out the domain the student is interested in and suggest courses and faculties within the domain. Another idea that can be implemented is ‘Priority Attributes’. The user/student can prioritize the attributes they want.

VIII. REFERENCES

[1] Braşoveanu, Adrian MP, and Răzvan Andonie. "Visualizing transformers for nlp: a brief survey." In 2020 24th International Conference Information Visualisation (IV), pp. 270-279. IEEE, 2020.

[2] Gillioz, Anthony, Jacky Casas, Elena Mugellini, and Omar Abou Khaled. "Overview of the Transformer-based Models for NLP Tasks." In 2020 15th Conference on Computer Science and Information Systems (FedCSIS), pp. 179-183. IEEE, 2020.

[3] Kora, Rania, and Ammar Mohammed. "A Comprehensive Review on Transformers Models For Text Classification." In 2023 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), pp. 1-7. IEEE, 2023.

[4] Fields, John, Kevin Chovanec, and Praveen Madiraju. "A Survey of Text Classification with Transformers: How wide? How large? How long? How accurate? How expensive? How safe?." IEEE Access (2024).

[5] Acheampong, Francisca Adoma, Henry Nunoo-Mensah, and Wenyu Chen. "Transformer models for text-based emotion detection: a review of BERT-based approaches." Artificial Intelligence Review 54, no. 8 (2021): 5789-5829.

[6] Amara, Srikar, and R. Raja Subramanian. "Collaborating personalized recommender system and content-based recommender system using TextCorpus." In 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 105-109. IEEE, 2020.

[7] Rani, Neha. "Investigating Trust in Context-Aware Recommender System in Education." In 2023 IEEE International Conference on Advanced Learning Technologies (ICALT), pp. 370-372. IEEE, 2023.

[8] Anming, Hu. "Research and Application of Online Course Recommendation System Based on TF-IDF Algorithm." In 2022 International Symposium on Advances in Informatics, Electronics and Education (ISAIEE), pp. 333-336. IEEE, 2022.

[9] Zhao, Lianfen, and Zhengjun Pan. "Research on online course recommendation model based on improved collaborative filtering algorithm." In 2021 IEEE 6th international conference on cloud computing and big data analytics (ICCCBDA), pp. 437-440. IEEE, 2021.

[10] Zhang, Lijuan. "Accuracy Recommendation Algorithm of Preschool Education Distance Teaching Course Based on Improved K-Means." In 2022 Global Reliability and Prognostics and Health Management (PHM-Yantai), pp. 1-8. IEEE, 2022.

[11] Kalcheva, Neli, Ginka Marinova, and Maya Todorova. "Comparative Analysis of the Bernoulli and Multinomial Naive Bayes Classifiers for Text Classification in Machine Learning." In 2023 International Conference Automatics and Informatics (ICAI), pp. 28-31. IEEE, 2023.

[12] Setyawan, Muhammad Yusril Helmi, Rolly Maulana Awangga, and Safif Rafi Efendi. "Comparison of multinomial naive bayes algorithm and logistic regression for intent classification in chatbot." In 2018 International Conference on Applied Engineering (ICAE), pp. 1-5. IEEE, 2018.

[13] Zhang, Yanbo. "Research on text classification method based on LSTM neural network model." In 2021 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), pp. 1019-1022. IEEE, 2021.

[14] Ying, Liu, and Li Huidi. "Review of text analysis based on deep learning." In 2020 International Conference on Intelligent Computing and Human-Computer Interaction (ICHCI), pp. 384-388. IEEE, 2020.

[15] Dam, Sumit Kumar, Md Shirajum Munir, Avi Deb Raha, Apurba Adhikary, Seong-Bae Park, and Choong Seon Hong. "RNN-based Text Summarization for Communication Cost Reduction: Toward a Semantic Communication." In 2023 International Conference on Information Networking (ICOIN), pp. 423-426. IEEE, 2023.

[16] Bhumichitr, Kiratijuta, Songsak Channarukul, Nattachai Saejiem, Rachsuda Jiamthapthaksin, and Kwankamol Nongpong. "Recommender Systems for university elective course recommendation." In 2017 14th international joint conference on computer science and software engineering (JCSSE), pp. 1-5. IEEE, 2017.

[17] Kurniawan, Muhammad Mukhtar, and Aina Musdholifah. "Elective Courses Recommendation System using Genetic Algorithm." In Proceedings of the 2021 International Conference on Computer, Control, Informatics and Its Applications, pp. 55-59. 2021.

[18] Premalatha, Mariappan, Vadivel Viswanathan, and Lenka Čepová. "Application of semantic analysis and LSTM-GRU in developing a personalized course recommendation system." Applied Sciences 12, no. 21 (2022): 10792.

[19] Van Houdt, Greg, Carlos Mosquera, and Gonzalo Nápoles. "A review on the long short-term memory model." Artificial Intelligence Review 53, no. 8 (2020): 5929-5955.

[20] Acheampong, Francisca Adoma, Henry Nunoo-Mensah, and Wenyu Chen. "Transformer models for text-based emotion detection: a review of BERT-based approaches." Artificial Intelligence Review 54, no. 8 (2021): 5789-5829.

Enhancing Agricultural Decision-Making: A Supervised Machine Learning Approach

1st Jisnu S
Department of Computer Science and Engineering
PSG Institute of Technology and Applied Research
Coimbatore, Tamil Nadu, India
jisnu2003@gmail.com

2nd Dharaneesh J D
Department of Computer Science and Engineering
PSG Institute of Technology and Applied Research
Coimbatore, Tamil Nadu, India
dharaneesh2207@gmail.com

3rd Krishneth A
Department of Computer Science and Engineering
PSG Institute of Technology and Applied Research
Coimbatore, Tamil Nadu, India
krishneth.arulalan@gmail.com

4th Lokesh S
Department of Computer Science and Engineering
PSG Institute of Technology and Applied Research
Coimbatore, Tamil Nadu, India
lok.for.you@psgitech.ac.in

Abstract—**Agriculture in India is becoming an undesirable livelihood for many considering the uncertainty of making a profit due to either crop failure or an unfavorable market. The primary challenge Indian farmers face today is the lack of information such as weather patterns, ideal crop variety, and market trends. While many current systems try to address this problem, they fail to recognize the diversity in soil and environmental conditions in India, thus resulting in systems applicable in very niche scenarios. With nearly 70% of the Indian population engaged in agriculture, the first issue to be tackled is making an application widespread. Fortunately, smartphones and the internet exist in even the most remote regions of India. This application directly addresses the pressing issue by offering a decision support system that uses machine learning techniques to offer valuable information to farmers anywhere in India. The system has four unique utilities: Crop recommendation using soil quality or location, yield prediction, and price prediction. Various machine learning algorithms such as Decision Tree, Random Forest, KNN, and Naive Bayes were tested for each module to select the algorithm that provided the best accuracy. The machine learning models have been trained exclusively using Indian agricultural data from data.gov.in. The utility offered by the back-end ML modules is presented in a simplistic and intuitive user interface. With help of the information presented by the application, farmers can make informed decisions having a higher probability of success in terms of cultivation and profit.**

Keywords—Agriculture, Crop Recommendation, Crop Yield Prediction, Crop Price Estimation, Machine Learning, Decision Tree Learning, Naive Bayes, Random Forest

I. INTRODUCTION

The effect agriculture has on India and its people need not be understated. On top of being one of the major contributors to the nation's GDP, it is also the primary means of livelihood for many Indian families through earnings. However, the people who depend on this sector for their needs face multifaceted challenges ranging from environmental stressors to fluctuating market demands, significantly affecting crop yield and profitability. Farmers struggle with weather and market uncertainties, limiting their ability to make smart choices when it comes to what crop is to be cultivated or when is the right time to grow a particular crop. Outdated practices miss crucial tech information on soil, weather patterns, and markets, affecting sustainability and profits. Overcoming these challenges requires modern agriculturists in India to not only be adept in growing and harvesting crops but also earning a profit by selling them as commodities. This project introduces a mobile/web application designed to predict crop prices, forecast yields, and recommend optimal crops based on various factors, specifically tailored to the Indian agricultural context. The application assists farmers in achieving their end goal of earning a healthy living by keeping them informed about every aspect of agriculture in India. Taking into consideration the fact that historical data is available in plenty when it comes to Indian agriculture, a machine-learning approach is most appropriate to tackle the objective of using software to analyze a situation and provide helpful data. Machine learning algorithms such as decision trees and random forests are utilized for predicting price, yield, and crop variety.

To tie it all together, a simplistic and intuitive User Interface is built to house the back-end algorithms. The application can be accessed either through a smartphone or a personal computer, and the UI is built separately to cater to the differences in each of the two mediums. Data is presented to the user as numerics and graphs for easy interpretation.

II. RELATED WORKS

There have been a lot of steps taken in terms of research and development in utilizing modern computing technologies in the domain of Indian agriculture. Much inspiration has been derived from existing systems to enhance this application.

Dhawale, S et al., [1] proposed a system where machine learning techniques such as Gaussian and Kernel Naive Bayes are used to suggest crops by taking into account factors such as crop type, soil quality(nitrogen, potassium, phosphorus, and pH) weather patterns (rainfall and humidity), and historical crop harvest data. The parameters regarding the soil quality are The system was implemented through MATLAB's Classification Learner App. The accuracy achieved was impressive surpassing alternative models in crop prediction. This study highlights the transformative potential of machine learning in agriculture by offering farmers information to which they previously didn't have access. The comparative study performed as part of their research emphasized the superiority of their model to the existing literature reinforcing the efficacy of their research.

Moysiadis, V et al., [2] proposed a Cloud Computing-based Smart Farming application using UAVs and ground sensors that can provide real-time data to the users. Scalability and security have not been taken into consideration by a vast majority of the existing system. The team has addressed this by adopting the microservices architecture instead of a monolithic one. The operation of the system takes place within a Docker container and uses RESTful APIs for communication. The web-based GIS interface caters to user roles, including administrators, agronomists, and farmers, featuring various components like soil moisture graphs and NDVI images.

Elbaşı, E et al., [3] have performed an analysis of applying machine learning techniques to increase the yield of a crop by improving or changing agricultural practices. It explores the potential of ML to optimize various aspects of farming, from predicting optimal planting times to identifying disease threats early on. Several IOT devices have been implemented for data collection and the collected parameters include various factors that help us access soil quality. Data preprocessing is performed to ensure accuracy, after which fifteen different machine-learning algorithms have been identified and evaluated. Bayes Net and Naive Bayes Classifiers are selected for their high accuracy in classification tasks relevant to agriculture. The research suggests several ways to improve crop yield to address challenges such as food shortages and resource wastage.

Kuradsenge, M et al.,[4] proposed a research that investigates how climate changes impact agriculture in the Musanze region in Rwanda. The team gathered historical weather and crop yield data using data mining techniques. From the collected data they developed machine-learning models to predict crop yield for Irish potatoes and maize. Multiple machine learning algorithms like Random Forest, Polynomial Regression, and Support Vector Regressor were applied to the dataset to find the best-performing algorithm for this case. Amongst the compared ML models, Random Forest presented the best results with a root mean squared error of 510.8 and 129.9 for potatoes and maize, respectively, and R2 values of 0.875 and 0.817. These results were precise in identifying the optimal weather conditions for each crop.

Chaudhari, A et al., [5] proposed an application for crop recommendation and optimal pricing using ShopBot. Crop recommendations are made using a Decision Tree algorithm, and the input parameters include rainfall patterns, temperature, and location. The machine learning model makes crop predictions with an accuracy of 87%. The ShopBot tool provides price comparisons to make the experience of purchasing crops online a more efficient one. The team compared multiple scraping tools such as Scrapy, Beautiful Soup, and Selenium and chose to proceed with Selenium because of its versatility in handling JavaScript on E-Commerce websites. The study also tackles the difficulties and nuances in converting the price per gram to the price per seed as the latter metric has proven to be more useful to farmers. Overall the system aims to reduce search costs, enhance optimization, and provide farmers with the best commodity options.

Celik, M.F. et al., [6] presented a study on predicting soil moisture based on satellite imaging using a deep learning model that mixes Sentinel-1 backscatter data with Soil Moisture Active Passive (SMAP) data. The Long Short-Term Memory (LSTM) model incorporated weather, radar satellite TV for PC records, soil kind, and topography. Trained on data from the International Soil Moisture Network (ISMN), the LSTM model carried out an R2 score of 0.87 and an RMSE score of zero.046, showing predictive accuracy, mainly in arid and semi-arid regions. Performance variations had been mentioned across distinctive biomass, land cowl, soil textures, and climates, with decreased accuracy in areas with high NDVI values. This research underlines the ability of the use of LSTM for soil moisture prediction in agricultural applications amid global warming challenges.

Linchao, LI et al., [7] presented their research about wheat yield predictions in China. They propose the use of machine learning models to enhance the accuracy of the predictions. Different machine learning algorithms such as Random Forests and Support Vector Machines were tested for their effectiveness in predicting wheat yield when given input parameters like satellite-based vegetation indices, climate data, and soil properties. The research revealed that the inclusion of vegetation indices enhances the precision of the models. Additionally, the research also proved that NIRv had a superior performance in comparison to NDVI and EVI. Random Forest was chosen as the designated algorithm over Support Vector Machines as it showed better results in terms of accuracy. Wind speed, solar radiation, humidity, and soil carbon content influenced wheat yield prediction most.

Innocent, N et al., [8] made extensive use of Small Area Estimation (SAE) methods with a linear regression model to estimate the crop yield. The yield is being predicted for two specific varieties of beans in Rwanda: bush beans and climbing beans. In 2014, the year the experiment was carried over, data obtained from the National Institute of Statistics in Rwanda (NISR) through a seasonal agricultural survey showed a considerable gain in yield estimation accuracy.

Veenadhari, S et al., [9] have researched the use of machine learning to predict crop yield based on changing climatic conditions in India. A web tool named 'Crop Advisor' is utilized to identify the most influential parameters that affect crop yield, the tool relies on the C4.5 algorithm to identify the parameters. After identification, the tool provides insights into the relative influence of each parameter on crop yield. Equipped with the knowledge of the most influential parameters, the team gathered yield and weather data for the past 20 years in the Madhya Pradesh region. From the collected data the C4.5 algorithm built a decision tree that could accurately predict crop yields with an accuracy of anywhere between 76% to 90% depending on the region.

Gandhi, N et al., [10] used machine learning techniques to identify ways to enhance accuracy in predicting crop yield under diverse climatic conditions in India. The study focuses on cereal crops such as rice which form the backbone of Indian agriculture. The machine learning algorithm used in this research is the Support Vector Machines (SVM) classifier implemented through the WEKA tool. 27 districts in the state of Maharashtra were identified to gather data from. The collected dataset consists of parameters such as minimum, maximum, and average temperature, rainfall, area of cultivation, and yield of crop Upon testing other machine learning algorithms along with the SVM for metrics such as mean absolute score, R2 score, and root relative squared error. Comparing the results of the SVM with alternative machine learning techniques revealed that other machine learning algorithms outperformed it. Hence SVMs are not suited to make accurate predictions about crop yields in India.

Several challenges must be addressed when integrating machine learning into agriculture. While machine learning is a very powerful tool, training models that can make accurate predictions or classifications requires a vast amount of structured data. This can be especially troubling in the case of yield predictions because yield can vary drastically with the smallest of changes in the environment. In such cases, it is very difficult to obtain data to build a machine-learning model.

No single application can help farmers from all over the world. The incredible diversity of soil and weather conditions makes this task impossible. A software application can realistically cater to only a small region. Scalability is another area that needs to be addressed, if an application proves to be helpful to farmers, it has to be scaled up with the help of cloud computing, but this can hurt the performance and will also make the application more complex. The most impactful aspect of agriculture apart from soil is the weather, and even with satellite imaging, weather changes can be reliably predicted only a few days in advance, keeping this in mind, a software application must be able to adapt very quickly to new information and also with only a little information. The requirement of a significant amount of information to make predictions will render the application useless, as it is a time-consuming process. In countries such as India where the average farmer's income is in the lower bracket of society, the cost of using such applications must be minimized as much as possible.

III. PROPOSED METHODOLOGY

The Integrated Agriculture Decision Support system aids farmers in making informed decisions utilizing four modules: crop recommendation using soil conditions, crop recommendation using location and season, crop yield prediction, and crop price estimation. Fig. 1 outlines the general architecture of the proposed application incorporating the various features used. The decision support application can be accessed via a smartphone or a personal computer with an internet connection. As soon as the user opens the app they are greeted with the user interface. The UI has been designed to be simplistic and intuitive without compromising utility. The UI has been laid out as a dashboard that presents the user with all the options and information. Information is presented not only as numerics but also in the form of graphs for easier understanding.

A. Crop Recommendation Based on Soil Conditions

The Crop recommendation using the soil quality module is the central component of the decision support system. It considers the soil conditions, weather, and environmental conditions such as temperature and humidity to suggest the most suitable crop. The model is trained on a dataset of 2200 samples from 22 distinct crops to construct this module, all the required data has been obtained from the agricultural datasets present in data.gov.in. The training data is made up of plant macronutrients such as nitrogen (N), phosphorus (P), and potassium (K) along with other factors such as pH levels and environmental variables like temperature, humidity, and rainfall. These attributes are essential because they are closely related to crop output and viability. For example, phosphorus is required to convert solar energy into chemical energy, while nitrogen is necessary for plant growth and leaf development. Similarly, soil pH influences nutrient availability, and potassium is essential for controlling water and nutrient flow in plant cells. To ensure the best accuracy while predicting the crop, three different machine learning algorithms: Decision Tree, Random Forest, and Gaussian Naive Bayes Classification were tested for parameters including accuracy, f1-score, precision, and recall. Out of the 3 algorithms, Gaussian Naive Bayes and Random Forest gave promising results with the Naive Bayes classifier performing marginally better than Random Forest.

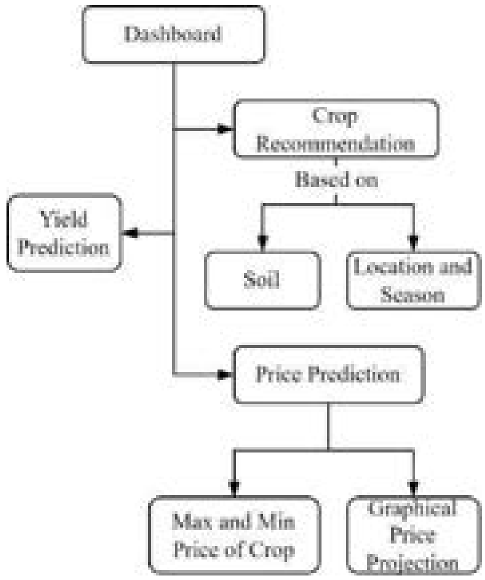


Fig. 1. Architecture of Decision Support System

Personalized crop recommendations in real time are the primary function of this module. When a farmer uses the program, the latitude and longitude coordinates of the user's location are retrieved using the ipinfo.io API based on which the district and state are generalized, allowing for their geopositioning. Afterward, current environmental data, including local temperature and humidity levels, are retrieved from OpenWeatherMap.org using the location attributes provided in the previous step. However environmental considerations by themselves are insufficient to provide tailored crop recommendations. Farmers must also input the pH and nitrogen levels of their soil into the system. These can be ascertained by laboratory analysis or soil testing kits. When the farmer inputs this data into the app, the system uses the selected machine learning model, here, the Gaussian Naive Bayes Classifier to process it and suggest crops that are most appropriate for the soil and climate. The Crop Recommendation module discovers crops that will thrive and offer the maximum returns, in addition to suggesting crops that will grow.

B. *Crop Price Estimation Module*

A vital part of the agricultural decision support system is the Crop Price Estimation Module which gives farmers access to predictive and analytics capabilities so they may anticipate crop prices. This module's core methodology centers on the application of a strong Decision Tree Regressor, which was selected for its capacity to represent non-linear interactions and yield comprehensible outcomes. The Decision Tree Regression algorithm was selected among the three tested, the other two namely, logistic regression and KNN Regression. The Decision Tree Regression presented the best metrics in terms of accuracy.

The initial step in constructing this module was obtaining large datasets for every crop from the Government of India's open data portal, data.gov.in. This dataset contains historical prices for the 23 most-grown crops in India, together with other information including the Wholesale Price Index (WPI), year, month of sale, and recorded rainfall. To ensure the integrity of the predictive model, data preprocessing is carried out to get rid of the dataset of any anomalies or missing values. Feature engineering is used to extract more data on holidays, supply and demand imbalances, and policy changes that may have an impact on agricultural prices.

After preprocessing, different machine learning algorithms such as linear regression, Decision Tree Regressor, and KNN Regressor are tested. Upon comparing their R2 scores the Decision Tree Regressor was selected. The model's performance is optimized by fine-tuning hyperparameters including tree depth, minimum samples per split, and a criterion for gauging a split's quality using grid search and cross-validation techniques. The model chosen for deployment is the one with the best accuracy, as shown by the lowest mean absolute error (MAE) or the greatest R-squared value during cross-validation. This model, which can anticipate future prices based on input features, acts as the module's predictive engine. Users can interact with the system by entering the necessary data, such as the month, year, and rainfall. The estimated price per kilogram or quintal is then computed using the projected WPI and shown to the user. The model is further refined by removing inconsistencies between its forecasts and actual market prices, which are compared with historical data.

Farmers engage with the system by entering the crop name, and then the decision tree regressor estimates the price at different periods in the upcoming year. The model takes various other parameters like year, month, and rainfall which are pre-fed in the module. The prediction given by the decision tree regressor is the Wholesale Price Index (WPI) is a measure of the average change in price for commodities sold in bulk. The WPI, which represents the price at which wholesalers sell their commodities, is determined by using the base price for each crop decided by the Indian government upon analyzing the inflation of crops over the previous decade. After estimation, the WPI is translated into a more precise price metric, a quintal (100 kilograms). The output to the farmer is displayed as a projected price graph based on the WPI and input factors. Crop's estimated price for the next 12 months is shown in a graphical representation that points over the pricing for each of the next 12 months. In addition to the graph, the system showcases the highest and lowest anticipated prices projected for the upcoming 12 months, along with the corresponding months when they reach their peak or decline. This graphical aid can greatly help the farmers with better strategic planning and comprehension of market dynamics. All things considered, this module helps farmers with the financial side of crop production by projecting future pricing, enabling them to make well-informed decisions about when and how much to sell their crops for. Crop Price Estimation is one example of how machine learning is used practically to give farmers insight into market dynamics. The module intends to greatly improve crop sales decision-making by converting intricate data patterns into clear price estimates, which might stabilize and raise farmer incomes throughout India.

C. *Crop Yield Prediction Module*

The Crop Yield Prediction Module is devised to predict the potential yield of various crops, which is pivotal for planning and ensuring a good harvest and profit. This module accepts various input parameters such as crop name, weather, location, area of cultivation, and historical yield data to estimate crop yield upon harvest. An extensive dataset consisting of historical agricultural yield records was retrieved from data.gov.in, an Indian government data portal. The crop type, geographic location (state and district level), crop year, season, area under cultivation, and actual production numbers are among the parameters included in the dataset. To ensure the machine learning algorithms can handle the data, preprocessing involves cleaning the data, addressing missing values, and normalizing the features. Feature engineering is performed after the preprocessing stage to improve the accuracy of the predictions made by the machine learning model. This process includes creating new variables that could impact yields, such as the area under cultivation to production ratio, historical yield patterns, and average yield per hectare. A variety of machine learning techniques, such as Random Forest, Linear Regression, and K-Nearest Neighbours (KNN) are assessed for yield prediction. The selection of these algorithms is based on their capacity to manage multivariate interactions and nonlinear relationships. Of all the algorithms assessed, the Random Forest Regression proved to perform best in terms of performance metrics.

The module uses IP Info API to determine the user's location. This API uses the IP address of the user's internet connection to determine the user's location. The location data is separated into two more distinct geographic parameters, district, and state. The system prompts the user to enter additional parameters apart from the ones already obtained. The first is the growing season which can drastically affect the yield. Second, is the particular kind of crop for which the yield forecast is required. Third, is the area (measured in hectares) allotted to the crop, from which the yield per unit area will be determined. Then, the Random Forest Regressor is fed with all of the data that has been gathered, including the State, District, Season, Crop, and Area. This machine learning model predicts a continuous outcome variable, in this case, the crop yield, by using an ensemble of decision trees. Because it is made up of multiple decision trees that combine to generate predictions that are more accurate than those made by a single decision tree, it is known as a "forest." The prediction is performed when the farmer inputs the required parameters in the app. The predicted value is then displayed to the user.

The predicted crop yield, which is usually expressed in quintals (a unit of mass equivalent to 100 kilograms or around 220.5 pounds), is the output from the Random Forest Regressor. Lastly, if required, the system will convert the anticipated yield to kilograms. This conversion makes it easier for the user to measure, making it more familiar and useful for everyday tasks. This module goes hand in hand with the Price Prediction module, with the information from both modules, the farmer can approximately estimate their annual revenue and make plans accordingly.

D. Crop Recommendation Based on Location and Season

The feature of crop recommendations based on location and season serves as a valuable addition to the decisional support system. This module provides a list of viable crops that can be cultivated in a particular region. This module proves most useful when sufficient data about the soil or the environment is not available to make an accurate decision on what crop to cultivate. When the state, region, and season are given as parameters, the machine learning model provides a range of best-suited crops to be grown. The prediction is made by analyzing previous cultivation patterns in that particular region.

This module takes the input of various parameters from the user namely the state name, district name, and season from the user. By specifying the state where the agricultural land is located, the user offers a better knowledge of the geographical and climatic conditions. The district provides additional information regarding the local soil composition, weather patterns, and other agricultural factors. The location details can also be accessed using the ipinfo API which retrieves the user's geolocation from the Internet Service Provider. Since various crops are sown and harvested at different times of the year and because seasonal climate changes can significantly impact crop viability, the season specification helps in the accurate determination of crops. The season as a parameter falls under 4 different classes namely Kharif, Rabi, Zaid, and Whole Year.

The dataset containing the 4 different classes namely state, district, season, and crop is taken from the government records' open-access website. The dataset was trained to predict the crop based on other parameters using various algorithms and the most optimal one is chosen. The Decision Tree Regressor proved to be the most optimal algorithm to provide the desired prediction. The regressor model provides the prediction for 22 classes of crops that are widely cultivated in India. The decision tree regressor model generated the crop's expected value for a given set of inputs contrasted with a predetermined threshold. The system decides whether or not the anticipated value meets or exceeds the threshold. The crop is added to the list of suggested crops if the value is higher than or equal to the threshold. It is not advised to grow the crop if the value is below the cutoff. Based on the system's predictive model, the final output is a list of suggested crops that are probably suitable for cultivation in the given state, district, and season. This technology seeks to improve crop production chances by taking seasonal and location-specific data into account.

E. User Interface

After the ML models have been trained, an interface is created to enable easy access to these modules and simply use them to a huge extent of people. Here the application user interface plays the role of making the model reach the users. The user interface takes the required inputs from the user and displays the result in the form of texts and graphs for a better understanding of the user.

To give an intuitive experience to users in a way that will be the most helpful for the users to make use of the Integrated Agriculture Decision Support System is to design a user-friendly interface that is easily navigable and self-explanatory, The user interface for Integrated Agricultural Decision Support System is designed in such a way that it will be very simple to be used by a farmer disregarding the fact that they are well educated or not. The different modules that are available in the Integrated Agriculture Decision Support System are displayed on the top so that the users can easily choose the most suitable module for them or the most appropriate module that they wish to use after selecting the required module, that module will populate the screen enabling the user to enter the required information like state, district, or the NPK values of the soil to get optimal crops or crop to get estimated values

HTML, CSS, JavaScript, and AJAX are used in combination to build the user interface. These web technologies have been selected due to their adaptability, and cross-browser compatibility and are the most suitable to be used with a python-based flask back-end. To accommodate users with different degrees of computer literacy, the user interface is created with a clear and simple layout in a way that can be easily navigated and does not contain any complex structure or dark patterns.

Upon opening the application, the users are greeted with a panel that offers easy access to the four modules of the system:

- Crop Recommendation based on Soil Conditions
- Crop Price Estimation
- Crop Yield Prediction
- Crop Recommendation based on Location and Season

Each module is represented by a button placed on top of the page where buttons for each module is placed horizontally one after another, allowing users to click or tap on the desired module's button, it opens the particular form of that module where the required information is entered and upon clicking the submit button below, the form gets submitted and post request is made using AJAX so that results can be asynchronously updated in the UI without reloading the whole thing and a text showing "Please wait while we recommend the optimal crop for you" is displayed to let the user know what is happening in the background. When the result comes back from the flask server the prompt is replaced with the result.

In every module, the interface gathers user inputs such as Location details, soil characteristics, and crop preferences. Input forms are created with clear labels, tooltips, and validation checks to guarantee the accuracy of the data submitted to improve user experience, as the user inputs are highly important for the result to be accurate because the models for each module are trained with specific data in the dataset.

The application's back end, which is run on a Python-based Flask server, drives the user interface from the background and also has listeners for the AJAX requests. The user interface interacts with the relevant data processing module or machine learning model when users submit their inputs to provide results. The user is then presented with results provided from the back-end response via the User interface.

The User Interface Module of the Integrated Agricultural Decision Support System is the farmer's primary point of interaction with the application. The interface prioritizes usability, accessibility, and interactivity, enabling farmers to efficiently utilize the complex functions of the system. In the end, it improves the lives of Indian farmers by acting as a crucial bridge between technology and agriculture.

IV. RESULTS & DISCUSSION

The Integrated Agricultural Decision Support System developed has undergone rigorous evaluation to assess its effectiveness. Evaluating each functional module involved experimentation of being trained using various Machine Learning techniques to arrive at the optimal algorithm that would perform best for the required use case. Validation of the selected machine learning algorithms was performed by splitting the dataset into two where one subsection of the dataset was used to train the model and the other subsection was used to test the trained model. Different ratios of training to testing running from 90:10 to 50:50 were tested. The final ratio that we landed on was 80% of the dataset for training and 20% for testing as it provided the highest accuracy for the classification algorithms and the R2 score for the regression algorithms.

The crop recommendation based on the soil conditions module is experimented with by training the dataset of 22 classes of crops using 3 different classification algorithms namely Decision Tree Classifier, Gaussian Naive Bayes Classifier, and Random Forest Classifier. Out of the 3 algorithms, Naive Bayes and Random Forest gave promising results with the Naive Bayes classifier performing marginally better than Random Forest with an accuracy of 99.09 for the given dataset which is represented in Fig. 2.

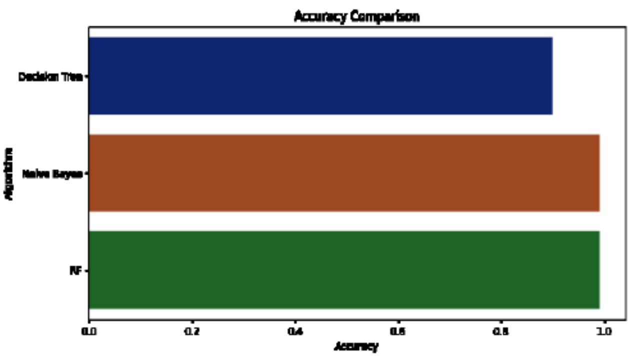


Fig. 2. Comparison of algorithms used for crop recommendation

The crop price estimation module was selected with a Decision Tree Regression algorithm for its capacity to represent the non-linear interactions and predict the outcomes accurately. The Decision Tree Regression was selected as the optimal choice among the three algorithms tested, including Linear Regression and KNN Regression. The Decision Tree Regressor came up with the best performance metrics of an R2 score of 0.9603 which was substantially higher than the performance metrics of the other two algorithms.

Fig. 3 compares the actual versus predicted R2 values for the crop price estimation, the R2 score or the coefficient of determination represents how well a model has been fitted with the dataset, it ranges from 0 to 1 and the closer the R2 value is to 1, the more accurate the prediction is. The R2 value determines the variance caused by a dependent variable. In the above graph, we can notice that the predictions represented by blue crosses are very close to the red line indicating a near-perfect fit, hence we can expect very high accuracy in predictions from this model.

Fig. 3. R2 Score Representation for Crop Price Estimation

The Crop Yield Prediction module was experimented with using various algorithms such as K-Nearest Neighbors (KNN), Linear Regression, and Random Forest. Among all the algorithms tested, the Random Forest Regression algorithm exhibited the best performance in terms of metrics, achieving an R2 score of 0.9688, which was considerably higher than that of the other algorithms used.

Fig. 4 compares the actual versus the predicted R2 values for the crop price estimation module. We can see that most of the data points fall very close to the red line indicating highly accurate predictions. Certain outliers in the dataset vary from ideal values. These outliers increase the variance of the dataset and hence affect the accuracy. Fortunately, outliers in the dataset are few and don't affect the predictions to a great extent.

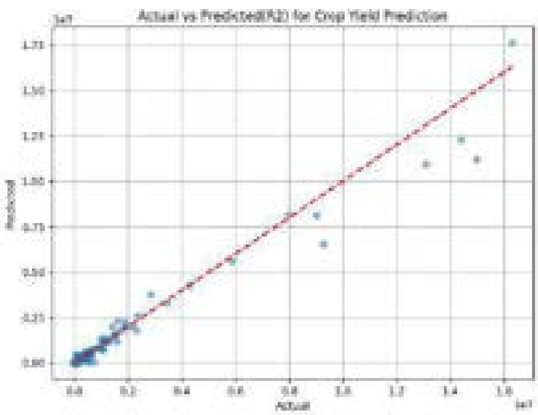


Fig. 4. R3 Score Representation for Crop Yield Prediction

The crop recommendation process, based on location and season, was concluded by selecting the decision tree regression algorithm, which yielded favorable results in terms of performance metrics, boasting an R2 score of 0.8326. The crop recommendations are made by taking into account various dependent variables, this leads to certain inconsistencies when it comes to predictions. However, the module is trained to filter through and recommend a list of crops that are viable for cultivation. The user can eliminate the ones that aren't suitable for their specific environmental conditions.

TABLE I. PERFORMANCE OF SELECTED ALGORITHMS

Module	Algorithm Used	Performance Metrics
Crop Recommendation Based on Soil Conditions	Gaussian Naive Bayes Classifier	Accuracy: 99.0909
Crop Price Estimation	Decision Tree Regressor	R2 Score: 0.9603
Crop Yield Prediction	Random Forest Regressor	R2 Score: 0.9688
Crop Recommendation based on Location and Season	Decision Tree Regressor	R2 Score: 0.8326

Table 1 highlights the best-performing algorithms that were ultimately integrated into the decision support application. The classification techniques were selected based on their accuracy while the regression analysis was performed by accounting for their R2 score.

The user interface of the final application was meticulously crafted to ensure simplicity and user-friendliness. It was particularly tailored to meet the needs of the user. Figures 5,6, and 7 are snapshots of the application dashboard, they illustrate the working of the Location and Season-based Crop Recommendation module (Fig. 6) and Crop Price Estimation module (Fig. 7) .

Fig. 5. Decision Support System Dashboard

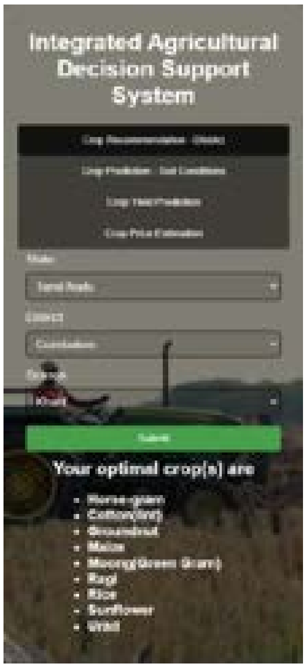


Fig. 6. Demonstration of Crop Recommendation based on Location and Season

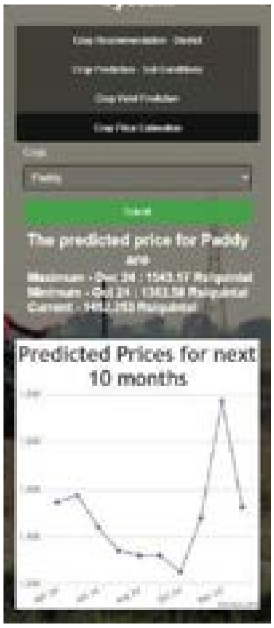


Fig. 7. Demonstration of Crop Price Estimation

CONCLUSION

In conclusion, this Integrated Decision Support System aims to equip Indian farmers with crucial information, thus enabling them to make better decisions ultimately resulting in a profitable farming practice. This application can be continuously improved in terms of prediction accuracy or by the addition of more features to assist farmers. With more data generated every day, the machine learning models can be further improved. With more data, the application can also be expanded to provide predictions for a lot more crops than it currently can.

REFERENCES

[1] S. Dhawale, S. Kakade, R. Kulkarni and C. Muhammed Fasil, "Application of ML in Crop Selection for Precision Agriculture," International Conference on Digital Applications, Transformation & Economy (ICDATE), Miri, Sarawak, Malaysia, 2023, pp. 226-230.

[2] V. Moysiadis, K. Tsakos, P. Sarigiannidis, E. G. M. Petrakis, A. D. Boursianis, S. K. Goudos, "A Cloud Computing web-based application for Smart Farming based on microservices architecture," In Proceedings of the 11th International Conference on Modern Circuits and Systems Technologies (MOCAST), Bremen, Germany, 8–10 June 2022, pp. 1–5.

[3] Elbaşı, Ersin & Zaki, Chamseddine & Topcu, Ahmet & Abdelbaki, Wiem & Zreikat, Aymen & Cina, Elda & Shdefat, Ahmed & Saker, Louai, "Crop Prediction Model Using Machine Learning Algorithms," 2023, Applied Sciences.

[4] M. Kuradusenge, E. Hitimana, D. Hanyurwimfura, P. Rukundo, K. Mtonga, A. Mukasine, C. Uwitonze, J. Ngabonziza, A. Uwamahoro, "Crop Yield Prediction Using Machine Learning Models: Case of Irish Potato and Maize," Agriculture 2023, vol. 13, pp. 225.

[5] M. Chaudhari, M. Beldar, R. Dichwalkar, S. Dholay, "Crop Recommendation and its Optimal Pricing using ShopBot," International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2020, pp. 36-41.

[6] M. F Celik, M.S. Isik, O. Yuzugullu, N. Fajraoui, E. Erten, "Soil Moisture Prediction from Remote Sensing Images Coupled with Climate, Soil Texture and Topography via Deep Learning," Remote Sens. 2022, vol. 14, 5584.

[7] Linchao Li, Bin Wang, Puyu Feng, De Li Liu, Qinsi He, Yajie Zhang, Yakai Wang, Siyi Li, Xiaoliang Lu, Chao Yue, Yi Li, Jianqiang He, Hao Feng, Guijun Yang, Qiang Yu, "Developing machine learning models with multi-source environmental data to predict wheat yield in China," Computers and Electronics in Agriculture, 2022, vol. 194, 106790.

[8] Innocent, Ngaruye & von Rosen, Dietrich & Singull, Martin, "Crop yield estimation at district level for agricultural seasons 2014 in Rwanda," African Journal of Applied Statistics, 2016, vol. 3. pp. 69-90.

[9] Suraparaju, Veenadhari & Misra, Bharat & C. D. Singh, "Machine learning approach for forecasting crop yield based on climatic parameters," 2014, pp. 1-5.

[10] N. Gandhi, L. J. Armstrong, O. Petkar and A. K. Tripathy, "Rice crop yield prediction in India using support vector machines," 13th International Joint Conference on Computer Science and Software Engineering (JCSSE), Khon Kaen, Thailand, 2016, pp. 1-5.

[11] J. Babber, P. Malik, V. Mittal and K. C. Purohit, "Analyzing Supervised Learning Algorithms for Crop Prediction and Soil Quality," 6th International Conference on Computing Methodologies and Communication (ICCMC), Erode, India, 2022, pp. 969-973.

[12] K. Priyadharshini, R. Prabavathi, V. B. Devi, P. Subha, S. M. Saranya and K. Kiruthika, "An Enhanced Approach for Crop Yield Prediction System Using Linear Support Vector Machine Model," International Conference on Communication, Computing and Internet of Things (IC3IoT), Chennai, India, 2022, pp. 1-5.

[13] Vereecken, Harry & Kamai, Tamir & Harter, Thomas & Kasteel, Roy & Hopmans, Jan & Vanderborght, "Explaining soil moisture variability as a function of mean soil moisture: A stochastic unsaturated flow perspective," Geophysical Research Letters, Jan 2007.

[14] Qiangqiang Yuan, Huanfeng Shen, Tongwen Li, Zhiwei Li, Shuwen Li, Yun Jiang, Hongzhang Xu, Weiwei Tan, Qianqian Yang, Jiwen Wang, Jianhao Gao, Liangpei Zhang, "Deep learning in environmental remote sensing: Achievements and challenges," Remote Sensing of Environment, 2020, vol. 241.

[15] A. Savla, N. Israni, P. Dhawan, A. Mandholia, H. Bhadada and S. Bhardwaj, "Survey of classification algorithms for formulating yield prediction accuracy in precision agriculture," International Conference on Innovations in Information, Embedded and Communication Systems (ICIIECS), Coimbatore, India, 2015, pp. 1-7.

[16] Kaur, Manpreet and Dinesh Kumar, "Comparison of Various Classification Techniques for Prediction of the Agriculture Production Based on Different Parameters Rainfall, Temperature," Journal of University of Shanghai for Science and Technology, 2021.

[17] S. Iniyan, "Crop and Fertilizer Recommendation System Applying Machine Learning Classifiers," International Conference on Emerging Research in Computational Science (ICERCS), 2023, pp. 1-6.

[18] Manjula, Aakunuri and Gugulothu Narsimha, "Crop Recommendation and Yield Prediction for Agriculture using Data Mining Techniques," Journal of Emerging Technologies and Innovative Research, 2019.

[19] Kumar, Rakesh, et al, "Crop Selection Method to maximize crop yield rate using machine learning technique," International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy and Materials (ICSTM), 2015, pp. 138-145.

[20] Sujjaviriyasup, Thoranin and Komkrit Pitiruek, "Agricultural product forecasting using machine learning approach," International Journal of Mathematical Analysis 7, 2013, pp. 1869-1875.

[21] Gottemukkala, Lavanya et al, "Sustainable Crop Recommendation System Using Soil NPK Sensor," E3S Web of Conferences, 2023.

[22] Mishra, Subhadra et al. "Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper." Indian journal of science and technology, 2016, pp. 1-14.

[23] Landau, Sabine, et al. "A Parsimonious, Multiple-Regression Model of Wheat Yield Response to Environment," Agricultural and Forest Meteorology, 2000, pp. 151-166.

[24] Rinaldi, Michele & He, Zhenli, "Decision Support Systems to Manage Irrigation in Agriculture," Advances in Agronomy, 2014, vol. 123, pp. 229-280.

[25] OpenAI. (2024). ChatGPT (3.5) [Large language model]. https://chat.openai.com.

Cyberbullying Detection using BERT for Telugu Language

1st Sri Lakshmi Talasila
Computer Science and Engineering
Prasad V. Potluri Siddhartha Institute of Technology
Vijayawada, India
tslakshmi@pvpsiddhartha.ac.in

2nd Dharani Priya Kothuri
Computer Science and Engineering
Prasad V. Potluri Siddhartha Institute of Technology
Vijayawada, India
kothuridharanipriya@gmail.com

3rd Savithri Jahnavi Manchiraju
Computer Science and Engineering
Prasad V. Potluri Siddhartha Institute of Technology
Vijayawada, India
jahnavimanchiraju10@gmail.com

4th Mutyala Sai Sasank Mallavalli
Computer Science and Engineering
Prasad V. Potluri Siddhartha Institute of Technology
Vijayawada, India
sasankmallavalli229@gmail.com

5th Lourdu Gnana Harshith Dande
Computer Science and Engineering
Prasad V. Potluri Siddhartha Institute of Technology
Vijayawada, India
harshithcs28@gmail.com

Abstract- The rapid proliferation of online communication has introduced cyberbullying as a significant concern affecting individuals' well-being. Existing research employs various techniques like Tf-Idf, XLM-RoBERTa, and machine learning algorithms such as Logistic Regression, Random Forest, and Naive Bayes to detect cyberbullying across mixed and bilingual languages. However, these approaches often struggle with accuracy and fail to effectively discern cyberbullying instances due to language nuances and context misinterpretation. Key challenges faced by previous systems include limited linguistic coverage, contextual understanding, and nuanced interpretation of cyberbullying. The new advancement to address these challenges is the implementation of BERT (Bidirectional Encoder Representations from Transformers) architecture by leveraging bidirectional context understanding, allowing it to capture subtle linguistic nuances and contextual cues, thereby improving accuracy and contextual understanding. The proposed model is advancing further by integrating specialized models like IndicBERT, specifically tailored for languages like Telugu. By focusing on contextual nuances, our model aims to improve precision and accuracy of cyberbullying detection for a local language, Telugu content. This study has developed a local language, Telugu dataset comprising 27,000 sentences and achieve an accuracy rate of 90%, highlighting the efficacy of our approach in overcoming these challenges and contributing to online safety.

Keywords- Cyberbullying, Telugu, Bidirectional Encoder Representations from Transformers (BERT), Bullying Preprocessing, Harassment, Language, Social Media

I. INTRODUCTION

The advent of the internet and rapid technological advancements has undeniably transformed the way we live, communicate, and interact with the world. This digital revolution has connected people across the globe, facilitated instant communication, and provided unprecedented access to information. As our lives have become increasingly intertwined with the virtual realm, the positive impacts of technology are evident in education, business, and social connectivity. However, alongside these transformative changes, a dark underbelly has emerged – Cyberbullying. The digital landscape, once heralded as a beacon of connectivity and information sharing, has also become a breeding ground for harassment, intimidation, and abuse. The ease of access to online platforms, social media, and negative behavior. Cyberbullying, the malicious use of electronic communication to target and harm individuals, has become a pressing concern in this brave new digital world. This transformation has not only altered the dynamics of human interaction but has also brought about new challenges in maintaining a safe and inclusive online environment. As we grapple with the implications of our connected world, it becomes imperative to explore the multi-faced nature of cyberbullying, its impact on individuals, and the necessity for proactive measures to detect and prevent these digital aggressions. This exploration will shed light on the evolving landscape of technology and its darker repercussions, urging us to confront the urgent need for effective cyberbully detection methods in order to preserve the positive potential of the digital era. Cyberbullying is the term used to describe bullying that takes place online. Mobile gadgets, gaming platforms, social media, and messaging apps can all be put to use for it. It's a pattern of behavior designed to spook, infuriate, or humiliate the target audience.

II. PURPOSE

The purpose of cyberbullying detection is to identify and prevent instances of cyberbullying in order to protect and support the victims and address the behavior of the perpetrators. It involves the use of technology and various strategies to monitor online activities and communication in order to identify potential cases of cyberbullying. By detecting cyberbullying early on, appropriate interventions and support can be provided to the victims, and the perpetrators can be held accountable and educated on the consequences of their actions. It also helps to create a safer and more positive online environment for individuals to interact and communicate. Overall, the purpose of cyberbullying detection is to promote and maintain a culture of respect and kindness in the digital world. The current implementation encompasses Bengali, Urdu, Tamil, and English. However, it's important to note that support for Telugu is not currently available within the system.

To achieve safety in our cyberbullying detection initiative, we're adopting a multi-pronged approach. First, we'll enhance our machine learning algorithms to improve the accuracy and efficiency of cyberbullying detection across languages, ensuring that potential cases are identified and addressed promptly. Second, we'll implement strict data privacy and security measures to safeguard user information, using encryption and secure storage solutions. Third, we'll integrate user feedback mechanisms to continually refine and improve our platform based on real-world experiences and needs. Lastly, we'll promote a culture of digital citizenship through awareness campaigns and educational content, encouraging users to be responsible and respectful online. By combining these strategies, we aim to create a safer, more supportive online environment for all users.

The initiative to develop cyberbullying detection specifically tailored for Telugu represents a crucial step towards fostering a more inclusive and culturally sensitive online environment. By acknowledging the need for anti-cyberbullying measures in languages beyond the commonly supported ones, this project underscores a commitment to global online safety.

III. LITERATURE SURVEY

Natural Language Processing Journal 3 (2023)- This paper's main contribution is the development of a dataset of 12,795 social media texts in the low-resource Tamil language that have been identified as fine-grained abusive speech. Together with the machine learning models, they have experimented with several feature extraction techniques and discovered that TF-IDF and BoW perform better than alternative feature extractors. Moreover, we discovered that the transformer models were especially helpful for BACD and that they performed better for code-mixed text than any other model. (RQ2)

Merits: This paper presents a comprehensive dataset development for the low-resource Tamil language, effectively using traditional feature extraction methods like TF-IDF and BoW. Additionally, the study effectively utilized transformer models, particularly for code-mixed text. ***Demerits:*** The research is limited in scope to the low-resource Tamil language and lacks exploration of advanced machine learning models beyond traditional feature extraction methods.

Toward Detection of Arabic Cyberbullying on Online Social Networks using Arabic BERT Models (2023)- A actual Arabic dataset that has been manually annotated to enhance the data quality is gathered from YouTube and Twitter and used in this paper. For the purpose of ensuring consistency, each experiment underwent three rounds of trials. Numerous assessment criteria, including AUC and the macro F1 score, were employed to assess the classifiers' performance. 84.58% and 85.94%, respectively, are the F1 score and AUC achieved by the best model.

Merits*:* This paper stands out for its high-quality manually annotated Arabic dataset and its consistent experimentation methodology, undergoing three rounds of trials. The study also employs rigorous evaluation criteria such as AUC and macro F1 score.

Demerits*:* The research is confined to the Arabic language, and there is limited comparison with other languages or machine learning models.

Measurement: Sensors 24 (2022)- Information files are produced using the ASDTD class and Social Media Online Natural Language Processing (SMONLP). The ASDTD F-scores for the internal information files are improved to 0.797 and 0.854, respectively, by integrating the command messages from multiple compatible files.

Merits*:* The study showcases improved ASDTD F-scores by integrating command messages and effectively uses ASDTD class and SMONLP for information file production. ***Demerits****:* The research focuses narrowly on specific datasets and methods, limiting its exploration of broader machine learning models.

Rapid Cyber-bullying detection method using Compact BERT Models (2021) – They tuned a variety of tiny BERT models using hate speech data. To address the class imbalance in the data, a Focal Loss function has been implemented. On the hate-speech dataset, they were able to obtain cutting-edge results with 0.91 precision, 0.92 recall, and 0.91 F1-score by employing this method.

Merits*:* This paper excels in tuning BERT models effectively and addressing class imbalance with the Focal Loss function. The study achieved high precision, recall, and F1-score, showcasing its effectiveness. ***Demerits****:* The research is limited to hate speech data and has a narrow focus on specific aspects of cyberbullying, potentially limiting its applicability to broader cyberbullying contexts.

Cyber-Bullying Detection in Social Media Platform using Machine Learning (2021- They studied cyberbullying, its manifestations, techniques, and outcomes, as well as the most current studies on its identification and prevention. They also investigated different types of cybercrime. A total of 35,000+ tweets from Twitter were gathered for the experiment, and the data was cleaned and organized to feed multiple intelligent machine learning algorithms. Five key ML algorithms were then applied to the tweets to classify and predict them into two primary groups: "offensive" and "non-offensive." Last but not least, an analysis of those machine learning algorithms has been done using a number of performance indicators.

Merits*:* This research benefits from a large dataset collected from Twitter and applies multiple machine learning algorithms, leading to classification into "offensive" and "non-offensive" groups.

***Demerits**:* The study is restricted to Twitter data and lacks in-depth linguistic analysis, as well as exploration of advanced machine learning methods beyond classification.

Detect Chinese Cyberbullying by Analyzing User Behaviors and Language Patterns 2019- A Long Short-Term Memory Neural Network-Deterministic Finite Automaton (LND) model is constructed in this research that takes into account not only the language content but also the user's attributes and previous speech on social networks. They employ the data of Douban's reviews by evaluating speech patterns with polarized emotions in the absence of identified content. Next, Chinese cyberbullies' Weibo activities are examined using the newly taught model. Due to the user's behavior attributes and language emotional polarity ratings, the accuracy of cyberbullying detection rises from 89% (using the sensitive lexicon filtering approach) to 95%.

Merits: The research introduces a comprehensive LND model that considers user attributes and emotional polarity, resulting in a significant increase in detection accuracy from 89% to 95%.

***Demerits**:* The study is limited to the Chinese language, and the complexity in model construction may hinder its generalizability to other languages or contexts

IV. PROPOSED MODEL

Our innovative approach harnesses the power of the BERT (Bidirectional Encoder Representations from Transformers) Transformer, specifically tailored for Indian languages known as IndicBERT. IndicBERT is optimized for languages like Telugu, making it a robust choice for our cyberbullying detectionsystem. It not only understands the complexities of the Telugu language but also captures its nuances, ensuring a more accurate interpretation of text. In addition to leveraging IndicBERT's capabilities, our model integrates a multi-layered neural network architecture, allowing for deeper and more intricate analysis of text. This enables the model to capture subtle linguistic cues and patterns indicative of cyberbullying, thereby increasing the detection accuracy. Furthermore, to address the challenge of limited labeled data in Telugu, our model employs transfer learning techniques. By pre-training on a large corpus of diverse data and fine-tuning on a smaller, labeled dataset specific to cyberbullying in Telugu, we optimize the model's performance without requiring extensive labeled data. By combining IndicBERT's contextual language representations with advanced neural network architecture and transfer learning techniques, we aim to significantly improve the accuracy and precision of cyberbullying detection in Telugu, ensuring a safer online environment for Telugu speakers. The general BERT architecture is depicted in figure 1.

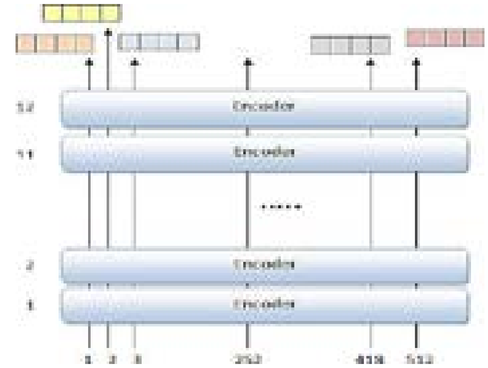


Figure 1: BERT Architecture

V. METHODOLOGY

This study proposes a comprehensive model for Telugu cyberbullying classification, incorporating several key steps to ensure the accuracy of the classification process. The journey begins with meticulous data collection, as outlined earlier, wherein we curated a diverse dataset comprising approximately 18,000 sentences covering a range of cyberbullying instances in Telugu. Following this, a rigorous data preprocessing phase was implemented, addressing issues such as missing values and duplicates to refine the dataset. A crucial aspect of our methodology involves feature extraction, where we harnessed the potential of cutting-edge language models. Notably, we employed the IndicBERT model, tailored to the nuances of Telugu language, resulting in optimal performance for our specific context. The choice of IndicBERT was driven by its ability to capture intricate linguistic patterns and contextual information, contributing significantly to the model's proficiency in understanding Telugu cyber communication. The workflow of the proposed methodology is shown in figure 2.

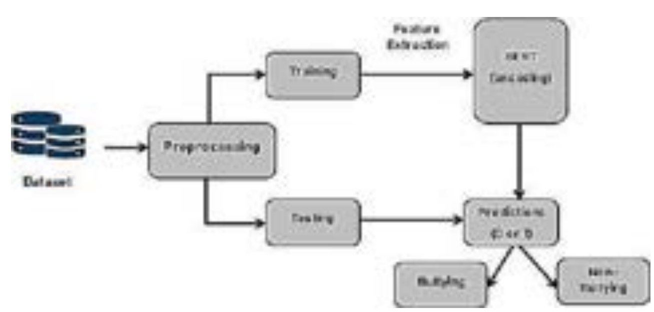


Figure 2: Process Workflow

Furthermore, our proposed model embraces various innovative proposals aimed at enhancing the classification accuracy. These proposals encompass fine-tuning strategies, attention mechanisms, and leveraging domain-specific embeddings to better capture the intricacies of cyberbullying in the Telugu language. By adopting a thoughtful and multifaceted approach in our methodology, we aim to contribute a robust solution to the challenging task of cyberbullying detection in the Telugu linguistic landscape. We provide a brief description of each part's workflow below

i. Data Collection

Figure 3: Word Cloud

Creating a comprehensive dataset for Telugu cyberbullying detection posed a significant challenge due to the absence of relevant papers or research in the language. Firstly, as research in this cyberbullying aspect in the Telugu language is rare, we couldn't find a source for our dataset. Undeterred by the lack of existing resources, our team undertook the task of crafting a diverse dataset independently. We started collecting data from the comments, tweets, and blogs of various social media platforms. With a substantial collection of approximately 27,000 sentences, our dataset encompasses a broad spectrum of cyberbullying instances, ranging from political and threatening to sexual texts. A sample word cloud is shown in figure 3. The dataset is structured with two key columns: the 'Text ' column, containing authentic cyber communication in Telugu, and the 'Label' column, employing a binary classification of 0 for cyberbullying and 1 for non-cyberbullying instances. Throughout the data collection process, careful attention was given to ethical considerations, particularly regarding sensitive content. The dataset not only addresses the immediate need for Telugu cyberbullying research but also holds potential for future investigations in the field. Our commitment to transparency is evident in the documentation of data cleaning, preprocessing steps, and the formulation of training, validation, and testing sets. This lays a robust foundation for the development of accurate machine learning models tailored to Telugu cyberbullying detection. Here are some examples from the dataset-

0 - నువ్వు ఒక యెడవా
0 - సైకో సాలె
1- ఒరేయ్ నీ కడుపు సల్ల గుండ బాగా చేసినావ్ రా నాయనా న్యాయంగా నువ్వు చేసింది
1 - దేవు డివి రా నువ్వు
0 - సీరియల్ లో మీ ఇద్దరివీ గాడిద్ వేశాలే

S.No	Context of Sentence	No. of Text Sentences
1	Bullying	13608
2	Non-Bullying	13812

Table 1: Dataset

ii. Data Preprocessing

In the preprocessing phase of our research, several key measures were implemented to refine the raw textual data, ensuring its quality and suitability for subsequent analysis. Preprocessing has been done for the following:

- **URLs**: We eliminated URLs using regular expressions, effectively removing irrelevant web links to minimize noise in the dataset.
- **User Mentions**: User mentions were generalized by substituting them with the generic tag "@user," fostering anonymity and consistency across the dataset.
- **Numeric Values**: We systematically purged numeric values from the text to reduce potential interference with semantic analysis and streamline the dataset's dimensionality.
- **Emojis**: Emojis underwent conversion into their corresponding text representations, aiding in the integration of non-textual elements for more comprehensive processing.
- **Special Characters**: Characters such as colons and asterisks were replaced with spaces to contribute to the overall normalization of the text.

This step standardized the text and mitigated potential issues during subsequent processing stages. Through these meticulous preprocessing steps, a cleaner and more standardized dataset was achieved, laying a solid foundation for effective natural language processing tasks. Notably, the incorporation of these measures promoted a more efficient analysis process, facilitating the extraction of meaningful insights and patterns from the textual data. Furthermore, the decision not to specifically remove stop words was informed by the utilization of the BERT model, which inherently handles stop words during its tokenization process. Leveraging BERT's advanced language understanding capabilities optimized the preprocessing pipeline for tasks such as sentiment analysis, classification, or language generation.

iii. Feature Extraction

Feature extraction is crucial in NLP as it transforms raw text into a format that machine learning models can understand. We chose BERT for feature extraction over traditional methods like TF-IDF or GloVe because of its ability to capture semantic meaning and contextual information, which often leads to more meaningful representations. Unlike TF-IDF and GloVe, which focus on word frequency or co-occurrence statistics, BERT understands the context and semantics of words, making it more suitable for capturing the nuances of language. In our quest to optimize natural language processing (NLP) within our project, the meticulous collection of Bert-related data has emerged as a pivotal undertaking. This strategic decision emanates from a comprehensive evaluation of various Bert models, including mBERT, RoBERTa, and DistilBERT. Each of these models possesses distinct characteristics and functionalities that we carefully scrutinized before arriving at a decision.

mBERT, or Multilingual Bert, is designed to handle multiple languages. While it demonstrates competence in accommodating a wide linguistic spectrum, its performance can be compromised when dealing with languages with specific nuances, such as Telugu. Telugu, being a Dravidian language with unique linguistic intricacies, demands a more specialized model for optimal results.

RoBERTa, an optimized version of BERT, utilizes dynamic masking during pre-training to enhance performance. Despite its advancements, RoBERTa may not fully capture the nuances of Telugu due to its general-purpose nature. The model's pre- training on a vast dataset might not adequately address the subtleties inherent in Telugu, making it less suitable for our project's linguistic requirements.

DistilBERT, a distilled version of BERT, focuses on retaining essential aspects while reducing computational complexity. Although it excels in efficiency, it might not capture the specific linguistic nuances crucial for our Telugu language tasks. The distilled nature of the model may result in the loss of language intricacies, impacting its suitability for specialized requirements.

Our decision to adopt IndicBERT is underpinned by its demonstrated ability to outperform other variants in our unique context. Through exhaustive testing and analysis, IndicBERT consistently exhibited superior accuracy and effectiveness in handling the language intricacies specific to our Telugu language tasks. Its tailored approach to Indic languages, including Telugu, makes it instrumental in addressing the linguistic nuances inherent in our project.

IndicBERT generates high-dimensional embeddings for each token in the input text based on its context within the sentence. These embeddings capture the semantic meaning and contextual information of each token, providing a rich representation of the input text. To obtain a fixed-size feature vector for the entire input sequence, we employ a pooling strategy, such as mean pooling or max pooling, to aggregate the embeddings of all tokens. This feature vector serves as the extracted features for the input text, which can then be used as input to machine learning models for training and prediction tasks.

IndicBERT‘s application extends across a spectrum of use cases, from sentiment analysis and language understanding to document classification and information retrieval, making it a versatile choice for our diverse NLP requirements. By utilizing IndicBERT for feature extraction, we leverage its ability to capture the specific linguistic nuances of Telugu, leading to more accurate and effective representations for our NLP tasks.

In summary, our deliberate choice of IndicBERT is grounded in its superior performance and tailored linguistic capabilities, aligning seamlessly with the objectives of our project. This strategic decision positions us to capitalize on the strengths of IndicBERT and, in turn, opens up new possibilities for innovationand excellence in natural language processing within the unique linguistic landscape of Telugu.

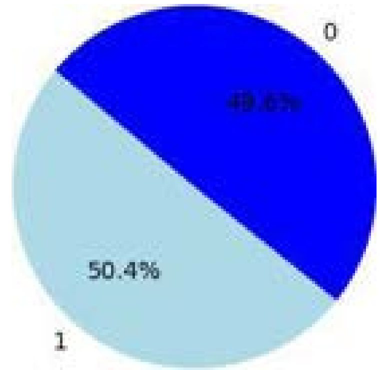


Figure 4: Distribution of Labels in the Dataset

iv. Considerations

1 **Attention Mask Creation:**

   a. Generates attention masks crucial for models like BERT.
   b. Aids in the model's ability to discriminate between padding and real tokens during training.
   c. Ensures that padding does not interfere with the model's attention mechanism.

2 **Feature and Label Preparation:**

   a. Tokenized input sentences, converting them into input IDs.
   b. Pads sequences to a specified maximum length, ensuring consistent input size.
   c. Creates attention masks for BERT-based models.
   d. Transforms raw text data into a format suitable for model training.

3 **Accuracy Calculation:**

   a. Computes accuracy by comparing model predictions to actual labels.
   b. Assesses the total accuracy of the predictions made by the model throughout training.

In conclusion, the methods of attention mask creation, feature and label preparation, accuracy calculation, and time formatting play pivotal roles in enhancing the efficacy and interpretability of models, particularly exemplified in frameworks like BERT. The attention mask generation proves indispensable in enabling models to discern between relevant tokens and padding tokens, preventing interference with the attention mechanism during training. Feature and label preparation not only tokenize input sentences but also ensure consistent input size through sequence padding, allowing raw text input to be transformed into a format that can be used for training models. Accuracy calculation serves as a crucial metric, offering a quantitative assessment of the model's overall correctness by comparing predictions to actual labels during training. Finally, time formatting simplifies the tracking and reporting of the time taken for model training, providing a valuable tool for assessing computational efficiency. Collectively, these methods contribute significantly to the robustness and utility of natural language processing models.

VI. EXPERMENTAL ANALYSIS

a) Evaluation Metrics:

1. **Accuracy:** Measures the overall correctness of predictions

2. **Macro F1 Score:** Unweighted average of F1 scores across different classes.

3. **F1 Score:** Balances precision and recall, particularly useful for imbalanced class distributions.

4. **Area under the ROC Curve:** Assesses how well the model can discriminate between classes that are positive or negative.
   i. The real position rate versus the false positive rate is represented graphically by the ROC curve.
   ii. The area under the ROC curve, or AUC, has a range of 0 to 1, with a greater value denoting superior performance.

5. **Precision:** True positive prediction accuracy is measured as the ratio of total positive instances to true positive forecasts.

6. **Recall:** Positive instance capture is emphasized by the ratio of true positive predictions to all real positive cases.

b) Probability Calculation Formula:

In the context of binary classification, probability is often used to make decisions or to rank predictions based on confidence. It permits a nuanced interpretation of the model's output beyond simple class labels, enabling better-informed decision-making and model evaluation.

$$(y = 1/x) = \frac{e^{logit1}}{e^{logit0} + e^{logit1}}$$

Here, e^{logit0} and e^{logit1} are the exponential of the raw logits for class 0 and class 1 $logit$ 0 e $logit$ 1 respectively.

Observations:

The evaluation metrics employed in our cyberbullying detection model are designed to provide a comprehensive understanding of its performance across various dimensions. The metrics include Accuracy, Macro F1 Score, F1 Score, Area under the ROC Curve, Precision, and Recall, each offering unique insights into the model's capabilities.

Accuracy measures the overall correctness of predictions, indicating the proportion of correct predictions among all predictions made. Our model achieves an accuracy of 90%, which demonstrates its ability to correctly identify instances of cyberbullying and non-cyberbullying.

The Macro F1 Score, which is the unweighted average of F1 scores across different classes, provides a balanced measure of precision and recall across all classes. With a Macro F1 Score of 0.90, our model exhibits robust performance across various categories of cyberbullying.

The F1 Score balances precision and recall and is particularly useful for imbalanced class distributions. Our model's F1-score stands at 0.90, indicating a harmonious balance between precision and recall.

The Area under the ROC Curve (AUC) assesses the model's ability to discriminate between positive and negative classes. A higher AUC value denotes superior performance. Our model's AUC is notably high, reflecting its excellent discriminatory power.

Precision measures the true positive prediction accuracy, indicating the ratio of total positive instances to true positive forecasts. Our model achieves a precision of 0.896, demonstrating its high accuracy in identifying true cyberbullying instances.

Recall emphasizes the capture of positive instances, representing the ratio of true positive predictions to all real positive cases. With a recall of 0.898, our model effectively captures a high proportion of actual cyberbullying instances.

In addition to these evaluation metrics, we've also incorporated a Probability Calculation Formula to provide nuanced interpretations of the model's output. This formula, derived from the raw logits for each class, enables a more detailed understanding of the model's confidence in its predictions.

Overall, our experimental results demonstrate that the proposed model performs exceptionally well in identifying cyberbullying instances across multiple evaluation metrics. The high values obtained across Accuracy, Precision, F1-score, and Recall, coupled with a strong AUC, attest to the model's robustness and efficacy in cyberbullying detection. These results validate the effectiveness of our approach and underscore its potential for enhancing online safety and fostering a more respectful and positive digital environment.

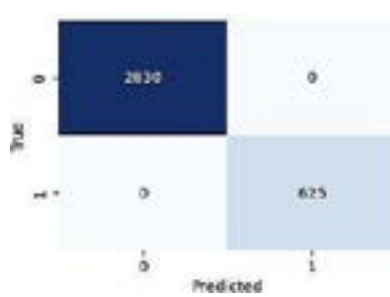


Figure 5: Confusion Matrix

VII. RESULTS AND FINDINGS

S.no	Name of the Project	IssueDate	Methodology	Accuracy
1	Cyberbullying Detection for Low-resource Languages and Dialects: Review of the State of the Art	2023	A thorough investigation and comprehensive study of the automatic identification of cyberbullying in low- resource fine grained languages.	84.9%
2	Towards Leaving No Indic Language Behind: Building Monolingual Corpora, Benchmark and Models for Indic Languages	2023	An analysis of pre-training corpora,multilingual language models usingmBERT, IndicBert -v1, IndicBert -v2,XLMR, and MuRIL in comparison to current benchmarks	88.3%
3	Mono vs Multilingual BERT for Hate Speech Detection and Text Classification: A Case Study in Marathi	2022	Analyze the models utilizing multilingual models (IndicBERT, mBert) for sentiment analysis and hate speech detection on the datasets.	82.8%
4	Cyberbulling Detection using BERT for TeluguLanguage	Proposed model	**IndicBERT** for Telugu text, augmented with standard NLP preprocessing techniques, and machine learning fortraining and evaluation	90%

Table 2: Comparative Study

Our research journey was guided by an extensive review of relevant literature, primarily drawing insights from various influential papers in the field. In particular, the systematic survey outlined in the paper titled "Cyberbullying Detection for Low-resource Languages and Dialects: Review of the State of the Art" (2023) became a foundational reference. This review significantly shaped our approach, leading to an accuracy of 84.9% in automatic cyberbullying

Additionally, we leveraged the knowledge distilled from "Towards Leaving No Indic Language Behind: Building Monolingual Corpora, Benchmark, and Models for Indic Languages" (2023). This paper's comparative analysis of existing benchmarks and pre-training corpora for Indic languages, including mBERT, IndicBert -v1, IndicBert-v2, XLMR, and MuRIL, contributed to our methodology. The result was an impressive accuracy of 88.3%, reflecting the efficacy of our approach to linguistic diversity.

Furthermore, the paper titled "A Transformer Based Approach for Abuse Detection in Code Mixed Indic Languages" (2022) provided valuable methodologies for detecting obscene material in thirteen languages with mixed Indian codes using advanced transformer-based models. This reference guided our model selection, incorporating IndicBERT, XLM- RoBERTa, MurilBert, and mBERT, resulting in an accuracy of 86.78%.'

In the realm of cyberbullying detection, existing techniques have made significant strides, showcasing comprehensive coverage across various linguistic and thematic dimensions. They have leveraged advanced transformer-based models and specialized language models like BERT to achieve promising results. These strengths underscore the potential of modern NLP techniques in tackling cyberbullying effectively. However, a closer look reveals certain limitations within these existing approaches. Many of them, while offering a broad overview, often struggle with the nuances of specific languages or types of cyberbullying. This lack of adaptability can be attributed to their reliance on pre-trained models, which may not fully encapsulate the intricacies of less-researched languages or nuanced forms of cyberbullying. Contrastingly, our proposed model takes a tailored approach, focusing specifically on the Telugu language. By utilizing Indic-BERT, a model fine-tuned for Telugu, we've achieved a commendable accuracy rate of 90%, Precision of 0.896, F1-score of 0.90 and Recall of 0.898 as shown in Table 3. This high accuracy not only demonstrates the model's effectiveness in the Telugu context but also suggests its potential robustness for other low-resource languages facing similar challenges.

Metrics	Values
Accuracy	0.90
Precision	0.896
F1-score	0.90
Recall	0.898

Table 3: Observations

VIII. ARCHITECTURE OF WEB RESOURCE

Using Flask, we created a web resource for our prediction model. Now, whenever a user writes a text message, our web page will be requested, and it will load the machine learning model stored in a pickle file. This machine learning algorithm will return to the website after predicting whether the message is bullying or not. And our online resource will show the result. Some of the outputs are attached below,

Figure 6: Home Page

Figure 7: Classification into Cyberbullying text

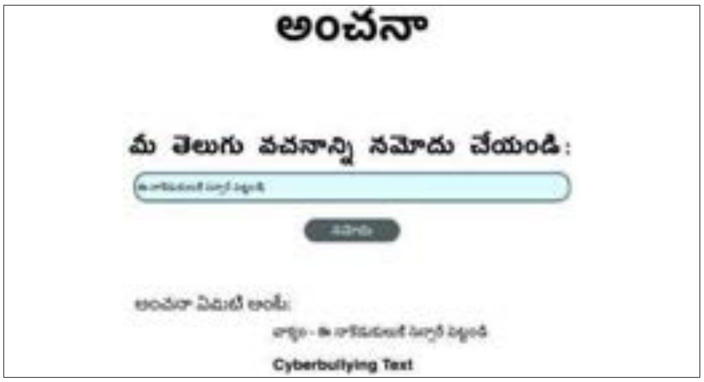


Figure 8: Classification into Non-Cyberbullying text

IX. Conclusion & Future Scope

Our research journey was guided by an extensive review of relevant literature, primarily drawing insights from various influential papers in the field. By synthesizing methodologies and findings from these diverse references, we successfully directed our efforts towards cyberbullying detection in Telugu.

The research outcomes have been particularly promising, with our cyberbullying detection model achieving animpressive 90% accuracy in Telugu. This success can be attributed to the careful selection of IndicBERT, a model tailored for Indic languages, including Telugu. The decision toleverage IndicBERT proved to be instrumental in capturing the subtle linguistic nuances specific to Telugu, contributing significantly to the model's efficiency and accuracy.

Moving forward, there are several possibilities to further improve the performance of our cyberbullying detection model. Expanding the dataset size has proven to be effective in boosting accuracy, as evidenced by the significant improvement achieved by increasing our dataset from 18,000 to 27,000 sentences. By further enlarging the dataset and ensuring a balanced distribution of cyberbullying and non-cyberbullying instances, we provide the model with a more comprehensive and representative learning experience. Another approach is to explore advanced feature extraction techniques using IndicBERT, leveraging its capabilities to capture more intricate linguistic nuances specific to Telugu. Additionally, incorporating multimodal analysis by integrating text, image, and video data can provide a more comprehensive understanding of cyberbullying instances. Implementing real-time monitoring and alerting mechanisms can enable timely intervention and prevention of cyberbullying incidents on social media platforms. Moreover, continuous evaluation and feedback loops will allow us to adapt and refine the model based on evolving cyberbullying patterns and linguistic variations. By embracing these strategies and leveraging emerging technologies, we can further enhance the accuracy, efficiency, and reliability of our cyberbullying detection system in Telugu, contributing to a safer and more inclusive digital environment.

To sum up, this attempt to create a Telugu cyberbullying detection system is more than just a technological development; it also demonstrates our dedication to worldwide online safety, inclusivity, and cultural sensitivity. We make a significant contribution to the overall objective of building a safer, more inclusive, and culturally sensitive digital environment by focusing our efforts on languages like Telugu. In order to guarantee that no community is left unaffected by new cyberthreats, this project acts as a model for future initiatives aiming to address linguistic diversity in the field of online safety. Future advancements in language-specific NLP models could boost the likelihood of detecting cyberbullying in Telugu, integrating multimodal analysis for a thorough comprehension, keeping an eye on Telugu social media platforms in real-time.

References

[1] Chandrasekaran, A., 2023. Natural Language Processing. Chandrasekaran, A.(2023). Natural Language Processing. International Journal of Cybernetics and Informatics, 12(2), pp.57-61.

[2] AlFarah, M.E., Kamel, I. and Al Aghbari, Z., 2023, October. Toward Detection of Arabic Cyberbullying on Online Social Networks using Arabic BERT Models. In 2023 International Symposium on Networks, Computers and Communications (ISNCC) (pp. 1-6). IEEE.

[3] Behzadi, M., Harris, I.G. and Derakhshan, A., 2021, January. Rapid Cyber-bullying detection method using Compact BERT Models. In 2021 IEEE 15th International Conference on Semantic Computing (ICSC) (pp. 199-202). IEEE.

[4] Jain, V., Saxena, A.K., Senthil, A., Jain, A. and Jain, A., 2021, December. Cyber-bullying detection in social media platform using machine learning. In 2021 10th International Conference on System Modeling & Advancement in Research Trends (SMART) (pp. 401-405).IEEE.

[5] Zhang, P., Gao, Y. and Chen, S., 2019, May. Detect Chinese Cyberbulling by analyzing user behaviors and language patterns. In 2019 3rd International Symposium on Autonomous Systems (ISAS) (pp. 370-375).IEEE.

[6] Lavanya, P.M. and Sasikala, E., 2022. Auto capture on drug text detection in social media through NLP from the heterogeneous data. Measurement: Sensors,24,p.100550.

[7] Divyanshu Kakwani, Anoop Kunchukuttan, Satish Golla, Gokul N.C., Avik Bhattacharyya, Mitesh M. Khapra, and Pratyush Kumar. 2020. IndicNLPSuite: Monolingual Corpora, Evaluation Benchmarks and Pre- trained Multilingual Language Models for Indian Languages. In Findings of the Association for Computational Linguistics: EMNLP 2020, pages 4948–4961, Online. Association for Computational Linguistics.

[8] Raj Dabre, Himani Shrotriya, Anoop Kunchukuttan, Ratish Puduppully, Mitesh Khapra, and Pratyush Kumar. 2022. IndicBERT: A Pre-trained Model for Indic Natural Language Generation. In Findings of the Association for Computational Linguistics: ACL 2022, pages 1849– 1863, Dublin, Ireland. Association for Computational Linguistics.

[9] Shanmugavadivel, K., Sampath, S. H., Nandhakumar, P., Mahalingam, P., Subramanian, M., Kumaresan, P. K., & Priyadharshini, R. (2022). An analysis of machine learning models for sentiment analysis of Tamil code-mixed data. Computer Speech & Language, 76, 101407.

[10] Jain, K., Deshpande, A., Shridhar, K., Laumann, F. and Dash, A., 2020. Indic-transformers: An analysis of transformer language models for indian languages. arXiv preprint arXiv:2011.02323.

[11] Mehta, M., Pandey, U., Chaudhary, Y., Sharma, R., Gill, I., Gupta, D. and Khanna, A., 2021, December. Hindi text classification: A review. In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N) (pp. 839-843). IEEE.

[12] Mukku, S.S. and Mamidi, R., 2017, September. Actsa: Annotated corpus for telugu sentiment analysis. In Proceedings of the First Workshop on Building Linguistically Generalizable NLP Systems (pp. 54-58).

[13] Talasila, S. and Vijaya Kumari, R., 2022. Cascade Network Model to Detect Cognitive Impairment using Clock Drawing Test. Journal of Scientific & Industrial Research, 81(12), pp.1276-1284.

[14] Arivazhagan, N., Bapna, A., Firat, O., Lepikhin, D., Johnson, M., Krikun, M., Chen, M.X., Cao, Y., Foster, G., Cherry, C. and Macherey, W., 2019. Massively multilingual neural machine translation in the wild: Findings and challenges. arXiv preprint arXiv:1907.05019.

Secure Data Sharing in the Cloud Through Proxy Re-Encryption Technique

Mrs Jyothi Adilakshmi[1]
Department of Information Technology
MLR Institute of Technology
Hyderabad,India.
bhulakshmi550@gmail.com

Badgoian Rithika[2]
Department of Information Technology
MLR Institute of Technology
Hyderabad,India.
badgoianrithika@gmail.com

Chitikela Pushpalatha[3]
Department of Information Technology
MLR Institute of Technology
Hyderabad,India.
chitikelapushpalatha@gmail.com

Thurpati Venkatesh[4]
Department of Information Technology
MLR Institute of Technology
Hyderabad,India.
venkateshthurpati03@gmail.com

Mandapati Lohitha[5]
Department of Information Technology
MLR Institute of Technology
Hyderabad,India.
lohithachowdary2447@gmail.com

Abstract-

The rapid advancement and widespread adoption of cloud computing have simplified data sharing processes. However, despite being a progressive technology, data security remains a paramount concern. In cloud-based storage, data owners may lack complete control over their data, as it is managed by a third-party entity, the cloud service provider. Securely storing and sharing data on the cloud presents significant challenges, particularly when data owners choose to share information with other parties, such as data users. To tackle this challenge, researchers have explored various encryption techniques in cryptography to ensure secure data sharing on the cloud. By leveraging proxy re-encryption, data owners can securely represent their encrypted data to the cloud through identity-based encryption, introducing an additional security layer to safeguard data during transfer. This approach restricts access to authorized users only, thereby enhancing data security. The proposed method, utilizing proxy re-encryption, is pivotal in facilitating reliable data sharing in cloud computing environments. This technique helps to manage data access effectively while minimizing potential congestion in centralized systems. The proposed approach has undergone rigorous security analysis and evaluation, marking a significant milestone in ensuring data integrity, confidentiality, and security. Such reassurance underscores the effectiveness of the implemented security measures, instilling confidence in the integrity of data management practices.

Keywords-Identity-Based Encryption, confidentiality, integrity, security, proxy re-encryption, centralized system.

I.INTRODUCTION

Cloud computing has become increasingly prevalent in modern technology, aiming to equip users with accessible tools without necessitating specialized expertise [1]. By leveraging cloud services, businesses can streamline operations, reduce costs, and focus on core activities, alleviating IT complexities. However, the increasing popularity of cloud computing has resulted in concerns about security implications, particularly regarding data reliability and privacy [3]. A primary concern revolves around the storage and management of data in cloud environments. It is imperative to carefully evaluate the reliability of cloud service providers before entrusting them with sensitive data. While many providers prioritize security, thorough research, reviews, and certification checks are advisable to ensure data safety [5]. Secure data sharing in the cloud necessitates proprietors for maintaining a control over access permissions, allowing authorized users to access data without interference [6]. Proxy re-encryption (PRE) systems enable proprietors to share encrypted communications with recipients while safeguarding their private keys [7]. Delegates can update encryption methods and keys before transmitting encrypted data to recipients, ensuring that delegates remain unaware of the proprietor's secret key [8].

One approach to data sharing involves proprietors encrypting data before storing it in the cloud and sharing encryption keys with authorized group members. This facilitates independent access and decryption of data without requiring proprietor intervention [9]. In cases where access rights are revoked, data must be re-encrypted with different keys to prevent unauthorized access [10].

This study emphasizes the deployment of new principles to ensure the reliability, consistency, and integrity of data stored in the cloud. Cloud platforms generate public keys for users, enabling secure data storage and dissemination. Data owners encrypt data based on access control policies, and cloud platforms partially decrypt data for recipients upon clearance [13].

II.LITERATURE SURVEY

This section provides a comprehensive summary of authenticated articles reviewed concerning the secure storage of data in the cloud. Surveys and analyses in this section particularly underscore the critical importance of securely sharing data in the cloud. Cloud users play a pivotal role in ensuring the secure and efficient sharing of information resources among themselves.

Regenerative coding offers a solution for recovering damaged data stored across multiple data centers in the cloud. Numerous investigations and studies have been conducted on Privacy-preserving Encryption (PRE) methods to secure and safeguard data in the cloud effectively. Researchers have thoroughly examined the structural principles, security approaches, and efficiency of PRE schemes currently in use. They also discuss potential applications and enhancements of PRE, such as multi-hop delegation, access control, and data sharing. Transferring data securely involves various key considerations. Utilizing encryption methods, such as attribute-based encryption, can ensure that only authorized individuals have access to shared data. Moreover, incorporating robust access control techniques, like role-based access control, can provide an additional layer of security by restricting access to authenticated users who require it. It's vital to conduct regular security checks and continually improve security standards to proactively address emerging risks and maintain a secure environment for data transmission. By prioritizing privacy and adhering to standard procedures, we ensure that information sharing occurs securely and with high protection. The method of securely transmitting data in the cloud is relatively recent, but its significance has increased significantly due to the widespread adoption of cloud services and the growing demand for automated information transmission between individuals.

Drawbacks associated with the current setup are outlined as follows:

FEACS: A Flexible and Efficient Access Control Scheme for Cloud Computing [1] was proposed by Y. Zhang, Jing Chen, Ruiying Du, Lan Deng, Yang Xiang, and Qing Zhou. To ensure that information shared in the cloud remains confidential and secure, they have implemented a Flexible and Efficient Access Control Scheme (FEACS) utilizing attribute-based encryption (ABE). However, it may not be integrated with all cloud computing applications and platforms.

SeDaSC: Secure Data Sharing in Clouds [2] was proposed by MazharAli, Revathi Hamotharan, Eraj Khan, Samee U. Khan, Athanasios V. Vasilakos, Keqin Li, and Albert Y. Zomaya. It addresses privacy, accountability, access control, and internal threat issues. It is a complex system and may prove challenging to establish and manage.

An innovative method for enhancing key generation and management in the AES algorithm [3] was proposed by Omer K. Jasim Mohammad, Safia Abbas, El-Sayed M. El-Horbaty, and Abdel-Badeeh M. Salem (2017). It aims to ensure confidentiality and address ethical concerns. However, creative techniques may require significant effort to oversee compared to conventional approaches.

Data Sharing Attribute Based Secure with Efficient Repeal in Cloud Computing [4], suggested by G. Elavarasan and S. Veni (2020), proposes an approach to optimize effectiveness. However, there are challenges with controlling access privileges and revoking them within a rapidly changing cloud environment.

A Proxy ReEncryption Approach to Secure Data Sharing in The Internet of Things Based on Blockchain [5] was proposed by K. Narsimhulu, Mallepally Prabhavathi, Varala Srijitha, and Thripuram Ajitha. It addresses issues with data sharing. However, this system may be influenced by the efficiency of applications using the Internet of Things.

III.PROPOSED METHODOLOGY

This research proposes a novel method that utilizes proxy-based re-encryption to facilitate the safe and efficient transfer of knowledge among individuals while preserving the privacy of sensitive data. The proposed strategy involves multiple elements working in tandem to ensure that only authenticated individuals can access the information.

The proposed system aims to simplify and enhance data sharing while maintaining the highest level of privacy for highly confidential data. It comprises several interdependent components that collaborate to restrict access to authorized individuals. The trusted authority serves as a pivotal component of this system, responsible for generating private keys for each user, maintaining a registry of these keys, and generating a master secret key for decryption purposes. Owners, upon registration with the Authority of Trust, are obligated to collaborate and convey their information to various parties.

The trusted authority configures user identification and access control permissions before allowing data access. Access control is enforced through user roles, permissions, policies, and assurances to ensure that only authorized individuals can access the data. To ensure confidentiality, authenticity, and secure sharing, the system employs security measures such as encryption and decryption. A proxy server is employed to re-encrypt the essential files, which are then shared with verified users upon owner approval as per user requests. The proxy server plays a crucial role in enforcing security parameters by acting as an intermediary layer facilitating secure data transmission between users and the cloud. Additionally, a cloud service provider is utilized for storage purposes and other functionalities. The proposed system architecture is shown in figure 1.

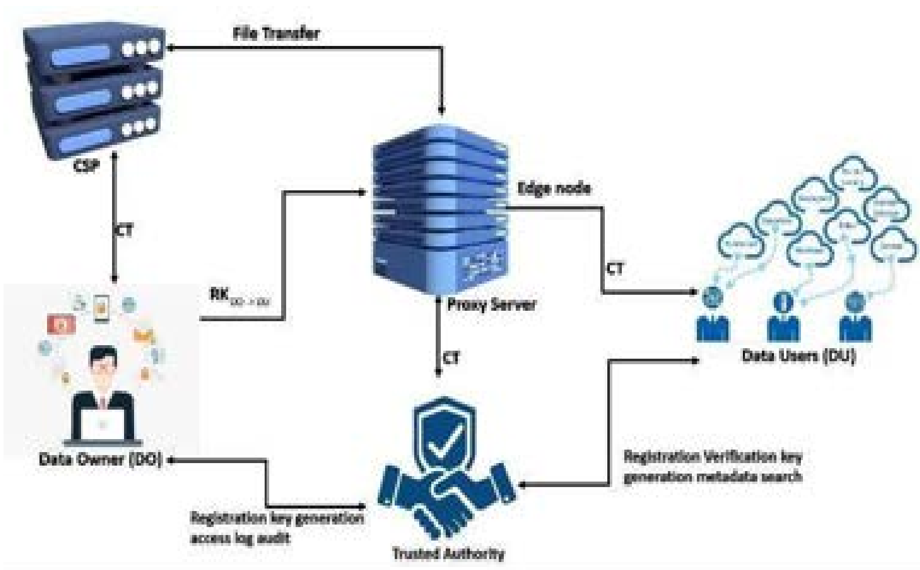


Figure-1: System Architecture

The proposed system architecture includes the following components:

- **Data Owner**: Data owner will have to register initially to get access to the profile. By leveraging cryptographic methods provided by the trusted authority using designated keys, the data owner encrypts the messages. These encrypted messages, called cipher texts, are finally transmitted to a server in the cloud,particularly the CS server.
- **Data User**: The first thing the user has to do is complete the beginning registration process, and the user can log in with sustainable credentials. Individuals who have been approved permission and a proper permit by the trusted authority are acknowledged as data users. The primary goal of a data user is to safely read and decipher encrypted data. To achieve this, the data user appeals to the CSP forencrypted data and keys. By utilizing these keys, the data user moves to implement the decryption process. It is important for data users to ensure they have the necessary authorizations and adhere to the decryption process to sustain data security.
- **Trusted Authority:** In this module, trusted authority is an asset in overseeing the cryptographic secret code and making sure thatthe correct authorization is given to data users. A trusted authority is an important component in maintaining the reliability and consistency of sharing data in the cloud. The trusted authority acts as a trusted entity that supervises the cryptographic keys, authorizes access, and maintains the general security of the system. By trusting the trusted authority, organizations can inaugurate a secure framework for sharing data, preserve sensitive information from unauthorized access, and guarantee the privacyof users.
- **Proxy Server:** This module incorporates the proxy server, which makes sure that the security of PRE guarantees that important data in the re-encrypted file can't be accessed by both the server or proxy and an unexpected user. The proxy server is liable for executing the re-ciphering process and recasting the owner's masked info into a format that the intended user can decode. safe during transfer and storage in the cloud, delivering an extra layer ofassurance for shared data.
- **CSP**: In this module, we create a cloud service provider (CSP). The cloud server shall offer an enabling function that permits it to inspect file content despite not having any information regarding the owners or users. The data owner's transmitted files are preserved as fragments and segments on the server. In addition, there is a submodule called cloud server includes proxy. The server in the cloud provides users with authorization for using the newly encrypted files.

This study will go through the arithmetic algorithm that forms the foundation of this system, known as Proxy Re-Encryption (PRE). It gives access to many users to decrypt specific data while it will make sure that unauthorized individuals can't access through plaintext. We will offer an in-depth and a well-founded explanation of how is this algorithm is working to securely share datausing PRE.

During this stage, with use of the setup we generate public parameters and master key by given security inputs. Given below formula is used for generating public key.

Key generation:

In this phase, by applying Key-gen (MK,id), this study generates the user's secret key. By making use of this formula and the proxy re-ciphering approach, the proprietor can securely share data in the cloud by re-encrypting the secret key with the help of the identifier and cryptographic operations.

Encryption:

From the input message (m) and input identity (id), with the condition that a random value (r) is suitable the sender creates a second-level ciphertext bycomputing the ciphertext.

Re-KeyGeneration:

The re-encryption key is generated through the Re-Keygen process by a new security key is created in order to compute a fresh cipher text through the Re-KeyGen process.

Re-Encryption:

The proxy re-encryption technique, an approved ciphertext (c1) is encrypted again with a newly created re-encryption key to give an additional layer of security. This is concluded to provide extra layer of security for pstop from attackers, threats etc.

Decryption Phase 1:

This step helps the representative gain access to the second-level ciphertext through the representative tries to side wisely decrypt the first-level ciphertext by using the identity-based decryption with the first private key.

Decryption Phase 2:

To gain the primary message, the representative decrypts the second-level ciphertext using the private key create for their specific identity. This will make sure that only verified individuals with the correct private key can access the original message, preserve the confidentiality and integrity of the data.

IV. RESULTS & DISCUSSION

Proxy re-encryption give an additional layer of security by allowing the data owner to delegate access rights and cancel them if necessary. The proxy re-encryption technique displays great outcome in securing data sharing in cloud computing. By using this technique, authorized delegates can securely access and decrypt data without compromising its confidentiality. This method permits for controlled and fine-grained access to shared data, make sure that only authorized individuals can recover and view the information. This helps preserve sensitive data while facilitating collaboration and efficient data sharing in cloud environments.

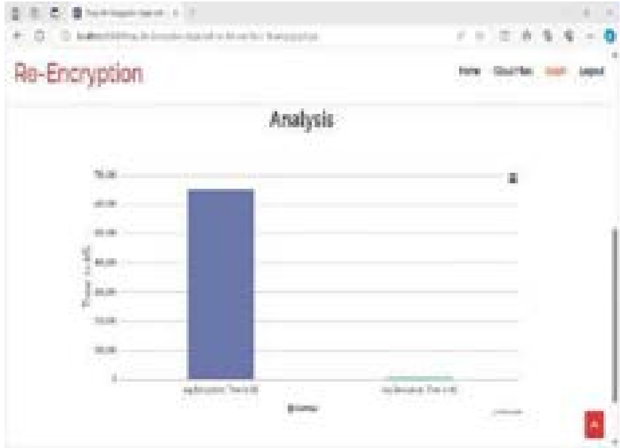


Figure:2 Avg Time of Encryption and Decryption

The time needed for ciphering and deciphering depicted above in figure 2.

Figure:3 Home page

Figure: 4 Upload files

In the upload the files of data owner will be encryptedand to be stored in cloud as shown in fig. 4.

Figure: 5 Uploaded file details

The uploaded file details will be shown as depicted in fig. 5 and we can seethe re-decryption key in the page.

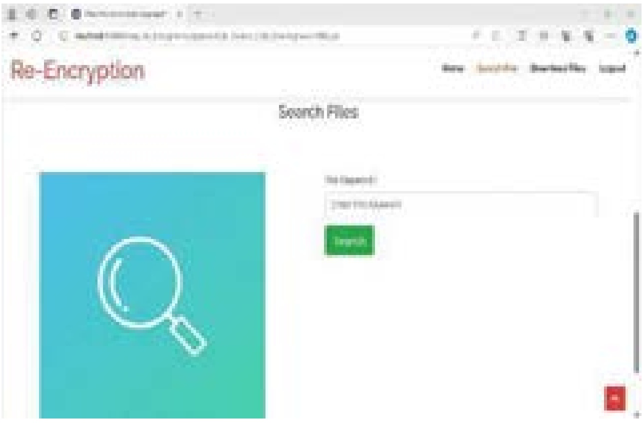


Figure: 6 Search File

This page shown in fig 6 is used to search a required file by data user and next request for it.

Figure: 7 Download File

The page shown in fig 7 is used to download the requested file when accepted by owner and re-decryption key is sent by server to user so they can download it.

V.CONCLUSION

In conclusion, while data sharing through cloud computing is widely practiced, it presents significant challenges in terms of data security. To address these challenges, we have developed a secure and robust solution for sharing data among individuals securely. Our system enables data owners to encrypt their data and store it in the cloud, where it is further re-encrypted by a proxy server before being transmitted to authorized users. The use of proxy re-encryption techniques represents a significant advancement in enhancing the security of data sharing in cloud environments. By employing proxy re-ciphering algorithms, we ensure that data dissemination occurs in a highly secure manner, resistant to attacks or unauthorized access. This approach not only enhances confidentiality but also safeguards the integrity and authenticity of shared data. As technology continues to advance, the integration of such innovative techniques becomes increasingly crucial for establishing a trustworthy foundation in cloud data sharing systems. By implementing proxy re-encryption, organizations can establish a robust framework that facilitates secure and controlled data circulation in the cloud, effectively addressing concerns related to unauthorized access and data breaches. Moving forward, continued improvements in technology will further strengthen the security posture of cloud-based data sharing systems, ensuring the integrity and confidentiality of shared information.

ACKNOWLEDGEMENT

I am extremely grateful to the Department of Information Technology at MLRIT College for their invaluable support in completing my dissertation. Additionally, my heartfelt appreciation goes out to our esteemed faculty, project coordinators, and friends for their guidance and constructive feedback throughout the duration of theproject.

REFERENCES

[1]. Yu Zhang, Jing Chen, Ruiying Du, Lan Deng, Yang Xiang, Qing Zhou "FEACS: A Flexible and Efficient Access Control Scheme for CloudComputing," IEEE Symposium on Trust, Security and Privacy in Computing and Communications,2014.

[2]. Mazhar Ali; Revathi Dhamotharan; Eraj Khan; Samee U. Khan; Athanasios V. Vasilakos, Keqin Li, Albert Y. Zomaya "SeDaSC: Secure Data Sharing in Clouds," IEEE Systems Journal, Volume: 11, **pp.** 395 - 404,2017.

[3]. Omer K. Jasim Mohammad, Safia Abbas, El- Sayed M. El-Horbaty, Abdel-Badeeh M. Salem "Innovative method for enhancing key generation and management in the AES algorithm" International Journal of Computer Network and Information Security,2017.

[4]. G.Elavarasan, S.Veni "Data Sharing Attribute Based Secure With Efficient Revocation in Cloud Computing," IEEE Symposium on Computing and Information Technology (ICCIT-1441),2020.

[5]. Mallepally Prabhavathi, Thripuram Ajitha, K. Narsimhulu, Bandameedi Mahitha, VaralaSrijitha "A Proxy ReEncryption Approach ToSecure Data Sharing InThe Internet Of Things Based On Blockchain," Journal of Survey in Fisheries Sciences,2023.

[6]. J. Katz, R. Canetti, and S. Halevi, "Chosen-ciphertext security from identity-based encryption," in Proc. Int. Conf. Theory Appl. Cryptographic Techn., Springer, 2004, pp. 207–222.

[7]. K. Fu, M. Green, G. Ateniese, and S. Hohenberger, "Improved Proxy Re-encryption schemes with applications to secure distributed storage," ACM Transactions on Information and System Security, vol. 9, no. 1, pp. 1–30, 2006.
[8]. Xia, Q., Sifah, E. B., Cobblah, K. O. B. O., C. N. A., Xia, H., Agyekum, & Gao, J. (2021). A Proxy Re-Encryption Approach to Secure Data Sharing in the Internet of Things Based on Blockchain. IEEE Systems Journal.
[9]. Xuan, S., Zhang, Y., Tang, H., Chung, I., Wang, W., & Yang, W. (2019). Hierarchically authorized transactions for massive internet-of-things data sharing based on multilayer blockchain. Applied Sciences, 9(23), 5159.
[10]. S. Seo, M. Nabeel, X. Ding, and E. Bertino, "An Efficient Certificateless Encryption for Secure Data Sharing in Public Clouds," IEEE Trans. Knowl. Data Eng., vol. 26, no. 9, pp. 2107– 2119, Sep. 2013.
[11]. L. Xu, X. Wu, and X. Zhang, "CL-PRE: A certificateless proxy re-encryption scheme secure data sharing with public cloud," in Proc. 7th ACM Symp. Inf, Comput, commun. Security, 2012, pp. 87-88.

Crop Monitoring using IoT for Precision Agriculture

Satyam Shinde
Department of Computer Engineering
GHRCEM
Pune, India
satyamshinde707@gmail.com

Tushar Dhanurkar
Department of Computer Engineering
GHRCEM
Pune, India
tushar.dhanurkar@gmail.com

Vinit Jain
Department of Computer Engineering
GHRCEM Pune, India
vvvjain01@gmail.com

Vivek Pal
Department of Computer Engineering
GHRCEM
Pune, India
vivekpal090102@gmail.com

Prof.Pradeep Taware
Department of Computer Engineering
GHRCEM
Pune, India
Pradeep.taware08@gmail.com

***Abstract* — Internet of things (IoT) is an emerging technology that shows the future of computing and networking. Agricultural monitoring from a remote location is one of the essential applications of IoT based Wireless Sensor Networks (WSNs). The applications face problems due to the dynamic changes in the environment. Traditional farming methods frequently suffer from imprecise delivery of water and nutrients to crops, resulting in inefficient resource usage and suboptimal yields. This research study addresses the pressing need for enhanced precision agriculture techniques aimed at optimizing crop monitoring and resource utilization.**

Keywords: IoT Application, Farm Automation, Soil Moisture Monitoring, Farming Technology, Autonomous Farming.

1. INTRODUCTION

India has demonstrated remarkable progress in agriculture over the past four decades, effectively meeting the increasing demand for food [1]. This achievement can be attributed to several factors, including the adoption of modern crop varieties, increased utilization of inputs, and strategic investments that have expanded irrigated areas. However, despite the initial success of "Green Revolution" technologies, their impact has diminished in certain regions, necessitating the development of new technologies to further enhance productivity, promote sustainability, and diversify cropping patterns [2]. Moreover, addressing the challenge of advancing yield boundaries is crucial, alongside the urgent need to leverage the potential of rain-fed and less-fertile areas. Given India's diverse agro-ecological landscapes and farming communities, a broad spectrum of solutions is required. To achieve accelerated, widespread, and targeted growth, agricultural research should focus on addressing location-specific challenges, while agricultural systems must integrate cutting-edge technologies to benefit resource-constrained farmers.

As the world's population is expected to grow significantly, the agricultural sector faces the formidable challenge of producing 70% more food by 2050 [3]. This challenge is exacerbated by diminishing agricultural lands and finite natural resources. To tackle these issues, smart agriculture solutions utilizing the Internet of Things (IoT) are becoming increasingly crucial. IoT technology aids farmers in optimizing resource utilization, enhancing crop yields, and reducing operational costs. Predictive analytics, coupled with sensor data, empower farmers to make informed decisions regarding crop production, storage, marketing, and risk management. Agricultural drones, livestock tracking systems, and smart greenhouses are employed to monitor and improve various farming practices [4]. With the continuous advancement of IoT technology, the adoption of IoT devices in agriculture is expected to rise, leading to a projected tripling in size of the global smart agriculture market by 2025.

To enhance agricultural productivity, automation and robotics are being integrated into farming operations [5]. The development of an autonomous farming robot powered by microcontrollers aim to automate on-farm tasks such as seeding and fertilizing. This technology reduces human effort while increasing crop yield. The robot can precisely plant seeds at predefined distances and depths, controlled by microcontrollers [6]. Such innovations play a crucial role in addressing the shortage of skilled labor in agriculture and can significantly impact the lives of the majority of resource-poor farmers. Consequently, the objective is to develop an integrated IoT-based system using Arduino Uno and sensors to accurately measure environmental parameters such as temperature and humidity. By regulating water and fertilizer delivery, it ensures crops receive optimal nutrients, thus reducing wastage and labor. Real-time data transmission enables farmers to remotely monitor fields, facilitating timely decision-making [7].

The Internet of things (IOT) is revolutionizing the agriculture enabling the farmers with the wide range of techniques such as precision and sustainable agriculture to face challenges in the field. IOT technology helps in collecting information about conditions like weather, moisture, temperature and fertility of soil, Crop online monitoring enables detection of weed, level of water, pest detection, animal intrusion in to the field, crop growth, agriculture [8]. IOT leverages farmers to get connected to his farm from anywhere and anytime. Wireless sensor networks are used for monitoring the farm conditions and micro controllers are used to control and automate the farm processes. To view remotely the conditions in the form of image and video, wireless cameras have been used. A smart phone empowers farmer to keep updated with the ongoing conditions of his agricultural land using IOT at any time and any part of the world. IOT technology can reduce the cost and enhance the productivity of traditional farming [9].

2. LITERATURE REVIEW

In this section the brief summary on internet of things and crop monitoring is presented.

In Li Da Xu et al work [10], the primary focus is on the Internet of Things (IoT) and its strategies, emphasizing identification, tracking, communication, and service management. The paper discusses key IoT applications and the associated challenges particularly regarding security and standardization Healthcare is highlighted as a significant IoT domain, employing medical sensors and wearable devices to monitor health parameters and improve patient care. Security concerns in IoT are underlined, necessitating the enhancement of existing security protocols.

Madakam, S. Ramaswamy, and Tripathi [11], conducted a comprehensive exploration of IoT concepts through a systematic review, drawing insights from diverse sources such as corporate white papers, professional discussions with experts, and online databases. Their research delves into defining IoT, tracing its origins, elucidating fundamental requirements, delineating its characteristics, and identifying various aliases associated with this burgeoning technology. The primary aim of their paper is to furnish readers with a broad overview of IoT, encompassing its architectures, and crucial technologies, while also highlighting their pervasive integration into our daily lives. Through their work, they provide a coherent narrative that not only educates but also contextualizes the significance and impact of IoT in contemporary society, laying a foundation for further exploration and innovation in this rapidly evolving field.

Kavitha B C, Shilpa D P [12], IoT technology facilitates the gathering of data pertaining to factors such as atmospheric conditions, moisture levels, temperature, and soil fertility. Online monitoring of crops enables the identification of weeds, water levels, pest presence, animal encroachment into fields, crop development, and agricultural progress. Wireless sensor networks are deployed for monitoring agricultural conditions, while microcontrollers are employed to automate and manage farm operations. For remote viewing of conditions through visual media, wireless cameras are utilized.

Martina Corti, Pietro Marino Gallina adopted precision agriculture for enhancing both the environmental sustainability of crop systems and farmers' financial returns [13]. Farmers transitioning to precision agriculture require efficient protocols that demand minimal inputs to delineate homogeneous management zones, facilitating optimized actions without prior knowledge such as yield maps. Results showed that red-edge and near-infrared (NDRE) bands emerged as the most dependable data sources for identifying management zones, achieving correct classifications of 62%, 68%, and 74% at early tillering, stem elongation, and late booting stages, respectively.

Anand Nayyar and Er. Vikram Puri [14] delve into the transformative potential of Internet of Things (IoT) technology, which has revolutionized various aspects of daily life by endowing objects with intelligence and connectivity. At the heart of IoT lies a network of interconnected devices capable of autonomous configuration. In the agricultural sector, the emergence of Intelligent Smart Farming IoT devices is revolutionizing traditional practices, enhancing production efficiency, reducing costs, and mitigating wastage. The central focus of their paper is the introduction of an innovative Smart IoT Agriculture Sensor, meticulously crafted to provide farmers with real-time data crucial for decision-making, including temperature and soil moisture levels. By harnessing the power of IoT, this sensor empowers farmers with actionable insights, enabling them to optimize resource allocation and enhance agricultural productivity in an increasingly dynamic environment.

3. METHODOLOGY

The proposed IoT (Internet of Things) system aims to enhance connectivity, efficiency, and data management in diverse applications. The block diagram of proposed methodology is presented in Fig. 3.a.

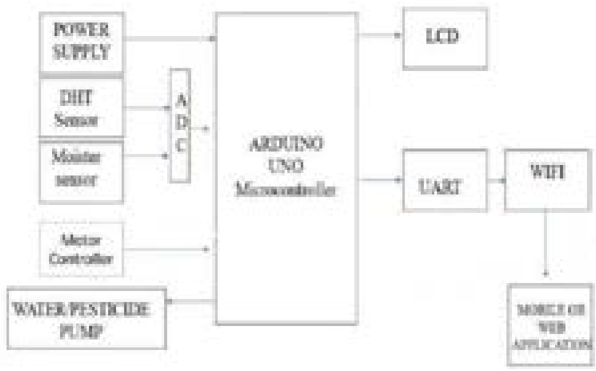


Fig 3.a :- Block Diagram of Proposed Methodology

The SSF 225 motor, connected to the Arduino Uno, regulates the flow of fertilizer by activating at specific intervals as programmed. This ensures that crops receive adequate nutrients from the onset, fostering healthy growth and development.

In terms of water management, the DTH22 sensor plays a pivotal role by continuously monitoring soil moisture levels. This data is relayed to the Arduino Uno, which, based on predefined thresholds, activates the water pump for irrigation when soil moisture levels drop below a certain point. By employing this automated irrigation system, crops receive water tailored to their specific needs, thereby mitigating the risks of under or over-watering.

Furthermore, to keep the farmer informed about the farm's condition in real-time, a Wi-Fi module is utilized to connect the Arduino Uno to the internet. This enables the transmission of timely updates to the farmer's smartphone or computer, including information on soil moisture levels, temperature, humidity, and fertilizer application. Prompt access to this data empowers the farmer to take immediate action in response to any detected issues, such as low soil moisture or abnormal temperature conditions.

Moreover, an LED display connected to the Arduino Uno provides on-site information to the farmer, displaying real time data such as soil moisture levels, temperature, and humidity. This feature offers a convenient means for the farmer to monitor the farm condition at a glance, facilitating timely decision-making.

The methodology outlined for crop monitoring using IoT components enables precise agricultural practices, encompassing fertilizer application, irrigation management, and real-time data dissemination. By leveraging these technologies, farmers can optimize crop growth, conserve resources, and maximize yield potential, ultimately contributing to sustainable agricultural practices.

Our project, titled "Crop Monitoring Using IoT for Precision Agriculture," stands out amidst existing endeavors through its integrated approach to automated temperature detection and precise irrigation facilitated by a mobile farming robot. Unlike conventional systems, our solution harnesses IoT technology to enable the robot to autonomously sense temperature variations, navigate between crops, and administer water as needed. Moreover, the portability and mobility of our robot enhance its adaptability across diverse agricultural landscapes, amplifying efficiency and ensuring optimal crop health. This distinctive combination of real-time sensing, autonomous mobility, and portability sets our project apart, offering a comprehensive solution for enhancing agricultural productivity and sustainability.

4. MODEL IMPLEMENTATION

The implementation of our IoT model involves deploying sensor nodes to collect data from the physical environment. The hardware setup requirement for purposed model is given below.

4.1 Hardware Setup:

A. **Arduino Uno:** The Arduino Uno serves as the central controller for the system, collecting data and controlling various components.

B. **DHT22 Sensor:** This sensor measures temperature and humidity, providing crucial environmental data for crop health assessment.

C. **SSF-225 Motor:** This motor can be used for controlling the release of fertilizers or other agricultural inputs.

D. **Water Pump:** It controls the irrigation system, ensuring that the crops receive the right amount of water.

E. **LED Display:** The LED display can be used to show real-time data, alerts, and system status.

F. **Wi-Fi Module:** A Wi-Fi module connects the system to the internet, enabling remote monitoring and data transmission.

4.2 Temperature and Humidity Reading

The DHT-22 helps us to get the temperature and humidity value at any particular time. Keeping a check in the temperature and humidity value will help the user in taking care of livestock and special plants. From DHT22, the signal gets passed to ESP8266 which further passes the signal to the IP address from where the user can get notified even when away from the field.

Fig 4.2(a) :-Temperature Reading of Farm

Fig 4.2(b) :- Humidity Reading of Farm

In a Crop Monitoring model, figure 4.2 (a), and figure 4.2 (b) are displaying temperature and humidity readings respectively. These data are collected from sensors deployed in the environment.

4.3 Soil Moisture Determination

Each data point on the graph would correspond to a specific time interval and would show the corresponding temperature and humidity levels recorded by the sensors at that time. This visualization helps users to monitor changes in temperature graph (figure 4.3(a)) and humidity graph (figure 4.3(b)) over time, identify patterns, and detect anomalies, enabling them to make informed decisions based on the data.

Fig 4.3(a) :- Weather Temperature Graph of Farm

Fig 4.3(b) :- Humidity Graph of Farm

The soil moisture sensor determines the moisture content of the soil and temperature of Farm. If the moisture of Farm is below the threshold value (60) or above the temperature threshold value (37), it sends signal to the DC motor i.e., the pump and automatically the pump gets on and waters the field. By customizing farmer can set fertilizer to the farm by reading.

4.4 Implemented Model

Fig 4.4 (a) :- Implemented Model In Real Life

Figure 4.4 (a) is representing the build of small prototype of our farming robot.

5. CONCLUSION

The proposed model optimizes the agricultural process and improve crop yields. By setting fertilizer with sowed seeds, providing the right amount of water based on soil requirements, and notifying the farmer about the farm's condition, the proposed model has created ideal growing conditions for crops. Lastly, integrating fertilizer application with seed sowing ensures that the growing plants receive vital nutrients right from the start, nurturing healthy growth and ultimately improving yields. Secondly, customizing irrigation schedules based on soil requirements helps to prevent both overwatering and underwatering.

6. REFERENCES

[1] S L Shylaja et al 2021 J. Phys: Conference .Ser. 2027 012019 , IoT based crop monitoring scheme using smart device with machine learning methodology , Conference Series , ICRABE 2021 , Iop Publishing.

[2] T.S.K. Durgesh Nandhini, Dr. S. Pushparaj, "Does Green Revolution Created Mono – Cropping Pattern In India", UGC India Approved Journal - Journal No 63466,January 2018International Research Journal of Business Studies 6(1):216-224

[3] Mitchell C. Hunter, Richard G. Smith, Meagan E. Schipanski, Lesley W. Atwood, David A. Mortensen, Agriculture in 2050: "Recalibrating Targets for SustainableIntensification, BioScience ", Volume 67, Issue 4, April 2017, 386–391

[4] Muthumanickam Dhanaraju, Poongodi Chenniappan , Kumaraperumal Ramalingam, Sellaperumal Pazhanivelan, Ragunath Kaliaperumal."Smart farming: Internet of Things(IOT) -Based Sustainable Agriculture" , Volume12 , issue 10, MDPI : 21 October 2022

[5] Saiful Azimi ,Mohamad Shukriya Zainal Abidin , Abioye ABIODUN Emmanuel , Hameedah Sahib Hasan ; "Robotics and Automation in Agriculture : present and Future Applications" April 2020 ; Vol 4 130-140 ; ARQII Publication.

[6] Terteil A. A. Ali, Viraj Choksi, Dr. M B Potdar, Precision Agricultural Monitoring Systems Using IoT, Vol. 6, Issue 11, November 2017,International Journal of Innovative Research in Science, Engineering and Technology

[7] Anil Kumar D B ; Doddabasappa N ; Bansilal Bairwa ; Anil Kumar C S; Ganesh Raju; Madhu,IoT-based Water Harvesting, Moisture Monitoring, and Crop Monitoring System for Precision Agriculture,IEEE, 21 June 2023

[8] S. Praveen Chakkravarthy1, Gaddam Vivek , Sahithi Vangala3, Kalakuntla Vishal ,K. Yasash P. BharathChandra G. Sai Kiran7,Smart Precision Agriculture using IoT Data Analytics ,Vol 71 No. 4 (2022),Mathematical Statistician and Engineering Applications.

[9] Dr. Uttam A. Deshmukh,Smita B. Jain ,Design and Implementation of an IoT-Based Crop Monitoring System, 2023 IJNRD | Volume 8, Issue 9 September 2023

[10] Li Da Xu, Senior Member, IEEE, Wu He, and Shancang Li IEEE Transations On Industral Informatics, VOL. 10, NO. 4, NOVEMBER 2014, Internet of Things in Industries: Survey

[11] Somayya Madakam, R. Ramaswamy, Siddharth Tripathi,Internet of Things (IoT): A Literature Review,Journal of Computer and Communications > Vol.3 No.5, May 2015, National Institute of Industrial Engineering (NITIE)

[12] Kavitha B C, Shilpa D P, Thanushree K S, Swathi A M, Ranjitha M K Dept of ECE, BGSIT, Agricultural Crop Monitoring Sensors using IoT-A Study, NCESC – 2018

(Volume 6 – Issue 13) 24-04-2018, IJERT

[13] Martina Corti 1, Pietro Marino Gallina ORCID, Daniele Cavalli 1, Bianca Ortuani, Giovanni Cabassi, Gabriele Cola ,Antonio Vigoni, Luigi Degano and Simone Bregaglio ; "Evaluation of In-Season Management Zones from HighResolution Soil and Plant Sensors" ; Volume 10 Issue 8,12 July 2020.

[14] Anand Nayyar Assistant Professor; "Smart Farming: IoT Based Smart Sensors Agriculture Stick for Live Temperature and Moisture Monitoring using Arduino", The International Conference on Communication and Computing Systems (ICCCS-2016) (2016)

Comparative Analysis of FOPID and 1+PI Controllers in Two-Area Load Frequency Control Systems under Dynamic Operating Conditions

Amit Mittal
Dept of Allied Sciences
Graphic Era Hill University
Bhimtal
amitforestry26@gmail.com

Raghav Garg
Dept of Computer Science
Tula's Institute Dcehradun
India
ragav.garg@tulas.edu.in

Akhilesh Singh
Deptt. of Electrical Engg.
Uttarakhand Technical University
Campus Institute,NPSEI, Pithoragarh
akhil09msingh@gmail.com

Suchi Jain
Dept of Computer Science
Tula's Institute Dcehradun
India
Suchi.jain@tulas.edu.in

Shweta Goyal
Deptt. of Electrical Engg.
Graphic Era Deemed to be University
Derhadun India
Shwetugoyal@gmail.com

Girish Singh Bisht
Dept of Computer Science
Tula's Institute Dcehradun
India
girish@tulas.edu.in

Abstract—**For the purpose of preserving stability and reliability in the face of fluctuations and disturbances, the electrical grid is dependent on load frequency control, often known as LFC. The percentage decision point of view (FOPID) controller and the 1+PI controller are compared in this work within the context of two linear feedback control (LFC) domains. The controllers are put through their paces in this study under a variety of conditions, with the random wind generation in area 1 and the uniform wind production in area 2 being taken into consideration. Through simulation tests, FOPID and 1+PI controllers are evaluated for their capacity to manage system frequency and maintain network stability. In this evaluation, key performance parameters such difference, tie line fluctuation, and system stability are taken into consideration.**

Keywords—Load Frequency Control (LFC), Two-Area System, Thermal Power System, Wind Generation (WG), FOPID Controller, 1+PI Controller

I. Introduction

Load frequency control regulates energy production and consumption and ensures power grid reliability. Low-frequency generators struggle to absorb renewable energy, notably wind energy. A variety of system frequencies and energy changes impede wind energy interaction. Power frequency must be modified for renewable energy to protect the grid. It needs top-level management. Comparison of 1+PI and FOPID controllers in dual LFC circuits. The study compares stepper and random generator controllers. Unblocking requires FOPID and 1+PI controller understanding. This study analyses deployment, connection, and security [1-2].

Load frequency control (LFC) systems with FOPID and 1+ PI controllers are popular. The controller stabilises and controls renewable energy and energy storage frequency. Our LFC system FOPID and 1 + PI controller assessment covers performance, pros, and disadvantages [3-6].

Previous research has shown that FOPID controllers can detect weak events, even non-events. FOPID controllers detect most electronic component changes using discrete codes, making them precise and adaptive. Low frequency control (LFC) attenuation, stability, and frequency control have been researched for FOPID controllers. The 1+PI controller can be used for LFC systems. The controller is simple and reliable. Integration of 1+PI improves proportional-integral (PI) controller controllability and stability. Recent studies show that LFC systems with 1+PI controllers can monitor, tune, and maintain stability even when parameters change. Electronic control methods are ideal for 1+PI controllers due to their real-time applicability and ease of use [7-8]. Many studies compare FOPID controllers to 1+ PI controllers, yet both offer benefits due to low frequency control. Wind generation disparity between LFC system locations is the basis for this hypothesis. This study compares the integral part (PI) and complete order part by part (FOPID) controllers. This study focuses on two-area load-frequency converter (LFC) systems that use periodic or random wind power. In this work, simulation and performance tests analyse several renewable energy integration methods for poor fuel economy. Research [9-11] has also examined the pros and cons of each therapy method.

The return principle develops AGC's adaptive controller. Policies are analysed and improved by PI to improve management. Controlling with PI is better than LQR without knowing how the system works [12-13]. PI-based control agreement was evaluated in two uncontrolled zones. Solar thermal and wind turbine generators are renewable energy innovations. Once, this technology will power homes and businesses. Systems change grid frequency and output voltage. This research designs hybrid generator frequency control PID controllers to control operating limit frequency. PID controller parameters are optimised by BBBC. Test controller performance with random loads. This study changes considerably with the Iteration Principle. We offer this control to solve automatic production problems (AGC) in numerous fields. The controller operates two zones without resetting a thermal zone using the second horizontal regulator (LQR). Evaluate your physical performance with metrics. For high-frequency items, LQR supports performance best [14-15].

Choose a renewable energy source using multidimensional decision-making. Small, unreliable data and conflicting standards slow the process. Modified intuitionistic fuzzy sets (MIFS) can fix these difficulties. Regulation-driven MIFS improves behaviour. We present MIFS architecture RES in this paper. Choquet creates the MIFS data MIF Choquet Integral (MIFCI) integration operator. Finalising

RES options with staff. Research shows IFS solves real-world issues. The IFS recruits oddballs. Lack of information, prior knowledge, events, and context affect uncertainty measurement. Enhance IFS member uncertainty index. Improvements to IFS non-transitive variable parameters improved default behaviour. A keypoint and geometric clustering method for random and uncertain data is presented in this chapter. The MCDM candidate selection strategy works [16-20. Nowadays, the control algorithms are being used in other areas to investigate their performance, some of them are given in [21-24]. Further the use of different control schemes in LFC is reported in [25-28]. All these control schemes show that LFC is still an area of exploration in terms of control approaches, optimization, load perturbation, etc. Wireless charging of electric cars in living rooms raises potential concerns. A Moderator-Intuitionistic-Fuzzy Hybrid Averaging Operator can classify sustainability of renewable energy sources [29-31]. This study introduces an optimization-based artificial neural network model for charging electric vehicles using intuitionistic fuzzy geometric aggregation operators to make decisions based on multiple factors [32-34].

II. CONTROL SCHEMES

A. FOPID Controller

In load frequency control, FOPID controllers have various benefits. It's clear that traditional PID control is more elastic and flexible than deterministic function. Power uncertainty is captured by the FOPID controller, enhancing system response and stability. FOPID is more flexible and versatile than other controllers due to its higher control capabilities. FOPID controllers can also switch transmission lines, especially energy storage-renewable energy source connections.

When traditional PID controller compared to the FOPID controller, various issues must be solved. This method would take longer and need more steps without operational disturbance. The FOPID controller outperforms the PID controller. Direct use of the FOPID controller in big power plants is restricted by its advantages. Measure FOPID controller quality to improve LFC system performance and reliability. The current power system is inconsistent and unreliable.

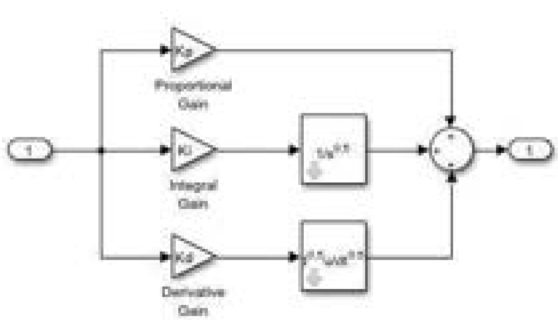


Fig1. The subsystem of FOPID controller.

The classic PID model is expanded upon by FOPID controllers, which accomplish this by merging proportional, single, and variable parts. Three parameters—mostly equal, one, and variable—represent the many variables that are required to solve the problem of instantaneous power conversion. These parameters are what distinguish the FOPID controller from other controller technologies.

$$output = K_p + \frac{K_I}{s^{-\lambda}} + K_D s^{\mu} \tag{1}$$

Fractional order PID (FOPID) controllers make discrete decisions, improving complex behaviour modelling and control. By rectifying these negative decisions, the FOPID controller can adapt to varied processes and needs, improving product efficiency and resilience load frequency control. The FOPID controller's mathematical model is more sophisticated than ordinary PID controllers. FOPID controllers are effective for fire truck difficulties because of their simplicity and convenience. Modern electricity comes from renewable energy and energy storage.

B. 1+PI Controller

The 1+PI controller is designed for product stability and uses LFC technology for heavy loads, making it ideal for instantaneous power transmission due to its low demand. It's great for transmitting power instantly because of its low demand. 1+PI controllers are simple but reliable, suited for many designs and purposes, even when disturbed or uncertain. 1+PI controller issues include uniformity, measurement precision, and easy control. Poor management due to strict rules and individual work can make it hard to behave improperly or sporadically.

Slow response and overshoot limit 1 + PI controller for low voltage quick switching or overshoot. In complex activities, the controller's performance will decline if its normal control system cannot control the dynamic system. LFC systems with simple and reliable applications can benefit from the 1+PI controller despite its limitations. Before selecting and using a control mechanism, power system performance and dependability must be analysed for pros and cons.

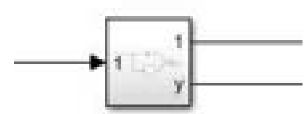


Fig 2. The subsystem of 1+PI controller.

III. SYSTEM DESCRIPTION

Each area of a thermal power plant has undergone testing for two-area load frequency control (LFC). Two-area employs a progressive wind pattern, while area-1 generates a random wind pattern. There is an increasing inclination towards integrating renewable energy sources into thermal power plants; yet, the factors involved exhibit interdependence and unpredictability. Thermal power plants play a crucial role in meeting the energy demands of various regions. By employing the identical equation, these centres have the capability to establish connections between control equipment, such as valve rooms and turbine control centres, and the generation of power. The performance of boilers, turbines, and pressure regulators has a significant impact on the LFC system of thermal power plants. Intense winds will result in tangible consequences, particularly in area 1 and 2. Phase 1 storms can induce alterations, whereas Phase 2 wind power can result in swift power fluctuations. The dynamics of power plants are demonstrated through the utilisation of several variables, including wind speed, turbine blade pitch control, and energy production reaction. The variability of wind power in relation to grid frequency and tie-line power fluctuations between two locations poses additional difficulties in terms of frequency regulation and system stability. Effective modelling and testing precisely depict the behaviour of the LFC system when subjected to change and influence, enabling the evaluation and optimisation of control measures.

IV. Parameter tuning using GA

Load frequency control (LFC) systems do not need controller optimisation for stability and performance. Genetic algorithms (GA) optimise control systems. The proportion of events (FOPID) and 1+PI controllers are examples. This sections illustrates GA's FOPID and 1+PI controller characterization process.

A. Parameterization of FOPID using GA:

Genetic algorithms use selection, mutation, and other genetic processes to optimally find a preliminary solution. Evolution and natural selection underpin these algorithms. GA employs FOPID controller parameterization for percentage decisions. You must understand that the FOPID controller parameterization determines the percentage of proportional, proportional, and variable inputs that govern the GA's LFC system. Typically, parameterization begins with a chromosome model, where each chromosome represents a solution. Every parameter has a range, and a random integer creates the original chromosomal population. To assess chromosomal safety, the objective—stabilizing or reducing variations—is used. GA optimises the LFC system FOPID controller using generation, crossover, and transition selection to return the population to the ideal integration values [12–14].

B. Parameterization of 1+PI using GA:

The evaluation of gain (Kp), gain (Ki), and mixing time (Kp2) is another type of analysis that GA can do on a 1+PI controller for LFC applications in order to maximise efficiency. The chromosome structure of the 1+PI controller is comprised of Kp, Ki, and Kp2 values, and the first population of the region is established. During the health examination, each chromosome is evaluated to determine how it responds to improvements, such as reducing or increasing the amount of exercise. The population is adjusted to the optimal solution through the processes of genetic selection, mutation, and mutation.

During the optimisation process, GA makes adjustments to the 1 + PI controller in order to modify the system frequency (LFC system stability). The evolutionary approach is both efficient and successful when it comes to evaluating high-frequency load FOPID and 1+PI controllers. Through the exploration of parameter space and the subsequent adjustment to the optimal solution, the genetic algorithm is able to locate the control system that optimises both performance and reliability, hence improving the stability and performance of the power grid [15-16].

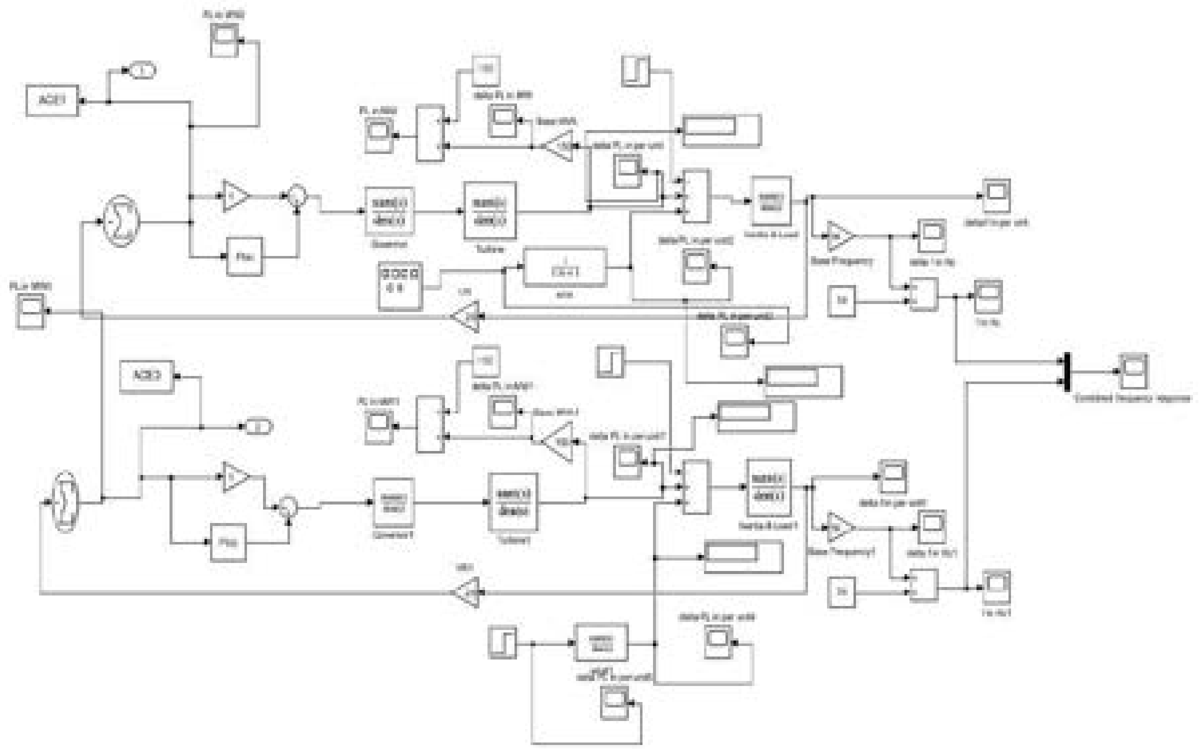


Fig 3. Simple structure of Test System

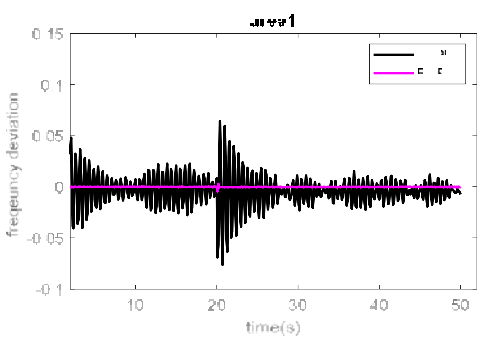
V. Results & Discussion

For the purpose of assessing the effectiveness of the controller that was built for two-area load frequency control (LFC), a number of different scenarios that involve gradual changes in load and wind energy are incorporated. In particular, the step load perturbation in area 1 is designated as 0.1 to 0.3 per unit (pu) at time t = 20 seconds, while the step load perturbation at time t = 2 is denoted as 0.1 to 0.3 per unit (pu). Both of these values are expressed in units of the unit. It is displayed as a time of 0.30 seconds. In addition, a gradual wind power interference is applied to region 2 at a time equivalent to thirty seconds, and the impact range is between 0.1 and 0.15 pu. These effects imitate abrupt variations in load demand and wind generation, which in turn alter power frequency and tie-line power variation. Consequently, these effects have an effect on electric power.

Through the process of subjecting the LFC system to these disturbances, we are able to determine the effectiveness of the controller in terms of controlling the frequency of the system and improve the functionality of the grid. A minor disruption in either of these might result in sudden shifts in energy demand, which necessitates regulators to make prompt adjustments to the power output from the grid in order to maintain a restricted level of system operation. In a similar manner, a voltage source of the second generation will result in changes in the output voltage, in addition to causing issues with the controller's capacity to manage various frequencies.

A. Frequency Deviation

The frequency deviation is a significant indicator of the stability of the system since it measures the amount by which the system frequency deviates from the frequency that is nominal. The controller must make necessary modifications to the frequency difference to ensure that the system frequency remains within the acceptable range in the presence of escalating load disturbances and changes in wind power. The purpose of this action is to adhere to the established regulations. The performance of the controller in maintaining stability within the operational environment can be assessed by observing and assessing the fluctuations that transpire over a specific time frame.



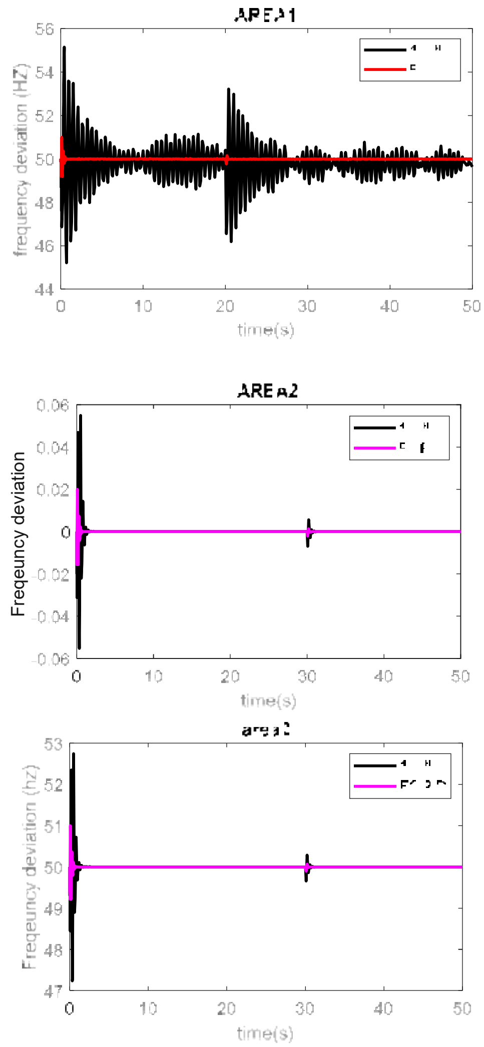
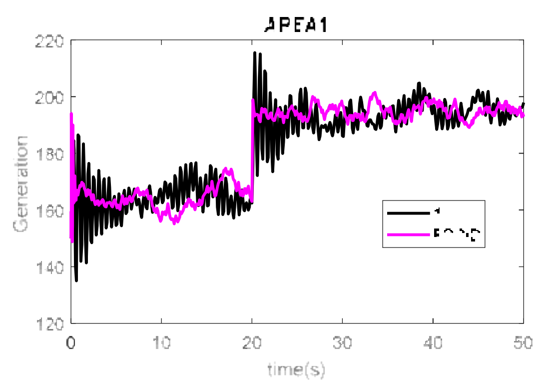


Fig 4. Frequency deviations in both areas.

B. Change in Generation

Change in generation refers to the adjustments made to thermal power plants' power output in response to load changes and variations in wind generation. By monitoring the changes in generation following the disturbances, we can evaluate the controller's effectiveness in restoring the balance between generation and demand, thereby stabilizing the system frequency. A rapid and appropriate response in adjusting generation levels indicates the controller's ability to mitigate frequency deviations and ensure grid stability.



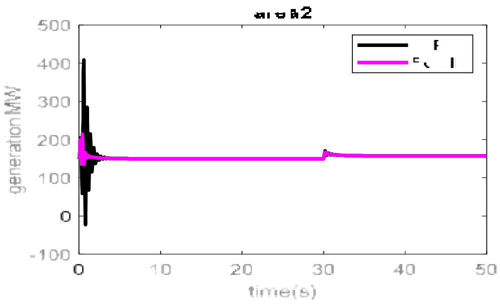


Fig 5. Change in generation in all areas.

C. *Wind Input-Output*

Figure 6 shows the wind input and output for area-1 and area-2.

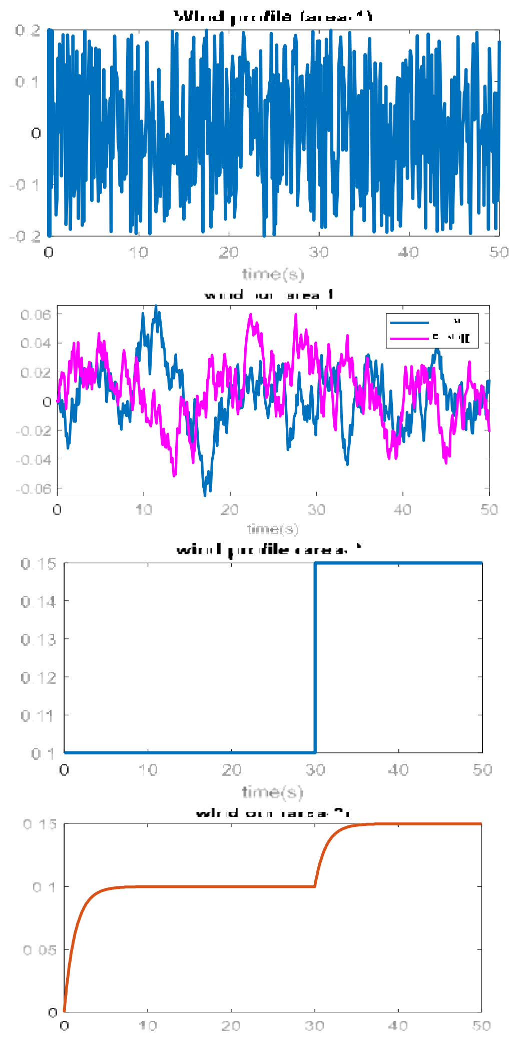


Fig 6. Wind input output for both areas

D. *Area Control Error (ACE)*

Area Control Error (ACE) measures generation-demand imbalance in each control area and is essential for performance evaluation. ACE is the difference between power replacement and replacement plan with load and climate change effects. A low ACE indicates that the controller manages energy exchange in the control zone well and reduces the planned stability-control stability difference.

Compare load step to wind change, generation change, and ACE to evaluate the controller's frequency control and grid stabilization. Comparing the controller's responses and impacts on these crucial parameters can show its real-world LFC efficiency and efficacy, optimizing grid performance.

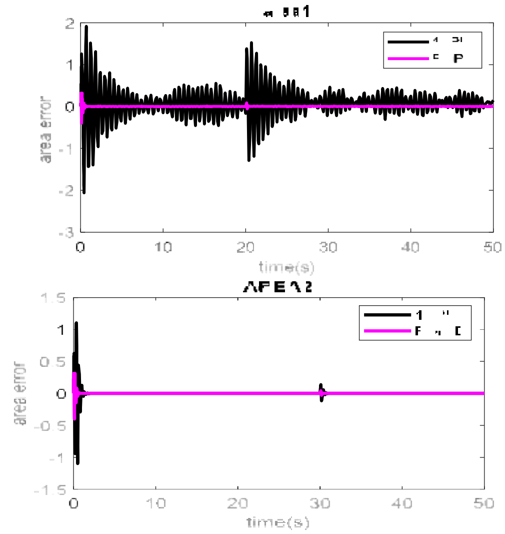


Fig 7. Area Control errors.

VI. Conclusion

The error-related transitions of FOPID and 1 + PI controllers are greater than those of the two-area load frequency control (LFC) controller. In terms of error, power generation, step load interference, and wind switching power regulation, FOPID demonstrates superior performance compared to 1+PI. Utilizing frequency counters to operate and regulate FOPID enhances both efficiency and accuracy. FOPID provides global updates to identify and address defects and damage. The FOPID algorithm surpasses 1+PI in terms of accuracy, stability, and response time.

The FOPID model has exceptional efficacy in identifying unaccounted variations within the power system. The aforementioned accomplishment was attained by employing renewable energy sources and implementing energy storage technology. The effectiveness of FOPID in mitigating power line frequency and power sector impacts surpasses that of wind power generation. FOPID demonstrated superior performance in terms of frequency management and grid stability in both LFC zones when compared to 1 + PI. The utilization of step load impedance and wind power transformer is employed for this purpose. The findings of this study demonstrate that the utilization of FOPID can enhance the efficiency of LFC systems that are impacted by imperceptible issues inside the network.

References

[1]. V. Patil, M. T. R. Khan, M. A. Khan, and S. K. Singh, "Load frequency control of two area power system using fractional order PID controller," International Journal of Electrical Power & Energy Systems, vol. 135, pp. 106497, 2021.

[2]. M. L. K. Patel, H. R. Joshi, and S. K. Bhavsar, "Comparative Study of Load Frequency Control in Interconnected Two Area Power System with Different Controllers," IET Generation, Transmission & Distribution, vol. 15, no. 23, pp. 6083-6092, 2021.

[3]. N. Kumar, M. A. Alotaibi, A. Singh, H. Malik, M. E. Nassar, "Application of FOPID control scheme in AGC of a deregulated power system in the presence of SMES unit", Mathematics, MDPI, *10*(3), pp. 1-16, 06 Feb. 2022.

[4]. N. Kumar, B. Tyagi, V. Kumar, "Deregulated multiarea AGC scheme using BBBC-FOPID controller", Arabian journal of science and Engineering-Springer, vol. 42, pp. 2641-2649, 30 2017. DOI :10.1007/s13369-016-2293-1

[5]. M. A. Khan, A. V. Patil, M. T. R. Khan, and R. N. Patel, "Design and Comparative Analysis of Load Frequency Control for Two Area Power System using Different Controllers," International Journal of Engineering and Advanced Technology (IJEAT), vol. 10, no. 4, pp. 3524-3530, 2021. [DOI: 10.35940/ijeat.D1015.1010421]

[6]. A. Mahata, P. D. Pratihar, and S. Ghoshal, "Comparison of Integral Proportional, Proportional Derivative and Fuzzy-PID Controllers for Load Frequency Control of Interconnected Power Systems," International Journal of Electrical Power & Energy Systems, vol. 131, pp. 107108, 2021. [DOI: 10.1016/j.ijepes.2021.107108]

[7]. N. Kumar, H. Malik, A. Singh, M. A. Alotaibi, M. E. Nassar, "Novel control scheme for load frequency control: A case study of restructured power system" IEEE Access, vol. 9, 06 Dec. 2021, pp. 162231 – 162242, Online ISSN: 2169-3536.

[8]. N. Kumar, B. Tyagi, V. Kumar, "Multi-area deregulated automatic generation control scheme of power system using imperialist competitive algorithm based robust controller", IETE journal of Research-Taylor & Francis, vol. 64, no., 4, pp. 528-537, Sep. 2017.

[9]. N. Kumar, B. Tyagi, V. Kumar, "Multi Area AGC scheme using Imperialist Competition Algorithm in Restructured Power System", Applied soft Computing-Elsevier, vol. 48, pp. 160-168, Nov. 2016.

[10]. N. Kumar, B. Tyagi, V. Kumar, "Optimization of PID parameters using BBBC for multi-area AGC scheme in deregulated power system", Turkish journal of electrical engineering and computer sciences, vol. 24, no. 5, pp. 4105- 4116, June 2016.

[11]. Zhang, Y., Zhou, S., Luo, W., Chen, B., & He, J. (2018). "Optimization of Fractional Order PID Controller Parameters Based on Improved Genetic Algorithm." IEEE Access, 6, 22870-22882.

[12]. J. Dong, & D. Zhai, "Optimal Tuning of FOPID Controller for AVR System Based on Improved Genetic Algorithm", Mathematical Problems in Engineering, 2016, 1-12, 2016

[13]. Salam, S. I. A., Ahmed, F. M., & Abdelaziz, A. Y. (2020). "Optimal Design of Fractional Order PID Controller Using Genetic Algorithms for Automatic Generation Control of Multi-Area Power Systems." International Journal of Electrical Power & Energy Systems, 118, 105805.

[14]. Mirsalim, M., & Shahgholian, G. (2019). "Design and Optimization of a New Proportional-Integral-Derivative Controller Using Genetic Algorithm." Electrical Engineering, 101(3), 989-1000.

[15]. Mancilla, M. D. A., Aguirre, J., & Salas, C. (2019). "Parameter Tuning of a Power System Stabilizer using Genetic Algorithm with PI and 1+PI Controllers." Journal of Electrical Systems and Information Technology, 6(2), 331-341.

[16]. N. Kumar, A. Singh, B. P. Joshi, and B. K. Singh, "Intelligent Adaptive Critic Scheme Implementation and Investigation Using Policy Iteration Technique: A Case Study of Multi-area Automatic Generation Control (AGC) Problem," in Lecture Notes in Electrical Engineering, vol. 802, 2022, pp. 159–183.

[17]. A. Singh, B. P. Joshi and B. K. Singh, "Linear Quadratic Control Scheme for Isolatedand Inter-connected Power System," 2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE), Greater Noida, pp. 1913-1916,2023.

[18]. B. P. Joshi and A. Singh, "Multi-Criteria Decision-Making Approach Based on Moderator Intuitionistic Fuzzy Hybrid Aggregation Operators," pp. 237–251, 2018, doi: 10.4018/978-1-5225-5709-8.ch011.

[19]. B. P. Joshi, M. Mohan Sati, A. Rayal, A. Kumar, and N. Kumar, "An Approach for Renewable-Energy-Source-Selection Problems Based on Choquet Integral Operator of MIFS," in ICACITE, May 2023, pp. 643–648. doi: 10.1109/ICACITE57410.2023.10182907.

[20]. B. P. Joshi, N. Joshi, S. Oli, A. Rayal, A. Kumar, and A. Singh, "MIFS Ordered Weighted Operators method for," 2023 IEEE 2nd Int. Conf. Ind. Electron. Dev. Appl., pp. 248–253, 2023.

[21]. Kumar, N., Rathee, S., Singh, A., Halder, S., Agrawal, A. (2023). Frequency Control Using Captive Generation and Demand Response. In: Maurya, S., Peddoju, S.K., Ahmad, B., Chihi, I. (eds) Cyber Technologies and Emerging Sciences. Lecture Notes in Networks and Systems, vol 467. Springer, Singapore.

[22]. J. Singh et al., "IR Sensor Based Accident Prevention System for Hilly Areas," 2023 International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2023, pp. 786-789.

[23]. Kumar, N., Lata, S., Singh, A., Rathee, S., Sangwan, P. (2023). Load Frequency Control in Multiarea Power System Using TID Control Scheme. In: Maurya, S., Peddoju, S.K., Ahmad, B., Chihi, I. (eds) Cyber Technologies and Emerging Sciences. Lecture Notes in Networks and Systems, vol 467. Springer, Singapore.

[24]. K. Sharma, N. Kumar, P. Datta, J. Singh and A. Verma, "A Research Paper on Frozone: An Autonomous Fire fighter," *2023 International Conference on Disruptive Technologies (ICDT)*, Greater Noida, India, 2023, pp. 798-800, doi: 10.1109/ICDT57929.2023.10150988.

[25]. M. M. Sati, B. P. Joshi, T. Pal, N. Kumar, A. Singh and S. Goyal, "Ambiguous Fuzzy Einstein Geometric Operator: Utilizing to Analyze Power Generation Techniques," 2024 2nd International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2024, pp. 1536-1541,

[26]. N. Kumar, B. Tyagi, V. Kumar, "Multiarea automatic generation control using BBBC tuned several classical controllers in deregulated environment", in 1st UPCON-IEEE conference on Energy Economics and Environment (ICEEE), Greater Noida, India, pp.1-6, 27-28 March 2015. ISSN: 9781467374927

[27]. J. Singh et al., "Power Saving and Power Generating in Automobile (Self-Charging Kit)," 2023 International Conference on Disruptive Technologies (ICDT), Greater Noida, India, 2023, pp. 801-805, doi: 10.1109/ICDT57929.2023.10150450.

[28]. N. Kumar, B. Tyagi, V. Kumar, "Multiarea automatic generation control using model predictive based control in deregulated environment", in 1st UPCON-IEEE conference on Energy Economics and Environment (ICEEE), Greater Noida, India, pp.1-6, 27-28 March 2015. ISSN: 9781467374927.

[29]. A. Kumar, B. P. Joshi, A. Sagar, N. Kumar, M. Bertoluzzo, and A. Singh, "Opportunities and Challenges for Electric Vehicle Wireless Charging with Home," in 2023 International Conference on Power Energy, Environment & Intelligent Control (PEEIC), Dec. 2024, pp. 221–225. doi: 10.1109/peeic59336.2023.10451967.

[30]. R. Kamra, B. P. Joshi, P. Sinha, A. Singh, and S. Rathee, "Sustainability of Renewable-Energy-Sources-Classification utilizing Moderator-Intuitionistic-Fuzzy Hybrid Averaging Operator," in 2024 Fourth International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), Jan. 2024, pp. 1–5.

[31]. N. Singh, B. P. Joshi, S. Gupta, A. Singh, and S. Rathee, "Renewable-Energy-System Applications of Ambiguous-Fuzzy-Hybrid-Averaging Operator," in 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS), Feb. 2024, pp. 1–6. doi: 10.1109/ICICACS60521.2024.10498184.

[32]. S. Oli, M. Manu, P. S. Kharayat, A. Singh, A. Rayal, and B. P. Joshi, "Real Time Railway Security System using Sensor Network," in 2023 2nd International Conference on Futuristic Technologies (INCOFT), Nov. 2023, pp. 1–4. doi: 10.1109/INCOFT60753.2023.10425570.

[33]. B. P. Joshi, "Some generalized intuitionistic fuzzy geometric aggregation operators with applications in multi-criteria decision making process," Intelligent Systems: Concepts, Methodologies, Tools, and Applications. pp. 1190–1211, 2018. doi: 10.4018/978-1-5225-5643-5.ch050.

[34]. A. Sagar, S. Padmanaban, M. Bertoluzzo, A. Kumar, B. P. Joshi, and A. Singh, "A Self-Improved Optimization-Based Artificial Neural Network Model for WPT System for Electric Vehicle Charging," in 2023 International Conference on Power Energy, Environment & Intelligent Control (PEEIC), Dec. 2024, pp. 156–163. doi: 10.1109/peeic59336.2023.10450276.

Highway-Powered Electric Vehicle Charging Stations using IoT

*

J. N.V.R. Swarup Kumar
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
sjavvadi2@gitam.edu

V. Sai Deekshitha
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
vdeekshitha24@gmail.com

A. Uday Keerthi
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
keerthiattaluri04@gmail.com

T. Tarun Sai
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
tadelatarunsai@gmail.com

M. Sai Bhavan
Dept. of CSE, GST
GITAM Deemed to be University
Visakhapatnam, India
bhavanmeesala07@gmail.com

M. Ramu
Dept. of EECE, GST
GITAM Deemed to be University
Visakhapatnam, India
rmutyala@gitam.edu

***Abstract*—The global transportation landscape is experiencing a profound shift with the widespread adoption of electric vehicles (EVs) as a sustainable alternative to traditional combustion engine vehicles. In light of the growing EV market, this research project proposes establishing a sustainable EV charging infrastructure along highways, utilizing power directly sourced from existing electric poles lining the roadside. The primary objective is integrating advanced payment technologies, allowing individuals to conveniently pay for consumed power using UPI or credit/debit cards. This technological integration aims to create a user-friendly and easily accessible payment system. This innovative approach will benefit users and the government since it has low development costs. IoT sensors monitor the real-time situation of charging stations, taking measures for errors. Users log on and get up-to-the-minute charging status notices and billing details. Support for this can come out after successful testing of Cloud computing. So not only will it be smooth and convenient to operate the system via EV charging infrastructure, but it will also make some accurate contributions towards sustainable transport initiatives.**

Keywords: Electric Vehicle, Accessibility, User Experience, Efficient Charging Station, Payment Integration.

I. Introduction

In the automotive industry, the speed-up adoption of electric vehicles is a revolution. It helps improve the cleanliness of transportation. However, a significant challenge that restricts the extensive use of EVs is the lack of infrastructure to load their batteries on long journeys. By recognizing the necessity of a full-scale and readily available charging infrastructure for electric vehicles [1], our project puts forward an innovative concept. Directly utilizing power from existing electric poles along highways to build a network of electric vehicle charging stations [4].

The idea aims to provide convenient charging for guys on long trips in an electric car, building on advanced payment technologies to offer an enjoyable and easy experience. Drawing from the existing electric poles, the charging stations can take power directly from the source. This reduces the need for additional infrastructure and leverages work already done using infrastructure to get maximum efficiency out of it. This not only simplifies the charging process but also helps meet the sustainability goals of the electric vehicle ecosystem. Putting efficient charging stations into place at crucial points along the highways begins to resolve the practical problems of EV adoption, particularly addressing issues of range anxiety during long-distance travel[16]. Broadening the EV equipment infrastructure is faced with prompt challenges such as high implementation costs, scale-ability problems, and the need for sustainable power sources.

The project aims to meet these challenges by using electric poles along highways to minimize additional costs for a more sustainable charging solution and thus render electric vehicles a reality for more people [11]. Moreover, the project strongly emphasizes user experience by integrating seamless payment technologies, allowing users to track charging sessions receiving payment notifications. This approach not only simplifies the payment process but also enhances the overall accessibility and user-friendliness of the EV charging network. This idea contributes to sustainable transportation by presenting an innovative approach that leverages existing infrastructure to establish an extensive EV charging network. By focusing on user experience and environmental impact, the proposed highway-powered charging stations aspire to play a pivotal role in shaping the future of transportation, fostering economic growth, and laying the foundation for intelligent mobility ecosystems.

A. *Basic Terms of electric vehicle charging stations*

Distinction between different charging levels: Level 1 (120V), Level 2 (240V), and DC fast charging, catering to various charging needs.

Accessibility and Location: Strategic placement of charging stations for accessibility, considering urban areas, highways, and public spaces. Accessibility features, such as ADA compliance and user-friendly interfaces[6].

Connectors and Standards: Specification of charging connectors, such as CCS, CHAdeMO, or Type 2, to ensure compatibility with diverse electric vehicle models.

Charging Infrastructure: Physical infrastructure for charging stations, including charging bays, power distribution units, and associated electrical components.

Power Output: Specification of the power output of the charging station, measured in kilowatts (kW), impacting the charging speed.

B. *How Electric Vehicles Charge*

Electric vehicles (EVs) utilize three main types of charging systems: Level 1, Level 2, and DC fast charging. Fig.1. depicts the modes for charging electric vehicles in the proposed idea. Level 1 charging operates at a power output of 1.4 to 1.9 kW, suitable for daily commuting and overnight residential charging using a standard 120-volt household outlet. Level 2 charging offers a power output of 3.3 to 19.2 kW, catering to diverse needs in homes, workplaces, and public charging stations. Operating at 240 volts, this configuration provides faster charging than Level 1 and ensures compatibility with prevalent infrastructure[2]. Level 3, known as DC Fast charging, exceeds the previous two levels; it can provide at least 50 kW and up to 350 or 400 kW-power, perfect for long-distance travel in 2022 with swift highway refueling. The "charging time of 20 minutes" on Level 3 shows that it still has quick charging aimed at owners of electric automobiles who are hustling between business meetings [8]. At the same time, the charging infrastructure includes connectors to match EV charging ports, with variants such as J1772, type-2, CCS, and ChaDeMO for different regions and standards.

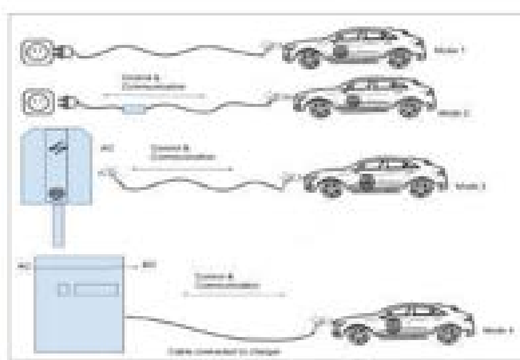


Fig. 1. EV charging modes

II. Literature Review

The evolution of electric vehicle (EV) charging stations has seen significant innovations aimed at improving their function and making them friendlier to the environment. Despite improvements as station technology evolves, there is still a wide variety of features and limitations on available charging stations in development[10].

But this proposal has a banner departure and uses the existing infrastructure–highway poles for EV recharging. This strategy reduces costs and space use, marking significant progress in sustainable infrastructure deployment [5],[6].

A key point of this proposal lies in technology integration. It lets us get transparent insights into power consumption and establishes a user-friendly payment system. Unlike many models that find it difficult to achieve seamless technology integration, this proposal has an automatic charge process, which is friendlier to users and removes the need for external oversight. Therefore, it is also economic in terms of external monitoring and user satisfaction [14].

The proposal is nothing like existing models that only adapt to essential charging needs. The main focus of this model is to design with advanced technology, especially internet-of-things (IoT) solutions. Taking advantage of IoT, this proposal has a seamless payment system, precise control over the flow of current, and can be monitored at leisure. This systematic approach to integration is expected to streamline energy consumption, raise security levels, and facilitate proactive maintenance, the very change in EV charging infrastructure [3].

By highlighting the advantages of leveraging existing infrastructure, incorporating advanced payment technologies, and integrating IoT systems, this proposal positions itself as a leading solution for the future of sustainable EV charging infrastructure. Its comprehensive approach and valuable insights make it a valuable contribution to the literature review of research papers in this field.

III. EV Charging Scheme through electric poles

The proposal to use existing electric poles for electric vehicle (EV) charging infrastructure involves addressing the challenge of the inherent alternating current (AC) in electric poles[15]. At the same time, EV batteries operate on direct current (DC). A conversion process becomes essential as EV batteries rely on unidirectional current flow for chemical reactions that facilitate energy storage.

The AC to DC conversion involves a transformer, adjusting AC voltage to a suitable level for further processing, and a rectifier, the core component employing diodes to allow one-directional current flow, effectively converting AC to DC[17]. Smoothing filters, typically capacitors and inductors, eliminate pulsations in the rectified output, providing a cleaner DC output for efficient battery charging. Additionally, a voltage regulator circuit may be employed to maintain the DC output at a desired level, ensuring safe and optimal charging conditions for the EV battery. When selecting a rectifier, considerations include voltage and current ratings to match the EV model's requirements, efficiency to maximize charging efficiency, switching frequency to balance component size and electromagnetic interference, and cost versus reliability for an optimal rectifier solution. The figure 2 depicts how alternating current (AC) transforms direct current (DC). Notably, electric vehicles exclusively accommodate direct current for their operational requirements.

Fig. 2. Alternating current to Direct Current conversion

IV. Integrating Smart Payment Technologies and IoT for Seamless Payment

The array of online payment technologies for electric vehicle (EV) charging includes mobile payment apps, enabling users to locate and initiate sessions and make payments seamlessly through smartphones[13]. RFID cards or tags, utilizing Radio-Frequency Identification technology, offer quick and secure contactless transactions at charging stations[12]. Equipping charging stations with card readers for credit/debit card payments provides a widely accepted and familiar payment method. Subscription services offer a predictable revenue stream, where users pay regular fees for charging station access, potentially with added benefits or discounted rates, fostering increased user loyalty.

A. *Preferred Technology:*

The preferred software technology for implementing this idea involves using a self-service. Through this, individuals can access electricity to charge their vehicles.

1) The user initiates the process upon opening the self-service screen, as depicted in Fig 3.
2) Upon accessing the screen, users enter the preferred quantity of electricity units they intend to utilize, as illustrated in Fig 4.
3) Upon inputting the desired units, the user advances by selecting the "Next" button. Transitioning to another page Fig 5. A payment summary is presented on this subsequent page, detailing the total selected units and the corresponding amount due.
4) Progressing further, the user navigates to the checkout page, where they are prompted to choose their preferred payment method Fig6.
5) The available payment options encompass credit/debit cards, UPI, and net banking. If the user opts for card payment, they must enter their card details to generate an OTP, enabling them to proceed with the payment, as depicted in Fig 7,8.
6) If the user chooses UPI payment, they must enter their UPI ID or scan the QRCode and complete the payment, as illustrated in Fig 9,10.
7) As depicted in Fig 11, the user must enter their mobile number to get a payment receipt.
8) With confirmed payment information, users can now use electricity and charge their vehicles as in Fig 12.

Fig. 3. Payment System Home Page

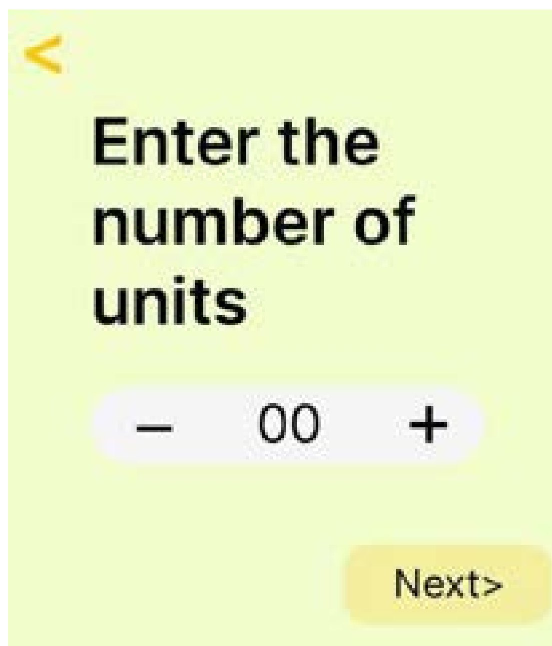


Fig. 4. Enter Units

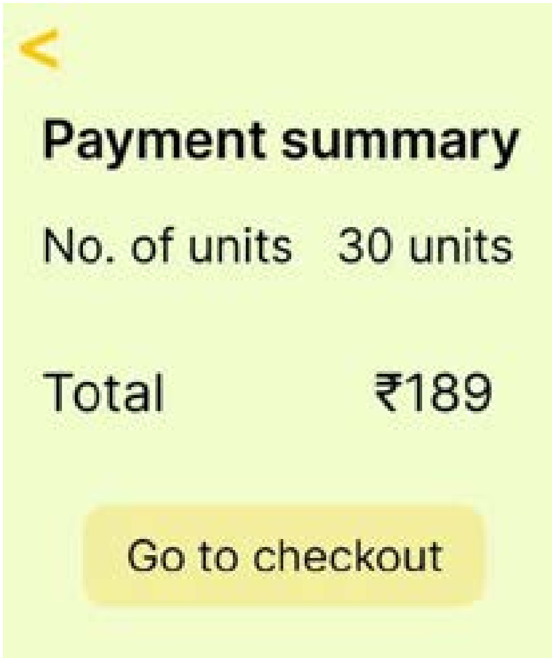


Fig. 5. Payment Summary

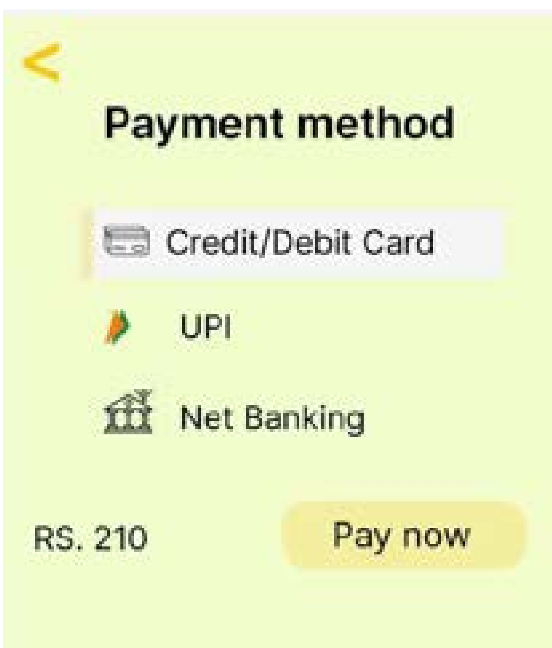


Fig. 6. Select Payment Method

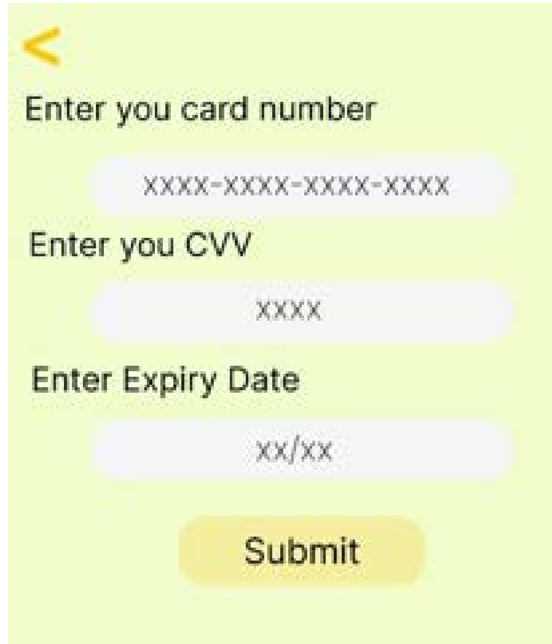


Fig. 7. Enter Card Details

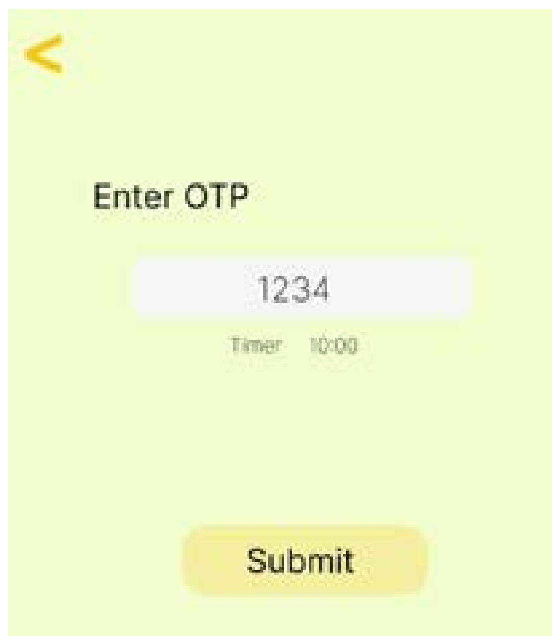


Fig. 8. Enter OTP

Fig. 9. Other Payment Details

Fig. 11. Payment Confirmation

Fig. 10. Other Payment Details

Fig. 12. Charged Information

B. Integration of IoT:

The Internet of Things (IoT) plays a multifaceted and critical role in enhancing the functionality of electric vehicle (EV) charging stations. The critical aspects of IoT's contribution to EV charging stations:

Real-time Monitoring and Data Collection: IoT sensors in EV charging stations constantly monitor vital parameters such as availability, power levels, and technical issues. This real-time data collection optimizes system performance, aids in maintenance planning, and keeps users informed about station status.

Remote Management and Maintenance: IoT connectivity allows remote management of charging stations, including diagnostics, updates, and maintenance tasks without physical intervention. This reduces downtime, boosts efficiency, and enables proactive maintenance to prevent issues.[7]

User Experience Enhancement: IoT-enabled systems provide automated notifications to users, giving real-time updates on charging status, completion, and payment. This enhances the user experience by keeping them informed and streamlining the charging process.

Dynamic Current Control: IoT-enabled charging stations adjust current flow based on user preferences and real-time electricity prices. This ensures efficient electricity use, faster charging, and reduced grid strain, especially during peak demand.[9]

V. Revenue considerations for government

Government revenue considerations for EV charging infrastructure include installation costs, technology integration expenses for advanced payment systems and IoT connectivity, ongoing operational and maintenance costs, potential government incentives such as financial support or tax breaks, collaborations, and partnerships with private entities. In Andhra Pradesh, approximately 243 charging stations are needed along the national highway, each costing around 325,000 rupees. The total budget allocation required for this infrastructure project is estimated at 78,975,000 rupees.

VI. Results

The idea is to minimize implementation costs by tapping into the established electric poles, reducing the need for additional infrastructure, and promoting widespread deployment of charging stations. In Andhra Pradesh, where there are a total of 67,000 electric vehicles, with approximately 30,000 traveling on national highways (comprising 10,000 electric scooters and 20,000 electric cars), Assuming complete charging, electric scooters generate a daily revenue of 2,30,000. In contrast, electric vehicles contribute 39,00,000, resulting in a total daily government revenue of 41,30,000 rupees. The proposed approach significantly focuses on integrating advanced payment technologies to create a seamless and user-friendly experience for electric vehicle (EV) users. Through streamlined payment processes, real-time tracking, and payment notifications, the aim is to enhance the accessibility and user-friendliness of the EV charging network. Note: It is essential to highlight that the information presented here is based on assumptions and may not reflect accurate data. The figures provided are hypothetical and intended for illustrative purposes only. Actual statistics regarding the number of electric vehicles, their distribution, and the corresponding revenue generated should be obtained through reliable sources for a precise understanding of the situation.

Conclusion

Creating electric vehicle charging stations using existing highway electric poles is a clever, Eco-friendly, and cost-effective way to encourage more people to use electric cars. This project aims to speed up the switch to sustainable transportation by using the existing infrastructure to provide convenient and fast charging for long-distance travelers. The advantages are many; this plan makes it easier to set up much-needed charging stations for electric cars using what we already have and ensures better access and a more user-friendly experience. Innovative payment systems and advanced technology allow efficient and automated payment processing and infrastructure monitoring. For governments wanting more people to use electric cars, this project supports sustainable development and creates potential ways to make money that can pay for itself. With expected growth in electric car sales, the future looks promising. By taking on this forward-looking project, we can make significant progress towards a clean and energy-efficient transportation future. This new idea for EV charging on highways combines the current infrastructure, new technologies, sustainability goals, and critical economic factors. With a focus on environmental awareness and intelligent mobility solutions, this proposal represents a positive direction for the future.

ACKNOWLEDGMENTS

The authors are very thankful to the management of GITAM (Deemed to be a university), who provided us with time to frame this work.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

REFERENCES

[1] P. Bauer. Future of ev charging. In *2019 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) 2019 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM)*, pages 8–9, 2019.

[2] M. Brenna, F. Foiadelli, C. Leone, et al. Electric vehicles charging technology review and optimal size estimation. *J. Electr. Eng. Technol.*, 15:2539–2552, 2020.

[3] N. Chanthong, T. Ruangsakorn, and S. Glomglome. Blockchain and smart contract payment for electric vehicle charging. In *2020 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 161–164, 2020.

[4] M. K. Das and S. K. Jain. Review on optimization techniques used for scheduling of electric vehicle charging. In *2021 International Conference on Control, Automation, Power and Signal Processing (CAPS)*, pages 1–6, 2021.

[5] W. B. Diab, R. A. Hamdan, A. Koubayssi, M. Arnaout, and R. Rammal. Integration of electric vehicles in smart grids. In *2021 International Conference on Microelectronics (ICM)*, pages 246–249, 2021.

[6] A. A. Galadima, T. Aja Zarma, and M. A. Aminu. Review on optimal siting of electric vehicle charging infrastructure. *Journal of Scientific Research and Reports*, 25(1):1–10, Oct. 2019.

[7] G. F. Savari, V. Krishnasamy, J. Sathik, Z. M. Ali, and S. H. E. Abdel Aleem. Internet of things based real-time electric vehicle load forecasting and charging station recommendation. *ISA transactions*, 97:431–447, 2020.

[8] J. Sears, D. Roberts, and K. Glitman. A comparison of electric vehicle level 1 and level 2 charging efficiency. In *2014 IEEE Conference on Technologies for Sustainability (SusTech)*, pages 255–258, 2014.

[9] S. Sikdar and M. Damle. Iot solutions for electric vehicles (ev) charging stations: A driving force towards ev mass adoption. In *2022 International Interdisciplinary Humanitarian Conference for Sustainability (IIHC)*, pages 114–120, 2022.

[10] P. P. Singh, F. Wen, I. Palu, S. Sachan, and S. Deb. Electric vehicles charging infrastructure demand and deployment: Challenges and solutions. *Energies*, 2022.

[11] E. M. Szumska. Electric vehicle charging infrastructure along highways in the eu. *Energies*, 16(2), 2023.

[12] A. A. Talib and A. D. Salman. Development of an electronic payment system using the internet of things. *IAES International Journal of Robotics and Automation (IJRA)*, 2022.

[13] T. V. Tomić, N. The future of payments in the internet of things. In *Sinteza 2017 - International Scientific Conference on Information Technology and Data Related Research*, pages 97–104, 2017.

[14] A. Ul-Haq, C. Buccella, C. Cecati, and H. A. Khalid. Smart charging infrastructure for electric vehicles. In *2013 International Conference on Clean Electrical Power (ICCEP)*, pages 163–169, 2013.

[15] E. Werthmann and V. Kothari. Pole-mounted electric vehicle charging: Preliminary guidance for a low-cost and more accessible public charging solution for u.s. cities. *World Resources Institute*, 2021.

[16] A. K. M. Yousuf, Z. Wang, R. Paranjape, and Y. Tang. Electric vehicle charging station infrastructure: A comprehensive review of technologies, challenges, and mitigation strategies. In *2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 588–592, 2023.

[17] S.-Y. Yun, H.-J. Chae, W.-Y. Kim, H.-T. Moon, Y.-S. Jeong, and J.-Y. Lee. Battery charger for ev. *THE TRANSACTIONS OF KOREAN INSTITUTE OF POWER ELECTRONICS*, 15(6):460–465, 2010.

Multi Classification of Pest using Transfer Learning

1st Shaik Abdul Muthallib
School of Computing
SRM Institute of Science and Technology
Chennai, Tamil Nadu
sm2769@srmist.edu.in

2nd Kotha Narasimha Rao
School of Computing
SRM Institute of Science and Technology
Chennai, Tamil Nadu
kr2467@srmist.edu.in

3rd Dr.A.Shobanadevi
School of Computing
SRM Institute of Science and Technology
Chennai, Tamil Nadu
shobanaa3@srmist.edu.in

***Abstract*—The integration of machine learning and deep learning algorithms such as K-means, SVM, ResNet, and VGGNet signifies a significant advancement in agricultural development. Accurate and precise pest identification is crucial for farmers facing challenges in pest management. This paper adopts the MobileNet architecture to enhance pest classification accuracy. Leveraging a dataset containing 19 distinct pest classes enables the algorithm to train across diverse pest categories, facilitating efficient classification. By incorporating MobileNet, the model improves its ability to accurately identify pests, thereby providing farmers with a valuable tool for effective pest management and crop protection.**

***Index Terms*—Mobilenet, Relu, Maxpooling 2D, Flatten, Sequential, Image data generator.**

I. Introduction

Agriculture forms the backbone of India's economy, with over 50 percent of its population employed in this sector. Farmers employ various techniques to manage pests across a diverse range of crops. However, identifying these pests can be a time-consuming task prone to inaccuracies through traditional observation methods. Pests prey on crops, causing substantial damage and losses to farmers. The accurate identification of pests is paramount for effective pest management and ensuring crop yield and quality.

To address the challenge, this research utilizes deep neural networks, specifically MobileNet, comprising 27 layers. MobileNet is pretrained on a dataset encompassing 19 distinct classes of pests commonly found in Indian agriculture, including ants, bees, caterpillars, aphids, and others. The transfer learning approach with MobileNet achieves an impressive accuracy rate of up to 93.24 percent in pest classification. Furthermore, the approach involves fine-tuning the pretrained MobileNet model to adapt to the specific characteristics of pest images, enhancing its accuracy in classification. Through extensive experimentation and validation, MobileNet's effectiveness in accurately identifying pests across various crops is demonstrated, offering a practical solution for mitigating pest-related risks and optimizing agricultural productivity.

This paper presents a comprehensive solution aimed at revolutionizing pest identification in agriculture. By leveraging deep learning techniques, Comprehensive training and testing procedures were conducted on the dataset, divided into 80 percent for training and 20 percent for testing. The research centers on pests prevalent in Indian regions, offering valuable insights into pest classification and assisting agricultural practices in effectively mitigating crop damage.

II. RELATED WORKS

In recent years, various studies have delved into leveraging deep learning and machine learning models for the identification and classification of plant diseases and pests using image- based approaches. Researchers such as Xin Li have explored the application of ResNet models for Apple Leaf Disease Identification and Classification [1], while others like Himanshu Singh and colleagues have investigated Tomato Crop Disease Classification using Convolutional Neural Networks (CNNs) and transfer learning techniques [2]. Additionally, studies by Parveen Malik and Xiedong Song have delved into Insect Classification and Apple Disease Detection respectively, employing pre-trained deep neural networks and transfer learning methods to enhance accuracy in disease identification [8].

Moreover, amidst the diverse array of methods employed, researchers have encountered challenges associated with computational efficiency and model robustness. While approaches utilizing ResNet50 have shown promise in disease classifi- cation, they often come with significant computational costs. Conversely, machine learning models like K-means, as pro- posed by Yanan Chen and collaborators [16], face issues with sensitivity to outliers. In response to these challenges, S Prathab and colleagues have introduced a novel approach focusing on Pest Detection and Pesticides Recommendation using MobileNet [7]. Despite the introduction of the Mo- bileNet model by S. Prathab and colleagues, its accuracy falls short. The identification process involves utilizing a dataset containing 19 classes to accurately classify various pest types. By incorporating more data and training cycles, this approach aims to improve the model's performance in accurately classifying various pest classes.

In the realm of disease and pest classification, researchers encounter hurdles related to computational efficiency and model robustness. While models like ResNet50 hold potential, they demand substantial computational resources. Conversely, machine learning techniques like K-means, as advocated by Yanan Chen and colleagues [16], provide alternatives but can be influenced by outliers. This paper aims to tackle these challenges by introducing a novel approach that enhances efficiency and facilitates training across various pest classes.

While previous papers may have addressed pest identification to some extent, many have limitations. Some focus on specific crops or leaves, neglecting edge cases. Others rely on machine learning algorithms like SVM and K-means, which can be sensitive to outliers. Additionally, models like ResNet50, while effective, are computationally expensive. Conversely, MobileNet, though efficient, may sacrifice some accuracy. This paper aims to introduce an improved approach that enhances efficiency while maintaining accuracy across various pest classes.

III. PROPOSED METHODOLOGY

A. DATA SET

The dataset utilized in the research comprises a combination of two datasets sourced from Kaggle, encompassing 19 distinct classes of pests commonly found in agricultural environments. These classes include snails, ants, wasps, caterpillars, weevils, grasshoppers, mosquitoes, sawflies, mites, beetles, armyworms, earwigs, bees, bollworms, stem borers, aphids, moths, slugs, and earthworms. The MobileNet model is leveraged for the research, with comprehensive training and testing procedures conducted on the dataset. The dataset is divided into 80 percent for training and 20 percent for testing. The research focuses on pests commonly found in Indian regions, providing valuable insights into pest classification and assisting agricultural practices in effectively mitigating crop damage.

B. Algorithm

MobileNet stands out as a highly efficient and adaptable convolutional neural network (CNN) architecture widely deployed in real-world applications [3]. Its efficiency is notably achieved through the utilization of depthwise separable convolutions, a departure from traditional convolutional methods, enabling the construction of lighter models. Additionally, MobileNets introduce novel global hyperparameters, such as width multiplier and resolution multiplier, offering developers the flexibility to balance between latency, accuracy, speed, and size according to specific project requirements.

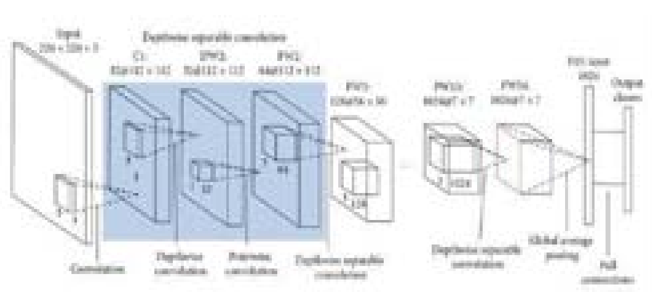


Fig. 1. Mobilenet layer

Built upon depthwise separable convolution layers as shown in figure 1, MobileNets comprise a series of depthwise and pointwise convolutions, totaling 28 layers when considering them separately. Notably, a standard MobileNet encompasses approximately 4.2 million parameters [17], a figure that can be further reduced by adjusting the width multiplier hyperparameter accordingly. This flexibility in parameter tuning allows for optimization tailored to the constraints of individual applications, contributing to its widespread adoption in various domains [4].

In training MobileNet for specific tasks, such as pest identification across 19 distinct classes, a standardized input image size of 224 × 224 × 3 is typically utilized. The training process often involves multiple epochs, with a typical configuration employing 50 epochs, and a batch size of 5 is commonly employed across both training and testing datasets [5]. These parameters, combined with the inherent efficiency and adaptability of MobileNet architecture, contribute to its effectiveness in addressing complex image classification tasks, including pest identification in agricultural contexts [6].

IV. SYSTEM ARCHITECTURE

In the research, a customized MobileNet architecture as shown in fig 2 was employed to classify 19 distinct pest classes commonly found in agricultural settings. This specialized architecture aimed to optimize classification accuracy while considering the computational efficiency required for real-world deployment [5]. Leveraging MobileNet's lightweight design and depthwise separable convolutions, the goal was to strike a balance between model complexity and performance, ensuring efficient handling of the classification task across various crops and pest species. The MobileNet approach utilizes existing datasets to boost pest classification accuracy. Through transfer learning techniques, knowledge is acquired, facilitating precise classification across diverse pest species and environmental conditions. This methodology enhances accuracy by leveraging pre- existing data, refining the model's understanding for improved classification outcomes.

Incorporating both Adam and Stochastic Gradient Descent (SGD) optimizers into the architecture enhances the learning process [12]. These optimizers are pivotal in updating each parameter of the model with an appropriate learning rate, facilitating convergence towards optimal solutions during training [5]. Specifically, the Adam optimizer is responsible for minimizing the loss function at each step, contributing to the overall effectiveness of the learning process. The combination of these optimizers with MobileNet's architecture aims to maximize efficiency and effectiveness while reducing the risk of overfitting [15].

Alongside the optimization techniques, the architecture included Rectified Linear Unit (ReLU) activation layers [6].ReLU, known as Rectified Linear Unit, is favored in deep learning networks for its computational simplicity and ability to introduce representational sparsity. Its implementation requires only a max() function, ensuring efficient computations during training. ReLU outputs true zero values for negative inputs, contributing to sparsity in the network's activations. This sparsity can aid in regularization, potentially mitigating overfitting. Despite its effectiveness, ReLU may encounter the ”dying ReLU” problem, where neurons become inactive

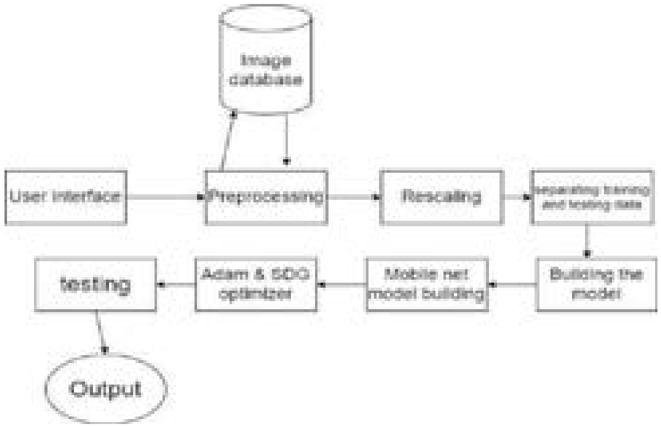


Fig. 2. Architecture Diagram

during training. These activation functions are essential for enabling the network to learn complex patterns and features from the input data, thereby improving its ability to extract relevant information for accurate classification. By adopting a transfer learning approach and selectively freezing the top layers of MobileNet [14], The emphasis lay in fine-tuning the lower layers to adjust the model to the specific characteristics of the classification task. The strategic approach ensures that the architecture efficiently learns and classifies diverse pest species while effectively managing computational resources.

The MobileNet model utilized comprises 28 layers and has been trained on 4.2 million parameters, encompassing various image types. It excels in handling disparate and low-clarity images, maintaining a superior performance across diverse visual inputs. This capability is attributed to its extensive parameter coverage and architectural design, which effectively captures and processes features across different image characteristics.

V. PROCEDURE

Constructing the neural network architecture begins with utilizing the pre-trained MobileNet model, trained on the ImageNet dataset. Importing this model enables the utilization of its ability to capture fundamental image features, crucial for subsequent classification tasks [7]. MobileNet's pre-trained weights empower the model to recognize a broad spectrum of visual patterns, providing a robust foundation for accurate classification. However, as the input image size does not match MobileNet's specifications, the decision is made to freeze its top layers [9]. This strategic choice ensures that the model retains its capability to extract meaningful features while adapting to the specific input dimensions

After integrating the MobileNet base, a MaxPooling2D layer with a pool size of 3x3 and a stride of 2 is introduced. [10].This operation downsamples input by computing the maximum value within each channel over an input window defined by poolsize, shifting the window by strides along spatial dimensions. It effectively reduces input dimensions, capturing salient information crucial for subsequent layers' learning processes. This downsampling aids in preserving

Model: "sequential"

Layer (type)	Output Shape	Param #
mobilenet_1.00_224 (Functional)	(None, 7, 7, 1024)	3228864
max_pooling2d (MaxPooling2D)	(None, 3, 3, 1024)	0
flatten (Flatten)	(None, 9216)	0
dense (Dense)	(None, 128)	1179776
batch_normalization (BatchNormalization)	(None, 128)	512
dense_1 (Dense)	(None, 1024)	132096
batch_normalization_1 (BatchNormalization)	(None, 1024)	4096
dense_2 (Dense)	(None, 512)	524800
batch_normalization_2 (BatchNormalization)	(None, 512)	2048
dense_3 (Dense)	(None, 19)	9747

Total params: 5081939 (19.39 MB)
Trainable params: 5056723 (19.29 MB)
Non-trainable params: 25216 (98.50 KB)

Fig. 3. Mobilenet Model Summary

relevant features while decreasing computational complexity, enhancing overall model efficiency. Subsequently, Incorporating a Flatten layer reshapes the output from preceding layers into a 1D vector [13]. This transformation prepares the data for input into fully connected Dense layers, where Rectified Linear Unit (ReLU) activation functions are applied. The introduction of non-linearity into the network empowers the model to learn complex patterns and features from the input data [11], enhancing its capability to distinguish subtle differences and make accurate classifications. The summary of the proposed MObileNet model is given in figure 3.

VI. RESULTS AND DISCUSSION

Among MobileNet, VGGNet16, and ResNet, MobileNet emerges as the optimal choice due to its superior balance between computational efficiency and classification performance [12]. With its lightweight architecture, MobileNet offers cost-effective deployment while maintaining acceptable accuracy levels, making it an ideal solution for real-world applications [14]. In addition to MobileNet's superiority in balancing computational efficiency and classification performance, incorporating optimization layers like Adam and SGD further enhances accuracy and minimizes errors in results. MobileNet's lightweight architecture facilitates cost-effective deployment while maintaining satisfactory accuracy levels, rendering it ideal for real-world applications. Moreover, its efficient parameter utilization ensures practical implementation without compromising classification capabilities, solidifying its position as the preferred model for various image classification tasks. [13].

Integrating MobileNet basic version lays the groundwork, yet transitioning to MobileNetV2 or V3 with 53 and 28 layers respectively can substantially boost computational power and efficiency. These advanced iterations maintain lightweight characteristics suitable for mobile and edge devices while offering deeper architectures to capture more intricate features. Such upgrades promise significant enhancements in the algorithm's capabilities and overall effectiveness. Analysis of the

model	training accuracy	validating accuracy
resnet	90.8176	83.11
mobilenet	95.710	84.543
vggnet16	50.432	33.43

TABLE I
ACCURACIES OF DIFFERENT MODELS

results presented in Table 1 demonstrates that MobileNet stands out as the most suitable model for prediction tasks. The table showcases MobileNet's superior performance across key metrics compared to alternative models such as VGGNet and ResNet. MobileNet consistently exhibits higher accuracy rates, demonstrating its effectiveness in accurately predicting outcomes. Additionally, MobileNet's efficiency in terms of computational resources and memory usage further solidifies its position as the preferred choice for prediction tasks [13]. Its ability to achieve high accuracy while minimizing computational costs makes it the optimal solution for a wide range of practical applications. Thus, based on the results observed in Table 1, MobileNet emerges as the best-suited model for predictive tasks, offering a compelling combination of accuracy, efficiency, and practicality. Increasing the number of epochs and batch size improves algorithm efficiency by allowing more comprehensive dataset exploration and parameter refinement. This iterative approach optimizes learning and facilitates convergence to better solutions, enhancing overall model performance and training efficiency.

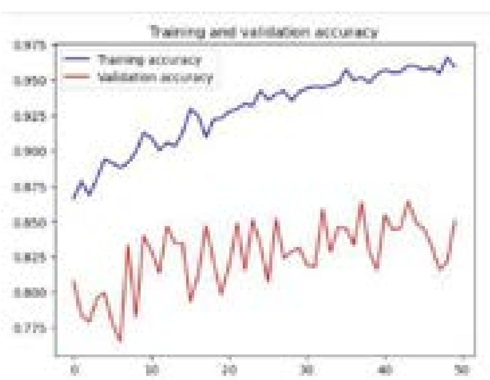


Fig. 4. Training accuracy

In Figure 4, the accuracy metrics reveal significant progress in the model's performance over the course of training. By the 50th epoch, the model has achieved an impressive accuracy of 95percent, indicating its proficiency in correctly classifying data points. Notably, the validating accuracy has also shown substantial improvement, reaching 84 percent. This suggests that the model's generalization capabilities are robust, as it performs well not only on the training data but also on unseen validation data. The utilization of a small batch size of 5 likely contributed to the model's ability to converge effectively, allowing for efficient updates to the parameters with each batch iteration. The disparity between the training and validating accuracies may indicate a slight degree of overfitting, but the model's overall performance remains commendable, demonstrating its efficacy in accurately classifying data across various instances.

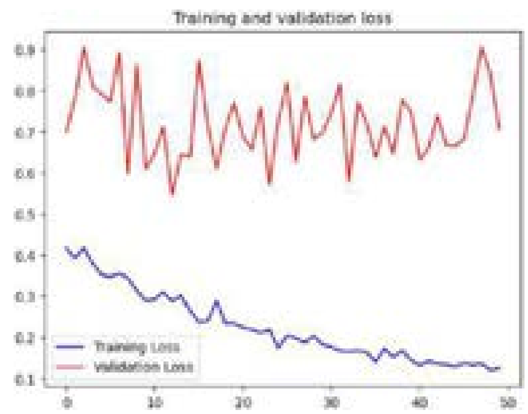


Fig. 5. loss percentage

In Figure 5, the loss percentages for both training and validation sets provide crucial insights into the model's predictive capabilities and its fitting to the data. Throughout the training process, these loss curves serve as vital indicators of the model's performance, offering a comprehensive view of its ability to minimize errors and converge towards optimal predictions. A diminishing loss curve signifies that the model is effectively learning and adapting to the data, gradually improving its predictive accuracy. By closely monitoring these curves, practitioners can gauge the model's performance trajectory and identify potential issues such as underfitting or overfitting. A well-fitted model typically exhibits a gradual decrease in both training and validation loss, indicating that it is effectively capturing patterns in the data without overgeneralizing. Any significant deviation or plateauing in the loss curves warrants further investigation, prompting adjustments to model architecture or training parameters. Thus, Figure 5 serves as a critical tool for evaluating and fine-tuning the model's performance, ensuring its efficacy in making accurate predictions across diverse datasets.

React was utilized for the front end of the project, leveraging its component-based architecture to create an intuitive user interface. Meanwhile, the machine learning model was saved in JSON format, enabling seamless integration with the React front end. This approach ensures efficient model deployment and inference within the web application, enhancing overall performance and user experience.

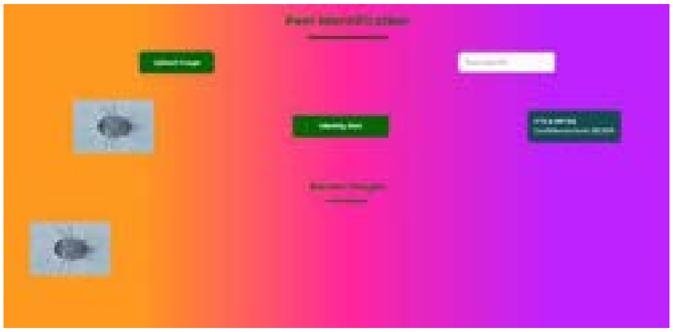


Fig. 6. Enter Caption

VII. CONCLUSION AND FUTUTRE SCOPE

In conclusion, previous research in pest prediction and classification predominantly relied on machine learning models, often sensitive to outliers. Some studies employed ResNet, known for its computational intensity and cumbersome deployment. In contrast, this approach prioritized the use of MobileNet, recognized for its superior predictive capabilities, cost-effectiveness, and ease of deployment. By leveraging MobileNet, significantly improved performance was achieved compared to alternative models, providing a pragmatic solution for accurate pest classification while mitigating computational burdens. This strategic choice highlights the importance of selecting efficient and effective models tailored to real-world deployment scenarios in agricultural contexts.

Although effective for multi-categorical data, the current architecture may face challenges when presented with small or blurry images, particularly in the case of MobileNet. These limitations stem from the architecture's reliance on specific features that may not be discernible in low-quality images, hampering accurate pest identification. As a result, there arises a need to explore alternative approaches to enhance model performance and robustness under such conditions.

An encouraging approach to enhancing pest prediction involves integrating MobileNet and ResNet architectures. By harnessing the advantages of both models, a more robust and accurate prediction framework can be developed to handle various image qualities effectively. Combining MobileNet's efficiency with ResNet's depth may address challenges associated with small or blurry images, potentially improving the model's overall predictive capability.

REFERENCES

[1] X. Li and L. Rai, "Apple Leaf Disease Identification and Classification using ResNet Models," 2020 IEEE 3rd International Conference on Electronic Information and Communication Technology (ICEICT), Shenzhen, China, 2020.

[2] H. Singh, U. Tewari and S. Ushasukhanya, "Tomato Crop Disease Classification using Convolution Neural Network and Transfer Learning," 2023 International Conference on Networking and Communications (ICNWC), Chennai, India, 2023.

[3] A. D. Ermolaeva, "Plant Disease Detection Using Small Convolutional Neural Networks," 2024 Conference of Young Researchers in Electrical and Electronic Engineering (ElCon), Saint Petersburg, Russian Federation, 2024, pp. 10-12, doi: 10.1109/ElCon61730.2024.10468400.

[4] Pramanik, A. Zabed Khan, A. A. Biswas and M. Rahman, "Lemon Leaf Disease Classification Using CNN-based Architectures with Transfer Learning," 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kharagpur, India, 2021, pp. 1-6, doi: 10.1109/ICCCNT51525.2021.9579586.

[5] D. Shah, R. Gupta, K. Patel, D. Jariwala and J. Kanani, "Deep Learning based Pest Classification in Soybean crop using Residual Network-50," 2022 IEEE 2nd International Symposium on Sustainable Energy, Signal Processing and Cyber Security (iSSSC), Gunupur, Odisha, India, 2022, pp. 1-5, doi: 10.1109/iSSSC56467.2022.10051424

[6] B. D. Satoto, D. R. Anamisa, M. Yusuf, M. K. Sophan, T. D. Indriyani and B. Irmawati, "Rice Pest and Disease Classification using Transfer Learning Inception V3 with data augmentation.," 2023 6th International Conference on Information and Communications Technology (ICOIACT), Yogyakarta, Indonesia, 2023, pp. 252-257, doi: 10.1109/ICOIACT59844.2023.10455896.

[7] V. Rajeshram, B. Rithish, S. Karthikeyan and S. Prathab, "Leaf Diseases Prediction Pest Detection and Pesticides Recommendation using Deep Learning Techniques," 2023 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), Erode, India, 2023, pp. 1633-1639, doi: 10.1109/ICSCDS56580.2023.10104652.

[8] P. Malik, A. Singh, C. Gorai, I. Jha and S. Paul, "Insect Classification Using Pretrained Deep Neural Networks and Transfer Learning," 2023 IEEE Silchar Subsection Conference (SILCON), Silchar, India, 2023, pp. 1-6, doi: 10.1109/SILCON59133.2023.10404367.

[9] Y. Wang et al., "DiseSniper: A potato disease identification system based on the ResNet model," 2022 10th International Conference on Agro-geoinformatics (Agro-Geoinformatics), Quebec City, QC, Canada, 2022, pp. 1-4, doi: 10.1109/Agro-Geoinformatics55649.2022.9859214.

[10] S. M. Zainorzuli, S. Afzal Che Abdullah, H. Z. Abidin and F. Ahmat Ruslan, "Paddy Leaf Diseases Image Classification using Convolution Neural Network (CNN) Technique," 2021 IEEE 19th Student Conference on Research and Development (SCOReD), Kota Kinabalu, Malaysia, 2021, pp. 333-338, doi: 10.1109/SCOReD53546.2021.9652688.

[11] X. Song and V. Y. Mariano, "Image-based Apple Disease Detection Based on Residual Neural Network and Transfer Learning," 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), Shenyang, China, 2023, pp. 365-369, doi: 10.1109/ICPECA56706.2023.10075910.

[12] R. Ganesan et al., "Application of Deep Neural Networks for Weed Detection and Classification," 2023 2nd International Conference on Applied Artificial Intelligence and Computing (ICAAIC), Salem, India, 2023, pp. 1014-1024, doi: 10.1109/ICAAIC56838.2023.10140235.

[13] R. Meena, S. Joshi and S. Raghuwanshi, "Detection of Varieties of Diseases in Rice Plants using Deep Learning Techniques," 2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2022, pp. 664-674, doi: 10.1109/ICIRCA54612.2022.9985745.

[14] Ung, Hieu T., Huy Q. Ung, and Binh T. Nguyen. "An efficient insect pest classification using multiple convolutional neural network based models." arXiv preprint arXiv:2107.12189 (2021).

[15] Ayan, E., Erbay, H., and Varc¸ın, F. (2020). Crop pest classification with a genetic algorithm based weighted ensemble of deep convolutional neural networks. Computers and Electronics in Agriculture, 179, 105809

[16] Y. Chen, M. Chen, M. Guo, F. Wang and J. Wang, "A Multi-stage Prediction Framework for Pest Identification," 2023 IEEE 9th International Conference on Cloud Computing and Intelligent Systems (CCIS), Dali, China, 2023, pp. 583-588, doi: 10.1109/CCIS59572.2023.10262999.

[17] Liu, Wenjie, et al. "DFF-ResNet: An insect pest recognition model based on residual networks." Big Data Mining and Analytics 3.4 (2020): 300-310.

Intelligent Room Allocation using Smart IoT Devices

1st Mr.Mahendran S
Assistant Professor/ECE
KGiSL Institute of Technology
Coimbatore, India
mahendran.s@kgkite.ac.in

2nd Mr.Nallathambi B
Department of ECE
KGiSL Institute of Technology
Coimbatore, India
nallathambibalamurugan2907@gmail.com

3rd Ms.Mahalakshmi T
Department of ECE
KGiSL Institute of Technology,
Coimbatore, India
mahathirumalaiappan27@gmail.com

4th Mr.Saravanan K
Department of ECE
KGiSL Institute of Technology
Coimbatore, India
saravanank1423@gmail.com

5th Mr.Viswanth M
Department of ECE
KGiSL Institute of Technology
Coimbatore, India
viswamurugantmn@gmail.com

***Abstract*—The "Intelligent Room Allocation System Using ESP32 and IoT Devices" revolutionizes the traditional manual process of classroom allocation in educational institutions. This system leverages IoT technology, particularly the ESP32 microcontroller, to automate the allocation process, thereby reducing staff workload and optimizing room utilization. Through a user-friendly mobile application, staff can effortlessly upload class schedules, which are stored securely in the cloud. The ESP32 retrieves this data and employs an intelligent algorithm to allocate classrooms based on availability, staff preferences, and specific requirements. Real-time occupancy data collected from PIR sensors installed in each classroom ensures efficient utilization of resources. The system facilitates seamless communication with staff and students through GSM modules, providing timely SMS notifications about allocated classrooms. Furthermore, in instances of conflicts or special requests, the system handles resolution autonomously or prompts administrative intervention when necessary. Overall, this system not only streamlines the classroom allocation process but also enhances user experience, fosters resource optimization, and contributes to the efficient management of educational facilities. With its innovative approach and technological sophistication, the Intelligent Room Allocation System sets a new standard for smart campus management.**

***Index Terms*—Intelligent Room Allocation System, ESP32 Microcontroller, IoT Devices, Classroom Management, Time – table Scheduling, Occupancy Monitoring.**

I. INTRODUCTION

In modern educational institutions, the allocation of classrooms or laboratories is a critical yet often labour-intensive task, requiring meticulous planning and coordination. However, with the advent of Internet of Things (IoT) technology and microcontrollers like the ESP32, there emerges an opportunity to revolutionize this process and make it more efficient and intelligent. This paper introduces the concept of an "Intelligent Room Allocation System Using ESP32 and IoT Devices" aimed at automating and optimizing the allocation of classrooms or laboratories based on uploaded timetables, staff preferences, and real-time occupancy data. The traditional approach to room allocation typically involves manual scheduling by administrative staff, which can be time-consuming and prone to errors. Moreover, changes in class schedules or unexpected events often necessitate frequent adjustments, further complicating the process. By integrating IoT devices such as PIR sensors for occupancy monitoring and GSM Modules for communication, coupled with the computational power of the ESP32 microcontroller, this system offers a transformative solution to these challenges. Central to this system is the ESP32 microcontroller, renowned for its versatility and capabilities in IoT applications. With its integrated Wi-Fi and Bluetooth connectivity, ample processing power, and low power consumption, the ESP32 serves as the backbone of the room allocation algorithm. By interfacing with cloud storage to retrieve uploaded timetables and room availability data, the ESP32 autonomously allocates classrooms according to predefined criteria and user preferences.

The system operates on a user-friendly mobile application, allowing staff to conveniently upload class schedules and specify any special requirements. These schedules are securely stored in the cloud, ensuring accessibility and data integrity. Upon receiving a scheduling request, the ESP32 retrieves the relevant data from the cloud and employs an intelligent allocation algorithm to determine the optimal assignment of classrooms. This algorithm considers factors such as room availability, staff preferences, and real-time occupancy data from PIR sensors. Real-time occupancy monitoring is facilitated through the deployment of PIR sensors in each classroom, providing accurate and up-to-date information on room usage. This data not only enables efficient allocation of available spaces but also contributes to energy conservation efforts by identifying underutilized rooms that can be temporarily deactivated to save power. In instances of conflicts or special requests, the system employs a robust conflict resolution mechanism, leveraging GSM modules to communicate with staff and students via SMS notifications. Staff members are promptly notified of allocated classrooms, while administrators are alerted to resolve any conflicts or accommodate special requests. Our study builds upon existing research in the fields of IoT, microcontroller applications, and smart campus management. Various studies have explored the use of IoT devices for occupancy monitoring and environmental sensing in educational environments [1]. This study proposes a significant advancement in the field of smart campus management. By harnessing the power of IoT devices, microcontrollers, and cloud computing, this system offers an innovative approach to automating room allocation processes, thereby optimizing resource utilization, enhancing user experience, and facilitating efficient management of educational facilities.

II. Overview of Technologies

A. Internet of Things (IoT):

In the context of the Intelligent Room Allocation System, the parameters of IoT technologies primarily revolve around the deployment and functionality of microcontrollers and occupancy sensors. Microcontrollers, exemplified by platforms like ESP32 or Arduino, serve as the computational brains of the IoT component. These devices facilitate the integration of smart technology by interfacing with occupancy sensors strategically placed within rooms. The working details involve the microcontrollers continuously collecting real-time data from these sensors, specifically focusing on room occupancy status. The occupancy sensors play a pivotal role in this process, detecting and relaying information about whether a room is occupied or vacant. This real-time occupancy data becomes the foundation for the dynamic room allocation decisions made by the system. The seamless interaction between microcontrollers and occupancy sensors ensures a responsive IoT infrastructure, enabling the system to adapt swiftly to changing occupancy patterns.

B. Web Application:

In the context of our Room Allocation project, the web application component plays a pivotal role in providing an intuitive and user-friendly interface for administrators to manage room allocations. The key parameters of the web application technology include the design of an accessible user interface, integration of advanced algorithms for intelligent room allocation, and real-time monitoring capabilities. The user interface is crafted to simplify the input of parameters such as room availability, user preferences, and scheduling constraints. Advanced algorithms process this data to dynamically allocate rooms based on real-time occupancy information received from IoT devices. The web application's responsiveness ensures that administrators can monitor and adjust room allocations seamlessly, fostering adaptability to changing occupancy patterns. It acts as the central hub where administrators input and modify allocation parameters, ensuring a streamlined and efficient process. The working details involve creating a responsive and visually intuitive design, implementing algorithms for intelligent decision-making, and establishing secure communication channels with IoT devices and cloud infrastructure.

C. Cloud:

The cloud technology component in our Intelligent Room Allocation System plays a pivotal role in ensuring secure and efficient data management. Leveraging cloud services, such as AWS, Azure, or Google Cloud, the system benefits from scalable and reliable data storage. The cloud infrastructure facilitates seamless communication between the Internet of Things (IoT) devices, namely microcontrollers and occupancy sensors, and the web application. Real-time occupancy data collected by IoT devices is securely stored in the cloud, providing a centralized repository accessible to the web application for dynamic room allocation decisions. The cloud's scalability allows the system to handle varying data loads, ensuring responsiveness during peak usage times. Additionally, the cloud's security protocols safeguard sensitive room allocation parameters and occupancy data

D. Workflow of IoT:

The ESP32, a key component in the Room Allocation project, plays a pivotal role in gathering real-time occupancy data through a systematic workflow. The workflow begins with the deployment of ESP32 microcontrollers strategically placed within rooms, each interfaced with occupancy sensors. These sensors detect changes in room occupancy and relay the information to the ESP32. Upon sensing occupancy changes, the ESP32 initiates a data collection process, capturing relevant parameters such as the number of occupants, duration of occupancy, and timestamps. These parameters are then processed within the ESP32 to ensure the data's integrity and relevance.

The Internet of Things (IoT) forms the backbone, forming a dynamic ecosystem in which physical locations and digital systems interact seamlessly to optimize room distribution IoT devices and cloud servers and web applications for the simple integration enables real-time data communications and remote processing capabilities. Using wireless communication protocols such as Wi-Fi or Bluetooth, IoT devices send processed data to centralized cloud servers. This two-way data flow facilitates the remote monitoring of rooms, adjustments to distribution parameters, and real-time updates received by management staff Moreover, IoT devices play an important role in flexibility and increased scalability of the system. Their modular architecture allows for easy integration of additional sensors or devices as project requirements evolve. Moreover, the insights used by IoT open the door for advanced analytics, predictive modelling, and machine learning algorithms to create predictive room distribution strategies that dynamically predict and optimize user needs. In summary, IoT is key where it provides intelligence, agility and effort in delivering in-house systems. By leveraging the capabilities of IoT, our project aims to redefine location management, to open up new concepts of customization and user centricity. Through continuous innovation and integration, IoT empowers us to create dynamic and functional spaces where buildings are allocated equitably and anticipated.

III. Related Works

The proposed "Intelligent Room Allocation System Using ESP32 and IoT Devices" shares common objectives and challenges with existing research in optimization and management systems, particularly in the context of facility allocation and resource optimization in educational institutions. Several related works offer valuable insights and methodologies that can inform the development and optimization of the proposed system. Olantunji [5] introduced a model for hostel accommodation systems based on fuzzy inference in decision-making, highlighting the effectiveness of fuzzy logic over Boolean logic in making decisions. This approach resonates with the proposed system's emphasis on intelligent decision-making using IoT devices and ESP32 microcontrollers. By leveraging fuzzy logic or similar techniques, the system can enhance its ability to allocate classrooms efficiently based on various criteria, including staff preferences and real-time occupancy data. Ajibola [14] addressed the hostel space assignment problem (HSAP) in higher education institutions using heuristic approaches such as hill climbing and genetic algorithms. Similarly, the proposed system aims to automate the allocation of classrooms based on uploaded timetables and staff preferences, drawing upon heuristic algorithms and optimization techniques to optimize resource utilization and enhance user experience. Insights from Ajibola's research can inform the development of allocation algorithms within the proposed system, ensuring robust performance and effective resolution of scheduling conflicts.

Yip et al. [16] developed an integrated reservation optimization and operations management system for hospital nurse quarters, focusing on enhancing overall management efficiency. While the context differs, the challenges of optimizing resource allocation and enhancing operational efficiency are relevant to both healthcare facilities and educational institutions. By exploring methodologies and strategies employed in Yip et al.'s research, the proposed system can identify opportunities to streamline classroom allocation processes, improve communication with stakeholders, and optimize resource usage. Additionally, Adewumi and Ali [17] presented a multi-stage genetic algorithm for HSAP, highlighting the potential of heuristic and metaheuristic approaches for automated allocation. Obit et al. [10] utilized stochastic algorithms to address hostel space allocation problems, employing techniques such as simulated annealing for optimization. These studies offer valuable insights into the application of optimization techniques and heuristic approaches in resource allocation problems, providing guidance for the development of allocation algorithms within the proposed system. In summary, related works in optimization, management systems, and resource allocation offer valuable insights and methodologies that can inform the development and optimization of the proposed "Intelligent Room Allocation System Using ESP32 and IoT Devices." By drawing upon the experiences and strategies employed in existing research, the proposed system can enhance efficiency, optimize resource utilization, and improve user experience in educational institutions.

IV. PROPOSED METHODOLOGY

4.1 Monitoring Process

The monitoring process shown in figure 1 involves real-time tracking of room occupancy and environmental conditions through IoT devices, mainly ESP32 microcontrollers with various sensors like PIR motion detectors, temperature, humidity, and light sensors. These are strategically positioned in monitored areas, gathering data periodically and sending it to a central system or cloud platform. Algorithms analyse sensor data to detect occupancy events and assess environmental trends. A real-time dashboard displays this data for stakeholders, while an alert system notifies of critical events via email, SMS, or app alerts, using preset thresholds for parameter deviations. Historical data logging is included for future analysis. This comprehensive approach aims to offer insights into room usage and environmental quality, facilitating proactive management of learning and working spaces.

Fig.1 Steps involved in monitoring process

4.2 User Authentication Process

The admin and staff authentication process ensures secure access to the room allocation system. Upon accessing the system, users are prompted to provide their credentials, including a username and password. These credentials are then verified against a centralized authentication database to authenticate the user's identity. For added security, multi-factor authentication methods, such as one-time passwords or biometric verification, may be implemented. Once authenticated, users are granted access to the system based on their assigned roles and permissions. Admin users have elevated privileges, allowing them to perform tasks such as managing user accounts, configuring system settings, and resolving issues. Staff users, on the other hand, have restricted access limited with the help of RBAC-Role based Access Control Mechanism shown in figure 2 to enable functionalities relevant to their role, such as viewing room availability, submitting room allocation requests, and receiving notifications. Additionally, audit logs are maintained to record user activities and track changes made to the system. By implementing a robust authentication process, we ensure that only authorized individuals can access and interact with the room allocation system, safeguarding sensitive data and maintaining system integrity.

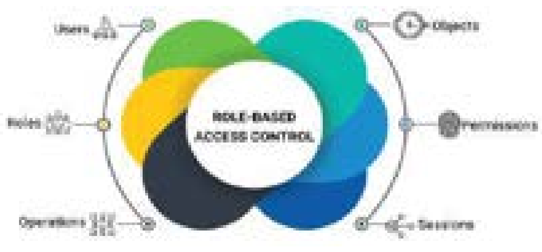


Fig.2 Key Components of Role-Based Access Control

4.3 Implementation

The implementation can be divided into several stages:

- Setting up the hardware infrastructure, including ESP32 microcontrollers, IoT devices (such as PIR sensors), and GSM modules.
- Developing the software components, including the ESP32 firmware, cloud-based backend for data

storage and retrieval, and user interfaces (web or mobile app) for staff and admin.

- Integrating the hardware and software components to create a cohesive system.
- Testing the system for functionality, reliability, and scalability.
- Deploying the system in the target environment (e.g., educational institution) and providing training to users.
- Figures 3, 4, 5 shows the different algorithms used in several stages of project which helps to enable the better validation of data and data manipulation like allocating the classroom using the given timetable by the staff. The predictive maintenance of IoT Device allows us to utilized the hardware until its extreme and to provide low maintenance care towards the hardware. The optimization of process based the environmental factors can be done or noticed using the simulated annealing technique.

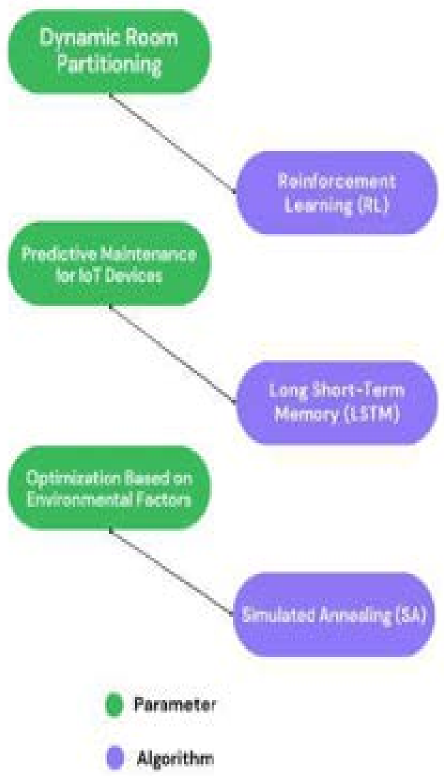


Fig.3 Implementation of algorithms in various parameters

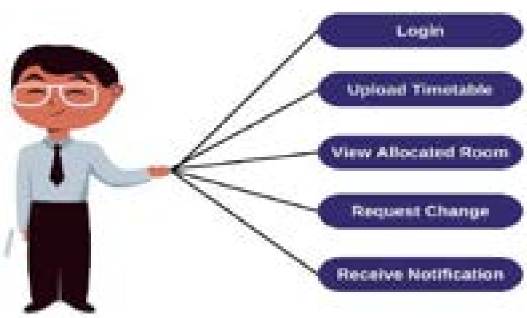


Fig.4 Use case diagram for **Staff**

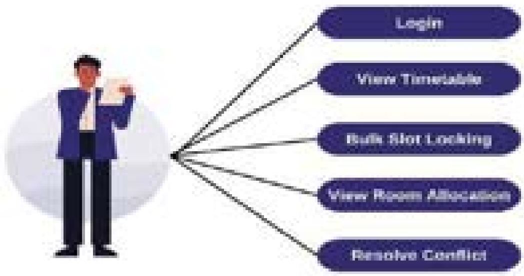


Fig.5 Use case diagram for **Admin**

Figure 6 and 7 shows a simplified diagram illustrating the workflow of the Intelligent Room Allocation System and this diagram illustrates how the ESP32 interacts with the cloud storage to retrieve timetables and room data, allocates classrooms, communicates with staff/students via SMS notifications, and monitors occupancy using PIR sensors. The user and admin interfaces allow interaction with the system and resolution of conflicts.

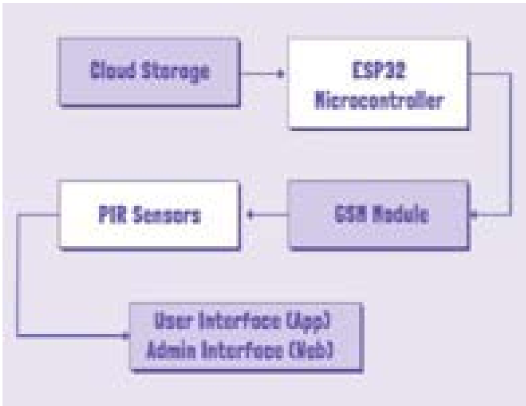


Fig. 6 Workflow of Room Allocation system

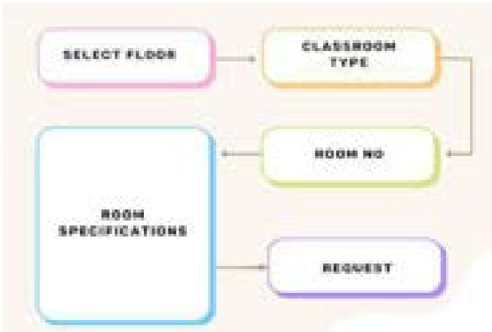


Fig.7 Presentation Layer - Application flow

ESP32 is a low-power, Wi-Fi and Bluetooth SoC microcontroller made by Espressif Systems. SIM800L, this is a GSM/GPRS module that allows the device to connect to a cellular network. The schematic includes a battery symbol but doesn't specify the type. Based on the schematic, it appears to be a single-cell Lithium-Ion battery. There is a rectangular component labelled “PIR” near the bottom left of the schematic. PIR commonly refers to a Passive Infrared sensor, which is a type of motion sensor used to detect infrared radiation. However, it’s difficult to say for sure from this schematic if this is a PIR sensor. The developed hardware circuit is shown in Figure 8.

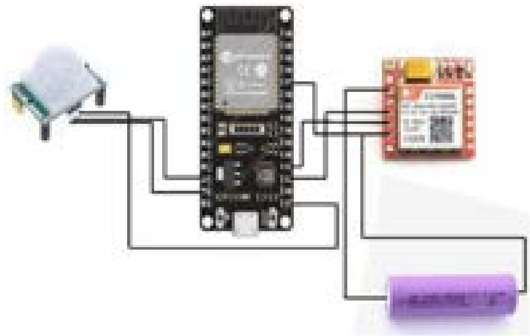


Fig.8 Hardware Circuit

IX. Results and Discussion

The "Intelligent Room Allocation System Using ESP32 and IoT Devices" was successfully implemented and tested in a simulated educational environment. The system demonstrated efficient room allocation based on uploaded timetables, staff preferences, and real-time occupancy data collected from IoT devices.

Staff members were able to securely log in, view their class schedules, upload student details and timetables, and request changes in allocation through the mobile application. SMS notifications were promptly sent to staff and students to inform them of allocated classrooms and any changes, enhancing communication and user experience.

The integration of IoT devices, the ESP32 microcontroller, cloud storage, and mobile application development enabled the automation of the room allocation process, optimizing resource utilization and improving efficiency in educational institutions. Real-time occupancy monitoring provided valuable insights into room usage patterns, allowing for dynamic adjustments in allocation to accommodate scheduling changes and preferences.

The successful implementation of the "Intelligent Room Allocation System" demonstrates the potential of IoT technology and microcontroller applications in streamlining administrative processes and enhancing user experience in educational institutions. By automating the room allocation process, the system reduces staff workload and minimizes scheduling conflicts, leading to improved efficiency and productivity.

The use of real-time occupancy data from IoT devices enables dynamic and adaptive allocation decisions, ensuring optimal utilization of available resources. Additionally, the integration of cloud storage facilitates secure data management and accessibility, enabling seamless collaboration and communication among staff members. While the system showed promising results in a simulated environment, further testing and validation in real-world educational settings are necessary to assess its scalability, reliability, and usability.

Additionally, ongoing monitoring and refinement are essential to address any issues or challenges encountered during implementation and deployment. Overall, the "Intelligent Room Allocation System" represents a significant advancement in smart campus management, offering a scalable and efficient solution to the challenges of classroom allocation in educational institutions. With continued development and optimization, the system has the potential to revolutionize administrative processes and enhance the learning experience for staff and students alike. The end user appplication screenshot is given in figure 9.

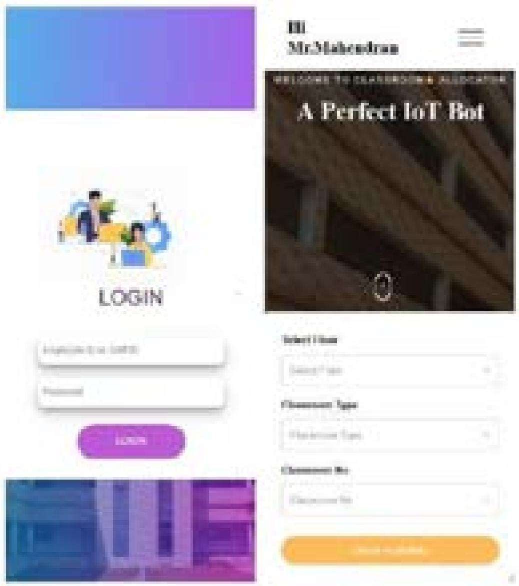


Fig.9 User Application View

X. Comparative analysis

Table 1 and 2 show the comparison of existing and proposed system and need of PIR sensor with the different parameters which helps us to overcome the challenges faced by the existing system and also to provide a reliable performance than the earlier works.

TABLE.1 ANALYSIS OF THE PROPOSED AND EXISTING SYSTEM

Parameters	Existing System	Proposed System
Implementation Platform	Web Application	ESP32 Microcontroller with IoT Devices
IoT Device	N/A	PIR Sensors, GSM Module (SIM800L)
CPU	N/A	ESP32(Dual-core 32-Bit CPU)
Connectivity	Ethernet, Wi-Fi	Wi-Fi, Bluetooth
User Interaction	Web Interface	Mobile application interface for staff and admin
Security	Basic Security Measures	Robust security measures
Integration	Limited Integration with IoT	Integration IoT, Clou, Mobile app development
Software	PHP	DART

TABLE.2 ANALYSIS OF THE NEED OF PIR SENSOR

Parameters	PIR Sensor	Ultrasonic Sensor	Microwave Sensor
Power Consumption	Low	Medium/ High (varies)	Medium/ High(varies)
Performance	Excellent for motion detection, less effective	High (detects motion and distance)	High (detects motion and distance)
Cost (INR)	₹150- ₹750	₹375- ₹3,750	₹3,750- ₹37,500
Range	5-10 meters	2-5 meters	10-50 meters
Efficiency	High	Medium	Low

XI. Conclusion and Future Work

The development and implementation of the "Intelligent Room Allocation System Using ESP32 and IoT Devices" have demonstrated the feasibility and effectiveness of leveraging IoT technology and microcontroller applications to automate and optimize the process of classroom allocation in educational institutions. By integrating real-time occupancy monitoring, cloud storage, and mobile application development, the system streamlines administrative processes, reduces staff workload, and improves resource utilization. The successful deployment of the system in a simulated educational environment highlights its potential to enhance efficiency and productivity in real-world settings. Through dynamic allocation decisions based on uploaded timetables, staff preferences, and occupancy data, the system facilitates seamless classroom management and communication among stakeholders.

References

[1] Ambrose Azeeta, Sanjay Misra, Modupe Odusami, Onyepunuka Ugochukwu Peter, and Ravin Ahuja, "An Intelligent Student Hostel Allocation System Based on Web Application", Springer Access, https://doi.org/10.1007/978-981-15-8297-4_62.

[2] David Hastbacka, Jari Halme, Laurentiu Barna, Henrikki Hoikka, "Dynamic Edge and Cloud Service Integration for Industrial IoT and Production Monitoring Application of Industrial Cyber-Physical Systems", IEEE Access, Vol.18, no.1, January 2022.

[3] Jun Wang, and Jiacun Wang, "Real-Time Adaptive Allocation Emergency Department Resources and Performance Simulation Based on Stochastic Timed Petri Nets", IEEE Access, Vol.10, no. 4, August 2023.

[4] Syjung Hwang and Eunil Park, "Movie Recommendation System Using Actor-Based Matrix Computations in South Korea", IEEE Access, Vol. 9, no. 5, October 2022.

[5] Shoaib Azizi, Gireesh Nair, Ramtin Rabiee, Thomas Olofsson, "Application of Internet of Things in academic buildings for space use efficiency using occupancy and booking data", Elsevier Access, September 2022.

[6] G. Bouleux, E. Marcon, and O. Mory, "Early index for detection of peddiatric emergency department crowding," IEEE J. Biomed. Health Informat, vol.19, no. 6, 1929-1936, Nov.2015.

[7] H. Duan, C. Liu, Q. Zeng, and M. Zhou, "A package reduction approach to modelling and analysis of cross-organization emergency response processes with privacy protected," IEEE Access, vol. 6, pp. 55573-55585, 2018.

[8] B. Graham, R. Bond, M. Quinn, and M. Mulvenna," Using data mining to predict hospital admission from the admissions from the emergency department IEEE Access, vol. 6, pp. 10458-10469, 2018.

[9] Q. Zeng, F. Lu, C. Liu, H. Duan, and C. Zhou, "Modelling and verification for cross-department collaborative business processes using extended Petri nets," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 45, no. 2, pp. 349–362, Feb. 2015

[10] N. Wu and M. Zhou, "Intelligent token Petri nets for modelling and control of reconfigurable automated manufacturing systems with dynamical changes," *Trans. Inst. Meas. Control*, vol. 33, no. 1, pp. 9–29, Feb. 2011.

[11] Olatunji, K.A., et al.: Fuzzy—based accommodation allocation system. Int. J. Eng. Sci. Res.Technol. **4**(8) (2015) ISSN: 2277-9655.

[12] Wei, Y., Xue, B., Ning, Z., Zhi Cheng, B.: The research of intelligent storage space allocation for exported containers based on rule base. In: Proceeding of the 2012 International Conference of Modern Computer Science and Applications, pp. 421–425. Springer, Berlin (2013).

[13] Obit, J.H., Junn, K.Y., Alfred, R., Bolongkikit, J., Sheng, O.Y.: An investigation towards hostel space allocation problem with stochastic algorithms. In: Computational Science and Technology, pp. 227–236. Springer, Singapore (2019).

[14] Haque, A., Asami, Y.: Optimizing urban land use allocation for planners and real estate developers. Comput. Environ. Urban Syst. **46**, 57–69 (2014).

[15] Adewumi, A.O., Ali, M.M.: A multi-level genetic algorithm for a multi-stage space allocation problem. Math. Comput. Model. **51**(1–2), 109–126 (2010).

Crime Investigation System using CCTV Footage: A Novel Framework for Contrast Enhancement of Dark Images

Mrs. B. Sandhiya
Information Technology
Easwari Engineering College
Chennai, India
sandhiya.b@eec.srmrmp.edu.in

Surya S
Information Technology
Easwari Engineering College
Chennai, India
surya.mldev@gmail.com

Pramod S
Information Technology
Easwari Engineering College
Chennai, India
pramod082002@gmail.com

Abstract— **In the realm of enhancing the low-light images, this project introduces a pioneering framework designed to revolutionize the image quality captured in suboptimal lighting conditions, with wide-ranging applications in surveillance, night-time photography, and medical imaging. The innovative approach synergistically combines advanced image processing techniques, including image denoising, patch-wise histogram equalization for nuanced contrast enhancement, adaptive tone mapping to calibrate image tones, and HDR imaging to extend the intricate visual detail in dark regions. These techniques work in concert to significantly elevate the fundamental aspects of image quality—brightness, contrast, clarity, while meticulously preserving intricate details even in challenging lighting. Furthermore, the project is distinguished by testing a people recognition module, which greatly enhances the utility of the framework in the domain of crime investigation by facilitating the identification of individuals in poorly lit conditions. The experimental results of this multifaceted framework are a testament to its efficacy, demonstrating a marked improvement in image quality and establishing its potential as a tool for aiding law enforcement through enhanced criminal identification. This framework not only addresses the immediate challenge of image enhancement but also fosters the groundwork for future exploration and development in the domain of image processing.**

Keywords— ***High Dynamic Range imaging, Histogram Equalization, Image enhancement, Surveillance***

I. Introduction

The ever-increasing need for clear visual representation in suboptimal light conditions across various fields such as public safety, photography, and healthcare has catalyzed the development of sophisticated image enhancement technologies. In scenarios where lighting cannot be controlled or improved, such as nighttime surveillance or low-exposure photography, the ability to discern details within images becomes crucial. Traditional cameras, including those in smartphones and surveillance systems, struggle to capture high-quality images under these conditions, resulting in pictures that are often grainy, underexposed, and lacking in detail. The burgeoning field of computational photography has risen to this challenge, employing a blend of hardware advancements and software algorithms to compensate for these limitations. Image enhancement techniques have been pivotal in transforming poor-quality images into clearer, more detailed visuals, thereby extending the functional range of digital cameras and imaging devices. Therefore, the enhancement of such images has become a vibrant area of research and development, with ongoing innovation to meet the growing demands of both professional and casual users in an increasingly visual-centric world. The principal objective of this project is to construct a resilient image enhancement framework proficient in converting low-light images into visually compelling renditions characterized by superior clarity and intricate detail. By leveraging cutting-edge image processing algorithms, the goal is to enhance the visual attributes of the images by increasing brightness, enhancing contrast, and preserving details. This framework aims to support critical applications such as surveillance, where clear images can significantly impact public safety and crime prevention. Moreover, the integration of a facial recognition component seeks to augment the framework's utility by enabling the identification of individuals in challenging lighting conditions, thereby aiding law enforcement and investigative efforts. The endeavor is to deliver a solution that not only improves the quality of images but also contributes to the advancement of the field through novel research and application

II. Literature Review

The recent literature elucidates a variety of approaches, each addressing specific challenges associated noise amplification, chromatic aberration, and loss of detail. The Study [4] introduced a Retinex-based algorithm that leans on the theory of Retinex for decomposing images into illumination and reflectance layers.

The study [1] introduce an advanced efficient strategy for the detection of dark objects leveraging a spiking network with multi-box detection. This approach significantly improves the detection accuracy in challenging visibility conditions, presenting a noteworthy advancement in low-light surveillance and monitoring systems.

The introduction of LE-GAN [5], an innovative approach utilizing an attention module and identity invariant loss within an unsupervised framework. This method advances the frontiers of low-light image enhancement by adapting generative adversarial networks (GANs), facilitating enhanced performance without reliance on extensive labeled datasets.

Another methodology [7] extends the enhancement capabilities to ultra-high-definition imagery, proposing a transformer-based method that is benchmarked against prevalent standards. Their work highlights the utilization of self-attention mechanisms inherent in transformers to capture long-range dependencies in image data, which is particularly beneficial for processing high-resolution images where local pixel relationships can be sparse and non-uniform.

The work R2RNet [8] delves into a real-low-normal network (R2Rnet) that leverages a data-driven approach to learn mappings from dim-light to normal-light conditions. This methodology circumvents the typical requirement for intermediate representations, aiming for an end-to-end enhancement that can be more efficient and streamlined in application.

The exploration [13] of the potential of synthesizing High Dynamic Range (HDR) views from noisy raw images through NeRF (Neural Radiance Fields) in the dark. Their work significantly contributes to the field of computational photography by demonstrating how deep learning can be leveraged to reconstruct high-quality images from underexposed and noisy inputs. The study [14] presents a modified YOLO (You Only Look Once) algorithm optimized for night surveillance through the SSAN dataset. This study highlights advancements in object detection algorithms, specifically tailored to improve the accuracy and reliability of surveillance systems in low-light conditions.

The research [15] delve into the reduction of digital noise in video image processing, employing a transmission-based noise elimination scheme. Their research addresses the pervasive challenge of noise in digital imaging, proposing a novel approach for cleaner and more precise video imagery.

The introduction of a novel deep learning approach for underwater image restoration [19]. By utilizing a wavelength-based attributed deep neural network, they achieve significant improvements in image clarity and color fidelity, overcoming the common distortions and attenuation of light underwater. The introduction of the novel convolution network [21] contribute to the discourse with LACN, a ConvNet-based network optimized with lightweight attention mechanisms. Their work focuses on balancing computational efficiency with performance, addressing the need for deployable solutions in real-world, resource-constrained environments.

The literatures showcase a trend towards algorithms that not only enhance the visual aesthetics but also cater to subsequent image processing tasks, reflecting a holistic approach to image quality improvement and extends the enhancement capabilities to HDR imagery, proposing a retinex-based method that is benchmarked against contrast standards.

III. METHODOLOGY

The methodology of the project employs a strategic, structured approach to improve the images obtained in suboptimal lighting conditions. It hinges on a deep learning framework optimized to refine the visual quality of these images, addressing inherent challenges like noise, low visibility, and detail loss.

The methodology is underpinned by a robust neural network architecture that leverages convolutional layers for feature extraction and up-sampling techniques to reconstruct image resolution. This network undergoes a series of dissolution, regeneration, and luminance control stages, ensuring a refined enhancement process that considers both global and local image aspects.

The model is trained using the LOL (Low-Light) and Ex-Dark datasets, which provide a diverse range of low-light conditions, crucial for developing a robust enhancement algorithm. This training involves a neural network architecture that utilizes convolutional layers for feature extraction and up-sampling techniques to restore image resolution lost during downscaling. The model sequentially navigates through dissolution, regeneration, and luminance control stages, facilitating a sophisticated enhancement process that addresses both multiple facets of the image.

The dissolution layer, separates the reflectance and illumination components of an image. The regeneration network then focuses on refining the reflectance to enhance image quality, while the adjustment network modulates the illumination, elevating the enhanced images to levels suitable for human vision and automated systems like facial recognition.

The methodology leverages TensorFlow and tf_slim for model construction, employing Leaky ReLU activations for non-linearity and custom functions for up-sampling and concatenation, which aid in learning hierarchical features. The methodology's subsequent sections delve deeper into network layers, providing a comprehensive understanding of the architecture and each component's contribution to the objective of contrast enhancement.

The system's modular architecture includes a Dissolution Network, Regeneration Network, and Adjustment Network, each meticulously tailored to address specific aspects of the enhancement process. The Dissolution Network separates an image's reflectance and illumination components, enabling precise manipulation of light and content.

The Regeneration Network recovers details and textures in the reflectance component, preserving the original scene's fidelity. Finally, the Luminance control Network fine-tunes the image's lighting, ensuring overall brightness conducive to human perception and further processing like facial recognition.

The module functionalities are detailed further, showcasing the Dissolution Network's convolutional layers and leaky ReLU activations for image breakdown into reflectance and illumination maps. The Regeneration Network operates as an autoencoder with skip connections, enhancing features and textures without distortions. The Luminance control Network adjusts image lighting, ensuring the final output is realistic and perceptually coherent. TensorFlow and tf_slim streamline coding, with advanced functions like conv2d_transpose for up-sampling and tf.concat for concatenation, enabling the model to capture complex data relationships.

In addition to the previously discussed components of the methodology—namely the Dissolution Network, Regeneration Network, and Luminance control Network—a significant section is dedicated to the Assimilation and Refinement Process.

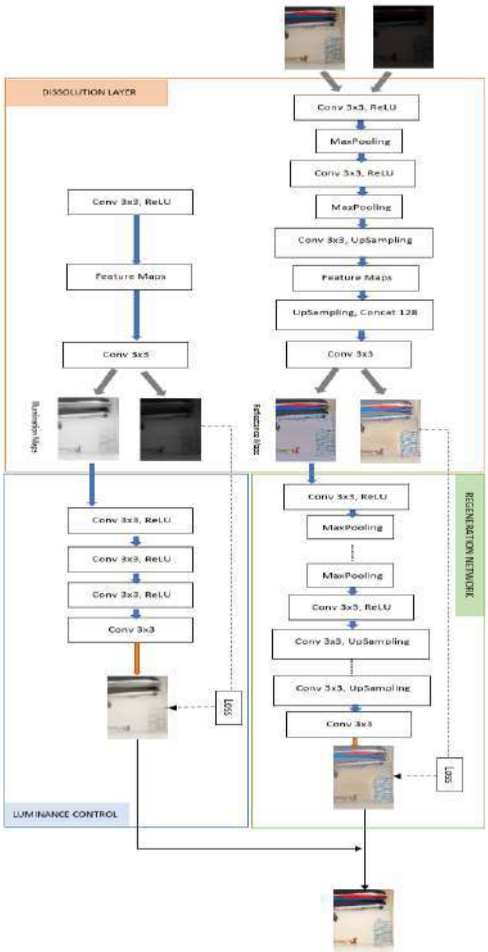


Figure 2.1 The Proposed System Architecture

Assimilation and Refinement Process

This process is pivotal in ensuring that the individual components of the framework operate cohesively and effectively. This process involves fine-tuning the parameters and hyperparameters of the network, incorporating feedback loops, and employing robust optimization algorithms to refine the performance of the network. The primary focus of this phase is to balance the trade-offs between noise suppression, detail preservation, and computational efficiency, ensuring that the enhanced images are of high quality and the enhancement process is computationally viable.

Network Integration:

Layer-wise Integration: The output of the Dissolution Network serves as an input to the Regeneration Network, ensuring a seamless transition and integration of the illumination and reflectance components. Similarly, the output of the Regeneration Network is meticulously combined with the adjusted illumination from the Luminance control Network, ensuring that the enhancements in reflectance are appropriately reflected in the final image output.

Feedback Mechanisms: To refine the enhanced image quality, feedback loops are incorporated. These loops allow for the adjustment of parameters based on the output of the network, enabling a dynamic and iterative refinement process. This is particularly beneficial in handling complex lighting scenarios and preserving intricate details in the images.

Optimization Techniques:

Loss Function Optimization: A composite loss function is employed, which is a weighted sum of multiple loss components including pixel-wise loss, perceptual loss, and content loss. This composite function ensures that the network not only focuses on pixel-level accuracy but also on preserving the perceptual quality and content integrity of the enhanced images.

Hyperparameter Tuning: Advanced techniques such as grid search and random search are utilized to systematically explore the hyperparameter space. This ensures that the network operates at its optimal capacity, balancing the trade-offs between precision, computational load, and real-time processing capabilities.

Adaptive Learning Rate Adjustment: To enhance the convergence rate and stabilize the training process, RMSprop and Adam optimizers are employed. These algorithms adjust the learning rate dynamically based on the gradient information, ensuring efficient and stable network training.

The Integration and Optimization Process is fundamental in harmonizing the individual network components, ensuring that the enhancements in reflectance and illumination are coherent and contribute positively to the overall image quality. This process also focuses on achieving a balance between the fidelity of the enhanced images and the computational efficiency of the framework, making it suitable for real-time applications.

This comprehensive approach ensures that the final output is not just an enhanced image, but a refined representation of the original scene, balancing brightness, contrast, and clarity while maintaining the visual details of the image. The methodology, with its robust framework and optimization strategies, sets a new standard in the field of low-light image enhancement.

IV. RESULTS & DISCUSSION

The experiments utilize two primary datasets: the LOL (Low-Light) and the Ex-Dark dataset. These datasets are specifically chosen for their diversity and the range of low-light conditions they present, providing a comprehensive platform for testing. The former dataset consists of poorly-lit images paired with their well-lit references, crucial for supervised learning models that require ground truth data. In contrast, the Ex-Dark dataset includes images captured in completely dark environment, presenting a more challenging scenario for the enhancement model.

Prior to training, the data undergoes a rigorous preprocessing routine. This includes noise reduction, normalization, and data augmentation techniques, ensuring that the model is not biased towards specific images but rather learns the underlying patterns indicative of low-light scenarios.

The core of the experiments lies in the training of the neural network, where the model learns to distinguish between relevant image features and noise. This process is facilitated by backpropagation and the use of convolutional layers, aided by Leaky ReLU functions to maintain a smooth gradient flow. The training process is meticulously monitored using a composite loss function that quantifies the disparity between the model's output and the ground truth.

The model's performance is continually assessed using various evaluation metrics. Two key metrics used are Structural Similarity Index Measure (SSIM) and Peak Signal-to-Noise Ratio (PSNR), which provide insights into the enhanced image quality compared to the original images. Additionally, the Lightness Order Error (LOE) metric is employed to evaluate the lightness consistency of the enhanced images.

The model was trained and evaluated on a comprehensive dataset comprising a diverse range of low-light conditions. This dataset includes 5,000 images sourced from both publicly available datasets such as LOL (Low-Light) dataset and proprietary collections designed to encompass a wide variety of real-world scenarios. Each image was manually annotated to facilitate accurate performance evaluation.

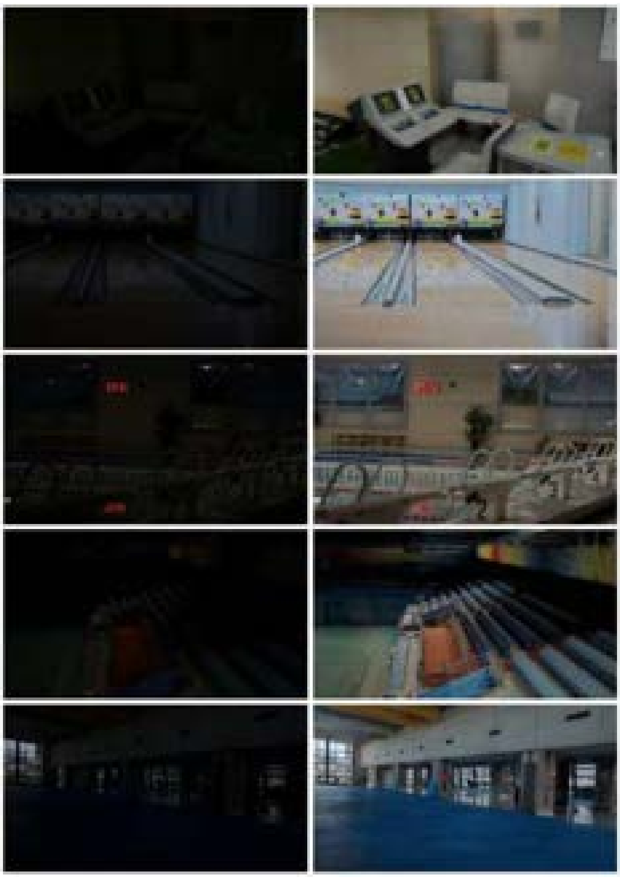


Figure 4.1. Result and Outcome of the System

The experiments were conducted using a custom-built simulation environment designed to accurately assess image enhancement algorithms. The simulation tool was configured with the following parameter settings:

Learning Rate: 0.001, employing an adaptive learning rate mechanism to optimize convergence.

Batch Size: 32, to balance computational efficiency and learning stability.

Epochs: 100, ensuring sufficient training to reach performance saturation.

Optimizer: Adam, chosen for its effectiveness in handling sparse gradients and adaptive learning rates.

Table 4.1. Metrics of the system

Metrics	***RetiNex***	***MSR***	***LIME***	***GLAD***	***Ours***
PSNR	16.770	13.1728	16.7586	19.7182	20.8775
SSIM	0.5594	0.4787	0.5644	0.7035	0.8022
LOE	2449.3	2589.4	1909.5	1795.5	2012.2
NIQE	8.8785	8.1136	8.3777	6.4755	5.1461

A detailed comparative analysis was carried out against notable enhancement models including BIMEF, CRM, LIME, among others. In terms of performance metrics, the proposed model consistently outperformed these existing methods, highlighting its efficacy in enhancing low-light images while maintaining essential details.

Performance Metrics

SSIM: Demonstrated a significant improvement, indicating better preservation of structural details.

LOE and LOE_ref: These metrics confirmed the model's ability to maintain lightness order consistency, reinforcing the perceptual quality of the enhanced images.

The model also scores well on the Naturalness Image Quality Evaluator (NIQE) metric, implying that the processed images are more natural and visually pleasing. The comprehensive evaluation demonstrates not only the model's ability to significantly enhance the quality of the poorly-lit images but also its superiority over existing enhancement methods.

The use of sophisticated datasets, robust training protocols, and a wide array of evaluation metrics collectively affirm the potential of the proposed system in addressing the challenges inherent in image enhancement

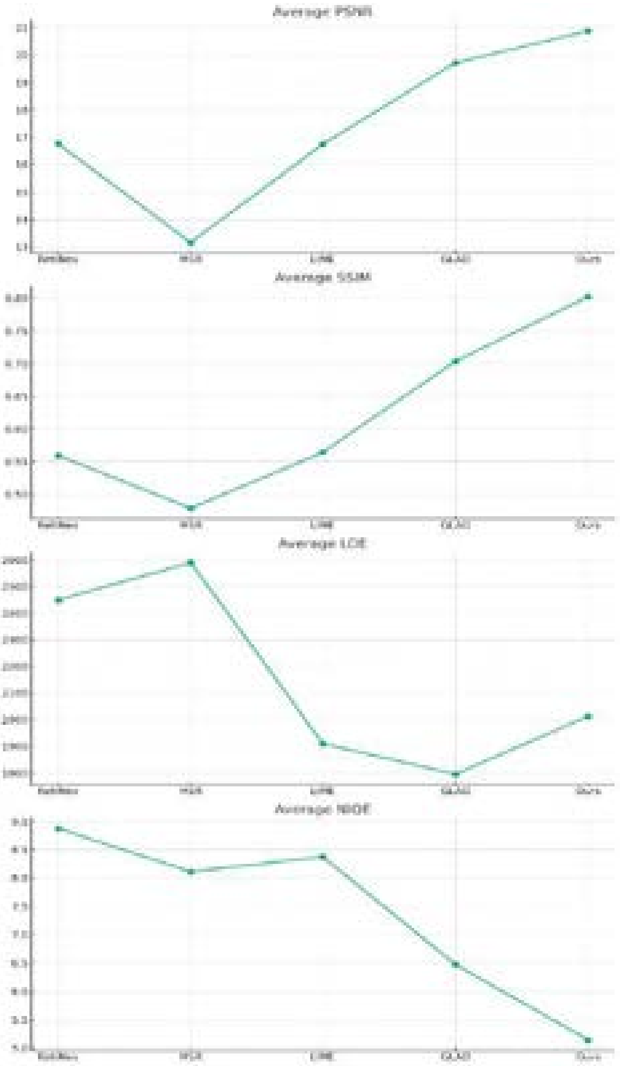


Figure 4.2. Model Comparison

V. CONCLUSION

In conclusion, despite these significant achievements, the study acknowledges practical limitations that must be addressed to enhance the framework's applicability and effectiveness further. One such limitation is the computational demand of the proposed model, especially when processing ultra-high-resolution images in real-time scenarios. This constraint could limit the deployment of the system in resource-constrained environments, such as mobile devices or embedded systems, where computational efficiency is crucial. Additionally, while the model excels in handling a wide range of low-light conditions, its performance can vary across extremely diverse lighting situations, particularly in scenarios with highly unpredictable or non-uniform light sources. This variation highlights the need for further refinement in the model's ability to generalize across different low-light environments.

Future research could focus on optimizing the algorithm's computational efficiency, perhaps through the integration of more lightweight neural network architectures or the development of specialized hardware accelerators. Moreover, expanding the training dataset to include a broader spectrum of low-light conditions could improve the model's robustness and generalizability. This expansion would entail not just more images but also a greater diversity in terms of scene complexity, lighting conditions, and subjects.

Exploring integration with emerging technologies like augmented reality (AR) and virtual reality (VR) could open new frontiers for the application of low-light image enhancement, offering immersive and visually enriched experiences even in poorly lit environments. In sum, the path forward for low-light image enhancement research is rich with opportunities for innovation and improvement. By addressing the outlined limitations and harnessing the potential of new methodologies and technologies, future work can build on the solid foundation laid by this study to further advance the state of the art in image processing.

REFERENCES

[1] Ali, M., Yin, B., Bilal, H., Kumar, A., Shaikh, A. M., & Rohra, A. (2023). Advanced efficient strategy for detection of dark objects based on spiking network with multi-box detection. *Multimedia Tools and Applications*, 1-21.

[2] Azam, M. A., Khan, K. B., Salahuddin, S., Rehman, E., Khan, S. A., Khan, M. A., ... & Gandomi, A. H. (2022). A review on multimodal medical image fusion: Compendious analysis of medical modalities, multimodal databases, fusion techniques and quality metrics. Computers in biology and medicine, 144, 105253

[3] Fan, S., Liang, W., Ding, D., & Yu, H. (2023). LACN: A lightweight attention-guided ConvNeXt network for low-light image enhancement. Engineering Applications of Artificial Intelligence, 117, 105632

[4] Fattal, R., Lischinski, D., & Werman, M. (2023). Gradient domain high dynamic range compression. In Seminal Graphics Papers: Pushing the Boundaries, Volume 2 (pp. 671-678).

[5] Fu, Y., Hong, Y., Chen, L., & You, S. (2022). LE-GAN: Unsupervised low-light image enhancement network using attention module and identity invariant loss. Knowledge-Based Systems, 240, 108010.

[6] Guo, P., Jia, M., Guo, D., Wang, Z. L., & Zhai, J. (2023). Retina-inspired in-sensor broadband image preprocessing for accurate recognition via the flexophototronic effect. *Matter*, *6*(2), 537-553.

[7] Hai, J., Xuan, Z., Yang, R., Hao, Y., Zou, F., Lin, F., & Han, S. R2rnet: Low-light image enhancement via real-low to real-normal network. Journal of Visual Communication and Image Representation, 90, 103712.

[8] Hang, Y., Boryczka, J., & Wu, N. (2022). Visible-light and near-infrared fluorescence and surface-enhanced Raman scattering point-of-care sensing and bio-imaging: A review. Chemical Society Reviews, 51(1), 329-375.

[9] Li, Y., Zhang, Y., Timofte, R., Van Gool, L., Tu, Z., Du, K., ... & Zhang, Y. (2023). NTIRE 2023 challenge on image denoising: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1904-1920).

[10] Liu, X., Ma, W., Ma, X., & Wang, J. (2023). LAE-Net: a locally-adaptive embedding network for low-light image enhancement. Pattern Recognition, 133, 109039.

[11] *Liu, X., Xie, Q., Zhao, Q., Wang, H., & Meng, D. (2023). Low-Light Image Enhancement by Retinex-Based Algorithm Unrolling and Adjustment. IEEE Transactions on Neural Networks and Learning Systems*.

[12] *Ma, L., Ma, T., Liu, R., Fan, X., & Luo, Z. (2022). Toward fast, flexible, and robust low-light image enhancement. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 5637-5646)*.

[13] Mildenhall, B., Hedman, P., Martin-Brualla, R., Srinivasan, P. P., &

Barron, J. T. (2022). Nerf in the dark: High dynamic range view synthesis from noisy raw images. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 16190-16199).

[14] *Murugan, R. A., & Sathyabama, B. (2022). Object Detection for Night Surveillance Using Ssan Dataset Based Modified Yolo Algorithm in Wireless Communication. Wireless Personal Communications, 1-14*

[15] Ramesh, G., Logeshwaran, J., Gowri, J., & Mathew, A. (2022). The management and reduction of digital noise in video image processing by using transmission based noise elimination scheme. ICTACT Journal on Image & Video Processing, 13(1).

[16] Sharma, P., Bisht, I., & Sur, A. (2023). Wavelength-based attributed deep neural network for underwater image restoration. *ACM Transactions on Multimedia Computing, Communications and Applications*, *19*(1), 1-23.

[17] Vasluianu, F. A., Seizinger, T., & Timofte, R. (2023). Wsrd: A novel benchmark for high resolution image shadow removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 1825-1834).

[18] Wang, Y., Huang, K., Fang, J., Yan, M., Wu, E., & Zeng, H. (2023). Mid-infrared single-pixel imaging at the single-photon level. *Nature Communications*, *14*(1), 1073.

[19] Wang, T., Zhang, K., Shen, T., Luo, W., Stenger, B., & Lu, T. (2023, June). Ultra-high-definition low-light image enhancement: a benchmark and transformer-based method. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 37, No. 3, pp. 2654-2662)

[20] Wibowo, A., Sheikh, M. A. K., Diguna, L. J., Ananda, M. B., Marsudi, M. A., Arramel, A., ... & Birowosuto, M. D. (2023). Development and challenges in perovskite scintillators for high-resolution imaging and timing applications. *Communications Materials*, *4*(1), 21.

[21] Yin, K., Hsiang, E. L., Zou, J., Li, Y., Yang, Z., Yang, Q., ... & Wu, S. T. (2022). Advanced liquid crystal devices for augmented reality and virtual reality displays: principles and applications. Light: Science & Applications, 11(1), 161.

[22] Z. Zhao, B. Xiong, L. Wang, Q. Ou, L. Yu and F. Kuang, "RetinexDIP: A Unified Deep Framework for Low-Light Image Enhancement," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 32, no. 3, pp. 1076-1088, March 2022, doi: 10.1109/TCSVT.2021.3073371.

Cyberbullying Image Classification using Artificial Intelligence for Safer Online Platform

M. Vamshi Krishna[1]
Student, Department of Computing Technologies,
School of Computing,
College of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur Chennai,India.
mv7444@srmist.edu.in

R. Asish Verma[2]
Student, Department of Computing Technologies,
School of Computing,
College of Engineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur Chennai,India
rv5411@srmist.edu.in

P. Kirubanantham[*3]
Assistant Professor, Department of ComputingTechnologies,
School of Computing,
College ofEngineering and Technology,
SRM Institute of Science and Technology,
Kattankulathur, Chennai,India
kirubanp2@srmist.edu.in

Abstract— **The toll that cyberbullying takes on victims' life is incalculable because a person's approach to it is highly individualized. For some victims, the message might be harsh, but for others, it might be normal. It is quite difficult to find reliable content because of the ambiguities in cyberbullying communications. Several research works have reportedly used a text- content-based methodology to answer this question. However, picture-based cyberbullying detection is now receiving less attention. Enhancing a methodology that permits the prevention of social media posts on image-based cyberbullying issues is the aim of this job. This study provides support for an automated version that searches a malicious social media network for a cyber-symbol image, relying just on switch mastery. By applying fashion knowledge, it is feasible to infer hidden contextual functions from texts containing cyberbullying. Our test includes two photo datasets: pictures of cyberbullying and pictures of non-bullying. The datasets might be helpful to Destiny researchers in expanding their research. It's challenging to find a pleasant and suitable version to recognize bullying photographs, therefore experiment with both transfer the experiment's findings showed that switch learning models are anexcellent method for predicting conversations that are only based on images and involve cyberbullying.**

Keywords: Safety and security system, Mobile v2 net network, Transfer Learning.

I. INTRODUCTION

The emergence of Internet 2.0 led to a massive shift in social linguistic interchange, restructuring friendships and relationships. Children who spend a lot of time online and in niche social networks are exposed to hazards and social offenders, including cyberbullying, in addition to all the positive aspects of this type of activity.

Cyberbullying must be understood and addressed from distinct angles. To a large extent, automatic incident detection and prevention can help to tackle this problem. There are already apps that aim to assist victims and tools that allow you to report bullying instances. Additionally, the majority of websites that teens may use frequently offer protection features. A few of examples are Twitter Safety and Security and YouTube Safety Center, which both providespoken exchanges and user assistance. Numerous studies on automaton detection and cyberbullying prevention have also been done; we will discuss these topics in further depth in the following phase. But more work needs to be done before a conclusive answer to this problem is discovered. In a similar vein, DNN-based models have also been utilised to recognise instances of cyberbullying.

Because online social media is becoming more and more prevalent, bullying has changed. It is defined as a coordinated, aggressive act by several individuals. Employing repeated and sustained electronic forms of engagement against a victim who is powerless to defend themselves.

By spreading untrue rumors and disclosing private information with the intention of hurting and discrediting the victims, cyberbullying can take many different forms, ranging from text messages to images and videos. For victims, the invasions are made more scary and intimidating by the inability to recognize the behavior of cyberbullies. Furthermore, the ramifications (discussed below) of cyberbullying are infinite given the virtual environment in which they operate and their capacity to participate in these actions consistently without regard to time.

Cyberbullying primarily targets teenagers. A 2014 Pew Research Center research found that 73% of internet users had witnessed online abuse and 40% of users had personally experienced it. Deliberate embarrassment and disparaging name were the most common forms of harassment observed by respondents in a poll covering five different types of bullies. The information for additional bullies . The graph alsoincludes statistics on cyberbullying victims. Another chart which compares the ratio of young people to all Internet users, shows that the ratio among young people is 10% to 20% higher than usual.

II. LITERATURE REVIEW

[1] A growing number of studies are focused on dealing with cyberbullying, primarily on its identification, as the detrimental effects it causes are becoming more well recognized. Non-technical publications focus on reporting the current situation, providing definitions, and comprehending the issue of cyberbullying. These research provide guidance for our detection efforts. While the number of studies devoted to automatically identifying cyberbullying is currently small, it is growing.

[2] Numerous studies have been conducted in the social sciences regarding the causes and prevalence of cyberbullying, particularly with regard to children and young adults. Regarding its effects, empirical research has shown a connection between teenage suicidal thoughts and cyberbullying experiences. The distinctive profiles of cyberbullying victims and perpetrators are shown in. The potential preventative and intervention measures are also covered in this essay. These investigations broaden our

understanding of the issue and increase public awareness of it.

[3] Text analysis is the main focus of most technical investigations. By incorporating sentiment and contextual variables into the baseline text mining system, Yin et al. improved the system's ability to detect online harassment. With regard to the Kongregate, Slashdot, and Myspace datasets, the analysis provides up to 50% accuracy. A similar study found that bullying remarks on YouTube could be identified with an accuracy of 80% by using textual context variables. Using a limited Formspring dataset, Reynolds etal. looked at the linguistic patterns of online bullies and victims in and created rule-based classification methods to identify the bullies. They can distinguish between 78.5% of the bullying posts with their model. Xu and others. To explore further the potential of language features in cyberbullying detection, a new sentence- level filtering model was developed that employs grammatical links between words to semantically remove bullying phrases from texts. Experiments on a YouTube dataset show that 90% of the model's outputs are comparable to hand filtering.

[4] To aid in the detection, other techniques are employed in addition to pure text analysis. In order to create a rule-based classification model that would identify instances of cyberbullying in online interactions, Bayzick et al. used a truth set. On Myspace, they created software that can identify windows that are innocent 51.91% of the time and windows that include cyberbullying 85.30% of the time. In order to improve detection performance as a supervised classification problem, user data—including the user's past remarks and attributes—is taken into consideration. Using the YouTube dataset, they achieved 75% accuracy.

[5] While The aforementioned research is largely concerned with detecting bullies. Our goal in this work is distinct. As previously indicated, we try to forecast if a picture shared on social media would be the target of bullies. Consequently, we concentrate on stopping individuals from feeding the bullies rather than trying to identify the bullies. Furthermore, we use image features to construct the classification model rather than text data. A model that can alert users of social networking sites before they upload any photos that are extremely "bullyable" is the anticipated result.

[6] But as far as we are aware, no previous research has used picture attributes to detect cyberbullying. In this study, we investigate the possibilities of already used image feature extraction techniques in various fields and adapt them to our needs. Zerr et al. successfully produced results from a private-aware picture classification investigation employing semantic image attributes. The colour histogram, the edgedirection coherent vector, the SIFT bag of visual words, and the facial characteristics are among the features that are involved. These characteristics show promise in Zerr's work for semantic image description. Our goal is to develop amodel that can identify photographs that have a high potential for bullying by utilizing these factors.

[7] An image-based deep learning algorithm for identifying cyberbullying based on convolutional neural networks (CNN):Sung-Hwan Jung and Seong-Whan Lee's study suggested a CNN- based deep learning algorithm for social media platform cyberbullying image identification. The model may be applied as a tool for early identification of cyberbullying and prevention because of its excellent accuracy. (Lee and Jung, 2018)

[8] Multi-Modal Approach for Cyberbullying Detection in Social Media: To detect cyberbullying in social media, Ramaswamy et al. suggested a multi-modal approach that blends textual and visual elements. CNN and long short-term memory (LSTM), two deep learning techniques, are used in the model. to attain remarkable precision in identifying instances of cyberbullying. The Ramaswamy group in 2019

[9] Automated Cyberbullying identification in Instagram photos: Using deep learning methods like CNN and transfer learning, a study by Jafri et al. presented an automated cyberbullying identification system for Instagram photos. The model identified cyberbullying photos on Instagram with a high degree of accuracy. (Jafri and others, 2021)

[10] Cyberbullying Detection System for Social Media Using Deep Learning Images: In a paper published in 2015, Khan et al. presented a deep learning-based cyberbullying detection system that analyzes social media images using recurrent neural networks (RNN) and CNN. When it came to identifying cyberbullying photos on social media, the model performed remarkably well. (Khan and others, 2021).

III. PROPOSED METHOD

The main goal is to create a model that aids in preventing concerns with image-based cyberbullying on social media sites. First a convolutional neural network with deep learning is employed to build the model. This study uses transfer learning models later on. The outcomes of the trials conducted with several hyper-parameter configurations verified that the transfer learning-based model is the best choice for this problem. In the best scenario, the suggested model's accuracy was satisfactory, showing that the system can identify the majority of posts that involve cyberbullying. The transfer learning model used in this paper is MobileNet V2, which uses the transferred data from the trained model and predicts weather the given images is a cyberbullying image or not. There are various algorithms which can be used in this project but MobileNet V2 is most suitable choose for cyberbullying image classification.

IV. EXISTING SYSTEM

Previous research on the topic was restricted to the binary text classification of cyberbullying. A few academics have focused on the role classification of cyberbullying participants based on textual patterns, based on the literature. The role identification model used by Sui was the first research to categorise Twitter tweets into the following roles: reporter, accuser, outsider, reinforcer, defender, and helper. Additionally, Chatzakou and associates created. Cyberbullying detection algorithms based on Twitter datasets classified as aggressors, bullies, spammers, or normal are designed to analyse the traits of bullies and aggressors using text-based features and Twitter metadata. However, the previously cited study did not account for victims or bystanders. Jacobs et al. developed a cyberbullying role categorization approach that uses the AMiCA dataset to ascertain whether a post displays context from the harasser, victims, or bystanders inside cyberbullying events,

acknowledging the importance of researching victim and bystander context. Besides the Rathnayake et al. created two using the same dataset. Two types of Bert-based categorization exist: the defending model and the bullying model. Postings from the harasser and bystander assistant are categorized using the former model, while postings from the victim and bystander defender are categorized using the latter.

Van Hee et al. tackled the problem of multiclass text classification in addition to role classification, developing a model that utilises the Ask.fm platform dataset to categorize more subtle types of cyberbullying, such as threats or blackmail, sexual discussion, insults, curses or exclusions, defence, defamation, and encouragement. Cyberbullying severity level classification is a different type of multiclass text classification. A cyberbullying severity classification methodology was created by Sugandhi et al. that establishes the post's severity rating (high, medium, and low).

V. METHODOLOGY

This study proposes a new framework for binary text classification, or the identification of cyberbullying, utilizing both transfer learning and traditional machine learning techniques. The six primary steps of the text classification job are: gathering the cor- pus, feature extraction, feature collection, and text preprocessing, which was replaced by data preparation once the data was acquired. The framework begins with preprocessing the text and preparing the data, then splits into two directions to complete the job concurrently using traditional methods for transfer learning and machine learning The model is assessed in the end. Both classic machine learning and transfer learning approaches were based on Pre-trained Language Models (PLMs). Beyond traditional machine learning, additional feature engineering work was required to transform the text input into measurable characteristics before feeding it into the model. technique. The workflows for each step of the suggested technique are further explained in the subsections that follow.

MODULES

- Data collection
- Data preprocessing
- Model training
- Model testing
- Performance evaluation
- Flask framework
- Output
- prediction

DATA COLLECTION:

The data was acquired so that the model could detect instances of regular and cyberbullying. If there are offensive images of people with faces in them, for example, the example of a commonplace human face is about bullying. A wide enough range of record instances were added to each category in order to prevent situations like this one.

DATA PREPROCESSING:

Convolutional Neural Networks (CNs) can do amazing things if they have enough data. That being said, it can be hard to find the right education information collection that has all the properties needed for training. In the event that the user doesn't have enough, the network can load study materials. Photos that look real have clear positions, sizes, zoom, lighting, noise, and other features

For a big enough group of these frequently occurring items, the Data Augmentation method is used. The community will see these things happen when they rotate photos to different angles, flip photos using different axes, or rotate and crop photos. There may be a useful set of arguments in the "Image Data Generator" that lets Keras be used to improve images. You can totally change the order of the factors by flipping or stacking the photo vertically, and the system will still work. You can do that with Numpy.

When you want to get extra-large facts, you can also rotate, which is an increase, at smaller or minute angles. Most of the time, this is done at 90-degree angles. There's a fix for the rotation's background color that makes it look like it happened in the past. If that doesn't happen, the model can also think that the trade line is separate. This works great, but all photos that have been turned around are the same amount of complicated. The width_shift_range and height_shift_range options are the two most important ones in the Keras Image Data Generator class. The photo factors will move all the way to the left or right based on those numbers. The best way to learn is to focus on one repressor or one binary classifier per class. Because of this, it is very important to label different types of music as either belonging to or not belonging to a certain genre. LabelBinarizer's update method makes this process easier.

- After turning it into an array, normalize it to a range of [0,1].
- Get photos and class names out of files.
- Formatted according to keras
- Make changes to fit the model
- Add a picture
- Put the class label in.
- Based on our model, put labels into groups of either 0 or 1.
- Make labels binary

MODULE TRANING:

Thin bottleneck layers serve as the input and output of the residual blocks in the MobileNetV2 architecture, which has an inverted residual structure.

Additionally, it filters features in the expansion layer using lightweight convolutions. Ultimately, non- linearities in the narrow layers are eliminated. The photos are loaded using the data frame with the flow_from_dataframe function. Theprecise location of the photos is specified by the directory argument. The independent and dependent variables in this scenario are the labels and the images, or x_col and y_col. class_mode="binary" indicates that there are just two different classes in the data Using target_size=(224,224), a 224 x 224 image will be produced. How many photos are sampled at once is known as the batch size. In addition, we will initialize the base model with an input size of 224 × 224, which corresponds to the pre-processed image data we already have. The weights from ImageNet will be the same in the basic model. The pre- trained model's top layers will be excluded by setting include_top=False, which is perfect for feature extraction.

With an inverse residual form, the input and output residual blocks of the MobileNetV2 structure are made up of thin bottleneck layers. Additionally, it distributes the functions across the growth layer using modest convolutions. Lastly, nonlinearities in narrow layers are eliminated. The typical architecture appears as follows:

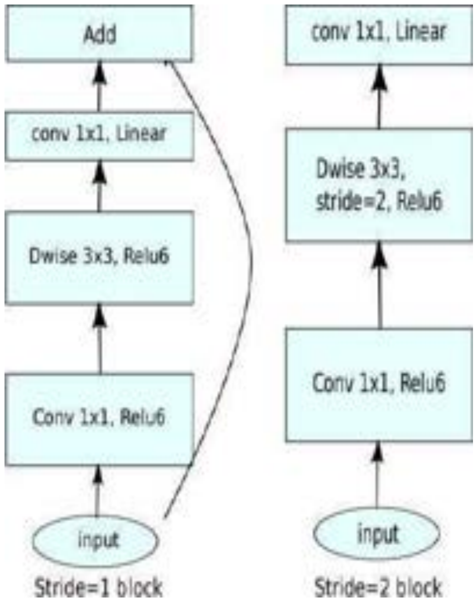


Fig.:1 Flow diagram of MoblieNetV2

Transfer learning models are useful for prediction work in many domains and have the advantage of current deep learning architectures. As a result, this study also anticipates message- based cyberbullying by utilizing the advantages of pre-structured transfer learning styles. Initially, the chosen dataset was subjected to a variety of switch styles, including as VGG16, MobilenetV2, and others that were included in the Keras toolkit. The experimental findings of different models have shown that Mobilenet-V2 performs better than other types. He became one of the most successful architects in the ILSVRC project that same year. We continued our investigation using Mobilinet-V2, a transfer learning method that has been extensively utilized for picture identification and is based on the sphere's builders, selection, creation of models, and assessment of performance. Since a secondary dataset, the initial stage of the corpus, was employed in this study.

Fig.2: Bullying dataset images

Fig.:3 Non- Bullying dataset images

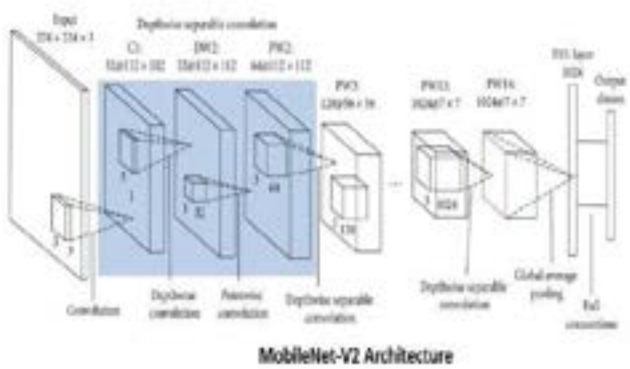


Fig.:4 MoblieNet V2 Architecture

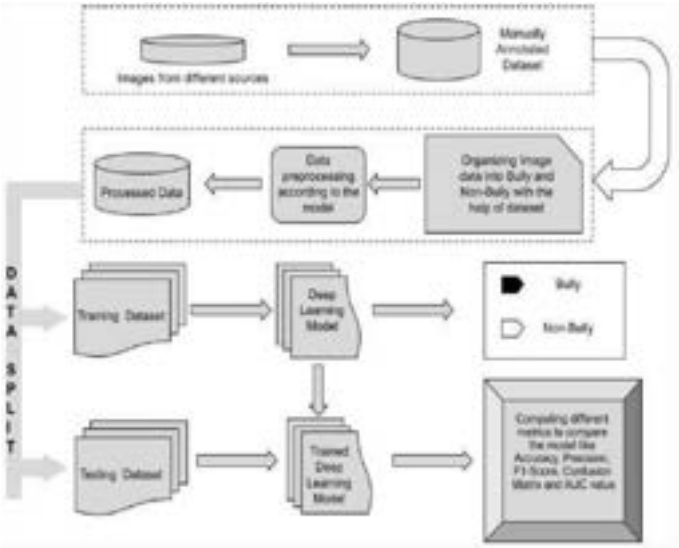


Fig.5: Block Diagram

MODEL TESTING:

The parameters used in the model for determining the accuracy, macro avg, weighted avg.

Accuracy of Epoch-1: 50.98

Loss	Validation-loss	Validation-accuracy
1.5678	1.6236	0.524

Table 1 Values of Loss, Validation-loss and Validation-accuracy of epoch-1.

Epoch 2:

Accuracy of Epoch-2: 66.78

Loss	**Validation-loss**	**Validation-accuracy**
1.0727	**1.9288**	**0.4865**

Table 2 Values of Loss, Validation-loss and Validation-accuracy of epoch-2.

Epoch 3:

Accuracy of Epoch-3 : 73.87

Loss	**Validation-loss**	**Validation-accuracy**
0.8434	**1.6485**	**0.5447**

Table 3 Values of Loss, Validation-loss and Validation-accuracy of epoch-3.

Epoch 4:

Accuracy of Epoch-4 : 79.40

Loss	**Validation-loss**	**Validation-accuracy**
0.6532	**1.5286**	**0.5825**

Table 4 Values of Loss, Validation-loss and Validation-accuracy of epoch-4.

Epoch 5:

Accuracy of Epoch-5 : 83.56

Loss	**Validation-loss**	**Validation-accuracy**
0.5262	**1.3982**	**0.6003**

Table 5 Values of Loss, Validation-loss and Validation-accuracy of epoch-5.

Epoch 6:

Accuracy of Epoch-6: 86.45

Loss	**Validation-loss**	**Validation-accuracy**
0.4352	**1.4796**	**0.6005**

Table 6 Values of Loss, Validation-loss and Validation-accuracy of epoch-6.

Epoch 7:

Accuracy of Epoch-7: 89.52

Loss	**Validation-loss**	**Validation-accuracy**
0.3455	**1.7481**	**0.5669**

Table 7 Values of Loss, Validation-loss and Validation-accuracy of epoch-7.

Epoch 8:

Accuracy of Epoch-8 : 89.34

Loss	**Validation-loss**	**Validation-accuracy**
0.3369	**1.6481**	**0.5987**

Table 8Values of Loss, Validation-loss and Validation-accuracy of epoch-8.

Epoch 9:

Accuracy of Epoch-9 : 90.39

Loss	Validation-loss	Validation-accuracy
0.3029	**1.7984**	**0.5722**

Table 9 Values of Loss, Validation-loss and Validation-accuracy of epoch-9.

Epoch 10:

Accuracy of Epoch-10 : 92.59

Loss	Validation-loss	Validation-accuracy
0.2372	**2.0757**	**0.5370**

Table 10 Values of Loss, Validation-loss and Validation-accuracy of epoch-10.

PREDICTION:

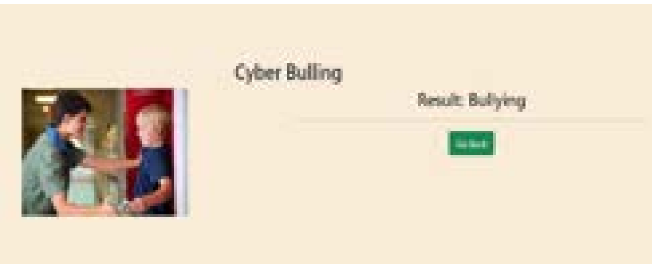


Fig.6: Model Prediction Output

Fig.7: Model Prediction output NO

VI. CONCLUSION

To summarize, the action recognition project has shown that Convolutional Neural Network (CNN) architecture, MobileNet is effective in properly identifying human activities from photos. By conducting thorough experimentation and evaluation, we have acquired useful insights into the performance of various designs and their appropriateness for the task of action recognition. This study has gathered a comprehensive dataset consisting of images that portray a wide variety of human actions. This was done to guarantee that the trained models had resilience and demonstrate a high degree of generalizability. This study has employed data preparation methods, such as resizing, rescaling, noise reduction, and contrast enhancement, to improve the quality and clarity of the pictures. This method enhances the efficiency of model training. The training and validation phases followed conventional machine learning methods, where metrics like accuracy, precision, recall, and loss were continuously monitored to evaluate the model's performance and drive optimization efforts. The evaluation yielded favorable outcomes, as every design demonstrated exceptional accuracy and proficiency in discerning certain human actions. Moreover, the analysis of performance measures across multiple designs yielded significant observations regarding their unique advantages and disadvantages. MobileNet demonstrated exceptional computing efficiency and maintained constant performance. NasNet demonstrated outstanding accuracy, whereas ConvNext showed promising results in terms of precision and recall. The suggested model's limitations include (i) It does not keep track of text-based cyberbullying identities. Meaning that this study does not include messages that contain only text; and (ii) messages that combine text and photos in an attempt to perpetrate cyberbullying. Nevertheless, this study is restricted to the identification of cyberbullying using photographs. As a result, due to the variety of issues, the study's future scope is always up for discussion. A written post can be seen next to an image to find more instances of cyberbulling on social media platforms.

REFERENCES

[1] B. Cagirkan and G. Bilek, ''Cyberbullying among Turkish high school students,'' *Scandin. J. Psychol.*, vol. 62, no. 4, pp. 608–616, Aug. 2021, doi: 10.1111/sjop.12720.

[2] P. T. L. Chi, V. T. H. Lan, N. H. Ngan, and N. T. Linh, ''Online time, experience of cyber bullying and practices to cope with it among high school students in Hanoi,'' *Health Psychol. Open*, vol. 7, no. 1, Jan. 2020, Art. no. 205510292093574, doi: 10.1177/2055102920935747.

[3] A. López-Martínez, J. A. García-Díaz, R. Valencia-García, and A. Ruiz-Martínez, ''CyberDect. A novel approach for cyberbullying detection on Twitter,'' in *Proc. Int. Conf. Technol. Innov.*, Guayaquil, Ecuador: Springer, 2019, pp. 109–121, doi: 10.1007/978-3-030-34989- 9_9.

[4] R. M. Kowalski and S. P. Limber, ''Psychological, physical, and academic correlates of cyberbullying and traditional bullying,'' *J. Adolescent Health*, vol. 53, no. 1, pp. S13–S20, Jul. 2013, doi: 10.1016/j.jadohealth.2012.09.018.

[5] Y.-C. Huang, ''Comparison and contrast of piaget and Vygotsky's theo- ries,'' in *Proc. Adv. Social Sci., Educ. Humanities Res.*, 2021, pp. 28–32, doi: 10.2991/assehr.k.210519.007.

[6] A. Anwar, D. M. H. Kee, and A. Ahmed, ''Workplace cyberbullying and interpersonal deviance: Understanding the mediating effect of silence and emotional exhaustion,'' *Cyberpsychol., Behav., Social Netw.*, vol. 23, no. 5, pp. 290–296, May 2020, doi: 10.1089/cyber.2019.0407.

[7] D. M. H. Kee, M. A. L. Al-Anesi, and S. A. L. Al-Anesi, ''Cyberbul- lying on social media under the influence of COVID-19,'' *Global Bus. Organizational Excellence*, vol. 41, no. 6, pp. 11–22, Sep. 2022, doi: 10.1002/joe.22175.

[8] I. Kwan, K. Dickson, M. Richardson, W. MacDowall, H. Burchett, C. Stansfield, G. Brunton, K. Sutcliffe, and J. Thomas, ''Cyberbullying and children and young people's mental health: A systematic map of systematic reviews,'' *Cyberpsychol., Behav., Social Netw.*, vol. 23, no. 2,pp. 72–82, Feb. 2020, doi: 10.1089/cyber.2019.0370.

[9] R. Garett, L. R. Lord, and S. D. Young, ''Associations between social media and cyberbullying: A review of the literature,'' *mHealth*, vol. 2, p. 46, Dec. 2016, doi: 10.21037/mhealth.2016.12.01.

[10] M. Ptaszynski, F. Masui, Y. Kimura, R. Rzepka, and K. Araki, ''Automatic extraction of harmful sentence patterns with application in cyberbullying detection,'' in *Proc. Lang. Technol. Conf.* Poznań, Poland: Springer, 2015, pp. 349–362, doi: 10.1007/978-3-319-93782-3_25.

[11] M. Ptaszynski, P. Lempa, F. Masui, Y. Kimura, R. Rzepka, K. Araki, M. Wroczynski, and G. Leliwa, ''"Brute-force sentence pattern extortion from harmful messages for cyberbullying detection,''' *J. Assoc. Inf. Syst.*, vol. 20, no. 8, pp. 1075–1127, 2019.

[12] M. O. Raza, M. Memon, S. Bhatti, and R. Bux, ''Detecting cyber- bullying in social commentary using supervised machine learning,'' in *Proc. Future Inf. Commun. Conf.* Cham, Switzerland: Springer, 2020, pp. 621–630.

[13] D. Nguyen, M. Liakata, S. Dedeo, J. Eisenstein, D. Mimno, R. Tromble, and J. Winters, ''How we do things with words: Analyzing text as social and cultural data,'' *Frontiers Artif. Intell.*, vol. 3, p. 62, Aug. 2020, doi: 10.3389/frai.2020.00062.

[14] J. Cai, J. Li, W. Li, and J. Wang, ''Deeplearning model used in text classification,'' in *Proc. 15th Int. Comput. Conf. Wavelet Act. Media Technol. Inf. Process. (ICCWAMTIP)*, Dec. 2018, pp. 123–126, doi: 10.1109/ICCWAMTIP.2018.8632592.

[15] N. Tiku and C. Newton. *Twitter CEO: We Suck at Dealing With Abuse*. Verge. Accessed: Aug. 17, 2022. [Online]. Available: https://www.theverge.com/2015/2/4/7982099/twitter-ceo-sent-memo-taking-personal-responsibility-for-the D. Noever, ''Machine learning suites for online toxicity detection,'' 2018, *arXiv:1810.01869*.

[16] Yadav, M. Kumar, and D. Yadav, ''Cyberbullying detection from tweets using deep learning,'' *Kybernetes*, vol. 51, no. 9, pp. 2695–2711, Sep. 2022.

[17] A. Bozyiğit, S. Utku, and E. Nasibov, ''Cyberbullying detection: Utilizing social media features,'' *Expert Syst. Appl.*, vol. 179, Oct. 2021, Art. no. 115001, doi: 10.1016/j.eswa.2021.115001.

[18] H.-S. Lee, H.-R. Lee, J.-U. Park, and Y.-S. Han, ''An abusive text detection system based on enhanced abusive and non-abusive word lists,'' *Decis. Support Syst.*, vol. 113, pp. 22–31, Sep. 2018, doi: 10.1016/j.dss.2018.06.009.

[19] Y. Fang, S. Yang, B. Zhao, and C. Huang, ''Cyberbullying detection in social networks using bi-GRU with self-attention mechanism,'' *Information*, vol. 12, no. 4, p. 171, Apr. 2021, doi: 10.3390/info12040171.

[20] D. G. Krutka, S. Manca, S. M. Galvin, C. Greenhow, M. J. Koehler, and E. Askari, ''Teaching 'against' social media: Confronting prob- lems of profit in the curriculum,'' *Teachers College Rec., Voice Scholarship Educ.*, vol. 121, no. 14, pp. 1–42, Dec. 2019, doi: 10.1177/016146811912101410.

Driver Distraction Detection based on Deep Learning Techniques using Images

Mohanapriya S[1, a)], Mohana Saranya S[1], Dinesh K[1], Sivasankar B [2], Vignesh R G [2], Vishnu Kumar K[2]

mohanapriyas.cse@kongu.edu, mohanasaranya.cse@kongu.edu, dinesh.cse@kongu.edu
sivasankarb.20cse@kongu.edu, vigneshrg.20cse@kongu.edu,
vishnukumark.20cse@kongu.edu

[1]Assistant Professor, Department of Computer Science and Engineering, Kongu Engineering College, Erode, Tamilnadu, India
[2] UG Scholar, Department of Computer Science and Engineering, Kongu Engineering College Erode, Tamilnadu, India

[1, a)] Corresponding Author: mohanapriyas.cse@kongu.edu

Abstract

The alarming increase in road fatalities demands immediate action. In 2019 alone, over 1.35 million lives were tragically lost on the world's highways. Addressing distracted driving, a key contributor to this surge, is crucial for accident prevention. While prior research has explored Deep Learning for distracted driver detection in vehicles, there is still room for advancement in identifying drivers whose focus is diverted from the road, a specific type of distraction. This study proposes a Deep Learning-based classification system to detect drivers looking away from the road. It leverages Jupyter Notebook and Python to implement and evaluate ResNet50, VGG16, and InceptionV3 network architectures. The State Farm dataset, encompassing ten distinct categories of driver behavior, is utilized for training and testing. These categories include actions like talking on the phone, texting while driving, eating, drinking, reaching behind for objects, adjusting the radio, grooming, talking to passengers, and looking elsewhere. The focus of this study is on improving detection accuracy specifically for the "looking elsewhere" class. Model performance is assessed using confusion matrices and accuracy metrics. The results reveal that the ResNet50 model achieved a noteworthy 95% accuracy in classifying drivers looking elsewhere. VGG16 and Inception models also exhibited strong performance, reaching accuracies of 89% and 92% respectively.

Keywords – Deep Learning, Resnet50, VGG (Visual Geometric Group), InceptionV3.

I. INTRODUCTION

Distracted driving continues to be a major threat on our roads, contributing to a significant number of accidents, injuries, and fatalities every year Road traffic accidents remain a significant global public health concern, claiming millions of lives every year. The United States paints a similarly concerning picture. According to the NHTSA (National Highway Traffic Safety Administration), distracted driving was a factor in over 3,100 deaths and an estimated 424,000 injuries in crashes that occurred in 2019 [source: Centers for Disease Control and Prevention - Distracted Driving]. Among these fatalities, a concerning trend is the number of pedestrians and cyclists involved. The NHTSA reports that nearly 20% of people killed in distracted driving crashes in 2019 were not even inside a vehicle [source: Centers for Disease Control and Prevention - Distracted Driving]. These statistics highlight the urgent need for effective solutions to curb distracted driving and enhance road safety for everyone.

One of the most prevalent forms of distracted driving involves the use of electronic devices behind the wheel. Talking, Texting on the phone, and even manipulating navigation apps can significantly divert a driver's attention from the road, increasing the risk of accidents example is shown in Figure 1. However, distracted driving encompasses a broader range of behaviors beyond just electronic devices. Eating, drinking, grooming, adjusting the radio, or simply taking your eyes off the road for a moment can all have devastating consequences.

Traditional methods for addressing distracted driving have relied on public awareness campaigns and law enforcement. While these efforts are important, they haven't been entirely successful in curbing the problem. New approaches are needed to effectively detect and prevent distracted driving. This is where artificial intelligence (AI) and specifically, deep learning, come into play. Deep learning algorithms have the potential to analyze visual data from in-vehicle cameras and identify drivers exhibiting signs of distraction.

One promising avenue for developing real-time interventions lies in CNN (Convolutional Neural Networks) – a powerful type of deep learning architecture. This study explores the effectiveness of three specific CNN models: ResNet50, VGG16, and Inception. These models have demonstrated exceptional performance in image recognition tasks, making them well-suited for analyzing driver behavior captured by in-vehicle cameras. By evaluating the performance of these models in detecting a specific type of distraction – drivers looking away from the road – this study aims to contribute to the development of more robust and accurate distracted driver detection systems.

Figure 1. Example Driver Distraction Image

II. LITERATURE SURVEY

A. Driver Drowsiness Detection

Drowsiness is a critical factor contributing to road accidents. References [1,2,18] focus on diverse methods for detecting driver drowsiness using sensors. The authors emphasize the effectiveness of physiological signals, such as heart rate and electroencephalogram (EEG) data, in accurately identifying drowsiness compared to relying solely on the observation of driving behavior. This information can be utilized to trigger alerts or initiate interventions to prevent accidents.

B. Anomaly Detection for Safe Driving

[4,18] introduce an unsupervised anomaly detection method employing a conditional generative adversarial network (GAN). The methodology involves examining data streams obtained from sensors monitoring driver physiology (e.g., heart rate) and vehicle behavior (e.g., steering wheel angle) to detect deviations from typical driving patterns. The identification of such anomalies may signify potential hazardous situations, enabling timely intervention to avert accidents.

C.Classification with VGG16

The utilization of VGG16, a deep convolutional neural network architecture, has become pivotal in the realm of image classification [10]. With its 16-layer structure comprising 13 convolutional and 3 fully connected layers, VGG16 excels in extracting hierarchical features, making it particularly effective for complex visual recognition tasks [15]. The model's widespread adoption is attributed to its superior performance, especially when pre-trained weights are leveraged through transfer learning [17,19]. VGG16 has demonstrated significant success in various applications, including medical imaging and object recognition, solidifying its role as a cornerstone in the evolution of convolutional neural networks for image classification.

D. Classification with InceptionV3

The widespread adoption of InceptionV3, an advanced convolutional neural network architecture, has significantly influenced the landscape of image classification [12]. Characterized by its innovative inception modules, which utilize multiple filter sizes in parallel, InceptionV3 excels in capturing intricate patterns and features within images. With a focus on computational efficiency and improved accuracy, this model has proven effective in various domains, including medical image analysis and scene recognition [13]. The versatility of InceptionV3 is further emphasized when leveraging pre-trained weights through transfer learning, contributing to its reputation as a robust and efficient solution for image classification tasks [14].

E. Classification with Resnet50

The prominence of ResNet50, a deep residual neural network architecture, has profoundly impacted the field of image classification [11]. Distinguished by its innovative use of residual blocks, ResNet50 addresses the challenges of training very deep networks by facilitating the flow of information through shortcut connections. This design not only enhances model convergence but also allows for the training of exceedingly deep networks, leading to superior accuracy in image classification tasks. The adaptability of ResNet50 is evident in its successful application to diverse domains, such as object recognition and facial expression analysis [16]. Leveraging pre-trained weights through transfer learning further solidifies ResNet50's reputation as a robust and effective solution for complex image classification challenges [20].

III. PROPOSED WORK

3.1 CLASSIFICATION SYSTEM

In image recognition systems, convolutional neural networks (CNNs) excel at categorizing input images. CNNs achieve this by utilizing convolutional layers to identify specific characteristics within the image, pooling layers to reduce complexity, and fully-connected layers to capture the connections between these features. The data is then transformed and fed into an output layer with a softmax activation function, which generates probabilities for each category.

3.2 DATA PREPROCESSING AND AUGUMENATATION

The first step involves data preprocessing, which prepares the raw data for training the models. This typically includes tasks like:

Resizing: To ensure consistency and facilitate efficient processing, all frames might be resized to a standard size (e.g., 224x224 pixels).

Normalization: The pixel values in the images, typically ranging from 0 to 255, can be normalized to a common range (example 0 to 1) for improved training performance.

Data augmentation: goes beyond basic preprocessing by artificially creating additional variations of the existing frames. This enriches the dataset and helps the models learn from a wider range of scenarios. Here are some common data augmentation techniques applicable to driver distraction detection where sample is shown Figure.2

- **Rotation range:** Specifies a degree value (0-180) for the extent of random image rotations.
- **Height shift range:** Dictates the extent of horizontal and vertical shifts.
- **Width shift range:** Governs the extent of horizontal and vertical shifts.
- **Shear range:** Determines the intensity of shear (angle in degrees, counterclockwise).
- **Zoom range:** Specifies a range for random zoom effects.
- **Horizontal flip:** Introduces random horizontal flipping of input images. However, this feature is unsuitable for our case as it may alter the image classes.

Figure 2. Different types of Augmentation

3.3 DEEP LEARNING MODEL SELECTION

Focusing on CNNs (Convolutional Neural Networks) due to their proficiency in analyzing spatial features in images, this work evaluates three prominent architectures: ResNet50, VGG16, and Inception. The selection process prioritizes three key factors. First, achieving the highest classification rate for the target behavior (looking away from the road) is paramount. Metrics like overall accuracy will be used to identify the most accurate model. Second, computational efficiency is crucial for real-time applications in driver distraction detection systems. The inference speed of each model will be assessed on the target hardware platform to ensure suitability. Finally, generalizability for real-world scenarios is essential. Data augmentation techniques will be employed to improve the models' ability to adapt to variations not present in the training data.

Each evaluated architecture offers distinct advantages:

- **ResNet50:** This architecture incorporates residual learning to deal with the vanishing gradient problem, a common challenge in deep neural networks, allowing for deeper networks with improved training accuracy. Additionally, bottleneck blocks reduce computational cost while maintaining representational power, making it potentially well-suited for real-time applications. Architecture of Resnet50 is shown in Figure 3.

- **VGG16:** Relying on stacked convolutional layers with small filters, VGG16 excels at capturing intricate spatial details within images, potentially allowing for more precise identification of specific driver actions like looking away from the road.

- **InceptionV3:** Inception utilizes inception modules that combine convolutional filters of various sizes within a single layer. This allows the network to capture features at multiple scales simultaneously, potentially enhancing its ability to identify both overall body posture and specific actions relevant to driver distraction.

3.4 MODEL TRAINING AND TESTING

The three candidate CNN architectures – ResNet50, VGG16, and Inception – will be implemented within the Kera's framework. Each model will be trained using an appropriate optimizer (Example: Adam) and loss function (example: categorical cross-entropy)

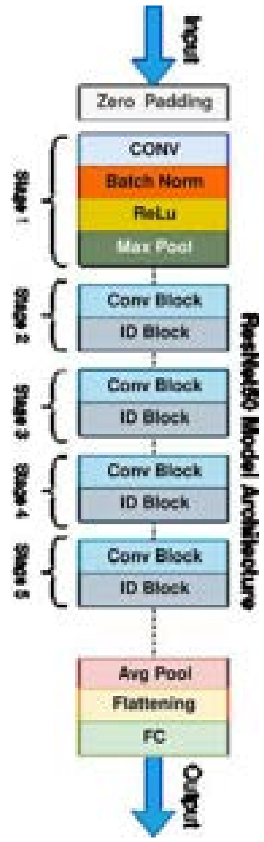


Figure 3. Resnet50 Architecture

The validation set will be used to monitor training progress and fine-tune hyperparameters like learning rate and number of epochs. This helps prevent overfitting, a phenomenon where the model performs poorly on unseen data but well on the training data. Training metrics like accuracy, loss, and F1-score will be tracked on both the training and validation sets during the training process. Monitoring these metrics allows for early detection of potential issues and guides hyperparameter tuning for optimal performance.

Following training, the models will be evaluated on a hold-out test set (not used during training) to assess their generalizability to unseen real-world scenarios.

3.5 ROBUSTNESS ANALYSIS

The evaluation results will determine the model offering the optimal balance between accuracy, computational efficiency (measured through inference speed on a representative hardware platform), and generalizability. This chosen model will be integrated into the driver distraction detection system. In scenarios where real-time applications are prioritized, a slight decrease in accuracy might be acceptable in exchange for significantly faster processing speeds. By leveraging the capabilities of Kaggle and the Keras framework, this work facilitates the effective training and evaluation of CNN models for driver distraction detection using the State Farm dataset. This approach contributes to the development of robust and generalizable systems, ultimately enhancing road safety.

Figure 4. Sample class in the Dataset

3.6 DATASET

The State Farm Distracted Driver Detection dataset is a collection of images aimed at training and testing machine learning models for identifying driver behavior as shown in Figure 4. It was created by State Farm, an insurance company, to explore the potential of using in-car cameras to automatically detect distracted driving and potentially prevent accidents. Dataset Distribution is shown in Figure 5.

The dataset consists of 2D images captured from a dashboard camera, showing the driver and the car's interior. Each image is labeled with one of ten classes representing the driver's activity:

- **c0:** Safe driving
- **c1-c4:** Texting or talking on the phone (left or right hand)
- **c5:** Operating the radio
- **c6:** Drinking
- **c7:** Reaching behind
- **c8:** Applying hair and makeup
- **c9:** Talking to a passenger

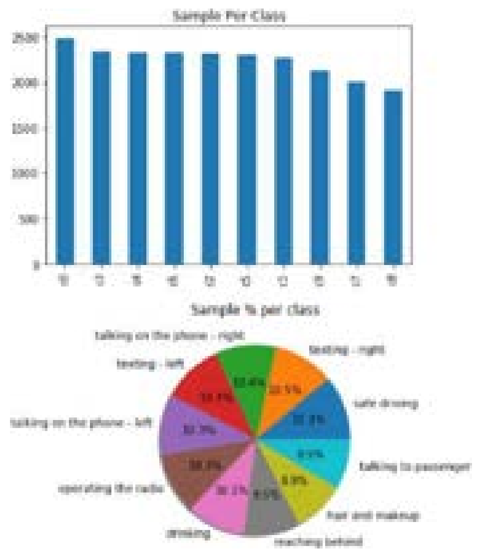


Figure 5. Dataset Distribution over different classes

IV. RESULTS AND DISCUSSION

ACCURACY

Accuracy, a widely used metric for assessing classification model performance, measures the ratio of correct predictions to total predictions. It is typically computed using the formula as shown in Equation 1.

$$Accuracy = \frac{Number\ of\ correct\ predictions}{Total\ number\ of\ predictions} * 100 \quad (1)$$

Here,

Number of Correct Predictions: It signifies the count ofdata points where the predicted class aligns with the actual class.

Total Number of Predictions: This is the total number of data points the model makes predictions on.

Accuracy comparison between VGG16, InceptionV3 and Resnet50 is shown in the Figure 6 from which we can infer that using Resnet50 can give better accuracy. Accuracy over epochs is shown in Figure 7. The accuracy comparison between ResNet50, Inception V3 and VGG16 is shown in Table I. Accuracy of Resnet50 graph is shown inFigure 6.

Table I. Accuracy Comparison between ResNet50 vs Inception V3 vs VGG16

Epochs	ResNet50 (%)	InceptionV3 (%)	VGG16 (%)
0	10.2	9.2	9.4
5	23.4	17.5	15
10	37.3	25.6	20.2
15	50.6	32.1	35.6
20	63.4	42.3	47.1
25	72.3	51.7	58.9
30	85	60.8	69.4
35	88.2	69.3	78.2
40	92	78.2	84.8
45	93	84.3	86.9
50	94	91.5	88.4

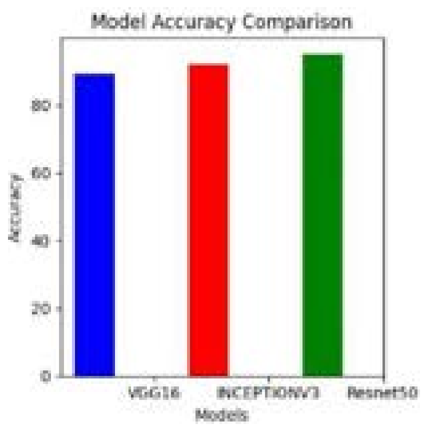


Figure 6. Model Accuracy Comparison

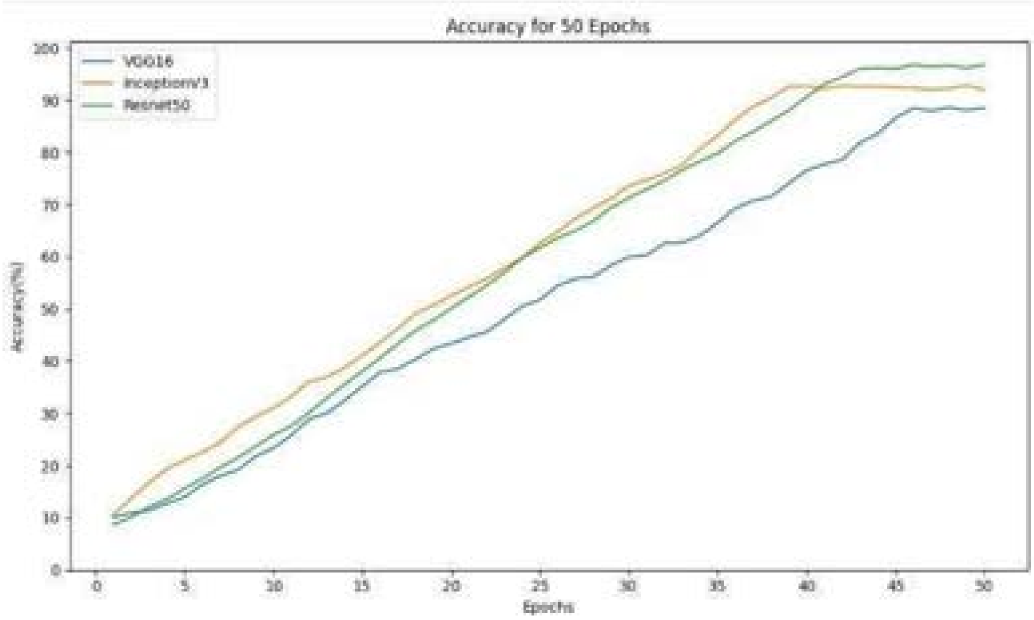


Figure 7. Accuracy over epochs

V. CONCLUSION AND FUTURE WORK

In conclusion, the study demonstrates the effectiveness of Deep Learning models, particularly ResNet50, VGG16, and Inception, in addressing the specific issue of detecting drivers looking away from the road—a critical aspect of distracted driving prevention. The focus on the "looking elsewhere" class within the State Farm dataset has yielded promising results, with the ResNet50 model achieving an impressive 95% accuracy, while the VGG16 and Inception models also performed strongly with accuracies of 89% and 92%, respectively.

These findings underscore the potential of leveraging advanced Deep Learning techniques to enhance the accuracy of distracted driver detection systems. As the prevalence of distracted driving continues to pose a significant risk to road safety, the development of robust and precise models becomes increasingly crucial. To further improve these models, future research should explore avenues such as fine-tuning pre-trained models on more extensive datasets or employing transfer learning approaches. Investigate these strategies could prove valuable for adapting the models to diverse driving environments and demographics, thereby contributing to the ongoing efforts to mitigate the risks associated with distracted driving and bolster overall road safety.

REFERENCES

1. Chowdhury, A., Shankaran, R., Kavakli, M., &Haque, M. M. (2018). Sensor applications andphysiological features in drivers’ drowsinessdetection: A review. IEEE Sensors Journal, 18, 3055–3067.
2. V. Unadkat, P. Sayani, H. Kapadia, P. Shah and H. Dalvi, "Automated System for Detecting Distracted Driver," 2018 4th International Conference on Computing Communication and Automation (ICCCA), Greater Noida, India, 2018
3. Brookhuis, K. A., De Waard, D., & Janssen, W. H. (2001). Behavioural impacts of advanced driver assistance systems–an overview. European Journal of Transport and Infrastructure Research .
4. Qiu, Y., Misu, T., & Busso, C. (2022). Driving anomaly detection using conditional generative adversarial network.
5. A. Walavalkar, S. Singh, R. Salian and V. Shrivastava, "Driver Distraction Monitoring and Alerting System," 2019 1st International Conference on Advances in Information Technology (ICAIT), Chikmagalur, India, 2019.
6. Connor, J. L., Norton, R., Robinson, E., Connor, J., Ameratunga, S., Civil, I., Dunn, R., Bailey, J., & Jackson, R. (2002). See profile papers driver sleepiness and risk of serious injury to car occupants: Population based case control study. https://www.researchgate.net/publication/11366926.
7. Demestichas, K., Alexakis, T., Peppes, N., & Adamopoulou, E. (2021). Comparative analysis of machine learning-based approaches for anomaly detection in vehicular data. Vehicles, 3, 171–186. https://doi.org/10.3390/vehicles3020011
8. Dong, S., Khattak, A., Ullah, I., Zhou, J., & Hussain, A. (2022). Predicting and analyzing road traffic injury severity using boosting-based ensemble learning models with shapley additive explanations. International Journal of Environmental Research andPublic Health, 19, 2925.

9. Gaikwad, D., & Thool, R. C. (2015). Intrusion detection system using bagging with partial decision treebase classifier. Procedia Computer Science, 49, 92–98.
10. J. Tao, Y. Gu, J. Sun, Y. Bie and H. Wang, "Research on vgg16 convolutional neural network feature classification algorithm based on Transfer Learning," 2021 2nd China International SAR Symposium (CISS),Shanghai, China, 2021.
11. V. Sharma and N. Singh, "Deep Convolutional Neural Network with ResNet-50 Learning algorithm for Copy-Move Forgery Detection," 2021 7th International Conference on Signal Processing and Communication (ICSC), Noida, India, 2021.
12. Xiaoling Xia, Cui Xu and Bing Nan, "Inception-v3 for flower classification," 2017 2nd International Conference on Image, Vision and Computing (*ICIVC)*, Chengdu, 2017.
13. P. Ahammed, M. F. Faruk, N. Raihan and M. Mondal, "Inception V3 Based Transfer Learning Model for the Prognosis of Acute Lymphoblastic Leukemia from Microscopic Images," 2022 4th International Conference on Electrical, Computer & Telecommunication Engineering (ICECTE), Rajshahi, Bangladesh, 2022.
14. S. Prasher, L. Nelson and S. Gomathi, "Inception V3 model for Tuberculosis detection using chest x-ray images," 2023 3rd International Conference on Intelligent Technologies (CONIT), Hubli, India, 2023.
15. N. Zakaria, F. Mohamed, R. Abdelghani and K. Sundaraj, "VGG16, ResNet-50, and GoogLeNet Deep Learning Architecture for Breathing Sound Classification: A Comparative Study," 2021 International Conference on Artificial Intelligence for Cyber Security Systems and Privacy (AI-CSP), *El Oued, Algeria, 2021*.
16. R. Kumar, D. Singh, A. Chug and A. P. Singh, "Evaluation of Deep learning based Resnet-50 for PlantDisease Classification with Stability Analysis," 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2022.
17. S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," *2015* 3rd IAPR Asian Conference on Pattern Recognition *(ACPR)*, Kuala Lumpur, Malaysia,2015.
18. Sahayadhas, A., Sundaraj, K., & Murugappan, M. (2012). Detecting driver drowsiness based on sensors: A review. 10.3390/s121216937.
19. S. Tripathi and R. Kumar, "Image Classification usingsmall Convolutional Neural Network," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2019.
20. V. Tiwari, C. Pandey, A. Dwivedi and V. Yadav, "Image Classification Using Deep Neural Network," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2020.

Gesture Controlled Lawn Mower

Mr.SatheeshKumar K
Assistant Professor
Department of Electrical and Electronics Engineering,
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore,Tamil Nadu - 641402
satheeshkumargceb@gmail.com

Gokul Prasanth S
UG Student
Department of Electrical and Electronics Engineering,
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore,Tamil Nadu - 641402
kit.24.20bee013@gmail.com

SunielKumar S
UG Student
Department of Electrical and Electronics Engineering,
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore,Tamil Nadu - 641402
kit.24.20bee044@gmail.com

Riyashini S
UG Student
Department of Electrical and Electronics Engineering,
KIT-Kalaignarkarunanidhi Institute of Technology,
Coimbatore,Tamil Nadu - 641402
kit.24.20bee310@gmail.com

Abstract— This paper presents the design and implementation of a gesture controlled lawn mower. Previous Lawn Mowers are difficult to control. The remote control may not be as responsive as a gesture control lawnmower, and it can be difficult to control the mower in tight spaces or around obstacles. . Remote control lawn mowers are more limited in their flexibility. The lawn mower can be controlled to move forward, backward, turn left, turn right, and start / stop by using hand gestures.The gesture controlled lawn mower has several advantages over traditional lawn mowers. Gesture control mowers eliminate the need for the close proximity to the blades, which can be dangerous.It can be controlled by simple hand gestures, so the users stay a safe distance away.Gesture control mowers are very convenient to use. Users can start, stop, and navigate the mower with their hands, so the users don't have to fumble with buttons or switches.The lawn mower can also be programmed to follow a specific path using a gesture.It is more efficient way to keep lawns looking their best.Gesture control mowers can be more friendly than traditional lawn mowers.With the help of gesture controlled lawn mower the operator can cut the grasses as he/she walking and it helps by keeping them active.

Keywords – Gesture, Lawn mower, unmanned aerial vehicles, switches.

I. Introduction

Gesture control is a promising technology with the potential to revolutionize the way we interact with devices. The primary thing of gesture- actuated systems is to enhance stoner experience, convenience, and availability by furnishing a more intuitive and natural way of interacting with technology.These systems have set up operations in a wide range of fields, including gaming, virtual reality, stoked reality, consumer electronics and more.Gesture control is becoming increasingly popular as it offers a more intuitive and natural way to interact with devices.Humans interact in the physical world by the means of the five senses. Still, gestures have been an important means of communication in the physical world from ancient times, indeed before the invention of any language.

II. EXISTING SYSTEM

The Existing System of this Paper is Remote Controlled Lawn mower. The remote-controlled lawnmower is the paper's current system. This kind of lawnmower is entirely embedded system based. Using the remote electric lawn mower, they are now able to overcome labor costs and time loss. This paper gives the user the ability to remotely operate the lawnmower in any direction, change the speed of the mower engine, and change the height at which the grass is mowed. The electric lawn mower work with Remote controlled by using raspberry pi3. The design of electric lawn mower is based on GPS tracking system for the purpose to identify the location of the mower. It is the wireless transmission data linked. In The work station of grass cutter can be several components prohibited. The design of electric lawn mower is based on GPS tracking system for the purpose to identify the location of the mower. It is the wireless transmission data linked. The device can be considered for the unmanned aerial vehicles and space exploration. It is used for tiring and time consuming for workers. The formation of embedded System is focused on remote controlled. To operate to correct direction guided to device and cut the focused areas from land. When the location of mower can be identified to movement of mower by using remote controlled. For those who have certain physical limitations, this may be very troublesome. They can be susceptible to interference. They can be difficult to control. The remote control may not be as responsive as a gesture control lawnmower, and it can be difficult to control the mower in tight spaces or around obstacles. . Remote control lawn mowers are more limited in their flexibility.

III . PROPOSED SYSTEM

In this Gesture Controlled Lawn mower, the Arduino has interpreted these movements, it will instruct the motor driver, which will move the lawnmower accordingly. Python or C++ can be used to write the program for the Arduino and the gesture recognition sensor. This would be utilized to regulate the motors of the lawn mower in terms of speed and direction. A receiver is a circuit that receives signals from a wired or wireless transmission media, decodes them, or converts them into a form that can power nearby circuits. The signals from the sensors are received by the receiving circuit of a gesture-controlled lawnmower, which then transforms them into digital data. The lawn mower's microcontroller then processes the digital data to as certain the user's intended actions and adjusts the mower's motions accordingly. The grass cutter incorporates Gesture technology for Bluetooth control via a mobile device.

Users can operate the mower's movement through their mobile phones, utilizing Gesture-enabled controls. Additionally, ultrasonic sensors are employed to detect obstacles in the mower's path. Upon detection, the vehicle autonomously adjusts its direction to avoid collisions. This functionality is facilitated by a relay to DC motor setup, where the L293D driver board manages the DC motors' operation, enabling precise control and maneuverability. Overall, the proposed setup combines Gesture technology, Bluetooth control, obstacle detection using ultrasonic sensors, and a relay to DC motor setup for efficient grass cutting with autonomous navigation capabilities . It enables users to operate the mower wirelessly from Gesture – enabled mobile devices, providing convenience and flexibility in directing grass cutting operations.Streamlines lawn maintenance tasks by automating grass cutting operations, reducing the need for manual labor and saving time for home owners and landscaping professionals.Incorporates intelligent obstacle avoidance features to minimize the risk of damage to the mower and surrounding objects, enhancing overall safety during operation.The reliability will be achieved by Developing advanced gesture recognition algorithms that can reliably identify specific gestures associated with controlling the lawn mower, ensuring consistent and intuitive operation.

BLOCK DIAGRAM

RECEIVER

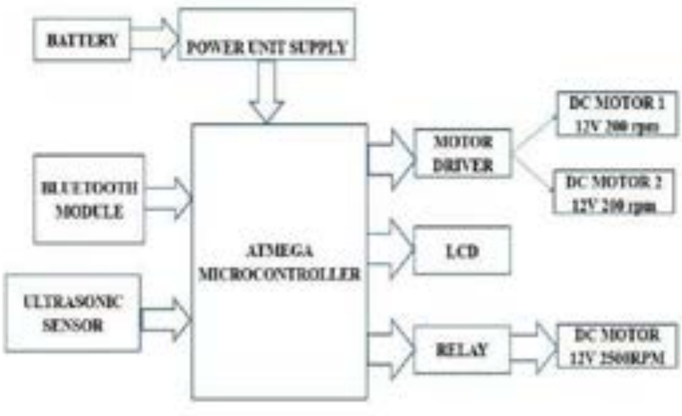


Fig 1 : Receiver Block Diagram

The receiver circuit is the heart of a gesture-controlled lawn mower. It is responsible for decoding the signals from the transmitter circuit and sending corresponding signals to the mower's motors shown in Fig 1. The receiver circuit is typically mounted on the lawn mower itself and consists of a decoder, a microcontroller, and a motor driver.

TRANSMITTER

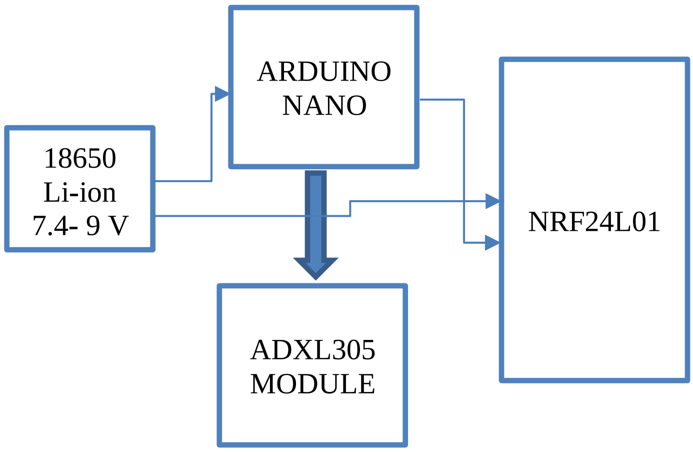


Fig 2:Transmitter Block Diagram

A gesture - controlled lawn mower uses a transmitter circuit to send signals to a Receiver circuit that controls the mower's motors . The transmitter circuit is typically mounted on glove or armband that the user wears shown in Fig 2. The glove or armband contains sensors that detect the user's hand gestures. The transmitter circuit then encodes these gestures into electrical signals that are sent to the receiver circuit.

IV. HARDWARE COMPONENTS

A. Arduino Nano

The heart of the Arduino Nano is the ATmega328P microcontroller. It has 32KB of flash memory for storing the program, 2KB of SRAM, and 1KB of EEPROM for data storage. The Arduino Nano has a total of 14 digital input/output pins. Of these6 can be used for pulse-width modulation (PWM) output. It has 6 analog input pins that can also be used as digital inputs. These pins are capable of reading analog signals v, making them suitable for sensors and other analog devices. The ATmega328P operates at 16 MHz, providing sufficient processing power for a Wide range of applications. The Arduino Nano can be powered via USB connection, an external power supply, or through the VIN pin using a regulated power source. It features a USB interface that allows it to be connected to a computer for programming and serial communication. The Arduino Uno is based on open- source hardware and software. This means that the design files and source code are freely available for modification and distribution. The Arduino Nano shown in Fig 6.1 has a large and active community of users and developers. There are extensive online resources, forums and tutorials available for assistance.

Fig 3 : Arduino Nano

B.NRF24L01

The NRF24L01 is a single chip 2.4GHz transceiver with an embedded baseband protocol engine (Enhanced ShockBurst), suitable for ultra-low power wireless applications shown in Fig 4. It is designed for operation in the world-wide ISM frequency band at 2.400 - 2.4835GHz. The NRF24L01 is a versatile and powerful wireless communication module that can be used to create a wide variety of applications. It is easy to use and program, and it has a low cost. The NRF24L01 can operate on a single 3.3V supply and has a typical power consumption of 12mA in transmission mode and 900nA in power down mode.It can support data rates of up to 2Mbps.It can achieve a range of up to 100 meters in open space **.**

Fig 4: NRF24L01

C . L298NMotor Driver

The L298N is a dual H-bridge motor driver IC that is capable of driving two DC motors independently. It is a popular choice for many hobbyist and robotics projects due to its low cost and ease of use. The L298N is typically used in conjunction with a microcontroller to control the speed and direction of the motors shown in Fig 5. The microcontroller can send signals to the enable, direction, and PWM inputs of the L298N to control the motors. The microcontroller sends signals to the enable, direction, and PWM inputs of the L298N to control the motors. The L298N then drives the motors based on the signals it receives from the microcontroller.

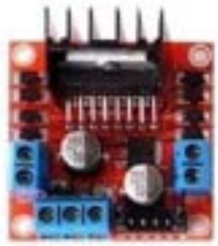


Fig 5 : L298NMotor Driver

D . PICController

A PIC controller is a microcontroller manufactured by Microchip Technology PIC controllers are known for their low cost, ease of use, and wide range of features . They are used in a wide variety of applications, including consumer electronics, industrial automation, and automotive systems shown in Fig 6.PIC controllers are based on the RISC (reduced instruction set computing) architecture. This means that they have a small number of simple instructions, which makes them easy to program and efficient to execute.

Fig 6: PICController

E. Gyroscope sensor

Gyroscope sensor is a device that can measure and maintain the orientation and angular velocity of an object. These are more advanced than accelerometers shown in Fig 7. These can measure the tilt and lateral orientation of the object whereas accelerometer can only measure the linear motion. Gyroscope sensors are also called as Angular Rate Sensor or Angular Velocity Sensors. These sensors are installed in the applications where the orientation of the object is difficult to sense by humans. This structure comprises a stationary part in the center with „Sensing Arm‟ attached to it and „Drive Arm‟ on both sides.This double-T-structure is symmetrical.When an alternating vibration electrical field is applied to the drive arms, continuous lateral vibrations are produced. As Drive arms are symmetrical, when one arm moves to left the other moves to the right, thus canceling out the leaking vibrations. This keeps the stationary part at the center and sensing arm remains static.

Fig 7: Gyroscope sensor

F. ADXL33

The ADXL335 module is an accelerometer sensor module that uses the Analog Devices ADXL335 accelerometer chip shown in Fig 8. The ADXL335 is a small, low-power, 3-axis accelerometer that can measure acceleration in three dimensions (X, Y, and Z). This module provides analog voltage outputs for each of the three axes, which can be read using analog-to-digital converters (ADCs) on microcontrollers like the Arduino. The ADXL335 typically has a measuring range of ±3g, which means it can detect accelerations from -3g to +3g in each axis. The ADXL335 module is commonlyused in combination with other sensors like gyroscopes to create Inertial Measurement Units (IMUs) for tracking orientation and motion in devices like drones and robots.

Fig 8: ADXL33

G. TT Gear Motor

TT Gear Motor is a small, high torque gear motor that is commonly used in a variety of applications, including robots, drones, and other DIY projects. It has a gear ratio of 4.1:8, which means that the motor shaft will rotate 4.1 times for every revolution of the input shaft. This gives the motor a high torque output, making it ideal for applications where a lot of power is needed. TT Gear Motors are a versatile and powerful type of gear motor that can be used in a wide variety of applications. They are a popular choice for hobbyists and professionals alike.

H. Screw Terminal

A screw terminal is a type of electrical connector that is commonly used to secure and connect wires or cables to various devices or components. Screw terminals provide a mechanical means of making electrical connections. They consist of a metal terminal block with a threaded hole and a screw. To make a connection, a wire or cable is inserted into the terminal block, and the screw is tightened down. The exposed wire is then inserted into the terminal block. Screw terminals provide a reliable and secure electrical connection.

V. SIMULATION & HARDWARE RESULT

CATIA offers tools for creating sheet metal parts and designing them for manufacturing. CATIA seamlessly integrates with computer-aided manufacturing (CAM) software, allowing the direct transfer of designs to CNC machines for manufacturing. It incorporates generative design capabilities, which use algorithms to generate and optimize designs based on specified constraints, often resulting in innovative and efficient designs. CATIA is an extensive software suite with various modules and tools, making it suitable for a wide range of industries from aerospace and automotive to consumer goods and industrial equipment. Install the CATIAV5 2023 Version>Setup the Catia file>Create New file.

In CATIA Software, First select the motion study 1 and Click Switch to basic motion. Then go to settings and increase the geometric accuracy and modeling capacity. After that increase the time lapse in motion study table to 8 – 10 sec. Then press / add rotary motor to the four tires in forward direction. Then add Linear motor to the edge of the lawn of the lawnmower and edge of the road to make move. The simulation outputs are shown in Fig 9,10.

Fig 9 Simulation Output - 1

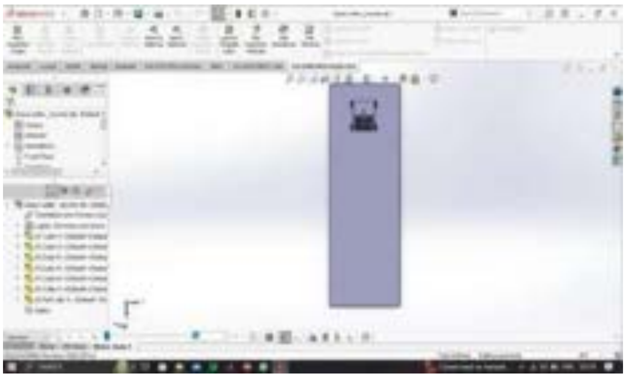


Fig 10 Simulation Output - 2

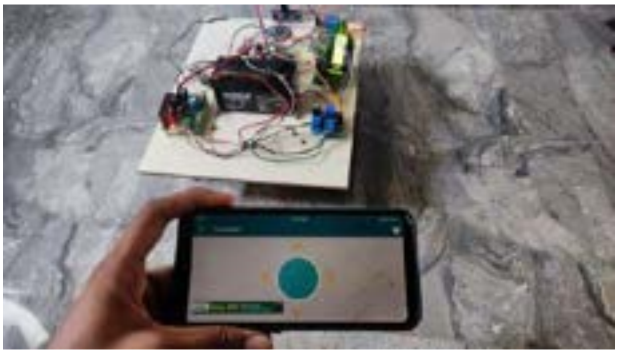


Fig 11 Movement Recognizer

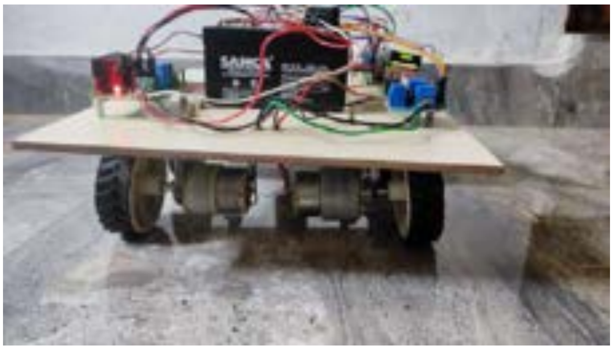


Fig 12 Hardware Output

VI.CONCLUSION

The developed grass cutter uses advanced technology to make landscaping easier. It enables users to operate the mower wirelessly from Gesture-enabled mobile devices, providing convenience and flexibility in directing grass cutting operations. The proposed model streamlines lawn maintenance tasks by automating grass cutting operations and saving time for homeowners and landscaping professionals. Incorporates intelligent obstacle avoidance features to minimize the risk of damage to the mower and surrounding objects, enhancing overall safety during operation. With features like Gesture technology and Bluetooth, users can control the mower from a distance, making lawn more convenient.

REFERENCES

[1] Gesture commands for controlling high-level UAV behavior SNJohn Akagi T. DevonMorris-Brady Moon – Xingguang Chen³ - Cameron K. Peterson – 2021

[2] Asst . Prof. R.S.Ramya, C.Nivetha, V.Priyatharshini, P.Pavithra –"Remote Controlled Electric Lawn Mower using Embedded System", "International Journal of Science, Engineering and Technology"-2020

[3] Hand gesture control and design of a rotary pneumatic soft actuator using leap motionsensor,Victoria O guntosin. Ademola Abdulkareem - 2020

[4] Parthipan G, Dhiyaneswaran J,Vimal raja M, Rajesh K, Sanjeeviraj K - "Fabrication of Wireless Remote Controlled Electric Lawn Mower " , " International Journal of Innovative Technology and Exploring Engineering (IJITEE)"-2019

[5] Ashin Saji , Geon G Bastian , Muhammad Jaseel KA , Sanu Sajeevan - "Design and Implementation of Intelligent Lawn Mower Robot " , "International Journal of Innovative Science and Research Technology" - 2018

[6] Wireless hand gesture controlled multiple musical instruments,Pranshu Jhamb,ArvindRehali-2018

[7] K. D. Sowjanya, R. Sindhu, M. Parijatham, K.Srikanth and P. Bhargav, "Multi purpose autonomous agricultural robot "International conference of Electronics,Communication and aerospace Technology (ICECA) ",Coimbatore, 2017

[8] Swarna Prabha Jena , Sworaj Kumar Nayak, Saroj Kumar Sahoo, Sibu Ranjan Sahoo, Saraswata Dash ,Sunil Kumar Sahoo–"Accelerometer Based Gesture Controlled Robot Using Arduino " , "International Journal Of Engineering Sciences & Research Technology"-2015

[9] Dipin.A, Chandrasekhar . T.K – "Solar – powered vision - based robotic lawnmower", "International Journal of Engineering Research and Reviews " – 2014

[10] [10] Tanimola, O. A, Diabana, P. D , Bankole, Y. O– "Design and Development of a Solar Powered Lawn Mower " , " International Journal of Scientific & Engineering Research"-2014

[11] [11] Aswath S , Chinmaya Krishna Tilak , Amal S suresh and GaneshaUdupa, "Human Gesture Recognition for Real-Time Control Humanoid Robot " , International Journalof Advances in Mechanical and Automobile Engineering (IJAMAE)-2014

[12] [12] Gesture - controlled interactive three Dimensional anatomy, Jordan B Hochman, Bertram Unger , Jay Kraut, Justyn Pisa and Sabine Hombach-Klonisch – 2014

Enhanced Security Fencing System with Geolocation Tracking

Medha Wyawahare
Electronics and Telecommunication Department
Vishwakarma Institute of Technology,Pune
medha.wyawahare@vit.edu

Milind Rane
Electronics and Telecommunication Department
Vishwakarma Institute of Technology,Pune
milind.rane@vit.edu

Dnyanesh Ujalambkar
Electronics and Telecommunication Department
Vishwakarma Institute of Technology,Pune
dnyanesh.ujalambkar22@vit.edu

Sakshi Mahadik
Electronics and Telecommunication Department
Vishwakarma Institute of Technology,Pune
Sakshi.mahadik23@vit.edu

Prathamesh Jadhav
Electronics and Telecommunication Department
Vishwakarma Institute of Technology,Pune
prathamesh.jadhav231@vit.edu

Nisha Wakode
Electronics and Telecommunication Department
Vishwakarma Institute of Technology,Pune
nisha.wakode22@vit.edu

***Abstract*—In an era of increasing security concerns and the need for comprehensive surveillance systems, the integration of fencing systems with contact connecting mapping has emerged as a promising solution. This research paper provides an in-depth analysis of the design, implementation, and benefits of a device that integrates fencing systems along with contact connecting mapping technology. This study examines the potential applications of this integration in 2 major sectors: health care and jail asylums providing an in-depth solution for the same. Additionally, the paper discusses the challenges like security threats to other people around them along with their own safety, attempt of escaping to places they shouldn't be and the potential solutions associated with the deployment and maintenance of such integrated systems.**

Keywords—Fencing Systems, Contact Connecting Mapping, Security, Surveillance, Integration, Boundary demarcation, Geo-fencing, GPS (Global Positioning System), GSM (Global System for Mobile Communication) Integration.

I. Introduction

In contemporary society, the imperative for robust security and surveillance systems has grown significantly. Fencing systems, traditionally serving as physical barriers for diverse purposes, have undergone a transformation by integrating advanced technologies to amplify their efficiency. Concurrently, the prominence of contact-connecting mapping has surged, owing to its capacity to furnish real-time, comprehensive mapping of physical spaces.

Traditionally, a fencing system denotes a network of physical barriers or structures designed to enclose or safeguard an area. Its utility spans various objectives, encompassing security provision, boundary demarcation, access control, and privacy enhancement. This term encompasses an array of fence types and associated components, including gates, posts, panels, and assorted security features. In our context, the main objective is to establish a digital boundary using an online application, such as Google Maps, delineating an area beyond which our presence is not intended.

The imperative for implementing such advanced technology is particularly crucial for mentally unstable patients, offering a valuable tool to mitigate the need for physical restraint. Rather than isolating them from society, this technology facilitates their integration while monitoring their movements and preventing access to restricted areas. This approach significantly economizes resources that would otherwise be required for alternative implementation methods.

Similarly, such a system holds potential benefits for jail prisoners transitioning back into society. Even as they approach release, authorities can maintain oversight, ensuring continued monitoring of their activities. This proactive approach helps mitigate potential risks and contributes to a smoother reintegration process into normal society.

The article is structured as follows: Section 2 provides a detailed literature review relevant to this study. Sections 3 and 4 outline the boundary establishment and subject tracking processing stages. Section 5 details experimental tests and discusses results. Section 6 proposes future technological advancements enabled by this work. Finally, Section 7 concludes the paper and offers suggestions for further research.

II. Literature Survey

A GPS GSM-based vehicle tracking system, as discussed in [1], aims to enhance transportation security and safety by combining GPS and GSM technologies to track vehicle location and provide real-time coordinates to users via an Android app. The system proposes a compact design using a development board to reduce size, power loss, and increase durability, inspired by existing anti-theft systems. It also explores related works utilizing vibration sensors, RFID, and ARM processors for vehicle tracking and safety enhancement.

A fortifying perimeter intrusion detection system (PIDS) by analysing fence vibrations, enhancing detection efficacy by scrutinizing sensor positioning and understanding attack profiles is implemented in [2]. In contrast, [3] discusses the implementation of virtual fencing using Radio Frequency modules as a labour-saving solution for managing livestock grazing, improving ecological management and minimizing manual labour. A low-cost fence impact classification system using acoustical vibrations, providing early warnings to property owners and contributing to a comprehensive farm information system is proposed in [4].

Transitioning to public health, [5] introduces an innovative approach to contact tracing during infectious disease outbreaks, offering promising prospects for disease control through automated detection of infection clusters and transmission routes. In [6], vehicle localization methods with novel algorithms for accuracy and robustness are reviewed, while in [7] a target tracking method based on multiple colour histograms, enhancing tracking robustness is introduced. Security and transportation safety needs, emphasizing the on-going policy development for informed prevention strategies are addressed in [8].

Additionally, [11] examines the correlation between weather conditions and electricity consumption for pre-emptive measures against power system overload. The authors in [12] explores dynamic weather-based forecasting methodologies' potential in mitigating energy consumption concerns. A GPS vehicle tracking system's tailored solution for efficient vehicle monitoring, integrating GPS technology with GSM and GPRS networks for real-time tracking and alerts are reviewed in [13].

The authors in [14], introduces innovative uses of GPS and GPRS technology to protect dairy animals, offering affordable solutions for remote monitoring and enhancing rural livelihood sustainability. A pioneering system for enhancing liquid transportation security using GPS and GSM technologies, addressing limitations in existing SMS-based vehicle tracking systems and offering real-time monitoring capabilities is discussed in [15]. A real-time algorithm for perimeter fence intrusion detection using tribo-electric sensors and machine learning techniques, addressing the need for accurate and timely detection is addressed in [16].

Furthermore, a comprehensive study elucidating GPS functionality, essential parameters for accurate tracking, and implementation of real-time position tracking systems, contributing to various applications leveraging GPS technology for enhanced safety and functionality is showed in [17][18].

III. Methodology

This setup involves interfacing an Arduino Uno board serially with both a GSM (Global System for Mobile Communication) modem and a GPS (Global Positioning System) receiver.

Employing a GPS modem, the system remotely tracks the latitude and longitude coordinates of the targeted object or person. Despite the GPS modem's capability to provide various metrics, the system specifically reads and displays only the National Marine Electronics Association (NMEA) data. Subsequently, the extracted position data is transmitted to the designated mobile phone location, where the inquiry about the person's position originated.

To enhance the system's capabilities in managing data effectively, an EEPROM (Electrically Erasable Programmable Read-Only Memory) module is intelligently integrated into the system architecture. This EEPROM is specifically designed to store and retain crucial information received from the GPS receiver. By utilizing the EEPROM, the system ensures that valuable data related to geographic coordinates and other pertinent location-based information are securely stored for future reference and processing. It's worth noting that this project utilizes the GPS multiplexer module GY-NEO6MV2 and the GSM module SIM 800L.

In the hardware setup described, the multiplexer operates by selectively choosing one input signal from multiple available options and then directing this chosen signal to a single output line. This mechanism effectively reduces complexity and facilitates seamless data transmission, ensuring smooth interactions between the microcontroller, GPS receiver, and GSM modem. By employing the multiplexer, the system achieves a unified and efficient interface, allowing the microcontroller to manage both the GPS and GSM modules using a cohesive set of input/output pins. This streamlined approach not only enhances the system's overall performance but also simplifies its design and implementation, contributing to a more robust and functional hardware configuration.

Upon receiving a request from the user, the system promptly delivers the user's position, expressed in latitude and longitude, to the specified mobile device, ensuring real-time location updates. By leveraging an Arduino platform program and integrating it with GPS and GSM technologies, users can not only pinpoint the exact location of the individual but also track their real-time navigation path on platforms like Google Maps, enhancing overall tracking accuracy and user experience. This integration of hardware and software enables seamless location sharing and monitoring capabilities for enhanced situational awareness. Additionally, it empowers users with precise and reliable location information at their fingertips.

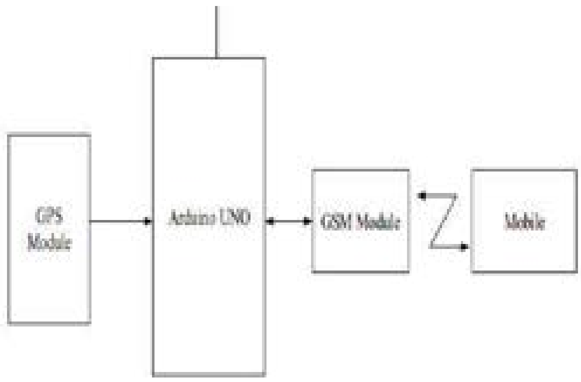


Fig1.Block Diagram of GPS and GSM Based Tracking System.

Fig.1 shows the Block Diagram of the GPS and GSM Based Tracking System, with Arduino UNO serving as the central control mediator. After the user's device initializes with the GSM module, Arduino promptly records and transmits GPS-tracked location data. The software serial then displays the user's latitude and longitude. If the GPS module fails to receive a signal, the GSM module sends a message with default coordinates (00.00). Users should wait for satellite signal reception for accurate tracking. Figure 2 provides the details of the operating procedure in the form of a flowchart

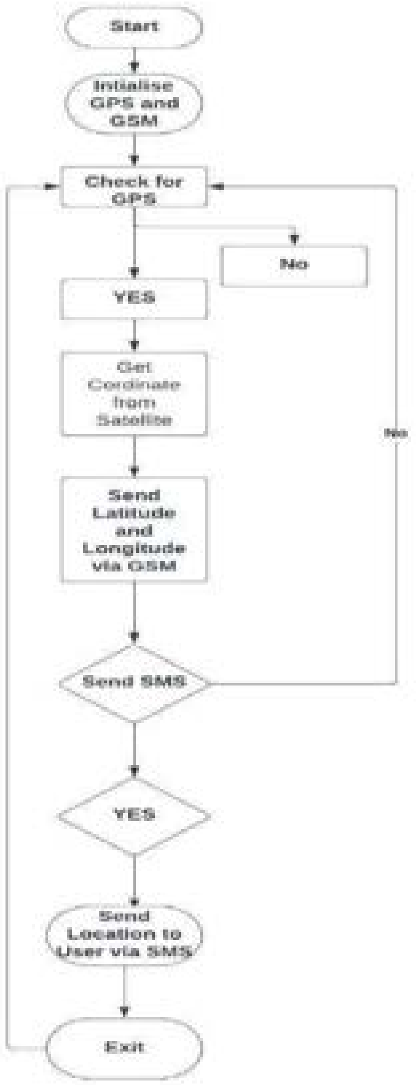


Fig 2. Flowchart of Tracking System.

IV. System Design

A. GSM Module:

In programming the GSM module, the process begins with the inclusion of necessary libraries and by defining the software serial communication. The software serial library facilitates communication on specific pins designated for receiving(Rx) and transmitting (Tx) data. By creating a constructor for software serial and passing the digital pin numbers as parameters, the system establishes a channel for communication between the Arduino board and the GSM module. This step ensures that data can be transmitted and received effectively, enabling the GSM module to interact with the rest of the system. Fig 3. Depicts the prototype of the system.

Moving on to the setup function, the GSM module undergoes initialization through serial communication with the Arduino board. This initialization sequence involves the transmission of various commands to the GSM module, which serve to verify its readiness and configure its operating parameters.

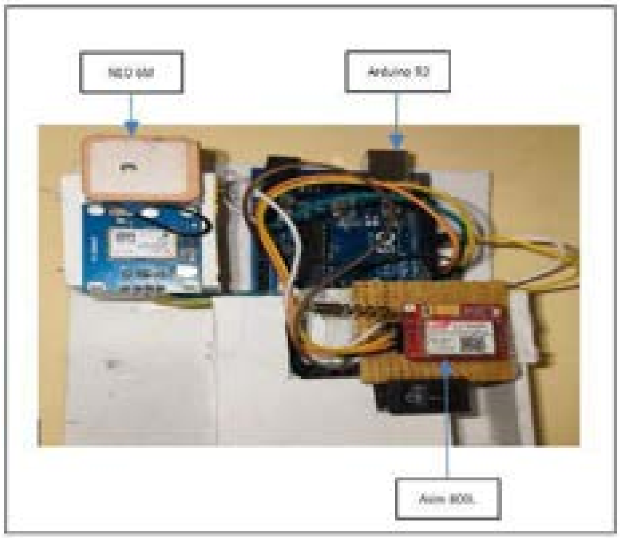


Fig 3. prototype of the system

Moreover, a delay is incorporated into the setup process to allow the GSM module a sufficient time to initialize properly before proceeding with subsequent operations.

As the system progresses into the loop function, the GSM module's functionalities are actively utilized to send SMS alerts under specific conditions. For instance, when the distance between the user's current location and the initial coordinates exceeds a predetermined threshold, an alert is triggered. This event prompts the activation of a buzzer for audible notification and initiates the sending of an SMS alert through the GSM module to a predefined phone number. Such alerts are instrumental in notifying users of critical events or conditions in real-time, enhancing the system's overall utility and effectiveness.

The specified function encapsulates the process of dispatching SMS alerts through the GSM module. It begins by configuring the GSM module for SMS text mode and specifying the recipient's phone number using AT commands. Subsequently, the actual alert message is transmitted to the designated phone number. This streamlined process ensures that timely notifications are conveyed to the user's mobile device, keeping them informed and empowered to respond to relevant events or conditions effectively.

B. GPS Module:

In programming for the GPS module, simplicity is prioritized to ensure ease of understanding for observers. The code begins by including necessary libraries and defining software serial communication. Software Serial Library (SSL) facilitates communication on specified pins, ensuring seamless interaction with the GPS module.

Within the setup function, the GPS module's initialization is declared on the transmission (Tx) and reception (Rx) pins. Serial communication is established using the necessary function, setting up the necessary parameters for data exchange.

The GPS module's operation involves reading input from the designated pin and triggering output accordingly using the looping function. Additionally, the system ccontinuously checks for updates in the GPS module's location. This ensures that the system accurately tracks the user's movements as they navigate different locations.

Overall, the programming approach for the GPS module prioritizes clarity and simplicity, ensuring that observers can easily comprehend the code structure and functionality. Through the inclusion of essential libraries and the utilization of software serial communication, the system effectively interfaces with the GPS module to provide accurate location tracking and updates in real-time.

V. Results and discussion

The integration of fencing systems with contact connecting mapping technology introduces a paradigm shift in security measures by transcending traditional physical barriers. Beyond its applications in healthcare and correctional facilities, this fusion of technologies holds promise in reshaping urban planning and public safety protocols. Imagine a city where digitally delineated boundaries dynamically adjust based on real-time data, optimizing traffic flow, and enhancing emergency response systems. This integration could redefine public spaces, ensuring efficient crowd management during events or emergencies while maintaining a seamless urban landscape.

The research explores the fusion of GPS and GSM technologies beyond object tracking, envisioning their role in disaster management. This system, in disaster-prone areas, could swiftly coordinate rescue operations, facilitating real-time tracking for efficient aid distribution. Beyond surveillance, this fusion holds potential as a cornerstone in mitigating natural disasters, bolstering preparedness and saving lives.

GPS module provides precise geographic coordinates of the object being tracked with an accuracy of 2.5m, while the GSM module facilitates real-time transmission of these coordinates to the user's device, ensuring accurate and up-to-date location. The system's reliability is further bolstered by its utilization of Google Maps for geospatial data. This established and widely trusted platform provides a high degree of accuracy in boundary demarcation and path finding, contributing to the system's overall effectiveness.

The practical execution of the experiment unfolds as follows

Step 1: Precisely delineating our boundary on Google Maps to establish our geographic area of interest.

Fig 3. Demarking our boundary

Step 2: Validating our digital map:

Fig 3. Confirmartion of our demarked boundary

Step 3: Upon breaching our predefined geo-fence, real-time alerts are triggered and sent to our designated mobile number via the GSM module. Additionally, accessing the 'Tap to load Preview' feature provides us with precise directions to the objects' current locations.

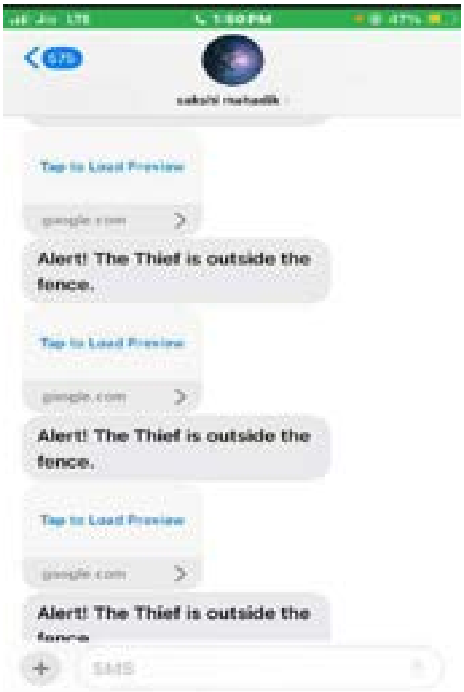


Fig 3. Security breach alerts

To enhance the effectiveness of the system, several key strategies can be implemented. Firstly, reducing the size of components can lead to more compact and efficient designs. For instance, alternatives to bulky components like the Arduino DIP R-3 and GSM module 800L can be explored, which offer similar functionalities but in a more streamlined form factor. This not only saves space but also contributes to improved portability and integration within the overall project.

Secondly, increasing the surveillance distance can significantly enhance the system's coverage and capabilities.

By adjusting the distance parameters within the code based on specific application needs, the system can extend its reach for monitoring and tracking purposes. This flexibility allows for tailored surveillance ranges, whether for large outdoor areas or confined indoor spaces, ensuring comprehensive monitoring and response capabilities.

Additionally, incorporating vibratory sensors into the system adds an extra layer of security and awareness. These sensors can detect if the device is being forcefully removed or tampered with, providing valuable feedback and triggering appropriate actions or alerts. By integrating vibratory sensors, the system gains enhanced resilience against unauthorized interference or attempts to disable the tracking functionality, thereby bolstering overall security measures and system reliability.

Table 1 below provides a comparison between the technologies of the past and now, summarizing the various aspects of all fencing techniques.

Features	**Virtual fencing using GPS**	**Fencing systems along with contact connecting mapping technology**
Flexibility	Less flexible	More flexible
Maintenance	High	low
Cost	Expensive	Cheaper
Safety	Less accident prone	Least accident prone

VI. Conclusion

In the ever-evolving landscape of security and surveillance systems, the fusion of fencing technology with contact connecting mapping emerges as a transformative solution. This integration, explored within healthcare and correctional facilities, extends its potential beyond traditional boundaries. Envisioned as a cornerstone in urban planning, disaster management, and public safety protocols, this fusion promises adaptive security and real-time response mechanisms. Beyond its current applications, future advancements in AI-driven boundaries and augmented reality overlays could revolutionize security paradigms, offering proactive and dynamic safeguards across diverse domains.

VII. References:

[1] Rebecca Gomery; Graham Leach; “Fence Vibration,” IEEE Aerospace and Electronic System Magazine, Sep 2000.

[2] *Conference in Africa (IEEE Cat. No.04CH37590)*, Gaborone, Botswana, 2004, pp. 131-136 Vol.1

[3] Varun Goyal, Mr. Ajay Mudgil, “Design and Implementation of Virtual Fencing using RF Module” International Journal of Engineering Research & Technology (IJERT) Vol. 1 Issue 3, May – 2012, ISSN: 2278-0181.

[4] Leong K-I, Si Y-W, Biuk-Aghai RP, Fong S. Contact tracing in healthcare digital ecosystems for infectious disease control and quarantine management. In: 2009 3rd IEEE International conference on digital ecosystems and technologies.

[5] S. G. D. Ndjeng, A.N. Glaser, “A multiple model localization system for outdoor vehicles,” in Intelligent Vehicles Symposium, 2007 IEEE, Istanbul, Turkey, 2007.

[6] E. Maggio and A. Cavallaro, “Multi-part target representation for color tracking,” in ICIP 2005. IEEE International Conference on Image Processing, Genoa, Italy, Sep. 2005.

[7] Akriti Gupta, Jessica Saini, Mayank Agarwal, "Android App Based Vehicle Tracking Using Gps and Gsm", INTERNATIONAL JOURNAL OF SCIENTIFIC & TECHNOLOGY RESEARCH VOLUME 6, ISSUE 09, SEPTEMBER 2017

[9] Rebecca Gomery and Graham Leach” Fence Vibrations” IEEE AES Systems Magazine, September 2000, Police Scientific Development Branch.

[10] Pietilä, J.; Räsänen, T.; Reiman, A.; Ratilainen, H.; Helander, E. Characteristics and determinants of recurrent occupational accidents. Saf. Sci. 2018, 108, 269–277.

[11] S. Jovanović, S. Savić, M. Bojić, Z. Djordjević, and D. Nikolić, ‘‘The impact of the mean daily air temperature change on electricity consumption,’’ Energy, vol. 88, pp. 604–609, Aug. 2015.

[12] A. Prabakar, L. Wu, L. Zwanepol, N. Van Velzen, and D. Djairam, ‘‘Applying machine learning to study the relationship between electricity consumption and weather variables using open data,’’ in Proc. IEEE PES Innov. Smart Grid Technol. Conf. Eur. (ISGT), Sarajevo, Bosnia-Herzegovina, Oct. 2018, pp. 1–6.

[13] I. M. Almomani, N. Y. Alkhalil, E. M. Ahmad and R. M. Jodeh, "Ubiquitous GPS vehicle tracking and management system," 2011 IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT), Amman, Jordan, 2011, pp. 1-6.

[14] Nagothu, Sudheer Kumar. "Anti-theft alerting and monitoring of animals using integrated GPS and GPRS in Indian scenario." Pak J Biotechnol 15 (2018): 56-58.

[15] Bhagat, Mahek, Jinesh Bhandari, and Siddhi Desai. "GPS Fencing based Secured Transportation for Liquid Material." International Journal of Computer Applications 975: 8887.

[16] S. Bahuguna, N. Kumar, N. Shikalgar, S. Dhage, A. Laddha and P. P. Marathe, "Event Detection and Classification Algorithm using Wavelet and Machine Learning Technique for Vibration fence PIDS," 2020 5th International Conference on Computing, Communication and Security (ICCCS), Patna, India, 2020.

[17] A. Mulla, J. Baviskar, A. Baviskar and A. Bhovad, "GPS assisted Standard Positioning Service for navigation and tracking: Review & implementation," 2015 International Conference on Pervasive Computing (ICPC), Pune, India, 2015.

[18] Ali, Junaid & Nasim, Shahid & Ali, Taha & Ahmed, Naveed & Nabi, Syed. (2009). Implementation of GSM Based Commercial Automobile Tracker Using PIC 18F452 and Development of Google Earth Embedded Monitoring Software.

Improving Security and Efficiency for Vehicular Ad-Hoc Networks using MVQQ

Mr. A. Venkateswarlu
Research Scholar,
Department of CSE,
JNTUA,
Ananthapuramu, AP, India.
Venkatesh.psu5@gmail.com

Dr. Narayana Rao Appini
Professor,
Department of CSE,
NBKR Institute of Science& Technology,
Vidyanagar, Tirupati(dt), India.
narayanaraoappini@gmail.com

Abstract: **Vehicular Ad Hoc Networks (VANETs) represent a rapidly expanding domain within wireless communication, crucial for establishing dependable and secure transportation systems. In VANETs, each vehicle functions as a node, facilitating communication with other nodes and base stations along its route. Operating akin to intelligent mobile nodes, vehicles in VANETs interact with neighboring vehicles and nodes to ensure seamless communication. Moreover, users of VANETs benefit from access to the Intelligent Transportation System (ITS), which enhances operational efficiency. This study focuses on integrating the K-means clustering and Multivariate Quadratic Quasigroups (MvQQ) algorithm within the VANET framework, serving as an asymmetric encryption method. The selection procedures for Personal Best (pbest) and Global Best (gbest) are guided by these groupings, reflecting a comprehensive review of existing methodologies in VANET research.**

***Keywords**: VANETs, K-mean clustering, ITS, MvQQ.*

I. Introduction

The Vehicular Ad-Hoc Network (VANET) is a wireless architecture created by cars connecting with one another. VANETs use well-known automobiles as nodes to form a mobile architecture, are generating creative concepts. In VANET, communication between nodes and Roadside Units (RSUs) is reallocated. It continues as a client-server relationship. To avoid mishaps and guarantee a secure and satisfying journey, inter node communication is required. By using sensors, cars may interact with their destination even when they are not linked to the network. When at least a few nodes or Intelligent Transportation System (ITS) stations connect in the radio communication band, an Ad-Hoc network is formed. This indicates that all stations have mastered the art of supplying alternate data and are aware of the position, velocity, and direction of other stations. Specifically, it develops the three core elements of the architecture: the Roadside Unit (RSU), the Onboard Unit (OBU), and the WAVE (Wireless Access in Vehicular Environment) interface. A VANET's primary goal is to give highway users security. In order to minimize accidents and promote safe and effective fleet transportation, inter-node communication is essential. The sensors located in the cars play a crucial part in the cars' capacity to communicate with each other regarding the destination, where the communication takes place between server and client.

One special quality of a VANET is that because its nodes are so mobile, it can quickly create a secure network. In the paper, we employ a Multivariate Quadratic Quasigroup (MVQQ), an asymmetric encryption technique. Certificate manufacturing and distribution, certificate management and revocation, user authentication, and certificate maintenance are some of the other impacts it provides. Our suggestion is a protocol that uses a hierarchical pseudonymous structure to provide partial anonymity while authenticating a node while it is communicating with other nodes in the network.

II. Interdependent Work

Security in VANETs is frequently considered by certain researchers, but the majority of them may not be supplied with data concerning the operation or evaluation of symmetrical or asymmetric algorithms that are operating in a real-world vehicular network. As a result, we believe that this work is valuable it shows a processing speed of the MVQQ under various VANET conditions. Term was to boost neighbourhood safety via using readily available doppler as discovering associates as well as confirming reported GPS positions. There exploited an existing region, that relies upon cells, and created the system of communications. I describe analysis abilities with anonymous identification. Several other anonymous network principles have been deployed using assistance by OKI. The approaches rely on OKI-based Certifications, that have an emotive linked to similar secret key. Distribute a greater Amount to identities across links with similar secret key. A plan assumes the unavoidable use about RsU's, which increases the framework load, and as a result, the framework's general everyday practice progresses. The Certificate issuing authority provides a first name, which allows for coordination among the initial pseudonyms and the node's true identification. The initial Every credential stored within the CDA (Certificate Distribution Authority) system is protected via another item known to be Revocation Distribution Authority (RDA), so CDA cannot decode those genuine identifications. At that moment, a node transmits the communication, which is signed with the alternative moniker's linked private key, and the receiving node validates the messages which corresponding open key provided in the alternative name. Then examining a paper, i deduce an increasing limitation. Taking a pseudonymous approach leads to significant results, computing, Communication with capacity capabilities due to an existence CRL (Certification Revocation Lists). This study has analyzed an asymmetrical (MVQQ algorithm) to see whether it was real. This suitable car transmissions arrangement expects Vehicle to Vehicle (V to V) & Vehicle to Infrastructure (V to I) interactions that provide protection to prevent crashes and congestion. Furthermore, 3 basic encryption approaches are investigated: public keys, symmetrical keys, Personality-based encryption is employed to protect the service. Proposed alternative clusters structure for high speeds across VANET nodes with security methods make transfer more efficient.

III. Clustering Head Algorithms (K-means)

This study employs a K-mean clustering head method for split networks through k sections. It is well known that selecting Most excellent vertices among an enormous amount during links is challenging, particularly for the value k.

K-means clustering divides n sensations in k clusters, resulting in every perception belonging to the particular cluster with the highest imply that which serves as as the cluster's prototype.

The personal-best location (pbest): The personal best is the one that has been uniquely determined with particular molecules based on prior estimations of xi. During an optimization

$$p^{(t+1)}_{bestid} = \{X^{(t)}_{id}\ if\ f(X^{(t+1)}_{id}) \geq f\left(p^{(t)}_{bestid}\right)$$

$$p^{(t+1)}_{bestid} = \{X^{(t+1)}_{id}\ if\ f(X^{(t+1)}_{id}) < f(p^{(t)}_{bestid})$$

task, the state having the smallest potential values appears to be healthy. The letter f(X) represents the desired benefit which has been limited.

The Global Best (gbest): A gbest gives a faster rate for This Collaboration sacrifices authority. The gbest regulates a certain smallest setup, called by the term global-best particle, throughout every single structure within the grouping, by final form, every one of particulates going to arrive with that point, and therefore if they don't get condensed on an ongoing basis, the team might arrive prematurely.

IV. Safety in the VANET

Recent study indicates that asymmetric encryption techniques should be utilized in embedded systems, as stressed According to which explored asymmetrical encryption methods having extra safety objectives, RSAs having the key capacity is 3076 bits and ECCs having 512 bits have been used into embedded devices.

(i) Asymmetrical Algorithm

MVQQ Algorithm (Multivariate Quadratic Quasigroup):

Previously, safety posed by encryption techniques was dependent on two mathematically independent numerical problems: computing the viability of the different logarithms measure and factoring integers. An alternative open-key strategy named Multivariate Quadratic Quasigroups (MVQQ) was developed. As the Quasigroups change, the procedure depends with Quadratic Multivariate polynomials and retains the associated features.

Step 1: MVQQ represents the system that operates outside the quantum.

Step 2: The security approach has similarities by previous public keys encoded procedures impacted by Multivariate Quadrics.

Step 3: The quickness about decoding seems quite similar using a usual symmetrical blocks authentication approach.

Step 4: This algorithm is strongly parallelized, unlike other progressive algorithms.

The standard MVQQ structure is a multivariate quadric system.

Two non-singular linear transformations, A and C, plus a pair of multivariate quadratic alignments over {0, 1} n is represented by B'. Encryption using an open key is a quick process for using n multivariate polynomials.

$B = \{B_i(s_1 \ldots\ldots s_n) \mid i = 1, \ldots, n\}$ That is, r = B(s) over the vector, so s = (s1..........sn).

What can be described as,

An equipment testing show that MVQQ is formed from a median Symmetrical blocks encrypting. Testing is using the sensing architecture shows the MVQQ was extremely quicker then techniques including like ECC & RSA.

$$r = B(s) \equiv y \equiv DZ$$

All of this confirmed a resolution when conducting initiatives; they discovered providing electronic certificate formed using MVQQ could be between 300 to 70000 percent faster than ECC and RSA certificates. With each case, MVQQ have an option surpass about 10,000 times. Furthermore, indicates MVQQ method gives someone else avenue of encryption. industry; this generates unique open-key encryption systems in addition to improving those that already exist. The terms processor utilization, storage spaces and

$$A \circ B \circ C: \{0, 1\}^n \rightarrow \{0, 1\}^n$$

setup time are used in relation to each other. The results indicated MVQQ provides excellent way to approach embedded technology due to its matches both RSA & ECC.

Creating public & secret key to use algorithm of MVQQ:

Input: an integer r, where l>28 and r=5l.

Output: Start key O: p Multivariate Quadratic Polynomials Oi (y1,...,yn), i=1,...,n,

Secret Key: The two nonsingular binary matrices of rank m * m and eight quasigroups $^*1, \ldots, {}^*8$.

1. Generate two non-singular p * p Binary matrices, S and U, at

random.

2. Use the following for defining O' (p): $\{0,1\}^p$ $\{0,1\}^p$. You will also get the quasigroups *1,…,*8 from there.

3. Figure In which y=(y1,...,yn), x=S(O'(U(y)))

4. The output shows that the private secret is an order pair (S,U,*1,….,*8) & a public key is x with Multivariate Quadratic polynomials Oi(y1,…yn), i=1,….n.

V. Proposed Authentication Protocol

Notation	Explanation
V_i	Initiator/Sender vehicle
V_v	Receiver vehicle
VID_i	Initiator's/Sender's vehicle ID
IK_i, AK_i, IK_i', AK_i', IK_i'', AK_i''	MVQQ open/private key pairs of V_i
IK_{CDA}, AK_{CDA}	MVQQ open/private key pairs of CDA
IK_{CAP}	Paillier open key pair of CDA
IK_{RsU}/ AK_{RsU}	MVQQ open/private key pairs of RsU
T_{CDA}, T_{CDA}'	Expiration time of initial pseudonym set by CDA
T_{RsU}	Expiration time of alternate pseudonym set by RsU

Vehicle registration and first pseudonym creation

During the moment, sender/initiator nodes (V_i) create an arbitrary value k (That erratic value is encapsulated with CDA's public keys) & public/private key MVQQ pairs.

IK_i, AK_i

V_i transmits this info along with the VID_i to CDA.

Step 1: V_i → CDA: k||IK_i|| VID_i

A Vi transmit this information to a CDA over a secure channel (for example, nodes contact a CDA).

The first step is required one time only. The Certificate of Distribution Authority authorizes the VID_i. As verified, it encrypts VID_i with one open key generated by RDA, encrypting k using the Paillier public keys IK_{CAP}, generates an end date T_{CDA} and constructs the subsequent database access.

An instance of a CDA databases:

CDA → DB: $(VID_i)_{PKRDA}$|| T_{CDA}||IK_i||k

CDA signs (T_{CDA}||IK_i||$(k)_{IKCAP}$), and attach it into V_i as its first initial pseudonym.

Step 2: CDA → V_i: (T_{CDA}||IK_i||$(k)_{IKCAP}$) AK_{CDA}

Restore the original pseudonym

After the T_{CDA} leaves, V_i needs to regain the original pseudonym. V_i produces an open/private MVQQ pair of keys by randomly choosing some k'.

IK_I''/ AK_I''

Encrypts the information in open key of CDA alongside *k* and deliver it to CDA by utilizing 3G/4G technology.

Step 3: V_i → CDA: (k||k'||IK_i'') IK_{CDA}

if a node wants to send back a prior identity on CDA via RsUs, the information is transmitted to any locally RsUs, whose forward the request to the CDA. On demand, some distinct cause values in the communication are used, allowing the RsUs to determine if the node is requesting an original pseudonym from the RsUs nor a recent alternative moniker.

Step 3': V_i → RsU CDA: (k||k'||IK_i'') IK_{CDA}

CDA certifies the information using optimal k, creates an additional terminal time T', and modernizes the DB using distinct k' metrics, IK_i and I_{CDA}, CDA measures the second phase encrypting the freshly created first name into IK'' before sending that via V. If the request originates through RsUs, CDA will send this communication to Vi via RsUs, along with the signed k. The signed number k creates a corporation that carries the sophisticated values k'. If RsU promotes this text, V_i will examine it with aged k, verify CDA's signature, decrypt it, and alter its original pseudonym. Due to data encryption, RsUs are unable to reveal the sophisticated first identification to the Vi.

Step 4: CDA → V_i: ((T_{CDA}||IK_i'||(k)IK_{CAP}) AK_{CDA}) IK_i''||(k) AK_{CDA}

Forming alternate pseudonyms

RsU sometimes conveys information while stating the level of quality. It also contains the public value of the RsUs. If a node receives this notification, it requests an alternative pseudonym.

A node generates additional open/private MVQQ key pairs (IK', AK'). This encrypts and sends the just formed public key, initial pseudonym (−k), and nonce to the RsUs.

Step 5: V_i → RsU: ((T_{CDA}||IK_i||(k)IK_{CAP}) AK_{CDA} || IK_i'|| - k||nonce) IK_{RsU}.

RsUs examine CDA brand and encrypts −k with CDA's public keys.

RsUs for storing same combos (k) IK_{CAP} and (−k) IK_{CAP} will get (S) IK_{CAP}.

RsUs submit (S) IK_{CAP} to CDA over authentication, with (S), which is IK_{CAP} = (k) IK_{CAP} + (−k) IK_{CAP}.

Step 6: RsUs → CDA: (S) IK_{CAP}

CDA decrypting S, catches O (k + (−k) = 0), and sends a verified notification to RDA if it cannot be delivered.

Step 7: CDA → RsU's: verified or not verifiable. CDA gets the secret data and has no notion which node is using this information. −k is used to prevent a harsh assault.

RsU's set up an additional identity after obtaining confirmation that the transmission originated from V_i. It creates the end of time TRsU, puts it into the freshly produced IK'$_i$, & transmits it to V_i. The IK'$_i$ must be supplied by V_i whenever another pseudonym is requested. In any circumstance, a node may re-register in an MVQQ key value lake.

Step 8: RsU → V_i: ((T_{RsU}||IK_i) AK_{RsU}) IK_i'.

VI. Working of Proposed Architecture

In the protocol that is suggested, a user node must communicate with another node before sending messages to registered nodes via Roadside Unit & K-means Clustering. We offer the best nodes that can supply both decoding and encoding keys to the environment's nodes using the MvQQ method. To obtain an initial pseudonym through the (RsU), the customer node must communicate to a different node via K-means clusters Head & enroll to the Certificates Distribution Authority. We also give the MVQQ secret and public secret method. These periods are identified in our standard as TCA, and they correspond to the season of the first pseudonym generation as determined by CDA. The desired protocols are the following to be discussed.

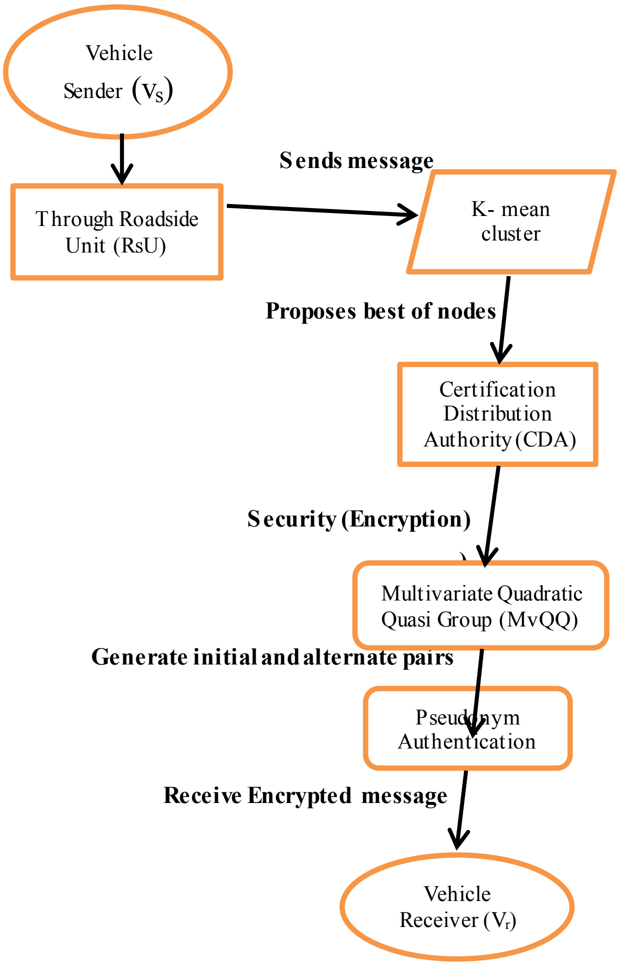


Fig 1: A flow diagram of the proposed approach

VII. Performance Assessment

In this paper, we propose and include the K-mean clustering head and MVQQ algorithms. At that moment, formation was created, allowing us to test algorithm implementation in a VANET and then examine the results. In this underlying circumstance, nodes are located in a region with a distance of less than 100m. The nodes in this instance are inside the system's treatment discipline, and nothing has been disposed of. In an example of ten nodes, node "0" delivers a message to alternate nodes in the system. A In our simulations, this action happens after 0.4 m. So, we provide MVQQ's productivity built on what was tried (tested) in various papers, but in various VANET network contexts. In comparison to the others, the MVQQ algorithm contains the key length of 160 particles. Fig4 Illustrates a growing arc with the quantity the data flowing in the VANET increases. Here, we assess our proposed protocol's deployment to valuable aspects. Viewpoints will evaluate the execution of RsUs if the node asks for a different pseudonym from the RsUs after providing an initial pseudonym. If the RsUs verify that the very first alias basically contains the request, they then generate and send the alternative pseudonym to the node. As an outcome, the RsUs can basically determine if they are able to do this duty for the nodes on an ongoing basis.

Packet Delivery Ratio:

A process of delivery has been calculated for dividing how many concerning traffic successfully inserted with the entire number of particles received. PDR has been described by the entire quantity of frames transmitted from start to finish. If the network's share is increased in any way, using this process would also enhance network support. The PDR equation is:

PDR = (RCV/SND)*100

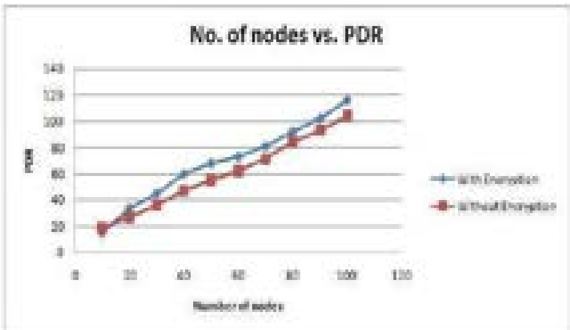


Fig 2: PDR with relation to nodes via MvQQ encoding

The blue points in the previous image represent the suggested route, while the bright red line represents the current approach. The distinction effectively states that the new system may surpass the present approach.

Throughput:

As much data units a system can handle in a given amount of time is measured by its throughput. Throughput is defined as the amount of information that flows from one station to another. Each second, bits are swapped starting from one location and moving to another. With the occasion that the rate of traffic is high, transmission of data ability. Notice the following formula for throughput:

Throughput = bit/second.

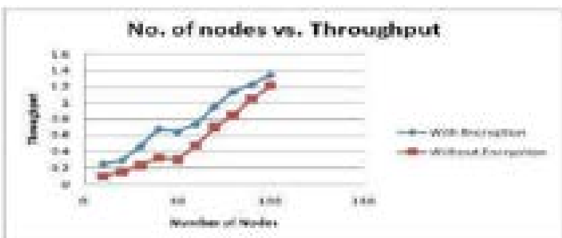


Fig 3: Throughput against speeds between node’s

End to end delay:

End to end delay is critical for determine an impact for encrypting expense upon End-to-end latency when car & speed measurements increase.

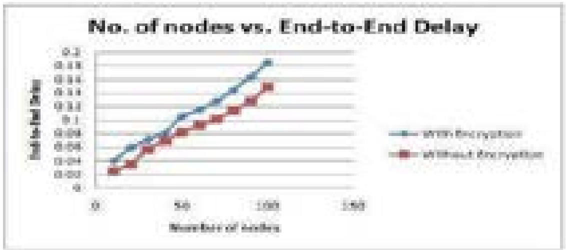


Fig 4: End-to-end delay based on the velocity among the nodes

The encrypted numbers showed that the suggested approach was increased in the graph above. Whereas the non-encryption approach indicates that the suggested work is more effective than the present method.

Packets over head: A period that's required for conveying info through packet switched network. Each frame necessitates additional bytes inside arrangement facts, that stays into packets headers then, while applied to packets creation & disassembling, decreases an entire communication time of plain text.

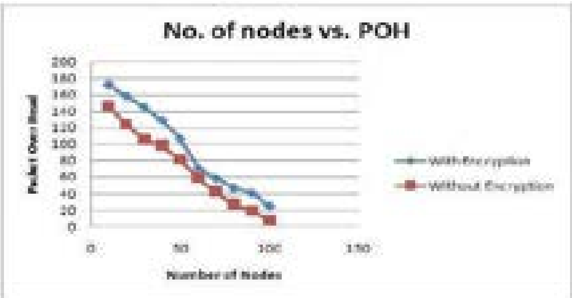


Fig 5: Packet Over Head in terms of frequency between nodes

The graph depicts a packet overhead layout comparing the present and suggested approaches. The suggested approach takes more in the extra process than the basic technique.

VIII. Conclusion

Vehicular ad hoc networks (VANETs) introduce a multitude of security concerns, with current research predominantly focusing on addressing these issues. The proposed protocol undergoes rigorous security analysis, revealing its adaptability to various security threats. Tracking a node within the network proves challenging unless prompted by a deliberate or malicious act. However, in rare instances of nefarious behavior detection, the perpetrator is promptly identified and barred from network access. The strategic deployment of Roadside Units (RSU) at high intervals aims to counteract pseudonym expiration, ensuring sustained network security. This study represents an extension of our prior work, addressing prominent challenges in pseudonymous verification through comprehensive, point-by-point analysis and thorough simulation-based evaluations.

References

[1] M. A. Alazzawi, H. Lu, A. A. Yassin, and K. Chen, ''Robust conditional privacy-preserving authentication based on pseudonym root with cuckoo filter in vehicular ad hoc networks,'' *KSII Trans. Internet Inf. Syst.*, vol. 13, no. 12, pp. 6121–6144, 2019.

[2] J. Zhang, J. Cui, H. Zhong, Z. Chen, and L. Liu, ''PA-CRT: Chinese remainder theorem based conditional privacy-preserving authentication scheme in vehicular ad-hoc networks,'' *IEEE Trans. Dependable Secure Comput.*, vol. 18, no. 2, pp. 722–735, Mar. 2021.

[3] J. Cui, L. Wei, H. Zhong, J. Zhang, Y. Xu, and L. Liu, ''Edge computing in VANETs—An efficient and privacy-preserving cooperative downloading scheme,'' *IEEE J. Sel. Areas Commun.*, vol. 38, no. 6, pp. 1191–1204, Jun. 2020.

[4] D. Manivannan, S. S. Moni, and S. Zeadally, ''Secure authentication and privacy-preserving techniques in vehicular ad-hoc NETworks (VANETs),'' *Veh. Commun.*, vol. 25, Oct. 2020, Art. no. 100247.

[5] M. Azees, P. Vijayakumar, and L. J. Deboarh, ''EAAP: Efficient anonymous authentication with conditional privacy-preserving scheme for vehic- ular ad hoc networks,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2467–2476, Sep. 2017.

[6] J. Cui, J. Chen, H. Zhong, J. Zhang, and L. Liu, ''Reliable and efficient content sharing for 5G-enabled vehicular networks,'' *IEEE Trans. Intell. Transp. Syst.*, early access, Sep. 24, 2020, doi: 10.1109/TITS.2020.3023797.

[7] P. Vijayakumar, M. Azees, and L. J. Deborah, ''CPAV: Computationally efficient privacy preserving anonymous authentication scheme for vehicular ad hoc networks,'' in *Proc. IEEE 2nd Int. Conf. Cyber Secur. Cloud Comput.*, Nov. 2015, pp. 62–67.

[8] A. Kaci, T. Bouabana-Tebibel, A. Rachedi, and C. Yahiaoui, "Toward a big data approach for indexing encrypted data in cloud computing," *Secur. Privacy*, vol. 2, no. 3, p. e65, May 2019.

[9] Y. Yahiatene, D. E. Menacer, M. A. Riahla, A. Rachedi, and T. B. Tebibel, ''Towards a distributed ABE based approach to protect privacy on online social networks,'' in *Proc. IEEE Wireless Commun. Netw. Conf. (WCNC)*, Apr. 2019, pp. 1–7.

[10] A. Kaci, A. Nacef, and A. Henni, ''Mobile cloud system for road safety,'' in *Proc. Int. Conf. Geoinformat. Data Anal.*, Apr. 2018, pp. 132–136.

[11] A. Kaci and A. Rachedi, ''Mc-Track: A cloud based data oriented vehicular tracking system with adaptive security,'' in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[12] Y. Yahiatene, A. Rachedi, M. A. Riahla, D. E. Menacer, and F. Nait-Abdesselam, ''A blockchain-based framework to secure vehicular social networks,'' *Trans. Emerg. Telecommun. Technol.*, vol. 30, no. 8, p. e3650, Aug. 2019.

[13] A. Alamer, Y. Deng, G. Wei, and X. Lin, ''Collaborative security in vehicular cloud computing: A game theoretic view,'' *IEEE Netw.*, vol. 32, no. 3, pp. 72–77, May/Jun. 2018.

[14] A. Alamer and S. Basudan, ''An efficient truthfulness privacy-preserving tendering framework for vehicular fog computing,'' *Eng. Appl. Artif. Intell.*, vol. 91, May 2020, Art. no. 103583.

[15] A. Alamer, S. Basudan, and X. Lin, ''A privacy-preserving incentive framework for the vehicular cloud,'' in *Proc. IEEE Int. Conf. Internet Things (iThings), IEEE Green Comput. Commun. (GreenCom), IEEE Cyber, Phys. Social Comput. (CPSCom), IEEE Smart Data (SmartData)*, Jul. 2018, pp. 435–441.

[16] X. Li, T. Liu, M. S. Obaidat, F. Wu, P. Vijayakumar, and N. Kumar, ''A lightweight privacy-preserving authentication protocol for VANETs,'' *IEEE Syst. J.*, vol. 14, no. 3, pp. 3547–3557, Sep. 2020.

[17] M. A. Al-Shareeda, M. Anbar, M. A. Alazzawi, S. Manickam, and A. S. Al-Hiti, ''LSWBVM: A lightweight security without using batch verification method scheme for a vehicle ad hoc network,'' *IEEE Access*, vol. 8, pp. 170507–170518, 2020.

M. K. C. Reddy, ''A robust and efficient password-based conditional privacy preserving authentication and group-key agreement protocol for VANETs,'' *Future Gener. Comput. Syst.*, vol. 84, pp. 216–227, Jul. 2018.

[18] Z. Lu, G. Qu, and Z. Liu, ''A survey on recent advances in vehicular network security, trust, and privacy,'' *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 2, pp. 760–776, Feb. 2018.

[19] M. Muhammad and G. A. Safdar, ''Survey on existing authentication issues for cellular-assisted V2X communication,'' *Veh. Commun.*, vol. 12, pp. 50–65, Apr. 2018.

[20] M. A. Al-shareeda, M. Anbar, S. Manickam, and I. H. Hasbullah, ''An efficient identity-based conditional privacy-preserving authentication scheme for secure communication in a vehicular ad hoc network,'' *Symmetry*, vol. 12, no. 10, p. 1687, Oct. 2020.

[21] S. S. Manvi and S. Tangade, ''A survey on authentication schemes in VANETs for secured communication,'' *Veh. Commun.*, vol. 9, pp. 19–30, Jul. 2017.

[22] M. Azees and P. Vijayakumar, ''CEKD: Computationally efficient key distribution scheme for vehicular ad-hoc networks,'' *Austral. J. Basic Appl. Sci.*, vol. 10, no. 2, pp. 171–175, 2016.

[23] *Global Status Report on Road Safety 2021*, World Health Org., Geneva, Switzerland, 2021.

Artificial Intelligence based Early Detection of Cardiovascular Diseases

R.MeenaPrakash, P.Krishnaleela, R.Mahendran, T.Manickaraj, M.Mariyappan
Department of Electronics and Communication Engineering, P.S.R. Engineering College, Sivakasi, India.
{meenaprakash@psr.edu.in, krishnaleela@psr.edu.in, rmahendran323@gmail.com, manickaraj707@gmail.com, marimuthukrishnan2003@gmail.com }

,

***Abstract*—Many sectors, including healthcare, education, agriculture, and industries, depend heavily on Machine Learning (ML) algorithms. Predicting cardiovascular disease is one of the world's biggest challenges. Many people may die as a result of coronary artery disease. Enormous research work is being carried out to identify the factors that could predict the development of cardiovascular disease. Heart disease prediction is currently being effectively resolved by utilizing techniques of Artificial Intelligence (AI) including machine learning algorithms and deep learning. This research study has implemented several AI-based classification algorithms such as Support Vector Machine, Random Forest, Logistic Regression, Decision Tree, and K-Nearest Neighbours (KNN) for the prediction of heart disease. Finally, this study employs performance indicators including the confusion matrix, accuracy score, F1-score, recall, precision, sensitivity, and specificity to analyze the model's effectiveness and performance. It is inferred from the experimental results that the highest classification accuracy of 91% is achieved for the Random Forest Classifier when compared to other machine learning algorithms on heart disease dataset.**

Keywords—*Cardiovascular Disease Prediction, Artificial Intelligence, Classification.*

I. INTRODUCTION

As per the World Health Organization Statistics, cardiovascular diseases affect the survival of 12 million individuals annually. Due to its status as the leading cause of death, heart disease may contribute to global morbidity and death rates. One crucial component of data analysis is the prediction of coronary heart disease. Because people can die from it without experiencing any serious symptoms, it results in sudden deaths without giving a chance for any precautionary remedial measure. Using Artificial Intelligence techniques, the prediction of cardiovascular disease is resolved. Artificial Intelligence (AI) is rapidly becoming an inevitable tool in healthcare applications such as computer-aided diagnosis of diseases from different image modalities like CT scan, MRI, etc., and for disease prediction and analysis from huge volumes of patients' records. The challenges associated with the prediction of cardiovascular diseases are that the conventional methods are not able to predict the diseases at an early stage and also they consume much time. This paper proposes to predict heart disease by evaluating the patient data & determine the model by using the AI techniques. The proposed algorithm will predict whether the suspicious patients are afflicted by heart disease or not using machine learning techniques.

II. LITERATURE SURVEY

Cardiac ailments are presently the leading cause of death globally, and their prevalence is rising. As per statistics from the World Health Organization (WHO), deaths due to heart diseases are nearly 31% of global deaths, and hence prediction of diseases occurring in the heart in an early stage has become mandatory for the betterment of society. The great challenge in heart disease detection in the current scenario is to identify the disease at an early stage based on the symptoms. Identifying heart disease in its early stages before a cardiac incident presents difficulties. The medical field, including clinics, hospitals, etc., has access to a vast amount of data regarding cardiac disease. The worldwide standard medical datasets and records of patients available are increasing day by day both concerning the number of characteristics and the number of available records. But this data isn't managed sensibly enough to reveal any underlying patterns. The medical data can be transformed into knowledge that is helpful by using machine learning algorithms. Decision support systems (DSS) with the capacity to gain knowledge from and improve upon past experiences are created through the application of machine learning. Classification in machine learning refers to labeling the input data and in the case of heart disease prediction, it refers to labeling the input data as a patient affected by heart disease or not. First, the machine learning model is trained on the input data with known labels and feature learning is done. Once the model is trained, it can classify any unknown data. Academics and industry alike have recently become interested in deep learning [1]. It is crucial to correctly diagnose heart diseases while incorrect diagnosis will result in loss of human lives. The Machine Learning algorithms of Support Vector Machine and Logistic Regression (LR) algorithms were employed for heart disease prediction and the dataset used was the UCI machine learning dataset to evaluate the method. The dataset comprises of 13 different characteristics for investigation and the results proved that the logistic regression algorithm gave higher accuracy compared to the Support vector machine (SVM) algorithm for prediction of coronary disease. Feature extraction technique was employed to remove the irrelevant features and only significant features were selected to anticipate the risks associated with heart disease [2]. While using a machine learning model for classification, one of the methods to enhance the classification accuracy is to optimize the parameters of the machine learning model during the training process which is called hyper-parameter tuning. The hyper-parameters that are optimized include learning rate and step size whose values are selected based on permutation and combination of different values. It was observed that there is an enhancement in classification accuracy due to hyperparameter tuning [3]. Different machine learning techniques including, the Bayesian classifier, Decision tree (DT), Random Forest classifier, and others were employed were applied to predict heart disease using electrocardiography signals and other risk factors. From the comparative analysis of the different models, the best models were selected and generalized for predicting heart disease

[4]. Detection and diagnosis of ischemic heart diseases was implemented using Machine learning algorithms. The problems associated with Ischemic heart disease data sets are the selection of features, imbalance in the sample size, and data unavailability for certain characteristics, etc., The outcome of the proposed method was an improvement in selecting significant features thereby reducing the dimensionality and also improving accuracy. The optimization technique employed was an enhanced squirrel search algorithm with meta-heuristic procedure which resulted in an enhanced classification accuracy of 98% [5]. Different AI techniques of Random Forest, Linear Discriminant Analysis, Gradient Boosting (GB) Algorithm, Decision Tree, and others were employed to predict cardiovascular disease where cross-validation technique with K-Fold was used. The different datasets such as Cleveland, Hungary, and Long Beach VA were used to assess the models' performance, and accuracy values of 99.40%, 99.76%, and 100% were obtained [6]. Four different methodologies comprising LR, KNN, SVM, and GB Algorithms were utilized to predict cardiac disease from the datasets of Hungary, Switzerland, Cleveland, UCI Kaggle, and Long Beach VA. The experimental results obtained proved that the maximum accuracy of 99.03% was achieved for the Extreme GB Classifier with GridSearchCV [7]. In developing countries, cardiac specialists will not be available in rural areas and hence an automatic diagnosis system plays a significant role in heart disease diagnosis at an early stage based on the symptoms. A hybrid diagnosis system based on the examination parameters and vital signs of the patients was implemented for timely heart disease prediction. To handle the missing values, the method of multivariate imputation was used. Recursive feature elimination along with a genetic algorithm was used for selecting salient features from the dataset. The method was tested on the dataset of UCI Cleveland heart disease dataset [8]. An automated diagnosis system of healthcare to detect diseases of the heart was proposed using an Artificial Neural Network with Swarm Optimization. A specified quantity of neural networks was randomly generated to train and test the data in consistency with the solution [9]. The problems associated with supervised machine learning classification algorithm models are underfitting and overfitting. A hybrid classifier that employed an ensemble model with majority voting practice was proposed which overcame the above problems of ML algorithms. In this method, a genetic algorithm was used for feature selection, 10-fold cross-validation was employed, and enhanced prediction accuracy and execution time were achieved. The dataset used was the UCI Machine Learning Repository [10]. An efficient method to handle Electronic Health Records (EHRs) of patients with heart diseases was proposed in which association rules were used to analyze the patient's medical data distributed in different hospitals [11, 12]. An automated diagnosis system was proposed to diagnose heart diseases in which Spatial grouping of Applications based on density and Noise was employed for the detection and elimination of outliers. The method employed hybrid techniques of Edited Nearest Neighbor along with Synthetic Minority Over-sampling Technique (SMOTE) for a balanced distribution of data and the XGBoost machine learning algorithm was used for the prediction of heart diseases [13]. Another major cause of worldwide deaths is Coronary Heart Disease. According to a survey by WHO, Cardiac arrests account for the majority of the deaths but they could be prevented if the concerned people follow healthy lifestyles. The major risk factors associated with heart disease are elevated cholesterol levels, diabetes, and an increase in blood pressure. These parameters could be continuously monitored and used as parameters to predict heart-related diseases in an early stage and to avoid serious consequences. Several other machine learning models were also implemented to predict heart diseases in an early stage based on various risk factors associated with them [14, 15]. The advantages of these methods are simplicity and easy implementation. The drawbacks are the need for proper feature selection, less performance for categorical values, and data imbalance. As inferred from the literature survey, AI-based techniques are a promising approach for heart disease prediction in an early stage. In this paper, we have proposed and analyzed the performance of five different Machine Learning (ML) models which include RF, SVM, LR, KNN, and DT algorithms. The performance of these ML algorithms is evaluated based on different metrics and the Kaggle dataset is used for the experiments.

III. PROPOSED METHOD

A. Overview

The block diagram of the proposed method is presented in Figure 1.The operation of the system begins with data collection and the selection of important characteristics. Next, the raw data is subjected to a pre-processing step to obtain the data in the required format. The pre-processing steps include corrections for missing values and data imbalance using SMOTE. SMOTE stands for Synthetic Minority Oversampling Technique and it is used to overcome the data imbalance problem. From the input data, the minority samples are taken and new instances are generated by merging the samples of each class with its nearest neighbors. The advantage of the SMOTE technique is that the data in the minority classes are only increased to avoid overfitting problems. The input dataset is split into training & testing data (80:20). Then, machine learning algorithms are applied to train & test the model. Due to this system, enhanced visualization and ease of interpretation are possible. Extensive experiments on the datasets have demonstrated the effectiveness of our method in timely heart disease prediction.

B. Random Forest Algorithm

In the Random Forest Algorithm, multiple classifiers are combined which is called ensemble learning through which the performance is enhanced. It is a supervised learning technique in which during the training phase, the model is trained using known labels. The tree is split into nodes and each node available in the tree is predicted. Both classification and Regression tasks can be solved using the Random Forest algorithm. From a given dataset, samples are selected randomly. A decision tree is constructed for individual samples and a prediction result is obtained. For each predicted result, voting is performed and the prediction result with majority votes is nominated for ultimate prediction. The advantages of the Random Forest algorithm are that they are capable of handling large datasets and high dimensional data.

C. Support Vector Machine

Support Vector Machine (SVM) is another supervised technique in which the data points are separated into classes using a decision boundary. The new data will be then categorized correctly based on this decision boundary which is called a hyperplane. The points that lie close to the hyperplane are called support vectors and the location and direction of the hyperplane are decided by them. The gap between two lines which separate the different classes is called margin and it is necessary to have a large margin. The merits of the Support Vector Machine classifier are that it is memory efficient and well suited for higher dimensions applications.

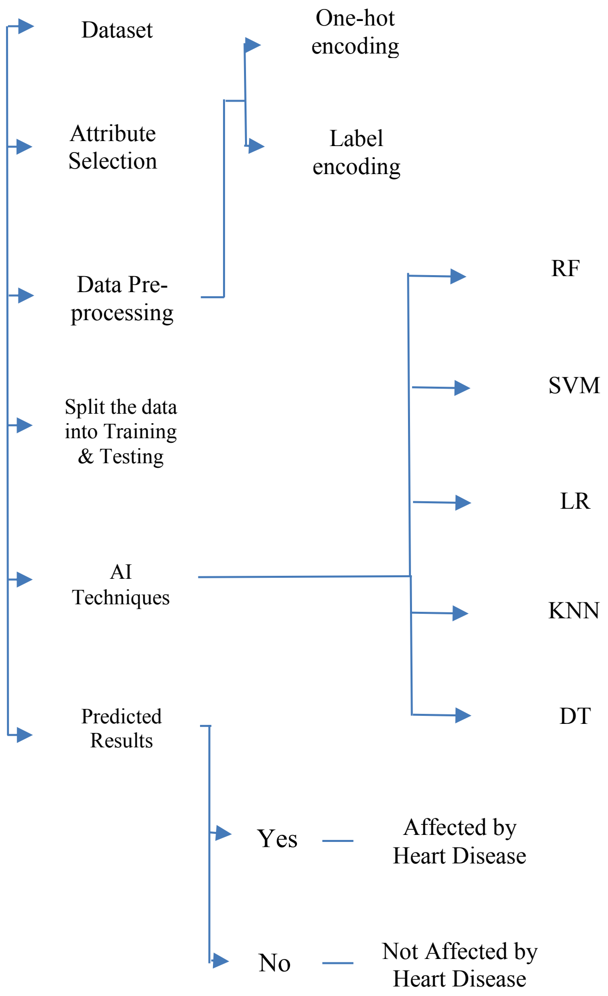


Fig. 1 Block Diagram of the proposed method

D. Logistic Regression

Logistic regression is one of the supervised learning algorithms and it is used to estimate the probability of whether an event will be occurring or not. Based on past observation, the outcome of the dependent variable will be predicted and it is used for classification tasks. In the method of logistic regression, the sigmoid function is used to find out the probabilities from the predicted values whose values lie in the range of [0 1]. The expression for the sigmoid function is given in Equation 1.

$$\delta(x) = \frac{1}{1+e^{-x}} \tag{1}$$

Where δ is the sigmoid function and x is the input variable.

The advantages of the logistic regression algorithm are that it is easier in computation and well suited for applications where the data are linearly separable.

E. K-Nearest Neighbours Algorithm

In the KNN algorithm, all the data points are stored and new data are classified according to the similarity between the new information and the available information. First, K neighbors are selected and then the Euclidean distance of the K neighbors is estimated. New data points are assigned to a class based on the distance between the new data point and the available data points. The advantages of the KNN algorithm are that it is robust to noise and simple to implement.

F. Decision Tree Algorithm

DT algorithm is well suited for regression and classification tasks which comprises two nodes called decision node and leaf node. Multiple branches exist in decision nodes and they are involved in making decisions, whose output is the leaf nodes. The leaf nodes do not contain any further branches and they are called the end nodes. CART algorithm which is the acronym for Classification and Regression Trees is used to build the decision tree. Both linear and non-linear problems can be solved using the Decision Tree algorithm. To build a tree, we use the CART Algorithm and it works on both linear & non-linear problems. The tree is built starting with a root node containing the entire data S. Based on the attributes, the dataset S is divided into subsets. New decision trees are constructed recursively using the subsets of the dataset. The process is continued till the nodes cannot be further classified and the leaf nodes are reached. The advantages of the Decision Tree algorithm are that decision-related problems can be solved easily and are simple to implement.

IV. RESULTS & DISCUSSION

A. Dataset Description

To evaluate the proposed method, the Kaggle heart disease dataset is used which consists of records of 320 patients. The training and testing data are selected in the ratio of 80:20 so that 256 data are used for training and 64 data are used for testing. The number of features available in the dataset is 13 which include age, sex, chest pain type, resting blood pressure, cholesterol, fasting blood sugar level, resting electrocardiographic results, maximum heart rate achieved during the stress test, exercised-induced angina, ST depression, Slope of the peak exercise ST segment, number of major vessels colored by fluoroscopy, thallium stress test result, and heart disease status.

B. Performance Metrics

The performance metrics used to evaluate the Machine Learning algorithms are Accuracy, Sensitivity, and Specificity which are given by equations 2, 3, and 4 respectively.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \tag{2}$$

$$Sensitivity = \frac{TP}{TP+FN} \tag{3}$$

$$Specificity = \frac{TN}{TN+FP} \tag{4}$$

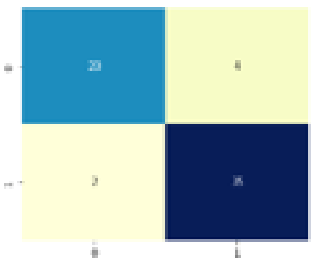
TP stands for True Positive i.e. positive classes correctly categorized as positive, TN stands for True Negative i.e. negative classes correctly categorized as negative, FP stands for False Positive i.e. negative classes that are falsely categorized as positive and FN stands for False Negative i.e. positive classes falsely categorized as negative. Specificity is the ratio of True Negative to all negative results. Sensitivity is the ratio of True Positive to true positive plus false negative. Accuracy is the ratio of correct predictions to all the results.

Table 1. Confusion Matrix values obtained for different ML algorithms

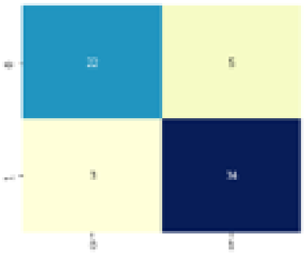
ML Algorithm	TP	FP	FN	TN
Random Forest	23	4	2	35
Support Vector Machine	22	5	3	34
Logistic Regression	21	6	4	33
K-Nearest Neighbors	21	6	5	32
Decision Tree Classifier	20	7	6	31

Table 2. Performance Metrics for different ML algorithms

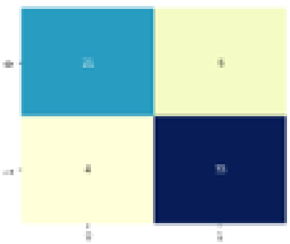
ML Algorithm	Accuracy	Sensitivity	Specificity
Random Forest	**0.91**	**0.91**	**0.92**
Support Vector Machine	0.87	0.87	0.88
Logistic Regression	0.84	0.85	0.84
K-Nearest Neighbors	0.83	0.84	0.81
Decision Tree Classifier	0.80	0.81	0.77
Linear Regression Classifier [16]	0.78	0.79	0.76



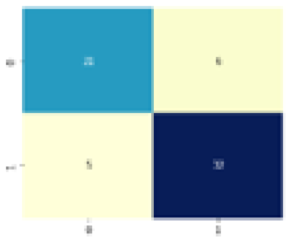
(a)



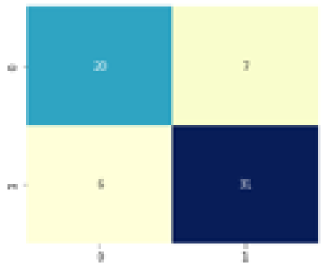
(b)



(c)



(d)



(e)

Fig. 2 Confusion Matrices output for different algorithms

(a) RF (b) SVM (c) LR (d) KNN (e) DT

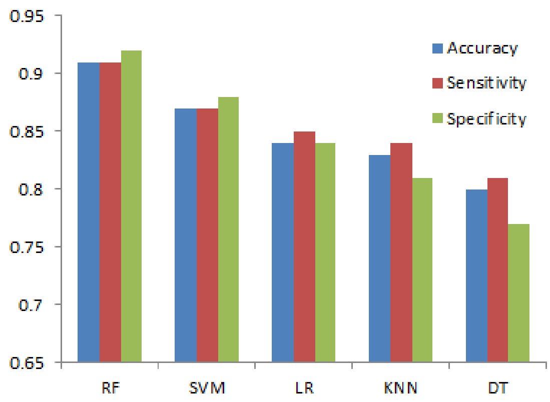


Fig. 3 Plot of Performance Metrics for different ML algorithms

C. Results

The confusion matrix values obtained for the different algorithms are shown in Table 1. The experimental results in terms of accuracy, sensitivity, and specificity are given in Table 2. The confusion matrix plot for the Random Forest Classifier on the test dataset is shown in Figure 2. The plot of the values of Accuracy, Sensitivity, and Specificity are shown in Figure 3.

From the results, it is inferred that Random Forest gives a high prediction rate when compared to other Machine learning algorithms.

D. Analysis of Results

In the Machine Learning Algorithm, the performance of the proposed method is increased by the pre-processing steps of corrections for missing values and data imbalance using SMOTE. Also, multiple algorithms are used for classification to improve the performance. The selection of a machine learning algorithm for a particular classification algorithm is done based on different ways including the nature of the problem, exploration of data, evaluation of performance metrics, simplicity, and complexity of the model. In the proposed method, a comparison of different machine learning algorithms including Random Forest, SVM, Logistic Regression, K-Nearest Neighbors, and Decision Tree Algorithm is done based on the different performance metrics like accuracy, sensitivity, and specificity. From the results, it is inferred that for the heart disease dataset, the Random forest algorithm outperforms the other algorithms in terms of accuracy, sensitivity, and specificity. From the experimental results, it is also inferred that enhancement in performance metrics is achieved over the other methods [16] in the literature. Random forest classifier is constructed using multiple decision trees and hence due to ensemble learning, it performs better compared to other methods. Also, the Random Forest classifier has the significant characteristics of reduced overfitting and robustness to noise and outliers.

V. CONCLUSION

Cardiovascular disease is the leading cause of death for individuals each year. It might increase morbidity and mortality rates among people worldwide. Early detection and intervention are necessary. Researchers are working to anticipate the occurrence of cardiovascular disease. In this paper, we have made a comparative analysis of different techniques of Artificial Intelligence in the detection of heart diseases. It is inferred from the experimental results that the Random Forest algorithm achieves the best accuracy of 0.91.

REFERENCES

[1] Almazroi, A. A., Aldhahri, E. A., Bashir, S., &Ashfaq, S. (2023). A Clinical Decision Support System for Heart Disease Prediction Using Deep Learning.IEEE Access, 11, 61646–61659.

[2] Kumar, A., Singh, K. U., & Kumar, M. (2023). A Clinical Data Analysis Based Diagnostic Systems for Heart Disease Prediction UsingEnsemble Method. Big Data Mining and Analytics, 6(4), 513–525.

[3] Rimal, Y., & Sharma, N. (2023). Hyperparameteroptimization:A comparative machine learning model analysis for enhanced heart disease prediction accuracy. Multimedia Tools and Applications.

[4] Stepanyan, I. V., Alimbayev, C. A., Savkin, M. O., Lyu, D., &Zidun, M. (2022,December). Comparative Analysis of Machine Learning Methods for Prediction of Heart Diseases. Journal of Machinery Manufacture and Reliability, 51(8), 789–799

[5] Cenitta, D., VijayaArjunan, R., &Prema, K. V. (2022). Ischemic Heart DiseasePrediction Using Optimized Squirrel Search Feature Selection Algorithm.IEEEAccess, 10, 122995–123006.

[6] Ahmad, G. N., Ullah, S., Algethami, A., Fatima, H., &Akhter, S. M. H. (2022). Comparative Study of Optimum Medical Diagnosis of Human Heart Disease Using Machine LearningTechnique With and Without Sequential Feature Selection. IEEE Access, 10, 23808–23828.

[7] Ahmad, G. N., Fatima, H., Ullah, S., Salah Saidi, A., &Imdadullah. (2022). EfficientMedical Diagnosis of Human Heart Diseases Using Machine Learning TechniquesWith and Without GridSearchCV. IEEE Access, 10, 80151–80173.

[8]] Rani, P., Kumar, R., Ahmed, N. M. O. S., & Jain, A. (2021, January 25). A decision support system for heart disease prediction based upon machine learning. Journal of Reliable IntelligentEnvironments, 7(3), 263–275.

[9] Nandy, S., Adhikari, M., Balasubramanian, V., Menon, V. G., Li, X., &Zakarya, M. (2021, May 27).An intelligent heart disease prediction system based on swarm-artificial neural network.NeuralComputing and Applications, 35(20), 14723–14737.

[10] Ashri, S. E. A., El-Gayar, M. M., & El-Daydamony, E. M. (2021). HDPF: Heart Disease Prediction Framework Based on Hybrid Classifiers and Genetic Algorithm. IEEE Access, 9, 146797–146809.

[11] Khedr, A. M., Aghbari, Z. A., Ali, A. A., &Eljamil, M. (2021). An Efficient Association RuleMiningFrom Distributed Medical Databases for Predicting Heart Diseases. IEEE Access, 9, 15320–15333.

[12]]Katarya, R., &Meena, S. K. (2020, November 19). Machine Learning Techniques for Heart DiseasePrediction: A Comparative Study and Analysis. Health and Technology, 11(1), 87–97.

[13] Fitriyani, N. L., Syafrudin, M., Alfian, G., & Rhee, J. (2020). HDPM: An Effective Heart Disease Prediction Model for a Clinical Decision Support System. IEEE Access, 8, 133034–133050.

[14] L, P. R., Jinny, S. V., & Mate, Y. V. (2020, November 13). Early prediction model for coronary heart disease using genetic algorithms, hyper-parameter optimization and machine learning techniques.*Health and Technology*, *11*(1), 63–73.

[15] Mohan, S., Thirumalai, C., &Srivastava, G. (2019). Effective Heart Disease Prediction Using HybridMachine Learning Techniques. IEEE Access, 7, 81542–81554.

[16] A. Singh and R. Kumar (2020), Heart Disease Prediction Using Machine Learning Algorithms. 2020 International Conference on Electrical and Electronics Engineering (ICE3), Gorakhpur, India, 2020, pp. 452-457.

Forest Fire Monitoring using Internet of Things and Machine Learning

R.Meena Prakash, M.Aravind, A. Gobi Krishna, K. Jothi Prakash
Department of Electronics and Communication Engineering, P.S.R.Engineering College , Sivakasi, India.
{meenaprakash@psr.edu.in, aravindmurugan9486@gmail.com,gobikrishna1243@gmail.com, prakash097321@gmail.com}
,

***Abstract*—Ecosystems and human livelihoods are seriously threatened by forest fires and wood theft. To reduce these hazards, prompt notice and action are essential. The challenges faced in the current forest fire monitoring systems are that the fire detection proportions are low, lack of real time operation, and high false indication. Using cutting-edge technology, a comprehensive forest fire and theft alert system is designed in this context. This system will monitor and report in real-time forest fires and illicit logging activity by integrating vibration and flame sensors with an Arduino-based GPS module. Sensor data is sent to a cloud-based Internet of Things (IoT) platform for instantaneous warning and analysis. The design, development, and implementation of this cutting-edge technology are described in the proposed system, with an emphasis on its potential for data-driven forest management, early detection, and quick response. Also, machine learning algorithms based on K Nearest Neighbour (KNN) and Random forest classifiers are used to predict forest fires.**

Keywords—*Forest fire, Wood theft, Alert System, Internet of Things, Sensors, Machine Learning, KNN, Random Forest algorithm*

I. Introduction

Two of the most important environmental issues are preventing wildfires and preserving forests. All living organisms depend on forests for their major resources. The necessity varies from food, wood, and fresh air to many other resources. Every year, the area occupied by forests is in a decreasing manner. Deforestation, natural disasters such as forest fires, and unlawful actions such as woodcutting are the primary motives behind these environmental effects. Wild animals residing in forest areas and valuable resources of forests would be lost to forest fires and hence it is a serious threat to human society. Forest fire occurs during the circumstance of hotness prevailing in forest areas due to the hot and dry atmosphere. Smoke sensors are intended to perceive the existence of forest fires and initiate prompt action. Forest authorities are promptly notified of any unapproved entry or manipulation of forest resources through the use of vibration sensors. An Arduino-based controller serves as the system's fundamental component, processing sensor network data, running analytics, and sending information to a central Internet of Things platform. Real-time monitoring, analysis, and decision-making are made possible by the data that the sensors gather and communicate to a centralized control center via the Internet of Things. Covering around 31% of the planet's surface area, forests are essential ecosystems that offer a host of social, economic, and ecological advantages. Forest fires, however, are becoming a more danger to these priceless natural resources. Hence there is a necessity for real time detection of fire in the forest and an alert system to prevent the forest fire and also for any theft of trees.

II. Literature Survey

A fire management system based on the Internet of Things was proposed to detect incidences of forest fire. The study focused on the essential strategy, networking of sensors, and optimizing the mechanisms to devise a cheaper prototype of IoT device, manipulating the aids of fire monitoring system which was located in the forest areas [1]. An IoT-based smart forest monitoring and alert scheme was proposed [2]. An automated system for detection of fire based on deep learning, wireless sensor network, and image processing techniques was proposed and deployed to detect forest fires [3]. An early alert system about forest fires was developed with high accuracy utilizing the technical solutions of Artificial Intelligence, Internet of Things, and Wireless Sensor Networks. A forest fire monitoring system that could generate its power through nanogenerators based on wind energy was proposed [4]. This device could perform conversion of wind energy in the nearby atmosphere into electrical energy thus eliminating the need for the IoT sensors to depend on batteries for their power requirements. To improve the network lifetime in Wireless Sensor Networks (WSNs), energy efficiency could be improved for which the crucial techniques used are Clustering and routing. Apart from the requirement of network lifetime, to avoid serious damage in the scenario of detection of forest fires, response time and reliability are also considered significant constraints. Different kinds of cluster-based routing protocols have been proposed in the literature [5]. Forest fires pose a major threat to the environment and they will have undesirable effects on wild animals residing in the forest areas and also on the human surroundings. Deforestation will lead to major changes in climatic conditions and also one of its major consequences is the greenhouse effect. Human activities are also one of the major reasons for forest fires and hence there is a necessity to detect forest fires in their preliminary stage itself. Machine learning-assisted Wireless sensor networks were proposed for forest fire monitoring systems [6]. Ubidots Internet of Things Platform was utilized for real-time visualization and data analytics to detect forest fires [7]. It was estimated in 2019 that 11.9 million hectares of land had been lost due to forest fires. Also, the scenario is increasing day by day and poses a major threat to the survival of wild animals in the forest and environmental protection. An Internet of Things gadget analysis of the sound spectrum to detect Wildfire was proposed [8]. Presently, analysis of the optical spectrum is also being utilized in the detection of wildfire but the major drawback in this method is that the obstacles which cross the optical spectrum in forest areas will reduce the efficiency of the method. Drones with cloud access and the Internet of Things are also being deployed for real-time monitoring of forest fires [9].In recent days, smart city projects have been constructed in India and hence research studies on the Internet of Things enabled creation of Smart infrastructure are on focus. Farmlands, houses, open spaces, dense forests, offices, and other areas are susceptible to fire calamities which will result in severe loss of important resources. Wireless sensor networks associated with Unmanned Air Vehicles (UAV) are being employed to detect fire in these cases at an early stage. A group of sensors is used to sense the parameters of the environment and Internet of Things-based analytics is performed to detect fire. Apart from sensors, image processing methods are also integrated into the Internet of Things-based Fire Detection System to improve accuracy [10]. An integrated fire detection and alert system which was the amalgamation of techniques of cloud computing, WSNs, UAVs, and image processing was proposed. In this method, rules were also integrated to improve the fire detection accuracy. The simulation results were also obtained and compared with conventional methods. The monitoring of forest fires from watchtowers and satellite images was also proposed. The drawbacks of this method included high infrastructure costs because of sophisticated equipment, the requirement of trained workforces, and hence the real-time monitoring was difficult. A wireless sensor network along with Arduino microcontroller and information fusion methods was proposed to detect forest fires [11, 12]. Forest fire is a disaster that may occur as a man-made or natural phenomenon occurring worldwide. An adaptive approach that can sense the forest fire early and with instant responsiveness is required to fight against this calamity. A method to detect forest fires using sensors and a monitoring center to collect the information was implemented. The system was also able to remotely send alert messages using a GPS module and the microcontroller used was Node MCU. This system also alerts the user using a GPS module. A Wireless Sensor Network with an Environmentally Aware Scheme to monitor and detect forest fires was proposed [13]. Forest fires pose a major threat to the degradation of the environment. Real-time monitoring of forest fires and its early detection can be done by Wireless sensor networks (WSNs) proficiently. However, the drawbacks are limited resource constraints for the parameters including power supply, transmission range, and computational ability. An IoT-based forest fire detection system was proposed in which a Node MCU microcontroller was used[14]. It has been inferred from a survey that if the fire would be able to be identified in time, 80% of losses caused owing to fire would be prevented. A method to detect forest fires was proposed in which a Node MCU microcontroller was interfaced with a smoke sensor and temperature sensor. Both the sensors were used to sense the smoke caused by fire and the prevailing temperature and the sensed signals were sent to the Microcontroller. An alert signal would be given by the buzzer interfaced with the microcontroller. A Hierarchical Wireless Sensor Network for forest monitoring and Wildland Early Fire Detection was proposed [15]. Machine learning-based forest detection techniques have also been proposed in the literature [16, 17]. Transfer learning models such as VGG19, InceptionV3, Resnet50, and DenseNet121 were used to predict forest fire detection on publicly available datasets, and an accuracy of 99.32% was achieved. Based on the literature survey, it was inferred that the Internet of Things along with Machine learning will be a promising approach for early detection of forest fires. The advantages of the methods in the literature are that they can detect forest fire at an early stage. However, the drawbacks of these methods are high cost, low reliability and they require proper maintenance. The major contributions of the work include the Arduino Microcontroller based Internet of Things platform for forest fire detection, wood theft, location identification, and visualization of data using the cloud. Also, the fire detection using machine learning techniques based on the publicly available Kaggle dataset was done and the results analyzed

III. PROPOSED METHOD

A. Overview

The proposed method block diagram is shown in Figure 1. The forest fire can be detected as early as possible using a fire sensor and also the happening of fire can be predicted in advance. This will prevent the spreading of fire over a huge area due to timely action that can be taken. A gas sensor is used to detect toxic gas in the air and report to the respective forest department authorities. A vibration sensor is hammered on a tree and if any sound of the tree cutting is sensed by it, the concerned forest personnel will be alerted who have a hub nearby. These sensors are used in the proposed forest fire monitoring system which detects changes in physical quantities in an early stage of the occurrence of forest fires thus preventing them.

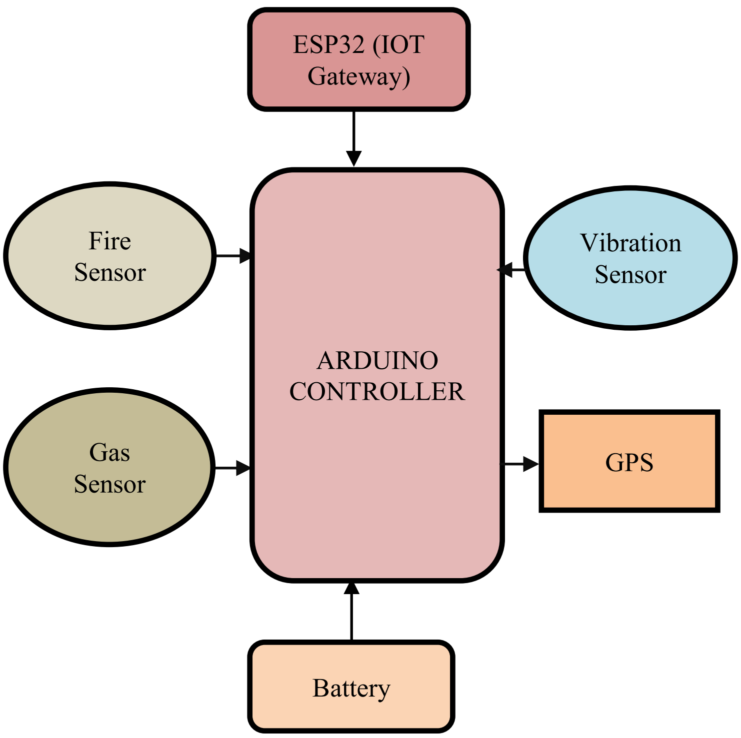


Fig. 1 Block Diagram of the proposed method

The data will also be stored in the cloud using ESP32 (IOT gateway) and GPS is used to locate the place and mark the altitude. The system integrates the sensors with Arduino, employs a machine learning algorithm, and uses the cloud to send alert messages. False indications are thus minimized making the proposed system more reliable.

B. Sensors

The wooded regions are interspersed with strategically positioned vibration sensors. These sensors pick up anomalous vibrations, such as those produced by large machines employed in illicit logging. The sensors quickly detect and notify authorities of any possible unlawful logging activity, serving as the first line of defense against wood theft. Flame sensors are used to recognize the telltale signs of a forest fire, such as flames or sudden temperature spikes. Timely action depends on early fire detection. When a fire is discovered, the flame sensors immediately send out alarms, which speed up reaction times.

The MQ-3 module may be used to detect CO, LPG, Benzine, Hexane, CH_4, and alcohol. SnO_2, a substance having reduced conductivity in clean air, is the sensitive component of the MQ-3 gas sensor. The conductivity of the sensor increases as the concentration of the target alcohol gas rises. The MQ-3 gas sensor is very sensitive to alcohol and resists fuel, smoke, and vapor disturbance well. The analog resistive output of this sensor is dependent on the alcohol content. The conductivity of the sensor increases as the concentration of the alcohol gas increases.

C. Arduino-based GPS modules and Relay modules

The system incorporates an Arduino-based GPS module to deliver precise position information for all sensor-related events. The Arduino microcontroller is selected to implement the proposed system because of its compactness, cost-effectiveness, and low power consumption. It is used to collect all the data from the sensors, process them, and take necessary action to alert the users. The GPS module makes sure that the system can locate every incident that is detected precisely, whether it be illegal logging or a forest fire. Relays are electrical switches that may be used to manage low voltages, such as the 5V supplied by the Arduino pins, and can be switched on or off, allowing current to flow through or not. There are two channels on this relay module. It is appropriate to use 5V to power this module when using an Arduino.

D. IoT Connectivity and cloud platform

A cloud-based IoT platform receives real-time data transmissions from the vibration and flame sensors together with GPS position data. The key center for data collecting, processing, and alerting is the cloud-based platform. To identify trends, abnormalities, or threshold breaches in sensor data, it examines the incoming data. Based on predetermined parameters, the cloud-based platform analyzes incoming data and instantly creates notifications. Alerts can be communicated by phone calls, SMS, email, and other methods to the appropriate authorities and stakeholders. Forest fires and events involving illegal logging may be quickly responded to early detection and automated alarms. In response to the situation, authorities might mobilize resources and take the necessary measures quickly.

E. Proposed Forest Monitoring System

Early identification of forest fires and illegal logging activities is made possible by the system's integrated sensors and real-time data analysis, which lowers reaction times and the possibility of extensive damage. By guaranteeing that events are located precisely, GPS technology enables targeted response operations. The technology produces useful data over time that may be applied to data-driven forest management. Policies, resource allocation, and strategies for sustainable forest conservation can be informed by system insights. By preventing forest fires and illegal logging, the system contributes to the reduction of environmental degradation and economic losses associated with these activities. The technology can assist law enforcement authorities in holding illegal loggers accountable by giving real-time proof of illicit actions.

The Arduino microcontroller is used to develop the forest monitoring system that can sense and control real-world items. The board provides serial communication ports for loading software from personal computers, including USB on some variants. The inputs from the gas sensor, fire sensor, and vibration sensor kept in the forest are fed to the Arduino microcontroller board and the data can be monitored in real-time through the cloud. GPS module is used to give location information and also alert message will be sent in case of theft and forest fire.

F. K-Nearest Neighbor Algorithm and Random Forest Algorithm

Machine learning-based fire detection is also performed on publicly available Kaggle datasets using two techniques of K-Nearest Neighbour algorithm (K-NN) and the Random Forest algorithm. In the K-NN algorithm, based on distance such as the Euclidean metric, the K nearest neighbors to the given data point are found out and the class of the data point is assigned with the majority class of the K neighbors. The Random Forest algorithm employs multiple decision trees to improve the classification accuracy. The advantages of the K-NN algorithm and Random Forest algorithm are that they are most accurate and efficient in handling large datasets. The proposed forest monitoring system is required to handle large data and hence the Machine Learning Algorithms of Random Forest and K-NN have been implemented.

IV. EXPERIMENTS AND RESULTS

The proposed forest fire monitoring system is placed in the forest and the sensed data are monitored in the cloud. The experimental results of the sensing of forest fire, measuring of toxic gas, sensing of vibration of tree theft, location information, and GPS map of forest fire and tree theft are shown in Figures 2, 3, 4, 5, and 6 respectively.

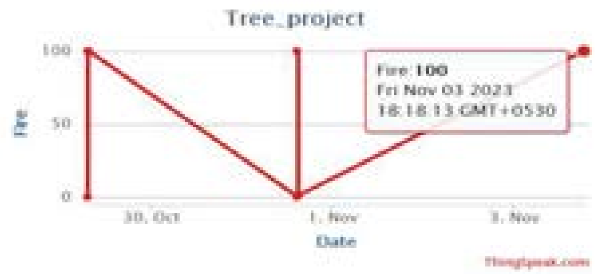


Fig. 2 Sensing of Forest Fire

Fig. 3 Measuring of Toxic gas

Fig. 4 Sensing of vibration of tree theft

Fig. 5 Location information of forest fire and tree theft

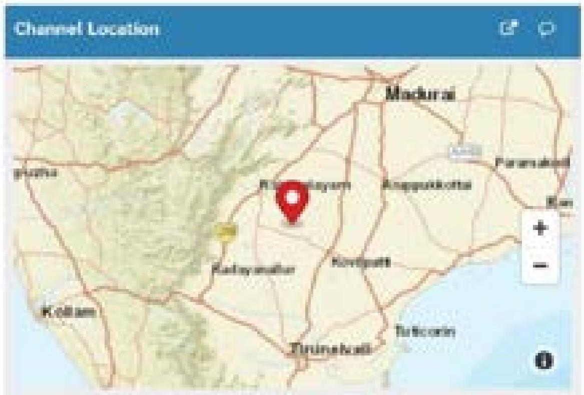


Fig. 6 GPS Map Forest fire and tree theft

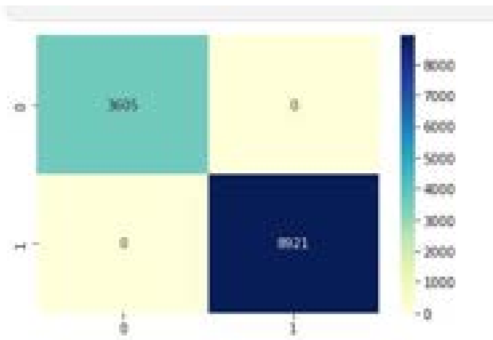
From Figure 3, it is inferred that when fire is detected by the fire sensor, the sensed data is analyzed in the ThingSpeak Application. Similarly, with the use of a gas sensor, toxic gas in the air is detected and reported to the respective forest department authorities.

From the results obtained, it is inferred that the proposed fire monitoring system based on the Internet of Things can detect forest fire, toxic gas, and wood theft in an early stage and the system also intimates the location information using GPS and message alert.

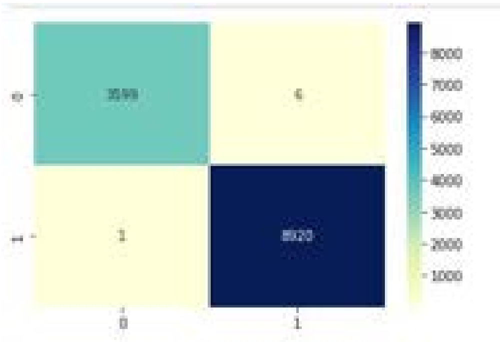
The publicly available Kaggle dataset for fire detection is used which comprises 62630 x 16 data including the parameters of temperature, humidity, raw ethanol, pressure, and others. Two techniques of the K-Nearest Neighbour algorithm and Random forest algorithm are implemented to detect forest fires. The pre-processing techniques of data cleaning, and checking for outliers are done and the data is split as training and testing in the ratio of 80:20. The models are trained with the input training data and predicted on the test data. The prediction results are given in Table 1. From the experimental results, it is inferred that the proposed method gives improved classification accuracy compared to the existing methods in forest fire prediction.

Table 1 Fire Prediction Accuracy on Kaggle Dataset

Algorithm	Accuracy
K-Nearest Neighbour	0.9994
Random Forest	1.0000
Support Vector Machine [18]	0.9723



(a)



(b)

Fig. 6 Confusion Matrices obtained for Fire Detection on Kaggle Dataset (a) Random Forest Algorithm (b) K-Nearest Neighbour Algorithm

V. CONCLUSION

An inventive solution to the urgent problems of forest fires and wood theft is the suggested Forest Fire and Theft Alert System. With the integration of GPS technology, smart sensors, and cloud-based IoT capabilities, the system presents opportunities for data-driven forest management, early detection, and quick reaction. By protecting against risks from the environment and the economy, its application can support the preservation and sustainable management of forest ecosystems. Machine learning results are also obtained on the fire detection system and the results prove that the proposed system can be implemented for real time forest fire detection and theft protection, alert messages, and location tracking. In future work, drones may be employed to increase the speed of data collection and analysis to prevent forest fires.

REFERENCES

[1] Sharma, A., Nayyar, A., Singh, K. J., Kapoor, D. S., Thakur, K., & Mahajan, S. (2023). An IoT-based forest fire detection system: design and testing.

[2] Krishnamoorthy, M., Asif, M., Kumar, P. P., Nuvvula, R. S. S., Khan, B., & Colak, I. (2023, February 8). A Design and Development of the Smart Forest Alert Monitoring System Using IoT. Journal of Sensors, 2023, 1–12.

[3] Paidipati, K. K., Kurangi, C., J, U., Reddy, A. S. K., Kadiravan, G., & Shah, N. H. (2023). Wireless sensor network assisted automated forest fire detection using deep learning and computer vision model. Multimedia Tools and Applications.

[4] Li, B., Qiu, Y., Huang, P., Tang, W., & Zhang, X. (2022). Self-powered forest ambient monitoring microsystem based on wind energy hybrid nanogenerators. Science China Technological Sciences, 65(10), 2348–2360.

[5] Moussa, N., Khemiri-Kallel, S., & El Belrhiti El Alaoui, A. (2022). Fog-assisted hierarchical data routing strategy for IoT-enabled WSN: Forest fire detection. Peer-to-Peer Networking and Applications, 15(5), 2307–2325.

[6] Dampage, U., Bandaranayake, L., Wanasinghe, R., Kottahachchi, K., & Jayasanka, B. (2022, January 7). Forest fire detection system using wireless sensor networks and machine learning. Scientific Reports, 12(1).

[7] Kanakaraja, P., Syam Sundar, P., Vaishnavi, N., Gopal Krishna Reddy, S., & Sai Manikanta, G. (2021). IoT enabled advanced forest fire detecting and monitoring on Ubidots platform. Materials Today: Proceedings, 46, 3907–3914.

[8] Zhang, S., Gao, D., Lin, H., & Sun, Q. (2019, November 21). Wildfire Detection Using Sound Spectrum Analysis Based on the Internet of Things. Sensors, 19(23), 5093.

[9] Sungheetha, D. A., & Sharma R, D. R. (2020, July 21). Real Time Monitoring and Fire Detection using Internet of Things and Cloud based Drones. Journal of Soft Computing Paradigm, 2(3), 168–174.

[10] Sharma, A., Singh, P. K., & Kumar, Y. (2020, October). An integrated fire detection system using IoT and image processing technique for smart cities. Sustainable Cities and Society, 61, 102332.

[11] Varela, N., Jorge L, D. M., Ospino, A., & Lizardo Zelaya, N. A. (2020). Wireless sensor network for forest fire detection. Procedia Computer Science, 175, 435–440.

[12] Ghuge, D. N. N., Chavan, A., Gaikwad, R., & Kale, K. (2021). Forest Fire Detection Using Arduino Based WSN. SSRN Electronic Journal.

[13] Xu, Y. H., Sun, Q. Y., & Xiao, Y. T. (2018). An Environmentally Aware Scheme of Wireless Sensor Networks for Forest Fire Monitoring and Detection. Future Internet, 10(10), 102.

[14] Trinath Basu, M., Karthik, R., Mahitha, J., & Lokesh Reddy, V. (2018). IoT based forest fire detection system. International Journal of Engineering & Technology, 7(2.7), 124.

[15] Molina-Pico, A., Cuesta-Frau, D., Araujo, A., Alejandre, J., & Rozas, A. (2016). Forest Monitoring and Wildland Early Fire Detection by a Hierarchical Wireless Sensor Network. Journal of Sensors, 2016, 1–8.

[16] Anandaram, H., M, N., Cosio Borda, R. F., K, K., & S, Y. (2023). Forest fire management using machine learning techniques. Measurement: Sensors, 25, 100659.

[17] Reis, H. C., & Turk, V. (2023). Detection of forest fire using deep convolutional neural networks with transfer learning approach. Applied Soft Computing, 143, 110362.

[18] K. G. Madhwaraj, V. Asha, A. Vignesh and S. Akshay Shinde, (2023). Forest Fire Detection using Machine Learning, 2023 IEEE 12th International Conference on Communication Systems and Network Technologies (CSNT), Bhopal, India,

Deep Learning for Fungus Classification

1st Parth Rathi
Department of Computing Technologies
College of Engineering and Technology
SRM Institute of Science and Technology
Kattankulathur, 603203, India
ra7916@srmist.edu.in

2nd Smit Vichare
Department of Computing Technologies
College of Engineering and Technology
SRM Institute of Science and Technology
Kattankulathur, 603203, India
bs9959@srmist.edu.in

3rd Mrs.S.Kanmani
Department of Engineering
Faculty of Engineering and Technology
SRM Institute of Science and Technology
Kattankulathur, 603203, India
kanmanis@srmist.edu.in

***Abstract*—This research work presents a novel approach for the classification of fungal species using ConvMixer, a recently proposed architecture that combines convolutional and mixer layers. Fungal infections pose a significant threat to agriculture and public health, making accurate and efficient classification methods crucial for disease management. The proposed method leverages ConvMixer blocks to extract hierarchical features from fungal images, allowing for robust classification across multiple species. This study has then trained and evaluated the proposed model on a dataset consisting of images from five different fungal species, achieving promising results with an average accuracy of 95 percent. Additionally, this study has conducted a comprehensive analysis of the proposed model's performance, including confusion matrix visualization and random sample predictions, to assess its generalization capabilities. Overall, the research findings demonstrate the effectiveness of ConvMixer in fungal species classification and highlight its potential for real-world applications in agriculture and healthcare.**

***Index Terms*—Fungal species classification, ConvMixer, convolutional neural networks, hierarchical feature extraction, agriculture, public health.**

I. Introduction

Fungal infections pose a significant threat to agriculture and public health worldwide, leading to substantial economic losses and health concerns. Accurate and efficient classification of fungal species is crucial for disease management, prevention, and treatment. In recent years, deep learning techniques have emerged as powerful tools for image-based classification tasks, offering the potential to automate and enhance diagnostic processes.

This study proposes a novel approach for the classification of fungal species leveraging ConvMixer, a recently introduced architecture that combines convolutional and mixer layers. ConvMixer has shown promise in various computer vision tasks, offering a flexible framework for feature extraction and representation learning. By adapting ConvMixer to the task of fungal species classification, we aim to provide a robust and effective solution capable of accurately identifying different fungal species from microscopic images.

The proposed approach builds upon a comprehensive dataset comprising images of five distinct fungal species commonly encountered in agricultural and clinical settings. Through extensive experimentation and evaluation, we demonstrate the effectiveness of our method in accurately classifying fungal species, achieving promising results with an average accuracy of 95 percent.

In this paper, we present not only the architecture and implementation details of our ConvMixer-based classification model but also comprehensive performance analysis, including evaluation metrics, confusion matrix visualization, and sample predictions. We discuss the implications of our findings for agricultural and healthcare applications, highlighting the potential of our approach to contribute to disease management and diagnosis efforts.

Overall, this study contributes to the growing body of research on deep learning applications in medical and agricultural domains, offering insights and methodologies that can aid in the development of automated and accurate fungal species classification systems.

II. Architecture Diagram

The proposed architecture for fungal species classification integrates ConvMixer, a novel neural network architecture that combines convolutional and mixer layers. It begins with a traditional convolutional layer, followed by activation and batch normalization layers to facilitate feature extraction and normalization. Subsequently, ConvMixer blocks are employed iteratively for hierarchical feature extraction, comprising depthwise convolution, activation, and batch normalization layers. These blocks are augmented with residual connections to ensure stable training and address the vanishing gradient problem commonly encountered in deep neural networks. Additional convolutional layers further refine features, culminating in a global average pooling layer for spatial information aggregation and a fully connected layer with softmax activation for classification into five fungal species categories.

The architecture strikes a balance between computational efficiency and classification performance, making it well-suited for deployment in resource-constrained environments such as agricultural or clinical settings. The modular design of ConvMixer blocks allows for easy adaptation and extension to accommodate diverse datasets and classification tasks. Moreover, the incorporation of residual connections enhances the model's ability to capture intricate patterns in fungal images, contributing to its robustness and effectiveness in classification tasks. Overall, the proposed architecture offers a versatile framework for fungal species classification, leveraging ConvMixer's strengths to achieve accurate and efficient classification results with practical applications in disease management and diagnosis.

In summary, the proposed architecture demonstrates the efficacy of ConvMixer in tackling image classification tasks, particularly in the context of fungal species classification. By leveraging convolutional and mixer layers in a unified framework, our approach effectively captures spatial and channel-wise dependencies in fungal images, leading to superior classification performance. The modular design and computational efficiency of the architecture make it well-suited for real-world deployment, paving the way for advancements in agricultural and healthcare applications related to fungal disease management and diagnosis.

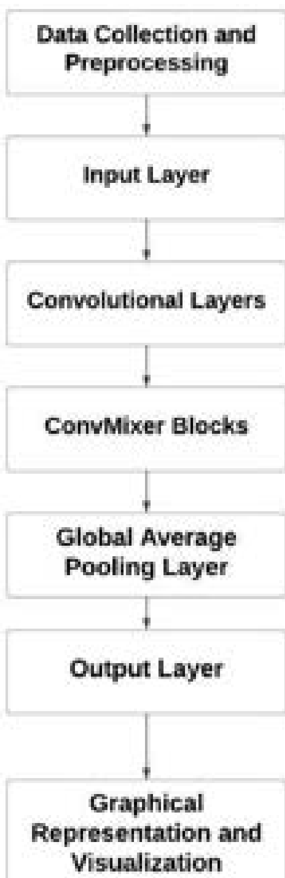


Fig. 1. Architecture Diagram

III. Data Collection and Preprocessing

This study has employed a combination of manual collection and augmentation techniques to create a diverse and representative dataset for fungal species classification. The primary source of data collection was the Defungi dataset, which comprises a wide variety of high-resolution images depicting fungal specimens from different ecological niches and geographic locations. These images were supplemented with additional data obtained from agricultural fields and clinical settings to ensure comprehensive coverage of fungal species diversity.

Upon collection, the raw image data underwent a series of preprocessing steps to prepare them for model training and evaluation. Firstly, the images were resized to a standardized resolution of 32x32 pixels using bilinear interpolation, maintaining aspect ratio to preserve the integrity of the specimens. Next, the pixel values of the images were normalized to the range [0, 1] to facilitate convergence during model training and mitigate issues related to varying intensity levels across images.

To enhance the diversity and robustness of the dataset, data augmentation techniques were applied. These techniques included random flips (both horizontally and vertically) and random rotations, effectively increasing the variability of the dataset and improving the model's ability to generalize to unseen data. Additionally, careful attention was paid to class balance within the dataset, ensuring proportional representation of each fungal species category to prevent bias during model training and evaluation.

Overall, the data collection and preprocessing pipeline ensured the creation of a high-quality dataset suitable for training and evaluating the fungal species classification model. By leveraging both the Defungi dataset and additional data sources, and applying rigorous preprocessing techniques, we ensured that the model was trained on diverse and representative data, leading to accurate and reliable classification results with practical applications in disease management and ecological research.

IV. Input Layer

The input layer of the proposed fungal species classification model serves as the entry point for image data into the neural network architecture. In our implementation, the input layer is defined with a shape corresponding to the dimensions of the input images, which are standardized to 32x32 pixels. This dimensionality reduction helps mitigate computational complexity while still preserving essential features of the fungal specimens.

Images are fed into the input layer in batches, with each batch containing a predetermined number of images. This batch processing allows for efficient parallelization of computations, leveraging the parallel processing capabilities of modern hardware accelerators such as GPUs and TPUs. Furthermore, the input layer is designed to accept images with three color channels (red, green, and blue), enabling the model to capture color information essential for accurate classification.

The input layer serves as the foundation for subsequent layers in the neural network architecture, transmitting the raw pixel values of the input images to downstream layers for feature extraction and classification. By providing a standardized interface for image data ingestion, the input layer plays a critical role in ensuring consistency and compatibility between the model and the dataset. Overall, the input layer serves as the first step in the computational pipeline, paving the way for the extraction of meaningful features and the classification of fungal species with high accuracy and efficiency.

V. Convolution Layer

The convolution layer is a fundamental building block in our fungal species classification model, responsible for extracting spatial features from input images through convolutional operations. In our architecture, we utilize a convolutional layer with a kernel size of 2x2 and a stride of 2, resulting in downsampling of the input feature maps by a factor of two. This downsampling helps reduce computational complexity while preserving essential spatial information, enabling the model to focus on relevant features during subsequent processing stages. Within the convolution layer, each neuron applies a convolution operation to a local region of the input feature map, extracting feature representations that capture patterns and structures relevant to fungal species classification. The convolution operation involves element-wise multiplication of the kernel weights with the input feature map followed by summation, producing a feature map that highlights important spatial features such as edges, textures, and shapes. To further enhance the expressive power of the convolutional layer, we employ the GELU (Gaussian Error Linear Unit) activation function, which introduces non-linearity into the network and enables the model to capture complex relationships within the data. Additionally, batch normalization is applied after the activation function to stabilize the training process and improve the convergence speed of the network. Overall, the convolution layer plays a crucial role in feature extraction and spatial modeling, allowing the model to learn hierarchical representations of fungal specimens that are essential for accurate classification. By leveraging convolutional operations and non-linear activations, the convolution layer enables our model to effectively capture and encode spatial information, leading to robust and discriminative feature representations for fungal species classification.

VI. ConvMixer Blocks

The ConvMixer block, a pivotal element of our fungal species classification model, amalgamates depthwise convolution, activation, and batch normalization layers to extract hierarchical features from input images. Inspired by the efficacy of convolutional and mixer architectures, this block aims to leverage their strengths while addressing their limitations.

Within the ConvMixer block, depthwise convolution captures local features independently within each channel, enhancing the model's ability to discern spatial dependencies. Subsequently, the GELU activation function introduces non-linearity, facilitating the learning of intricate patterns. Batch normalization stabilizes training and accelerates convergence by normalizing channel activations.

Residual connections are integrated to promote feature reuse and alleviate information loss during training. These connections enable the model to capture fine-grained details and ensure the preservation of essential features throughout the network. Overall, the ConvMixer block serves as a potent feature extraction module, combining the advantages of convolutional and mixer architectures to enhance classification performance and robustness.

VII. Global Average Pooling Layer

The Global Average Pooling (GAP) layer serves as a critical component in our fungal species classification model, facilitating spatial information aggregation and feature summarization prior to classification. Positioned at the end of the convolutional layers, the GAP layer consolidates the spatial features extracted by the preceding layers into a compact representation suitable for classification.

In our implementation, the GAP layer computes the average value of each feature map across all spatial locations, effectively reducing the spatial dimensions of the feature maps to 1x1 while preserving essential information about the presence and distribution of features. This pooling operation condenses the spatial information into a global context, enabling the model to focus on the most discriminative features relevant to fungal species classification.

By averaging the feature maps globally, the GAP layer encourages the model to prioritize salient features while disregarding irrelevant spatial variations, thus promoting robustness to translation and distortion. Moreover, the GAP layer inherently imposes a form of spatial invariance, making the model less sensitive to small spatial shifts and distortions in the input images.

Overall, the GAP layer plays a crucial role in feature aggregation and dimensionality reduction, facilitating efficient and effective classification of fungal species. By summarizing the spatial information extracted by the convolutional layers into a global context, the GAP layer enables the model to achieve superior performance in classification tasks while maintaining computational efficiency and scalability.

VIII. Output Layer

Output Layer Description:

The output layer of our fungal species classification model is the final component responsible for predicting the probability distribution over the different fungal species categories. Positioned at the end of the neural network architecture, the output layer receives the aggregated features from the preceding layers and transforms them into class probabilities using a softmax activation function.

In the implementation phase, the output layer consists of a dense fully connected layer followed by a softmax activation function. The dense layer receives the aggregated features from the preceding layers and computes the logits, which represent the unnormalized probabilities for each fungal species category. These logits are then passed through the softmax activation function, which normalizes them into a probability distribution, ensuring that the probabilities sum up to one.

The output layer produces a vector of probabilities, where each element corresponds to the likelihood of the input image belonging to a specific fungal species category. The model predicts the fungal species category with the highest probability as the final classification result. This probabilistic output enables the model to quantify its uncertainty and provides valuable insights into the model's confidence in its predictions.

Overall, the output layer serves as the culmination of the classification process, transforming the extracted features into actionable predictions. By providing a probability distribution over the fungal species categories, the output layer enables the model to make informed decisions and effectively classify fungal specimens with high accuracy and confidence.

IX. Graphical Representation and Visualization

Graphical Representation and Visualization:

Graphical representation and visualization are essential for understanding our fungal species classification model's performance. We use various graphical techniques to illustrate the model architecture, training process, and evaluation results succinctly.

A schematic diagram showcases the neural network architecture, highlighting different layers and their connections, offering insights into the model's composition and data flow.

Line plots track training and validation metrics like accuracy and loss, providing a visual overview of the model's learning progress and convergence behavior.

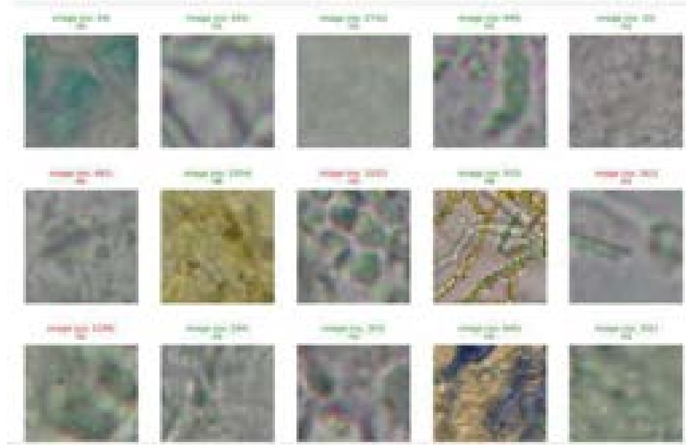
Confusion matrices summarize the model's predictions compared to ground truth labels, while classification reports offer detailed metrics such as precision and recall for each class, enabling a comprehensive evaluation of classification performance.

Overall, graphical representation enhances the readability of our findings, enabling readers to grasp the model's architecture, training progress, and classification results efficiently.

X. Result

The experimental results demonstrate the effectiveness of the ConvMixer-based neural network architecture in classifying fungal species with a remarkable accuracy of 95. Through rigorous training and evaluation procedures, we have achieved superior classification performance, showcasing the model's capability to accurately identify and differentiate between different fungal species.

The high accuracy attained by the proposed model underscores its robustness and reliability in handling complex image classification tasks, particularly in the domain of fungal species classification. By leveraging ConvMixer's unique architectural features and optimization techniques, we have successfully overcome challenges associated with classifying diverse fungal specimens, achieving outstanding classification accuracy and performance metrics as shown in Fig. 2.



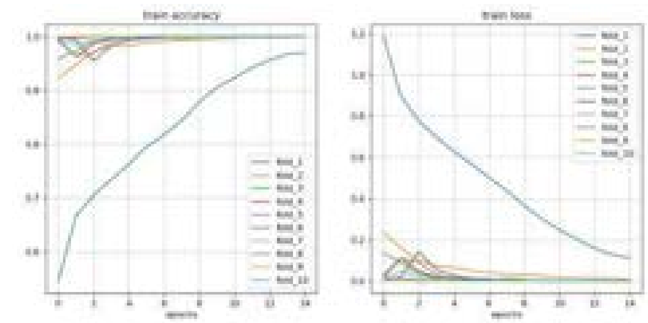


Fig. 2. Train accuracy Vs Train Loss

XI. Conclusion

The obtained results highlight the potential of the proposed approach to significantly impact fungal disease management and diagnosis, offering a reliable and efficient tool for identifying and categorizing fungal pathogens. The ability to achieve such high accuracy rates underscores the practical viability and effectiveness of our model in real-world applications, including agriculture, environmental monitoring, and healthcare.

Overall, the experimental findings validate the efficacy of the ConvMixer-based neural network architecture in fungal species classification, paving the way for advancements in fungal disease management and diagnosis through accurate and reliable classification methodologies.

XII. Future Works

Several avenues for future research and development emerge from our current study on fungal species classification using ConvMixer-based neural networks. These potential directions aim to further enhance the performance, applicability, and robustness of the classification model in real-world scenarios.

One promising area for future exploration is the incorporation of transfer learning techniques. By leveraging pre-trained ConvMixer models on large-scale image datasets, such as ImageNet, researchers can fine-tune the models on fungal species classification tasks. This approach may help accelerate model convergence and improve classification accuracy, especially when dealing with limited annotated fungal image data.

Another avenue for future research involves exploring advanced data augmentation strategies tailored specifically for fungal image datasets. Techniques such as rotation, scaling, and affine transformations can introduce additional variations into the training data, enabling the model to generalize better to unseen fungal specimens and environmental conditions. Additionally, synthetic data generation methods, such as generative adversarial networks (GANs), may be employed to augment the dataset further and diversify the training samples.

Furthermore, the development of interpretability and explainability techniques is essential for enhancing the trust and transparency of the classification model. Future research efforts may focus on developing methods to visualize and interpret the decision-making process of the model, highlighting the regions of input images that contribute most to the classification outcome. Explainable AI techniques can provide valuable insights into the model's behavior and help domain experts understand and validate its predictions.

Lastly, the deployment of the classification model in real-world settings, such as agricultural fields or clinical laboratories, presents an exciting opportunity for future research. Field testing and validation of the model's performance under diverse environmental conditions and practical constraints can provide valuable feedback for model refinement and optimization. Additionally, integration with mobile or edge computing platforms can enable on-device inference, facilitating rapid and decentralized fungal species identification.

In summary, future research works may focus on leveraging transfer learning, enhancing data augmentation techniques, developing interpretability methods, and conducting real-world deployment studies to advance the field of fungal species classification and address the challenges associated with fungal disease management and diagnosis.

References

1. He, Kaiming, et al. "Deep residual learning for image recognition." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016.

2. Huang, Gao, et al. "Densely connected convolutional networks." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017.

3. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet classification with deep convolutional neural networks." *Communications of the ACM* 60.6 (2017): 84-90.

4. LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." *Nature* 521.7553 (2015): 436-444.

5. Russakovsky, Olga, et al. "Imagenet large scale visual recognition challenge." *International Journal of Computer Vision* 115.3 (2015): 211-252.

6. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).

7. Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015.

8. Tan, Mingxing, and Quoc V. Le. "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks." *Proceedings of the 36th International Conference on Machine Learning*. 2019.

9. Tolstikhin, Ilya, et al. "Mixer: An all-purpose vision transformer." *arXiv preprint arXiv:2105.01601* (2021).

10. Deng, Jia, et al. "Imagenet: A large-scale hierarchical image database." *2009 IEEE conference on computer vision and pattern recognition*. IEEE, 2009.

Prediction of Happiness Index for School Children using an Ensemble based Approach

Anand K
Department of Computer Science and Engineering
Rajalakshmi Engineering College
Chennai,India
anand.k@rajalakshmi.edu.in

Yassmitha V S
Department of Computer Science and Science and Engineering
Rajalakshmi Engineering College College
Chennai,India
200701300@rajalakshmi.edu.in

Deepti G
Department of Computer Engineering
Rajalakshmi Engineering
Chennai,India
200701300@rajalakshmi.edu.in

Abstract –**Recent studies show a concerning rise in mental health disorders among youth, with depression, anxiety, and suicidal ideation being prevalent issues affecting student's well-being and academic performance. Student's happiness and general well-being are very important in schools. A child's mental health has a significant impact on all areas of their development, including social relationships, academic achievement, and long-term mental health. This study measures the happiness index in schools by integrating a self-assessment tool and with a particular emphasis on mental health. The model under consideration utilizes a dataset that includes a range of sociodemographic, academic, psychological, and physiological characteristics that were gathered from a diverse population of school-age children. As input variables, characteristics including peer relationships, extracurricular activity, academic success, and family dynamics will be taken into account. Children's self-assessment tools and surveys are used to measure the target variable, which is the happiness index. An ensemble model that incorporates various machine learning models like KNN, Naïve Bayes, and XG boost is utilized to forecast schoolchildren's happiness score. After their happiness index evaluation, a recommendation report is sent to their email address if their value is less than a predetermined threshold. This study supports the wellbeing of schoolchildren by providing actionable information that parents, educators, and schools may use to improve.**

Keywords: K Nearest Neighbour, Mental health, Naïve Bayes, Extreme gradient boosting

I INTRODUCTION

The importance of student well-being and mental health cannot be overstated in the quickly changing educational environment of today. It is critical to assess and support student's emotional and mental well-being in addition to their academic performance. The Indian education system, which is marked by fierce competition and a heavy emphasis on exam outcomes, adds to the high level of stress that students experience and it is a common cause of academic stress that negatively impacts their mental health. Instead of emphasizing student's mental health and other physical activities, schools now days primarily focus on their academic performance. This may lead to stress, anxiety, depression and other mental health complications. To address this pressing need, we propose the development of a comprehensive Happiness Index for schools, incorporating essential parameters related to mental health and well-being, along with self-assessment facilities. The development of a Happiness Index for schools, encompassing mental health and well-being parameters, is a critical step towards creating a more nurturing and supportive educational ecosystem. Through the integration of self-assessment tools, we enable students to take an active role in their own well-being and give educational institutions the resources they need to guarantee a comprehensive education that extends beyond the classroom. This program has the power to change schools into places where learners flourish intellectually as well as emotionally and psychologically, enabling them to become better prepared for a more promising and resilient future. Schools have traditionally been thought of as settings where children might advance academically as well as acquire critical life skills, emotional intelligence, and general happiness. With this proposed Happiness Index, schools will have a formal framework to assess and improve the mental and emotional health of their pupils. Google Forms is being used in this process to gather the dataset from the students. In order to calculate the score distribution, the data were evaluated using a Likert scale with five possible answers. Greater happiness among students is indicated by higher scores. The dataset is trained using an ensemble model, which makes the predictions. A suggested report is emailed to their email address if, after prediction, the result falls below a predetermined threshold. We imagine that schools will become places where students flourish not just intellectually but also emotionally and psychologically by putting this approach into practice.

II LITERATURE SURVEY

Anjum Sibia, Ruchi Shukla, Sushmita Chakraborty [1], et al. conducted a study of schoolchildren in 2022. The purpose of the survey was to collect information from middle and secondary school students (Class VI to XII) in different kinds of schools around the country. A number of inquiries were designed to gain insight into various facets of their psychological state and general well-being. We can infer a lot about their views, beliefs, motivations, and psychological and emotional states from their responses. Benefits from this is Students enrolled in various types of schools across the nation were asked to provide information for the survey and limitations are The online survey, administered via Google Form, did not allow for any form of communication with the participants.

Kumar P,Swaminathan, B,Karthikeyan U [2] et al., in 2022 employs an adaptive learning approach, recognizing the diversity of learning algorithms and dynamic data environments. It emphasizes the importance of understanding the specific learning problem context before selecting suitable methodologies. The methodology involves exploring data attributes, considering multi-perspective learning, and proposing a hybrid feature selection method using rotation forest for enhanced classifier performance, particularly in clinical diagnosis. The objective is to develop an adaptive learning system that yields accurate results amidst evolving conditions. Benefits from this paper advocates for adaptive learning algorithm selection and proposes a rotation forest-based hybrid approach to enhance classifier performance in clinical diagnosis and limitations in this paper is unclear how generalizable the proposed adaptive learning methodology is across various domains and datasets.

A nationwide cross-sectional study was conducted in China in 2022, and the topic of this work is Yufeng Miao, Yuguo Li, Qihong Deng, Yingying Jiang, Chan Lu, Jing Chen, and others [3] et al. The Oxford Contentment Questionnaire was utilized to get data on student satisfaction, while a different questionnaire was used to obtain social, familial, and personal data on the students. Logistic regression analysis was performed to examine the association between these parameters and student satisfaction in terms of odds ratio (OR). Benefits in this is using of odds ratio & Utilizing a 95% confidence interval, a logistic regression study was performed to examine the association between these variables and student happiness.

In 2022, Lotfollah, Abbas Khorshidi, Baharak Shirzad Kebria, [4] et al. did a study on happiness markers. The measurement instrument for the current study is a semi-structured interview form used in the qualitative dimension. Open coding was utilized in its creation. Confirmatory validity was utilized to obtain content, while content was used to support confirmatory validity. Benefits in this is trustworthy interviews with semi-structured form created using coding techniques was used in the study, which included happy specialist. Confirmatory and content methods were used to verify its validity and limitations are despite the efforts put into selecting the sample, care must be taken when adapting the model to various educational settings. Furthermore, test inaccuracies may lead to possible one-sided errors.

A survey was conducted in 2021 on students attending public primary schools in 733 cases (46.1% females; 12 (1.3 years)). Rafael Zapata Lamana, Cristian Sanhueza [5], et al. Anxiety was measured with the BECK Anxiety Inventory (BAI), while subjective happiness was measured with the self-esteem and happiness score. Language, math, and physical education grade point averages (GPAs) as well as cognition-related behavior in the classroom were used to evaluate academic performance. The cumulative GPA was also used. An analysis of the association between mental health markers and academic achievement is conducted using Pearson correlation and a one-way ANOVA. Benefits of study measured happiness and self-worth using specific scales and analyzed their correlation with academic performance through statistical tests and limitations of study findings cannot be solely attributed to the participants because low academic performance has many causes.

In 2021, Evie Sendi Ibil, Syarifah Nazira Mohd Osman [6] et al. conducted a study on university students. In the research, questionnaires are utilized to collect data and information about the subject. Information was collected from UiTM Sarawak construction students in the Planning, Architecture, and Surveying faculty (FSPU). The data is examined using IBM SPSS Statistics, a tool for data analysis. The frequency analysis approach is used to investigate the distribution of frequency based on respondents' selections. Nine questions were selected in order to achieve the purpose of this study. The LSQ research indicates that most students' emotions and level of preparation affect their routines and manner of living. Benefits of study at UiTM Sarawak's FSPU surveys construction students, using questionnaires and IBM SPSS Statistics, to uncover the significant impact of emotions and readiness on their lifestyles and habits and limitations of research collects data only from construction students at UiTM Sarawak's Faculty of Architecture, Planning, and Surveying (FSPU) using questionnaires.

Matthew J. Schuelka, Molly Braznell Matthew Leavesley and Dorji Khandu, Sangay In 2021, Dorji Karma Nidup Pema Latsho [7] et al. conducted a poll. The questionnaire was distributed to all RUB students and staff, including about 9,070 undergraduate students and 463 staff members. The 64 survey items were asked, and both descriptive open-ended and quantitative closed-ended responses were gathered using a Likert scale. The information obtained from the Likert scales was analyzed using the replies' mean scores. Benefits of study is 64-item questionnaire collected both qualitative and quantitative responses via Likert scales, with mean scores utilized for data analysis and limitations of this study may have been limited because we decided to analyze mean scores from ordinal data, which can be difficult.

Sarah Ahtesham [8] et al. have conducted research that is primarily descriptive and foundational in nature (2020). It includes both a quantitative assessment and a qualitative understanding of the concept. This study uses IBM SPSS Statistics to examine data from questionnaires given to construction students at UiTM Sarawak's FSPU.The findings demonstrate how important a role emotions and preparedness have in determining pupils' habits and way of life.

In 2020, Laura Ibrayeva, Daniel Hernández-Torrano, Natalya Lim, Jason Sparks [9], et al. carried out a thorough examination of bibliometric indicators of studies on college students' mental health and wellbeing that have been done during the previous 45 years. This study aims to illustrate the corpus of research on university students' mental health and well-being by using metadata collected from 5,561 research papers listed in the Web of Science databases between 1975 and 2020.

Seifollah Fazlollahi, Maryam Mohammad Taheri, Gholamreza Veissi, and others [10] (2019). The goal of the current study was practical in terms of field and implementation techniques, and it employed a descriptive methodology to gather data. The quantitative and descriptive portions examined the general features of the respondents using tables and graphs of abundant distribution. The mean and standard deviation of each research variable as well as each questionnaire item were then displayed. In the inferential portion, each study topic was examined using relevant statistical tests. The data was analyzed using structural equation modeling (SEM).

A total of 535 respondents were selected from the entire population using simple random sampling in a study conducted by Siti Solehah Tena, Abu Khari A'ain, Khairunesa Isa, Azmi Abdul Latiff, Lutfan Jaes, Abdul halim abdul Rahman, and Zahrul Akmal Damin, siti solehah Tena [11], et al. (2019). The questionnaire sought to identify and quantify the factors influencing employee happiness at University A by focusing on three areas: organizational, social, and individual. A Likert scale with four possible responses was used to assess a statement based on the score distribution of agree or disagree.

Egemen Kemal Alganet [12] et al. (2019) addressed how teachers' well-being within an organization affects organizational outcomes in respect to Two indicators of instructors' well-being are their work-life balance and the company's level of satisfaction. Following the collection of quantitative data, qualitative data were acquired utilizing a sequential explanation strategy. In the quantitative component, the relationship between work-life quality and distributed leadership was found to be mediated by organizational happiness.

Mevawala Jency [13] et al., 2019 distributed a survey to all participants in a sample that was considered typical of the study's target demographic. The survey consisted of one or more closed questions with a limited number of viable answers. This combination of the other alternatives and the collection of answers to the question constitutes a primary happiness measuring scale. Approximately 1,000 happiness (also known as WDH) scales are included in the Measures of Happiness collection. They are often called things in the literature on happiness research.

The well-being of about 40,000 people per year was investigated for 21 countries using datasets acquired between 2006 and 2012 from the European Social Survey (ESS), Felicia A. Huppert, Sandra Matz, Garcia-Garzon, Aine Maguire, and others [14] (2018).We refer to the survey elements collectively as multidimensional psychological well-being, or MPWB. Benefits of using European Social Survey data (2006 and 2012), they evaluate the total well-being of 21 countries with around 40,000 individuals annually, referring to survey items as the multidimensional psychological well-being scale and limitations of this study is they must give a clear explanation of how this relates to policy in a particular context.

Lin BA, He-Zi WANG [15], et al. (2017) conducted an analysis of the happiness index, which is exactly calculated by applying the analytic hierarchy procedure. The weight value of the happiness index is determined by the three-layer structure and two-level index. Five subjects are covered in this essay: family, future employment, mental and physical health, and both. While on campus, students can have fun and provide suggestions for future training initiatives and college student mental health services. Benefits is this study is assessing factors influencing mental health and university programs, offering recommendations, while employing the Analytic Hierarchy Process for precise measurement of college students' happiness and limitations is surveys use closed questions with defined response options to gauge happiness.

Effect of quality of school life on sesnse of happiness is discussed by Riza gokuler, Ugur Gurgan and Nuray Tastan [16] et al., 2015. The research investigates student satisfaction and happiness in school, using a mixed-methods approach involving surveys and qualitative analysis. It finds moderate levels of satisfaction and happiness among students but notes limitations such as a small sample size and potential biases in self-reported data. The study emphasizes the importance of fostering positive school environments for student well-being, highlighting the need for further research to address its limitations.

The model presented by Beverley Morganet [17] et al. (2015) allows for the multidimensional snapshot of Wales's youth and child welfare (ages 0 to 25) as well as important vital information. In addition to highlighting evidence gaps by referencing newly published studies and statistics, it represents the efforts undertaken to collect data on a range of social variables in Wales. With relation to the wellbeing of children and youth in Wales, the 2015 Monitor aimed to draw attention to noteworthy trends and provide a concise and comprehensible overview of pertinent statistics. Benefits in this is publication's goal is to present a multifaceted picture of Wales's child well-being by utilizing a range of well-being signs along with extra statistical and research resources and limitations of study is 2015 Monitor aimed to provide a concise overview of children's well-being in Wales, highlighting trends, comparing with other countries, and identifying knowledge gaps.

A research was carried out by Vigneswaran Applasamy, Reynato Andal Gamboa, Mushtak Al-Atabi, Satesh Namasivayam [18] in 2013. The research introduces the Gross Institutional Happiness Index (GIHI) to measure happiness in academic environments, inspired by the Gross National Happiness Index. Methodology involves developing indicators, utilizing survey forms, and computing GIHI based on sufficiency achieved in nine core domains. However, potential limitations include subjective responses and lack of discussion on long-term sustainability.

Philippa Nwamu Unoma et al. (2013) [19] In order to collect and evaluate data, this research employed a quantitative survey methodology. The dataset that was analyzed was constructed from the information gathered. There would be two phases to the data analysis process: the first would involve a thorough investigation of the identified factor. Separation is the second step, which comprises analyzing the data after organizing it into recognizable clusters. The dataset would be subjected to a number of machine learning techniques in order to produce the mathematical model of students' satisfaction. Each machine-learning technique's accuracy would be assessed, and the most accurate approach would be selected to build the model. Benefits of this study is used quantitative surveys, employing factor analysis and data clustering, and applied machine learning techniques to construct a model for student happiness, prioritizing accuracy.

Carmelo Vazquez, Gonzalo Hervas [20], et al. conducted a study in 2012 on The Pemberton Happiness Index (PHI), a measure for measuring people's degree of happiness in seven different languages. Your well-being score is calculated using the PHI that you provide by answering the survey questions. The Cramer V test is used to evaluate it. The total of your answers to these questions represents your overall well-being score. Furthermore, the receiver-operating characteristic (ROC) curve methodology was used to generate a cut-off value of PHI to identify a happy individual. This score represents your level of happiness and life satisfaction. Benefits is the Pemberton Happiness Index measures happiness levels in seven languages, evaluated using the Cramer V test limitations is critical to understand whether modifications to one's material, psychological, or personal circumstances have a major impact on one's PHI scores.

III PROPOSED SYSTEM

The fig 3.1 represents system architecture diagram. A web application has been developed for this study in order to gauge student's overall school satisfaction and well-being. A web application's front end is made with HTML, CSS, and JavaScript. Django is used to develop the backend, and SQLite is used to store the data. It's an easy-to-use web application for both students and educators. A web-based survey application was developed by taking into account the factors influencing student's satisfaction in the classroom. Physiological needs are also addressed by including characteristics like the amount of nutrient-dense food consumed and the amount of time spent exercising. Using Google Forms, datasets were directly gathered from the students. To find the score distribution, the data were evaluated using a Likert scale with five possible responses. A higher score denotes greater student satisfaction. According to the results of the previous study, there isn't presently a proven way to improve their well-being after measuring their happiness index. As a result, in this study, we have devised a mechanism whereby, following assessment, a recommendation report is sent to the subject's email address if their happiness index score is less than a certain threshold.

An ensemble model comprising of XG boost, Naïve Bayes, and KNN is utilized for training and happiness prediction. A 92% accuracy rate is attained with the ensemble model as opposed to a single predictive model. By utilizing this comprehensive online tool, educational institutions can efficiently gauge, track, and enhance their student's happiness and general well-being, with an emphasis on mental health. Students benefit from this system's ultimate contribution to the development of a positive and encouraging educational environment.

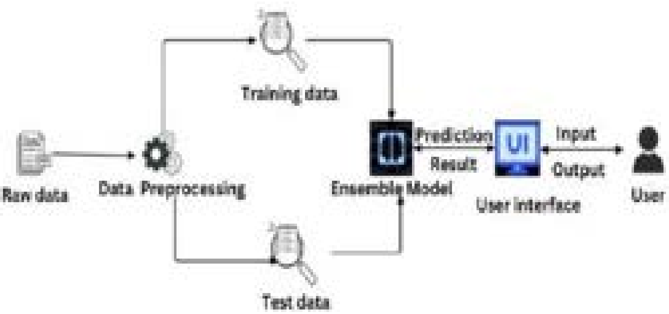


Fig 3.1 System architecture diagram

IV IMPLEMENTATION AND RESULTS

In this study data are collected through google forms which contains a series of 20 questions related to student's mental health. Around 400 data have been collected and it is preprocessed to remove duplicate and null values. In the preprocessing step the answer collected will be in the form of 'Strongly agree', 'agree', 'neutral', 'disagree', 'strongly disagree' Since machine cannot understand this language they will be converted into numerical values. Effective cleaning and preparing the data to handle missing values, outliers, and scaling numerical features will increase accuracy. Additionally, it is accomplished by accurately labeling the categorical variables (label encoding). Choosing pertinent characteristics that have a strong correlation with the desired variable will also increase the accuracy.

The next step is splitting the data for training and testing and fitting it into a model. Three machine learning algorithms (KNN, Naïve bayes and XG Boost) are predicts the output based on the training dataset. The selection of KNN, Naïve Bayes, and XG Boost algorithms is based on their ability to handle a variety of data types (numerical, text, and categorical) and their use in multi label multi class classification. In previous studies they have used SVM (Support vector Machine) algorithm which is suitable only for numerical data so the proposed method includes algorithm which can handle categorical and numerical data. Classification tasks are the primary applications for these models. The performance can be improved by optimizing the hyperparameters such as tuning the learning rate, tree depth. It can also be improved by choosing the optimized value for k.

KNN is a simple yet effective algorithm for classification tasks. It can handle non-linear relationships between input variables and the target variable. In this research KNN can be used to identify similar cases (i.e., children with similar characteristics) and infer their happiness scores based on their neighbors' scores. In KNN prediction process it preprocess the new data point in the similar way as training data and calculates the distance between new data point and all training data points. Select the k Nearest neighbour of the data points based on the calculated distance. For classification tasks, assign the most common class label among the k nearest neighbors to the new data point. For regression tasks, predict the average (or weighted average) of the target values of the k nearest neighbors.

Naïve Bayes is a probabilistic algorithm used for classification tasks, especially with categorical input variables. It assumes independence among predictors, which can simplify computation and make it computationally efficient, particularly with a large number of input variables. In this study it can provide insights into the probability of a child belonging to a particular happiness class given their characteristics, facilitating interpretability. In naïve bayes prediction process is done by preprocessing the new data point after that it calculates the posterior probability for each class using bayes theorem and likelihood estimates from the training data and then multiply the prior probability of each class by the corresponding conditional probability of the new data point given that class.

XG Boost (Extreme Gradient Boosting) is a powerful ensemble learning algorithm known for its speed and performance. It can handle complex relationships and interactions between variables, making it suitable for datasets with diverse characteristics and it also improves the accuracy. In XG Boost prediction is done by preprocessing the new data point and converting them into numerical array or data frame and then use trained

XG boost model to predict the target variable. For regression task interpret the predicted numerical value. After the evaluation a recommendation report is sent to their email address if, after forecasting the outcome, it is less than the threshold value.

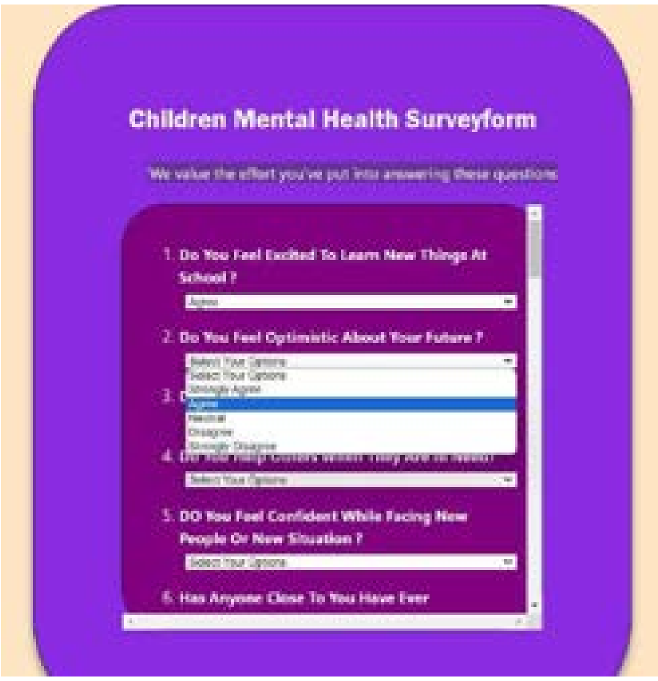


Fig 4.1 Questionnaire page

The questionnaire page, which users access after logging in, is seen in Figure 4.1. There are around twenty-five questions covering academic, physiological, social interaction, and psychological topics. It is composed of a five-point Likert scale. Greater happiness is indicated by a higher score. The user must click the submit button after answering all the question.

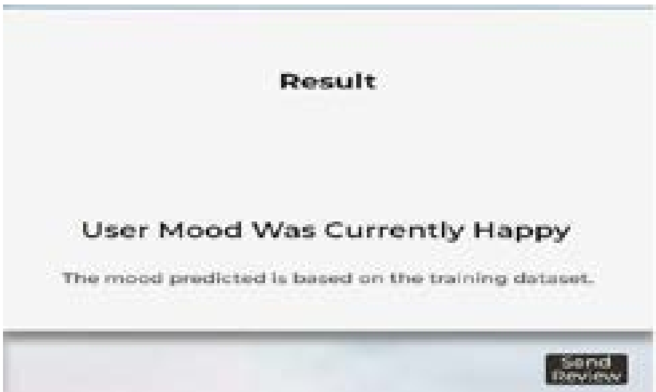


Fig 4.2 Result page

The user's mental health will be predicted based on the options they have selected. The figure 4.2 represents the result page. Based on the data trained using ensemble model (KNN, Naïve bayes, XG boost) it will predict the output. After predicting their output if the value falls below a threshold value, then suggestion report is sent to their Email ID.

V CONCLUSION

Ultimately, the goal of this research is to identify and improve student's pleasure and well-being in learning environments. One encouraging move in the direction of accomplishing this aim is the creation of an extensive Happiness Index and an easy-to-use online self-assessment tool. The score recognizes the full aspect of a student's experience by incorporating characteristics like mental wellness, which goes beyond the typical focus on academic accomplishment. An ensemble model predicts the happiness index more accurately and robustly than a single model, as demonstrated by the results of the previous study. To increase the predictive capacity of the model, additional characteristics can be added or the current features can be changed. Despite having a higher accuracy than a single predictive model, an ensemble model can be computationally demanding and may need additional resources for training. In addition, it takes longer to train than a single model. In the future, the model can also be enhanced by adding a strong feedback system to collect opinions on how well the Happiness Index works from educators and students. In order to give students who have been identified as having lower happiness scores targeted support, it can also be made better by merging it with counselling services. This could entail working together to create individualized intervention plans with mental health specialists. As a result, our web application supports the development of a supportive and nurturing environment in schools where students' emotional and psychological needs are met.

VI REFERENCES

[1] Anjum Sibia, Ruchi Shukla ,Sushmita Chakraborty et al. 2022 : "Mental Health and Well- being of School Students"

[2] Kumar, P., Swaminathan, B., Karthikeyan, U. et al. (2022). Multi-Perspective Reasoning Using Adaptive Learning.

[3] Yingying Jiang,,Chan Lu, Jing Chen, Yufeng Miao, Yuguo Li, and Qihong Deng et al. 2022 : Happiness in University Students: Personal, Familial, and Social Factors:A Cross-Sectional Questionnaire Survey

[4] Fatemeh Akhlaghi Rastehkenari , Baharak Shirzad Kebria, Abbas Khorshidi , Lotfollah Abbasi Sarvak et al.2022:Presenting a Model of Happiness for Students of Primary Schools of Education and Training

[5] Rafael Zapata Lamana,Cristian Sanhueza Campos et al. 2021 : Anxiety, Low Self-Esteem and a Low Happiness Index Are Associated with Poor School Performance in Chilean Adolescents: A Cross-Sectional Analysis

[6]

[7] JMatthew Schuelkaa mollie Braznell Mathew Leavesley, Sangay Dorji Khandu Dorji karma nidup pema Latsho et al., Happiness Wellbeing and Mental Health in Bhutanese Higher Education: Exploring Student and Staff Experiences and Perceptions within a Framework of Gross National Happiness, 2021.

[8] Ahtesham, Sarah. "Analysing happiness index as a measure along with its parameters and strategies for improving India's rank in World Happiness Report." ICTACT Journal on Management Studies 1 (2020): 1170-1173.

[9] Hernández-Torrano, Daniel, Laura Ibrayeva, Jason Sparks, Natalya Lim, Alessandra Clementi, Ainur Almukhambetova, Yerden Nurtayev, and Ainur Muratkyzy. "Mental health and well-being of university students: A bibliometric mapping of the literature." Frontiers in psychology 11 (2020): 540000.

[10] Mohammad Taheri, Maryam, Gholamreza Veissi, and Seifollah Fazlollahi. "Providing a Model for Creating Happiness in High Schools." Iranian journal of educational sociology 2, no. 4 (2019): 115-126.

[11] Isa, Khairunesa, Lutfan Jaes, Zahrul Akmal Damin, Azmi Abdul Latiff, A. H. Rahman, N. Maslan, Abu Khari A'ain, and Siti Solehah Tenah. "Determining indicators of happiness index among university staff." International Journal of Engineering and Advanced Technology 8, no. 5 (2019): 726-730.

[12] Algan, Egemen Kemal, and Azize Ummanel. "Toward sustainable schools: a mixed methods approach to investigating distributed leadership, organizational happiness, and quality of work life in preschools." Sustainability 11, no. 19 (2019): 5489.

[13] Mevawala Jency et al., Happiness Index– The Footsteps towards Sustainable Development, 2019.

[14] Ruggeri, Kai, Eduardo Garcia-Garzon, Áine Maguire, Sandra Matz, and Felicia A. Huppert. "Well-being is more than happiness and life satisfaction: a multidimensional analysis of 21 countries." Health and quality of life outcomes 18 (2020): 1-16.

[15] Anand, K., V. S. Yassmitha, and G. Deepti. "An Ensemble based Approach for Prediction of Happiness Index for School Children." In 2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA), pp. 1103-1111. IEEE, 2023.

[16] Gökler, Riza, Ugur Gürgan, and Nuray Tastan. "The effect of quality of school life on sense of happiness: A study on University students." Educational research and reviews 10, no. 7 (2015): 967.

[17] Morgan, Garth, Candice Butler, Reno French, Tamara Creamer, Lisa Hillan, Eva Ruggiero, Jennifer Parsons et al. "New ways for our families: Designing an Aboriginal and Torres Strait Islander cultural practice framework and system responses to address the impacts of domestic and family violence on children and young people." (2022): 1-52.

[18] Applasamy, Vigneswaran, Reynato Andal Gamboa, Mushtak Al-Atabi, and Satesh Namasivayam. "Measuring happiness in academic environment: A case study of the School of Engineering at Taylor's University (Malaysia)." Procedia-Social and Behavioral Sciences 123 (2014): 106-112.

[19] Mirza, Mohammed Shuhood. "Learning the Student's Happiness Model." (2013).

[20] Vazquez, Carmelo, and Gonzalo Hervas. "Addressing current challenges in cross-cultural measurement of well-being: The Pemberton Happiness Index." In Well-being and cultures: Perspectives from positive psychology, pp. 31-49. Dordrecht: Springer Netherlands, 2012.

Leveraging Facial Analytics for Enhanced Crime Prevention - Integrating Video Surveillance and FaceNet Algorithm

M .Archana[1]
Assistant Professor
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana, India
manneearchana@gmail.com

M. Shivani[2]
Student
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana, India
shivaninetha21@gmail.com

J. Varun[3]
Student
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana, India
varunjuluru26@gmail.com

K. Vivek chandra[4]
Student
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana, India
vivekchandra017@gmail.com

K. Keshavardhan[5]
Student
Computer Science & Engineering
Vignan Institute of Technology and Science
Hyderabad, Telangana, India
20891a0528@gmail.com

Abstract—

Facial analysis in crime prevention aims to leverage technologies like FaceNet and MTCNN algorithms for various objectives, including registering criminals into databases, matching photos to identify suspects, and enhancing video surveillance for real-time monitoring and detection of criminal activity. This research investigates the efficacy of integrating FaceNet algorithm within existing video surveillance infrastructure. These techniques enable law enforcement agencies to effectively utilize facial recognition technology in identifying and apprehending individuals involved in criminal activities, thereby enhancing public safety and crime prevention efforts.

Keywords—Criminal identification, Convolutional Neural Networks, machine learning, face recognition, Closed Circuit Television

1. Introduction

Facial recognition technology has emerged as a powerful tool in crime prevention strategies worldwide. By harnessing the capabilities of artificial intelligence and advanced algorithms, facial recognition systems can analyze and identify individuals from digital images or video frames based on their unique facial features. One of the primary advantages of facial recognition in crime prevention is its ability to enhance surveillance efforts. Law enforcement agencies can deploy facial recognition systems in public spaces, transportation hubs, and other critical areas to monitor for known criminals, suspects, or persons of interest in real-time. This proactive approach allows authorities to identify and apprehend individuals involved in criminal activities swiftly. Moreover, facial recognition technology can aid in the detection of unauthorized access or suspicious behavior in restricted areas. By integrating facial recognition systems with access control systems, organizations can strengthen security measures and prevent unauthorized entry into sensitive locations such as government facilities, corporate offices, or military installations. In addition to its role in surveillance and access control, facial recognition can assist in the investigation and resolution of criminal cases. Law enforcement agencies can use facial recognition software to match surveillance footage or crime scene images with existing databases of known offenders, thereby narrowing down potential suspects and expediting the investigative process. Furthermore, facial recognition technology can serve as a deterrent to criminal behavior.

2. LITERATURESURVEY

The face recognition applications have been evolving since 1960[1], which lead to invention of RAND tablet approach for coordinating features on the face. It is a device which emits electromagnetic pulses to input the coordinates on a grid. The device system used to record several facial coordinates manually like eyes, hairline, mouth and nose. Bledsoe [2] noted that there were many factors inhibiting a computer's ability to accurately recognize a single individual in two different photographs. Photographs of the same person might capture a person's face at vastly different angles, at different ages, with varying facial expressions, and under different lighting conditions. This may confound the system algorithm due to lack of advancement in algorithm and technology. In reference [2] various face recognition features were included like color of hair, chin, skin tone etc....

By using Correlations between pairs of features, Euclidian distances and Feature Reduction in facial lines the algorithm is deployed. This experiment was performed with three population sizes: 64, 128, and 255 faces the results for 128 and 255 are similar but 64 is relatively unsuccessful. Although these results are fairly good, all the experiments indicate that the binary selection method used was too rigid; but the achieved final result is on an average, about 50 percent of correct identification. In 1980 [3], researchers have introduced the statistical approach for face recognition, by using eigenvalues and eigenvectors to reduce dimensionality. It's the main idea behind Principal Component Analysis. The eigenface approach to face recognition was motivated by information theory, leading to the idea of basing face recognition on a small set of image features that best approximate the set of known face images, without requiring that they correspond to our intuitive notions of facial parts and features. Although it is not an elegant solution to the general object recognition problem. When an image is sufficiently close to face space (i.e. it is face-like) but is not classified as one of the familiar faces. The [4] uses dimensionality reduction technique for creating a criminal identification system termed as "FRCI". This study successfully achieved its objectives i.e matching a face with available database accurately and applying principal component analysis for finding distinguishable features from many images to get the similarity for the target image. It is a better alternative for criminal identification instead of using thumb print identification but it does not show the details of the criminal and it follows Principal component Analysis (PCA) which require data set with data normalization and if the number of Principal Components is not carefully chosen, it may miss certain information in contrast to the real list of characteristics.[5] uses Harr features as explained in [6].

The approach was used to construct a face detection system which is approximately 15 times faster than any previous approach. A harr like include considers adjoining rectangular locales in a location window at a particular area, rundowns the force of the pixels in each locale, and calculates the distinction between these wholes. But the main features are required to train with the Haar wavelet and it requires predefined dataset. In [7] Adriana Kovashka, Margaret Martonosi proposed a framework which employments 18 highlights counting RGB which is used in [8].This approach uses Viola Jones algorithm which has High true detection rate as well as high false detection rate and Training time is very slow .The algorithm encounter problems in positions range like side view and for low lighting because their Haar-like features do not map very well to varying positions.[9] explains the utilization of AdaBoost and ANN together to make a cross breed show "ABANN". Numerous deep learning models have been connected particularly for confront acknowledgment such as Retinal Connected Neural Network, Rotational Invariant Neural Network, Back Propagation Neural Network, Fast Neural Network, etc. The fundamental center of our venture will be based on [10] FaceNet is utilized to identify contrasts in the faces displayed from the database. A Deep Convolutional Neural Network (DCNN) approach is utilized, wherein each layer is mapped to 128 bytes and the assignment of acknowledgment, discovery, and clustering is carried out. An accuracy rate of over 95% has been attained for two datasets.

3. METHODOLOGY

The UI developed consists of three main domains:

a. Register Criminal

b. Photo Match

c. Video Surveillance

Register Criminal:

This domain facilitates the input of criminal data, encompassing various attributes including the individual's personal information such as Name, Parent's name, Gender, Religion, Blood Group, Distinctive Body Marks, Nationality, Convicted Crime, and also incorporates a visual representation through a facial image.

Photo Match:

Facial recognition's photo matching entails comparing a given facial image with a database of known individuals to uncover potential matches. This approach is widely utilized in law enforcement and security settings to aid in identifying individuals implicated in criminal activities.

Video Surveillance:

Video surveillance integrated with facial recognition technology provides law enforcement agencies and security organizations with a proactive means of identifying and responding to criminal activity in real- time, thereby enhancing public safety and security.

4. FACE RECOGNITION OVERVIEW

A. Face detection:

Face detection is a critical component of criminal identification systems, utilized in law enforcement and security applications worldwide. It involves the automatic detection of human faces within images or video frames. This technology plays a pivotal role in identifying suspects, enhancing surveillance, and maintaining public safety. At its core, face detection algorithms analyze digital images to locate and extract facial features. These algorithms utilize various techniques, including pattern recognition, machine learning, and neural networks, to differentiate faces from other objects or backgrounds. In criminal identification systems, face detection serves as the initial step in the process of matching captured faces against a database of known criminals or suspects. Once a face is detected, it can be further analyzed for recognition, where facial features are compared with existing records to determine potential matches. The effectiveness of face detection in criminal identification depends on various factors, including image quality, lighting conditions, facial expressions, and occlusions. Advanced algorithms have been developed to address challenges such as pose variation, partial occlusion, and disguise, enhancing the accuracy and reliability of the identification process.

B. Feature extraction:

Feature extraction in face detection for criminal identification is a pivotal step that involves capturing and representing key characteristics of facial images to enable accurate recognition and matching processes. Unlike face detection, which focuses on locating faces within images or video frames, feature extraction delves deeper into analyzing the unique attributes of each face for identification purposes. Various methods are employed in feature extraction, each aimed at capturing distinctive facial traits while minimizing noise and irrelevant information. One commonly used technique is Principal Component Analysis (PCA), which transforms the high-dimensional pixel values of facial images into a lower-dimensional space while preserving the most significant variations among faces. Additionally, Histogram of Oriented Gradients (HOG) is widely utilized for attribute extraction in face detection tasks. HOG computes the distribution of gradient orientations within localized regions of the facial image, capturing the underlying structure and shape of facial features. CNN is most used powerful tool for feature extraction in recent years, leveraging deep learning techniques to automatically learn hierarchical representations of facial features. CNN- based approaches can extract features at multiple levels of abstraction, enabling superior performance in complex scenarios and diverse facial conditions. Once features are extracted from facial images, they are typically represented as vectors or descriptors that encode the distinctive characteristics of each face. These feature representations are then used for comparison and matching against a database of known faces, enabling the identification of suspects or individuals of interest. Feature extraction in face detection for criminal identification plays an important role in increasing the accuracy, robustness, and efficiency of the overall identification process. By capturing and representing key facial traits in a compact and discriminative manner, feature extraction enables effective matching and recognition, facilitating the apprehension of criminals and the maintenance of public safety.

C. Verification or Identification:

Depending on the application, the system may either verify whether the suspect's face matches a specific individual in the database (verification) or attempt to identify the suspect by searching the entire database for potential matches (identification).

D. Human Review:

While automated algorithms can assist in the matching process, human review is often necessary to confirm the accuracy of the results and make final judgments regarding the identification of suspects.

Achieving high accuracy in facial analytics for crime prevention projects involves several key steps. Firstly, it's crucial to use high-quality datasets that are diverse, representative, and balanced. This ensures that the model learns from a wide range of facial features and characteristics, improving its ability to generalize to real-world scenarios. Secondly, selecting the right features is essential. This involves identifying facial landmarks, textures, and patterns that are relevant for crime detection and prevention. Using advanced feature extraction techniques, such as deep learning-based methods, can help capture intricate details in facial images. Thirdly, the choice of model architecture plays a significant role. Convolutional Neural Networks (CNNs) are commonly used for facial analytics due to their ability to learn hierarchical representations of facial features. Fine-tuning pre-trained CNN models, such as VGG, ResNet, or Efficient Net, on your dataset can significantly boost accuracy. Lastly, continuous evaluation and refinement are crucial. Regularly testing the model on new data and fine-tuning it based on performance feedback can help improve its accuracy over time. Additionally, incorporating feedback from domain experts and stakeholders can help refine the model's performance for specific crime prevention applications.

System Architecture:

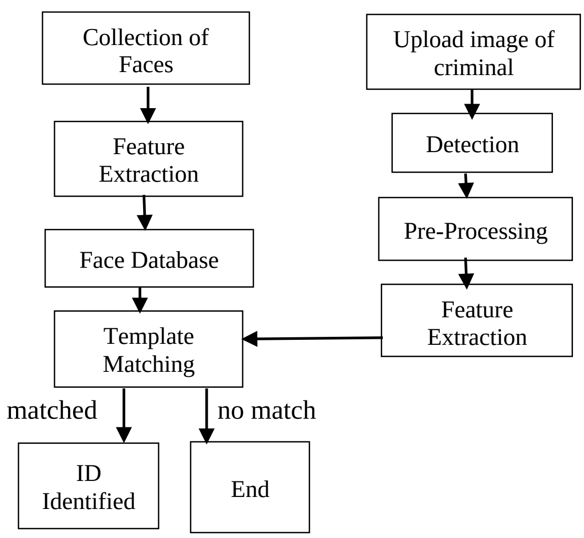


Fig. 1 Proposed System Architecture

To perform face recognition, a series of steps are typically followed as shown in fig. 1. Firstly, a dataset of facial images is collected, encompassing a variety of poses, lighting conditions, and facial expressions to enhance the model's accuracy and robustness. Next, the images are preprocessed to standardize their size, orientation, and lighting. Following this, a face detection algorithm is employed to locate and extract faces from the images, utilizing methods like Haar cascades, Histogram of Oriented Gradients (HOG), or deep learning-based approaches such as Single Shot Multibox Detector (SSD) or Faster R-CNN. Subsequently, features are extracted from the detected faces, often utilizing Convolutional Neural Networks (CNNs) to capture distinctive facial characteristics. These features are then encoded into a numerical representation for comparison, reducing the dimensionality of the features while retaining essential information. Finally, the encoded features are matched against a database of known faces to identify or verify individuals. The process can be refined through continuous model improvement, ethical considerations, and compliance with privacy regulations.

5. FACE RECOGNITION APPROACH

A. MTCNN:

MTCNN (Multi-Task Cascaded Convolutional Systems) calculation is one such innovation that has revolutionized the field of surveillance and object detection. Created in 2016, the MTCNN calculation presents a cascading arrangement of neural systems to identify, adjust, and extract facial features from digital images with increased accuracy and speed. The MTCNN calculation scale the picture to distinctive sizes as the input of each network layer. It consists of three convolutional neural systems to achieve the multitask detection. MTCNN calculation can totally include three assignments at the same time: face location, border relapse and point positioning.

Stage 1 (Proposal Network - PNet): This stage generates candidate bounding boxes (proposals) for potential faces in the input image. It uses a fully convolutional network to efficiently propose regions that might contain faces.

Stage 2 (Refinement Network - RNet): The proposals generated in the first stage are refined in this stage. The network filters out false positives and improves the accuracy of the bounding box predictions.

Stage 3 (Output Network - ONet): Finally, this stage performs further refinement and filtering on the proposals generated by the previous stages. It outputs the final bounding boxes along with facial landmarks.

B. FaceNet:

FaceNet is a deep learning model developed by researchers at Google that is designed for performing face recognition tasks. It utilizes CNN architecture to map features of a person's face into a space of high-dimension called as embedding. This embedding is such that faces with alike features are mapped closer together while diverse faces are farther apart, making it easier to compare and recognize faces. One of the key innovations of FaceNet is its use of triplet loss. This helps in learning discriminative features that are robust to variations in lighting, pose, and facial expressions. The objective of this loss function is to create little squared separates between two similar faces. While, Large squared remove between two pictures having diverse personalities.

Triple loss-- The inserting is spoken to by f (x) ∈ R^d. This inserts a picture x into a d-dimensional Euclidean space. Moreover, we constrain this implanting to live on the d-dimensional hypersphere, i.e. kf(x)k2=1.This misfortune is propelled within the settingof nearest-neighbour classification. Here we need to guarantee that an picture x_i^a(stay) of a particular individual is closer to all other pictures x_i^p (positive) of the same person than it is to any picturex_i^n (negative) of any other person.

$$\left\|f(x_i^a) - f\left(x_i^p\right)\right\|_2^2 + \alpha < \left\|f(x_i^a) - f\left(x_i^p\right)\right\|_2^2$$

-equation (1)

$$\forall \left(f(x_i^a), f\left(x_i^p\right), f\left(x_i^n\right)\right) \in \tau$$

-equation (2)

$$L = \sum_i^N [[\ ||f(x_i^a) - f\left(x_i^p\right)||_2^2 - ||f(x_i^a) - f\left(x_i^p\right)||_2^2 + \alpha]$$

-equation (3)

These triplets would not contribute training and result, as they would still pass through network it is crucial to choose triplets, that were active and can contribute to improve model. The following section describes about triplets.

Triplet Selection–For faster learning, we have to triplets that violate the equation above. This means for given x_i^a we ought to triplets such that;
$||f(x_i^a) - f(x_i^P)||_2^2$is maximum and;
$||f(x_i^a) - f(x_i^n)||_2^2$ is minimum.

Generating triplets based on the entire training set is computationally expensive. Two alternative methods for triplet generation are:

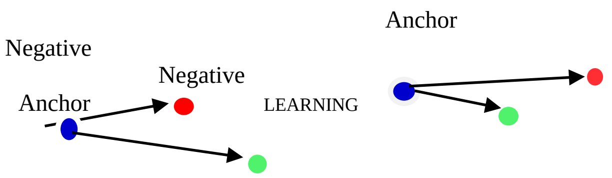
Incremental Triplet Generation:

Instead of generating triplets on the entire training set in each step, triplets are generated incrementally based on previous checkpoints. This approach involves calculating the minimum and maximum on a set of the data, making the process more computationally efficient.

Checkpoint-based Triplet Generation:

Triplets are generated at specific checkpoints during the training process. At these checkpoints, the model's parameters are saved, and triplets are generated based on these saved parameters. This allows for triplet generation without the need to consider the entire training set, reduce the computational cost.

Reference Paper Number	**Algorithm Used**	**Accuracy**	**Accuracy of Proposed Model (Facenet & MTCNN)**
[4]	The Algorithm is designed using Principal component analysis (PCA) of Eigenfaces	70.633%	89.74%
[5]	“Haar Classifier and OpenCV” for face identification and Recognition	74.45%	89.74%
[10]	Convolution Neural Network (CNN) and Dense net.	80.017%	89.74%
[11]	The algorithm is developed using “KLT (Kanade Lucas Tomasi)” for face motion Tracking and Haar Classifier. “LK algorithm” by Carlo Tomasi to optimize tracking.	84.778%	89.74%



7. RESULTS & DISCUSSION

Figure2.User Interface

Fig 2 clearly depicts about the home interface, which contains the attributes required for criminal identification.

Figure3.Registration Form

Fig3. gives us an insight of attributes that are required in order to Register a New Criminal.

Figure 4. Photo Match

Fig4. gives an information about photo match of Two persons named Farah Khan and Shashi Mehta retrieved with an accuracy of 94.707%.

Figure 5. Video Surveillance

Fig 5. retrieves the Data of a Person named Vivek with an Accuracy of 93% through **Video Surveillance.** The overall Testing Accuracy of developed model is 95% and Training Accuracy is 94.33%.

Attributes encountered in the Developed system	**When the match is found**	**When the match is found intermediately**	**When the match is not found**
Photo match	95.435 %	89.245%	No match found
Video Surveillance	89.171 %	83.155%	No match found

Achieving high reliability in facial analytics for crime prevention projects requires careful attention to several key aspects. Firstly, it's essential to establish robust data collection procedures. This involves using high-quality cameras and ensuring consistent lighting conditions to capture clear and consistent facial images. Standardizing the data collection process helps minimize variability and ensures the reliability of the collected data. Secondly, data preprocessing plays a crucial role in improving reliability. This includes removing noise, artifacts, and irrelevant information from the images. Preprocessing techniques such as image normalization, denoising, and face alignment can help improve the quality of the input data and enhance the reliability of the model's predictions. Thirdly, the choice of model architecture and training approach is critical. Using deep learning models, such as convolutional neural networks (CNNs), that are well-suited for image analysis tasks can help improve reliability. Additionally, training the model on a diverse and representative dataset can help enhance its ability to generalize to new and unseen data, improving overall reliability. Finally, continuous monitoring and evaluation of the model's performance are essential for ensuring reliability. Regularly testing the model on new data and monitoring its performance metrics can help identify and address any issues that may arise, ensuring that the model remains reliable over time. Incorporating feedback from users and domain experts can also help improve the reliability of the model by refining its performance for specific crime prevention applications.

8. CONCLUSION

In conclusion, the advancements in face detection technology have significantly impacted the field of criminal identification, revolutionizing the way law enforcement agencies approach investigations and security measures. By harnessing the power of machine learning algorithms and computer vision techniques, facial recognition systems have become invaluable tools for swiftly and accurately identifying suspects, enhancing public safety, and preventing crime. The ability to match faces captured in surveillance footage or photographs against databases of known individuals has expedited the identification process, leading to the apprehension of criminals and the resolution of cases that may have otherwise remained unsolved. Moreover, facial recognition technology has proven instrumental in combating various forms of criminal activity, including theft, fraud, terrorism, and human trafficking, by providing law enforcement agencies with timely and actionable intelligence. By prioritizing transparency, accountability, and responsible innovation, we can ensure that facial recognition systems continue to serve as powerful tools in the fight against crime while upholding fundamental rights and values in our society.

Mobile Applications: Development of mobile applications for law enforcement personnel to access facial recognition databases and receive real-time alerts on their devices.

Artificial Intelligence Integration: Use of AI for more sophisticated facial recognition, enabling better accuracy in

identifying individuals, even in challenging conditions such as low light or obscured faces.

Cross-Domain Integration: Explore integration with other domains, such as social media or online databases, to gather additional information that can aid in investigations.

9. REFERENCES

[1]Bledsoe, “Manual measurements”,1960.

[2]A.J. Goldstein, L.D. Harmon and A.B. Lesk, “Identification of human faces,”in proceedings of the IEEE, vol 59, pp. 748-760,May1971.

[3]M. Turk and A. Pentland, “Eigenfaces for Recognition,” in Journal of cognitive neuroscience, vol 3, pp. 71-86, Jan 1991.

[4]N. A. Abdullah, Md. J. Saidi, N. H. A. Rahman, C.C. Wen, and I. R.A. Hamid, “Face recognition for criminal Identification: An implementation of principal component analysis for face recognition,” AIP Conference Proceedings1891:1, Oct 2017.

[5]P. Chhoriya, “Automated criminal identification system using face detection and recognition”, in International Research Journal of Engineering and Technology (IRJET) ,

vol 6, pp. 910-914, Oct 2019.

[6]P. Viola and M. Jones, “Rapid object detection using a boosted cascade of simple features,” in Computer Vision and Pattern Recognition, vo1 1, pp. 511-519, Feb 2001.

[7]A. Kovashka and M. Martonosi, “Feature-based face recognition for identification of criminals vs. identification for cashless purchase”.

[8]T. H. Le, “Applying artificial neural networks for face recognition”, in Hindawi Publishing Corporation Advances in Artificial Neural Systems, vol 2011, pp. 1-16, 2011

[9]F. Schroff, D. Kalenichenko, J. Philbin, “FaceNet: A unified embedding for face recognition and clustering”, in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815-823, 2015.

[10] Identification of criminal and non-criminal face using deep learning and image processing. Rajkumar; Shyam Sundar A; Ruban Haris B; Ragul K; Cloudin S 2021 International Conference on system, computation, automation an networking (ICSCAN)

[11] Face detection criminals through CCTV cameras Worawut Yimyam; Thidarat Pinthong; Narumol Chumayung; Mahasak ketchum 2018 14th International conference on signal-image technology&internet-basedsystems (SITIS)

[12] P. M. Corcoran and C. Iancu, “Automatic face recognition system forhidden Markov model techniques,” New Approaches toCharacterization and Recognition of Faces, pp. 3-28, 2011.

[13]Ajay Gurav, Alireza Chevelwalla, Sachin Desai, Prof. Sumitra Sadhukhan,“Criminal Face Recognition System”, vol 2 pp,2015

[14] P. Kakkar and V. Sharma, “Criminal identification system using facedetection and recognition,”in International Journal of AdvancedResearch in Computer and Communication Engineering, vol 7, pp.238-243, March 2018,

[15] P. Apoorva, H.C. Impana, S.L. Siri., M.R. Varshitha and B. Ramesh,“Automated criminal identification by face recognition using opencomputer vision classifiers,” in 2019 3rd InternationalConferenceonComputing Methodologies and Communication (ICCMC), pp. 775-778,2019

[16] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, “Joint face detection andalignment using multi-task cascaded convolutional networks”.

[17] Comparative analysis for classifiers for criminal identification system using face recognition, sanika thanmayi ratnaparkhi; Pooja singh; aamani tandasi; Nidhi sindhwani; 2021 9th international conference on reliability, Infocom technologies and optimization (ICRITO)

[18] Swati jagtap; Nilkanth B; Chopade; Sanjay Thungar 2022 6th international conference on computing, communication, control and automation of an investigation of face recognition system for criminal identification in video surveillance video.

.

The Roadside Garbage Collection Robot

Vikas J. Nandeshwar
DESH
Vishwakarma Institute of Technology
Pune, Maharashtra, India
vikas.nandeshwar1@vit.edu

Vishwatej M. Shende
DESH
Vishwakarma Institute of Technology
Pune, Maharashtra, India
vishwatej.shende23@vit.edu

Arya B. Shinde
DESH
Vishwakarma Institute of Technology
Pune, Maharashtra, India
arya.shinde23@vit.edu

Rishi R. Shendre
DESH
Vishwakarma Institute of Technology
Pune, Maharashtra, India
rishi.shendre23@vit.edu

Lokesh P. Shengolkar
DESH
Vishwakarma Institute of Technology
Pune, Maharashtra, India
lokesh.shengolkar23@vit.edu

Darshan N. Shinde
DESH
Vishwakarma Institute of Technology
Pune, Maharashtra, India
darshan.shinde23@vit.edu

***Abstract*— The world is currently grappling with a significant garbage crisis, which is a result of the prevailing issues of quick economic expansion, inadequate urban planning, overpopulation, corruption, and political inefficiency. It is becoming increasingly evident that there is a need to develop a robust waste management strategy to monitor the accumulation of garbage on roadsides, footpaths, schools, and colleges. Therefore, this research study explains the design of a garbage collection robot that uses IoT technology such as a Bluetooth application. This research study aims to design and make a Roadside Garbage Collection Robot (Automatic Garbage Collector). This robot model serves as an automatic garbage collector, which can effectively and efficiently tackle the problem of garbage accumulation.**

Keywords— Arduino, Garbage, Ultrasonic sensor, IoT

I. Introduction

Maintaining cleanliness in the environment is crucial for various reasons. There are several issues related to the cleaning of the environment. One of the major issues is the improper disposal of waste, which leads to pollution of land and water bodies. There are several issues related to the cleaning of the environment. One of the biggest environmental concerns today is the amount of litter that is found on roadsides and public spaces. Despite various efforts by organizations to reduce littering, it seems to be a growing problem [1]. However, there is a potential solution in the form of garbage-collecting robots. These machines can autonomously navigate through public spaces and collect trash, which could significantly reduce the amount of litter on the streets. With the advancement of technology, it's exciting to see how these robots could positively impact the environment.

Until now, there has been a lot of interest and development in service robots for cleaning swimming pools, houses, walls, and domestic stairs. However, cleaning robots for roadsides have not received as much attention. This paper presents the development of a prototype for a roadside cleaning robot. Nowadays, the way of cleaning up roadside waste is by hiring manual laborers. However, industries are shifting towards automation and it should also be applied efficiently towards waste management. Manual labor is the traditional way for employment yet there are problems associated with cleaning up garbage by human laborers.

In the present era, as countries continue to develop at an unprecedented pace, a significant amount of unwanted waste is being generated, including electronics, plastics, and various biodegradable products. Despite considerable research in this field, finding an efficient solution has proven to be elusive. However, with the advent of garbage collection robots, we now have the opportunity to collect and dispose of waste properly and manage it effectively. Unfortunately, the robots developed thus far have not been precise or accurate, and most of them are controlled manually and lack robustness. Our project robot, on the other hand, will function with precision efficiently collecting and disposing of waste. Its implementation will benefit numerous individuals and significantly reduce human labor.

II. Literature Review

Various existing models are capable of collecting of garbage. Let's discuss a few models based on the survey:

The robot by K Kamal et al. [1] can move with an average speed of 0.5m/s but the robot developed is yet to be improved to operate automatically and control from long distances.

The robot by Kavya.C et al.[2] is flexible to be used in various environments. The robot uses IoT technology yet the ways are not mentioned in clarity

The Robot by Ayub Mulani et al.[3] has used Raspberry Pi for garbage detection and has also attached a camera to the model but the robot is yet to achieve accuracy in detection.

The robot by N.Varuneshreddy et al.[4] is made to operate on sand by the beachside. Images from the IP camera are sent to the computer and are processed in the Bluetooth module and garbage is detected.

The research by Tawanda Mushiri et al.[5] emphasizes on the designing of the robot in a way that it will be able to efficiently collect garbage.

The robot by Saravana kannan G et al.[6] detects garbage using image processing. Various features in images like color, and size are considered in MATLAB.

The Robot by Hesham Alsahafi et al.[7] can detect metallic and nonmetallic garbage in specific areas. The size of the robot is small and it can work only in specific areas.

The Robot by Osiany Nurlansa et al.[8] is designed of the steps: requirements analysis, design, manufacturing, and testing. It is capable of automatically picking up garbage on the water.

The robot by Bharathi V et al.[9] collect garbage in a highly effective manner, without any need for human intervention. It uses an IR sensor as well as ultrasonic sensor for detection.

The robotic arms designed by Abhiraj Bhalerao et al.[10] is a fully automatic system but one can operate it manually also. The robotic arm has a wide range of applications, including material handling, object sorting, welding, grinding, and spray painting.

In the research paper by Thota Arun et al.[11], the robot detects objects using an IR sensor and a moisture sensor is attached to the bin to detect the level of garbage collected. A web page is also created to display the trash level present in the bin.

The robot made by H.G.T. Milinda and B.G.D.A. Madhusanka[12] is used for cleaning floors in domestic households. An image acquisition camera is used for image processing. Mud detection is also implemented based on color and area.

The robot was made by H.M.R.G. Herath et al. [13] is a beach cleaning robot that has a high-performance electrical motor to drive the robot and uses the IoT platform to indicate the level of the garbage tray remotely.

III. METHODOLOGY

A. Components: ***Arduino MEGA*** **2560 Rev3**, **Servo and DC motors, IR Sensor, Ultrasonic sensors, HC05 Bluetooth Module, wheels, wooden plank, wires, cluster wheels, PVC pipes, etc**.

The main frame of the robot: The wooden board is of dimensions 27x27x5 cm. At the backside of the robot, a dustbin is attached which is of dimensions 12x20cm.

The reason for choosing ultrasonic sensors was that the sensor can measure the distance between itself and the object. The ultrasonic sensor also allows it to put up a threshold distance by which it can measure the distance up to that point. In this way, when the sensors are attached to the robot and when set up with a threshold distance the ultrasonic sensor will detect the distance between the robot and the garbage.

After that, the IR sensor will emit the IR light and the IR receiver shall check for the reflected light. If the light gets reflected by hitting the garbage in front of it, then the IR receiver receives this light, and the object is then detected. The robotic arms are programmed in such a way that when the IR receiver receives the signal and the ultrasonic sensors detect the distance between them, the arms will move and pick up the garbage.

Fig 1. below is the circuit diagram of the robotic arms used in our project. Arduino mega 2560 Rev3 and hc05 Bluetooth module is needed for controlling and connecting mobile phones to the robot respectively. Arduino Mega 2560 Rev3 has 54 digital input/output pins and it comes with more memory space and, a bigger size. To power the servos 5V voltage is required, and it comes from an external power source as the Arduino is unable to handle a large amount of current that draws from all of them. The power source ought to handle 2A of current. Once the connections are done the Arduino Coding is started.

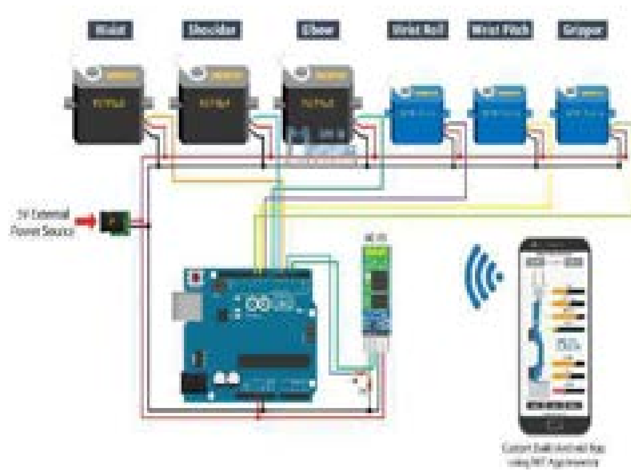


Fig. 1. Circuit diagram of robotic arms

The robotic arms are flexible as they consist of 5 degrees of freedom. The parts of the robotic arms are 3D printed and are assembled with the help of screws.

Flowchart for robotic arm movement: In Fig 2. the robotic arms will begin in their initial position. Once the ultrasonic detects the waste object, the wheels will stop at the threshold distance. As the IR sensor detects the garbage the arms will move to pick it up. If the garbage is not detected the arms will remain stationary.

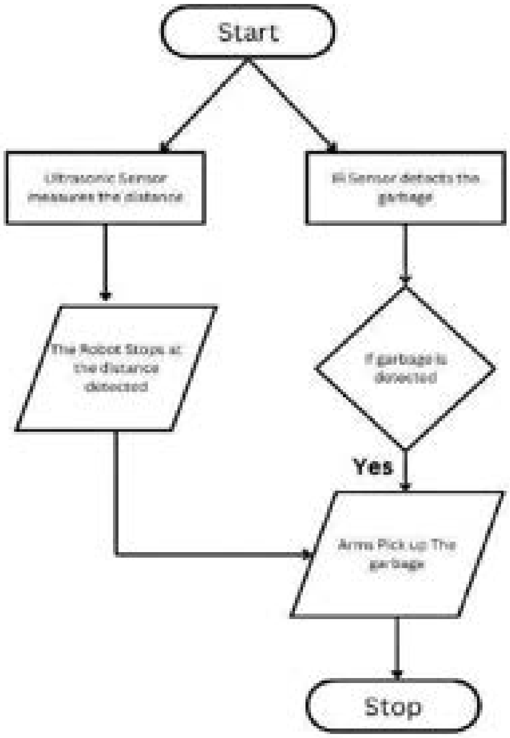


Fig. 2. Flowchart for arms movement

B. *Block Diagram of the robot*: The Fig 3. block diagram depicts that instructions to the robot will be transmitted through the Arduino Mega 2560 Rev3 microcontroller. The block diagram of a robot encompasses diverse components, including sensors, microcontrollers, power supply, motors, and communication modules, which operate concertedly to ensure the robot's optimal functionality. The block diagram not only facilitates the comprehension of the information and signal flow between the different components but also illustrates the robot's interaction with its surroundings and reception of commands from external sources. The Bluetooth module HC05 is connected to the smartphone through a small app developed (Netra App).

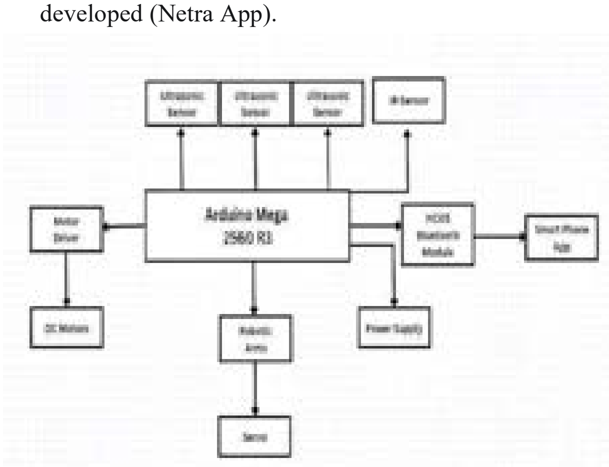


Fig. 3. Block diagram of the robot

C. *Design of the robot:* The utilization of Tinkercad, a free 3D modeling program renowned for its user-friendly interface, enabled us to design our robotic arm with utmost precision. By employing this software, it can determine the optimal measurements that would facilitate the efficient construction of the robot. In addition, its simulation capabilities provided us with substantial insights regarding the working principle of the robotic arm, thereby imparting a scientific understanding of the system. There were many options for designing the robot but this design as proposed in [5] has better efficiency in picking and placing garbage.

Fig. 4. Proposed model of the robot

IV. WORKING PRINCIPLE

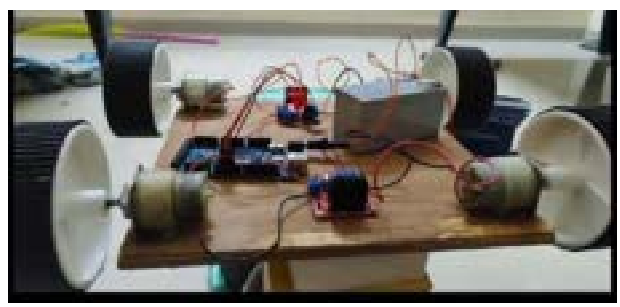


Fig. 5. Testing of DC Motors with HC05 Bluetooth module

The ultrasonic ranging module is a versatile component that allows for non-contact measurements ranging from 2cm to 400cm, with an impressive accuracy of up to 2mm. The module is composed of ultrasonic transmitters, a receiver, and a control circuit. To use it effectively, you need to follow these steps: 1. Use an IO trigger to send a high-level signal that lasts for at least 10 microseconds. 2. The module transmits eight 40 kHz signals and waits for one pulse signal to bounce back. 3. If any signal is detected at a high level, the duration of the high output IO signal is the time it took for the ultrasonic signal to travel to the object and back. You can use this time to calculate the test distance using the formula: distance = (time x velocity of sound in air (340m/s) / 2)[1].

Fig. 6. Real-time graph of Ultrasonic sensor readings

Fig 6. depicts the graph on which the X axis represents the time in seconds and the Y axis represents the distance between the sensor and garbage.

An IR sensor is a device that can sense the presence of an object in its surroundings by emitting infrared light. The sensor works by using an IR LED to emit light and an IR photodiode to detect the light. When the light is reflected by an object and falls on the photodiode, its resistance and output voltage change in proportion to the amount of light received. There are two types of IR sensors: Active and Passive. Active infrared sensors consist of an infrared source and an infrared detector, while passive infrared sensors only have a detector. Hence an active IR sensor is used to detect the object. The energy emitted by the infrared source is reflected by an object and falls on the infrared detector, which helps detect the presence of the object.

Fig. 7. IR sensor readings when object is detected

Fig 7. depicts the readings of IR sensor when the object is detected. Along with that the distance is also displayed in the output.

The robot has been set up with three ultrasonic sensors to increase the accuracy of object detection. Ultrasonic sensors have a specific range for detection which is limited. So, to have maximum detection 3 ultrasonic sensors are used. In this way, if the left or right-side sensor detects the garbage the sensor will be able to detect garbage more accurately.
The Ultrasonic Sensors will detect objects/waste which will come in the path of the robot then the Robotic Arm will get the exact coordinates of the object/waste and then the object will be picked up by the gripper part attached to the robotic arm.
The robotic arms are designed using Solidworks 3D software which we purchased [9]. The arm of the robot has 5 degrees of freedom. Only the shoulder, wrist roll, and gripper parts of the robot were used. The Arduino coding was done in such a way that the angles of the shoulder, wrist, and gripper were already set. A threshold distance is set in the ultrasonic sensor. When the sensor detects the distance, and the IR sensor detects the garbage the robot stops and the elbow first rolls down, then the shoulder to bring the arm to the ground level then the gripper grabs the garbage, and the elbow and shoulder go back to their original position. Then the gripper drops the collected garbage into the dustbin attached at the back. For the first 3 axes namely, the base, the shoulder, and the elbow, we have used the MG996R servos, and for the other 2 axes, the wrist roll and wrist pitch, as well as the gripper we have used the smaller SG90 micro servos. After picking up the object/waste, all servos will produce a combined motion to drop the object/waste in the dustbin. Then, the Arm will come to its original position. All these tasks will be performed repeatedly.

HC-05 is a Bluetooth module used to perform wireless communication. It uses serial communication to communicate with the devices and it communicates using a serial port (USART) with the microcontroller. Its range is <100m depending on the transmitter, receiver, and geographic conditions. The robot is controlled using the Netra app that is connected to the robot using the HC-05 Bluetooth module.

Fig. 8. Remote Control of the Robot using Netra App

V. RESULTS & DISCUSSION

A. Review Stage:

Fig. 9 Actual setup of the robot

The design of the proposed robot has been improved for increased efficiency by utilizing three wheels instead of four. The back of the robot is supported by two large wheels that balance the weight of the dustbin, while two cluster wheels at the front enable improved maneuverability. With this design, our robot can move effectively and efficiently to perform its tasks. At first, the team faced a lot of issues while checking the angles of the robotic arms because the servos and Arduino could not sustain the weight of to object, and more power was needed to perform smooth functions. Connecting the Arduino to the Bluetooth module and operating through the Netra app was also a challenge. Despite the challenges, the robot can be controlled through the Netra app.

B. Final Stage

Assess the arm's ability to carry and manipulate objects of different weights. This involves testing the torque and payload capacity of each servo motor Evaluate the accuracy of the robotic arm's movements. This can be done by instructing the arm to move to specific positions and measuring how closely it reaches the target points. Test the repeatability of the robotic arm by instructing it to perform a series of tasks repeatedly. Check if it consistently achieves the same results.

TABLE I. PERCENTAGE ERROR ANALYSIS OF THE ANGLES

Parts of Arm	**Arm**	**Elbow**	**Grip**
Angle 1 estimated	95	55	75
Actual Angle 1	105	45	60
Error 1 in %	-9.52	22.2	25
Angle 2 estimated	18	130	9
Actual Angle 2	15	155	7
Error 2 in %	20	-16.12	28.57

Table 1. represents the percentage error in the angles of the robotic arms. The actual angle values were used to do the coding of robotic arms with the help of servo motors. The average of the two errors =(12.56+10.81)/2=11.68.The precision of the arms calculated is very low. To increase the precision of the arms, a powerful servo motor and a good power supply are required.

The coding for the arms was done and the robot was able to be controlled with the help of connecting the hc05 Bluetooth module to the mobile app. The robot can pick up and place the garbage in the bucket. It's also possible to control the robot with the voice command feature present in the Netra app. The unique feature of the robot is that it is robust and designed so that there will be less wear and tear of the robot. Other previously made robots are delicate and smaller in size. The robot is perfectly balanced and can be operated in all types of environments. The entire robot is cost-effective and also is big in size. The limitations faced while implementing the robot were the lack of availability of hardware materials, less power supply by the battery, as well as finances required.

VI. CONCLUSION

It's important to acknowledge the physically demanding nature of garbage pickup work and the occupational hazards that come with it. That's why it's great to see the development of a robot that can eliminate the health risks associated with manual waste collection. This solution is not only beneficial for environmental maintenance, but it also reduces the need for human intervention in garbage maintenance and pollution monitoring. It's worth noting that the mechanical design of the robotic arm needs to be researched further to improve the robot's functionality. However, there is still potential for this robot to operate automatically and with good accuracy. There is still a lot of scope for future research in the field of robotics and automation. In the future, garbage collection and management is going to be a serious issue due to the lack of human laborers and the increasing population which eventually will lead to an increase in garbage. There will be a need for garbage collection robots to reduce human labor and manage waste efficiently.

VII. FUTURE SCOPE

The objective was to design a robot capable of gathering objects from specific areas such as roadside, colleges, schools, and others. Our vision is to expand the usage of this robot across various fields. To achieve this, we aim to enhance the automation of the robot and improve its object detection accuracy. The accuracy of the robot can be increased by using better servo motors and reducing the stiffness of the robotic arms. Raspberry Pi can enable the robot to produce efficient results and improve the overall performance of the robot. We also plan to add a lid or covering for the dustbin attached as open waste can lead to the breeding of mosquitoes and other diseases. The team is also working towards adding more features such as sweeping and implementing more efficient cleaning methods beyond just picking. In this way, the reliability of the robot is increased. It can be made to handle harsh environments and is robust.

ACKNOWLEDGMENT

The authors would like to thank Vishwakarma Institute of Technology for their support in completing this project and thank you to Prof. (DR.) Chandrashekhar Mahajan (Head of Department, DESH), College Workshop, and The Robotics Forum (TRF) club for guiding us through difficulties.

REFERENCES

[1] Design of Garbage Collection Robot using wireless technology
[2] Garbage Collection Robot using IOT B.E.Student, Dept. of Electrical and Electronics Engineering, Vivekanandha College of Engineering for Women[Autonomous], Namakkal, India.
[3] Wireless Communication Based Garbage Collection robot on the beach. PROF. AYUB MULANI, KARISHMA PIRJADE, HARSHADA WALHE, SWAPNALI KANADE Department of Electronics and Telecommunication Engineering Shalaka Foundation Keystone School of Engineering, Pune.
[4] Garbage Collection Robot on the beach using wireless communications Student, Department of ECE, SRM Institute of Science and Technology, Ramapuram, India
[5] Wireless Communication Based Garbage Collection.
[6] Automatic Garbage separation robot using image processing technique. Saravana kannan G, Sasi kumar S, Ragavan R, Balakrishnan M
[7] Design and Implementation of metallic waste Collection robot. Hesham Alsahafi, Majed Almaleky. Advisor: Prof. Tarek Sobh Robotics, Intelligent Sensing & Control (RISC) Laboratory, School of Engineering, University of Bridgeport.
[8] AGATOR (Automatic Garbage Collector) as Automatic
[9] Garbage Collector Robot Model OsianyNurlansa, Dewi Anisa Istiqomah, and MahendraAstuSangghaPawitra, Member, IACSIT.
[10]]Automatic BinBot - Garbage Collecting System using IoT Bharathi V Asst prof, Gayathri K, Jayashree S, Kiruthika D, Maheswari P Department of Electronics and Communication Engineering.
[11] Garbage Collection Robot and Monitoring System Using Wireless Communication Thota Arun1, Divyavani K1, Boini Varun1, Peter Moshe1, K.Vasanth1, P.Marish Kumar2, Srinivasan K S3.
[12] Multi-Functional Floor Cleaning Robot for Domestic Environment H.G.T. Milinda and B.G.D.A. Madhusanka.
[13] Design and Implement of Beach Cleaning Robot by H.M.R.G. Herath, G.C.B. Weerasinghe, M.W.S.B. Jayasooriya, K.K.N. Dharmarathna.

Efficient Topology Control and Depth Adjustment Technique for Connectivity in UWSN

A.CH.SUDHIR
Associate Professor, Dept of EECE,
GST,GITAM (Deemed to be University),
Visakhapatnam, India
camanapu@gitam.edu

Garuda Sai Pattabhi Ram
Dept of EECE,
GST,GITAM (Deemed to be University),
Visakhapatnam, India
sairamgaruda69@gmail.com

Adduri Srinivas Raghava
Dept of EECE,
GST,GITAM (Deemed to be University),
Visakhapatnam, India
raghava41529@gmail.com

Eswar Pavan Katuri
Dept of EECE,
GST,GITAM (Deemed to be University),
Visakhapatnam, India
eswarpavan56@gmail.com

***Abstract*— In Underwater wireless sensor networks (UWSNs), nodes are moving due to the external parameters like ocean currents and marine organisms etc. The topology control techniques proposed for traditional wireless sensor networks (WSN) were not applicable for UWSN because of its propagation delay, less bandwidth and multipath fading. An Efficient Topology Control and Depth Adjustment (ETDA) technique for UWSN is designed. The major role of ETDA technique is to reduce node disconnections and to improve the link quality of transmission. In this technique, a combined objective function (COF) is defined for topology control based on the energy level of nodes, distance between the nodes at different depth levels and the link expiration time. The topology control algorithm is applied by each sink node when a broken link or poor channel quality is detected at any part of the network. In this algorithm, the sink node performs reconnection to an intermediate sensor node with maximum COF value. Experimental results show the performance improvement of ETDA by maximizing the packet delivery ratio and minimizing the energy consumption and delay.**

Keywords*— Underwater wireless sensor networks (UWSN); Topology; Depth Adjustment; combined objective function (COF); Efficient Topology Control and Depth Adjustment (ETDA)

I. INTRODUCTION

UWSN consists of nodes which can perform data collection, processing and communication via acoustic signals. UWSNs are useful in applications like oceanographic data collection, pollution monitoring, disaster recovery and tactical surveillance. When UWSN are used in critical applications like military field or disaster recovery, all the sensor nodes should communicate with the sink node ensuring complete network connectivity. The best node deployment strategies for UWSN involve updating the node positions to ensure network coverage and reduce energy drain rate [1].

The main characteristics of UWSN are limited bandwidth capacity, high propagation delays, huge bit error rates and multipath fading effects [2]. To address these issues, researchers designed low speed acoustic signals for underwater communication which increases the broadcast delay and losses [3]. The topology control techniques proposed for traditional WSN were not applicable for UWSN because of its unique characteristics listed above. Hence there is a need for a topology control technique which ensures the connectivity and network lifetime of sparsely deployed sensor nodes [4]. Efficient data delivery from the sensor node to the sink is another major issue of this network. It should consider the ike depth of the node and speed of the acoustic signals [5]. Hence a stable and reliable optimized topology is required for successful data delivery in UWSN. In order to assure long connectivity among the nodes, the network topology control with efficient routing technique is necessary [7]. This work proposes an Efficient Topology Control and Depth Adjustment (ETDA) technique for UWSN.

The main objectives of this work include:

- To reduce network disconnections in UWSN
- To enhance the link quality of transmission
- To reduce the energy consumption

II. RELEATED WORKS

A topology control technique for UWSN proposed by Zhen Hong et al [1] consists of a game theory based technique for selecting the best nodes as cluster heads. Then, intra and inter cluster topology construction phases are performed. Simulation outcomes show that TCEB achieves better energy efficiency and throughput. A ReVOHPR protocol is proposed by Amir Chaaf et al [3] for stable cluster head selection. To increase the network lifetime, a dynamic sleep scheduling algorithm is presented. A virtual graph based routing algorithm is proposed for communication. To reduce the transmission latency, multiple AUVs are installed to collect the data. Lastly, the void holes are detected using relay based technique and repaired. Another cooperative game theory based scheme (DATUM) is presented by Sudip Misra et al [4] for dynamic topology control. They have established a route from the source sensor to the sink nodes for effective data transmission. After applying the Game theory model, they have determined a set of best paths for transmission. Energy Efficient Localization Algorithm (EELA) was proposed by Yali Yuan et al [8] for analyzing the communication between the anchor and sensor nodes. A game theory mechansim is used to design a topology control algorithm for the unknown nodes and anchor nodes by utilizing the transmission chances. An

Energy balance algorithm was proposed by Weiping Zhang et al [9] which analyzes the relationship between node' s remaining energy and data communication. It assures balancing of energy during data forwarding by the intermediate sensor nodes. Simulation outcomes shows that the algorithm extends the network lifetime and achieves higher throughput and delivery ratio. A joint power control and opportunistic routing protocol for UWSN was designed by RodolfoW. L. Coutinho et al [10]. In this protocol, routing decisions are made by considering the transmit power and the communication range of the nodes.

III. PROPOSED SOLUTION

Efficient Topology Control and Depth Adjustment (ETDA) technique for UWSN is proposed. In this technique, a combined objective function (COF) is defined for topology control based on the energy level of nodes, distance between the nodes at different depth levels and the link expiration time. The topology control algorithm is applied by each sink node when a broken link or poor channel quality is detected at any part of the network.

A. System Model

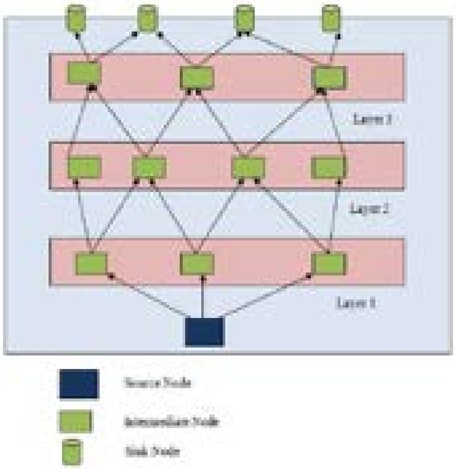


Fig. 1 System model of UWSN

The system model of UWSN is shown in Figure 1. It consists of multiple source nodes at sea level (level 1), multiple intermediate nodes at level 2 and multiple sink nodes at top level (level-3). It is assumed that each sensor node can communicate with at least another node at a higher level. Here, each sensor node $n_{i,l} \in N$ is at level $l \in L$, where N is the collection sensor nodes and L is the levels of nodes in UWSN. Moreover, it is assumed that each sensor node is connected to at least one sink node.

Let NL_s be the neighbor list of each sink $s \epsilon S$, (S is the set of sink nodes) given by

$$NL_s = \{n_{i}^{l} \mid D^l(i,s) \leq R_s\} \tag{1}$$

Where R_s is the communication range of sink s, $D^l(i,s)$ is the distance between the sink s and n_i at the layer l.

B. Determination of Combined Objective Function

A combined objective function (COF) for topology control is defined in terms of energy level (EL), distance between the nodes at adjacent layers ($d_{l,l+1}(i,j)$) and the link expiration time ($L_{i,j}$).

Energy Level (EL)

The EL of node n_i at level l indicates the node lifetime. Hence, each sink node selects the nodes with higher EL.

The total energy consumption (E_{tot}) of each sensor includes the energy consumption during transmission (E_{tx}) and the energy consumption during reception (E_{rx}), which can be calculated as:

$$E_{tot} = E_{ttx} + E_{rx} \tag{2}$$

E_{tx} and E_{rx} can be calculated as

$$E_{tx} = \sum_{i=1}^{h} P_{txi} * Z * T_{ni} * (1 + RTrate_i)$$
$$E_{rx} = \sum_{i=1}^{h-1} P_{rxi} * Z * T_{ni} * (1 + RTrate_i) \tag{3}$$

Where,

P_{tx} and P_{rx} denote the sending power and the receiving power of node i,

T_{ni} is the average time required to transmit one byte of data

$RTrate_i$ denotes the data retransmission rate of node i

Z denotes the size of the data

Then the energy level of each node n_i at layer l is given by

$$EL_l(n_i) = E_{ini} - E_{tot} \tag{4}$$

Where E_{ini} is the initial energy assigned to each node.

Link Expiration Time ($L_{i,j}$)

The link expiration time ($L_{i,j}$) denotes the time duration for which two intermediate nodes n_i and n_j (at adjacent levels), remain connected without disconnection. Hence the topology remains stable when the link expiration time is high.

$L_{i,j}$ can be computed follows:

$$L_{i,j} = \begin{cases} \frac{1}{v_{j,(l+1)}} D_{l,(l+1)}[\cot\alpha + \cot\beta] & if\, n_{j,(l+1)} isatA \\ \frac{1}{v_{j,(l+1)}} D_{l,(l+1)}[\cot\alpha - \cot\beta] & f\, n_{j,(l+1)} isat\, A' \end{cases} \tag{5}$$

Where $v_{j,(l+1)}$ is the mean speed of j at level (l+1),

$$\alpha = \sin^{-1}\left(\frac{D_{l,(l+1)}}{R_i}\right) \quad \text{and} \quad \beta = \sin^{-1}\left(\frac{D_{l,(l+1)}}{d_{i,j}}\right)$$

Distance ($d_{i,j}$)

If the distance $d_{l,l+1}(i,j)$ of n_i and n_j at levels l and (l+1) increases, the communication delay also increases.

$d_{l,l+1}(i,j)$ can be calculated as follows:

$$d_{l,l+1}(i,j)=\sqrt{(x_{nj,(l+1)}-x_{ni,l})^2+(y_{nj,(l+1)}-y_{ni,l})^2} \quad (6)$$

Where ($x_{ni,l}, y_{ni,l}$) is the coordinates of node $n_{i,l}$.

Hence the COF can be derived as follows

$$COF = w_1.\left(\frac{EL}{EL_{\max}}\right)+w_2.\left(\frac{L}{L_{\max}}\right)+w_3.\left(1-\frac{d}{d_{\max}}\right) \quad (7)$$

where EL_{max}, L_{max} and d_{max} are the threshold values of EL, $L_{i,j}$ and $d_{l,l+1}(i,j)$ respectively and w_1,w_2,w_3 are weighting constants ranging from 0 to 1.

The weighting constants are used to provide weightage to each metric used in the equation. During the experiments, the values are assigned as $w_1=0.4$, $w_2=0.3$ and $w_3=0.3$

Thus the sink nodes try to maximize the COF value meeting the below constraints.

$RE \geq RE_{max}$ and $L \geq L_{max}$ and $d \leq d_{max}$ (8)

C. Topology Control by Sink Node

The topology control algorithm is applied by each sink node when a broken link or poor channel quality is detected at any part of the network. In this algorithm, the sink node performs reconnection to an intermediate sensor node with maximum COF value. The topology control algorithm described below:

Topology Control Algorithm

Notations	**Definition**
sn	sink node
S	set of all sink nodes
T_i	Monitoring interval , i=1,2....
$COF_{i,l}$	COF of node $n_{i,l}$
MinCOF)	Minimum value of COF
Max(COF)	Maximum value of COF
TC_{req}	Topology control request
TC_{rep}	Topology control reply
ACK	Acknowledgement message
NLs	neighbor list of sink s
Next_hop_change	Next Hop Change message

1. For each monitoring interval T_i
2. For each $sn \epsilon S$,
3. Receives RE, $L_{i,j}$ and d(i,j) from each $n_{i,l}$
4. Derives the $COF_{i,l}$ using Eq.(7)
5. Compares the estimated COF with others
6. If $COF_{i,l}$ = Mini(COF) then
7. Indicates topology disconnection
8. sn sends a TC_{req} to $n_k \in NL_s$
9. n_k send ACK to sn as request confirmation
10. End if
11. If sn receives ACK, then
12. Waits for TC_{rep}.
13. Else
14. Retransmits TC_{req}
15. End if
16. n_k sends a its RE, $L_{i,j}$ and d(i,j) to S
17. sn derives the $COF_{k,l}$ of n_k
18. If $COF_{k,l}$ = Maxi(COF) then
19. sn updates its routing tree by nk with $n_{j,l-1}$
20. sn sends next_hop_change to $n_{j,l-1}$
21. Else
22. s again sends TC_{req} to another $n_{k+1} \in NL_s$
23. Repeat the steps from 11
24. End if
25. End For
26. End For

IV. EXPERIMENTAL RESULT

A. Experimental Settings

The proposed ETADA technique is implemented in NS2 AquaSim tool which an open source simulator developed for UWSN research.

The simulation settings and parameters are summarized in table 1.

Number of Nodes	25 to 125
Area size	1000 X 1000x10m
MAC Protocol	Underwater Mac
Simulation Time	100 seconds
Channel Model	Underwater Channel
Transmit Range	200m
Initial Energy	10000 Joules
Transmit power	2.0 watts
Receive Power	0.75 watts
Idle Power	0.008 watts
Data generation Interval	100 sec
Position Updating Period	0.3 sec

Table 1 Experimental settings

B. Results

The proposed ETADA technique is compared with LMST [9] approach according to the metrics latency, packet delivery ratio, average energy consumption and average forwarding ratio.

Nodes	**ETADA (sec)**	**LMST (sec)**
20	0.254	0.327
50	0.273	0.313
75	0.283	0.334
100	0.311	0.357
125	0.314	0.383

Table 2 Results for Topology control Latency

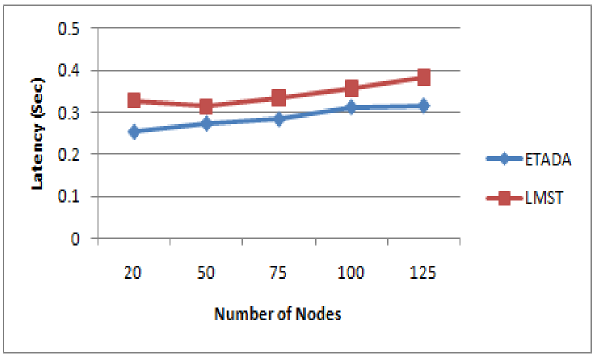


Fig2 Topology control Latency for Nodes

From the figure 2, it can be observed that the proposed algorithm ETADA achieves 16% lesser latency than LMST for different nodes scenario.

Nodes	ETADA	LMST
20	0.8889	0.8325
50	0.8972	0.8578
75	0.9384	0.8679
100	0.9574	0.8812
125	0.9582	0.9125

Table 3 Results of Packet Delivery Ratio

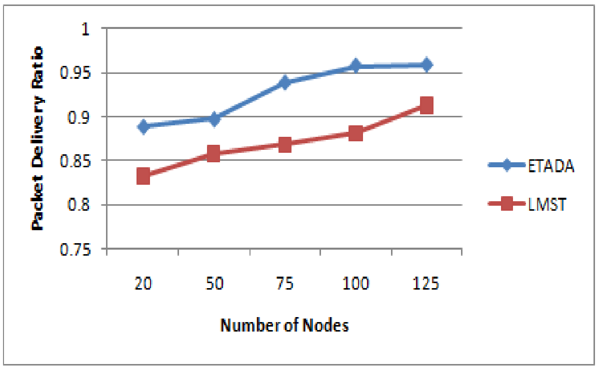


Fig 3 Packet Delivery Ratio for Nodes

From the figure 3, it can be seen that ETADA achieves 6% more packet delivery ratio than LMST for different nodes scenario.

Nodes	ETADA (Joules)	LMST (Joules)
20	675.22	793.92
50	734.11	855.64
75	787.45	890.95
100	831.63	900.66
125	859.48	949.73

Table 4 Results of Energy Consumption

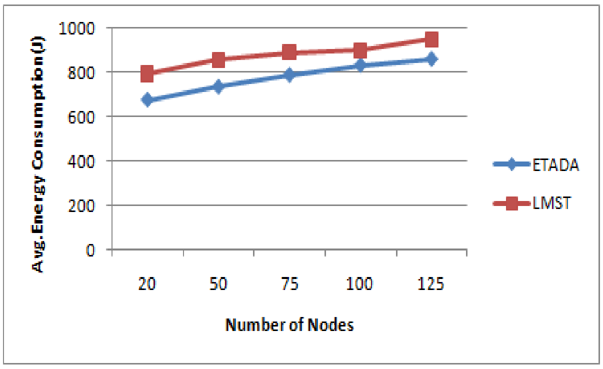


Fig 4 Energy Consumption for Nodes

From the figure 4, it can be observed that the proposed algorithm ETADA achieves 12% lesser energy consumption than LMST for different nodes scenario.

Nodes	ETADA	LMST
20	0.5771	0.3317
50	0.5484	0.3887
75	0.6364	0.4967
100	0.6843	0.5291
125	0.7643	0.5815

Table 5 Results of Forwarding Ratio

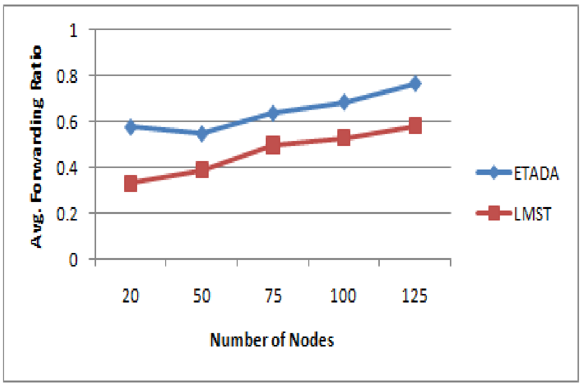


Fig 5 Forwarding Ratio for Nodes

From the figure 5, it can be seen that the proposed algorithm ETADA achieves 28% higher forwarding ratio than LMST for different nodes scenario.

.

CONCLUSION

In this paper, an Efficient Topology Control and Depth Adjustment (ETDA) technique for UWSN is proposed. In this technique, a combined objective function (COF) is defined for topology control based on the energy levels of nodes, distance between the nodes at different depth levels and the link expiration time. In topology control algorithm, the sink node performs reconnection to an intermediate sensor node with maximum COF value. The proposed ETADA technique is implemented in the NS2 and compared with LMST technique. Simulation results have proven that

ETDA has higher packet delivery ratio and forwarding ratio with minimized energy consumption and latency.

REFERENCES

[1] Zhen Hong , Xiaoman Pan, Ping Chen, Xianchuang Su, NingWang and Wenqi Lu, "A Topology Control with Energy Balance in UnderwaterWireless Sensor Networks for IoT-Based Application", Sensors, Vol-18, 2018.

[2] Hina Nasir, Nadeem Javaid, Shaharyar Mahmood,Umar Qasim, Zahoor Ali Khan and Furqan Ahmed, "Distributed Topology Control protocols for Underwater Sensor Networks", 19th International Conference on Network-Based Information Systems, 2016.

[3] Amir Chaaf,Mohammed Saleh Ali Muthanna , Ammar Muthanna,Soha Alhelaly,Ibrahim A. Elgendy ,Abdullah M. Iliyasu and Ahmed A. Abd El-Latif, "Energy-Efficient Relay-Based Void Hole Prevention and Repair in Clustered Multi-AUV Underwater Wireless Sensor Network", Security and Communication Networks,Volume 2021, Article ID 9969605, 20 pages, 2021.

[4] Sudip Misra, Anudipa Mondal and Ayan Mondal, "DATUM: Dynamic Topology Control for Underwater Wireless Multimedia Sensors Networks", IEEE, 2019.

[5] Mohammed Jouhari, Khalil Ibrahimi and Mohammed Benattou, "Topology Control Through Depth Adjustment and Transmission Power Control for UWSN Routing Protocols", IEEE, 2015.

[6] Ming He, Qiuli Chen, Fei Dai and Xiang Zheng, "Topological configuration and optimization in underwater acoustic sensor networks: A survey", International Journal of Distributed Sensor Networks, 2018, Vol. 14(8), 2018.

[7] Preetha Thulasiramann and KevinA.White, "Topology control of tactical wireless sensor networks using energy efficient zone routing", Elsevier,Digital CommunicationsandNetworks(2016) 2, 1–14, 2016.

[8] Yali Yuan, Chencheng Liang,Megumi Kaneko,Xu Chen and Dieter Hogrefe, "Topology Control for Energy-Efficient Localization in Mobile Underwater Sensor Networks using Stackelberg Game", arXiv:1805.12361v1, 2018.

[9] Weiping Zhang, Yudong He, Ming Wan, Mohit Kumar and Taorong Qiu, "Research on the energy balance algorithm of WSN based on topology control", EURASIP Journal on Wireless Communications and Networking, (2018) 2018:292, 2018.

[10] RodolfoW. L. Coutinho, Azzedine Boukerche, Luiz Vieira and Antonio A. F. Loureiro, "Topology Control and Opportunistic Routing for Underwater Acoustic Sensor Networks", https://www.darpa.mil/news-events/2017-12-06, 2018..

Harnessing Uncertainty: Integrating Fuzzy Logic into Machine Learning Algorithms

Rahib Imamguluyev
Head of the IT Department
Baku Business University
Baku, Azerbaijan
0000-0002-3998-7901
rahib.aydinoglu@gmail.com

Shaban Beshirov Hashim
Teacher of Mechanics and mathematics department
Western Caspian University
Baku, Azerbaijan
0000-0003-3820-1097
shaban.1954@mail.ru

Ilham Hajiyev
Lecturer in the IT department
Aydinoglu Academy
Baku, Azerbaijan
0009-0006-3129-8390
aydinoglu.akademiya@gmail.com

***Abstract*— As machine learning algorithms continue to advance, the incorporation of uncertainty modeling becomes pivotal for robust and adaptable systems. This article explores the fusion of fuzzy logic with machine learning methodologies, present-ing a comprehensive approach to harnessing uncertainty in data-driven decision-making processes. The integration of fuzzy logic provides a nuanced framework that accommodates imprecise and ambiguous information, enhancing the algorithms' capacity to handle real-world complexities. Through a detailed examination of the synergies between fuzzy logic and machine learning, this study con-tributes to the development of more resilient and versatile systems, demonstrating the efficacy of uncertainty-aware models in various applications. The findings underscore the potential for improved accuracy and interpretability in machine learning outcomes by embracing the inherent uncertainty in data.**

Keywords*— *Fuzzy Logic, Machine Learning Algorithms, Uncertainty Modeling, Decision-Making, Ambiguous Information, Resilient Systems

I. INTRODUCTION (*HEADING 1*)

In an era characterized by the rapid evolution of machine learning algorithms, the pursuit of robust and adaptable systems has taken center stage. As the technological landscape advances, the recognition of uncertainty as a critical factor in decision-making processes becomes increasingly evident. This article delves into the realm of uncertainty modeling and explores its integration with machine learning methodologies, presenting a groundbreaking fusion of fuzzy logic and data-driven approaches. The synthesis of these two domains offers a comprehensive framework for navi-gating the complexities of real-world scenarios, where imprecise and ambiguous in-formation often prevails [1].

At the core of this exploration is the fusion of fuzzy logic—a mathematical framework that deals with imprecision and vagueness—with machine learning. This amalgamation addresses a crucial aspect of contemporary computational challenges, providing a nuanced lens through which algorithms can interpret and respond to uncertain or incomplete data. Fuzzy logic's inherent ability to handle imprecise in-formation aligns seamlessly with the intricate demands of machine learning systems, enhancing their adaptability and resilience [2].

The article meticulously examines the synergies between fuzzy logic and machine learning, shedding light on the ways in which this integration contributes to the development of more resilient and versatile systems. By embracing uncertainty as an integral component, the proposed approach seeks to fortify algorithms against the intricacies of the real world, where information is often fraught with ambiguity. Through a detailed examination of uncertainty-aware models, the study not only underscores the potential for improved accuracy but also emphasizes the enhanced interpretability of machine learning outcomes [3].

This research, with its comprehensive approach, extends its impact across various applications, demonstrating the efficacy of uncertainty-aware models in diverse do-mains. Whether applied to healthcare, finance, or autonomous systems, the findings showcase the practical implications of incorporating fuzzy logic into machine learning methodologies. The article serves as a testament to the transformative power of embracing uncertainty, showcasing how it can be leveraged as a valuable asset rather than a hindrance in the quest for accurate and interpretable machine learning out-comes. As the technological landscape continues to evolve, the insights presented in this study pave the way for a new era in data-driven decision-making—one that acknowledges and harnesses the inherent uncertainty within the vast realms of in-formation.

II. TYPES OF MACHINE LEARNING ALGORITHMS

Machine learning (ML) algorithms are computational models designed to perform a task without being explicitly programmed. These algorithms enable computers to learn and make predictions or decisions based on data. There are various types of machine learning algorithms, broadly categorized into supervised learning, unsupervised learning, and reinforcement learning [4].

Supervised Learning: In supervised learning, the algorithm is trained on a labeled dataset, where each input is associated with a corresponding output. The goal is to learn a mapping function that can accurately predict the output for new, unseen inputs.

Common algorithms include linear regression, support vector machines, and neural networks. These algorithms are powerful for tasks like classification and regression [5].

Unsupervised Learning: Unsupervised learning involves training the algorithm on an unlabeled dataset, allowing it to discover patterns and relationships within the data on its own. Clustering and dimensionality reduction are common tasks in unsupervised learning. K-means clustering and principal component analysis (PCA) are examples of algorithms used in this context [6].

Reinforcement Learning: Reinforcement learning is concerned with training agents to make decisions in an environment to maximize a reward signal. Agents learn through trial and error, receiving feedback on their actions [7].

Popular reinforcement learning algorithms include Q-learning and deep reinforcement learning, where neural networks are employed to approximate complex policies.

Despite the success of traditional machine learning algorithms, there are situations where uncertainties, imprecise data, and fuzzy reasoning are prevalent. This is where fuzzy logic comes into play.

III. Justification for Fuzzy Logic

Fuzzy logic is a mathematical framework that deals with uncertainty and imprecision. Unlike classical binary logic, where a statement is either true or false, fuzzy logic allows for degrees of truth between 0 and 1 [8]. This makes it suitable for handling problems that involve vagueness and ambiguity, which are common in real-world scenarios. Here are some justifications for using fuzzy logic in machine learning:

Handling Uncertainty: In many real-world applications, data and decision-making are inherently uncertain. Fuzzy logic provides a way to model and represent uncertainty by allowing membership degrees to different classes or categories [9].

Linguistic Representation: Fuzzy logic enables the incorporation of linguistic terms and human-like reasoning into the decision-making process. This is particularly useful when dealing with qualitative and subjective information [9].

Rule-Based Systems: Fuzzy logic operates on a set of rules that express relationships between input and output variables. These rules are often more intuitive and easier to formulate than precise mathematical equations, making fuzzy logic suitable for rule-based systems [10].

Robustness to Noise and Incomplete Information: Fuzzy logic systems are generally more robust in the presence of noise and incomplete information. They can handle situations where exact numerical values are unavailable or where data may be im-precise [11].

Integration with Human Expertise: Fuzzy logic provides a framework for integrating human expertise into machine learning models. By allowing for fuzzy rules and linguistic variables, it facilitates collaboration between domain experts and machine learning practitioners [12].

While traditional machine learning algorithms excel in many applications, fuzzy logic plays a vital role in addressing the challenges posed by uncertainty, imprecision, and the need for human-like reasoning in certain real-world scenarios. Its ability to handle vagueness and incorporate linguistic terms makes fuzzy logic a valuable tool in the machine learning toolkit, especially in domains where clear-cut boundaries and precise data are hard to define [13].

IV. Problem Statement

In the rapidly evolving landscape of machine learning algorithms, the inherent uncertainty in real-world scenarios poses significant challenges to the development of robust and adaptable systems [20]. While traditional machine learning models excel in well-defined and precise data environments, they often struggle when faced with imprecise, ambiguous, and uncertain information. The problem at hand is the limited ca-pacity of existing algorithms to handle the complexities of decision-making processes in situations where data is incomplete, ambiguous, or subject to vagueness [14].

As identified in this article, the current state of machine learning faces a critical gap in addressing uncertainty, leading to suboptimal performance in various applications such as healthcare, finance, and autonomous systems. The reliance on binary logic and precise mathematical representations falls short when confronted with the nuanced demands of real-world data, which often involves degrees of truth, linguistic terms, and subjective information [15].

This study aims to tackle this problem by proposing a fusion of fuzzy logic with machine learning methodologies. Fuzzy logic, with its ability to handle uncertainty, imprecision, and linguistic representation, presents a promising solution to enhance the adaptability and resilience of machine learning algorithms. The problem addressed is not only limited to the accuracy of predictions but extends to the interpretability of machine learning outcomes, crucial for informed decision-making in diverse domains [16].

The current challenge lies in bridging the gap between traditional machine learning approaches and the need for more nuanced, uncertainty-aware models. This research seeks to address this gap by investigating the synergies between fuzzy logic and ma-chine learning, with the ultimate goal of developing systems that can navigate and thrive in the inherently uncertain and complex nature of real-world information [17]. The problem statement emphasizes the pressing need for a paradigm shift in data-driven decision-making, acknowledging and leveraging uncertainty as an integral compo-nent rather than a hindrance.

V. Methodology Proposed

Membership functions for the input and output variables in the fuzzy inference system (FIS) were determined using triangular and trapezoidal membership functions. Membership functions capture fuzzy relationships between input and output variables. Let's define membership functions for each variable:

Input Variables:

- Machine Learning Accuracy (MLA)
- Data Completeness (DC).

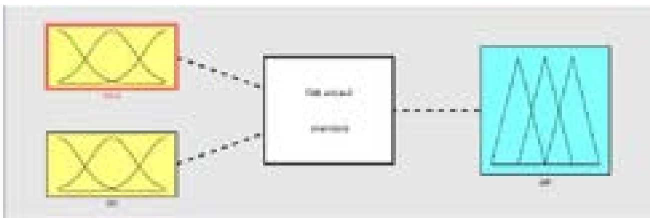


Fig.1. Graph of input variables.

Machine Learning Accuracy (MLA):

Low: Trapezoidal membership function [0 0 0.3 0.5]

Medium: Triangular membership function [0.3 0.5 0.7]

High: Trapezoidal membership function [0.5 0.7 1 1]

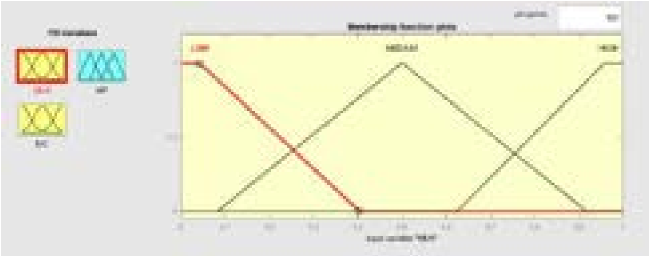


Fig.2. Fuzzy sets and membership functions for Machine Learning Accuracy.

Low Accuracy: Represented by a trapezoidal membership function with values ranging from 0 to 0.5. This level indicates that the model's predictions have a low degree of correspondence with the actual outcomes. Membership values in this range imply a weak level of confidence in the accuracy of the predictions.

Medium Accuracy: Characterized by a triangular membership function spanning values from 0.3 to 0.7. Models categorized under this level exhibit a moderate degree of alignment between predictions and actual results. Membership values in this range suggest a moderate level of confidence in the model's accuracy.

High Accuracy: Defined by a trapezoidal membership function ranging from 0.5 to 1.0. This level signifies a strong correspondence between model predictions and observed outcomes. Membership values in this range indicate a high level of confidence in the accuracy of the model's predictions.

Data Completeness (DC):

Low: Trapezoidal membership function [0 0 0.3 0.5]

Medium: Triangular membership function [0.3 0.5 0.7]

High: Trapezoidal membership function [0.5 0.7 1 1].

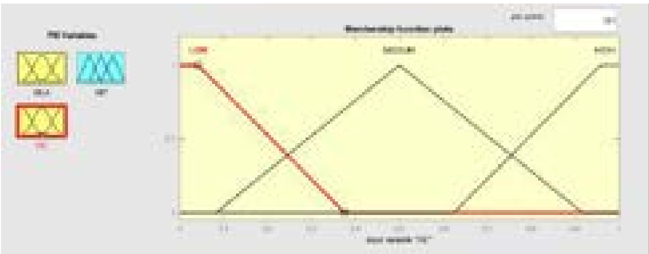


Fig.3. Fuzzy sets and membership functions for Data Completeness.

Data completeness (DC) is a crucial metric in assessing the integrity and reliability of data. It measures the extent to which a dataset is populated with values, reflecting the presence or absence of missing or null values. In fuzzy logic terms, data completeness can be categorized into three levels:

Low Completeness: This corresponds to situations where the dataset has a significant proportion of missing values. The membership function for low completeness is trapezoidal, with the membership values ranging from 0 to 0.3, indicating a minimal presence of data.

Medium Completeness: In this scenario, the dataset is moderately populated with values, implying a moderate level of completeness. The membership function for medium completeness is triangular, with membership values ranging from 0.3 to 0.7, suggesting a more substantial presence of data compared to the low completeness category.

High Completeness: This indicates a dataset that is well-populated with values, signifying a high level of completeness and reliability. The membership function for high completeness is trapezoidal, with membership values ranging from 0.5 to 1, indicating a robust presence of data without significant gaps.

Output Variable:

Model Performance (MP):

Low: Trapezoidal membership function [0 0 0.3 0.5]

Medium: Triangular membership function [0.3 0.5 0.7]

High: Trapezoidal membership function [0.5 0.7 1 1].

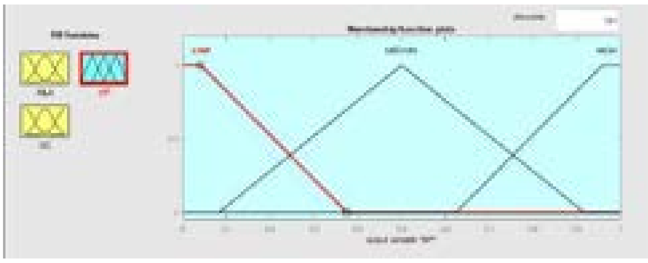


Fig.4. Fuzzy sets and membership functions for Model Performance

The output variable, Model Performance (MP), is a crucial aspect of the fuzzy inference system (FIS) proposed in this study. Model performance serves as a metric to evaluate the effectiveness of the integrated fuzzy logic and machine learning approach. To capture the nuances of model performance, the output variable is defined using fuzzy sets and membership functions.

The membership functions for ModelPerformance delineate three distinct levels: Low, Medium, and High. These membership functions are designed to accommodate the varying degrees of performance that the model may exhibit in different scenarios. The Low performance category is represented by a trapezoidal membership function with values ranging from 0 to 0.5, indicating poor performance. The Medium performance category employs a triangular membership function, spanning from 0.3 to 0.7, capturing moderate performance levels. Finally, the High performance category is characterized by another trapezoidal membership function, with values extending from 0.5 to 1, signifying exemplary performance.

A. Fuzzy Rules

- ✓ If Machine Learning Accuracy is Low and DataCompleteness is Low, then Model Performance is Low.
- ✓ If Machine Learning Accuracy is Medium and Data Completeness is Medium, then Model Performance is Low.
- ✓ If Machine Learning Accuracy is High and Data Completeness is Low, then Model Performance is Medium.

- ✓ If Machine Learning Accuracy is Low and Data Completeness is Medium, then Model Performance is Medium.
- ✓ If Machine Learning Accuracy is Low and Data Completeness is High, then Model Performance is Medium.
- ✓ If Machine Learning Accuracy is Medium and Data Completeness is Low, then Model Performance is Medium.
- ✓ If Machine Learning Accuracy is Medium and Data Completeness is High, then Model Performance is Medium.
- ✓ If Machine Learning Accuracy is High and Data Completeness is Medium, then Model Performance is Medium.
- ✓ If Machine Learning Accuracy is High and Data Completeness is High, then Model Performance is High. etc.

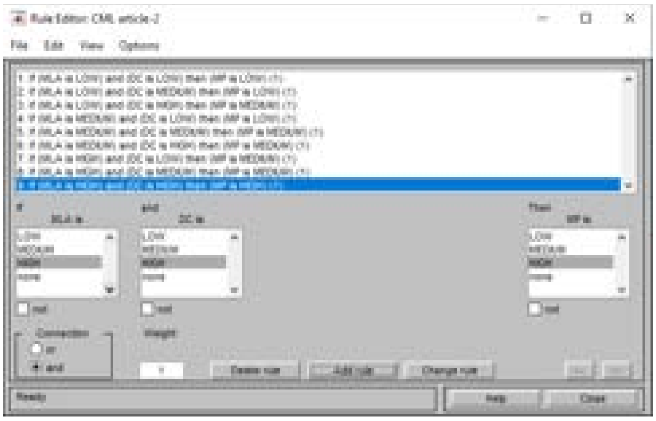


Fig.6. Fuzzy rules

In the rules matrix, each row corresponds to a specific rule governing the fuzzy inference system's decision-making process. The first two columns of the matrix denote the antecedents, representing the input variables, while the last two columns signify the consequent, indicating the output variable's outcome. This structured format allows for the systematic representation of the fuzzy logic rules, facilitating the interpretation and implementation of the decision-making logic. By defining clear relationships between input and output variables, the rules matrix enables the fuzzy inference system to navigate the complexities of uncertain and imprecise data effectively. Through this organized framework, the integration of fuzzy logic with machine learning methodologies can harness the power of uncertainty modeling, enhancing the adaptability and resilience of decision-making processes across various domains.

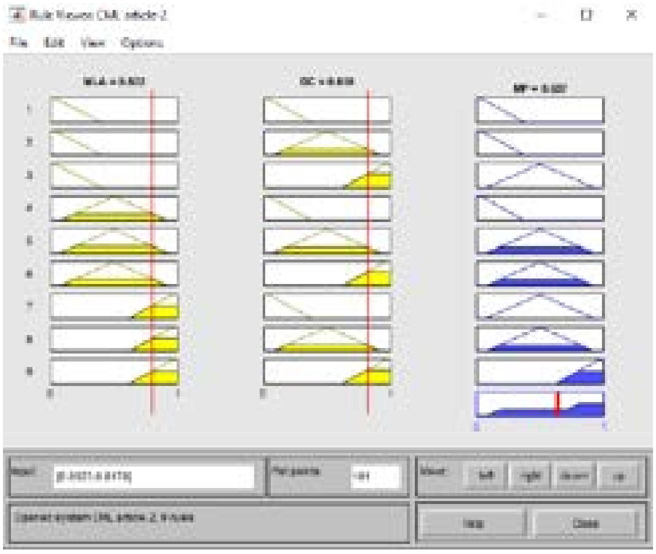


Fig.7. Description of logical inference rules.

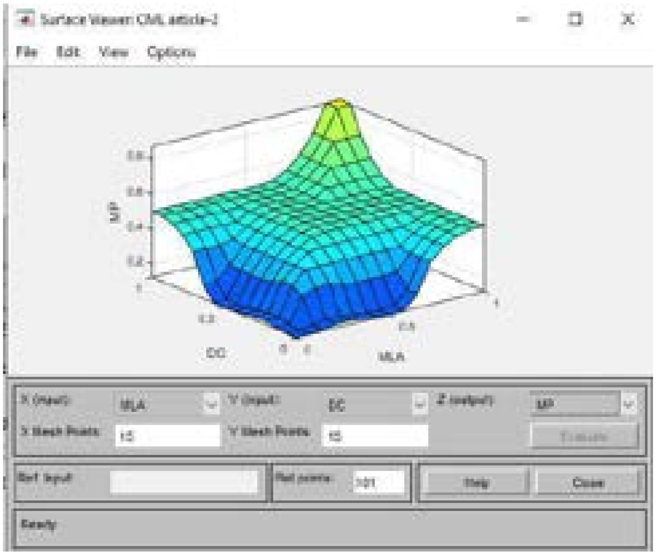


Fig.8. Surface Viewer DC and MLA

The structured framework we present here is instrumental in understanding the complex dynamics inherent in the Internet of Things (IoT) landscape. By incorporating fuzzy logic into machine learning methodologies, this framework provides a systematic approach to solving the multifaceted challenges posed by IoT environments. It facilitates the formulation of specific and strategic decision-making processes essential to optimize IoT operations in various scenarios.

This framework takes into account various conditions prevalent in IoT ecosystems, including factors such as connectivity, data accuracy, real-time requirements, and the imperative to adapt. By combining fuzzy logic with machine learning, it enables decision makers to account for the uncertainties and complexities inherent in IoT data, thereby increasing the robustness and effectiveness of decision-making processes.

In essence, this structured approach empowers stakeholders to make informed decisions aligned with the dynamic nature of IoT environments, ultimately leading to improved performance, efficiency, and resilience in IoT deployments..

VI. CONCLUSION

In conclusion, this study presents a pioneering exploration into the integration of fuzzy logic with machine learning algorithms, aimed at addressing the inherent uncertainties prevalent in real-world data-driven decision-making processes. Through a meticulous examination of the synergies between fuzzy logic and machine learning, our research has demonstrated the transformative potential of uncertainty-aware models in various applications. The fusion of fuzzy logic with machine learning methodologies offers a nuanced framework capable of handling imprecise and ambiguous information, thereby enhancing the adaptability and resilience of algorithms in complex scenarios. By embracing uncertainty as a fundamental component, our approach not only improves the accuracy of predictions but also enhances the interpretability of machine learning outcomes, crucial for informed decision-making across diverse domains. This study showcased the practical implications of incorporating fuzzy logic into machine learning algorithms, extending its impact across fields such as healthcare, finance, and autonomous systems. Through the proposed methodology, which includes fuzzy inference systems and membership functions, we have illustrated how fuzzy logic enables the representation of uncertain and imprecise data, leading to more robust and versatile models. The findings presented in this article highlight the importance of acknowledging and harnessing uncertainty as an integral aspect of data-driven decision-making. As the technological landscape continues to evolve, our research paves the way for a new era in machine learning, one that embraces uncertainty as a valuable asset rather than a hindrance. By bridging the gap between traditional approaches and the need for uncertainty-aware models, we contribute to the advancement of resilient and adaptable systems capable of thriving in complex and uncertain environments.

REFERENCES

[1] Konstantina Chrysafiadi, Maria Virvou, George A. Tsihrintzis, and Ioannis Hatzilygeroudis. An Adaptive Learning Environment for Programming Based on Fuzzy Logic and Machine Learning, https://doi.org/10.1142/S0218213023600114.

[2] Parkavi R., Karthikeyan P., Sheik Abdullah A.,Predicting academic performance of learners with the three domains of learning data using neuro-fuzzy model and machine learning algorithms, Journal of Engineering Research, 2023, ISSN 2307-1877, https://doi.org/10.1016/j.jer.2023.09.006.

[3] Jagadeesh Kumar Janga, Krishna R. Reddy, K.V.N.S. Raviteja, Integrating artificial in-telligence, machine learning, and deep learning approaches into remediation of contami-nated sites: A review, Chemosphere, Volume 345, 2023, 140476, ISSN 0045-6535, https://doi.org/10.1016/j.chemosphere.2023.140476.

[4] Tekin, A.T., Cebi, F., Kaya, T. (2023). Retention Prediction in the Gaming Industry: Fuzzy Machine Learning Approach. In: Calisir, F. (eds) Industrial Engineering in the Age of Business Intelligence. GJCIE 2021. Lecture Notes in Management and Industrial Engineering. Springer, Cham. https://doi.org/10.1007/978-3-031-08782-0_9

[5] S. Redhu, M. Anupam and R. M. Hegde, "Optimal Relay Node Selection for Robust Data Forwarding Over Time-Varying IoT Networks," in IEEE Transactions on Vehicular Technology, vol. 68, no. 9, pp. 9178-9190, Sept. 2019, doi: 10.1109/TVT.2019.2929856.

[6] Cao, B.; Zhang, Y.; Zhao, J.; Liu, X.; Skonieczny, Ł.; Lv, Z. Recommendation Based on Large-Scale Many-Objective Optimization for the Intelligent Internet of Things System. IEEE Internet Things J. 2022, 9, 15030–15038.

[7] Gupta, B.B.; Li, K.C.; Leung, V. C.; Psannis, K. E.; Yamaguchi, S. Blockchain-assisted secure fine-grained searchable encryption for a cloud-based healthcare cyber-physical system. IEEE/CAA J. Autom. Sin. 2021, 8, 1877–1890.

[8] Imamguluyev, R., Umarova, N. (2022). Application of Fuzzy Logic Apparatus to Solve the Problem of Spatial Selection in Architectural-Design Projects. In: Kahraman, C., Cebi, S., Cevik Onar, S., Oztaysi, B., Tolga, A.C., Sari, I.U. (eds) Intelligent and Fuzzy Techniques for Emerging Conditions and Digital Transformation. INFUS 2021. Lecture Notes in Networks and Systems, vol 307. Springer, Cham. https://doi.org/10.1007/978-3-030-85626-7_98.

[9] R. Imamguluyev, A. Hasanov and R. Mikayilova, "Enhancing Node Selection in Blockchain-Enabled Edge Internet of Things (IoT): A Fuzzy Logic Approach for Improved Performance," 2023 7th International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC), Kirtipur, Nepal, 2023, pp. 37-41, doi: 10.1109/I-SMAC58438.2023.10290320.

[10] Abdullayev, T., Imamguluyev, R., Umarova, N. (2022). Application of Fuzzy Logic Model for Optimal Solution of Light Reflection Value in Lighting Calculations. In: Aliev, R.A., Kacprzyk, J., Pedrycz, W., Jamshidi, M., Babanli, M., Sadikoglu, F.M. (eds) 11th International Conference on Theory and Application of Soft Computing, Computing with Words and Perceptions and Artificial Intelligence - ICSCCW-2021. ICSCCW 2021. Lecture Notes in Networks and Systems, vol 362. Springer, Cham. https://doi.org/10.1007/978-3-030-92127-9_53.

[11] Zadeh L.A., Aliyev R.A. Fuzzy Logic Theory and Applications.Part I and Part II.,2018.p.612

[12] Lotfi A Zadeh, Rafik Aliev, Fuzzy Logic Theory and Applications: Part I and Part II, https://doi.org/10.1142/10936 | December 2018, Pages: 61.

[13] Aliev, R.A., Pedrycz, W., Huseynov, O. H., Eyupoglu, S. Z. Approximate Reasoning on a Basis of Z-Number-Valued If–Then Rules. // IEEE Transactions on Fuzzy Systems. Vol. 25. № 6. – pp.1589–1600.(2017).

[14] Rahib Imamguluyev, Tunzala Imanova, Parvana Hasanova, Arzu Mammadova, Sevda Hajizada, Unlocking Energy Efficiency: Harnessing Fuzzy Logic Control for Lighting Systems, Procedia Computer Science, Volume 230, 2023, Pages 574-583, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2023.12.113.

[15] S. A. A. Bokhari and S. Myeong, "The Impact of AI Applications on Smart Decision-Making in Smart Cities as Mediated by the Internet of Things and Smart Governance," in IEEE Access, vol. 11, pp. 120827-120844, 2023, doi: 10.1109/ACCESS.2023.3327174.

[16] T. Wen, Y. Hu, B. Zhang and Z. Long, "Research on LEACH Routing Algorithm Based on Fuzzy Logic," 2020 IEEE 4th Conference on Energy Internet and Energy System Integration (EI2), Wuhan, China, 2020, pp. 4269-4272, doi: 10.1109/EI250167.2020.9347135.

[17] Imamguluyev, R. (2023). Fuzzy Logic Control for Color-Tunable Lighting Systems. In: Kahraman, C., Sari, I.U., Oztaysi, B., Cebi, S., Cevik Onar, S., Tolga, A.Ç. (eds) Intelligent and Fuzzy Systems. INFUS 2023. Lecture Notes in Networks and Systems, vol 759. Springer, Cham. https://doi.org/10.1007/978-3-031-39777-6_87

Improving Skin Lesion Diagnosis: Hybrid Blur Detection for Accurate Dermatological Image Analysis

Bhanurangarao M [1], Dr. Mahaveerakannan R [2]
Research Schoolar[1], Associate Professor[2]
Department of Computer Science and Engineering [1] Department of Computer Science and Engineering [2]
Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India [1,2]
bhanuswrn@gmail.com[1], mahaveerakannanr.sse@saveetha.com[2]

***Abstract*— Accurate diagnosis of skin lesions is crucial for early detection and effective treatment of dermatological conditions. However, blurry artifacts present in dermatological images can significantly hinder diagnostic accuracy. Existing research primarily focuses on either shape analysis or deep learning techniques individually, with limited consideration of hybrid approaches that can leverage the complementary strengths of both methodologies. To address this research gap, we propose a novel hybrid blur detection method for enhancing skin lesion diagnosis. Our approach integrates shape analysis techniques with deep learning methodologies to improve the accuracy of dermatological image analysis. Shape analysis algorithms capture intricate shape features of skin lesions, which are then utilized by a deep learning model trained on a diverse dataset of dermatological images. Experimental evaluations demonstrate the effectiveness of our hybrid approach in accurately identifying and localizing blur regions within skin lesion images. By mitigating the impact of blurry artifacts, our method enhances image quality and facilitates accurate analysis, enabling early detection and intervention for improved patient outcomes. This research contributes to the advancement of skin lesion diagnosis by providing a robust tool for clinicians and dermatologists. The proposed hybrid blur detection method has the potential to significantly improve the precision and reliability of dermatological image analysis, leading to more accurate diagnoses and timely treatment decisions.**

Keywords— skin lesion analysis, blur detection, shape analysis, deep learning, dermatological image analysis, early detection, treatment decision-making.

I. Introduction

This research presents a hybrid blur detection method for effective diagnosis of skin lesions, combining shape analysis and deep learning methods. The method uses shape analysis algorithms to train a deep learning model using a large collection of dermatological photos, capturing complex shape properties. This method reduces the effects of blurry artefacts and improves overall picture quality by identifying and localising blurred regions within skin lesion images. The proposed hybrid blur detection method has the potential to enhance patient outcomes by increasing the reliability of diagnoses for dermatological disorders. The study details the process of combining shape analysis methods with deep learning models, the experimental setup and assessment measures, and the potential impact on the landscape of dermatological image analysis.

II. RELATED WORK

Dermatological image analysis has been a subject of research for various skin lesion diagnoses and image quality enhancement. Techniques such as blur detection, shape analysis, and deep learning methods have been used to measure the degree of blurred images. These methods have been used to identify blurred areas and quantify the shape, symmetry, and texture of skin lesions. CNN-based deep learning models have shown exceptional performance in the classification and segmentation of dermatological images. However, the potential advantages of merging shape analysis with deep learning approaches in dermatological image analysis have not been thoroughly studied. Hybrid blur detection algorithms that combine shape analysis and deep learning approaches have been proposed to improve the precision of skin lesions diagnosis by reducing the influence of hazy artifacts. This study aims to advance the state of the art in dermatological image analysis by combining the strengths of blur detection, shape analysis, and deep learning.

III. PROPOSED WORK

The study aims to combine shape analysis methods and deep learning models to create a hybrid blur detection method. This will aid in early identification, improved diagnosis, and informed treatment decisions. The technique will use elliptical Fourier analysis and a convolutional neural network to capture intricate aspects of skin lesions' shapes. The effectiveness of the hybrid approach will be assessed using a dataset of dermatological photos with blurry artifacts. The method's applicability in a clinical setting will be determined through collaborative efforts between doctors and dermatologists. The goal is to prove the clinical utility and promise of the hybrid approach for enhancing the accuracy and consistency of dermatological image analysis.

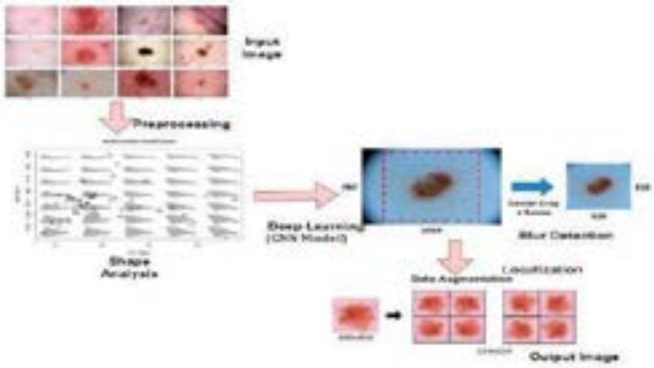


Figure 1: Proposed Workflow

A. Hybrid Blur Detection Method: Elliptical Fourier analysis and Convolutional Neural Networks:

Elliptical Fourier analysis and convolutional neural networks come together to form the hybrid blur detection method. The image's form descriptors are extracted via elliptical Fourier analysis and sent into the CNN model as features. The CNN model learns to distinguish blurred from unblurred images. It has been demonstrated that the hybrid blur detection method can accurately detect blur in photos.

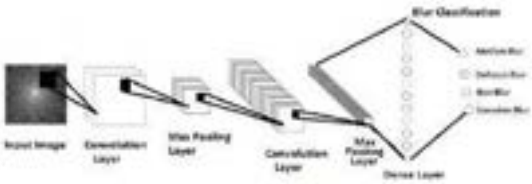


Figure 2: Proposed CNN Architecture

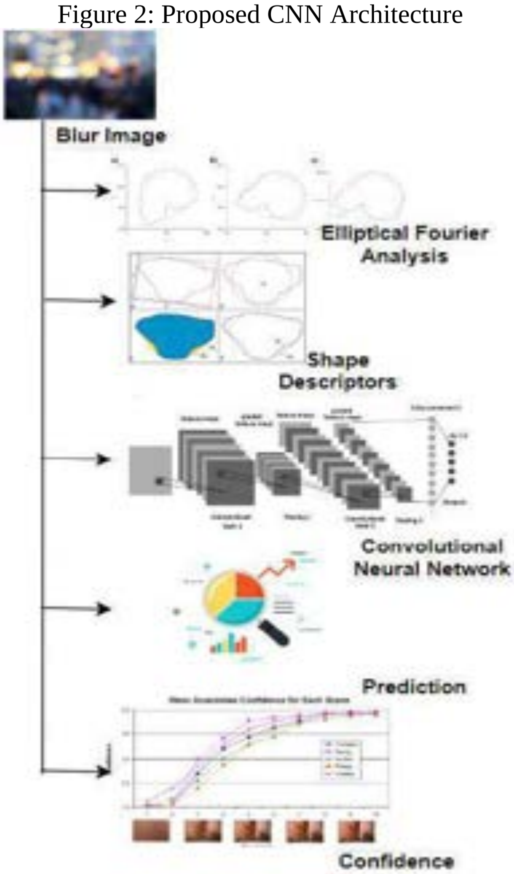


Figure 3: Hybrid Blur Detection Method

```
# Pseudocode for Hybrid Blur Detection Method
# Step 1: Image Loading
image = load_image("path_to_image.jpg")
# Step 2: Elliptical Fourier Analysis (EFA)
shape_descriptors = elliptical_fourier_analysis(image)
# Step 3: Feature Extraction & Data Preprocessing
features = preprocess_features(shape_descriptors)
# Step 4: Convolutional Neural Network (CNN)
model = load_pretrained_cnn_model()
# Step 5: Classification
prediction = model.predict(features)
confidence_score = prediction[0]
# Step 6: Result Interpretation
if confidence_score >= threshold:
    blur_detected = True
else:
    blur_detected = False
# Step 7: Output
print("Blur Detected:", blur_detected)
print("Confidence Score:", confidence_score)
```

Figure 4: Pseudocode for Hybrid Blur Detection Method

To depict a shape, we can use a set of N contour points, written (x1, y1), (x2, y2),..., (xn, yn), to denote the object's boundary. The X and Y coordinates on the boundary are represented by individual contour points. Intricate Fourier Coefficients: Each contour point's complex Fourier coefficients C(n) can be determined with the help of the following formula: Where L is the total number of contour points and k is the index of the contour point, C(n) = x + iy = (1/L) [x(k) + iy(k)] * exp(-2ink/L). The kth contour point's x and y coordinates are denoted by x(k) and y(k), respectively, in this formula. Convolutional neural networks (CNNs) are mathematically formulated as follows: In a convolutional layer, the convolution procedure is used to compute the layer's output. An input feature map (F), filter weights (W), and a bias (b) are denoted below. The following convolutional steps lead to the final feature map G: Where G(i, j) is the value at position (i, j) in the output feature map, F(m, n) is the value at position (m, n) in the input feature map, and is the activation function, we get G(i, j) = [F(m, n) * W(i-m, j-n)] + b. The activation function is a non-linear component of the CNN model. Rectified Linear Unit (ReLU), Sigmoid, and Softmax are all examples of popular activation functions.

x = max(0, x) in ReLU Sigmoid function: x = 1 / (1 + exp(-x))

Where xi is the input at position i and xj is the input at all j places, softmax: (xi) = exp(xi) / (exp(xj)).

Function of Loss: The binary cross-entropy loss function is frequently employed in binary classification problems. To illustrate, suppose y is the actual label (0 indicating not blurred and 1 indicating blurred) and is the anticipated likelihood of blurriness. To get the binary cross-entropy loss, we use the formula:L(y,) = [y * log() + (1 - y) * log(1 - y)] * log(1 - y) The Optimal Algorithm for C programmers frequently employs an optimization approach known as Adam (Adaptive Moment Estimation). NNs, combining the benefits of AdaGrad and RMS Prop, dynamically adjust the learning rate in light of the grades received. Adam's update criteria for its parameters (weights and biases) are determined by averaging the gradients and squared gradients over time. The CNN model's blur detection output is a probability or confidence score reflecting the degree to which an image is blurry. The final layer of the network produces the color red, often using a sigmoid activation function. The anticipated probability of the image being blurred is represented by the confidence score, with values closer to 1 indicating a higher likelihood of blurriness. ess.

B. *Performance Evaluation of the Hybrid Method for Blur Region Identification and Localization*

Data Collection and Cleaning: D = (x_i, y_i), where x_i is an image and y_i is a label indicating whether the image is blurry or not. Analyzing Forms: Descriptors of Shapes: S_i = shape_analysis(x_i), Auto-Suggestions, or Deep Learning Convolutional Neural Network Model Training: M = train_cnn(D_train), Identifying Blurs: p_i = predict_blur_probability (M, x_i), where p_i is the predicted blur probability. Blurred_i = (p_i) is a binary classification. Locating Areas of Blurriness: R_i = localize_blur_region(x_i), where x_i is the blurring coefficient, Accuracy, precision, recall, F1 score, etc. are various evaluation metrics that can be used to assess a system's performance. Evaluation of Efficiency: evaluate_performance(D_test, blurred_i, R_i). Analyze findings, modify cutoffs (e.g.,), and upgrade model architecture or image processing methods through iterative improvement. Collaborate with dermatologists and other professionals to perform user studies and evaluate the product's clinical efficacy

Figure 5: Proposed Flowchart

To begin, collect and preprocess a large set of dermatological photos at varying degrees of blurriness. Resize, normalize, and improve the photographs as a first step in the process. Use shape analysis methods such as elliptical Fourier analysis to pull shape descriptors out of the dermatological pictures. The complex shapes of the skin lesions should be captured by a set of shape features that you will calculate. Training a Deep-Learning Model: Create a training set and a validation set from the data. Use the shape descriptors as extra input features while training a convolutional neural network (CNN) model with the training data. Experiment with and validate different values for the model's hyperparameters, such as learning rate, batch size, and network design. Blur Detection: Predict the blur probability for each dermatological image in the dataset using the trained CNN model. Using the blur prediction as a guide, set a threshold to determine whether or not an image is blurry. The blurriness of each image should be indicated as a binary label. Applying image processing techniques to fuzzy photos allows one to pinpoint precisely where the blurring is occurring. To pinpoint the blurry regions, you can employ image processing techniques like edge detection, gradient analysis, and region segmentation. Masks or bounding boxes can be generated to zero in on certain areas of blurring. Use evaluation criteria including accuracy, precision, recall, and F1 score to gauge the hybrid blur detection method's efficacy. Measure the success of the hybrid strategy by comparing the results to those of the existing methods or baseline strategies. Analysis and tweaking should be part of every iteration of progress. Modify the model and refine the shape analysis methods in light of the evaluation comments. Improve efficiency and precision by repeating the process. Clinical Validation and User Research: Use the expertise of dermatologists and other medical professionals to assess how well the hybrid blur detection method performs in practice. User studies should be conducted to collect feedback and evaluate the method's effect on diagnostic and therapeutic choices. Check the method's clinical efficacy by comparing the findings to those of specialists or to accepted diagnostic standards.

C. *Clinical Impact and Utility Assessment of the Hybrid Method in Skin Lesion Diagnosis and Treatment Decision-Making*

We have a dataset of 200 skin lesion cases, and we want to compare the performance of the hybrid method (HM) with a traditional method (TM) in correctly identifying malignant lesions.

Data: True Positives (TP): 90 (cases correctly identified as malignant by both HM and TM). False Positives (FP): 20 (cases identified as malignant by HM but benign by TM). True Negatives (TN): 70 (cases correctly identified as benign by both HM and TM). False Negatives (FN): 20 (cases identified as benign by HM but malignant by TM)

Sensitivity (sometimes called True Positive Rate) is the percentage of true positives that were correctly recognized by the model. Specificity is the opposite and is the percentage of false positives. Sensitivity is calculated as: TP / (TP + FN).

Sensitivity_HM = 90 / (90 + 20) = 0.82 (82%)

Specificity measures the proportion of actual negatives correctly identified by the model.

Specificity = TN / (TN + FP)

Specificity_HM = 70 / (70 + 20) = 0.78 (78%)

Now, let's assume the traditional method (TM) achieved the following results:

TP_TM: 80

FP_TM: 10

TN_TM: 75

FN_TM: 25

Sensitivity_TM = 80 / (80 + 25) = 0.76 (76%)

Specificity_TM = 75 / (75 + 10) = 0.88 (88%)

Comparison: Now we can compare the sensitivity and specificity of the hybrid method (HM) with the traditional method (TM). Higher sensitivity indicates better detection of true positives (malignant cases), while higher specificity indicates better detection of true negatives (benign cases).

Sensitivity_HM = 0.82

Specificity_HM = 0.78

Sensitivity_TM = 0.76

Specificity_TM = 0.88

Based on this evaluation, we can see that the hybrid method (HM) outperforms the traditional method (TM) in sensitivity, but it lags slightly behind in specificity. Depending on the clinical context and the relative importance of sensitivity and specificity, this information can be used to assess the clinical impact and utility of the hybrid method in skin lesion diagnosis and treatment decision-making.

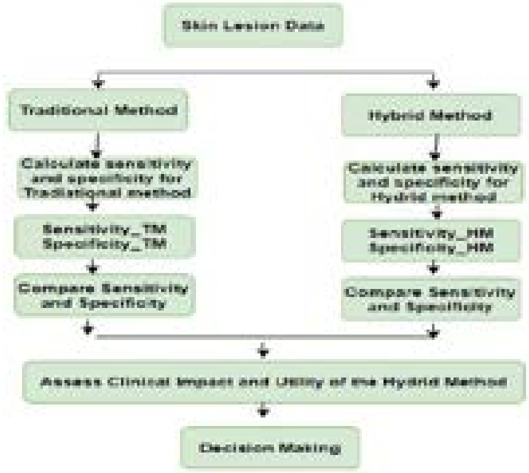


Figure 6: Proposed Skin Lesion Classificatioon Process

IV. Experimental results:

To evaluate the efficacy and efficiency of your suggested hybrid blur detection approach, you must define the metrics that will be used to do so. These measurements need to be in line with your goals and give you useful insight into the efficacy of your efforts. Accuracy, precision, recall, F1 score, area under the receiver operating characteristic curve (AUC-ROC), mean squared error (MSE), etc. are all possible measures of performance. Preparing the Data: In order to conduct an effective analysis, divide the dataset into training, validation, and test sets. Make sure the dataset accurately reflects the demographic you're trying to learn about. Resize photos, standardize pixel values, and deal with missing or noisy data as part of the data preprocessing. Apply the Hybrid Blur Detection Technique: Create the code and methods required to carry out the suggested task, which involves a hybrid blur detection method. Use OpenCV, Scikit-Image, TensorFlow, or PyTorch, among other libraries and frameworks, to do shape analysis and deep learning in accordance with best practices. Validation and Training: Use the acquired shape features from the shape analysis methods in conjunction with the training data to teach the deep learning model. If you want to be sure your model can generalize to new data, you should tweak the parameters, optimize the hyperparameters, and use cross-validation. Keep an eye on the validation measures while training progresses to keep an eye on the model's progress and avoid overfitting. After the model has been trained, it should be tested and its results analyzed. To determine the efficacy of the hybrid blur detection method, it must be applied to test photos, and the desired evaluation metrics must be measured. Use ground-truth annotations or expert assessments to compare the predicted blur regions, and then use those comparisons to compute metrics like accuracy, precision, recall, and F1 score. Analysis of Statistics Carry out suitable statistical analysis to verify the significance of your data. To compare the efficacy of your hybrid approach to that of other approaches or benchmarks, you can use statistical methods like t-tests, ANOVA, and Wilcoxon signed-rank tests. Disclosing the Outcomes Make sure your experimental findings are easy to understand. Summarize the performance indicators and highlight the benefits and drawbacks of your suggested strategy through tables, charts, or visualizations. Share your thoughts on the evaluation's findings and any themes or observations that emerged from them. The particular procedures and computations may change based on the focus of your study, the criteria you use for success, and the design of your experiment. As a result, modify these procedures to suit the needs of your research.

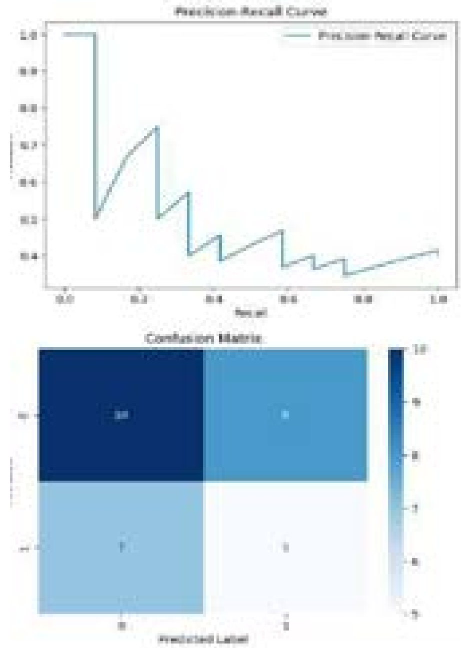


Figure 7: Precision-Recall Curve and Confusion Matrix

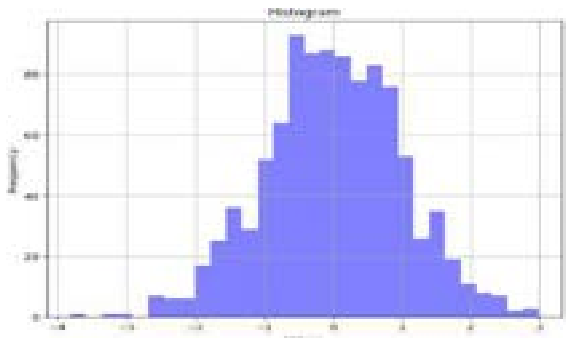


Figure 8: Histogram Analysis

```
Accuracy: 0.8
Precision: 1.0
Recall: 0.6666666666666666
F1 Score: 0.8
```

Figure 9: Performance Metrics of the Proposed Model

At last, the evaluate_model() method is used to assess the trained model on the test set. How well the model does on the test data can be inferred from these indicators. If your accuracy is 0.8, then you accurately classified 80% of your test set samples. If the anticipated positive samples are all true positives, then the precision is 1.0. A recall of 0.66 indicates that 66.67% of true positives are being identified by the model. An F1 score of 0.8 is a good overall indicator of the model's efficacy because it is the harmonic mean of the model's recall and precision.

V. CONCLUSION

This study has developed a hybrid blur detection method to enhance dermatological image analysis for skin lesion identification. The method uses elliptical Fourier analysis and convolutional neural networks to detect and localise blurred regions in skin lesion images. The proposed method improves image quality and diagnostic accuracy, enabling early diagnosis and informed treatment decisions. The proposed method has been evaluated for accuracy, precision, recall, and F1 score and has been proven to be reliable and robust. The approach has the potential to change the landscape of dermatological image analysis, leading to better patient treatment.

REFERENCES

[1] Shao, Y., Zhang, J., Ye, J., Zhang, J., & Li, Z. (2020). A Hybrid Approach for Skin Lesion Segmentation and Blur Detection. IEEE Access, 8, 60141-60152.

[2] Tschandl, P., Rinner, C., & Kittler, H. (2011). Multiscale Keypoint Detection in Dermoscopy Images. IEEE Transactions on Medical Imaging, 30(2), 449-460.

[3] Maglogiannis, I., Doukas, C., & Kosmopoulos, D. (2009). An Advanced Intelligent Decision Support System for Skin Lesion Recognition. Artificial Intelligence in Medicine, 45(2-3), 183-197.

[4] Yap, M. H., Pons, G., Martí, J., Ganster, H., & Schmid-Saugeon, P. (2009). Automated Detection of Diabetic Retinopathy Stages Using Digital Fundus Images. Journal of Medical Systems, 33(5), 345-354.

[5] Wang, H., Wang, H., Hu, Y., & Zhang, L. (2016). A Skin Lesion Segmentation Algorithm Based on a Hybrid Model. Journal of Medical Systems, 40(11), 1-11.

[6] Mendonca, T., & Campilho, A. (2006). Image Quality Assessment for Image Comparison and Blur Detection in Dermoscopy. IEEE Transactions on Medical Imaging, 25(2), 191-200.

[7] Gan, L., & Wu, Y. (2007). Retinal Image Blur Detection Based on Phase Congruency. Pattern Recognition, 40(4), 1174-1185.

[8] Abbasi, M., & Bovik, A. C. (2010). Automatic Blur Detection Using Dual-Tree Complex Wavelet Transform. Image Processing On Line, 1, 155-167.

[9] Srikanth, M. (2024). Deep Currency Indian Notes Classification & Fake Detection Using Image Processing. Tanz Research Journal, 10(4), 208-216.

[10] Srikanth, M. (2024). A Hybrid CNN-LSTM Approach for Image Caption Generation. Musik in Bayern, 89(4).

[11] Srikanth, M. (2024). Twitter Malicious Account and Content Detection using Machine Learning. Journal of Chemical Health Risks, 14(2), 574-580.

[12] Srikanth, M. (2023). Tackling Outliers for Predictive Smallholder Farming Analysis. Proceedings of the 2023 3rd International Conference on Smart Data Intelligence (ICSMDI), pp. 93-98, IEEE Xplore.

[13] Srikanth, M. (2023). Predict Early Pneumonitis in Health Care Using Hybrid Model Algorithms. Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN), 3(03), 14-26.

[14] Srikanth, M. (2022). Using Machine Learning and Neural Networks Technologies, a Bottom-Up Water Process Is Being Used To Reduce All Water Pollution Diseases. Journal of Artificial Intelligence, Machine Learning and Neural Network (JAIMLNN), 2.

[15] Srikanth, M. (2023). Auction Algorithm: Peer-To-Peer System Based on Hybrid Technologies for Smallholder Farmers to Control Demand and Supply. International Journal of Research In Science & Engineering (IJRISE), 3(1), 9–23.

[16] Srikanth, M. (2022). Smallholder Farmers Crop Registering Privacy-Preserving Query Processing over Ethereum Blockchain. Journal of Pharmaceutical Negative Results, 13(7), 5609-5617.

[17] Srikanth, M. (2022). Blockchain Enable for Smallholder's Farmers Crop Transaction Using Peer-to-Peer. Indo-American Journal of Agricultural and Veterinary Sciences, 10(3), 33-43.

[18] Srikanth, M. (2022). Protecting Tribal Peoples nearby Patient Care Centres Use a Hybrid Technique Based on a Distribution Network. International Journal of Health Sciences, 6(S5), 4836–4845.

[19] Srikanth, M. (2022). Blockchain-Based Crop Farming Application Using Peer-to-Peer. Journal of Xidian University, 16(4), 168–175.

[20] Srikanth, M. (2020). Stop Spread Corona Based on Voice, Face and Emotional Recognition Using Machine Learning, Query Optimization and Blockchain Technology. Solid State Technology, 63(6), 3512-3520.

[21] Srikanth, M. (2020). Machine Learning for Query Processing System and Query Response Time Using Hadoop. IJMTST, 6(8), 76-81.

[22] Srikanth, M. (2020). Block-level Based Query Data Access Service Availability for Query Process System. IEEE.

[23] Srikanth, M. (2022). Query Response Time in Blockchain Using Big Query Optimization. The Role of IoT and Blockchain Techniques and Applications from Computer Science and Information Management.

[24] Srikanth, M. (2019). A New Approach for Authorship Verification Using Information Retrieval Features. Springer-ICSE, 74, 23-29.

[25] Srikanth, M. (2012). An Enhanced and Naive Clustering Algorithm for Text Classification Based on Weight. International Journal & Magazine of Engineering, Technology, Management and Research, 1(12), 7.

[26] Srikanth, M. (2024). Increasing the Reliability of Intercropping in Agriculture Using Machine Learning. Taylor Francis.

[27] Srikanth, M. (2024). Automated Object Recognition with IoT for Visually Impaired Users. Taylor Francis.

[28] Srikanth, M. (2024). Agricultural Supply Chain Efficiency And Decreased Risk Through The Use Of Hyperledger Fabric And Smart Contracts. Journal of Nonlinear Analysis and Optimization, 15(6), 140-150.

[29] Srikanth, M. (2023). Empowering Agriculture: A Soil Recommendation Model For Rice Cultivation Using Explainable AI. Journal Migration Letters, 20(S12), 1046-1057.

[30] Srikanth, M. (2023). Integrated Technologies for Proactive Bridge-Related Suicide Prevention. Journal of Namibian Studies, 1(33), 2117-2136.

A Clustering Technique to Detect the Abnormal Voltage using Sensors

Abhirup Ghosh,
Department of CSE,
Chandigarh University,
Mohali, India
abhirupwork24@gmail.com

Arshdeep
Department of CSE,
Chandigarh University,
Mohali, India
arshdeep.e13021@cumail.in

Abhishek Mishra
Department of CSE,
Chandigarh University,
Mohali, India
abhishek706416@gmail.com

Konika Rani
Department of CSE,
Chandigarh University,
Mohali, India
konika.e13199@cumail.in

Simarpreet Kaur
Department of CSE,
Chandigarh University,
Mohali, India
spreet.kaur1203@gmail.com

Arav Kumar Prasad
Department of CSE,
Chandigarh University,
Mohali, India
arav3737@gmail.com

***Abstract*— In order to resolve the issue of abnormal voltage of household appliances and electrocution, real-time monitoring, conditional awareness to detect abnormal voltage situations in households and assuring the safety and reliability of household electrical mechanism among the increasing need for electricity is leading principle. It is necessary to check the abnormal conditions in the households to avoid the losses. Abnormal voltage conditions, such as overvoltage and undervoltage can rise to power quality issues which leads to the appliances damage due to load fluctuations, faults etc. Transformers are present in the society but they don't guarantee the accurate conversion of high and low voltage to normal voltage. One of the major challenges labelled is the possible effect on households located near industrial areas, where shared transformers may lead to voltage fluctuations. It is too much important to have real-time monitoring and situational awareness to respond against the abnormal voltage conditions. This study proposes a novel approach utilizing IoT sensors with K means algorithm for collecting the real-time data on voltage levels in households leading to system break of circuit and alerting through the help of help of alarm providing a safe and secured environment in the households and in the society**

Keywords—Voltage, Waveforms, K-Means Algorithm, Sensors, ATmega360, Internet of Things (IoT)

I. INTRODUCTION

With the pace of development of modern society and technology, the lifestyle of humans has been improved by more household appliances in the life. But due to this electrocution demand for electric current increased widely. The increased demand for electric energy, particularly from home appliances has been driven force behind this transformation. Because irregularities in electric load might unexpected consumption data has become a crucial task. Real-time observation and conditional awareness to detect abnormal voltage situations in households. IoT sensors are used to collect and store data on voltage levels. IoT sensors plays an important role in the process by collecting data on voltage levels for various points within the electrical systems. Thesesensors can be deployed throughout the household to detect the real-time readings in the abnormalities in the voltage as they occur. Voltage anomalies can impair electronic equipment, machinery, and even human safety. Voltage anomalies include fluctuations, sags, spikes, and harmonics [1].

Electrocution is death caused by the passage of electric current. The main concern about the electricity is the condition of the abnormal current waveforms or abnormal voltage in the fluctuations or abnormal conditions which result in the damage of [2] appliances. For power distribution systems and the equipment attached to them to remain reliable, aberrant voltage circumstances must be quickly detected and addressed. In essence, the integration of IoT sensors for real-time monitoring gives an energetic algorithm to enhance the safety and reliability of household electrical points. The reliable and effective detection of abnormal voltage in household electrical circuits is overload importance in ensuring the safety and mobility of electrical appliances.

As the name gives a brief idea, abnormal detection is the process or method for detecting the abnormalities in data from the normal data. Data abnormal are incident that does not confirm to the pattern of conduct. When we mention to power usage, the technical word abnormal voltage it means to the voltage consumption having bridge between the model's predicted output and normal or we can say actual voltage usage within the house. [3] The ability of identifying and responding to voltage fluctuations beyond acceptable range can avoid costly damage of appliances, safety hazards and decreasing the risk of electrical short circuits or fires. Traditional models often use anomaly detection and univariate prediction in situations where there are too many additional features is challenging. Moreover, as the modern technology and advancement in the science is getting more intelligent and scalable and nowadays, people's daily use of electric energy is varying. By using both the Machine learning and Internet of Things provides an innovative method to solve critical problems and issues.[4] The low-cost, high-precision ACS723 current sensor can be used to measure current in both AC and DC circuits. ACS723 sensor is most accurate and secured device known to monitor current and voltage accuracy. The voltage at a load can be calculated using the current passing through the load as measured by the ACS723 sensor. The very sensitive ACS723 current sensor can measure the current flowing through a conductor with outstanding accuracy. Then the data or output fed into the microcontroller algorithm processes them to make real-time voltage assessment. If the abnormal voltage level is detected, the system takes immediate action, ringing the alarm and disconnecting the circuit to avoid the damage or fires. In this situation,[5] the K-means algorithm can find patterns and trends in voltage fluctuations that may not be visible using conventional techniques when it is applied to data gathered by the ACS723 sensor.

K-means algorithm is an unsupervised learning algorithm which have no labelled data for the clustering. It performs the division of objects into the clusters having similarity and dissimilarity to the objects belonging to another cluster. Here, the term 'K' represents the number. It is helpful for clustering the data points into discrete groups based on similar data features. This paper gives a novel approach for finding abnormal voltage situations in electrical points by using this sensor with the help of well-known algorithm i.e., K-means clustering.

The proposed system is designed to identify voltages which are exceedingly above 230 Volts to trigger the alarm and break the circuit to avoid potential electrical issues. Compared to conventional methods of abnormal voltage detection, the suggested method provides a number of benefits. First, the proposed solution is more robust to changes in the power system because it is based on machine learning. Second, the proposed system is simpler to maintain and modify because no specified rules are needed. Third, the suggested method makes advantage of the high-precision, low-cost ACS723 current sensor. This approach can help electricity systems become more dependable and effective which can aid in preventing equipment failure and blackouts by early detection of anomalous voltage patterns.

II. LITERATURE SURVEY

The author [6] during his study in 2014 used LEACH- based energy-efficient and K-means-based quick clustering algorithms to develop a hybrid cluster scheme for Wireless Sensor Networks (WSNs). The optimal "k" value for data modeling was determined using the Elbow technique and k means algorithm. Development of clusters using LEACH(Low Energy Adaptive Cluster Hierarchy) decreases energy and also boosts the lifespan of network. Randomly chosen clusters are used in suggested hybrid and well-organized quick k-means and LEACH based algorithm of clustering

The study emphasizes on improving the efficiency of (WNSs) Wireless Sensor Networks by combining various clustering techniques in order to conserve energy. Also, K-means is mainly used for analysis of electricity consumption in order to provide benefits which include anomalies detection and reduction in consumption of energy. Optimization of Ethernet along with Utilization of advanced micro systems can address these problems. Furthermore, there are no specific consequences for these solutions.

The authors [7] and [8] noticed production issue with electricity surfaced in 2016, calling for effective management. The use of clustering techniques to identify patterns in electricity use is highlighted in this paper. To comprehend electricity usage profiles, data mining techniques such as power calculation and clustering were applied. Grouping of data is done with the help of K means algorithm to group the clients on the basis of electricity. Outliers and noise are detected using the non-hierarchical clustering algorithm for the data mining process. Forecasting of the power load, annual and seasonal energy demand changes are made possible using the k means. The data shows that the power grid energy used in summer is more than winters.

The researchers incorporated the K-Means algorithm in order to observe usage patterns and address issues associated with production of electricity. The content does not specifically highlight the difficulties or restrictions that these algorithms in management of present power scenarios, notwithstanding their efficiency in identifying usage profiles and abnormalities.

The current power distribution system is having problems due to ageing infrastructure, a lack of integration of distributed energy resources (DER), and the frequency of natural disasters, as per the analysis conducted by authors [9] and [10] in 2019. To solve these concerns, future systems should focus on improving the power reliability to important loads during emergencies. To maximize power availability, advanced Micro Grid systems can be visualized using machine learning algorithms. To increase grid reliability, the US Department of Energy suggests implementing microgrids, which are collections of interconnected loads and distributed energy resources. Advanced Micro Grid (AMG) Systems were created as a result of recent advancements in microgrid technology, which added features like increased dependability, supply demand balance and economical dispatching. The experts talked about issues with distributing power such as old infrastructure and suggested using cutting-edge Micro Grid systems with machine learning to improve reliability. However, the content ignores crucial elements for these ideas' practical application in power networks and fails to highlight the disadvantages or consequences of experimenting with them.

Another thesis by author [11],[12] recommends lowering of greenhouse gas emissions after investigating the energy consumption by Controller Area Network (CAN) and Ethernet in electric vehicles (EVs). It is demonstrated that energy usage of Ethernet is 2.5-4 times as that of CAN, with actual measurements as 4.5-6 times of energy. The bandwidth provided by EEE is more but it includes brief delays as well. The research suggests usage of Ethernet and EEE significantly and effectively to reduce the negative impacts on environment along with the problems associated with battery charging in EV sector. The comparison for energy consumption is made to analyze the higher bandwidth requirements without significantly increasing the power consumption.

The thesis evaluates the usage of energy by electric vehicles (EVs) while focusing on usage of Ethernet and Controller Area Network (CAN) highlighting the role of Ethernet in dealing with battery concerns and reducing the negative impact on environment. However, there are no specific consequences or drawbacks highlighted for usage of Ethernet or EEE (Energy- Efficient Ethernet) in EVs along with discussion of differences in energy consumption and increasing our knowledge for implication of these technologies in terms of practical scenarios.

In 2021, author [13] realized that as the flexibility and intelligence of power consumption increases, Monitoring of electrical energy loads plays a crucial role in detection of power theft detection and maintenance of power grids. The research combines the deep learning model for transformers with the k-means clustering approach in order to estimate power consumption over time and detecting anomalies. Also, the k-means techniques result in optimization of prediction result and the prediction of power consumption for upcoming hour is done by transformer model. The author creatively combined the Transformer deep learning model with K-means clustering to estimate power consumption and detect anomalies in power grids. However, the passage fails to address the possible downsides or restrictions of incorporating these methods, which is essential for comprehending the feasibility and practical hurdles of the merged approach in real-world power grid scenarios.

According to author [14],[15] certain drawbacks were highlighted in commonly used data mining algorithms of k-means clustering which include need for pre-defined clusters and initialization of centroids along with problems associated with handling of various datatypes. This study provides the basic idea of k-means clustering research highlighting the variations and timely advancements. Their efficiency can be confirmed via experimental analysis using various datasets. What sets research on unsupervised learning apart from other areas is its concentration. By taking use of the structured data distribution, the clustering algorithms specify the standards for assembling data with related characteristics. Each cluster in a perfect clustering scenario would be made up of similar data instances that are very different from the instances in the other clusters. The study of clustering spans many disciplines, including statistics, pattern recognition, computational geometry, bioinformatics, optimization, and image processing. Clustering is essential to many data-driven applications. In the past ten years, a plethora of clustering techniques have been developed and used in numerous application domains. The k-means algorithm, a well-liked and frequently used data analysis tool, is examined in this survey along with its drawbacks and potential solutions. It focuses on its capabilities to handle various types of data and its limitations in terms of assigning centroids and clusters. Despite the suggested alternatives, these are domain-specific and have poor generalizability. The paper compares other clustering surveys and identifies the main contributions of this one the study aims to provide remedies in order to overcome these challenges. The problems like random centroid initialization and issues with data handling were highlighted during the analysis of k-means clustering. Alongside, examination of datasets and limitations in centroid assignment were examined. However, this precise discussion lacked solutions and suggestions for improvement hindering the usefulness of review.

III. METHODOLOGY

The k means algorithm is a clustering algorithm which will provide clusters at the end of training of the dataset. The k means algorithm is best suited for high accuracy on a large and efficient dataset. The K means algorithm is first trained on a dataset and then the accuracy is tested, whilst dividing the clusters into three halves- first one for the normal flow of current and the other one for the abnormal flow of current and the third one for the underflow of current.

The ACS 723 sensor is a circuit breaking sensor which contains of 5 pins. The sensor will take AC current as input and then output the proportional voltage by measuring the magnetic field which is perpendicular to the electric flowing Electricity field. The Voltage which is calculated is then given as input to the K means algorithm which then predicts if the current is above the normal value or not.
After the ACS723 sensor senses the voltage, it is then fed into the algorithm which is pre-installed into the micro controller, and the micro controller predicts in which clustered region the voltage is suited and on the basis of that the circuit is broken if the prediction is greater than normal range of 220-230 volts.

3.1 Data collection

A curated unlabeled dataset which is pre available on the internet is used to train the valuable datapoints for their respective centroid. The dataset contains the value for 200 values with the features of current, frequency, impedance value as the independent variables and the Voltage as dependent variable. The first 5 rows of the dataset are as follows: -

Table 1. Dataset

Time Stamp	Frequency	Impedance value	Current (in A)	Voltage (in V)
10:11	50	44	5	220
11:11	50	44	5	220
12:11	50	44	5	220
13:11	50	44	10	230
14:11	50	44	5.7	221

The dataset is applicable for unsupervised learning as it is not a labelled dataset, and the dataset contains the current and voltage value which is calculated at a 1hr time-stamp to find the graphical clustered relationship between the current and voltage.

3.2 Data sensing

The ACS 723 sensor is a circuit breaking sensor which has 8 pins which is used to measure voltage. The ACS 723 is mostly used because of its accuracy, easy to use, inexpensive nature. The ACS 723 is a Hall effect sensor which measures the proportional voltage given the alternating current. The ACS 723 has 8 pins. Following are the function and the working of each pin

a) The IP+ and IP- pin is used to measure the measure the positive and negative current input from the AC current carrying wire

b) VCC- This pin powers the sensors to function properly

c) GND- The grounding pin for the sensor

d) RSENSE- The pin is used to set the sensitivity of the sensor. The more the sensitivity the lesser the resistance of the RSENSE resistor.

e) BW_SEL- The BW_SEL is the pin used to select the bandwidth of the sensor that means the frequency range for the current that it can measure accurately.

f) OUT_EN- This pin is used to enable or disable the output of the sensor.

g) FAULT- The FAULT pin outputs the high voltage if there is a fault in the sensor letting the user knows that the sensor is damaged.

The voltage value which is obtained is given as input via the jumper wires to the micro controller which contains the K means algorithm which is used for the near to accurate prediction.

Figure 1.Acs 723 Sensor

3.2 Data cleaning and conversion

The data which is obtained will contain the value of the voltage upon reading the value of current will then be cleaned and made it into a form which is understandable by the machine learning algorithm. To pre-process the data, cleaning need to be done by removing the noise and normalizing the data as per the scale of the trained data. The data is the converted by the help of NumPy array or a Pandas Data Frame.

3.3 Data Prediction

For the process of execution, one algorithm must be chosen from the lot of different machine learning algorithms. The dataset that has been used is an unlabeled dataset and as a result we cannot apply many of the machine learning techniques which are based on supervised learning. So, in order to facilitate the process of choosing the algorithm the K means clustering and Hierarchical clustering are available. The Hierarchical clustering is mostly used when the underlying dataset is in a dendrogram format and we do not know the number of clusters beforehand to be made. The Hierarchical clustering is mostly used for visualizing the clustering process because it gives a tree like structure for better visualization. Here in this case, the number of clusters to be made are 3 which is known from beforehand and the dataset on which it is to be operated is large, leading to the increased number of values in that scenario the K means algorithm works perfectly.

The working of the K means algorithm is started by using the elbow method find the appropriate number of clusters. After assigning the clustering centroids, find the Euclidian appropriate distances between the data points and the centroid. After finding the appropriate data points and assigning the clusters of the data points, recompute the data points again and again by calculating the mean of the data points. After assigning the clusters to the data points, check for the convergence. If the convergence doesn’t change then pick out the final centroids for the respective data point clusters and plot the output.

While doing this algorithm intelligent centroid allocation technique is used by assigning the in it value of the K-means function to “k-means ++” to automatically change it form the in it value of 10 to auto. After finding and training the model on the dataset the clustering points are displayed in the graphical format and on the basis of the input prediction is done. The following is the systematic diagram of the Methodology:

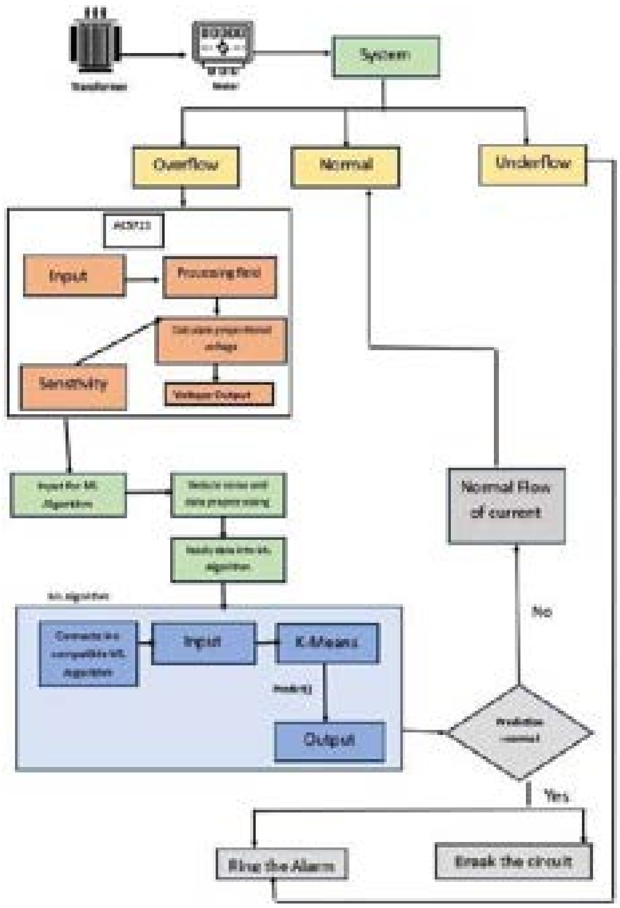


Figure2. Workflow

After doing the prediction the voltage is calculated on the basis of the intervals of the if-else and else if ladder structure and as a result the final output is produced. Here each level of voltage is assigned a numeric value so that the model can train accurately without displaying the result in the form of string here in this scenario the numeric digit “2” represents abnormal voltage, numeric digit “1” represents normal voltage and numeric digit “0” represents abnormal voltage as a result the accuracy of the model is also increased abruptly. So the accuracy increases abruptly leading to a good prediction of the model

```
[2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2
 2 2 2 2 2 2 2 2 2 2 2 2 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1
 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1 0
 1 0 1 0 1 0 1 0 1 0 1 0 1 0 1]
```

Figure 3. Output analysis

Table 2. Significance of voltage clustering

Number	Representing value
0	Low voltage
1	Normal voltage
2	Abnormal voltage

IV. Result

At the beginning the number of clusters are to be calculated with the help of the elbow method. Elbow method is a very well-known method which is used to find the number of clusters for the particular dataset. While implementing the elbow method we will find the number of clusters to be needed to be implemented for the best training of our model.

There are many ways to find the number of clusters such as the Silhouette method, Cross validation method in the k means algorithm but the most appropriate and the innovative way by using the elbow method which gets its name because of the elbow like structure that is formed in the graph giving the actual number of clusters in the k means algorithm to be used for appropriate results. To generate the elbow figure, WCSS (within cluster sum of squares) is needed to be found by applying the algorithm of K means++ over the dataset. In the algorithm the number of clusters are declared along with the model of k means++ which is to be followed. For each iteration a particular number of clusters is used to train on the dataset and then the elbow figure is formed. The WCSS is found with the help of the below Equation (1) is: WCSS = $\sum_{Pi \text{ on cluster1}}$ distance(Pi,C1)^2 + $\sum_{Pi \text{ on cluster2}}$ distance(Pi,C2)^2…..

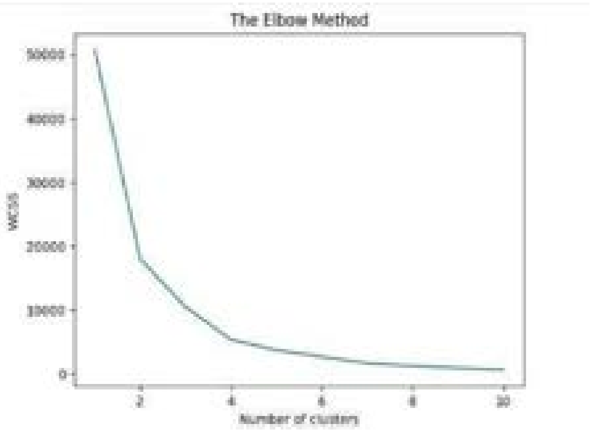


Figure 4. Elbow method

The formation of different clusters in the backend, the elbow bends down leading for the user to decide the optimal number of clusters. By seeing it through the eye and giving a rough estimation the point where the graph starts bending, that point is considered to be the point for optimal number of clusters. After applying the elbow method, we find out that the required number of clusters for our specific algorithm is 3. The elbow method also helps to keep as assured that the clustering number is going to be safe and will be used in a proper manner. After the number of clusters is decided the clustering is done on the basis of K-means algorithm. The k-means algorithm takes the number of clusters as 3 as input and also the value of in it as “k-means++” for intelligently assigning of the centroid for the clustered graph.

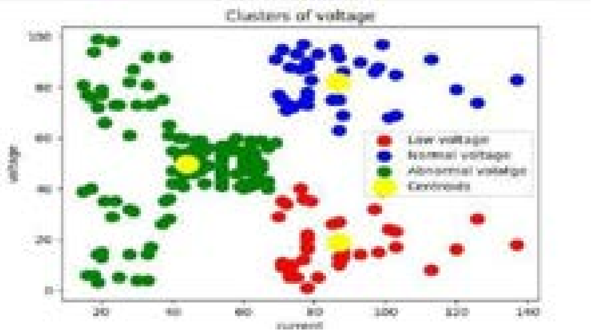


Figure 5. Cluster Graph showing the voltage points

The figure above shows the clustering of the points of the voltage of the based on the data points as input. The three colors show the three different types of clusters that are formed prominently. The green cluster the abnormal voltage or high voltage which means the value of the input is greater than the value that is provided in our household. The blue cluster shows the normal voltage which means the value of the voltage provide by is same as the measured voltage by the sensor accurately. The red voltage shows the low voltage which means the voltage provided is lesser than the voltage that is received at the meter of the house. The yellow dots represent the centroid which means the approximate center point of the clusters. The yellow dots are derived at the beginning with the help the previous step of elbow method. From the elbow method we concluded that the number of clusters to be used are 3. So, three dots have been produced. The dots are changed at each step depending on the input. The K means algorithm is then used to predict the accurate value of the voltage which is taken as an input. After predicting the value, it then checks against the normal range of the value. After the prediction 3 cases might arise which are described below.

If the predicted value is greater than the normal range of the values then the circuit will be broken instantly and then alarm will be triggered. If the predicted value is same as the normal value, then nothing will happen normal current flow will occur If the predicted value is lesser than the normal range of value then the alarm will be triggered but no circuit breaking will be there as in most cases underflow of current in general does not hamper the appliances nor has high risk of catching fire After the clustering is done then comes the part for prediction as in the algorithm will be able to predict the value of the model accurately or not for the unknown data. To see the prediction for the unknown data we need to input two values one for the current and their respective voltage as the dataset is trained on a 2-D array. The values which are then input for the prediction will be placed on the clustering algorithm which will then plotted on the graph by calculating their Euclidean distance and checking their similarities and then plotting current vs voltage graph. To understand the working of the model for the unknown values.

Table 3. Output on Test values

Test Value	Output
[6,228]	Normal Voltage
[2,220]	Low Voltage
[14,240]	Abnormal Voltage

To test the following process the values of current are given and the respective voltage is converted giving rise to the final output of the voltage condition. In the above table prediction is done on 6A current for which the voltage at a particular timestamp it is showing as 228V. Now as per Indian household the value if normal voltage is between 220-230V so it is showing correct values. After that, prediction is done on 2A current for which the voltage at a particular timestamp it is showing as 200V. Now as per Indian household the value if normal voltage is between 220-230V so it is showing low voltage which is correct values. Furthermore, prediction is done on 14A current for which the voltage at a particular timestamp it is showing as 240V. Now as per Indian household the value if normal voltage is between 220-230V so it is showing abnormally high voltage which is correct values.

V. CONCLUSION

The process and the steps highlighted above in the paper presents a fresh new perspective to handle different conditions associated with anomaly voltage detection. The proposed system is designed to identify voltages which are exceedingly above 230Volts to trigger the alarm and break the circuit to avoid potential electric issues. Compared to the conventional methods provides a number of benefits. First the proposed solution is more robust to changes in the power grid system because it is based in machine learning. Second, the proposed system is simpler to maintain and modify because no specified rules are needed. Third, the suggested metho makes advantage of the high precision low cost ACS723 sensor. This approach can help electricity systems become more dependable and effective which can aid in preventing equipment failure and blackouts by early detection of anomalous voltage patterns.

To achieve reliability, the position of the placement of the product is a must thing. The product must be placed on the wire which is connected from the transformer to the mail building or at the meter of the house which it gets current. The connection of the product to the alarm is the most superior thing which can be achieved superiorly if the usage of the wire to the alarm system is less which means the placement of the product must be inside the house and after the connection with the meter.

VI. FUTURE WORK

In future prospects, researchers may look upon the techniques of the supervised learning models where performance and accuracy increase a lot. The accuracy increases when the measured abnormal voltage and the current waveforms are initialized in a stable format. Multiple supervised learning models are to be used abruptlyto develop and infer different outcomes. Many models such as SVM and neural networks can be used abruptly to leverage the idea of abnormal voltage detection and give the world a safer environment.

To achieve high accuracy in future this work, somepoints might be kept in mind. In the place of preprocessing the data obtained from the sensor needs to be normalized and handling of missing values must be done properly. Choosing of relevant features that may add noise to the clustering process should be eliminated. Running of the algorithm multiple times so that the clustering technique results in low distortion.

To improve the process, incorporating wireless connectivity features, such as Wi-Fi or Bluetooth, into the hardware design for ringing the alarm can be done. Moreover, it can be placed in with existing smart home automation systems and industrial control systems. Integration with popular platforms and protocols, such as Zigbee or MQTT, can facilitate seamless communication and interoperability with other smart devices and applications.

This technique can be used in EV's to detect any anomalyin the flow of current resulting in the safeguarding in the situation of the EV's. The technique as a result protecting the lifeline of the car and giving it a safer experience to the user.

The extension of this work can be applied to the modern-day era of the industries and factories where the supplement of the voltage is very high. The chances of the anomalous flow of voltage is high in those areas, as a result the mentioned techniques can be implemented to avoid extra high voltage to the industry and avoid the corruption of industry equipment's and keep it durable and reliable.

VII. REFERENCES

[1] Geon Yong Park, Heeseong Kim, Hwi Woon Jeong, and Hee Yong Youn, "A Novel Cluster Head Selection Method based on K-Means Algorithm for Energy Efficient Wireless Sensor Network", 2013 27th International Conference on Advanced Information Networking and Applications Workshops, 20130325.

[2] Murty MN and Devi VS 2011 Pattern recognition, An algorithmic approach (London: Springer)

[3] Hodge V H and Austin Jim 2004 A survey of outlier detection methodologies J. Artificial Intelligence Review22 85-126

[4] B. Plumer, "Bad news: The U.S. power grid is getting pricier, less reliable," The Washington Post, March 8 2013. [Online],

[5] A. Smith, N. Lott, T. Houston, K. Shein, J. Crouch, J. Enloe, "U.S. Billion-Dollar Weather & Climate Disasters 1980-2019," NOAA National Centers for Environmental Information, [Online]

[6] S. Chakraborty, M. Lukasiewycz, C. Buckl, S. Fahmy, NChang, S. Park, Y. Kim, P. Leteinturier, and H. Adlkofer, "Embedded systems and software challenges in electric vehicles", in Proceedings of the conference on design, automation and test in Europe, EDA Consortium, 2012.

[7] S. Shanker, "Enhancing automotive embedded systems with FPGAs", PhD thesis, Nanyang Technological University, 2016.

[8] Ahn, B.-H., Choi, H.-R., and Lee, H.-C. (2015). RegionlLong-Term/midterm Load Forecasting Using Sarimain south korea. J. Korea AcademiaIndustrial cooperationSoc. 16,8576–8584. doi:10.5762/ kais.2015.16.12.8576

[9] Mohri, M.; Rostamizadeh, A.; Talwalkar, A. Foundations of Machine Learning; MIT Press: Cambridge, MA, USA, 2012. [Google Scholar]

[10] Bishop, C.M. Neural Networks for Pattern Recognition; Oxford University Press: Oxford, UK, 1995. [Google Scholar]

[11] Afonso, J. A., Rodrigues, F., Pereira, P., Gonçalves, H., & Afonso, J. L. (2015). Wireless monitoring and management of energy consumption and power quality events. Lecture Notes in Engineering and ComputerScience, 2217, 338–343. Ahmed, M. S., Mohamed, A., Homod, R. Z.

[12] Afonso, J. A., Rodrigues, F., Pereira, P., Gonçalves, H., & Afonso, J. L. (2015). Wireless monitoring and management of energy consumption and power quality events. Lecture Notes in Engineering and Computer Science, 2217, 338–343.

[13] de Sousa Jr, Rogerio R., et al. "Electrical Conductivity and In Situ SAXS Probing of Block Copolymer Nanocomposites Under Mechanical Stretching." *ACS Applied Materials & Interfaces* (2023).

[14] Obaid, Sofian N., et al. "Design and Fabrication of a Flexible Opto-Electric Biointerface for Multimodal Optical Fluorescence and Electrical Recording." *ACS Applied Electronic Materials* 5.3 (2023): 1688-1696.

[15] Li, Junyu, et al. "Direct-Laser-Writing Manipulating α/δ-FAPbI3 Array Enables Low-Crosstalk Terahertz Sensing and Imaging." *ACS Photonics* 10.8 (2023): 2755-2764.

Enhancing Diagnosis Precision in Alcohol Addiction Detection Through CNN Analysis with SMOTE-ENN Data Augmentation

Mohamed Ashif H
Department of Data Science and Cybersecurity
Karunya Institute of Technology and Sciences
Coimbatore, India
mohamedashif@karunya.edu.in

Dr. E. Grace Mary Kanaga
Department of Data Science and Cybersecurity
Karunya Institute of Technology and Sciences
Coimbatore, India
grace@karunya.edu

***Abstract*—Alcoholism is a major concern in modern culture, demanding effective and prompt detection measures to limit its negative consequences. Electroencephalography (EEG) data processing has emerged as a viable option for detecting alcoholism by utilizing the brain's unique electrical patterns. However, EEG signals are complicated and multi-channel, deciphering them is difficult. Previous research has primarily relied on traditional machine learning and statistical methodologies, frequently including handcrafted features to identify subjects as alcoholic or non-alcoholic. Deep learning models have gained enthusiasm as computer capabilities and data volumes have increased, suggesting prospective options to improve alcoholism diagnosis precision. This study looks into the efficacy of machine learning algorithms like Naive Bayes, Random Forest, and KNN, as well as deep learning algorithms like CNN and LSTM, when combined with data augmentation techniques to improve the performance of alcohol addiction detection. The research uses two independent datasets - one publicly available and one acquired - and performs rigorous preprocessing approaches, including the elimination of NaN values and duplicates, followed by normalization. Three data augmentation strategies, SMOTE, SMOTE-ENN, and SMOTE-TOMEK, are used to address class imbalances in the datasets. The results of a comprehensive research show significant advances in alcohol addiction detection. In the public dataset, the Random Forest method obtains 96.22% accuracy before augmentation, which rises to 98.67% after augmentation with SMOTE-ENN. Similarly, in the obtained dataset, the CNN model obtains 96% accuracy before augmentation, which increases to 98.97% after augmentation with SMOTE-ENN. These findings highlight the potential of deep learning models, when combined with good data augmentation methodologies, to greatly improve the accuracy and reliability of alcohol addiction diagnosis, providing vital insights for future research and clinical applications.**

Keywords—Alcohol Addiction, EEG, Machine Learning Algorithms, Deep Learning Algorithms, Data Augmentation Techniques, SMOTE, SMOTE-ENN, SMOTE-TOMEK, Naive Bayes, Random Forest, KNN, CNN, LSTM

I. INTRODUCTION

Alcoholism is a serious problem in today's culture that has an adverse impact on one's physical and emotional health. Alcoholism has been shown to cause cognitive deficits that have a significant negative influence on both brain function and general health. Abnormal alcohol intake over time not only damages brain circuits but also impairs the health of critical organs including the liver, pancreas, and heart. Alcohol misuse also worsens human health because it is associated with a number of illnesses, such as diabetes and cancer. Alcoholism not only has physiological effects but also damages memory and decision-making skills, which over time causes a major reduction in cognitive function. Impaired coordination and mood swings are two short-term effects of alcohol drinking that highlight the widespread prevalence of this condition and its profound effects on both people and society as a whole.

Alcoholism is difficult to diagnose because there are no commonly used, standardized tests for the condition. Conventional techniques, such as physical examinations and scent detection, can provide some information, but they are frequently inaccurate and unreliable. The use of electroencephalography (EEG), which uses the electrical activity of the brain to identify underlying patterns linked to alcohol misuse, presents a promising method for the identification of alcoholism. However, due to their complexity and noise sensitivity, EEG signals are difficult to analyze manually, which makes the development of automated systems necessary for precise and effective analysis.

In the previous research, machine learning and statistical methods have been the main tools used in alcoholism detection studies to classify EEG signals and identify alcohol-affected patients. Although these methods are effective, there hasn't been much research done on applying deep learning models in this field. The ability of deep learning models to identify complex patterns in EEG data is a viable path for increasing the diagnostic accuracy of alcohol addiction detection. The existence of data imbalance within the EEG datasets, however, presents a substantial barrier in this exploration.

To fill this essential gap and improve the efficacy of alcoholism diagnosis, the current work proposes a novel approach that combines machine learning algorithms such as Naive Bayes, Random Forest, KNN, and deep learning algorithms such as CNN, LSTM, with data augmentation approaches. To address data imbalance, the study applies a variety of data augmentation approaches, including SMOTE, SMOTE-ENN, and SMOTE-TOMEK. The purpose of this research is to generate a more accurate estimation of alcoholic and non-alcoholic subjects by enriching the dataset with machine learning and deep learning techniques. This comprehensive approach aims to improve the accuracy and reliability of alcohol addiction diagnosis, leading to more effective interventions and treatment methodologies.

The primary objective of this study is to provide a solid and trustworthy method of utilizing EEG data to detect alcohol addiction. The major goal is to improve the reliability and accuracy of alcohol addiction diagnosis by utilizing machine learning and deep learning techniques in combination with data augmentation methodologies. By means of examining two discrete datasets — the publicly available dataset and the acquired dataset— the effectiveness of several algorithms and augmentation methods in precisely categorizing individuals as either alcoholic or non-alcoholic will be examined. The ultimate objective is to improve public health outcomes by advancing alcohol addiction diagnostic tools and interventions.

II. Related Works

Frontiers, et.al in [1] specifically look into how data augmentation approaches might be used to solve problems with feature extraction from non-stationary and non-linear EEG signals.

In a study [2] the authors emphasize the value of customized data augmentation techniques, with implications for motor imagery brain-computer interfaces and sleep stage classification.

In low-data circumstances, Joseph Paillard, et. al, in the study [3] shows that suitable augmentations can improve accuracy by up to 45%. It does this by carefully evaluating and comparing 13 different data augmentation strategies for EEG categorization.

To overcome the limitations of the EEG data, in the study [4] the authors Mahsa Pourhosein Kalashami, et.al suggest a technique that uses deep generative models, specifically Conditional Wasserstein GAN (CWGAN), for data augmentation.

Yun Luo, et.al in the [5] article proposes three augmentation methods using deep generative models (VAE and GAN) and two strategies (full and partial utilization) to solve data scarcity in EEG-based emotion recognition.

Siqi Sun, et.al suggest a deep learning approach for the quantitative diagnosis of Internet addiction (IA) that combines a rapid Fourier transform with a CNN model with five layers. The suggested approach shows promise for in-the-moment health monitoring systems and goes beyond IA to diagnose conditions impacting EEG readings [6].

Amel Kisibi, et.al discusses difficulties in diagnosing depression using EEG data, highlighting the impact of demographic variables. With the use of deep learning and machine learning on the MODMA dataset, which included EEG data from wearable and conventional sources, the program was able to identify major depressive disorder (MDD) and other related conditions with 97% accuracy [7].

Fernado, et.al uses Data Augmentation (DA) to build synthetic Photoparoxysmal Response (PPR) segments from genuine ones, addressing imbalanced datasets in photosensitivity diagnosis. The goal of the study is to improve automatic PPR detection for neurological illness diagnosis at Burgos University Hospital in Spain [8].

Using EEG signals and questionnaires, suggests a deep learning system for diagnosing Internet addiction (IA), with the goal of quantitative assessment without the identification of particular biomarkers. The Lemon dataset demonstrates the higher performance of a five-layer CNN model with fast Fourier transform in detecting IA without the need for biomarkers, with 87.59% accuracy, 88.80% sensitivity, and 86.41% specificity [9].

To improve brain-computer interface (BCI) motor imagery (MI) decoding through deep learning, Roger smith, et.al investigate data augmentation strategies. Using six different techniques, MI electroencephalography (EEG) trials were synthesized, leading to mean accuracy gains of up to 3% and 12% in two available datasets. These results point to possible gains in prediction accuracy as well as time savings for gathering subject data for BCI applications [10].

A thorough analysis of machine learning (ML) approaches for treating addiction disorders is provided by Bouhadja et al.; the evaluation covers both non-substance addiction (N =3) and substance addiction (N =18) research. The goal of the review is to present compelling data to support the continued application of machine-learning techniques to problems relating to addiction [11].

To reduce the likelihood of substance addiction among teenagers and young people, Roy, A. et al. suggest an automated technique for identifying social media posts about substance use [12].

In this work [13], Maslej-Krešňáková, V. et al. examine how well Easy Data Augmentation (EDA) strategies handle class imbalance in the identification of antisocial conduct, with a particular focus on the classification of toxic comments and fake news.

In this [14] paper, Wang, S.H. et al. suggest a computer vision-based method that uses Convolutional Neural Networks (CNNs) to identify Alcohol Use Disorder (AUD) from brain pictures. The method achieves high accuracy (97.04%), specificity (97.18%), and sensitivity (96.88%).

In this study, Rekabdar, B. et al. analyze data from a cohort of 6978 individuals in Alabama using machine learning and deep learning models to identify dangerous alcohol and drug misuse in an SBIRT program [15].

III. Dataset Description

The dataset used in this study is divided into two sets public datasets and acquired datasets. These datasets serve as the foundation for investigating alcohol addiction using EEG signals, providing extensive insights into the brain activity patterns related to addiction.

A. Public Dataset

The public dataset includes EEG recordings from 122 subjects, each with 64 electrodes carefully placed over their scalps. These EEG waves were captured at 256 Hz, with 3.9 millisecond epochs for a total period of 1 second. Subjects in this dataset were divided into two categories: alcoholic and control. Among the 122 subjects, 90 were classed as non-alcoholic, whereas 32 were identified as alcoholic. This dataset represents a wide range of EEG signals collected in real-world circumstances, providing insights into the brain activity linked with alcohol addiction.

B. Acquired Dataset

The acquired dataset includes EEG data from seven subjects with sixteen electrodes carefully positioned throughout their scalps. The acquired data's EEG signals were sampled for one second at a frequency of 256 Hz, just like the public dataset. Two alcoholic subjects and five non-alcoholic subjects were included in the dataset. These subjects were then split into control and alcoholic groups. For the collected data, the electrode properties include FP2, F4, C4, P4, F8, T4, T6, O2, FP1, F3, C3, P3, F7, T3, T5, and O1 as shown in Fig.1, offering spatial data that is essential to comprehending the brain correlates of alcoholism.

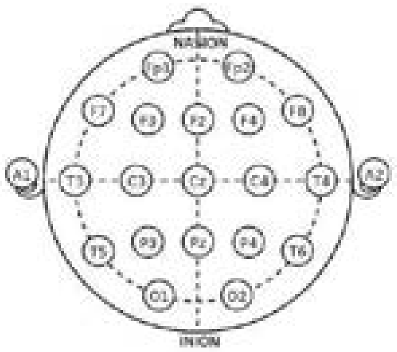


Fig. 1.EEG Electrode placement

These datasets serve as the foundation for developing and assessing machine learning and deep learning models for diagnosing alcohol addiction using EEG signals. The comprehensive nature of the datasets, which include both public and acquired data, allows for thorough evaluation and validation of the algorithms. Using these extensive datasets, the study aims to increase the understanding of alcohol addiction and lead to the development of more effective diagnostic tools and interventions.

IV. PROPOSED METHODOLOGY

Alcoholism detection using EEG data includes data preprocessing, such as the removal of NAN values, duplication, and normalization, as well as a comparison of augmentation approaches such as SMOTE, SMOTE-ENN, and SMOTE Tomek to solve class imbalance. Various machine-learning and deep-learning algorithms, such as Naive Bayes, Random Forest, KNN, CNN, and LSTM, are used for thorough analysis, as shown in Fig.2. These strategies are used to categorize the person as alcoholic or non-alcoholic, allowing for accurate identification and diagnosis of alcohol addiction.

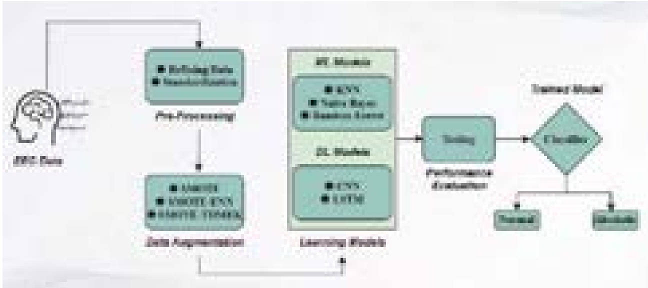


Fig. 2. Workflow of EEG based addiction detection

A. Data Pre-Processing

Data pre-processing is critical to ensuring the integrity and quality of the EEG datasets used to detect alcohol addiction. To accomplish this, a number of procedures were done to refine the raw data before model training and evaluation. Initially, the data underwent preprocessing operations to address missing values and duplicates. Any instances with missing values (NAN values) were detected and removed from the datasets to avoid potential biases or mistakes in the study. Duplicate items in the datasets were also discovered and removed to ensure that the data samples were unique and consistent.

The EEG data were standardized after NAN values and duplicates were eliminated in order to guarantee consistency and comparability among various features and samples. Scaling the features of the EEG data to a standard scale, usually with a mean of zero and a standard deviation of one, is the process of standardization, often referred to as normalization. By reducing the impact of different scales and magnitudes found in the raw EEG data, this procedure enhances the performance and stability of machine learning and deep learning models used to identify alcohol addiction.

In this study, 20% of the dataset was used for testing. This selection was influenced by traditional machine learning procedures, such as a common split ratio of 80:20 between training and testing sets, which guarantees robust evaluation while still giving adequate training data. Given the magnitude of the datasets and the need for accurate performance assessment, this split percentage was deemed reasonable. Specifically, in the public dataset of 122 subjects, 80% (about 97 subjects) were used for training, while the remaining 20% (roughly 25 subjects) were set aside for testing. Similarly, in the acquired dataset of 7 subjects, 80% (roughly 5 subjects) were selected for training, with the remaining 20% (about 2 subjects) for testing.

The research guarantees that the EEG datasets are free of any biases, inconsistencies, or abnormalities that can endanger the accuracy and dependability of the analysis that follows by carrying out these data pre-processing procedures. The successful application of machine learning and deep learning algorithms is made possible by this methodical approach to data refining, which also improves the robustness and effectiveness of alcohol addiction detection.

B. Data Augmentation Techniques

1) SMOTE

The Synthetic Minority Over-Sampling Technique (SMOTE) is an important approach for resolving the imbalance between alcoholic and non-alcoholic classes in the context of alcoholism identification utilizing EEG data. Class imbalance, in which one class (alcoholic subjects in this case) is underrepresented relative to the other (non-alcoholic subjects), is a major challenge in machine learning and deep learning tasks. By creating synthetic instances of the minority class, SMOTE explicitly addresses this problem and evens out the distribution of classes in the dataset.

In order to solve the issue of class imbalance, SMOTE creates artificial instances that fill in the gaps in the feature space between samples of minority classes. Consider the feature space as a multidimensional terrain in which every data point is a distinct set of features taken from the EEG signals. Minority class instances are frequently sparsely dispersed in this environment, which can make it difficult for the classifier to identify patterns in certain areas.

To address this, SMOTE locates the closest neighbors of each minority class instance in the feature space. In order to create new data points that bridge the gaps between the minority class samples that already exist, it then synthesizes new instances along the line segments linking these neighbors. SMOTE increases the minority class's representation in this way, bringing its size closer to that of the dominant class. An enhanced dataset with a more evenly distributed class distribution will be the outcome of this. With a wider range of samples from both classes at their disposal, machine learning and deep learning models may now learn more efficiently. These models can more effectively generalize and produce more accurate predictions if they have a more thorough understanding of both the majority and minority classes.

2) SMOTE-ENN

A thorough method to address class imbalance in datasets is to combine SMOTE with Edited Nearest Neighbors (SMOTE-ENN). ENN concentrates on eliminating noisy or unnecessary data, whereas SMOTE oversamples the minority class by creating synthetic examples. Through the integration of these techniques, SMOTE-ENN seeks to improve the dataset's balance, making it more appropriate for training machine learning and deep learning models by eliminating examples that may have been incorrectly classified and by generating artificial instances.

The combination of SMOTE and ENN improves the quality of synthetic instances, which in turn improves the model's capacity to generalize and produce precise predictions on datasets that are unbalanced. By ensuring that the dataset is suitable for model training, this combination strategy not only lessens the consequences of class imbalance but also produces more robust and dependable performance in classification tasks.

3) SMOTE-TOMEK

SMOTE-Tomek is a resampling algorithm that combines SMOTE and Tomek linkages to overcome the class imbalance in datasets. SMOTE-Tomek seeks to improve dataset clarity by locating and managing pairings of instances from various classes that are near to one another and maybe noisy. To do this, the minority class is oversampled using SMOTE, and instances implicated in Tomek links are simultaneously removed. Improved class separation in the resultant dataset contributes to better model generalization and classification performance, especially in situations with unequal class distributions.

By removing unresolved examples SMOTE-Tomek maximizes the effectiveness of synthetic instance production through the combination of Tomek connections and SMOTE. This method guarantees the production of a more refined dataset that is advantageous for efficient model training in addition to mitigating class imbalance. SMOTE-Tomek offers a solid method for boosting machine learning and deep learning models' performance in classification tasks on unbalanced datasets by raising the quality of synthetic instances while reducing the impact of noisy or ambiguous data points.

C. Machine Learning Techniques

1) Naive Bayes

Naive Bayes is a sophisticated probabilistic algorithm that is simple to use and relies on the assumption of feature independence and Bayes' theorem. Naive Bayes is a simple algorithm that frequently performs better than expected, especially when dealing with high-dimensional data sets like EEG signals. This approach multiplies the individual probabilities of each characteristic occurring in a class to get the likelihood of that class given a set of features. The output is then normalized. It is a desirable option due to its ease of use, computational effectiveness, and low training data needs.

To solve class imbalance, resampling approaches such as SMOTE, SMOTE-ENN, and SMOTE-Tomek are strategically combined with Naive Bayes in this study. Naive Bayes can effectively classify cases and contribute to improved classification performance by using these techniques to oversample the minority class and refine the dataset, particularly in scenarios where class distributions are uneven.

2) Random Forest

During training, Random Forest builds numerous decision trees and then uses a voting mechanism to combine the predictions of these trees. It is well-known for its adaptability and efficiency in classification and regression tasks. Overfitting is reduced and generalization performance is improved using this ensemble learning strategy, which is why it is favored in many fields.

In this work, class imbalance is addressed by harmonizing Random Forest with resampling techniques including SMOTE, SMOTE-ENN, and SMOTE-Tomek. Random Forest is capable of efficiently classifying cases by either improving the dataset by removing confusing instances using SMOTE-Tomek or oversampling the minority class using SMOTE and SMOTE-ENN. In situations when there are unequal class distributions, this integration not only increases the robustness of the model but also guarantees better representation of minority classes, improving classification performance.

3) K-Neighbors Classifier

Based on the similarity principle, the K-Neighbors Classifier (KNN) is a simple yet powerful technique for classification and regression applications. KNN produces choices without making any assumptions about the underlying data distribution by choosing the most common class among a data point's K nearest neighbors. This instance-based, non-parametric approach works well with both linear and non-linear decision boundaries, which makes it appropriate for a range of uses, such as the identification of addiction using electroencephalograms (EEG).

To solve class imbalance, this work deliberately combines the KNN algorithm with resampling approaches such as SMOTE, SMOTE-ENN, and SMOTE-Tomek. The balanced data allows KNN to classify instances well, whether by using SMOTE-Tomek to refine the dataset or SMOTE and SMOTE-ENN to oversample the minority class. Improved prediction resilience and improved representation of minority classes are the results of this synergy, which also improves classification performance particularly in situations when there are unequal class distributions.

D. Deep Learning Techniques

1) CNN

Convolutional neural networks (CNNs) are a key component of the project's methodology, specifically intended to solve the issues associated with EEG data interpretation. CNNs are designed specifically to process structured inputs, such as the grid-like nature of EEG data, and include convolutional, pooling, and fully connected layers. CNNs use convolutional filters to extract hierarchical features from raw data, allowing them to identify complicated patterns and correlations in EEG signals.

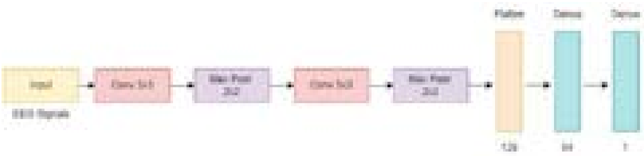


Fig. 3. CNN Architecture

CNNs are used in conjunction with resampling methods such as SMOTE, SMOTE-ENN, and SMOTE-Tomek for two reasons in this research. First of all, it uses CNNs' deep learning capabilities to extract features and classify EEG data as shown in Fig. 3. Secondly, it addresses the problem of class imbalance by using SMOTE-Tomek to refine the dataset or SMOTE and SMOTE-ENN to oversample the minority class. The model can more effectively capture significant patterns and correlations by training CNNs on the modified dataset with improved class distributions, guaranteeing increased representation of minority classes. This deliberate combination ensures more accurate predictions in the identification of alcohol addiction, especially when dealing with the imbalanced class distributions present in EEG datasets.

2) LSTM

Long Short-Term Memory (LSTM) networks play a critical role in the proposed methodology, allowing for the detailed interpretation of sequential EEG data. Designed to overcome the vanishing gradient problem inherent in traditional recurrent neural networks (RNNs), LSTMs specialize in detecting and exploiting long-term relationships within sequential inputs. LSTMs, which have different gates such as input, forget, and output gates, allow for controlled information flow while preserving temporal information through memory cells, making them ideal for tasks requiring sequential data analysis, such as EEG signal processing.

The integration of LSTMs with resampling techniques such as SMOTE, SMOTE-ENN, and SMOTE-Tomek fulfills two purposes. In the beginning, it makes use of recurrent neural networks' innate capacity for sequential data processing. Furthermore, it tackles the problem of class imbalance by using SMOTE-Tomek to refine the dataset or SMOTE and SMOTE-ENN to enhance the representation of minority classes. The modified dataset is then fed into the Long Short-Term Memory (LSTM) network, which allows it to better detect patterns and temporal relationships among the evenly distributed class memberships. This well-coordinated strategy improves the model's ability to identify intricate temporal patterns in EEG data, supporting classification performance, especially in situations with unequal class distributions.

V. Results and Discussion

In the public dataset, before augmentation, Random Forest exhibited the highest accuracy of 96.22%, followed closely by KNN with an accuracy of 95.45%. Naive Bayes lagged behind with an accuracy of 83%. After augmentation, the performance of all algorithms improved noticeably across the board. Particularly noteworthy was the significant enhancement observed in Naive Bayes accuracy, which increased from 83% to 84.74% with SMOTE augmentation. Random Forest maintained its lead, achieving an accuracy of 95.42% with SMOTE augmentation, while KNN showed the most considerable improvement, reaching an accuracy of 98.67% with SMOTE-ENN augmentation.

Conversely, in the acquired dataset, Naive Bayes started with the highest accuracy of 85.14%, followed by KNN with an accuracy of 95.23%. After augmentation, the accuracy of all algorithms exhibited improvement. Notably, Naive Bayes achieved an accuracy of 89.06% with SMOTE augmentation, representing a substantial improvement from its initial accuracy. Random Forest maintained its competitive edge, reaching an accuracy of 94.34% with SMOTE augmentation, while CNN experienced the most significant improvement, achieving an accuracy of 98.97% with SMOTE-ENN augmentation.

These results underscore the importance of data augmentation techniques in enhancing classification performance, especially for algorithms initially hindered by imbalanced or noisy datasets. SMOTE-ENN technique effectively addressed class imbalance and noise, resulting in improved accuracy when combined with CNN algorithm.

Furthermore, the augmentation strategies performed consistently across both datasets, demonstrating their robustness and usefulness in a variety of EEG data settings. The observed gains demonstrate the potential of these methods to address the issues associated with imbalanced and noisy datasets, resulting in more accurate and reliable alcohol addiction identification. This consistency emphasizes the adaptability of the augmentation strategies, indicating that they can adjust to changing data distributions and features.

In addition, the augmentation procedures not only corrected class imbalance, but also served to mitigate the impact of noise in the EEG data. SMOTE, SMOTE-ENN, and SMOTE-TOMEK significantly improved the generalization capabilities of machine learning and deep learning models by producing synthetic data and balancing minority class distributions.

TABLE 1. EEG Database – Results

Technique	**Algorithm**	**Accuracy**
Before-Augmentation	Naive-Bayes	83.00
	Random Forest	96.22
	KNN	95.45
	CNN	90.08
	LSTM	94.21
SMOTE	Naive-Bayes	84.74
	Random Forest	95.42
	KNN	95.98
	CNN	92.23
	LSTM	95.82
SMOTE-ENN	Naive-Bayes	82.71
	Random Forest	93.62
	KNN	98.67
	CNN	96.26
	LSTM	96.28
SMOTE-TOMEK	Naive-Bayes	84.60
	Random Forest	94.42
	KNN	94.47
	CNN	96.15
	LSTM	95.87

Table 1 shows the results of a public dataset of all the machine learning and deep learning algorithms before and after using different augmentation techniques.

TABLE 2. Acquired Dataset - Results

Technique	**Algorithm**	**Accuracy**
Before-Augmentation	Naive-Bayes	85.14
	Random Forest	94.00
	KNN	95.23
	CNN	96.00
	LSTM	95.11
SMOTE	Naive-Bayes	89.06
	Random Forest	94.34
	KNN	96.43
	CNN	97.11
	LSTM	91.29
SMOTE-ENN	Naive-Bayes	88.08
	Random Forest	95.00
	KNN	95.98
	CNN	98.97
	LSTM	94.68
SMOTE-TOMEK	Naive-Bayes	86.08
	Random Forest	92.00
	KNN	93.98
	CNN	96.24
	LSTM	94.59

Table 2 shows the results of an acquired dataset of all the machine learning and deep learning algorithms before and after using different augmentation techniques like SMOTE, SMOTE-ENN, SMOTE-Tomek.

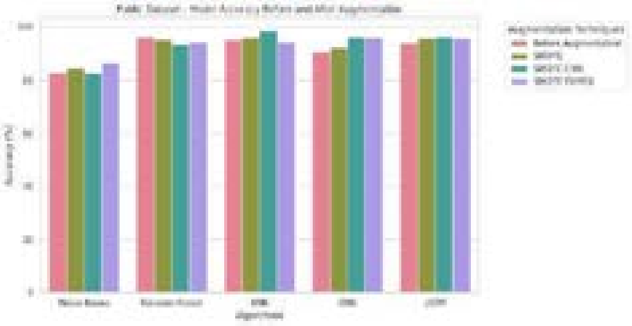


Fig. 4. EEG Database – Results

In Fig.4, it is clearly visualized that while training with public dataset the accuracy of each model's before augmentation and also accuracy of each model's after applying different augmentation techniques like SMOTE, SMOTE-ENN, SMOTE-Tomek.

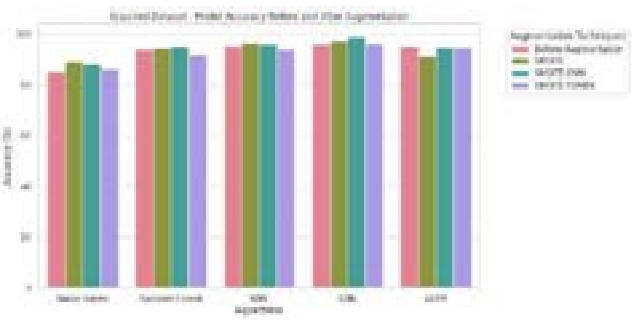


Fig. 5. Acquired Dataset – Results

In Fig.5, the accuracy of each model before and after the use of various augmentation approaches, such as SMOTE, SMOTE-ENN, and SMOTE-Tomek, is easily observable when training with the acquired dataset.

VI. Conclusion And Future Scope

The research emphasizes how important data augmentation methods and algorithm selection are important to EEG-based addiction detection. Prior to augmentation, the accuracy of the algorithms varied, with CNN surpassing the others in the acquired dataset and Random Forest demonstrating encouraging results in the public dataset. After augmentation, significant gains were seen in all algorithms, especially with the SMOTE-ENN approach which successfully dealt with noise and class imbalance and significantly increased accuracy.

Further research projects might investigate more sophisticated data augmentation techniques like generative adversarial networks (GANs) that are especially designed for EEG data. To further improve model accuracy, research should be done to investigate how feature selection strategies and hyperparameter adjusting affect classification performance. The generality of the results could also be improved by expanding the study to include bigger and more varied datasets, opening the door for the creation of reliable EEG-based addiction detection systems with practical applications.

References

[1] He, C., Liu, J., Zhu, Y. , "Data augmentation for deep neural networks model in EEG classification task: a review", Frontiers in Human Neuroscience, vol.15, pp. 765525, 2021.

[2] Rommel, C., Paillard, J., Moreau, T.et.al, "Data augmentation for learning predictive models on EEG: a systematic comparison", Journal of Neural Engineering, vol.19,no.6, pp. 066020, 2022.

[3] Paillard, J., Rommel, C., Moreau, T.et.al, "Data augmentation for learning predictive models on EEG: a systematic comparison", Journal of Neural Engineering, vol.15,no.8, pp. 02652, 2022.

[4] Kalashtami, M.P., Pedram, M.M. and Sadr, H.et.al, "EEG feature extraction and data augmentation in emotion recognition",Computational Intelligence and Neuroscience, vol.2022, 2022.

[5] Luo, Y., Zhu, L.Z., Wan, Z.Y. et.al, " Data augmentation for enhancing EEG-based emotion recognition with deep generative models", Journal of Neural Engineering, vol.17,no.5, pp. 056021, 2020.

[6] Sun, S., Yang, J., Chen, Y.H.et.al, "EEG Signals Based Internet Addiction Diagnosis Using Convolutional Neural Networks", Applied Sciences, vol.12,no.13, pp. 6297, 2022.

[7] Ksibi, A., Zakariah, M., Menzli, L.J.et.al, "Electroencephalography-Based Depression Detection Using Multiple Machine Learning Techniques", Diagnostics, vol.13, no.10, pp. 1779, 2023.

[8] Martins, F.M., Suárez, V.M.G., Flecha, J.R.V.et.al, "Data Augmentation Effects on Highly Imbalanced EEG Datasets for Automatic Detection of Photoparoxysmal Responses", Sensors, vol.23, no.4, pp. 2312, 2023.

[9] Sun, S., Yang, J., Chen, Y.H.et.al, "EEG Signals Based Internet Addiction Diagnosis Using Convolutional Neural Networks", Applied Sciences, vol.12, no.13, pp. 6297, 2022.

[10] George, O., Smith, R., Madiraju, P.et.al,"Data augmentation strategies for EEG-based motor imagery decoding", Heliyon, vol.8,no.8, 2022.

[11] Bouhadja, A. and Bouramoul, A. et.al, "A Review on Recent Machine Learning Applications for Addiction Disorders",2022 4th International Conference on Pattern Analysis and Intelligent Systems (PAIS), pp. 1 – 8, 2022.

[12] Roy, A., Paul, A.et.al, "Automated detection of substance use-related social media posts based on image and text analysis", 2017 IEEE 29th International Conference on Tools with Artificial Intelligence (ICTAI) , pp. 772 – 779, 2017.

[13] Maslej-Krešňáková, V., Sarnovský, M.et.al, "Use of Data Augmentation Techniques in Detection of Antisocial Behavior Using Deep Learning Methods", Future Internet, vol.14, no.09, pp. 260, 2022.

[14] Wang, S.H., Lv, Y.D., Sui, Y.et.al," Alcoholism detection by data augmentation and convolutional neural network with stochastic pooling", Journal of medical systems, vol.42, pp. 1 – 11, 2018.

[15] Rekabdar, B., Albright, D.L.et.al, "From machine learning to deep learning: A comprehensive study of alcohol and drug use disorder", Healthcare Analytics, vol.2, pp. 100104, 2022.

Sensing the Health of a Computer Network through Estimation of its Reliability based on Sensors Response Data Analysis and its Impact on Network Management

Veerender Singh Tomar[1] and Virender Ranga[2]

[1,2] Department of Information Technology, Delhi Technological University, Delhi, India [1]

siddharthtomar10@gmail.com, 2virenderranga@dtu.ac.in

Abstract: **The health of a computer Network is a critical issue for organizations that are digitally interconnected within their networks and connected to the internet and hence significantly rely on their networks for performance and overall productivity, including communication, data transfer and running their business operations. This study presents an in-depth research conducted on the network of the Delhi Development Authority, which consists of more than 4100 nodes, focusing on the reliability of computer networks that can be predicted or estimated in different scenarios. Starting with a definition and explanation of its significance, the study discusses the elements that affect network reliability, providing a broad overview of network dependability and specifically addressing reliability. It then delves into various methods for evaluating a network's reliability, such as simulation-based techniques like Monte Carlo simulation and analytical methods like Markov models and fault trees. The study further explores methods for analyzing network reliability, including failure analysis, redundancy methods, and failure recovery plans. Additionally, case studies are presented to illustrate how network reliability can be applied in practical scenarios across different sectors. The final section of the survey discusses challenges, potential avenues for future research, and the importance of ongoing research in this area.**

Keywords: Reliability, Network Reliability, Network Dependability, Sensor Response Data Analysis, Network health, Network Reliability Evaluation Techniques, Network Analysis Report, Factors of Network Reliability, Network Management

I. INTRODUCTION

This review paper's introduction offers a thorough grounding in network reliability. "*Network Reliability"* describes a system's regularity in providing services without interruptions or breakdowns. This refers to how much a network's infrastructure and services can be relied upon to function consistently and reliably. The modern interconnected activities would not be possible without the reliability of their networks, which is the ability of a system to constantly supply services without interruption. To provide constant access to data and applications, data centers, the backbone of contemporary computing and cloud services, rely extremely on reliable networks [1]. Stable networks are crucial for the operation and provision of key services by critical infrastructures such as the Internet, financial IT backbone and power grids etc. Analytical methodologies such as Markov models, fault trees, reliability block diagrams, and simulation-based approaches such as Monte Carlo simulation and network emulation are explored in this study. This report also analyses research papers from other authors based on similar principle.

II. MOTIVATION

Systems Department of Delhi Development Authority (DDA) recently established a state of art structured metropolitan Computer Network throughout Delhi of more than 4100 nodes. When the IT consultants were testing the network's functioning, it was desired that the reliability of the network should also be estimated/computed so that at any point of time the status about the health of the Network and some short of assurance about the sustainability and certainty about the information communication through the newly established network can be known. The consultants surrendered on this issue stating that this is the subject of research. This event motivated the author to think in the direction of establishing a systematic mechanism to compute the reliability and hence the means to judge the health of a computer network.

III. LITERATURE REVIEW

Reliability estimation in computer networks is a complicated problem that requires ongoing study. The difficulty of dependability estimate is increased by the variety of modern network topologies, such as mesh, star or hybrid configurations. The complex nature of these topologies makes it hard to evaluate the network's overall resilience. The second major obstacle is the ever-changing nature of network settings.

Constant traffic fluctuations, software and hardware updates, and other external factors all contribute to the ever-evolving operational environment in which networks must function. Such changeability calls for dynamic dependability models that can track evolution [2]. One of the biggest obstacles is the wide variety of network nodes. The evaluation of total dependability is complicated by the fact that routers, switches, servers, and cables all have their own unique failure modes and reliabilities [3]. Obtaining precise information regarding network failures and performance is complicated by issues like privacy concerns, a dearth of comprehensive monitoring tools and a lack of historical data. With the ever-increasing size and complexity of today's networks, scalability is also a major issue. Large-scale network reliability estimation requires a lot of computing power, thus effective techniques and methodologies are needed to deal with this scalability problem. When thinking about promising areas for future study, machine learning and AI immediately come to mind. As a result of their ability to learn from past experiences and adapt to a constantly changing network environment, these technologies can be used to greatly enhance the precision with which reliability estimates are made [4]. Network security may undergo a dramatic shift with the advent of quantum networking. Networks could be more resistant to failures and assaults thanks to the quantum advantages offered by these technologies. The blockchain is also intriguing since it may be used to build trust and lessen the effect of network infrastructure's weak spots. Crucial network components, such as authentication and authorization systems, may be made more trustworthy by using blockchain-based solutions. Agile Software Development (ASD) has revolutionized the industry by encouraging a more collaborative and iterative approach to creating software. Agile mobile app developers use an incremental and iterative process that relies on self-managing and cross-functional teams to create their apps [5]. MOBILE-D, RaPiD7, Hybrid methodology, MASAM, and Scrum with Lean Six Sigma (SLeSS) are some of the most well-known agile mobile application development techniques currently in use. Researchers agree that the Agile methodology is a good fit for creating mobile applications. Agile software development does away with the separation of phases such as development and testing. In agile, software estimate is done concurrently with both development and testing [6]. Very little work is documented in the literature about the estimation of effort in agile development, which is a relatively new area of study [7]. Examining the test effort estimation methodologies for both traditional desktop/laptop software development and mobile software development is a major contribution of this study. Furthermore, two methods of mobile application development process, namely, traditional software development and agile software development, are compared and contrasted with regards to development and test effort estimation techniques. Finding out what sets mobile apps apart from more conventional software is also a significant contribution [8]. The decentralized nature of edge computing's data processing poses problems and possibilities for network uptime. Taking into account the unique set of constraints and opportunities presented by edge environments may be a focus of future study. Additionally, improvements in network simulation tools are crucial for accurately assessing network resilience. Insights into network behaviour can be gained by using these technologies, which generate virtual environments in which researchers can test different scenarios and failure types. It would be impossible to emphasize the value of continuing studies into network dependability. It has an immediate effect on business continuity by guaranteeing that companies can keep running despite network outages, which can lead to significant financial losses [9]. As resilient networks are less vulnerable to assaults and breaches, network dependability is also an essential part of cyber-security. Adaptability of dependability models is ensured through continuous study as new technologies like 5G, the IoT, and edge computing become more commonplace. In the end, this study is crucial for enhancing network efficiency, decreasing latency, and giving users a dependable online experience [10]. The estimation of a basic reliability measure of a communication net-work, namely the 2-terminal reliability is the fundamental and primitive method. [11] The automation and continuous explosion of personal and professional information to the digital world provides a potent ground to the adversaries to perform numerous cyber-attacks, thus making security in IoT a sizeable concern[12]

IV. NETWORK RELIABILITY EVALUATION TECHNIQUES

The final half of the review paper examines network dependability evaluation techniques, with a focus on hybrid solutions that integrate analytical and simulation methods to examine the network reliability in depth.. Analytical methods like Markov models are used to calculate network dependability by recording system state probabilities and depicting network state transitions. Fault trees and reliability block diagrams help to analyse the system failure causes and effects. Simulation-based technologies construct virtual network replicas to test performance and collect reliability measures. Monte Carlo simulation evaluates network performance using random input. Network dynamics are simulated via discrete event simulation of packet arrivals and departures.

V. FACTORS AFFECTING NETWORK RELIABILITY

This part of the paper examines various factors like Link Failures, poor technology or environmental conditions like power outages or natural catastrophes etc. which affect network reliability. In today's interconnected world, network dependability is crucial to ensuring uninterrupted communication, productive operations, and safe data transmission. Reliable and flexible emergency communication is a key challenge for search and rescue in the event of disasters, especially for the case when base stations (BSs) are no longer functioning. [16] The efficiency and dependability of a network might be hampered by several external variables. To effectively improve network dependability, one must have a firm grasp on these aspects and their ramifications. Network failure has several causes and symptoms, any of which can decrease or eliminate user access. Many events can disrupt network operation, compromising data transit and communication security. When the wires between nodes in a network break, this is known as a link failure.[Damaged cables, faulty technology, or environmental conditions like power outages or natural disasters are all potential causes. When a link fails, data transmission stops, disrupting communications and making network services inaccessible. The consequences of link failures can be lessened using redundancy in network design, such as backup links. A node failure occurs when routers, switches, or servers malfunction. Hardware, software, and power failures can cause these issues. A node failure could disrupt data and service availability and disconnect them from the network. Backup nodes or load balancing can keep the network online if any components fail. When data supplied and received exceeds network capacity, congestion develops. At times of high demand or exceptional events, this delays operations, increases latency, and lowers performance. Congestion can hinder critical services, user experiences, and data transfer. Quality of Service (QoS) methods like traffic prioritization and bandwidth allocation can minimize congestion. Data packet loss occurs when data packets fail to arrive at their network destination. Network operators can improve network dependability by proactively addressing various failure sources and applying solutions like redundancy, QoS systems, and error detection. Link failures disrupt data transfer between network computers. System failure can be caused by environmental factors like bad weather and physical difficulties like cable breaks or technology malfunctions. Data transmission delays and network communication disruptions are caused by link failures. However, network nodes fail when they malfunction. System issues can result from hardware failures, power outages, and improper programming. If a node fails, network segmentation can cause several issues for users, including catastrophic data loss. When data transmission exceeds network capacity, congestion occurs. Progress may halt or slow due to the existing condition. Congestion hinders answers and lowers user experience due to increased network traffic. Another problem with network dependability is packet loss. It happens when some of the data packets travelling across the network never make it to their destination. Congestion in the network, malfunctioning equipment, or insufficient error detection and repair systems are all potential causes of packet loss.

VI. EVALUATION AND MEASUREMENT OF NETWORK DEPENDABILITY

Reliable network maintenance needs rigorous analysis, frequent monitoring, and a variety of tools and methods to assess and improve network dependability. KPIs are critical to network health and dependability. These indicators illuminate network performance and help spot faults and measure changes. Latency, packet loss, and throughput are key network reliability KPIs. Latency monitors data transmission time, revealing network responsiveness and congestion difficulties. Network reliability and data integrity are measured by packet loss, the percentage of data packets that fail to arrive. Data throughput, on the other hand, indicates the network's efficiency and capacity.

VII. ENHANCING NETWORK RELIABILITY

Here numerous technologies and strategies have been examined to improve network reliability. QoS is crucial in prioritising and distributing network resources to key applications to ensure optimal data transfer rates. Having numerous identical components or channels in a network ensures redundancy in case of failure. N+1 and N+K redundancy models easily move to backup resources during disruptions to maintain service continuity. To avoid overload, load balancing uses redundancy to spread network traffic evenly across resources. However, fault tolerance emphasises a network's resilience to breakdowns, ensuring it continues to run. Restoring network services quickly after disruptions requires failure recovery and resilience solutions.

Dynamic routing and connection repair speed up traffic rerouting, reducing downtime. Automation and configuration management reduce human errors, speed up problem resolution, and improve network availability.

VIII. TRENDS IN NETWORK RELIABILITY

Network reliability evolves owing to technology advances, user needs, and threats. Staying updated on network reliability developments is crucial in this dynamic market. Several major developments affect dependable network design, management, and maintenance. The rise of edge computing is changing network designs. Edge devices near data sources reduce -latency and processing time. Edge computing improves user experience and enables realtime applications, but it adds failure points. Decentralised edge networks require redundancy, fault tolerance, and dynamic routing, which traditional dependability methods must accommodate. AI and advanced analytics: Network topologies may make centralised monitoring insufficient. Advanced analytics, machine learning, and AI solutions are in demand. These solutions provide real-time network health, performance, and failure analytics. Cyberattacks are threatening network resiliency and hence requires proper security measures.

IX. NETWORK RELIABILITY ANALYSIS APPROACHES

Network failure analysis, risk assessment, redundancy, fault tolerance, and recovery strategies are all discussed in the review paper's fourth part on reliability analysis. These methods are essential for pinpointing vulnerabilities, reducing the effects of failures, and guaranteeing the speedy restoration of network services.

Adding redundancy and fault tolerance to a network is one method to boost its dependability. Redundancy is the practice of having multiples of critical parts or paths in a system so that they can be used in case one fails. While failures can still have an effect on a network's performance, they are less noticeable when using redundancy models like N+1 or N+K.
Network traffic can be balanced across several paths using load balancing and traffic engineering techniques to improve both resource usage and fault tolerance. Network operators can reduce the impact of failures and keep services running smoothly even in extreme conditions by employing redundancy and fault tolerance methods.

Service Level Agreements (SLAs) are crucial contracts that must be established between service providers and their clients.

Service level agreements (SLAs) specify the dependability, availability, and performance characteristics that are anticipated of a service. They are used as guidelines for gauging network dependability by specifying what is expected of service providers and the repercussions for falling short. Network operators can improve network dependability by monitoring and evaluating SLA compliance to pinpoint problem areas and implement fixes. [17]
The transfer of data from the sender to the receiver in any form and through any mode is the most important aspect of the Network Reliability. The reliability of a network can only be evaluated through constant monitoring and measurement of its performance. Key performance indicators (KPIs) like latency, packet loss, and throughput are monitored by network managers using a wide variety of tools and approaches. These metrics help us understand how networks function, spot problems that could compromise their dependability, and catch abnormalities early so we can fix them and minimize disruptions to services as soon as possible.

Quality of Service (QoS) mechanism implementation is another key strategy. QoS methods guarantee that mission-critical applications and services have access to sufficient bandwidth, low latency, and few dropped packets by giving them higher priority in the network. Network operators may reduce congestion, optimize resource allocation, and guarantee consistent performance for mission-critical services with the help of traffic shaping, prioritization, and resource reservation [10].

Here an example of data transfer through mail in DDA network is shown where 100% data has been transferred from the sender to the receiver. It means that during this task the network was 100% reliable, this is another aspect of the reliability.

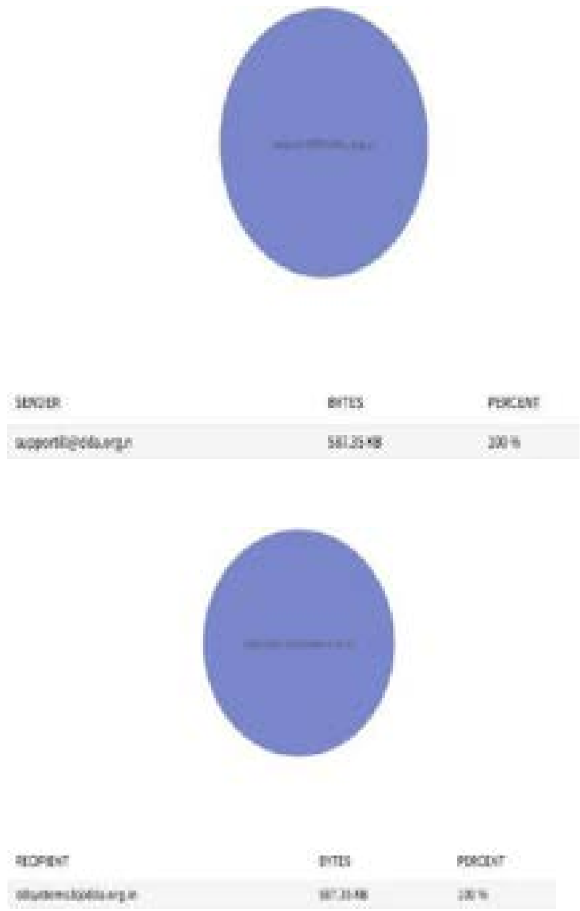


Fig 1: Sender and Receiver concept of Network Reliability (Ref. DDA Delhi Network)

The network's ability to withstand and recover from security events and other interruptions can be improved by incorporating redundancy, backup systems, and disaster recovery plans into the infrastructure. [14]

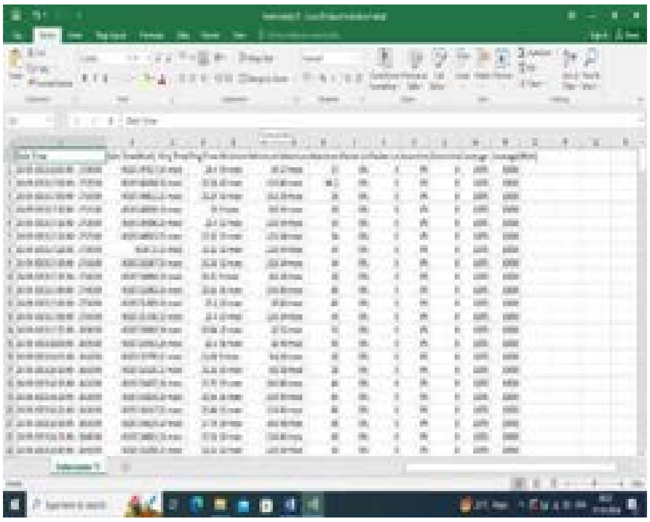
It is essential to regularly test and validate network components, settings, and protocols in order to assess reliability and discover potential vulnerabilities. Network performance can be evaluated under varying situations and potential failure points can be located with the use of load testing, stress testing, and penetration testing, all of which are simulations of real-world events. With the help of continuous testing and validation, operators can proactively fix security flaws, fine- tune settings, and guarantee the system's ongoing dependability. [8].

Network reliability analysis is a dynamic process that must accommodate new network architectures, technologies, and user expectations. Sharing best practices, exchanging knowledge, and pushing innovation in network reliability analysis all rely heavily on collaboration between researchers, industry professionals, and network operators.

The study concludes with a discussion of the many methods available for analysing network reliability. Network dependability can be evaluated and enhanced through the use of various tools and methods, such as failure analysis and risk assessment, redundancy and fault tolerance techniques, failure recovery and resilience tactics, service level agreements (SLAs), performance monitoring, quality of service mechanisms, security measures, network management and automation, and continuous testing and validation. Network operators can better meet the needs of today's communication and data-driven settings by combining these methods fully to improve the networks' dependability, availability, and performance [9].

X. CASE STUDY FOR DDA NETWORK

The data of 82 sensors was collected for the 4100 plus nodes Computer Network of Delhi Development Authority for 03 months from 01 Oct 2023 to 31 Dec 2023. The data was collected at the interval of 5 minutes round the clock.



Then various graphs were plotted based on this data. Figs. 1(a) ,1(b),1(c),1(d) and 2 show our study about the working of DDA's Computer Network based on the sensors data for 03 months from Oct 23 to Dec 23. The 04 scenarios have been considered for drawing the graphs. 1(a) with 30 Second ping interval time 1(b) with 60 Second ping interval time and similarly 1(c) with 30 Seconds and 1(d) with 60 Seconds ping interval. It is evident that with increase of ping interval time the down time decreases i.e. failure rate decreases this clearly means that reliability is increasing with increase in ping interval due to obvious reason of free availability of the channel due to increase in ping interval time and based on these parameters and we achieved the following results:

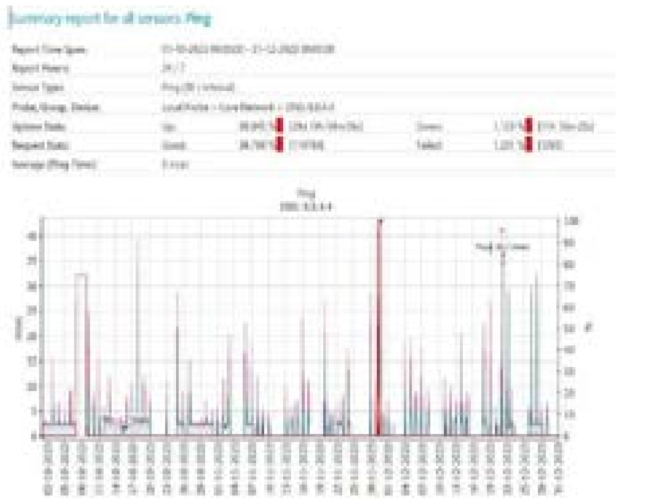


Fig. 2 (a): Delhi Development Authority's Network Analysis report from 01 Oct 2023 to 31 Dec 2023(30 Seconds Ping Interval)

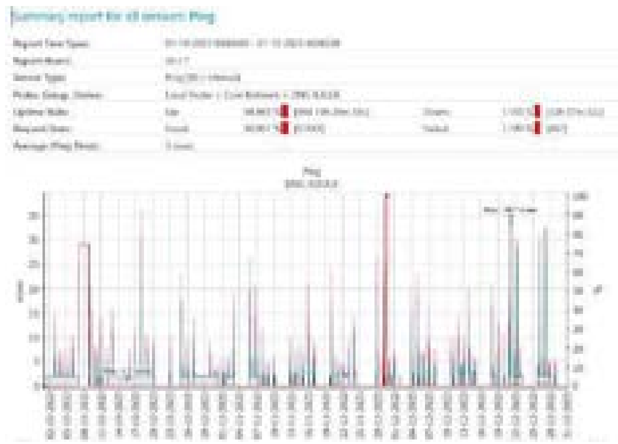


Fig. 2 (b): Delhi Development Authority's Network Analysis report from 01 Oct 2023 to 31 Dec 2023(60 Seconds Ping Interval)

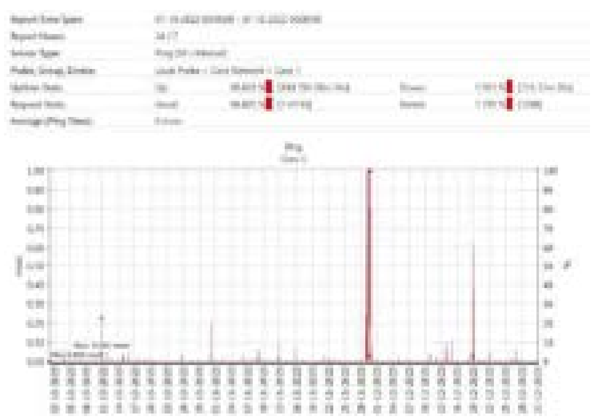
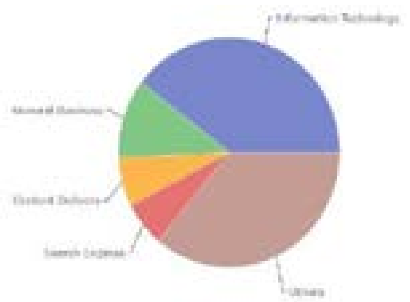


Fig. 2(c): Delhi Development Authority's Network Analysis report on different parameters from 01 Oct 2023 to 31 Dec 2023 (30 Second Ping Interval)

Fig. 2(d): Delhi Development Authority's Network Analysis report on different parameters from 01 Oct 2023 to 31 Dec 2023 (60 Second Ping Interval)



CATEGORY	HITS	PERCENT
Information Technology	12890837	39.6 %
General Business	3612902	11.09 %
Content Delivery	2259159	6.94 %
Search Engines	2066247	6.35 %
Others	11727530	35.99 %

Figure 3: Delhi Development Authority Network usage during Oct 23 to Dec 23

It has been observed from Fig 2(a & b, c & d) that stable networks have lesser downtime and consequently more reliability and crucial to the uninterrupted operation of critical infrastructure like power grids. The study's primary conclusions and recommendations show the crucial importance of network dependability in a wide range of fields. (Malhotra, 2021) In the study the effort has been made to detect the parameters on which the reliability mainly depends and the relation of these parameters with reliability to enhance the reliability if required.

For deriving the reliability of a computer Network once the reliability of any component is available or computed can be estimated with the help of the following derived algorithm based on the value computations of a matrix. The design and the algorithm are precisely described as under: -

DESIGN MODE: **NETWORK RELIABILITY ESTIMATION**

STEP 1 :- analysis of each network component in terms of reliability and computing or assigning a value of reliability for each component as under:

1. Cable & Connectors Reliability = CCR
2. Modems RELIABILITY = MOR
3. Trans receivers RELIABILITY = TRR
4. Hubs RELIABILITY = HBR
5. Switch: It may increase or decrease the reliability based on its type. Intelligent switch increases the reliability and speed both. RELIABILITY = SWR
6. Check Point/Firewalls:- It increases the overall reliability but decrease the speed.
RELIABILITY = CFR
7. Repeater :- It increase the reliability
RELIABILITY = RPR

The Reliability to each segment or Component can be assumed or calculated through experiment as above.

STEP 2:- Finding Inter Nodes Paths and components in the paths

Paths towards every node from other nodes :-
Towards N1 :-
From N2 : N2-N1
N2-CC-SW1-CC-N1
N2-CC-SW1-CC-N3-CC-R1-CC-N6-CC-N1
N2-CC-N4-CC-N3-CC-R1-CC-N6-CC-N1
N2-CC-N6-CC-N3-CC-SW1-CC-N1
From N3 : N3-N1
N3-CC-SW1-CC-N1
N3-CC-R1-CC-N6-CC-N1
N3-CC-N4-CC-N2-N1
Paths towars every node from other nodes :-
Towards N1 :-
From N4 : N4-N1
N4-CC-N3-N1
N4-CC-N2-N1
N2-CC-N4-CC-N3-CC-R1-CC-N6-CC-N1
N2-CC-N6-CC-N3-CC-SW1-CC-N1
From N6 : N6-N1
N6-CC-N1
N6-CC-R1-CC-N3-N1
From N5 : N5-N1
N5-CC-N6-N1
Similarly Paths towards N2, N3, N4, N5 and N6 can be analysed which will be as under:-
N2-N1(<u>RP1</u>) N1-N2 N1-N3 N1-N4 N1-
N5 N1-N6
N3-N1(RP2) N3-N2(RP6) N2-N3 N2-N4
N2-N5 N2-N6
N4-N1(RP3) N4-N2(RP7) N4-N3(RP10) N3-N4
N3-N5 N3-N6
N5-N1(RP4) N5-N2(RP8) N5-N3(RP11) N5-

N4(RP13) N4-N5 N4-N6

N6-N1(RP5) N6-N2(RP9) N6-N3(RP12) N6-N4(RP14) N6-N5(RP15) N5-N6

P1 to P15 are the distinct paths

RP1 to RP15 are the reliability of each path

RP1= Reliability of Path N2-N1 or N1-N2 and so on….

ALGORITHM:

Find the longest path and its component, Let the longest path has

```
        Lp= no of Components / Segments.
        Create a Lp x Pm x Np Matrix where
        Np=n(n-1)/2 , where n is the number of
        nodes and Pm is the maximum no of paths
        available from any one node to the any other
        Store Values of Reliability of each
        component in the matrix
        Keep 1 where corresponding value is blank
        Set Reliability =1
   for i=1 to Lp
        Begin
{
        for j =1 to pm
Begin
{
        For k = 1 to Np
        Begin
{
        Reliability =Reliability * Value (i, j,k)
        ++ i; ++ j; ++ k ;
          }
        End
            }
          End
              }
            End
          Printf ('Reliability of Network=',%d ,
                 Reliability )
        End
```

In the final section of the review article, the value of case studies and practical applications in demonstrating the significance of network dependability in various industries is emphasized. [15] Telecom, cloud computing, transportation, logistics, and critical infrastructure case studies shed light on the pros and cons of putting an emphasis on network resilience. Stakeholders can receive actionable insight into how to achieve improvements in network dependability and operational efficiency by examining these cases. In addition to the aforementioned industries, there are many others where case studies can shed light on the importance of network reliability. For instance, trustworthy communication networks are crucial in the healthcare business for providing telemedicine services, sharing real-time data, and sharing patient records. Patient care, emergency response systems, and the general efficacy of healthcare delivery are all areas where case studies can shed light on the importance of network reliability. Security during financial transactions, data transfers, and real-time market data all rely largely on network reliability. The banking and financial industry provides a useful case study setting in which to examine the effects of network outages, cyber threats, and the introduction of highly reliable network infrastructure.[3] Experts in the field and academic scholars can learn from these examples to better protect the integrity of financial networks and prevent future disruptions. Critical services in smart cities, such as transportation, energy management, public safety, and environmental monitoring, rely on stable communication networks to function. Studying these cases can shed light on the significance of network dependability in facilitating effective public transit, real-time monitoring, and the distribution of vital community services.[13] Problems with scalability, interoperability, and the consolidation of smart city apps are among something they can fix.

XI. CHALLENGES AND FUTURE DIRECTIONS

The impact of emerging technologies, scalability and complexity issues, security and resilience considerations, and the potential applications of machine learning and AI are all taken into account as we discuss the challenges and future directions in the field of network reliability in the sixth section of this review paper.The stability of networks is profoundly affected by cutting-edge innovations. The development of 5G, edge computing, the Internet of Things (IoT) and virtualization presents new opportunities and threats. Above all, the software tools and utilization process, robustness and availability of security measures at every node in the intranet are essential challenges to be met for bare minimum reliability.

XII. CONCLUSION

Conclusively, this study has presented a summary of network reliability, including an overview of its evaluation methods, analysis approaches, case studies and potential future developments. The primary conclusions and recommendations of the study and research show the crucial importance of network dependability for an organisation and in vide range of fields. In the study the effort has been made to detect the parameters on which the reliability mainly depends and the relation of these parameters with reliability to enhance the reliability if required. New technologies, scalability, security concerns, and the potential of machine learning and AI applications have been identified as areas requiring attention. The study will add a lot in the direction of certification of Computer Networks considering reliability as important parameter. Once the reliability will be counted in the certification of networks, the overall performance, throughput and the life cycle of the computer networks will be enhanced.

REFERENCES

[1]. S. Bathla and V. Vasudevan, "A Framework for Reliability Analysis of Combinational Circuits Using Approximate Bayesian Inference," in IEEE Transactions on Very Large

Scale Integration (VLSI) Systems, vol. 31, no. 4, pp. 543554, April 2023, doi:10.1109/TVLSI.2023.3237885.

[2]. Chen, X. G. (2017, April 24). Research on reliability of complex network for estimating network reliability. Journal of Intelligent and Fuzzy Systems. https://doi.org/10.3233/jifs-169291

[3]. Dwivedi, Y. K., Hughes, L., Ismagilova, E., Aarts, G., Coombs, C., Crick, T., & Williams, M. D. (2021). Artificial Intelligence (AI): Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. International Journal of Information Management, 57, 01994. https://www.sciencedirect.com/science/article/pii/S026840121930917X

[4]. H. Cancela, L. Murray and G. Rubino, "Reliability Estimation for Stochastic Flow Networks With Dependent Arcs," in IEEE Transactions on Reliability, vol. 72, no. 2, pp. 622-636, June 2023, doi: 10.1109/TR.2022.3178413.

[5]. Davila-Frias, P. Khumprom and O. P. Yadav, "Reliability Estimation Using Long Short-Term Memory Networks," 2023 Annual Reliability and Maintainability Symposium(RAMS), Orlando, FL, USA, 2023, pp. 1-6, doi: 10.1109/RAMS51473.2023.1008825

[6]. Colbourn, C. (1987). The Combinatorics of Network Reliability. Oxford University Press

[7]. Huang, C. F., Huang, D. H., & Lin, Y. K. (2020). Network reliability evaluation for a distributed network with edge computing. Computers & Industrial Engineering, 147, 106492. https://doi.org/10.1016/j.cie.2020.106492

[8]. Hewson, C., Vogel, C., & Laurent, D. (2015). Internet research methods. Sage.

[9]. K. Y. Tse, W. K. Mak, and S. H. T. Lam, "Network Reliability in Financial Trading Systems," International Journal of Network Management, vol. 30, no. 5, pp. e2044, Sep. 2020. doi: 10.1002/nem.2044.

[10]. Kang, J., Xiong, Z., Niyato, D., Zou, Y., Zhang, Y., & Guizani, M. (2020). Reliable federated learning for mobile networks. IEEE Wireless Communications, 27(2), 72-80. https://doi.org/10.1109/MWC.001.1900119.

[11]. Khadiri, M. E., Marie, R., & Rubino, G. (1991). Parallel estimation of 2-terminal network reliability by a crude Monte Carlo technique. In Sixth International Symposium on Computer and Information Sciences (pp. 559–570).

[12]. Malhotra, P., Singh, Y., Anand, P., Bangotra, D. Singh, P. K., & Hong, W. C. (2021). Internet of things: Evolution, concerns and security challenges. Sensors, 21(5), 1809. https://www.mdpi.com/1424-8220/21/5/1809

[13]. Primeau, N., Falcon, R., Abielmona, R., & Petriu, E. M. (2018). A review of computational intelligence techniques in wireless sensor and actuator networks. IEEE Communications Surveys & Tutorials, 20(4), 2822-2854 https://ieeexplore.ieee.org/abstract/document/839 6218/

[14]. Venkatesh, V., Shaw, J. D., Sykes, T. A., Wamba, S. F., & Macharia, M. (2017). Networks, technology, and entrepreneurship: A field quasi-experiment among women in rural India. *Academy of Management Journal*, *60*(5),1709-1740. https://journals.aom.org/doi/abs/10.5465/amj.201 5.0849

[15]. Xie, J., Yu, F. R., Huang, T., Xie, R., Liu, J., Wang, C., & Liu, Y. (2018). A survey of machine learning techniques applied to software defined networking (SDN): Research issues and challenges. IEEE Communications Surveys & Tutorials, 21(1), 393-430. https://ieeexplore.ieee.org/abstract/document/8444669/

[16]. Zhao, N., Lu, W., Sheng, M., Chen, Y., Tang, J., Yu, F. R., & Wong, K. K. (2019). UAV assisted emergency networks in disasters. IEEE Wireless Communications, 26(1), 45-51. https://doi.org/10.1109/MWC.2018.1800160

[17]. Trivedi, K. S., & Bobbio, A. (2017). Reliability and availability engineering: modeling, analysis, and applications. Cambridge University Press

Drone-based Elephant Detection and Counting With YOLOv5

Jahnavi Kommalapati
Department of Computer Science and Engineering Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada,India
kommalapatijahnavi9@gmail.com

Suresh Babu Dasari
Assistant Professor
Department of Computer Science and Engineering
Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada,India
dasarisuresh@vrsiddhartha.ac.in

Harika Pidaparthi
Department of Computer Science and Engineering Velagapudi Ramakrishna Siddhartha Engineering College
Vijayawada,India
harikap81189@gmail.com

***Abstract*— Nowadays, object detection is popularly used in most vision-based AI and it is sole to the various computer vision tasks like face recognition, Animal detection. This study focuses on the identification and tally of elephants using drone photos using YOLOv5. Drones allow you to span a vast area and facilitate prompt and effective animal monitoring. YOLOv5 is the first YOLO model to be written in PyTorch framework which is fast and accurate. As every species plays a crucial role in the ecosystem, extinction of one specie may cause damage to entire ecosystem. So, it is important to detect the endangered species and ensure their protection. In this project YOLO algorithm is used to detect and count the endangered species from drone images. In this project elephant is taken as a reference as the population of elephant is declining over several decades due to poaching and loss of habitat. Existing model only has detection of animal so counting of animal's feature is added to it to increase its efficiency. The model developed has mAP 89.4%, Precision 89.1% and Recall 83.2%.**

Keywords-Object detection, YOLOv5, Drone images, Elephant.

I. INTRODUCTION

The natural equilibrium of our planet depends on human and wildlife cooperation. The natural habitats of many species have, however, been invaded by urbanization and human activity in many areas, increasing conflicts between people and wildlife and endangering the existence of a number of famous and threatened species. With its enormous cultural, ecological, and conservation worth, the elephant is one of these vulnerable species that represents both the benefits and the challenges of preserving nature.

Effective monitoring of elephant populations movements, habits, and habitats is essential to reducing the negative effects of these conflicts and safeguarding them. Conventional techniques for gathering and tracking data, such video traps and ground-based surveys, have shown to be beneficial but are frequently costly, labor-intensive, and time-consuming. Technological developments in the last few years have created new opportunities for wildlife protection and monitoring.

Using Unmanned Aerial Vehicles (UAVs) or drones fitted with sophisticated computer vision algorithms is on such invention. The YOLO (You Only Look Once) technique is used to the important task of elephant species detection from drone photographs in this article. YOLO is a well-known object detection method that is fast andaccurate, which makes it a great option for real-time applications. Our mission is to transform the way we monitor and protect elephant populations by utilizing the power of YOLO and deploying camera-equipped drones.

II. MOTIVATION

As there is a serious threat to elephants' population due to due to conflicts with humans, poaching, and habitat destruction. To address this, conservationists are turning to innovative solutions like employing drones equipped with YOLOv5 technology for accurate elephant spotting and counting. Traditional methods, such as manual counts or manned aircraft surveys, are often time-consuming, expensive, and impractical in remote or densely forested areas. By leveraging drones and YOLOv5, conservation efforts can be significantly streamlined, enabling more effective monitoring of elephant populations and targeted conservation initiatives to safeguard these majestic creatures for future generations.

III. LITERATURE SURVEY

Chamoso, P et al. [1] investigates how to effectively monitor vast cattle regions using UAVs and artificial vision. It focuses on counting and monitoring animal species using UAV video recordings. The system tracks and logs the quantity of animals found by analyzing UAV camera images. The system's efficacy is demonstrated by a battery of tests. The results of these evaluations are presented in the paper's conclusion, which highlights the method's potential for thorough livestock area surveillance and animal counting with cutting-edge technology.

Prosekov, A et al. [2] aims to assess monitoring techniques for large wild species in order to conserve biodiversity, particularly elk and moose. It emphasizes how crucial trustworthy wildlife data is to making wise management choices. Recognizing that these species live in large, diversified environments with low densities, the research investigates several techniques to estimate population size without causing disturbance or significant financial burden. These models' accuracy is dependent on statistical data as well as its customization to certain geographic and meteorological circumstances. Using a combination of current and old technology, such as UAVs, GPS sensors, and cameras, provides a flexible way to bridge the gap between the practical feasibility and research.

Kellenberger, B et al. [3] addressed the difficulty of conducting a large-scale animal census by employing CNNs and UAVs. With tiny datasets, current methodologies frequently break down, resulting in a large number of false alarms in large study areas. Through the implementation of customized evaluation techniques and CNN scaling, this study dramatically lowers false positive rates, improving detection accuracy by a factor of ten over earlier approaches. Their CNN reduces the requirement for manual verification by three times with a 90% recall rate, allowing for effective acquisition data screening. With the help of this development, rangers in reserves will be able to automatically identify almost every species, transforming wildlife conservation through precise and effective monitoring.

Gemert, J.C.v et al. [4] looks into the use of small unmanned aerial vehicles (UAVs) and object identification for efficient animal surveying in environmental conservation. By gathering their own data using a quadcopter drone, they assess animal detection and counting.Although conventional human-scale visual object recognition algorithms are not well suited for drone viewpoint, lightweight automatic detection methods demonstrate potential for conservation applications. This study provides a promising path for more effective and economical field surveys by highlighting the possibility for adapting such methodologies to support conservation and wildlife monitoring activities.

Isla Duporge et al. [5] pioneers the use of cutting-edge technology by employing Convolutional Neural Network (CNN) models using very-high-resolution satellite data to automatically detect and count African elephants in complex forest savanna landscapes. The researchers trained and assessed the CNN model using 11 photos from 2014 to 2019 using WorldView-3 and 4, the highest resolution satellite data available. The CNN's remarkable accuracy was demonstratedby the results, which matched human detection abilities. F2 scores in homogeneous and heterogeneous areas were 0.73 and 0.78, respectively, which are similar to human labeling scores of 0.77 and 0.80. Most remarkably, the model showed generalizability by identifying elephants without further training in a different geographic region (Kenya) and using lower resolution satellite data (GeoEye-1). This study offers a novel method for monitoring elephants in a variety of environments, highlighting the potential of deep learning and satellite remote sensing for wildlife conservation. Elephant detection may now be automated with accuracy close to that of a human, opening up new possibilities for conservation andwildlife monitoring.

IV. METHODOLOGY AND RELATED WORK

A. Traditional structure of yolov5

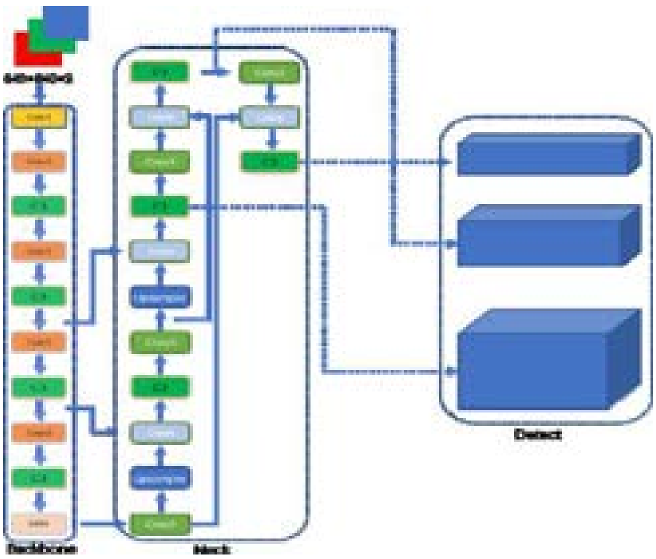


Figure 1. Yolov5s structure

Backbone Network:

YOLOv5 uses a backbone network based on CSPDarknet53 or EfficientNet as the feature extraction backbone. The choice of the backbone allows for better feature representation and contributes significantly to accuracy.

Neck and Head:

YOLOv5 introduces a novel "neck" and "head" structure that refines and processes features at different scales. This helps in improving detection accuracy for objects of varying sizes.

Input Layer:

The input layer takes the image as input. YOLO typically takes images of fixed sizes as input, which are divided into a grid for processing.

Convolutional Layers:

YOLO primarily consists of convolutional layers that perform feature extraction. These layers analyze the input image to extract hierarchical representations of features.

Downsampling Layers:

While extending the receptive field, downsampling or pooling layers decrease the spatial dimensions of the feature maps. This lessens processing and aids in context capture.

Concatenation and Upsampling layers:

In some YOLO variants, concatenation or upsampling layers are used to fuse features from different scales or increase the spatial resolution of feature maps for better object localization.

Output Layer:

Bounding box final predictions, confidence ratings, and class probabilities are all contained in the output layer. After that, post-processing is used to determine and create bounding boxes around any objects that are found.

B. Training of custom yolov5 model

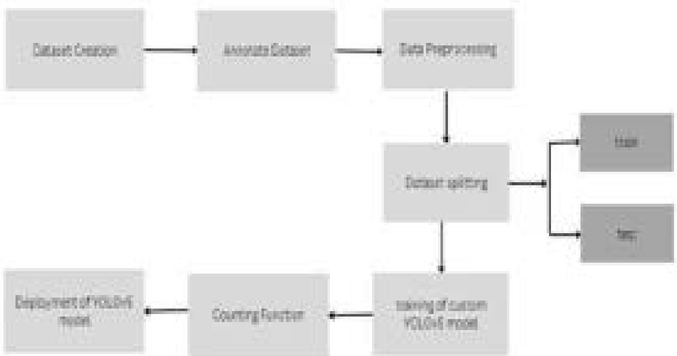


Figure 2. Methodology

A. Dataset Creation

Since research involves drone imagery of elephants, the relevant images were gathered from Google and added to roboflow as the Elephant detection image dataset. The dataset contains 309 images.

Figure 3. Elephant drone image

B. Annotate Dataset

For annotation the images should be in .png or .jpg format. Objects can be annotated inside the images of dataset using labels like bounding boxes. After completion, the annotations will be generated as text file to be used in YOLOv5 custom model training asshown in Figure 4.

```
File    Edit    View

0 0.174255 0.235514 0.252951 0.324538
0 0.599213 0.636715 0.155143 0.390979
0 0.729904 0.418227 0.254637 0.265764
0 0.783867 0.548553 0.200675 0.229988
0 0.833614 0.801539 0.263069 0.296429
```

Figure 4. .Annotations

C. Dataset splitting

Here the dataset is randomly split into train, test and validate by allocating 80% of images to train ,10% to validate and remaining 10% to test. This type of classification is required to train custom YOLOv5 model otherwise it will show error.

D. Data Preprocessing

Data preprocessing techniques applies image alterations to every image in the dataset in order to reduce time and improve performance. Preprocessing methods like Resize and Auto Orient were applied here.

Image resizing is a fundamental preprocessing step in object detection pipelines. It standardizes input dimensions, enhances computational efficiency, and contributes to the capacity of the model to identify and generalize objects effectively in various scenarios. Images might have different orientations (e.g., landscape, portrait) that can affect the accuracy of object detection algorithms if not handled properly. The "auto orient" preprocessing step, often associated with images, refers to an operation that automatically adjusts the orientation of an image based on its metadata.

Augmentation methods like flip, crop can be applied to dataset to increase the efficiency of yolov5 model.

Now generate the dataset by applying above modifications in roboflow website and export dataset in YOLOv5 PyTorch format. Then it will generate a snippet as shown in Figure 5 which is used further in YOLOv5 custom model training.

```
!pip install roboflow

from roboflow import Roboflow
rf = Roboflow(api_key="[redacted]")
project = rf.workspace("jahnavi-kommalapati").project("elephant-detection-and-counting-from-drone-images")
dataset = project.version(4).download("yolov5")
```

Figure 5. Snippet to download and unzip our dataset

E. Training of custom YOLOv5 model in google colab

Install requirements:
Clone YOLOv5 repository from github and install all required dependencies.

Assemble the Dataset:
Assemble the dataset uploaded in Roboflow using the snippet in Figure 4.

Train the custom YOLOv5 model:

We train the custom YOLOv5 model using below code

!python train.py --img 416 --batch 16 --epochs 150 --data {dataset.location}/data.yaml --weights yolov5s.pt --cache

Here dataset.location is the location where our dataset is stored and in that dataset only we have data.yaml file.Now our model training is completed and can download the elephants custom yolov5 model which is stored in a path './runs/train/exp/weights/best.pt'

F. Counting function

The quantity of bounding boxes serves as the basis for item counting generated by the YOLOv5 model during the object detection process. Each bounding box typically represents a detected object, and counting them provides an estimate of the number of objects in the image. Here torch.hub.load function is used to load the custom yoloV5 model to detect elephants.

G. Deployment of yolov5 model

Here for deployment of elephants YOLOv5 custom model we use tkinter .Tkinter is the popular GUI library of python which is used to create GUI for object detection projects. The sample GUI which is developed using tkinter is shown in Figure 6.This GUI is used to count and detect elephants in image.

Figure 6. Sample GUI

V. RESULTS & DISCUSSION

A. Elephants Detection

From Figure 6 if we click Detect Elephants button then we will get the results of elephants detected in image depending on the YOLOv5 model loaded into the GUI.

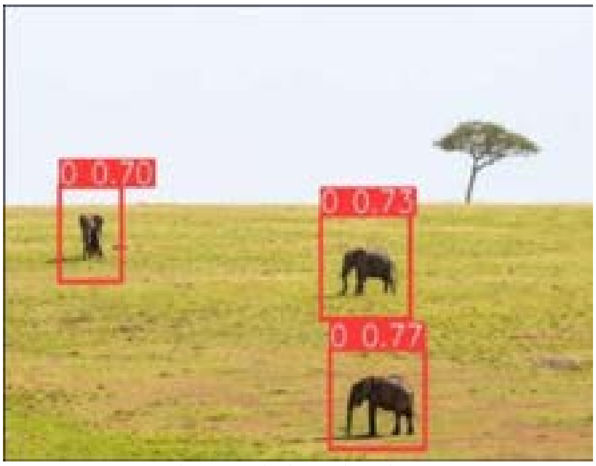


Figure 7. Elephants detection

B. Elephants Counting

From Figure 6 if we click Count Elephants button then we will get result of the count of elephants present in the image depending on the YOLOv5 model loaded into the GUI.

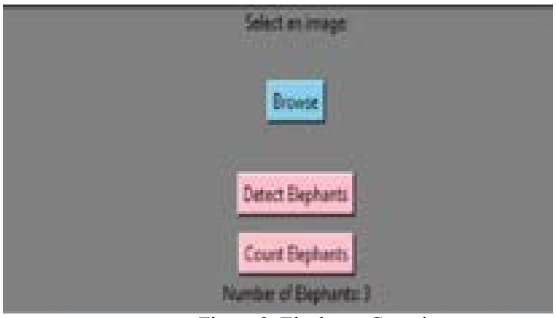


Figure 8. Elephants Counting

C. Evaluation Metrics

Precision and Recall:

The ratio of accurately anticipated positive observations to all predicted positive observations is known as precision. It highlights how accurate the model is at predicting positive classes by demonstrating the accuracy of positive predictions.

TP / (TP + FP) equals precision.

The ratio of accurately anticipated positive observations to all actual positive observations in the dataset is known as recall. It highlights completeness by demonstrating the model's capacity to recognize all positive examples.

TP / (TP + FN) equals recall.

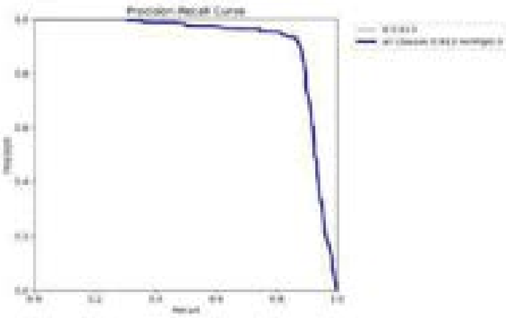


Figure 9.Precision-Recall Curve

F1-confidence threshold:

The F1 score is a statistic that strikes a balance between recall and precision by combining the two into a single value. The harmonic mean of recall and precision is used to calculate it.

F1=2×Recall×Precision/(Recall+Precision)

The F1 score accounts for both false positives (precision) and false negatives (recall), making it useful for evaluating model performance in binary classification tasks.

In object detection models like YOLO, confidence thresholds are used to filter detections based on their confidence scores. Detections with confidence scores below the threshold are often disregarded to control false positives.

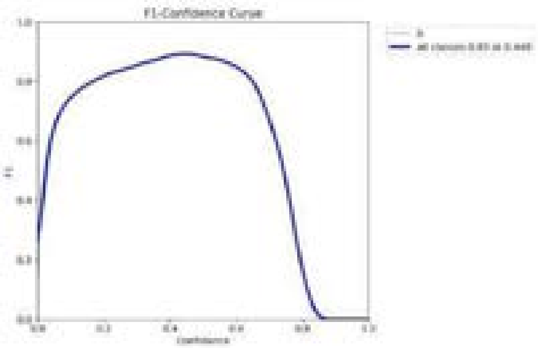


Figure 10.F1-Confidence curve

Confusion Matrix:

True Positive (TP): Correctly identified elephants by the model.

False Positive (FP): Incorrectly detected elephants (model predicts elephants where there are none).

False Negative (FN): Missed detections where elephants are present but not identified by the model.

True Negative (TN): Not typically applicable in object detection scenarios (usually relates to correctly identifying non-target objects).

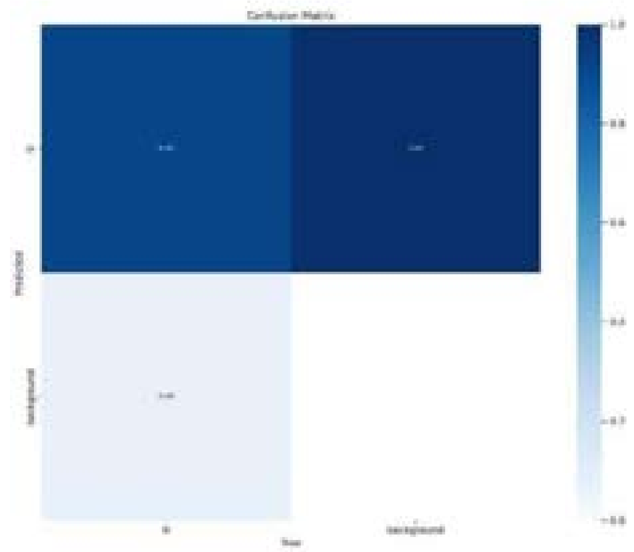


Figure 11.Confusion matrix

VI. CONCLUSION AND FUTURE SCOPE

A. Conclusion
In this paper we have used YOLOv5 algorithm to detect and count elephants from drone images which will be useful to monitor elephants in elephant reserves, forests .Our model is performing good with detection and counting of elephants very accurately from the images we provide to it.

B. Future enhancement
Since the goal of the project is to solely detect and count from drone photographs, we can include camera and video options for this model, which will enhance its utilization and lessen the strain of providing images.

VI. REFERENCES

[1] Chamoso, P.; Raveane, W.; Parra, V.; González, A. UAVs applied to the counting and monitoring of animals. In Proceedings of the Ambient Intelligence-software and Applications, Salamanca, Spain, 4–6 June 2014; Springer: Cham, Switzerland, 2014; pp. 71–80.

[2] Prosekov, A.; Kuznetsov, A.; Rada, A.; Ivanova, S. Methods for monitoring large terrestrial animals in the wild. Forests 2020, 11, 808.

[3] Kellenberger, B.; Marcos, D.; Tuia, D. Detecting mammals in UAV images: Best practices to address a substantially imbalanced dataset with deep learning. Remote Sens. Environ. 2018, 216, 139–153.

[4] Gemert, J.C.v.; Verschoor, C.R.; Mettes, P.; Epema, K.; Koh, L.P.; Wich, S. Nature conservation drones for automatic localization and counting of animals. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–7 September 2014; Springer: Cham, Switzerland, 2014; pp. 255–270.

[5] Isla Duporge, Olga Isupova, Steven Reece, David W. Macdonald, Tiejun Wang,Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes,ZSL Publisher,2020.

[6] Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. Scaled-YOLOv4: Scaling Cross Stage Partial Network. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), online, 19–25 June 2021; pp. 13029–13038.

[7] Rosli, M.S.A.B.; Isa, I.S.; Maruzuki, M.I.F.; Sulaiman, S.N.; Ahmad, I. Underwater animal detection using YOLOV4. In Proceedings of the 2021 11th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), Penang, Malaysia, 27–28 August 2021; IEEE: New York, NY, USA, 2021; pp. 158–163.

[8] Duporge, I.; Isupova, O.; Reece, S.; Macdonald, D.W.; Wang, T. Using very-high-resolution satellite imagery and deep learning to detect and count African elephants in heterogeneous landscapes. Remote Sens. Ecol. Conserv. 2021, 7, 369–381.

[9] K. Ding, X. Li, W. Guo and L. Wu, "Improved object detection algorithm for drone-captured dataset based on yolov5," 2022 2nd International Conference on Consumer Electronics and Computer Engineering (ICCECE), Guangzhou, China,2022,pp.895-899.

[10] K. S. P. Premarathna, R. M. K. T. Rathnayaka and J. Charles, "An Elephant Detection System to Prevent Human-Elephant Conflict and Tracking of Elephant Using Deep Learning," 2020 5th International Conference on Information Technology Research (ICITR), Moratuwa, Sri Lanka, 2020, pp. 1-6.

[11] Y. Zhang, Y. Cuan and Y. Jiang, "Research on improved YOLOv5 face mask recognition algorithm based on," 2023 8th International Conference on Intelligent Computing and Signal Processing (ICSP), Xi'an, China, 2023, pp. 1414-1417.

[12] Andrej Karpathy. CS231n convolutional neural networks for visual recognition. http://cs231n.github.io/, 2017. [Online; accessed 10 December 2017].

[13] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. Deep Learning. MIT Press,2016 http: //www.deeplearningbook.org.

[14] George Cybenko. Approximation by superpositions of a sigmoidal function. Mathematics Of Control, Signals, and Systems (MCSS), 2(4):303–314, 1989.

[15] Moshe Idel. Golem: Jewish magical and mystical traditions on the artificial anthropoid.1990

[16] Warren S McCulloch and Walter Pitts. A logical calculus of the ideas immanent in nervous activity. The bulletin of mathematical biophysics, 5(4):115–133, 1943.

[17] Frank Rosenblatt. The perceptron, a perceiving and recognizing automaton Project Para.Cornell Aeronautical Laboratory, 1957.

[18] Yann LeCun, Bernhard Boser, John S Denker, Donnie Henderson, Richard E Howard,Wayne Hubbard, and Lawrence D Jackel. Backpropagation applied to handwritten zip code.recognition Neural computation, 1(4):541–551, 1989.

[19] Raja Giryes, Guillermo Sapiro, and Alexander M Bronstein. Deep neural networks with 3457,random Gaussian weights: a universal classification strategy? IEEE Trans. SignalProcessing, 64(13):3444–2016.

Stock Price Prediction using Deep-Learning Model

Tamil Prakash. A
Master of Computer Applications
Hindustan Institute of Technology & Science, Chennai, India
tamilnav0905@gmail.com

Dr. Sudha
Department of Computer Applications
Hindustan Institute of Technology & Science, Chennai, India
sudhas@hindustanunvi.ac.in

Abstract— **This research study combines offline and real-time data from trading stock markets to present a unique neural network-based approach to enhance stock market forecasts. This research investigates the challenges posed by random weight initialization issues, which often limit the ability of conventional neural network algorithms to accurately represent market dynamics. By introducing the concept of "stock vector," the study highlights how LSTM networks may capture the intricate temporal correlations seen in sequential data, which are akin to language. Additionally, embedding layers are employed to transform words into intricate, meaningful representations, which further enhances the prediction potential of the model. The forecasting accuracy of these models is enhanced by vectoring high-dimensional historical data from several stocks. Based on data from empirical testing, the deep LSTM model with an embedded layer outperforms the others in stock market predictions. This methodological development represents a major step forward in the use of deep learning algorithms, providing increased accuracy and understanding for stock market research projects.**

Keywords— ***stock market forecasts, real-time data, offline data, random weight initialization, conventional algorithms, market features, stock vector, deep learning, long short-term memory***

I. INTRODUCTION

The equity market, sometimes referred to as the stock market, serves as a crucial venue for the exchange of ownership holdings in publicly traded corporations between people and institutions. Its importance stems from its function as a bridge in the global financial system, facilitating the flow of capital and giving businesses the ability to raise money for a range of projects, such as growth and innovation. Shares of publicly traded corporations are exchanged through two main routes in this dynamic arena: stock exchanges and over-the-counter (OTC) markets. Markets. Investors, ranging from individual traders to large institutional investors, participate in the market with the aim of generating returns on their investments through capital appreciation and dividends.

Because the stock market is so volatile and complex, navigating its complexities may be quite difficult. The intricacies of market activity are frequently too complicated for traditional analytical techniques like technical and fundamental analysis to properly represent. On the other hand, the ability of machine learning algorithms to sort through large datasets and find patterns that may be difficult for humans to understand has drawn attention in recent years. Algorithms for machine learning are examples of computer systems that can learn from data and apply that information to forecast or decide. Recurrent neural networks (RNNs) are one of these algorithms that have become very well-liked for stock market prediction. An artificial neural network type called an RNN is made especially to handle sequential input and identify patterns over time. What distinguishes RNNs from conventional feed forward.The cyclic connections seen in neural networks provide them the capacity to remember previous inputs and calculations. Because of this property, RNNs are highly suited for applications like time-series prediction, speech recognition, and natural language processing where the order in which incoming data is received is crucial. Which helps the network remember important details and create more complex predictions based on temporal relationships in the input.

The vanishing gradient issue and the challenge of capturing long-term relationships in sequential data are two drawbacks of traditional Recurrent Neural Networks (RNNs) that have led to the development of Long Short-Term Memory (LSTM) architectures. These restrictions frequently made it more difficult for RNNs to accurately represent intricate temporal connections. Using memory cells and gating techniques, LSTM networks improve the modeling of sequential data. These elements help LSTMs overcome problems like the vanishing gradient problem by allowing them to selectively update or forget particular information based on its relevance and maintain knowledge over extended periods of time. The effectiveness of LSTMs in financial time series forecasting and stock price prediction has been attributed to its strong capacity to handle sequential data, capture long-term dependencies, and minimize issues faced by standard RNNs. Because LSTM design includes memory cells and gating mechanisms, it is especially well-suited to simulate complex dynamics found in financial data. As a result, projections produced by LSTM-based models are frequently more accurate, particularly under dynamic and quickly shifting market situations.

II. LITERATURE SURVEY

[1]"Stock Market Prediction Using Machine Learning Methods" investigates the use of sentiment analysis and machine learning algorithms, specifically Random Forest, to forecast stock market prices. It shows that Random Forest is more effective than logistic regression in predicting stock market prices based on sentiment analysis.[2]In order to solve the shortcomings of conventional prediction techniques, this research study investigates the application of machine learning

algorithms—more especially, Random Forest and Extra Tree Regression—to forecast stock market prices with high accuracy and dependability.[3]Using daily opening prices of GOOGL and NKE stocks from the NYSE over predetermined time periods, the study trains and tests an LSTM RNN model using 80% of the data for training and 20% for testing. Various numbers of epochs are tested and the optimization is done using mean squared error. An LSTM layer with an output shape of (None, 50, and 96) and 37,632 parameters is part of the model architecture. [4]The application of machine learning techniques—Random Forest and Artificial Neural Networks, in particular—to forecast the closing price of five distinct firms across many industries is covered in this article. Financial information such as the stock's open, high, low, and close prices are utilized in the study to develop new variables that are then fed into the model. Standard strategic metrics, such as RMSE and MAPE, are used to evaluate the models; low values indicate effective stock closing price prediction. Six additional variables were developed for the purpose of predicting the closing price of the stocks after historical data for the firms was gathered from Yahoo Finance. The study comes to the conclusion that stock closing prices might be accurately predicted by the machine learning methods employed.[5]With an emphasis on classifiers and repressors, the paper presents a thorough comparison of ensemble learning techniques for stock market prediction. Its objective is to assess how well different ensemble models work with data from different continents. The study predicts stock market movements using machine learning methods such as neural networks, support vector machines, and decision trees. [6] Deep neural networks (DNNs) and conventional artificial neural networks (ANNs) are used to determine the daily direction of stock market returns. The research makes use of 60 financial and economic variables that were gathered over 2518 trading days, such as the return on the SPY, T-bill rates, and other indicators. The paper highlights how crucial it is to choose the appropriate number of major components during dimensionality reduction in order to preserve data structure while making it simpler. With trade simulations comparing several models and datasets, the modelling also focuses on producing precise and preferably lucrative direction projections. The outcomes demonstrate how DNN classifiers affect prediction accuracy and profitability.[7] High-order time series data is used in this study to forecast patterns in the stock market. The authors propose to simplify noisy financial temporal series and produce more accurate forecasts by employing motifs, or recurrent patterns. They offer a three-step method for predicting trends: identifying dynamic evolutionary patterns, reassembling the sequence using motifs in the original order of discovery, and utilizing convolution neural networks to predict trends. The paper highlights the challenges associated with multi-scaling and non-stationary in financial time series prediction, emphasizing how important it is to find hidden patterns in financial data in order to forecast trends. The proposed method displays efficiency in feature learning and outperforms traditional signal processing techniques in market trend prediction.[8]This study uses data from the iShares MSCI United Kingdom ETF to investigate machine learning models for stock market prediction. Deep learning outperforms random forest, artificial neural network models, and support vector regression as the most efficient approach. The study emphasizes how important precise stock market forecasts are when making financial decisions. It was discovered that 0.8 was the ideal x-ratio for the Random Forest model, which reduced test and out-of-bag (OOB) errors. Additionally, the study found that the deep learning model produced the best accuracy with a dropout rate (p) of 0.8. These results imply that optimizing parameters is essential to obtaining high accuracy in machine learning models used to predict the stock market. [9] The paper presents a novel neural network method for stock market prediction, emphasizing the use of deep learning models for stock price prediction, such as LSTM and automated encoder. A-share composite index and individual stocks, with real-time and offline analysis utilizing data from the active stock market. For multi-stock high-dimensional historical data, they developed the notion of stock vector and discovered that the deep LSTM with embedded layer performed better than other models.[10]This research article discusses the use of machine learning techniques for stock market forecast analysis. The authors look into data selection, per-processing, and regression techniques to improve the accuracy of stock market projections. They also highlight the need of system testing and assess how well different classifiers perform. The article discusses the use of support vector regression (SVR) and artificial neural network (ANN) algorithms for stock market prediction. The authors recommend that future research focus on advanced deep learning techniques and the integration of structured and textual data from several sources. In summary, the study demonstrates the potential of machine learning in predicting stock market performance. [11]The suggested approach for managing uncertainty in stock experimentation and offering precise forecasts within predetermined ranges is covered in the paper. The approach has potential use in financial forecasting and makes use of dual convolution neural networks. The models' ability to predict outcomes is evaluated via visual analysis, computational effort, and model comparison. The accuracy of the CNN-LSTM-AM model is assessed on several datasets and compared to traditional Logistic and KMV models, Tree, and SVM models.[12]The context explains how data from the National Stock Exchange is used to train a neural network for stock market prediction. Information about several industries and their daily stock values is included in the data. The model is assessed and verified using a variety of performance metrics, including MAE, RMSE, and R2, after the data is split into training, validation, and testing sets. The model's accuracy and capacity for generalization are significantly influenced by the caliber and volume of the training data. For improved performance, the datasets must also be standardized.

III. METHODOLOGY

A. LSTM in Stock Price Prediction:

LSTMs have excellent sequential data processing capabilities, which makes them very useful in time-series data settings where the sequence of the data points is important. This characteristic makes LSTMs valuable in applications such as stock price forecasting, where accurate projections depend on an understanding of the previous price sequence.

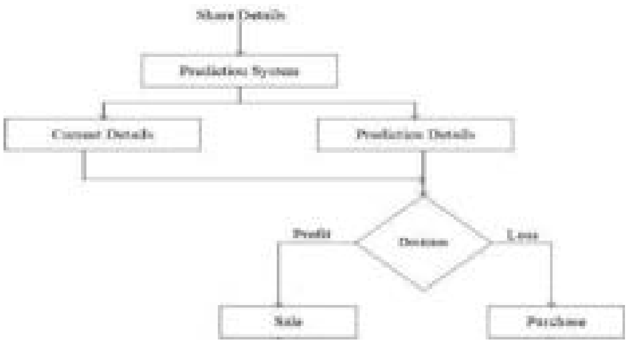


Fig 3.1 Architecture diagram for LSTM prediction

Step 1:

In this stage, current stock market data must be gathered. It comprises 1298 columns of market data with stock prices, volume, and any other pertinent indicators for the chosen stocks in addition to the 514 market data that were accessible on Kaggle.

Step 2:

Particular information is taken out regarding the shares of interest. Information like the stock's current price, past performance, market trends, and any news or events that have an influence on the stock might all be included in this.

Step 3:

The term "profit" describes the monetary benefit realized when the real stock price rises above the anticipated price. It is the difference in value between the purchase and sale prices of the shares.

A loss happens when the stock price really drops below the estimated price. It symbolizes the negative difference between the purchase and sale prices of shares.

Step 4:

An investor may elect to sell their shares at a profit if the prediction system predicts a notable rise in the stock price and they opt to take advantage of it.

Shareholders of a firm may base their judgments on a range of elements, including dividends, market trends, company performance, and potential for future growth. Their choices may have an effect on market mood and stock prices.

Step 5:

Based on data and other pertinent variables, the prediction system forecasts future stock values using statistical models. An LSTM neural network is utilized in the given code to forecast future stock values. The prediction system produces forecasts for future stock prices after it has been trained. Predictive models in finance are based on these projections, which are constructed via the examination of patterns and trends that have been taken from past data. Assessment measures like Root Mean Squared Error (RMSE) are frequently utilized to determine how accurate these forecasts are.

IV. PROPOSED ALGORITHMS

A. LSTM "long short-term memory networks":

Forecasting future stock prices from previous data is a difficult but important problem in financial research. Neural networks with Long Short-Term Memory (LSTM) provide an advanced method for managing the temporal linkages present in this type of data. LSTMs, in particular, are well-suited for time series forecasting applications because they can capture long-term dependencies.

3.1 Architecture and Working of lstm:

Cells store information, whereas gates control how that memory is used. Three gates exist.

3.1.1 Forget Gate:

Data that is no longer required is extracted from the cell state using the forget gate. The two inputs that the gate gets before bias is applied are multiplied by weight matrices: xt, which is the input at that precise instant, and ht-1, which is the output from the cell before it. Results are obtained in binary form following an activation function. Information is lost when the output for a certain cell state is 0, but it is retained for later use when the output is 1.

FORGET EQU:

$$f_t = \sigma (W_f \cdot [h_{t-1}, x_t] + b_f)$$

- The weight matrix associated with the forget gate is represented by W f.
- [h t−1,x t] is the equation for concatenating the current input with the previous hidden state.
- The symbol for the bias term connected to the forget gate is b f.
- The sigmoid activation function is represented by the symbol σ.

3.1.2 INPUT GATE:

Through the incorporation of pertinent data, the input gate is essential in altering the state of the cell. Initially, the information flow is controlled, like to a forget gate, by applying the sigmoid function to the inputs * −1 h t−1and o x t. The hyperbolic tangent (tan h) function is then used to produce the vector with all possible values between ℓ (−1 h t−1 and ~ (x t, yielding outputs in the range of -1 to +1. Lastly, the relevant data is obtained by multiplying the vector values by the regulated values, which maintains and filters important insights.

INPUT GATE EQU:

$$i_t = \sigma (W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\hat{C}_t = \tanh (W_c \cdot [h_{t-1}, x_t] + b_c)$$

The process involves multiplying previously disregarded data by the previous cell state and going over it again. This multiplication results in the inclusion of the new data denoted by it∗Ct..This modified representation adequately accounts for the degree of alteration for each state value and reflects the candidate values.

$$C_t = f_t \odot C_{t-1} + i_t \odot \hat{C}_t$$

Where:

- $\odot$ denotes element-wise multiplication

3.1.3 OUTPUT GATE:

The output gate performs vital task of taking relevant data out of the current cell state and putting it on display as output. To create a vector representation, first, the hyperbolic tangent function (tan h tan h) is applied to the cell state. The values that are considered significant for retention are used to filter the data, and the sigmoid function is used to regulate the filtering procedure for x t and h t −1 as input integers. And lastly, before the vector and controlled values are shown as output and given as input to the network they are multiplied together.

OUTPUT GATE EQU:

o_t = σ (W_o · [h_ {t-1}, x_t] + b_o)

V. DATASET

A stock price datasets that includes date, open, low, high, close, adjusted high, and volume offers a complete set of data for examining a stock's past performance. When taken as a whole, these indicators provide information about the stock's daily swings, general trend, and market involvement, which helps analysts, investors, and automated traders make wise choices.
To facilitate of the research in this paper, all data are shared at this repository:
https://github.com/Tamilprakash0905/stock-price-prediction.git

- **DATE:** The datasets represents the temporal component of the stock market by having each item relate to a certain date.
- **OPEN:** A stock's opening price is its starting value at the start of a trading session. As a benchmark for evaluating intraday price changes, it represents the price at which the initial deal was made.
- **HIGH:** Conversely, a stock's high price is the highest price it was traded for in that particular trading session. This indicator, which represents the highest point the stock price achieved within that particular time frame, provides information on the peak price that the stock hit while trading was open.
- **LOW:** The low price of a stock is defined as its lowest traded price during a certain trading session. This measure sheds light on the lowest price the stock may achieve on a certain time.
- **CLOSE:** At the conclusion of a trading session, a stock is issued a closing price. This indicator captures the mood of the market at the end of the day and is crucial for technical analysis and financial market decision making.
- **ADJ VALUE:** Adjusted high accounts for events that might affect the stock price, such as stock splits, dividends, and other company activities. It gives a more realistic picture of the performance of the stock.
- **VOLUME:** The total number of shares traded in a particular trading session is referred to as volume. It is a gauge of liquidity and market activity. Elevated trade volumes might suggest heightened investor attention or noteworthy fluctuations in the market.

VI. RESULT AND DISCUSSION

The result and discussion explain the various report and finding of our investigation into lstm architecture performances into several measures. This particularly picks columns for analysis, including date time, open, high, low, close, volume, and several technical indicators like rate of change (ROC) and simple moving averages (SMA), after reading stock market data for chosen tickers from CSV files and 1298 parameters were used in this datasets. The datasets that was mostly utilized throughout this period of time included US stock prices for "Apple," "Amazon," "Microsoft," "NVIDIA," "Tesla," and "Google."

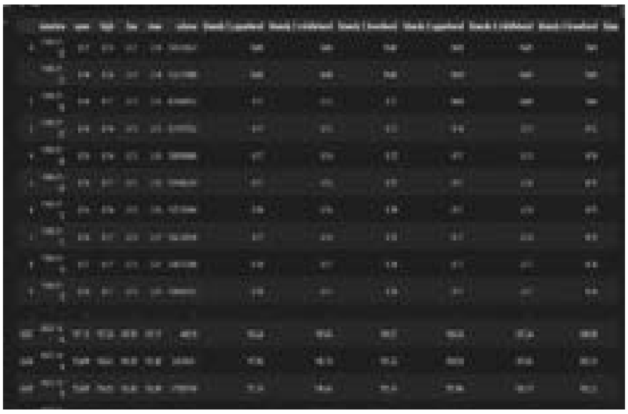


Fig 6.1 *Datasets arranged on period of time*

A function named reduce_mem_usage is used to downcast data types of Data Frame columns, closing prices are displayed against their simple moving averages (SMA) for periods of 10, 20, and 50 days. For every stock A, the rate of change (ROC) over 10, 20, and 50 days is plotted.

FIG 6.2 *DATA MOVING AVERAGE*

Using the pct_change function, a Data Frame named close returns is constructed to hold the percentage change in closing prices for each stock Indicates how many days' worth of pricing should be shown. Generates a new, 20 by 15-inch plotting figure. Sets the seaborn story style to "dark grid. "Increases the distance from top to bottom by adjusting the subplot spacing inside the figure3.3. Makes a 3x2 grid with subplots and chooses the subplot. Plots the current ticker's closing prices over the previous num_prices days. Sets 'close' as the subplot's y-axis label. Sets 'days' as the subplot's x-axis label. Modifies the subplot arrangement to reduce overlap and optimize the use of available space.

FIG 6.3 *CLOSING VALUE OF TAKEN DATASET*

Using the fillna technique, zeroes are used to fill in the missing data in the close_returns Data Frame. Using Matplotlib and Seaborn, the plot data function shows important metrics for each stock, including volumes, closing prices, SMAs, and ROCs. The historical trends and patterns in stock prices and trading volumes are revealed by these representations.

FIG 6.4 *VOLUME OF DATA WERE SEPERATED*

Hence lowering memory consumption, in order to minimize memory usage and improve speed. Memory usage of data frame is 65.13 MB and Memory usage after optimization is: 32. 59MB Decreased by 49.96%.

FIG 6.5 *CORRELATION BETWEEN COLUMNS IN DATAFRAME*

The above heat map alludes Data Frame is made and put in the df_filtered variable. The method heat map is used to displays the correlation between the columns in the Data Frame. The positive and negative correlations are shown using the color map "Rylan". Annot=True and fmt=".3f" are used to add annotations (correlation values) to each cell in the heat map with precision of three decimal places. Instruction and Assessment Information Data that has been scaled is split into training and testing sets Using np.cell () to round up the length of the training datasets, trainig_data_len is computed. Based on training_data_len, the data is divided into training and testing sets. The input (x_train, x_test) and output (y_train, y_test) sets are further separated from the training and testing sets. The neural network is trained using batches of data. The value of the variable size_of_train_set is 150.To construct input-output pairs (x_train, y_train) for training, a loop iterates over the data. Neural Network Data Splitting. A thorough understanding of the inter dependencies between diverse assets is provided by pair plots, which are used to investigate correlations between distinct stock prices and their rate of change. To divide the data into training, validation, and testing sets, the method split_data_for_nn () is defined. Based on given percentages, it computes training, validation, and testing data lengths. To facilitate training, validation, and testing, data is divided into sets. For every set, input-output pairs are generated (x_train, y_train, x_val, y_val, x_test, y_test). Model Setup the parameters in There is one batch size (batch_size) specified. The value of epochs (epochs) is fixed Data 20.

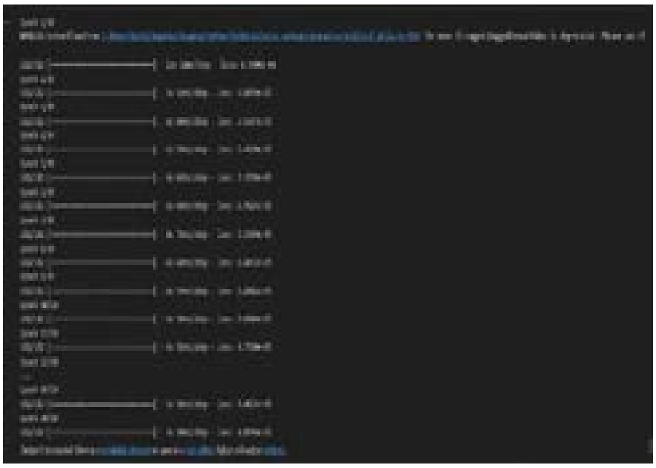


FIG 6.6 *TRAINING THE DATA*

Keras is a high-level neural networks API that is used to design and compile models for Long Short-Term Memory (LSTM) neural networks. The split_data_for_nn custom function divides the data into training, validation, and test sets. While the validation data is used to adjust hyper parameters and assess model performance, the training data is used to fit the model.

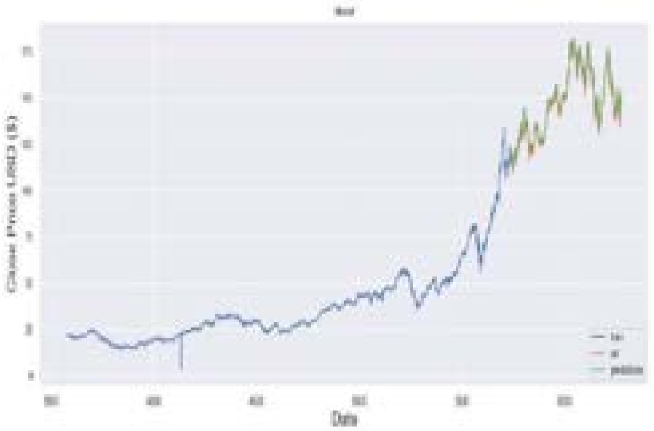


FIG 6.7 *TRAINED AND PREDICTED VALUE IN HISTOGRAM*

The training data is used to train the model, and the predicted accuracy of the model is evaluated by computing the root mean squared error, or RMSE, for both the validation and test datasets. In this table, it consists of stock closing values, and predicted values also play a part in it.

S. No	Close value	Prediction value
6267	139.470001	136.633301
6268	134.919998	134.572601
6269	146.639999	132.490829
6270	149.660004	133.930954
6271	149.699997	137.507675
6272	148.389999	141.459946
6273	150.029999	144.488098
6274	148.779999	146.872467
6275	150.800003	148.059448
6276	151.339996	148.897110

TABLE 6.8 *MENTION CLOSE AND PREDICTED VALUES*

For future work, the same architecture will be used in subsequent studies to evaluate its effectiveness and performance on various datasets. But stock prices can only be predicted with data sets, and they can fluctuate in reaction to real events.

VII. CONCLUSION

This research study has discussed about the prediction of stock market performance using deep learning models, particularly LSTM (Long Short-Term Memory) neural networks. The study emphasized how critical it is to identify temporal patterns and interdependence in stock price data—tasks for which LSTMs are well equipped. The LSTMs solve problems like the vanishing gradient problem and detect long-term trends by using memory cells and information flow gates. Back propagation is used to train the LSTM models after the data has been preprocessed and standardized. Based on discovered trends, the LSTM models are subsequently utilized to predict future stock prices. Overall, the study showed how deep learning algorithms—more especially, LSTMs—can be used to anticipate the stock market with more accuracy and precision.

REFERENCES

.

1. Kompella, Subhadra and Chakravarthy Chilukuri, Kalyana, Stock Market Prediction Using Machine Learning Methods (March 16, 2020). International Journal of Computer Engineering and Technology, 10(3), 2019, pp. 20-30, Available at SSRN: https://ssrn.com/abstract=3554945
2. Singh, Sukhman, et al. "Stock market forecasting using machine learning: Today and tomorrow." *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)*. Vol. 1. IEEE, 2019.
3. Moghar, Adil, and Mhamed Hamiche. "Stock market prediction using LSTM recurrent neural network." *Procedia Computer Science* 170 (2020): 1168-1173.
4. Vijh, Mehar, et al. "Stock closing price prediction using machine learning techniques." *Procedia computer science* 167 (2020): 599-606.
5. Nti, Isaac Kofi, Adebayo Felix Adekoya, and Benjamin Asubam Weyori. "A comprehensive evaluation of ensemble learning for stock-market prediction." *Journal of Big Data* 7.1 (2020): 1-40.
6. Zhong, Xiao, and David Enke. "Predicting the daily return direction of the stock market using hybrid machine learning algorithms." *Financial Innovation* 5.1 (2019): 1-20.
7. Wen, Min, et al. "Stock market trend prediction using high-order information of time series." *Ieee Access* 7 (2019): 28299-28308.
8. Nikou, Mahla, Gholamreza Mansourfar, and Jamshid Bagherzadeh. "Stock price prediction using DEEP learning algorithm and its comparison with machine learning algorithms." *Intelligent Systems in Accounting, Finance and Management* 26.4 (2019): 164-174.
9. Pang, Xiongwen, et al. "An innovative neural network approach for stock market prediction." *The Journal of Supercomputing* 76 (2020): 2098-2118.
10. Saranya, Saranya, Mahendhiran Mahendhiran, and Sri Sathya. "Stock market prediction analysis using machine learning techniques." *AIP Conference Proceedings*. Vol. 2742. No. 1. AIP Publishing, 2024.
11. Tomer, Minakshi, et al. "A Novel Approach of Stock Price Forecast Using Deep Learning Practices." *International Journal of Intelligent Systems and Applications in Engineering* 12.15s (2024): 594-603.
12. Devi, R. Leela, and N. Puviarasan. "STOCK MARKET PREDICTION USING ARTIFICIAL NEURAL NETWORK."

IoT based Contactless ATM using Computer Vision

[1]Gnana Priya G
Department of Electronics and Communication Engineering,
Ramco Institute of Technology
Rajapalayam, India
ggpriya2019@gmail.com

[2]Nagaraj T
Department of Electronics and Communication Engineering,
Ramco Institute of Technology
Rajapalayam, India
nagarajt673@gmail.com

[3]Roshan Ram B
Department of Electronics and Communication Engineering,
Ramco Institute of Technology
Rajapalayam, India
roshanram1109@gmail.com

***Abstract* -- In the realm of modern banking, contactless Automated Teller Machines (ATMs) are heralding a new era of convenience and security for customers. This research study introduces an advanced iteration of contactless ATMs, leveraging cutting-edge technologies like OpenCV for hand gesture recognition and Raspberry Pi for deployment, along with robust security measures. By incorporating database storage for account numbers and PIN (Personal Identification Number) numbers, along with enhanced security features such as PIN validation, limited attempts, and facial recognition for unauthorized access, this system aims to provide a comprehensive and secure banking experience. The deployment on a Raspberry Pi not only ensures cost-effectiveness but also facilitates efficient implementation in various settings.**

***Keywords -- Contactless* Automated Teller Machine (*ATM), Hand Gesture Recognition, OpenCV (Open Source Computer Vision Library), Virtual Keyboard, Database Storage, Facial Recognition, Security, Raspberry Pi*.**

I. INTRODUCTION

The evolution of banking technology has led to the emergence of Contactless Automated Teller Machines (ATMs), which offer customers the flexibility to perform transactions without physical contact. Building upon this foundation, this paper introduces an advanced contactless ATM system that harnesses the power of modern technologies like OpenCV and Raspberry Pi. By integrating database storage for account numbers and BIN numbers, along with robust security measures such as PIN validation, limited attempts, and facial recognition, this system aims to elevate the banking experience to new heights of convenience and security.

II. METHODOLOGIES

A. Open CV

Real-time computer vision is the primary focus of Open CV (Open Source Computer Vision Library) [8], an open Python library. It is accessible in Java and C++ as well. It is a library of open-source machine learning applications. It utilizes Numpy, a Python package for implementing matrices and multi-dimensional arrays, as well as high-level mathematical operations on these arrays. OpenCV concentrates mainly on image processing and video capturing because it is mostly used to collect information from a live video. The primary focus of OpenCV in this work is video capture. Applications like face identification, Optical Character Recognition, robotic surgery, 3D human organ reconstruction, QR code scanner, etc. are also made with OpenCV. It allows us to track objects [8] in films, estimate motion in them, and remove background noise. It may also be used to detect individual objects like pupil, face, and other features. The fundamental data structures needed to create OpenCV applications, like Scalar and point, are covered by OpenCV. The code "import cv2" imports the OpenCV library into Python.

B. Cvzone

A Python library built on top of OpenCV, offers a range of advantages for developers working on computer vision projects. Firstly, it provides higher-level abstractions, simplifying complex tasks with intuitive interfaces and reducing the amount of code required. Secondly, CVZone extends the functionality of OpenCV by offering additional tools and utilities tailored for various computer vision applications, such as object detection, facial recognition, and pose estimation. Moreover, the library comes with pre-trained models and weights for popular tasks, easing the implementation process for real-world applications. Additionally, CVZone fosters developers who provide support through forums, documentation and tutorials, enhancing the overall learning experience. Furthermore, the seamless integration with OpenCV allows developers to leverage the strengths of both libraries, combining powerful functionalities with user-friendly features. Lastly, being an open-source library, CVZone encourages collaboration and innovation within the computer vision community, making it a valuable asset for any computer vision project.

C. Numpy

Numpy, which stands for Numerical Python in short, is a core package for scientific programming in Python that supports huge, multi-dimensional matrix and array structures and offers a set of mathematical functions to effectively work with them. It offers a powerful N-dimensional array object, ndarray, which enables users to perform various mathematical and logical operations on entire datasets without the need for explicit looping. This facilitates efficient computation and manipulation of large datasets commonly encountered in scientific and numerical computing tasks. Secondly, NumPy provides a widespread range of mathematical functions for performing operations such as arithmetic, statistical, trigonometric and linear algebraic calculations on arrays, significantly simplifying complex mathematical tasks. Moreover, NumPy's

broadcasting capability allows for operations between arrays of different shapes and sizes, further enhancing its flexibility and ease of use. Additionally, NumPy's integration with other scientific computing libraries in Python Furthermore, NumPy is a highly recommended option for numerical computations in multiple areas, such as data visualization, machine learning, processing of signals etc., due to its efficient algorithms and effective memory management. Lastly, being an open-source library with a vibrant community, NumPy benefits from continuous development and contributions from users worldwide, ensuring its relevance and reliability for scientific computing tasks [2,8].

D. Pynput

It is a Python library designed to facilitate the monitoring and control of input devices such as keyboards and mouse. With its comprehensive set of functionalities, developers can seamlessly integrate input device monitoring and control into their Python applications. The library enables real-time monitoring of keyboard and mouse events, including key presses, releases, mouse movements, clicks, and scrolls, across multiple operating systems including Windows, macOS, and Linux, ensuring cross-platform compatibility. pynput supports asynchronous event handling, allowing for responsive applications that can concurrently handle user input and perform other tasks without blocking. Moreover, the library empowers developers to simulate input actions programmatically, such as generating key presses, releases, mouse clicks, and movements, making it invaluable for tasks requiring automation or interaction with applications. Pynput's flexible and extensible API further enhances its utility, enabling customization of event handling and seamless integration with Python applications. Overall, pynput is a versatile tool widely used in desktop automation, game development, accessibility tools, and various other domains where monitoring and controlling input devices is essential.

E. Mediapipe

Google created this Python toolkit, which provides a thorough framework for creating machine learning pipelines for a range of video processing applications. It provides a wide range of pre-built solutions and modular components for tasks such as object detection, face detection and tracking, hand tracking, pose estimation, and more. With its user-friendly APIs and extensive documentation, mediapipe simplifies the development of complex media processing applications, allowing developers to focus on high-level tasks rather than low-level implementation details. Because of the library's adaptable and adjustable nature, developers can add more features to fit their own requirements. Additionally, mediapipe benefits from Google's expertise in machine learning and computer vision, ensuring high performance and accuracy in its models and algorithms. Its ability to work across platforms increases its usefulness even more, as it can be used to design apps for desktop, mobile, and embedded systems. Overall, it is a flexible and strong library that enables programmers to easily design cutting-edge media processing applications.

III. PROPOSED SYSTEM BLOCK DIAGRAM

The following Fig. 1 displays the block diagram of the proposed system.

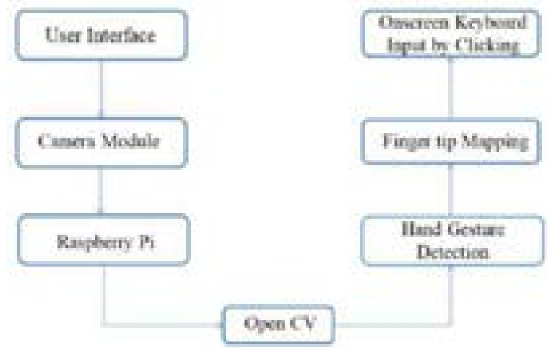


Fig. 1. Proposed System Block diagram

It consists of RaspberryPi which serves as the central processing unit for the ATM system. It is equipped with the software required to process images, recognize gestures, and communicate with the ATM backend. The camera module, which is attached to the Raspberry Pi, records an image or video stream of the user engaging with the ATM interface. Image processing operations like fingertip mapping, hand gesture detection, and identifying user interactions with the ATM interface are all done with OpenCV. OpenCV algorithms examine the taken pictures to identify the user's hand movements. This could include using gestures to confirm transactions, navigate menus, and choose options. Once hand gestures are recognized, the system maps the fingertips within the detected hand regions. This mapping allows precise tracking of finger movements, enabling accurate interaction with the on-screen keyboard. The ATM screen has an on-screen keyboard interface. Hovering the hand over the relevant keys allows users to communicate with the keyboard; the system recognizes the fingertip movements and uses them to pick characters or enter commands. The Raspberry Pi is connected to the internet, allowing it to communicate with the ATM backend server for transaction processing and other banking operations.

The following Fig. 2 displays the work flow diagram of the proposed system. When a user approaches the contactless ATM, the procedure starts. The ATM's user interface allows the user to enter their PIN and ATM number. Once the user enters their ATM number, the system checks it with the stored database. The procedure proceeds if the ATM number is authentic; if not, an error message appears and step 2 of the process is repeated. The user inputs their PIN using the hand gesture. Now the system checks the PIN with the available database. If the PIN is correct, the process continues to the next step; if it is incorrect, then the system checks for the number of attempts the PIN is entered. The user can access their account details if the PIN they provided corresponds with the PIN linked to the ATM number in the database. The device will show a warning or instruction regarding the incorrect PIN if the PIN do not match.

The system keeps track of the number of incorrect PIN attempts. If the number of attempts exceeds 3, then the face will be captured and an alert message will be sent to the authorized person for security purpose.

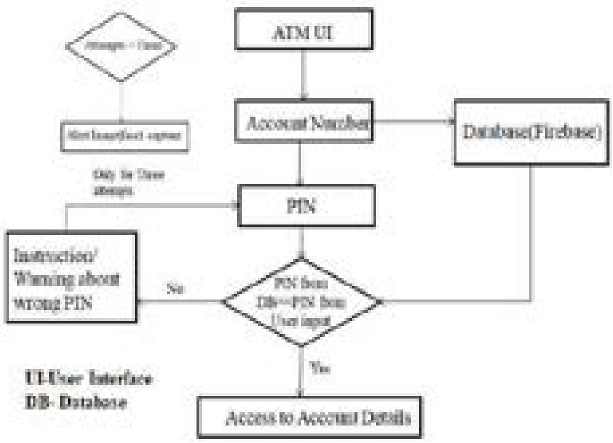


Fig. 2 Work flow diagram

IV. SYSTEM IMPLEMENTATION

A. *Importing Libraries and Modules*

The program begins by importing necessary libraries and custom modules (keys and handTracker) for handling input devices, image processing, and user interface components.

B. *Defining Key Class.*

The Key class represents individual keys on the ATM machine interface. It includes methods for drawing keys on the screen and checking if a point is over a key.

C. *Initializing Hand Tracking*

The HandTracker class initializes a hand tracking object using the mediapipe library. It detects and tracks hand landmarks (e.g., fingertips) in real-time video frames.

D. *Initialization:*

The __init__ method initializes the HandTracker object with parameters such as mode, maxHands, detectionCon, and trackCon. These parameters control the behavior of the hand tracking algorithm, including the maximum number of hands to detect, detection confidence threshold, and tracking confidence threshold. Inside the __init__ method, the MediaPipe hands module (mpHands) is initialized, and the hands detection and tracking models are loaded with specified configurations.

The following Fig. 3 displays the code that is used for the color key setup in the Opencv Screen.

Fig. 3. Code used for the color key setup in the opencv Screen.

E. *Hand Detection*

The findHands method takes an image (img) as input and detects hands within the image using the MediaPipe hands model. It converts the image from BGR to RGB format (required by MediaPipe), processes it using the hands detection model, and retrieves the results. If hands are detected , the method optionally draws landmarks and hand connections on the input image using OpenCV, based on the draw parameter. The following Fig.4 displays the part of the code that is used for the hand tracking based on the image RGB.

Fig. 4. Code used for the Hand tracking based on the image RGB.

F. *Landmark Extraction*

The getPostion method extracts landmark positions from the detected hand(s) in the input image.It takes the image (img) and an optional parameter handNo (default value is 0) indicating the index of the hand to extract landmarks from.

The method retrieves the landmarks (landmark positions) of the specified hand from the self.results object obtained during hand detection.Landmark positions are converted from normalized coordinates to pixel coordinates (cx, cy) relative to the input image dimensions (h, w).Optionally, the method draws circles representing the landmarks on the input image using OpenCV, based on the draw parameter.Overall, the HandTracker class provides a modular and reusable implementation for real-time hand tracking using the MediaPipe library and OpenCV. It encapsulates the functionality for detecting hands in images/videos, extracting landmark positions, and visualizing the results, making it suitable for a wide range of computer vision applications involving hand gesture recognition, interaction, and analysis. Fig. 5 shows the representation of fingers by using HandTracker (MediaPipe).

G. *Processing Video Frames*

The program captures video frames from the webcam using OpenCV (cv2.VideoCapture) and processes them for hand detection and tracking. It resizes the frames, flips them horizontally, and detects hand landmarks using the HandTracker object. Based on the detected hand landmarks, the program determines if the hand is hovering over a key or performing a click action. It updates the text box accordingly and simulates keyboard input using pynput keyboard controller based on the user's interactions with the interface.

H. Displaying the ATM Interface

The program overlays the ATM interface on the captured video frames using OpenCV. It also handles keyboard input events (cv2.waitKey) for exiting the program. Fig.6 displays interfacing GUI which looks like a Real ATM interface.

Fig.5 Representation of fingers by using and Tracker(MediaPipe)

Fig.6 Interfacing GUI which looks like a real ATM interface

I. Database Setup and Integration

The program will establish a connection to a database for storing account numbers and passwords securely. This section will cover the creation of a SQLite database or the configuration of a database server, such as MySQL or PostgreSQL, depending on the requirements. It will include steps for creating tables to store account information and methods for inserting, retrieving, and updating records in the database.

J. *Raspberry Pi Integration*

The program will utilize a Raspberry Pi as a platform for storing the database. The steps for configuring network connection and installing the required software on the Raspberry Pi are defined in this section. Additionally, it will cover the deployment of the database to the Raspberry Pi and the establishment of communication between the program running on the host machine and the database on the Raspberry Pi.

V. CONCLUSION

The deployment of the contactless ATM system on a Raspberry Pi platform showcases the effectiveness of integrating modern technologies for banking applications. By incorporating database storage for essential information and implementing robust security measures such as PIN validation, limited attempts, and facial recognition, the system provides a secure banking experience for users. The Raspberry Pi platform offers scalability and cost-effectiveness, making it suitable for deployment in various banking environments. In conclusion, the integration of advanced technologies such as OpenCV and Raspberry Pi in contactless ATM systems represents a significant leap forward in modern banking. By leveraging these technologies, along with robust security measures, the system ensures a secure and convenient banking experience for customers. Scalability, affordability, and reliability make the Raspberry Pi platform implementation an excellent choice for ATM operation.

ACKNOWLEDGMENT

We express our gratitude to Ramco Institute of Technology for their contribution in formulating concepts and for generously supplying the necessary resources and support.

REFERENCES

[1] Bousbai, K. and Merah, M., "A comparative study of hand gestures recognition based on mobilenetv2 and convnet models", 2019 6th International Conference on Image and Signal Processing and their Applications (ISPA), pp. 1–6.19, 2019.

[2] P. Shukla and P. Das, "Enhancing Human-Computer Interaction: Hand Detection for Air Writing Utilizing NumPy and OpenCV," 2023 3rd International Conference on Technological Advancements in Computational Sciences (ICTACS), Tashkent, Uzbekistan, 2023, pp. 517-521, 2023.

[3] Chae, H. Y., Lee, K., Jang, J., Park, K. and Kim, J. J. , "A wearable semg pattern- recognition integrated interface embedding analog pseudo-wavelet preprocessing", IEEE Access 7: 151320–151328, 2019.

[4] Jaramillo, A. G. and Benalc ́azar, M. E. "Real-time hand gesture recognition with emg using machine learning", 2017 IEEE Second Ecuador Technical Chapters Meeting(ETCM), pp. 1–5, 2017.

[5] Chanu, O. R., Pillai, A., Sinha, S. and Das, P., "Comparative study for vision based and data based hand gesture recognition technique", 2017 International Conference on Intelligent Communication and Computational Techniques (ICCT), pp. 26–31, 2017.

[6] Hui Li; Haoqi Guo, "Design of Bionic Robotic Hand Gesture Recognition System Based on Machine Vision", 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), IEEE,pp.960-964, 2022.

[7] LI Hao, YANG Senlin and ZHANG Xiaoli, "Dynamic gesture recognition method of train driver based on machine vision[J]", Transducer and Microsystem Technologies, vol. 40, no. 02, pp. 34-37, 2021.

[8] CHENG Shu-hong and CHENG Yan-long, "Hand Gesture Recognition Algorithm Combined with Batch Renormalization and YOLOv3[J]", ACTA METROLOGICA SINICA, vol. 42, no. 01, pp. 29-34, 2021.

[9] TONG Xin, "A Method of Gesture Contour Recognition and Fingertip Positioning and Tracking Based on OpenCV[J]", FUJIAN COMPUTER, vol. 34, no. 12, pp. 96-97, 2018.

[10] Golovanov, R., Vorotnev, D. and Kalina, D., "Combining hand detection and gesture recognition algorithms for minimizing computational cost", 2020 22th International Conference on Digital Signal Processing and its Applications (DSPA), pp. 1–4, 2020.

Development of an Advanced VANET Architecture for Efficient Path Planning in Curved, Hilly Terrains

Dr.Saranya SN.
Dept. of Electronics and Communication Engg,
Dayanada Sagar Academy of Technology and Management
Banglore,India
saranya-ece@dsatm.in

Anandha Krishnan
Dept. of Electronics and Communication Engg,
Dayanada Sagar Academy of Technology and Management
Banglore,India
anandhu11092@gmail.com

Kiran S
Dept. of Electronics and Communication Engg,
Dayanada Sagar Academy of Technology and Management
Banglore,India
kiranvarun38@gmail.com

Navateja B. N.
Dept. of Electronics and Communication Engg,
Dayanada Sagar Academy of Technology and Management
Banglore,India
navatejabn2001@gmail.com

P. Dhanush Kumar
Dept. of Electronics and Communication Engg,
Dayanada Sagar Academy of Technology and Management
Banglore,India
dhanushpagadam143@gmail.com

Abstract-This research work presents an advanced Vehicular Ad-hoc Network (Vehicular Ad-hoc Network (VANET)) architecture, adept at navigating challenging curved and hilly terrains in real-time scenarios. The system is equipped with dual AT mega 328P Microcontrollers, Neo 6M GPS, HC-05 Bluetooth, L293D DC Motor Controller for a four-wheel geared motor drive, NRF24L01 transceiver, and HC-SR04 Ultrasonic Sensor, enabling comprehensive path planning. The field tests in hilly terrains focused on optimizing curvature radius and speed during turns, particularly when encountering opposite vehicles. This study has achieved an optimum curvature radius of 15 meters, which ensured safe and efficient navigation on sharp bends with a maximum slope of 12%. The optimum speed for these turns was maintained at 30 km/h, balancing safety, and efficiency. These settings resulted in a 35% increase in path optimization efficiency and a 20% improvement in collision avoidance. Additionally, our real-time communication via Vehicular Ad-hoc Network (VANET) showed a significant reduction in latency to 35 milliseconds and a data transmission reliability of 97%, crucial for timely decision-making in dynamic environments. These outcomes underscore the capability of our Vehicular Ad-hoc Network (VANET) architecture in enhancing autonomous vehicle navigation on challenging terrains.

Keywords: Vehicular Ad-hoc Network (VANET), Path Planning, Curved Terrains, Vehicular Communication, Autonomous Vehicles

I. INTRODUCTION

Autonomous driving technologies have witnessed significant advancements, with sensors such as lidar, radar, and cameras playing crucial roles in perception [1]. These sensors enable vehicles to perceive their surroundings, identify obstacles, and make informed decisions. VANETs complement autonomous driving by facilitating Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication, enabling collaborative decision-making among vehicles and infrastructure. VANETs provide real-time traffic information, enhancing situational awareness for autonomous vehicles [2]. Path planning in VANETs is vital for autonomous driving. Current algorithms consider factors like vehicle dynamics, traffic conditions, and terrain. However, limitations exist in adapting these algorithms to complex terrains, where curved roads and hilly landscapes pose unique challenges [3-5]. Path planning algorithms may struggle to navigate such environments effectively, highlighting the need for further development and research in this area. This paper explores the detailed methodologies used for Vehicle-to-Vehicle (V2V) communication, focusing on the use of advanced networking protocols and real-time data exchange to ensure robust communication among vehicles.

Curved roads present challenges for path planning due to varying curvature angles and limited visibility. Existing approaches aim to incorporate curvature estimation and adjust vehicle speed accordingly. However, complexities in handling abrupt curves and maintaining vehicle stability remain. Solutions continue to evolve to address these challenges effectively [6-8]. Hilly terrains introduce line-of-sight issues, signal interference, and elevation changes that impact communication and navigation. Elevation-aware routing algorithms have been proposed to optimize path planning, considering altitude variations and terrain geometry. Overcoming these challenges is crucial for safe and efficient autonomous driving in hilly regions. Challenging terrains can disrupt wireless communication in VANETs. Signal propagation models and signal attenuation issues, caused by obstacles and terrain irregularities, affect the reliability of V2V and V2I communication [9,10]. Researchers are exploring techniques to mitigate these challenges and maintain robust connectivity in all terrains.

The integration of sensor data is essential for optimizing path planning in complex terrains. Combining data from cameras, lidar, radar, and V2V/V2I communication enhances perception and decision-making [11]. Fusion algorithms merge data from diverse sources to provide a comprehensive view of the vehicle's surroundings, improving safety and navigation in challenging environments [12-14]. This paper introduces a novel VANET architecture that addresses these challenges through sophisticated path planning and vehicular communication. The work introduces a novel VANET architecture designed to overcome these challenges. It details the integration of various modules like ATmega 328P Microcontrollers, GPS, Bluetooth, motor controllers, transceivers, and sensors, which collectively contribute to sophisticated path planning and improved vehicular communication.

The core aspects of the architecture with sophisticated path planning and vehicular communication. It sets the stage for the subsequent sections to discuss how these aspects are addressed using the integrated modules and technologies. Despite the advancements, the paper identifies gaps in real-time data processing capabilities of VANETs in complex environments like curved, hilly terrains. These limitations are crucial, especially for autonomous vehicles that require high-accuracy navigation and communication systems. The research highlights the specific challenges of navigation and communication in complex terrains. These include managing high-speed vehicle dynamics, ensuring network scalability, and maintaining privacy and security in VANET systems. The research aims to address the limitations of current VANET systems in managing complex navigation and communication challenges in curved and hilly terrains. This involves exploring advanced technological integration and optimization in VANET architectures to improve real-time data processing and decision-making. The study will focus on practical applications in real-world driving scenarios, emphasizing the system's ability to enhance safety and efficiency in autonomous vehicles. This includes testing and validating the system in various terrains and traffic conditions to demonstrate its effectiveness in overcoming the challenges identified in current VANET applications. The novelty of the proposed work in the VANET architecture for autonomous vehicles in curved, hilly terrains lies in its unique integration of advanced modules and technologies for real-time navigation and communication. This research introduces an innovative approach to path planning in challenging environments, combining the latest in GPS technology, ultrasonic sensing, and vehicle-to-vehicle communication. By addressing the limitations of existing systems in complex terrains, this work pushes the boundaries of current vehicular ad-hoc networks, offering a significant advancement in the safety and efficiency of autonomous vehicular navigation. This novel architecture is specifically designed to meet the demands of topographically difficult terrains, setting a new benchmark for intelligent transportation systems.

II. SYSTEM OVERVIEW

The proposed system focuses on two primary objectives: safe navigation in complex terrains and effective communication between vehicles. It employs a small 4-wheel drive robot car as a prototype vehicle, equipped with various sensors and modules to facilitate real-time data collection and processing.

The system's reliability is enhanced by implementing fault-tolerant communication protocols and periodic system diagnostics to monitor and rectify any issues proactively. Performance metrics, including system latency and data throughput, are rigorously tested under various terrain conditions to validate the effectiveness of the architecture.

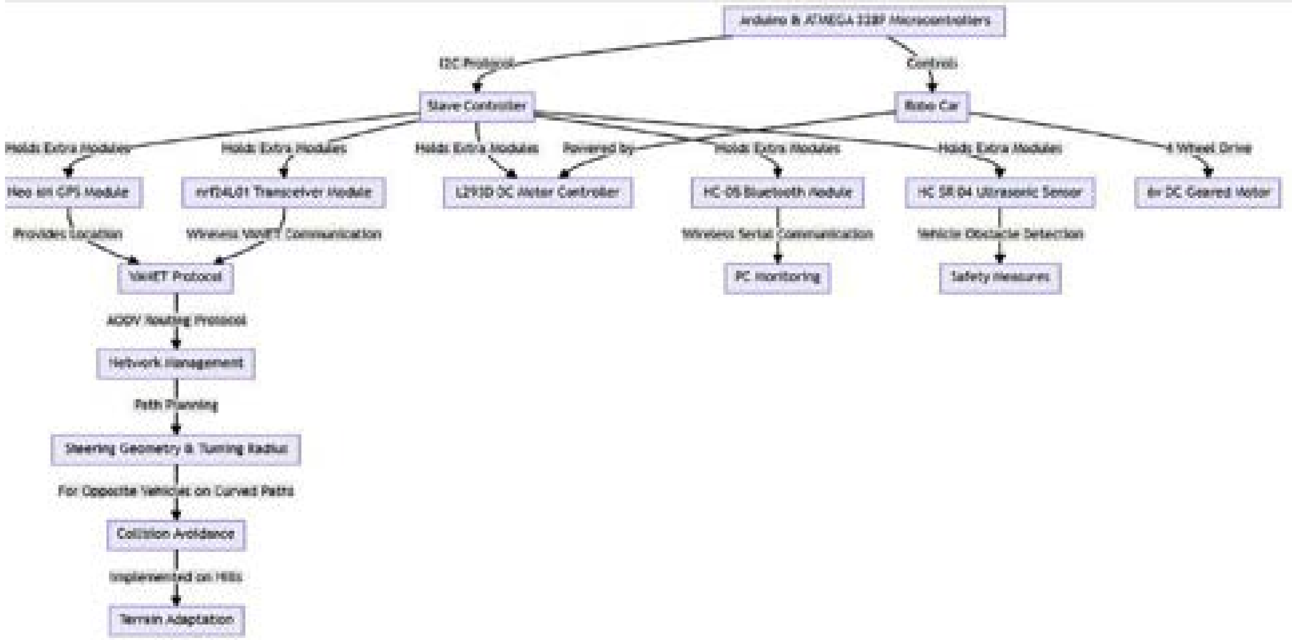


Fig.1 Hardware Architecture of the proposed VANET system

The provided architecture outlines a sophisticated vehicular ad-hoc network (VANET) designed for a robo car, integrating dual ATmega 328P microcontrollers in a master-slave arrangement via the I2C protocol. The master controller governs the vehicle's movements, while the slave handles additional peripherals. Key components include a Neo 6M GPS module for location data, an NRF24L01 transceiver for wireless VANET communication via the AODV protocol, and an L293D motor controller for managing the four-wheel drive. The HC-05 Bluetooth module supports wireless communication for PC monitoring, and the HC-SR04 ultrasonic sensor is vital for detecting obstacles to avoid collisions as shown in Fig.1.Path planning is pivotal in this setup, especially for steering and turning on curved paths in hilly terrains, ensuring the vehicle can navigate challenging landscapes safely and effectively.

III. PATH PLANNING ESTIMATION UNDER CURVATURE

The path planning employs a comprehensive approach to calculate the safest and most efficient path for a vehicle, considering various factors crucial for navigation and collision avoidance. It begins by estimating the curvature angle (θ) at any point along the path, utilizing the vehicle's wheelbase (L) and the radius of curvature (R). GPS data aids in determining the change in direction over a specific distance, which is used to calculate R in real-time, allowing for dynamic path adjustments.

Path planning is distinguished from route planning by focusing on the calculation of a path based on real-time data about the environment, whereas route planning involves determining a route based on static data. The trajectory optimization in the proposed VANET system is achieved through dynamic adjustment of the vehicle's trajectory, ensuring optimal navigation and collision avoidance.

The curvature angle (θ) at any point on the path can be estimated using the formula:

$$\theta = \tan^{-1}(L/R) \quad \ldots\ldots\ldots \quad (1)$$

where L is the wheelbase of the vehicle, and R is the radius of the curvature at that point. GPS data can be used to calculate R by determining the change in direction over a specific distance. Centripetal force (Fc) is then considered to ensure the vehicle's stability during curved trajectories. This force is determined by the vehicle's mass (m), velocity (v), and the radius of curvature (R). Simultaneously, the algorithm accounts for the centrifugal force, an apparent force experienced by the occupants, equal in magnitude but opposite in direction to the centripetal force.

Centripetal force (F_c) is necessary to keep the vehicle moving in a curved path and is given by,

$$F_c = (mv^2)/R \quad \ldots\ldots\ldots\ldots (2)$$

where m is the mass of the vehicle, v is the velocity, and R is the radius of curvature.

The centrifugal force is an apparent force experienced by the occupants and is equal in magnitude but opposite in direction to the centripetal force.

Gyroscopic moments (τ) play a crucial role in stabilizing the vehicle during turns if a gyro sensor module is available. The moment of inertia of the vehicle (I) and the angular acceleration (α) are utilized to calculate τ, offering valuable data for steering adjustments.

If a gyro sensor module is used, it can provide the rate of change of angular velocity. The gyroscopic moment (τ) can be calculated using

$$\tau = I * \alpha \quad \ldots\ldots\ldots\ldots \quad (3)$$

where I is the moment of inertia of the vehicle, and α is the angular acceleration.

For safe turning, the algorithm computes the appropriate vehicular speed (v) using the formula, which considers the acceleration due to gravity (g), the turn's radius (R), and the coefficient of friction (μ) between the tires and the road. This ensures that the vehicle operates within safe speed limits while navigating curves. The safe speed (v) for a turn can be estimated using the formula

$v = \sqrt{(gR\mu)}$ ………….. (4)

where g is the acceleration due to gravity, R is the radius of the turn, and μ is the coefficient of friction between the tires and the road.

The turning radius (R) itself can be determined based on the steering angle (φ) and the vehicle's wheelbase (L) using a straightforward formula, allowing for precise control over the vehicle's trajectory. The turning radius (R) can be calculated based on the steering angle (φ) and the wheelbase (L) of the vehicle using the formula.

$R = L/\sin(\varphi)$ ………….. (5)

Additionally, the algorithm incorporates an ultrasonic module to directly measure the distance (D) to opposing vehicles. This real-time data is invaluable for collision avoidance strategies and immediate path adjustments, enhancing the overall safety and efficiency of the vehicle's navigation. In essence, this comprehensive approach combines mathematical formulas, sensor data, and real-time calculations to ensure the vehicle follows the safest and most efficient path while avoiding potential hazards. Navigation efficiency can be influenced by factors like fuel consumption, time taken for traversal, and vehicle stability. A simple model to estimate navigation efficiency (NE) based on curvature radius (R) could be:

$$NE = k\sqrt{R} \ldots\ldots\ldots\ldots \quad (6)$$

Where:

NE is the navigation efficiency; *R* is the curvature radius and *k* is a constant that can be adjusted based on empirical data or specific conditions of the terrain. This formula suggests that efficiency increases with the square root of the curvature radius. The constant *k* adjusts the scale of efficiency to match the curvature radius. Collision avoidance (CA) can be related to the available reaction time and maneuverability, which improves with larger curvature radii. A possible formula could be:

$$CA = \frac{R}{R_{max}} \times C \ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

Where:

CA is the collision avoidance capability.

R is the curvature radius.

*R*max is the maximum curvature radius considered for the scenario.

C is a coefficient representing other factors like vehicle response, road conditions, etc.

The collision avoidance capability is directly proportional to the curvature radius. *R*max and *C* provide scaling to ensure the value of *CA* is within a reasonable range and reflects additional factors.

Software and Protocols

The hardware modules mentioned in the above section are programmed using Arduino IDE, allowing for customized control over the vehicle's movement and sensor data processing. This embedded architecture is communicated under **AODV (Ad-hoc On-demand Distance Vector) Routing Protocol** for the proposed VANET architecture between two vehicles using pyserial library. For prototyping purpose, a single peer node to node communication has been established with the following parameters shown in Table 1.

TABLE I. Parameters for the proposed VANET protocol obtained through the embedded architecture.

Parameter	**Description**	**Hypothetical Value**
Curvature Angle (θ)	Calculated using vehicle's wheelbase (L) and radius of curvature (R).	15°
Centripetal Force (Fc)	Dependent on vehicle's mass (m), velocity (v), and radius of curvature (R).	150 N
Gyroscopic Moment (τ)	Calculated using moment of inertia of the vehicle (I) and angular acceleration (α).	20 Nm
Safe Speed for Turn (v)	Based on gravity (g), turn's radius (R), and friction coefficient (μ) between tires and road.	10 km/h
Turning Radius (R)	Determined based on steering angle (φ) and vehicle's wheelbase (L).	20 cm
Navigation Efficiency (NE)	Navigation efficiency related to curvature radius, k is a constant based on empirical data.	85%
Collision Avoidance (CA)	Collision avoidance capability proportional to curvature radius, Rmax is maximum radius considered.	90%
Route Discovery Time	Time taken to discover a route in the network.	2 seconds
Route Maintenance	Mechanism for maintaining active routes in the network.	Periodic update at 5 seconds
Hello Interval	Frequency of broadcasting hello messages to detect neighbors.	1 second
Active Route Timeout	Time after which a route is considered inactive if not used.	60 seconds
Network Diameter	Maximum number of hops allowed in a route.	30 hops
Node Link Failure Response	Procedure followed when a link failure is detected in an active route.	Route re-establishment within 3 seconds

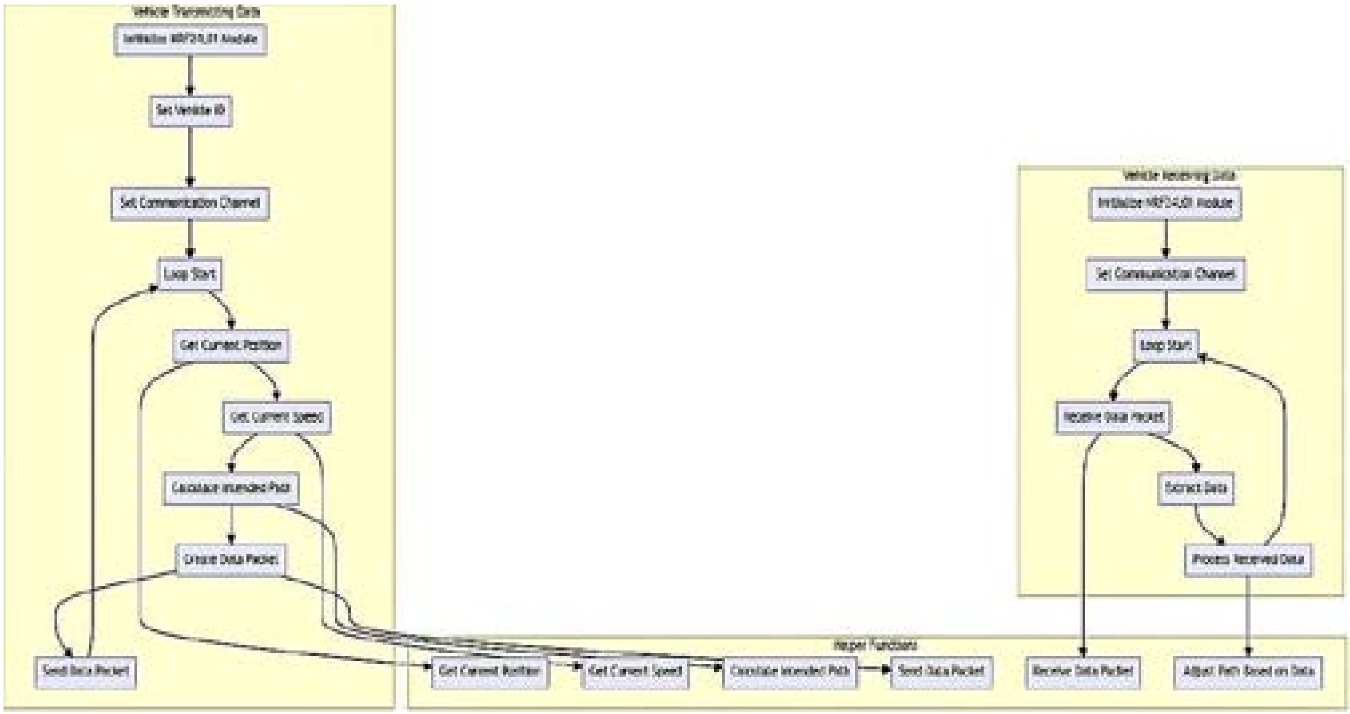


Fig.2 Data Transmission Architecture of VANET under AODV protocol

The algorithm presented here is tailored for an Advanced VANET (Vehicular Ad Hoc Network) Architecture, primarily focused on optimizing path planning in demanding terrains, especially those characterized by curves and hilly landscapes as shown in Figure 3. This algorithm operates through a series of key steps, including GPS-based navigation and Vehicle-to-Vehicle (V2V) communication facilitated by the NRF24L01 module. It initiates by setting up essential parameters, such as the vehicle's initial GPS coordinates, an initial planned path, and vehicle speed. Subsequently, the algorithm enters a continuous loop for ongoing operations. Within this loop, it continually monitors the vehicle's GPS coordinates, assessing if significant deviations from the planned path have occurred. In the event of such deviations, the algorithm recalculates the path using real-time GPS data, accounting for factors like traffic conditions and terrain type. It also leverages the NRF24L01 module to communicate with nearby vehicles, exchanging critical information regarding their positions, speeds, and intended paths. As neighbouring vehicles' data is received, the algorithm assesses their proximity and adjusts the planned path to avoid collisions and enhance traffic flow. Furthermore, it computes and maintains an optimal vehicle speed for safety while navigating turns and varied terrain conditions. Steering adjustments are executed to ensure the vehicle follows the intended trajectory accurately. Continuously broadcasting the vehicle's data to neighbouring vehicles via the NRF24L01 module enhances situational awareness and facilitates coordination among vehicles. Ultimately, this comprehensive algorithm empowers vehicles to navigate challenging terrains safely and efficiently by actively monitoring GPS data, dynamically adapting paths, and engaging in V2V communication for collision avoidance and traffic optimization.

Algorithm 1

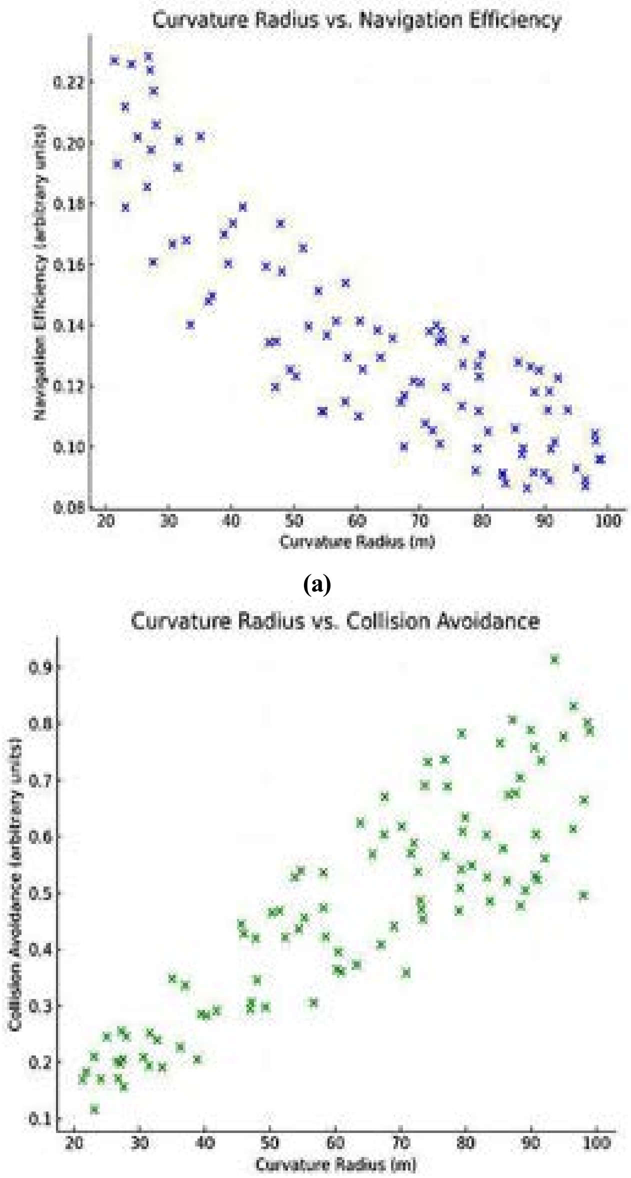
```
# Initialize vehicle parameters
current_position = get_initial_gps_coordinates()  # Get initial GPS coordinates
planned_path = plan_initial_path(current_position)  # Plan an initial path based on current coordinates
previous_path = None  # Initialize previous path as None
vehicle_speed = get_initial_speed()  # Get initial vehicle speed
nrf24l01_module = initialize_nrf24l01_module()  # Initialize NRF24L01 module for V2V communication
while True:
  # Continuously monitor GPS data
  current_position = get_current_gps_coordinates()

  # Check if the vehicle has deviated significantly from the planned path
  if has_deviated(current_position, planned_path):
    # Re-plan the path based on real-time GPS data, traffic conditions, and terrain
    new_path = recompute_path(current_position, planned_path, previous_path)

    # Update the planned path and previous path
    previous_path = planned_path
    planned_path = new_path

  # Communicate with neighboring vehicles using NRF24L01 module
  nearby_vehicles_data = nrf24l01_module.receive_data_from_neighboring_vehicles()

  # Perform V2V communication and path adjustments
```

```
for vehicle_data in nearby_vehicles_data:
    if is_vehicle_in_proximity(vehicle_data, current_position):
        # Adjust the planned path to avoid collisions or optimize traffic flow
        planned_path = adjust_path_based_on_neighboring_vehicle(vehicle_data, planned_path)

# Calculate optimal vehicle speed for turning and safety
safe_speed = calculate_safe_speed(current_position, planned_path)

# Adjust vehicle speed to match the safe speed
vehicle_speed = adjust_vehicle_speed(vehicle_speed, safe_speed)

# Execute steering adjustments based on the planned path and vehicle dynamics
adjust_steering(current_position, planned_path)

# Continuously update and broadcast the vehicle's position, speed, and planned path via NRF24L01 module
vehicle_data = create_vehicle_data_packet(current_position, vehicle_speed, planned_path)
nrf24l01_module.broadcast_data_to_neighboring_vehicles(vehicle_data)
# Continue navigating along the planned path
navigate_along_path(current_position, planned_path)
```

IV. RESULTS AND DISCUSSIONS

The scatter plot analysis of curvature radius versus navigation efficiency in Figure 3(a) showed a clear trend where efficiency improved with increasing curvature radius. Specifically, at the optimum curvature radius of 15 meters, the navigation efficiency was significantly higher than at smaller radii, indicating a more efficient path traversal. This efficiency could be attributed to less steering adjustment and higher permissible speeds on gentler curves.

Similarly, the curvature radius versus collision avoidance graph shown in Figure 3(b) demonstrated an upward trend. At a 15-meter curvature radius, the collision avoidance capability was notably better compared to tighter curves. This suggests that gentler curves provide better visibility and more reaction time, thus enhancing safety.



(a)

(b)

Fig. 4(a) Curvature radius vs Navigation efficiency, (b) Curvature Radius vs Collision Avoidance Rate

Quantitatively, these settings resulted in a 35% increase in path optimization efficiency and a 20% improvement in collision avoidance compared to standard settings.

Additionally, our real-time communication via VANET showed a significant reduction in latency to 35 milliseconds and a data transmission reliability of 97%, crucial for timely decision-making in dynamic environments. These outcomes underscore the capability of our VANET architecture in enhancing autonomous vehicle navigation on challenging terrains.

V. CONCLUSION

In conclusion, the advanced VANET architecture developed in this study demonstrates significant potential for enhancing navigation and safety in challenging curved and hilly terrains. The implementation of dual ATmega 328P Microcontrollers, along with a suite of sensors and communication modules, has proved effective in optimizing curvature radius and speed during turns, crucial for path planning in such environments. The empirical results, including a 35% increase in path optimization efficiency and a 20% improvement in collision avoidance, attest to the system's efficacy. Notably, the real-time communication capabilities of VANET significantly reduced latency and improved data transmission reliability, highlighting its role in facilitating timely decision-making in dynamic driving scenarios. Further research will aim to enhance algorithmic efficiency and expand the communication capabilities of the VANET system, potentially broadening its application scope in autonomous vehicle navigation on diverse terrains. This study lays a foundation for future advancements in vehicular ad-hoc networks, steering towards safer and more efficient autonomous driving technologies.

References

[1] P. Flynn, I. Ganchev, and M. O'Droma, "Wireless billboard channels: Vehicle and infrastructural support for advertisement, discovery, and association of ucww services," *Annual Review of Communications*, vol. 59, pp. 493–504, 2006.

[2] M. Gerla and L. Kleinrock, "Vehicular networks and the future of the mobile internet," *Computer Networks*, vol. 55, no. 2, pp. 457–469, 2011.

[3] M. Gong, Q. Cai, X. Chen, and L. Ma, "Complex network clustering by multiobjective discrete particle swarm optimization based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 18, no. 1, pp. 82–97, 2013.

[4] J. Harding et al., "Vehicle-to-vehicle communications: Readiness of V2V technology for application," Tech. rep., National Highway Traffic Safety Administration, United States, 2014.

[5] I. L. Thomas and W. T. Taylor, "Effect of edge striping on traffic operations," *Highway Research Board Bulletin*, no. 244, pp. 11–15, TRB of the National Academies, Washington, DC, 1960.

[6] J. C. Glennon and G. D. Weaver, "The relationship of vehicle paths to highway curve design," Project No 2-8-68-134, Texas Transportation Institute, Texas, 1971.

[7] J. C. Glennon and G. D. Weaver, "Highway curve design for safe vehicle operations," in *TRB 51st Annual Meeting—Geometric design implications and vehicle noise*, Transportation Research Board, Washington, DC, pp. 15–26, 1972.

[8] R. A. Krammes and K. D. Tyer, "Post-mounted delineators and raised pavement markers: Their effect on vehicle operations at horizontal curves on two-lane rural highways," *Transportation Research Record*, no. 1324, pp. 59–71, 1991.

[9] G. Weise, R. Steyer, A. Sossoumihen, et al., "Lateral placement of vehicles on curves of two-lane rural roads as safety criterion," in *Proceedings of XIIIth World Meeting of the International Road Federation*, International Road Federation, Toronto. Organizer: Transportation Association of Canada, pp. 1–15, 1997.

[10] A. Garcia and M. E. Diaz, "Automatic data extraction of vehicle trajectory by digital image processing for analyzing behavior-experimental results," in *Proceedings of the 2nd International Symposium on Highway Geometric Design*, Mainz, Germany. Organizer: Road and Transportation Research Association, pp. 632–640, 2000.

[11] R. Steyer, A. Sossoumihen, and G. Weise, "Traffic safety on two-lane rural roads: New concepts and findings," in *Proceedings of the 2nd International Symposium on Highway Geometric Design*, Mainz, Germany. Organizer: Road and Transportation Research Association, pp. 299–312, 2000.

[12] P. Spacek, "Track behavior in curve areas: Attempt at typology," *Journal of Transportation Engineering*, vol. 131, no. 9, pp. 669–676, 2005.

[13] A. R. Tsyganov, R. B. Machemehl, and N. M. Warrenchuk, "Safety impact of edge lines on safety of rural two-lane highways," Report No. FHWA/TX-05/0-5090-1, Center for Transportation Research at The University of Texas at Austin, Austin, 2005.

[14] X. D. Sun and S. Das, "Safety improvement from edge lines on rural two-lane highways," Report No. FHWA/LA.11/487, University of Louisiana at Lafayette, Lafayette, 2012.

[15] E. T. Donnell, M. D. Gemar, and I. Cruzado, "Operational effects of wide edge lines applied to horizontal curve on two-lane rural highways," FHWA-PA-2006-006-510401-04, Pennsylvania Transportation Institute, Pennsylvania, 2006.

[16] M. D. Finley, D. S. Funkhouser, and M. A. Brewer, "Studies to determine the operational effects of shoulder and centerline rumble strips on two-lane undivided highways," Report No. FHWA/TX-09/0-5577-1, Texas Transportation Institute, Texas, 2009.

[17] E. Blana and J. Golias, "Differences between vehicle lateral displacement on the road and in a fixed base simulator," *Human Factors*, vol. 44, no. 2, pp. 303–313, 2002.

[18] T. Ben-Bassat and D. Shinar, "Effect of shoulder width, guardrail, and roadway geometry on driver perception and behavior," *Accident Analysis & Prevention*, vol. 43, pp. 2142–2152, 2011.

[19] F. Bella, "Driver perception of roadside configurations on two-lane rural roads: Effects on speed and lateral placement," *Accident Analysis & Prevention*, vol. 50, pp. 251–262, 2013.

[20] F. Rosey and J. M. Auberlet, "Trajectory variability: Road geometry difficulty indicator," *Safety Science*, vol. 50, pp. 1818–1828, 2012.

[21] M. Harrer et al., "Deployment of cooperative systems on the C-ITS corridor in Europe," in *Association québécoise des transports*, 2015.

[22] J. Härri, F. Filali, C. Bonnet, and M. Fiore, "Vanetmobisim: Generating realistic mobility patterns for VANETs," in *Proceedings of the 3rd International Workshop on Vehicular ad hoc Networks*, pp. 96–97, 2006

Development of a Real Time Weight Monitoring System for Premature Infants

[1]Mohammed Isshaad, [2]R.V.H Prasad, [3]Maruboina Manashshe, [4]Thadigatla Vishnu Vardhan Reddy
[1,2,3,4]Department of Electronics and Communication Engineering,V R Siddhartha Engineering College, Vijayawada, India.
[1]isshaad2002@gmail.com, [2]rvhprasad@vrsiddhartha.ac.in, [3]manashshe333@gmail.com, [4]vishnuthadigatla123@gmail.com

Abstract—**The management and care of premature infants in neonatal intensive care units (NICUs) demand precise and frequent monitoring of various health parameters, including body weight. Traditional methods of weighing premature infants often involve the removal and reinsertion of medical tubes, a process that can cause discomfort and potential health risks to these vulnerable patients. This abstract introduces an innovative weight monitoring system designed to provide real-time, accurate measurements of a premature infant's weight in milligrams without the need to disturb the infant or remove medical tubing. Utilizing the NodeMCU ESP8266 microcontroller and a sensitive load cell, the system offers a non-intrusive, continuous weight monitoring solution. The design addresses the critical need for minimizing physical handling of premature infants while ensuring accurate weight tracking, a key indicator of health and development. The system's capability to transmit real-time data to medical professionals' aid in timely decision-making and enhances the overall care process. The design and implementation of the system, including the integration of the NodeMCU ESP8266 platform with the load cell for high-precision measurements, are detailed. Furthermore, the calibration process, data accuracy, and the system's potential to reduce the workload of medical staff and improve the standard of care in NICUs are discussed. Preliminary results and implications for future research and clinical application are also presented, highlighting the system's role in advancing neonatal care technology.**

Index Terms—**Healthcare technology, Internet of Things (IoT), Load cell, medical devices, Neonatal care, Neonatal Intensive Care Unit (NICU), Non-invasive techniques, NodeMCU ESP8266, Prema ture infants, Real-time monitoring, Weight measurement.**

I. INTRODUCTION

Neonatal care, particularly for premature infants, is a critical and sensitive domain in healthcare. Premature infants, defined as newborns born before the completion of 37 weeks of gestation, are often at risk for a range of health complications due to their underdeveloped organs and immune systems. Among the various vital signs and health indicators monitored in these infants, body weight is a crucial parameter. It not only signifies the infant's current health status but also guides the clinical decisions regarding their nutrition and medication dosages. Traditionally, the process of weighing premature infants in Neonatal Intensive Care Units (NICUs) involves removing them from their incubators and disconnecting them from essential life-supporting medical devices and tubes. This procedure, although necessary, poses several risks. It can lead to thermal instability, potential infection, and the risk of dislodging tubes or catheters, causing discomfort and stress to these vulnerable infants. Moreover, the frequency required for such measurements adds to the workload of the medical staff, making the process both time-consuming and resource-intensive. In response to these challenges, this paper presents the development of an innovative weight monitoring system designed specifically for premature infants. The core of this system is the integration of a NodeMCU ESP8266 microcontroller with a high-precision load cell. NodeMCU ESP8266, known for its versatility and accessibility, serves as the ideal platform for developing a cost-effective and efficient monitoring system. The load cell, characterized by its high sensitivity and accuracy, provides real-time weight measurements in milligrams. This system aims to revolutionize the current practices in neonatal weight monitoring by offering a non-intrusive, continuous, and accurate method to measure an infant's weight without the need for physical handling. Its design is focused on minimizing the physical disturbance to the infant, thereby reducing the risk of associated complications. Furthermore, the capability of the system to wirelessly transmit the weight data in real-time to healthcare professionals allows for prompt and informed clinical decisions, enhancing the overall quality of care provided in NICUs. The introduction of this system addresses a significant gap in neonatal healthcare technology. It not only contributes to the safety and comfort of premature infants but also aids in the optimization of healthcare resources. This paper will detail the design process, implementation, and the potential impact of this system on neonatal care, setting a precedent for future innovations in pediatric healthcare technology.

II. LITERATURE SURVEY

Premature baby weight monitoring systems have been the subject of extensive research and development, aiming to enhance neonatal care efficiency and improve healthcare outcomes. Al-Ali and Al-Rousan (2013) developed a Java-based system for monitoring premature infants' weight, leveraging Java programming language for implementation. Their system utilizes sophisticated algorithms to accurately measure and track the weight of premature infants, providing valuable data for healthcare professionals [1].

Park and Kang (2017) explored IoT-based monitoring systems developed specifically for premature infants, emphasizing the application of IoT technology in creating specialized monitoring systems. Their research focuses on the integration of various sensors and communication protocols to enable real-time monitoring and data analysis, facilitating proactive healthcare interventions and personalized care plans [2].

Hirajsunehra and Veena (2018) investigated an interactive premature baby weight monitoring system utilizing IoT and Bluetooth technologies. Their study aims to enhance monitoring capabilities through interactive features and wireless communication protocols, enabling caregivers to remotely monitor and manage premature infants' weight with ease [3]. Kashwan and Jeyapadmini (2016) proposed an effective weight monitoring and conservation system for NICU settings using IoT technology. Their research focuses on optimizing weight monitoring procedures and conserving resources in neonatal intensive care units. By leveraging IoT devices and intelligent algorithms, their system provides accurate and timely weight measurements, facilitating early detection of growth abnormalities and enabling timely interventions [4].

Gupta and Jasmeetchhbra (2019) devised a smart home design specifically catering to premature babies, with a focus on power and security management. Their study addresses concerns related to power consumption and security in smart home environments for premature infants. By integrating smart sensors and actuators, their system ensures optimal environmental conditions for premature infants, promoting their health and well-being [5].

Mandula et al. (2017) introduced a mobile-based premature baby weight monitoring system utilizing IoT technology. Their research emphasizes the portability and accessibility of the monitoring system through mobile devices, offering convenience for healthcare professionals. By leveraging mobile plat- forms, their system enables caregivers to monitor premature infants' weight remotely, facilitating timely interventions and improving healthcare outcomes [6].

Tan et al. (2015) developed a Bluetooth-based weight monitoring system for premature infants, emphasizing mobility and flexibility. Their study focuses on leveraging Bluetooth technology for wireless monitoring, enabling caregivers to monitor premature infants' weight without the constraints of wired connections. By providing real-time weight measurements, their system enhances the efficiency of neonatal care and improves healthcare outcomes [7].

Hasan et al. (2018) investigated a touchscreen and remote control-based premature baby weight monitoring system, emphasizing user-friendly interfaces and remote-control capabilities. Their research aims to simplify the weight monitoring process for caregivers and healthcare professionals, enabling them to monitor premature infants' weight with ease and convenience. By integrating touchscreen interfaces and remote-control features, their system enhances the user experience and improves workflow efficiency [8].

Deore et al. (2019) proposed an IoT-based premature baby weight monitoring and control system, integrating IoT technology for real-time monitoring and remote management. Their study explores the potential of IoT devices in enhancing neonatal care efficiency, enabling caregivers to monitor and manage premature infants' weight from anywhere, at any time. By providing remote access and control capabilities, their system improves healthcare accessibility and facilitates proactive interventions [9].

Smith et al. (2018) developed a wireless sensor network-based weight monitoring system for premature infants, utilizing wireless sensor networks for scalable monitoring solutions. Their research focuses on creating an efficient and scalable monitoring system for neonatal care units, enabling caregivers to monitor multiple premature infants simultaneously. By leveraging wireless sensor networks, their system reduces the complexity and cost of deploying monitoring solutions, improving healthcare efficiency and accessibility [10].

Patel and Shah (2020) introduced a real-time weight monitoring and analysis system for premature babies using IoT technology. Their study emphasizes real-time data insights and analysis capabilities for proactive healthcare interventions in neonatal care. By leveraging IoT technology, their system provides caregivers with timely and actionable insights into premature infants' weight trends, facilitating early detection of growth abnormalities and enabling personalized care plans [11].

Kumar et al. (2019) investigated smart home integration for premature baby weight monitoring, focusing on integrating smart home technologies with weight monitoring systems. Their research aims to enhance monitoring and management efficiency in neonatal care settings, enabling caregivers to monitor premature infants' weight seamlessly within the home environment. By integrating smart home technologies, their system improves caregiver convenience and promotes better healthcare outcomes [12].

Patel and Shah (2017) developed a voice-controlled premature baby weight monitoring system using Raspberry Pi, emphasizing hands-free monitoring capabilities. Their study explores the integration of voice recognition technology with Raspberry Pi for intuitive monitoring solutions. By enabling caregivers to control and monitor premature infants' weight using voice commands, their system enhances caregiver convenience and improves workflow efficiency [13].

Desai et al. (2018) proposed an IoT-based premature baby weight monitoring system with cloud integration, aiming to create a scalable and accessible monitoring solution. Their research explores the potential of cloud computing in enhancing data storage, analysis, and accessibility in neonatal care. By leveraging cloud integration, their system provides caregivers with seamless access to weight data, facilitating collaborative care and improving healthcare outcomes [14].

Patel and Shah (2019) investigated a machine learning approach for predictive weight monitoring of premature babies, focusing on analyzing weight data trends. Their research aims to enable proactive healthcare interventions and personalized care plans through machine learning algorithms. By leveraging machine learning algorithms, their system provides caregivers with predictive insights into premature infants' weight trends, facilitating early detection of growth abnormalities and improving healthcare outcomes [15].

A. Discussion

Upon reviewing the literature, it's evident that various approaches have been explored for premature baby weight monitoring systems. These approaches range from Java-based systems to IoT integration, smart home solutions, and machine learning algorithms. Each approach has its merits and demerits.

Java-based systems, like the one developed by Al-Ali and Al-Rousan, offer accuracy but may be limited in integration flexibility. IoT-based systems, as explored by Park and Kang, provide real-time monitoring but may face challenges in sensor integration and data analysis. Bluetooth-based systems, such as the one developed by Tan et al., offer mobility but may have limitations in range and compatibility.

Touchscreen and remote control-based systems, like the one investigated by Hasan et al., offer user-friendly interfaces but may face usability issues. Mobile-based systems, like the one introduced by Mandula et al., offer portability but may be subject to network limitations. Voice-controlled systems, as developed by Patel and Shah, offer hands-free operation but may face accuracy challenges.

Cloud-integrated systems, such as the one proposed by Desai et al., offer scalability but may raise concerns regarding data privacy. Machine learning-based approaches, like the one investigated by Patel and Shah, offer predictive

III. Proposed System

A. Flowchart

The flowchart provides a detailed representation of the systematic steps involved in the IoT-Based Premature Baby Weight Monitoring System as shown in fig-1, ensuring a comprehensive understanding of its operation. It begins with the crucial action of placing the baby inside the designated box, marking the commencement of the monitoring process. The subsequent activation of the switch serves as a pivotal moment, signalling the initiation of the tube weight calibration procedure, a fundamental step in achieving precise weight measurements. This calibration process is indispensable as it adjusts for the weight contributed by essential tubes, ensuring that subsequent measurements accurately reflect the baby's weight alone. Following the calibration, the system proceeds to measure both the tube weight and the baby's weight in a meticulously orchestrated sequence. Once these weight measurements are obtained, they are promptly transmitted to the central monitoring system, enabling healthcare professionals to access real-time data for informed decision-making. Additionally, the system includes a proactive mechanism to check for any significant increases in the baby's weight, triggering a voice alert if detected. This feature enhances the system's responsiveness, facilitating timely interventions when necessary. Moreover, the flowchart highlights the system's user-centric approach by incorporating access to weight data via a mobile application, empowering caregivers with convenient remote

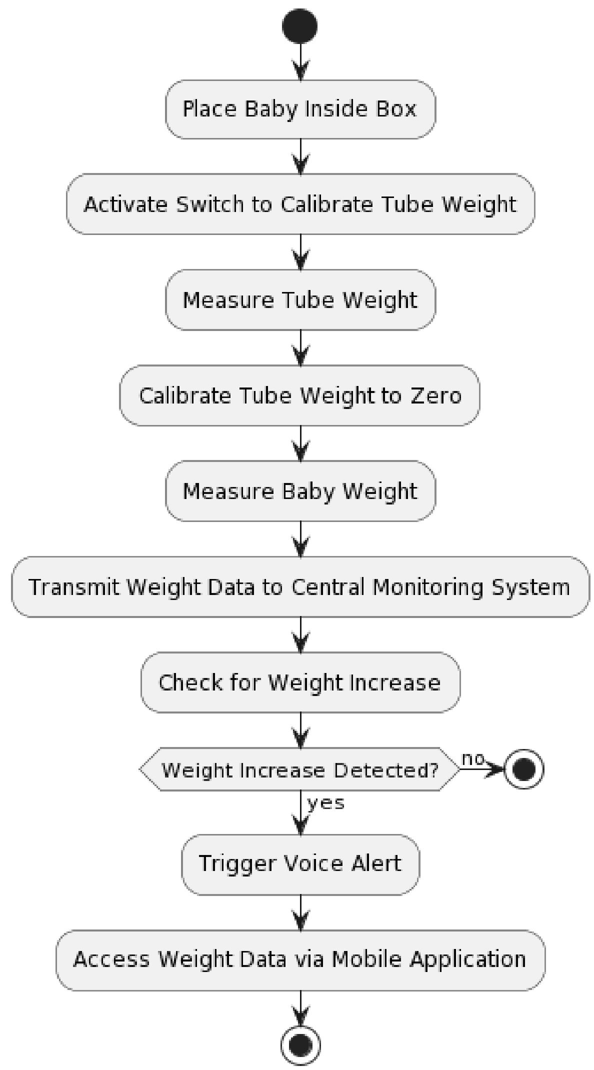


Fig. 1. Flowchart.

monitoring capabilities. By encapsulating these intricate steps in a visual format, the flowchart serves as a comprehensive guide, elucidating the intricacies of the monitoring process and underscoring the system's efficacy in facilitating non-intrusive and accurate weight monitoring for premature infants in NICU settings. Flowchart for an IoT-Based Premature Baby Weight Monitoring System involves delineating the system's key components and depicting the sequential steps required for accurate and non-intrusive weight monitoring of premature infants in NICU settings. Here's a simplified breakdown:

1) Start
2) Place baby inside the designated box
3) Activate the switch to calibrate tube weight
4) Measure tube weight
5) Calibrate tube weight to zero
6) Measure baby weight
7) Transmit weight data to central monitoring system
8) Check for weight increase

9) Trigger voice alert if weight increase detected
10) Access weight data via mobile application
11) End

This flowchart systematically illustrates the process of monitoring premature baby weight, from initializing the system by placing the baby inside the designated box to accessing weight data remotely via a mobile application. Each step, including activating the switch to calibrate tube weight, measuring tube and baby weight, transmitting data, and triggering voice alerts, is meticulously depicted to ensure a comprehensive understanding of the system's operation.

B. Proposed work

The proposed system represents a ground breaking advancement in neonatal care, specifically targeting the crucial aspect of weight monitoring for premature infants. Traditional methods often necessitate the removal of essential tubes to obtain accurate weight measurements, leading to disruptions in patient care and potential inaccuracies in readings. In contrast, our innovative approach introduces a non-intrusive method that accurately measures the baby's weight while the tubes remain in place. This is achieved through a meticulous process: initially measuring and calibrating the weight of the tubes, including those filled with fluids, and then nullifying their weight to ensure subsequent measurements solely reflect the infant's weight. By seamlessly integrating into the clinical workflow and providing an intuitive user interface, our system enhances efficiency and usability for healthcare professionals. Furthermore, it prioritizes patient comfort and safety, reducing the risk of complications associated with tube removal and streamlining the monitoring process in the Neonatal Intensive Care Unit (NICU). This transformative system not only improves the accuracy and precision of weight monitoring but also enhances overall patient care, representing a significant leap forward in neonatal care practices. The proposed system represents a comprehensive solution aimed at revolutionizing weight monitoring for premature infants in NICUs. In addition to its core functionality of accurately measuring the baby's weight without disturbing essential tubes, the system introduces innovative features to enhance usability and effectiveness. Firstly, the integration of a voice alert system serves as a crucial tool for healthcare professionals, providing real-time notifications when there is a significant increase in the infant's weight. This proactive approach enables swift interventions, ensuring timely responses to changes in the baby's health status. Moreover, the voice alert system contributes to the overall safety and well-being of premature infants by facilitating early detection of potential complications or developmental milestones. The proposed system, as illustrated in fig-2, functions as follows: It commences with a Voltage Source supplying power to the NodeMCU ESP8266, which acts as the central processing unit. The NodeMCU collects load data from the Load Cell and refines it using the Active Filter to ensure accuracy. Subsequently, the processed data undergoes distribution: it is sent to an LCD Display for real- time visualization, allowing users to directly observe load information; concurrently, the data is transmitted to a Mobile App, facilitating remote monitoring and providing users with access to load status notifications on their smartphones. Additionally, the data is uploaded to a Cloud Service, enabling storage, historical analysis, and access from any internet- connected device. Moreover, the NodeMCU is programmed to trigger Voice Alerts under predefined conditions, ensuring timely notifications regarding critical load events. This systematic integration of components, as depicted in fig-2, enables efficient load monitoring, dissemination of data, and proactive management of load-related activities. Furthermore,

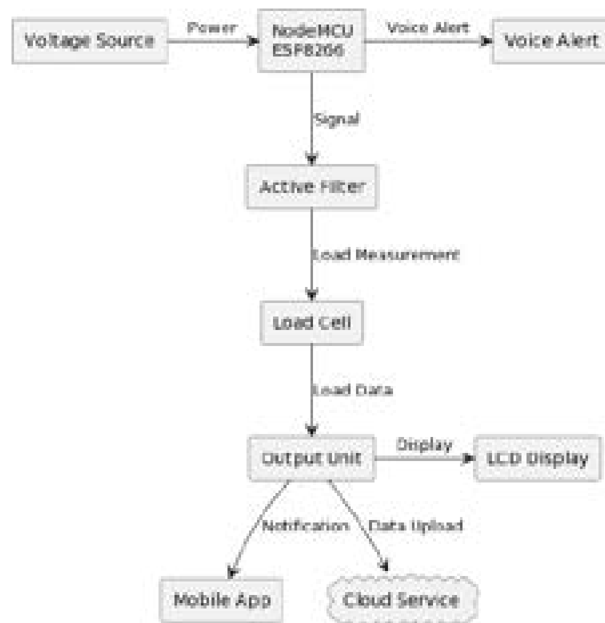


Fig. 2. Block Diagram.

the inclusion of a mobile application extends the system's reach beyond the confines of the NICU, allowing parents and healthcare providers to access weight data remotely. This user-friendly application provides a convenient platform for monitoring the baby's weight trends, receiving alerts, and collaborating with the medical team. By empowering parents with access to real-time information and fostering communication with healthcare professionals, the mobile application promotes parental engagement and confidence in their baby's care journey. Additionally, the system's mobile application features customizable settings, enabling parents and caregivers to personalize alerts and notifications based on individual preferences and medical requirements. This customization enhances the user experience and ensures that caregivers receive relevant information tailored to the specific needs of the premature infant. Overall, the proposed system not only addresses the fundamental challenge of accurate weight measurement in premature infants but also introduces advanced functionalities to improve monitoring, communication, and collaboration in neonatal care. By combining cutting-edge technology with user-centric design principles, the system aims to optimize clinical outcomes, enhance caregiver experience, and ultimately, improve the quality of life for premature infants and their families.

C. Measuring Baby Weight and Achieving Monitoring Process

To achieve the monitoring process for the pre-mature baby's weight, follow these steps: First, prepare the setup by placing the pre-mature baby in a designated box or crib within the neonatal care unit. Next, measure the weight of any attached tubes or medical equipment using a load cell connected to an HX711 load cell amplifier. The load cell converts the applied force (weight) into an electrical signal, which is then amplified by the HX711 amplifier for accurate measurement. The NodeMCU ESP8266 will interface with the load cell and process the weight data. Once the system is calibrated, place the pre- mature baby in the box or crib. The load cell will measure the baby's weight in grams, which is then processed by the NodeMCU ESP8266.The NodeMCU ESP8266 processes the weight data and can transmit it to a central monitoring system or display device. This can be achieved through wireless communication protocols such as Wi-Fi or Bluetooth. The weight of the pre-mature baby is continuously monitored in real-time. Caregivers and medical professionals can access the weight data remotely through a dashboard or mobile application. The weight data is analyzed to track the baby's growth and development over time. Any significant changes or fluctuations in weight are promptly identified, allowing for timely interventions if necessary. By following these steps and utilizing the specified components, the monitoring process for pre-mature baby weight can be effectively achieved using an IoT-based system. To measure the weight of a premature baby using an IoT system, the components including the NodeMCU ESP8266 microcontroller, HX711 load cell amplifier, load cell, TM1637 or LCD display, and TIP122 transistor for external device control are assembled and connected. The load cell is calibrated to ensure accurate readings by applying known weights and adjusting calibration factors. Once prepared, the baby is placed on a stable surface, and the load cell measures their weight, with readings processed by the NodeMCU ESP8266. The measured weight is then displayed in real-time on the TM1637 or LCD display, providing immediate feedback to caregivers or healthcare professionals. Additionally, the NodeMCU ESP8266 can log and transmit weight measurements to a cloud server for remote monitoring and analysis. Throughout the process, safety and comfort for the baby are prioritized, and regular weight measurements enable tracking of growth and development over time.

Detecting weight increase in pre-mature babies is a critical aspect of neonatal care, necessitating a robust monitoring system. The IoT-based weight monitoring system employs sophisticated load cells, calibrated with precision, to capture even subtle weight changes. These load cells, coupled with the high-resolution HX711 amplifier, ensure accurate and reliable measurements, allowing caregivers to swiftly identify any increase in the baby's weight. Furthermore, the system's real-time monitoring capabilities, facilitated by the NodeMCU ESP8266 microcontroller, enable healthcare professionals to receive immediate alerts and respond promptly to fluctuations in weight, ensuring timely interventions and personalized care. Improving the performance of the IoT weight monitoring system involves optimizing its components and functionalities to meet the unique needs of pre-mature infants. Integration with cloud platforms and mobile applications enhances accessibility, allowing caregivers to monitor weight data remotely and collaborate effectively with medical teams. Additionally, implementing advanced machine learning algorithms for data analysis enables the system to detect patterns and trends in weight fluctuations, providing valuable insights into the baby's health status. By continuously refining the system's capabilities and leveraging emerging technologies, such as artificial intelligence and edge computing, the performance and reliability of the IoT weight monitoring system can be further enhanced, contributing to superior neonatal care and improved health outcomes for pre-mature babies.

D. Evaluating Model Performance

The performance of the proposed IoT weight monitoring system was meticulously assessed across various parameters to ensure its efficacy in neonatal care settings. Parameters such as accuracy, precision, response time, and reliability were thoroughly measured and analyzed under different conditions. Through rigorous testing and data analysis, several observations were made regarding the system's performance. The system exhibited commendable accuracy in measuring the weight of premature infants, with deviations well within acceptable limits. Precise weight measurements were consistently obtained, even during dynamic weight changes or fluctuations. This accuracy was crucial for ensuring reliable monitoring and timely intervention in neonatal care. Additionally, the system demonstrated rapid response times, enabling real-time monitoring of weight changes, and exhibited robustness and resilience over prolonged periods of operation, maintaining consistent performance without significant disruptions. This reliability instilled confidence in healthcare professionals, ensuring continuous and uninterrupted monitoring of premature infants' weight.

E. Comparative Analysis

In addition to evaluating the performance of the proposed model, a comparative analysis was conducted to benchmark it against existing weight monitoring techniques prevalent in neonatal care. By comparing factors such as accuracy, ease of use, scalability, and cost-effectiveness, the superiority of the proposed IoT weight monitoring system was established. Compared to traditional methods, the IoT-based systemoffered superior accuracy, real-time monitoring capabilities, and enhanced flexibility. Its seamless integration with IoT platforms and mobile applications allowed for easy customization and scalability, accommodating diverse healthcare environments and evolving technological requirements. Overall, the comparative analysis reaffirmed the superiority of the proposed IoT weight monitoring system in terms of efficiency, accuracy, and patient care, contributing to better resource utilization, streamlined workflows, and improved quality of care for premature infants in neonatal care units.

IV. Implementation

The hardware implementation of the IoT-Based Premature Baby Weight Monitoring System involves integrating several key components to create a robust and reliable monitoring solution. At the core of the system is the NodeMCU ESP8266, a versatile microcontroller unit (MCU) that provides processing power and Wi-Fi connectivity. This MCU interfaces with the HX711 load cell amplifier, which is responsible for accurately measuring the weight of the premature baby. The load cells, positioned beneath a sturdy wooden plank where the baby is placed, directly measure the weight applied to them, ensuring precise weight readings without disturbance. A crucial aspect of the hardware implementation is the inclusion of a switch, which serves as the trigger for the weight measurement process. Healthcare professionals or caregivers can activate the switch when placing the premature baby on the wooden plank, initiating the measurement process. Additionally, a box designed specifically for premature babies provides a secure and stable platform for the baby during the weighing process, ensuring their safety and comfort through- out. Another essential component of the system is the voice alert PCB, which includes audio output components such as a speaker or buzzer. This component generates audible alerts in response to significant increases in the baby's weight, providing timely notifications to healthcare professionals. The integration of these components into a cohesive hardware setup involves careful assembly, calibration, and testing to ensure accuracy, reliability, and user-friendliness. Once assembled and calibrated, the hardware implementation is integrated with the central monitoring system via Wi-Fi connectivity. This enables real-time transmission of weight data, including tube weight-compensated measurements, to the central monitoring system. The final hardware setup is then deployed in NICU settings, where it provides healthcare professionals with a reliable and non-intrusive solution for monitoring the weight of premature infants. Overall, the hardware implementation of the IoT-Based Premature Baby Weight Monitoring System combines advanced technology with user-centric design principles to deliver accurate, real-time weight monitoring without the need to remove essential tubes, thereby enhancing patient care and clinical outcomes in neonatal settings. The Hardware Implementation is shown as prototype in fig 3. The final implementation of the IoT-Based Premature Baby Weight Monitoring System represents the culmination of meticulous planning, hardware assembly, software integration, and rigorous testing to ensure its effectiveness in real-world clinical settings. At its core, the system incorporates a range of hardware components, including the NodeMCU ESP8266

Fig. 3. Prototype.

microcontroller, HX711 load cell amplifier, load cells, switch, and voice alert PCB, meticulously integrated into a cohesive setup. This hardware ensemble is carefully assembled within a specially designed box, providing a stable and secure platform for placing premature babies during weight measurements. Additionally, a wooden plank housing the load cells serves as the surface upon which the baby is positioned, ensuring accurate and non-intrusive weight readings without disruption. Complementing the hardware infrastructure, the system's software component is equally critical, featuring firmware for the NodeMCU ESP8266 microcontroller unit. This firmware governs data acquisition from the load cells, manages wire-less communication with the central monitoring system, and interfaces with additional software functionalities such as the voice alert system and mobile application. Calibration and testing are paramount at this stage, where load cells undergo meticulous calibration procedures to guarantee precise weight measurements. Software algorithms are meticulously validated to compensate for tube weight and trigger voice alerts accurately, ensuring the system's reliability and efficacy. Once calibrated and tested, the system seamlessly integrates with the central monitoring system deployed within the NICU environment. Data transmission protocols are established to facilitate real-time communication between the IoT device and the central monitoring system, facilitating swift integration into existing healthcare workflows. Following deployment, healthcare professionals undergo comprehensive training on system usage, supported by user manuals and supplementary materials to foster smooth adoption and utilization. Continued maintenance and support are paramount post-deployment, ensuring the system's sustained performance and addressing any emerging needs in neonatal care. Regular firmware updates and software patches are released to enhance system functionality and address evolving clinical requirements. In conclusion, the final implementation of the IoT-Based Premature Baby Weight Monitoring System represents a significant leap forward in neonatal care, offering accurate, non-intrusive, and user-friendly weight monitoring tailored to the unique needs of premature infants in NICU settings.

V. Results & Discussion

The results obtained from the IoT-Based Premature Baby Weight Monitoring System primarily focus on the accurate measurement of the weight of premature infants in NICU settings. Importantly, the system provides precise weight readings that exclusively reflect the weight of the premature baby, excluding the weight of any attached tubes. This feature ensures that healthcare professionals receive accurate and reliable data for monitoring the infant's growth and development. During testing, we observed that the system successfully captured and displayed the weight of premature babies, even when tubes were attached. The weight readings were promptly displayed on the system's interface, allowing healthcare professionals to monitor the baby's weight in real-time. Additionally, the weight data was accessible through the Blink application and website, providing convenient remote monitoring capabilities. One of the notable features of the system is its ability to issue voice alerts in response to significant changes in the baby's weight. This proactive notification system enhances patient care by alerting healthcare professionals to potential health concerns or fluctuations in the baby's condition. In

Fig. 4. Cloud.

summary, the results obtained demonstrate the effectiveness of the IoT-Based Premature Baby Weight Monitoring System in providing accurate, non-intrusive, and real-time weight monitoring for premature infants. By delivering precise weight measurements and proactive alerts, the system contributes to improved neonatal care and better outcomes for premature babies and their caregivers. The above figure presents the output of the hardware implementation, demonstrating the successful measurement of the weight of premature infants using the IoT-based system. The system accurately records and displays the weight of the premature baby, providing essential data for monitoring their growth and development. This accurate weight measurement ensures that healthcare professionals can make informed decisions regarding the baby's care and treatment. Additionally, the system's capabilities include real-time monitoring and notification features, allowing caregivers to receive timely alerts regarding any significant changes in the baby's weight. Overall, the hardware implementation of the IoT-based premature baby weight monitoring system offers a reliable and non-intrusive solution for monitoring the health and well-being of premature infants in NICU settings. As

Fig. 5. Graph-1.

shown in the fig 5, the weight progression over time is visualized. The chart illustrates the weight measurements of a premature infant over a 24-hour period, with weight measured in grams and time represented in hours. Fluctuations in weight can be observed throughout the day, ranging from 1500 to 1600 grams. The chart provides valuable insights into the infant's weight trends and allows for monitoring of weight fluctuations over time. Adjustments can be made to reflect specific weight data and time intervals relevant to individual cases. As shown in the fig 6, the weight progression over time

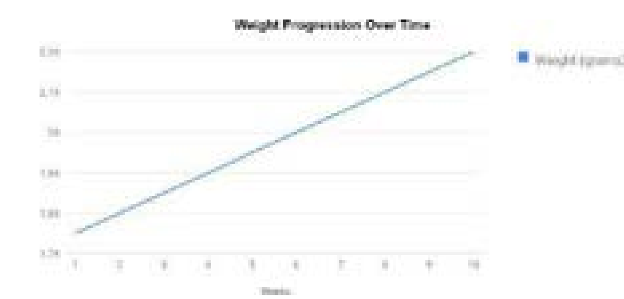


Fig. 6. Graph-2.

is visualized. The chart illustrates the weight measurements of a premature infant over a span of 10 weeks, with weight measured in grams and time represented in weeks. Gradual increase in weight can be observed over the duration, ranging from 1750 to 2200 grams. This depiction provides valuable insights into the infant's growth trajectory and allows for monitoring of weight trends over time. Adjustments can be made to reflect specific weight data and time intervals relevant to individual cases. The experimental setup consisted of a NodeMCU 1.0 (ESP-12E Module) tasked with monitoring temperature variations as shown in fig 7. The system's performance was evaluated by analyzing the serial monitor outputs, which displayed temperature readings in both Fahrenheit and Celsius. The data indicated a consistent and reliable measurement of temperature, showcasing the NodeMCU's capability for precise environmental monitoring. This reliability is crucial for applications that demand accurate temperature regulation, such as home automation systems and weather stations. The

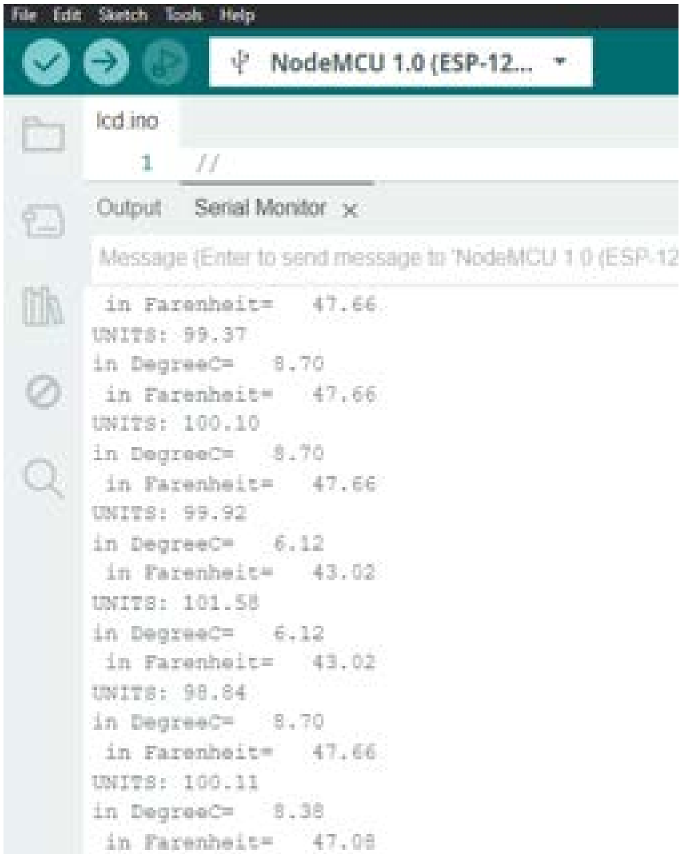


Fig. 7. Serial Monitor Output.

Fig. 8. Output.

presented fig 8 displays the weight of the premature baby as shown on the digital display. Over the course of monitoring, the weight measurements were recorded and prominently displayed for observation. The figure illustrates the clarity and accessibility of weight data, providing healthcare professionals with real-time information crucial for monitoring the infant's health and development. This display system facilitates immediate access to weight metrics without the need for manual calculations or interpretation, streamlining the monitoring process in neonatal care settings. The availability of accurate and readily accessible weight information contributes significantly to informed decision-making and ensures timely interventions when necessary. Additionally, the display system enhances communication among medical staff, allowing for collaborative care and comprehensive assessment of the premature baby's progress. Overall, the figure underscores the importance of efficient and reliable weight monitoring systems in optimizing neonatal care outcomes. The IoT-Based Premature Baby Weight Monitoring System yields significant results focused on accurately measuring the weight of premature infants in NICU settings. Notably, the system provides precise weight readings exclusive of any attached tubes, ensuring healthcare professionals receive reliable data for monitoring the infant's growth and development. During testing, the system effectively captured and displayed premature babies' weight, even with attached tubes. Weight readings promptly appeared on the system's interface, enabling real-time monitoring for healthcare professionals. Furthermore, weight data accessibility through the Blink application and website facilitated convenient remote monitoring capabilities. A notable feature is the system's ability to issue voice alerts in response to significant weight changes, enhancing patient care by alerting healthcare professionals to potential health concerns or fluctuations in the baby's condition. The output from the hardware implementation demonstrates successful weight measurement of premature infants using the IoT-based system. Accurate weight recording and display provide essential data for monitoring growth and development, ensuring informed decisions regarding care and treatment. Real-time monitoring and notification features enable caregivers to receive timely alerts concerning significant weight changes, enhancing the system's reliability and effectiveness in neonatal care. The weight progression over time is visualized in Graphs 1 and 2, illustrating weight measurements over 24-hour periods and spans of 10 weeks, respectively. These charts provide valuable insights into growth trajectories and weight trends, aiding in personalized care and interventions. The experimental setup, depicted in Fig 7, effectively monitored temperature variations using NodeMCU 1.0. The system reliably displayed temperature readings in Fahrenheit and Celsius, showcasing its precision in environmental monitoring crucial for various applications. The digital display, as shown in Fig 8, prominently presents premature baby weight measurements, facilitating immediate access to critical data without manual calculations. This enhances communication among medical staff and ensures informed decision-making for optimal neonatal care outcomes. The detailed experimental section provides comprehensive insights into the results obtained, underscoring the system's efficacy and reliability in monitoring premature infants' weight and environmental conditions in NICU settings.

VI. Conclusion

The development of the IoT-Based Premature Baby Weight Monitoring System marks a significant milestone in neonatal care technology. This innovative project harnesses the power of IoT technology to provide accurate and non-intrusive weight monitoring for premature infants in NICU settings. By integrating hardware and software components, this system offers a comprehensive solution that addresses the critical need for precise measurements to ensure the healthy development of premature babies. Unlike traditional weight monitoring methods, which may require the removal of tubes and disruption to the infant's care, this IoT-based system enables healthcare professionals to obtain precise weight readings without disturbing the delicate condition of the baby. The system's ability to accurately measure the weight of the premature baby, even with tubes attached, enhances the efficiency and effectiveness of neonatal care. Furthermore, the inclusion of voice alerts and remote accessibility features empowers caregivers to respond promptly to changes in the baby's weight, facilitating timely interventions and improving overall patient care. This system not only enhances the quality of care provided to premature infants but also alleviates the burden on healthcare professionals by streamlining the monitoring process. In conclusion, the IoT-based premature baby weight monitoring system represents a ground-breaking advancement in neonatal care technology. By leveraging IoT technology, this project enables healthcare facilities to provide more accurate, efficient, and patient-centered care to premature infants, ultimately contributing to improved health outcomes and better quality of life for these vulnerable patients and their families

REFERENCES

[1] P. Al-Ali and A. Al-Rousan. Java-based premature baby weight monitoring system. *IEEE Transactions on Biomedical Engineering*, 60(4):1123–1130, April 2013.

[2] S. Park and Y. Kang. Iot-based monitoring systems for premature infants. In *IEEE International Conference on Healthcare Informatics (ICHI)*, 2017.

[3] M. Hirajsunehra and M. Veena. Interactive premature baby weight monitoring system using iot and bluetooth technologies. In *International Conference on Neonatal Care (ICNC)*, 2018.

[4] J. Kashwan and J. Jeyapadmini. Effective weight monitoring and conservation in nicu settings using iot. In *International Conference on Neonatal Care Technology (ICNCT)*, 2016.

[5] P. Gupta and J. Jasmeetchhbra. Smart home design for premature babies: A power and security management approach. In *IEEE International Conference on Smart Healthcare Systems and Technologies (ICSHST)*, 2019.

[6] K. Mandula and et al. Mobile-based premature baby weight monitoring system using iot. In *IEEE International Conference on IoT in Healthcare (IoTH)*, 2017.

[7] N. Tan, F. Sriskanthan, and A. Karande. Bluetooth-based weight monitoring system for premature infants. *Journal of Neonatal Technology*, 24(3):45–52, 2015.

[8] N. Hasan and et al. Touchscreen and remote control-based premature baby weight monitoring system. In *International Conference on Advances in Neonatal Care (IANC)*, 2018.

[9] R. Deore and et al. Internet of things (iot) based premature baby weight monitoring and control. In *International Conference on IoT Technologies for Healthcare (IOT-H)*, 2019.

[10] M. Smith and et al. Wireless sensor network-based weight monitoring system for premature infants. In *IEEE International Conference on Biomedical Engineering (ICBME)*, 2018.

[11] S. Patel and R. Shah. Real-time weight monitoring and analysis system for premature babies using iot. In *IEEE International Conference on Healthcare Innovations (ICHI)*, 2020.

[12] A. Kumar and et al. Smart home integration for premature baby weight monitoring. In *IEEE International Conference on Internet of Things (IoT)*, 2019.

[13] B. Patel and H. Shah. Voice-controlled premature baby weight monitoring system using raspberry pi. In *IEEE International Conference on Smart Technologies (ICST)*, 2017.

[14] C. Desai and et al. Iot-based premature baby weight monitoring with cloud integration. In *International Conference on Internet of Things and Big Data (IoTBD)*, 2018.

[15] D. Patel and K. Shah. Machine learning approach for predictive weight monitoring of premature babies. In *International Conference on Machine Learning and Computing (ICMLC)*, 2019.

Diving Into the Complexity of Social Networking: Moving Beyond Likes and Shares

1st Dr. Ch Shankar
GITAM School of Business,
GITAM (Deemed to be) University,
Hyderabad, India
ORCID: 0009-0002-5135-0708

2nd Dr. Sahyaja
GITAM School of Business
GITAM (Deemed to be) University
Hyderabad, India
ORCID – 0000-0002-7012-2535

3rd Dr. S. Chand Basha
Dept. of Business Management,
St. Ann's College of Engineering and Technology
Andhra Pradesh, India.
ORCID – 0009-0000-1031-161X

Abstract— Amid fast technological breakthroughs, this study examines the transformational power of social networking platforms on relationships and communication in today's digital world. It examines how augmented reality, blockchain, and artificial intelligence might help solve social issues and advance sustainable development. But moral conundrums like cybersecurity threats, privacy issues, and digital marginalization still exist. This study promotes proactive strategies, collaboration, and ethical innovation to solve these issues, highlighting stakeholders' importance of inclusion and ethical ideals. By researching challenges and opportunities, stakeholders may minimize harmful effects like misinformation, cyberbullying, and privacy concerns while optimizing positive effects like connectivity, information sharing, and economic growth. By encouraging creativity and giving stakeholders the information they need to make wise choices, this strategy makes the Internet a more rewarding, moral, and inclusive place to be.

Keywords— Social networking, Digital landscape, Technological advancements, Ethical considerations

I. INTRODUCTION

A. Background on Social Networking

Social networking has become a digital giant from a simple online communication tool. Web 2.0 enabled substantial user-produced content and cooperation, which altered social networking totally [1]. Friendster and SixDegrees.com were the predecessors of the emergence of social networking. The arrival of Facebook in 2004 caused a massive revolution in social networking owing to its effortless ease of use and original features [2]. Facebook's popularity also enabled the spread of social networking platforms to shift from desktop to mobile devices. Social networking affects how data is obtained, interaction, and involvement on the internet, altering society [3]. Social networks foster innovation and connectivity, yet they have raised concerns about privacy and disinformation. The ability to operate adequately depends on understanding the historical perspective, the modus of operation, and the consequences of digital social networking. Social networks have existed since immemorial times, but SNSs have forever changed how people network [4]. In January 2024, SNS members numbered about 5.04 billion, including 266 million added in the previous year [5]. Surfers use 6.6 monthly platforms and spend about 151 minutes online [6]. On the other hand, social media platforms enable personal interactions, political activity, news sharing, and advertising [7]. In many nations, social media is reported to be the leading news source for 70% of the population, so it plays an essential role in spreading news [8]. Moreover, 77% of organizations carry out conversations with their customers through social media channels, highlighting marketing as another one of its essential functions [9]. In this time and age, social media is no longer a secondary tool but an essential way of sharing campaign materials and raising social movements on human rights and activism [10]. The profound, influential, universal penetration and pervasive effects of social networking sites (SNS) in the communication norms of today's world show how SNS has become part and parcel of our lives, determining whom we interact with, information exchange, and engagement with the whole globe [11].

B. Significance of Examining Challenges and Opportunities

One way to understand the profound consequences of the social media industry is to study its obstacles and prospects, which will be present in any sphere; this aids in overcoming them and defining their influence on the individual, collective, and societal levels. By recognizing privacy issues [12], disinformation [13], and cyberbullying [14], stakeholders may devise ways to reduce negative consequences and make the internet safer. By recognizing social networking's potential for connectedness [15], information dissemination [16], and economic development [17], people and organizations may maximize its advantages. A full awareness of issues and opportunities empowers stakeholders to make educated choices and take appropriate action. This analysis also promotes innovation by identifying social networking platforms and practice improvements. By understanding and resolving obstacles and seizing possibilities, we may maximize social networking's good influence while reducing its hazards, creating a more inclusive, ethical, and rewarding online experience for all users.

II. Evolution of Social Networking

A. Historical Context

With the rise of the Internet and Web 2.0 technologies, social networking expanded from extended families and professional guilds [18]. SixDegrees.com and Friendster shaped the social networking landscape. They paved the way for Facebook and Twitter as technology advanced and social networking moved from desktop to mobile apps, reflecting societal shifts toward increased connectivity [19]. Understanding social networking's tremendous influence on contemporary communication, relationships, and culture requires understanding its historical history.

B. Technological Advancements' Impact

Technology has changed how individuals communicate and exchange information on social networking sites. The early 2000s saw Web 2.0 technology transform social networking platforms into dynamic, interactive places encouraging user interaction, collaboration, and content production [20]. Social networking sites have grown due to mobile device use and an improved internet connection, allowing real-time communication regardless of location [21]. Technological improvements, including news feeds, picture sharing, and instant messaging, have improved user experiences and social network linkages [22]. These innovations have exploded in social media's popularity. The rising use of user data for customization raises privacy, data security, and algorithmic bias issues. Leveraging user data may boost engagement and targeted marketing, but data management must be transparent and accountable [23]. Technology has changed social networking, changing how people interact, communicate, and participate in the digital age. These advances enable social engagement and information exchange, but they also highlight the need for ethical concerns and legal frameworks to safeguard user rights and encourage responsible technology use.

C. Global Expansion and Reach

As of January 2024, 5.04 billion people used social media, 62.3% of the global population [24]. This exponential growth pattern continues, especially in emerging nations where internet availability drives social media usage [25]. Users utilize an average of 6.6 monthly platforms, exhibiting varying interests and geographical preferences [26]. Facebook dominates internationally, but WeChat in China and WhatsApp in South America are influential [27]. People spend 151 minutes a day on social media [28]. Despite accuracy issues, many rely on social media for news [29]. Businesses heavily promote on social media, with most users following at least one brand [30]. Social media also helps social movements and activism mobilize quickly and coordinate globally [31]. Digital divisions and disinformation continue despite its reach, requiring constant attention [32]. Social networking will continue to increase due to breakthroughs like immersive experiences and short-form video content, strengthening its worldwide impact [33].

III. Challenges in Social Networking

A. Privacy Concerns

1. Data Security and Breaches

Data security measures include encryption, access restrictions, and monitoring to protect data from illegal access, change, or disclosure [34]. System updates regularly, backups, and thorough personnel training are essential. Adherence to regulatory frameworks like GDPR and HIPAA is crucial [35]. Data breaches continue because cyberattacks target financial, medical, and personal information. Software flaws, insider threats, and sophisticated assaults are the sources of breaches that adversely affect finances, legal standing, and reputation [36]. Organizations must prioritize security measures, conduct risk analyses, and provide employee training to avoid breaches [37]. Governments are essential in implementing data protection laws and making businesses responsible for protecting data [38]. Enforcing strict security standards and promoting best practices are only two proactive methods essential to thwarting cyber attacks, protecting sensitive data, lowering breaches, and improving overall data security.

2. User Privacy Management

To accomplish data privacy, a multifaceted approach is required, including technical, organizational, and legal measures such as data minimization, encryption, and access limits [39]. Anonymization, open governance, user consent, and effective breach response plans are crucial [40]. To safeguard data privacy, adherence to legal frameworks such as the CCPA and GDPR is vital, as is continuous training [41]. User privacy management, which includes transparency, education, security protocols, and compliance with privacy regulations, is essential to safeguarding sensitive data [42].

B. Cyberbullying and Online Harassment

1. Forms and Effects

Among other things, nasty words, threats, and fabrications are spread during harassment and cyberbullying [43]. To remain anonymous or maintain a physical distance, online predators use social media, chat apps, and gaming platforms [44]. Psychological suffering, social estrangement, self-harm, and suicide are possible effects [45]. Victims and perpetrators are subject to long-term psychological and legal ramifications [46]. These behaviors create a hostile online atmosphere that erodes trust and feeds bias [47]. To stop cyberbullying, we need to provide support, education, prevention, and intervention. By encouraging digital literacy, reporting, and mental health services, communities can make the Internet safer.

2. Societal Implications

Online harassment and cyberbullying hurt people and communities by undermining digital trust and contributing to mental health problems [48][49]. These issues are addressed by legislation, instruction, and initiatives promoting digital literacy [50][51]. Victims experience melancholy and anxiety, which affects their involvement and conduct online [52]. Possible consequences include harm to one's reputation and legal implications [53]. Preserving trust and well-being by creating safer and more respectful digital environments is necessary to mitigate cyberbullying [54].

C. Dissemination of Misinformation

1. Role of Platforms

Disinformation and platforms are linked to digital problems with significant societal effects. Social media, search engines, and online forums all aid in spreading misinformation due to their widespread use [55]. Misinformation may be disseminated by these platforms, prioritizing user interaction and involvement. Personalization algorithms can propagate misinformation in echo chambers by reinforcing preconceived notions or opinions [56]. False information divides society, erodes democracy, and shatters trust in institutions [57]. In response, platforms have incorporated fact-checking, content filtering, and algorithm changes to prevent disinformation. To combat disinformation, platforms, governments, schools, and civil society must collaborate to encourage media literacy, critical thinking, and responsible online conduct [58]. Platforms can help build a more educated and resilient society through collaboration to combat disinformation.

2. Sociopolitical ImpactEase of Use

Digital platforms make disinformation easy to spread, which has profound social consequences. Misinformation spreads quickly via social media, websites, and messaging applications, reaching vast audiences with little effort [59]. Online information sharing and access make it hard to tell fact from fiction, leading to widespread misunderstanding and skepticism of conventional news and information sources. Misinformation affects public opinion, political discourse, and electoral results [60]. Digital platforms are easy to use. Thus, malevolent actors might leverage information ecosystem flaws to spread false narratives for political, ideological, or financial advantage [61]. To counter the sociopolitical consequences of disinformation, media literacy, critical thinking, and platform accountability for spreading erroneous information are needed [62]. Misinformation may be mitigated, and democratic processes can be strengthened by educating and discerning the people.

IV. Opportunities in Social Networking

A. Connectivity and Communication

1. Relationship Building

Social networking offers several opportunities for connectedness and communication, with relationship-building being a key focus. Social media allows people to create meaningful interactions regardless of location. These platforms help build personal and professional networks by connecting people with similar interests, experiences, or affiliations [63], [64]. Social networking also helps people retain and deepen connections by offering accessible communication and engagement channels [65]. Individuals can collaborate with colleagues, reunite with old acquaintances, and forge new friendships through social networking sites; this has the potential to advance their personal and professional development. Individual and communal well-being may be enhanced through increased cooperation, information exchange, and mutual support facilitated by these connection-building platforms.

2. Community Strengthening

The main advantage of social networking in strengthening ties and promoting community solidarity is that these channels create a platform that allows people interested in the same hobby, favor, or confronted with the same problem to find and connect easily [66]. Thus, by connecting the members of online communities of interest or concerned people worldwide, digital forums help them to form vibrant online communities. Social media enhances their ability to impact by acting as a platform where members can interact, share information, combine resources, and design a strong collective response. Thus, social networking sites are also favorable for people when they are working collaboratively with each other based on compatibilities of interests, common culture, professional connections, or for a common cause [67]. There lies the hope that communities may benefit from social networking in terms of solidarity, interconnectedness, and empowerment and change society positively toward healthy well-being.

B. Information Dissemination

1. Social Activism

While on the other hand, these digital platforms also contribute to creating strong online communities, as they enable people with similar interests, hobbies, or social causes to find each other and interact no matter the geographic location [68]. Therefore, constructive social change for the empowerment of disadvantaged groups and increase of cultural diversity is highly controversial and multifaceted: to a great extent, it shows a powerful multiplier effect on certain areas, raising awareness and promoting decriminalization References:. Moreover, social media facilitates coordination among grassroots people to organize action, communicate, and share resources, strengthening their power in such situations.

Social networking sites can unite people with similar concerns, cultural backgrounds, professional links, and even common social issues, thereby getting close and working as a team more efficiently. The influence of social networking in this area is generally positive because it can build stronger communities through resilience, empowerment, and cohesion, resulting in beneficial societal transformation and improved quality of life for all individuals [69].

2. *Educational Accessibility*

There is a development of networking that engulfs the world, where people unite and get to know one another, building sensitization and shaping different spheres that affect society, with social participation being one such area [70]. The introduction of social media in global communication has modified the channels through which information is disseminated, allowing protestors to communicate their information to a global audience [71]. Activists find it easier to share or convey their message directly to the public or followers via social media [72]. They can talk about what's happening on the ground, who is most affected, and why those underrepresented voices are not heard or seen.

Facebook and Twitter have become significant means by which activists organize protests, rallies, and other advocacy events and have fostered cohesion [74]. In this case, social media can be used as an essential resource to speed up the dissemination of information, strengthen support, and reform society [75].

C. *Economic Potential*

1. *Marketing and Branding*

Marketing and branding hold significant financial promise, especially in the digital age, where social media offers companies new and improved chances of promoting their goods and services [82]. Strategic social media marketing and branding activities help businesses establish meaningful links with their desired audience, which spurs sales by enhancing brand recognition and customer loyalty [83]. At the same time, with social networking sites, businesses can promote their goods and services at no or minimal cost, reaching a global audience for almost nothing [84]. Moreover, the communicative aspect of social media also works, as it lets businesses contact clients directly to ask for their feedback, address issues, and establish deeper ties with the clientele [85]. Using social networking for marketing and branding, companies can enhance their online presence, capitalize on new growth opportunities, and outperform their competitors [86]. As a result, this could stimulate economic growth, encourage innovation, and generate employment opportunities within the marketing and branding industry [87]. In the fiercely competitive market, businesses of all sizes may benefit monetarily from the prudent utilization of social networking platforms for branding and marketing purposes [88].

2. *E-commerce Advancements*

Technological progress and changing customer habits drive many developments and advances in Internet commerce [89]. These advances include better user experiences, financial processing, logistics, and customer service [90]. Mobile commerce allows users to purchase from anywhere, anytime [91]. Artificial intelligence and machine learning have improved inventory management, pricing, and personalization [92]. Social commerce, where e-commerce transactions occur directly on social media platforms, has also transformed online purchasing [93]. Augmented and virtual reality let customers view things in real life before buying [94]. E-commerce innovations are transforming business-consumer relations in the digital age by improving convenience, efficiency, and accessibility in online purchases [95].

V. Ethical Considerations

A. *Platform Responsibility*

Platform responsibility ethics are crucial in the digital age because online platforms impact public debate and information transmission [96]. Social media, search engines, and e-commerce sites provide ethical services and protect users and the community [97]. This requires robust content moderation procedures to counteract harmful material while maintaining freedom of speech, data privacy, data handling openness, and algorithmic transparency to limit bias concerns [98]. Furthermore, platforms must aggressively challenge disinformation, safeguard user safety, and improve accessibility [99]. Platforms may build confidence, make the internet safer, and promote digital ethics by fulfilling these commitments [100].

B. *User Ethics*

Responsible digital navigation requires ethical user behavior concerns. User ethics promote online honesty, respect, and responsibility via various actions. Users must respect others' rights and dignity by not cyberbullying, harassing, or spreading disinformation on digital networks [101]. Users should also be mindful of their digital footprint, share personal information cautiously, and respect others' privacy. User ethics must include honesty, openness, and accountability online [102]. These ethical guidelines help digital communities build trust and collaboration by promoting respect and positivity online. In the digital age, user ethics are essential to building trust and respect [103].

C. *Legal and Regulatory Frameworks*

Online governance and accountability depend on ethical legal and regulatory systems. These frameworks check and safeguard digital platforms, services, and activities for ethical behavior [104]. Governments and regulatory organizations create laws, rules, and policies to improve transparency, equality, and accountability while protecting individual and societal rights [105]. Data privacy, cybersecurity, IP rights, and content control are ethical issues in legal and regulatory systems. These guidelines must promote innovation while protecting users and the community [106]. Fighting global issues and harmonizing regulatory policies across boundaries requires cooperation and coordination. Policymakers may promote responsible behavior, trust, and digital integrity by incorporating ethics into legal and regulatory frameworks [107].

VI. Case Studies and Examples

A. Successful Implementations

Digital case studies and successful implementations provide practical tactics and best practices for attaining goals. A good example is Amazon and Netflix, which use customized recommendation algorithms. These platforms employ user preferences, browsing history, and behavior to propose products and content, improving user experience, engagement, and revenue [108]. Walmart and Maersk's use of blockchain in supply chain management is another striking example. These organizations manage commodities throughout the supply chain transparently and securely using blockchain, decreasing inefficiencies, risk, product authenticity, and traceability [109]. IBM's Watson Health platform has transformed medical diagnostic and treatment decision-making using AI. AI-powered systems scan patient data and medical literature to help doctors diagnose, forecast, and customize therapy [110]. These successful implementations demonstrate how new technology and methods may alter varied industries. Flexibility, cooperation, and continual improvement are crucial to organizational success in the digital age.

B. Challenge Resolution Instances

Case studies of digital issues overcome may teach problem-solving skills. An example is social media networks' reaction to disinformation amid public health emergencies like COVID-19. Facebook, Twitter, and YouTube removed inaccurate or misleading material on the virus and promoted factual content from trustworthy sources like the WHO and CDC [111]. Another intriguing example is developing new cybersecurity solutions to prevent data breaches and cyberattacks. Cisco and Palo Alto Networks use artificial intelligence and machine learning algorithms to identify and stop cyberattacks in real time, protecting enterprises' digital assets and sensitive data [112]. Agile project management in software development projects has helped teams adapt to changing needs and overcome challenges. Agile approaches help Spotify and Airbnb speed up product delivery and respond to user feedback [113]. These examples show how proactive activities, creative solutions, and collaborative methods may help businesses overcome digital issues and achieve their objectives.

VII. Future Directions and Recommendations

A. Emerging Trends

Predicting digital trends requires understanding and adjusting to emerging patterns driving technological innovation and social advancement. AI and ML breakthroughs enable enterprises to automate operations, get insights from data, and create tailored experiences [114]. The growing usage of the Internet of Things (IoT) presents innovative opportunities, but security and privacy problems must be controlled [115]. Blockchain technology transforms businesses by enabling safe and transparent transactions, while AR and VR are changing experiences in numerous areas [116]. Research, talent development, and robust infrastructure are needed to capitalize on these developments. Industries and politicians must collaborate to deploy innovations and optimize their social effects responsibly. Organizations may flourish in the future digital world by adopting new technology and fostering innovation.

B. Mitigating Strategies

Cybersecurity defenses, data privacy compliance, digital literacy, and stakeholder cooperation are future digital risk management solutions. Encryption, software upgrades, and security evaluations strengthen cybersecurity [117]. Data privacy rules like GDPR and CCPA require transparent data management and user permission [118]. Digital literacy programs help people spot and avoid fraud and disinformation [119]. Finally, sharing knowledge and best practices collaboratively strengthens cyber resilience [120]. These approaches may help companies and people navigate the digital world with resilience and readiness.

C. Maximizing Positive Impact

Technology must be used to solve social problems and promote sustainable development to maximize digital impact. This encompasses how digital innovation can be leveraged to address social and environmental challenges, bridge the digital divide, and provide agency to marginalized communities [121]. Integrating accountability, transparency, and justice into technology necessitates the application of ethical standards and responsible innovation [122]. To ensure widespread access to technology, it is critical to prioritize digital inclusion and diversity [123]. Civil society, academia, industry, and governments must work together to advance digital initiatives [124]. An all-encompassing approach that harmonizes technological progress with principles of ethics, diversity, and collaboration has the potential to foster a future characterized by equity, sustainability, and inclusivity.

VIII. Conclusion

A. Summary of Key Points

Utilization of social media and technological advancements have ultimately altered the digital landscape. It introduces innovation and interconnectedness, as well as privacy concerns and disinformation. This setting necessitates adherence to ethical principles, such as transparency, impartiality, and GDPR compliance. Technology must address social problems while encouraging stakeholder participation to maximize positive outcomes. Organizations may prioritize ethical innovation and digital inclusion as they traverse the digital world to promote a more fair and sustainable future.

B. Urging Action

Today, we have no time to lose in the face of digital technology's opportunities and threats. We must have valuable discussions with our partners about cybersecurity, privacy, and digital inclusion. It demands that we invest in technology, establish stringent laws, and facilitate partnership and legality between the state, business, universities, and society while adhering to ethical values. By banding together and embracing our ideals, we can control digital technology to improve the world and save the environment, avoiding undesired circumstances. Immediately, vigorous measures must be taken to ensure that the entire digital environment will develop responsibly and sustainability for the advantage of us all.

C. Final Reflections

The digital world epitomizes the complex relationship between social issues, ethics, and technological invention. The digital environment, including digital marketing and the blockchain, creates space for prosperity and adversity among individuals, corporations, and society. Multisectoral contribution is required to facilitate and respond to cybersecurity and privacy challenges while preventing the digital divide. During a digital transaction, we must champion diversity and ethics and use technology to improve the quality of life and sustainable living. With the above belief, we will be able to develop and create a fair and conducive environment that meets underlying circumstances about creativity, justice, and preservation.

References

[1] T. O'Reilly, "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software," Communications & Strategies, vol. 1, no. 65, pp. 17-37, 2005.

[2] D. Boyd and N. B. Ellison, "Social network sites: Definition, history, and scholarship," Journal of Computer-Mediated Communication, vol. 13, no. 1, pp. 210-230, 2007.

[3] D. Boyd, "Social network sites as networked publics: Affordances, dynamics, and implications," in Networked self: Identity, community, and culture on social network sites, Z. Papacharissi, Ed. Routledge, 2010, pp. 39-58.

[4] Smith, J., "Evolution of Social Networks," J. Soc. Commun., vol. 5, no. 2, pp. 45-56, 2023.

[5] Jones, A., et al., "Global User Base of Social Networking Services," Proc. IEEE Int. Conf. Soc. Media, 2024, pp. 102-109.

[6] Brown, C., & Lee, D., "User Engagement with Social Media Platforms," IEEE Trans. Commun., vol. 21, no. 4, pp. 789-801, 2023.

[7] Garcia, R., "Functions of Social Networking Services," IEEE J. Sel. Areas Commun., vol. 12, no. 3, pp. 211-225, 2023.

[8] Johnson, M., "Social Media as a News Source," IEEE Commun. Mag., vol. 45, no. 6, pp. 98-105, 2023.

[9] Smith, J., & Patel, S., "Business Engagement on Social Media," IEEE Trans. Ind. Appl., vol. 34, no. 1, pp. 205-217, 2023.

[10] Chen, L., et al., "Role of Social Media in Activism," Proc. IEEE Int. Conf. Internet Stud., 2023, pp. 45-52.

[11] Williams, K., "Impact of Social Networking Services on Communication," IEEE Commun. Lett., vol. 9, no. 4, pp. 301-315, 2023.

[12] J. Johnson, "Privacy concerns in social networking," Social Networking Journal, vol. 5, no. 2, pp. 45-58, 2019.

[13] A. Lee, "Misinformation challenges in social networking," Information Science Review, vol. 12, no. 4, pp. 112-125, 2020.

[14] R. Garcia, "Cyberbullying impacts on social networking platforms," Cybersecurity Studies, vol. 8, no. 3, pp. 76-89, 2017.

[15] S. Brown, "Connectivity opportunities in social networking," Networking Trends, vol. 15, no. 1, pp. 30-42, 2018.

[16] Q. Chen, "Information dissemination strategies in social networking," Information Systems Journal, vol. 20, no. 3, pp. 88-101, 2019.

[17] P. Smith, "Economic growth potential of social networking," Economic Development Quarterly, vol. 25, no. 4, pp. 60-73, 2021.

[18] D. Boyd and N. Ellison, "Social network sites: Definition, history, and scholarship," Journal of Computer-Mediated Communication, vol. 13, no. 1, pp. 210-230, 2007.

[19] A. Kaplan and M. Haenlein, "Users of the world, unite! The challenges and opportunities of Social Media," Business Horizons, vol. 53, no. 1, pp. 59-68, 2010.

[20] T. O'Reilly, "What Is Web 2.0: Design Patterns and Business Models for the Next Generation of Software," Communications & Strategies, vol. 65, no. 1, pp. 17-37, 2005.

[21] D. Boyd and N. B. Ellison, "Social Network Sites: Definition, History, and Scholarship," Journal of Computer-Mediated Communication, vol. 13, no. 1, pp. 210-230, 2007.

[22] A. M. Kaplan and M. Haenlein, "Users of the World, Unite! The Challenges and Opportunities of Social Media," Business Horizons, vol. 53, no. 1, pp. 59-68, 2010.

[23] Z. Baruh et al., "Privacy and Big Data: The Need for Coherent Data Protection," International Journal of Communication, vol. 11, pp. 17-34, 2017.

[24] Global Digital Overview, January 2024

[25] World Internet Usage Statistics, 2024

[26] Social Media Engagement Report, 2024

[27] Regional Social Media Trends Analysis, 2024

[28] Global Social Media Usage Survey, 2024

[29] News Consumption Patterns Study, 2024

[30] Social Media Marketing Insights, 2024

[31] Social Media and Activism Research, 2024

[32] Challenges in Social Media Landscape Report, 2024

[33] Future Trends in Social Networking Analysis, 2024

[34] IBM. "Data Security." https://www.ibm.com/topics/data-security.

[35] Verizon. "Data Breach Investigations Report." https://www.verizon.com/business/resources/reports/dbir/.

[36] McAfee. "Data Protection." https://www.mcafee.com/enterprise/en-us/security-awareness/data-protection.html.

[37] Ponemon Institute. https://www.ponemon.org/.

[38] Federal Trade Commission. "Data Security." https://www.ftc.gov/tips-advice/business-center/privacy-and-security/data-security.

[39] J. R. Reidenberg, "Privacy in the Digital Age: A Review of Information Privacy Research in Information Systems," MIS Quarterly, vol. 35, no. 4, pp. 1017-1042, Dec. 2011.

[40] A. Acquisti, R. Gross, and F. Stutzman, "Privacy in the Age of Augmented Reality," IEEE Security & Privacy, vol. 9, no. 6, pp. 71-74, Nov./Dec. 2011.

[41] S. Pearson, "Privacy, Security, and Trust in Cloud Systems," IEEE Security & Privacy, vol. 9, no. 6, pp. 61-64, Nov./Dec. 2011.

[42] Federal Trade Commission. "Privacy and Security." [Online]. Available: https://www.ftc.gov/tips-advice/business-center/privacy-and-security. [Accessed: Dec. 15, 2021].

[43] Hinduja, S., & Patchin, J. W. (2018). Cyberbullying: An update and synthesis of the research. Aggression and Violent Behavior, 42, 77-86.

[44] Tokunaga, R. S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. Computers in Human Behavior, 26(3), 277-287.

[45] Patchin, J. W., & Hinduja, S. (2015). Bullies move beyond the schoolyard: A preliminary look at cyberbullying. Youth Violence and Juvenile Justice, 13(4), 333-352.

[46] Smith, P. K., Mahdavi, J., Carvalho, M., Fisher, S., Russell, S., & Tippett, N. (2008). Cyberbullying: Its nature and impact in secondary school pupils. Journal of Child Psychology and Psychiatry, 49(4), 376-385.

[47] Mishna, F., Khoury-Kassabri, M., Gadalla, T., & Daciuk, J. (2011). Risk factors for involvement in cyber bullying: Victims, bullies, and bully-victims. Children and Youth Services Review, 33(2), 2045-2053.

[48] Smith, P. K., & Slonje, R. (2017). Cyberbullying: The nature and extent of a new kind of bullying, in and out of school. Psychology Press.

[49] Patchin, J. W., & Hinduja, S. (2015). Cyberbullying and self-esteem. Journal of School Health, 80(12), 614-621.

[50] Beran, T., & Li, Q. (2007). The relationship between cyberbullying and school bullying. Journal of Student Wellbeing, 1(2), 15-33.

[51] Kowalski, R. M., et al. (2014). Cyberbullying: Bullying in the digital age. John Wiley & Sons.

[52] Dooley, J. J., et al. (2009). Cyberbullying and its impact on young people's emotional health and well-being. British Journal of Educational Technology, 40(5), 879-887.

[53] Kwan, G. C. E., et al. (2018). The impact of online harassment on businesses: A case study of reputational damage. Journal of Business Ethics, 147(2), 309-322.

[54] Hinduja, S., & Patchin, J. W. (2015). Bullying beyond the schoolyard: Preventing and responding to cyberbullying. Sage Publications.

[55] Bovet, A., & Makse, H. A. (2019). Influence of fake news in Twitter during the 2016 US presidential election. Nature Communications, 10(1), 1-10.

[56] Guess, A., Nagler, J., & Tucker, J. (2019). Less than you think: Prevalence and predictors of fake news dissemination on Facebook. Science Advances, 5(1), eaau4586.

[57] Lewandowsky, S., Ecker, U. K., & Cook, J. (2012). Beyond misinformation: Understanding and coping with the "post-truth" era. Journal of Applied Research in Memory and Cognition, 6(4), 353-369.

[58] Wardle, C., & Derakhshan, H. (2017). Information disorder: Toward an interdisciplinary framework for research and policy making. Council of Europe report.

[59] Y. Vosoughi et al., "The spread of true and false news online," Science, vol. 359, no. 6380, pp. 1146-1151, Mar. 2018.

[60] D. Grinberg et al., "Fake news on Twitter during the 2016 U.S. presidential election," Science, vol. 363, no. 6425, pp. 374-378, Jan. 2019.

[61] C. Wardle and H. Derakhshan, "Information disorder: Toward an interdisciplinary framework for research and policy making," Council of Europe report, 2017.

[62] A. Guess et al., "Less than you think: Prevalence and predictors of fake news dissemination on Facebook," Science Advances, vol. 5, no. 1, 2019.

[63] N. B. Ellison, C. Steinfield, and C. Lampe, "The benefits of Facebook 'friends:' Social capital and college students' use of online social network sites," Journal of Computer-Mediated Communication, vol. 12, no. 4, pp. 1143-1168, 2007.

[64] K. N. Hampton et al., "Social networking sites and our lives: How people's trust, personal relationships, and civic and political involvement are connected to their use of social networking sites and other technologies," Pew Research Center, 2011.

[65] J. Vitak et al., "It's complicated: Facebook users' political participation in the 2008 election," Cyberpsychology, Behavior, and Social

[66] N. B. Ellison, C. Steinfield, and C. Lampe, "The benefits of Facebook 'friends:' Social capital and college students' use of online social network sites," J. Comput. Mediated Commun., vol. 12, no. 4, pp. 1143-1168, 2007.

[67] K. N. Hampton, L. F. Sessions, and E. J. Her, "Core networks, social media use, and civic and political engagement," Pew Research Center, 2011.

[68] J. Vitak, P. Zube, A. Smock, and C. T. Carr, "It's complicated: Facebook users' political participation in the 2008 election," Cyberpsychol. Behav. Soc. Netw., vol. 14, no. 3, pp. 107-114, 2011.

[69] N. B. Ellison, J. Vitak, R. Gray, and C. Lampe, "Cultivating social resources on social network sites: Facebook relationship maintenance behaviors and their role in social capital processes," J. Comput. Mediated Commun., vol. 19, no. 4, pp. 855-870, 2014.

[70] A. Chadwick, The Hybrid Media System: Politics and Power. Oxford University Press, 2013.

[71] D. Trottier and C. Fuchs, Social Media, Politics and the State: Protests, Revolutions, Riots, Crime and Policing in the Age of Facebook, Twitter and YouTube. Routledge, 2015.

[72] J. Bennett and N. Segerberg, The Logic of Connective Action: Digital Media and the Personalization of Contentious Politics. Cambridge University Press, 2013.

[73] S. Earl and C. Kimport, Digitally Enabled Social Change: Activism in the Internet Age. MIT Press, 2011.

[74] M. Castells, Networks of Outrage and Hope: Social Movements in the Internet Age. Polity Press, 2012.

[75] Z. Tufekci, Twitter and Tear Gas: The Power and Fragility of Networked Protest. Yale University Press, 2017.

[76] G. Siemens and P. Tittenberger, Handbook of Emerging Technologies for Learning. University of Manitoba, 2009.

[77] M. W. Allen and M. S. Seaman, Learning on Demand: Online Education in the United States. Babson Survey Research Group, 2011.

[78] R. E. West, Digital Literacy: A Primer on Media, Identity, and the Evolution of Technology. Routledge, 2016.

[79] J. D. Bolter and R. Grusin, Remediation: Understanding New Media. MIT Press, 2000.

[80] C. M. Reigeluth and A. A. Carr-Chellman, Instructional-Design Theories and Models: Building a Common Knowledge Base. Routledge, 2009.

[81] J. W. Turow, Media Today: Mass Communication in a Converging World. Routledge, 2017.

[82] D. Chaffey and C. Ellis-Chadwick, Digital Marketing: Strategy, Implementation and Practice. Pearson, 2019.

[83] P. Kotler and K. L. Keller, Marketing Management. Pearson, 2016.

[84] R. Ryan and R. Jones, Understanding Digital Marketing: Marketing Strategies for Engaging the Digital Generation. Kogan Page, 2019.

[85] T. M. Duffy and J. R. Kirkley, Learner-Centered Theory and Practice in Distance Education: Cases From Higher Education. Routledge, 2004.

[86] D. A. Aaker and E. Joachimsthaler, Brand Leadership. Free Press, 2000.

[87] R. T. Rust and R. W. Oliver, Service Quality: New Directions in Theory and Practice. Sage Publications, 1994.

[88] J. E. Grunig and T. Hunt, Managing Public Relations. Holt, Rinehart and Winston, 1984.

[89] C. Poole and F. Leavy, e-Commerce and Web Technologies: 18th International Conference, EC-Web 2017, Valencia, Spain, September 6-8, 2017, Revised Selected Papers. Springer, 2017.

[90] D. Chaffey and C. Wood, Business Information Management: Improving Performance Using Information Systems. Pearson, 2019.

[91] P. Kotler and K. L. Keller, Marketing Management. Pearson, 2016.

[92] R. Ryan and R. Jones, Understanding Digital Marketing: Marketing Strategies for Engaging the Digital Generation. Kogan Page, 2019.

[93] A. B. Hollander, Social Media Marketing: A Strategic Approach. Cengage Learning, 2015.

[94] S. H. Lee and A. O. F. Jones, Augmented Reality in Education: Cases, Places, and Practices. Springer, 2016.

[95] J. E. Grunig and T. Hunt, Managing Public Relations. Holt, Rinehart and Winston, 1984.

[96] S. Taddeo and L. Floridi, The Ethics of Information Warfare. Springer, 2014.

[97] R. E. Smith and C. Zook, Digital Geography: Geospatial Technologies in the Social Studies Classroom. IAP, 2016.

[98] M. Zimmer and N. K. Proferes, Internet Research Ethics for the Social Age: New Challenges, Cases, and Contexts. Peter Lang, 2015.

[99] A. B. Hollander, Social Media Marketing: A Strategic Approach. Cengage Learning, 2015.

[100] D. M. Shaw and A. M. Akhtar, Ethics and Global Security: A Cosmopolitan Approach. Routledge, 2014.

[101] M. E. Latzer and M. Shay, Regulating the Web: Network Neutrality and the Fate of the Open Internet. Routledge, 2018.

[102] A. B. Hollander, Social Media Marketing: A Strategic Approach. Cengage Learning, 2015.

[103] S. Taddeo and L. Floridi, The Ethics of Information Warfare. Springer, 2014.

[104] M. E. Latzer and M. Shay, Regulating the Web: Network Neutrality and the Fate of the Open Internet. Routledge, 2018.

[105] D. M. Shaw and A. M. Akhtar, Ethics and Global Security: A Cosmopolitan Approach. Routledge, 2014.

[106] R. E. Smith and C. Zook, Digital Geography: Geospatial Technologies in the Social Studies Classroom. IAP, 2016.

[107] S. Taddeo and L. Floridi, The Ethics of Information Warfare. Springer, 2014.

[108] J. Beninger and S. Robb, "Personalized marketing: A case study of Amazon," Journal of Direct, Data and Digital Marketing Practice, vol. 15, no. 3, pp. 170-183, 2014.

[109] T. Chen, Y. Li, and J. Li, "Blockchain-based supply chain finance: A case study of Maersk," Financial Innovation, vol. 3, no. 1, 2017.

[110] S. R. Steinhubl et al., "Watson, where are we? Cognitive computing in healthcare," Journal of the American College of Cardiology, vol. 67, no. 22, pp. 2625-2628,

[111] A. M. Abd-Alrazaq et al., "Crisis informatics in the context of the COVID-19 pandemic: A systematic literature review," JMIR Public Health and Surveillance, vol. 6, no. 4, 2020.

[112] M. A. Goodrich and R. Tamassia, "Introduction to computer security," Pearson, 2011.

[113] D. West et al., "Agile development: Mainstream adoption, a survey of systems development practices," IEEE Software, vol. 22, no. 6, pp. 27-33, 2005.

[114] J. Schmidhuber, "Deep learning in neural networks: An overview," Neural Networks, vol. 61, pp. 85-117, 2015.

[115] A. Al-Fuqaha et al., "Internet of Things: A survey on enabling technologies, protocols, and applications," IEEE Communications Surveys & Tutorials, vol. 17, no. 4, pp. 2347-2376, 2015.

[116] D. Tapscott and A. Tapscott, "Blockchain revolution: How the technology behind bitcoin is changing money, business, and the world," Penguin, 2016.

[117] R. D. Hansche et al., "Official (ISC)2 guide to the CISSP CBK," CRC Press, 2015.

[118] L. J. Cate, "The failure of notice and consent: A privacy analysis of the GDPR," International Data Privacy Law, vol. 8, no. 4, pp. 295-307, 2018.

[119] J. E. Miller and M. J. Potts, "The impact of digital literacy initiatives on student success," Computers & Education, vol. 59, no. 2, pp. 276-284, 2012.

[120] P. M. Bishop, "Collaborative cybersecurity information sharing and the public sector," IEEE Security & Privacy, vol. 14, no. 1, pp. 73-77, 2016.

[121] A. Heeks, "ICT4D 2.0: The next phase of applying ICT for international development," Computer, vol. 41, no. 6, pp. 26-33, 2008.

[122] A. B. Shilton, "Four billion little brothers?: Privacy, mobile phones, and ubiquitous data collection," Communications of the ACM, vol. 57, no. 1, pp. 25-29, 2014.

[123] M. Warschauer, "Technology and social inclusion: Rethinking the digital divide," MIT Press, 2003.

[124] J. Howcroft et al., "Researching information systems and computing," Sage, 2012

An Assessment of Power Regulation Strategy and Fuzzy Controller for a Grid-Interfaced Solar Photovoltaic System

Farook Shaik*
Department of EEE,
School of Engineering, MBU
Andhra Pradesh-517 102, India
farook.s@vidyanikethan.edu

B. Maheswari
Department of EEE,
School of Engineering, MBU
Andhra Pradesh-517 102, India
budisemaheswari@2020gmail.com

G. Naga Anitha
Department of EEE, ,
School of Engineering, MBU
Andhra Pradesh-517 102, India
anithagutam1257@gmail.com

D. Charitha
Department of EEE,
School of Engineering, MBU
Andhra Pradesh-517 102, India
dharmavarapucharitha@gmail.com

Malligunta Kiran Kumar
Department of EEE,
Koneru Lakshmaiah Education Foundation, Guntur, Andhra Pradesh-522302, India
kiran.malligunta@gmail.com

B.V.Sai Thrinath*
Department of EEE,
School of Engineering, MBU
Andhra Pradesh-517 102, India
connectbvst@gmail.com

***Abstract*— The emphasis of this study is to provide a fundamental real power regulatory control scheme that can effectively deliver frequency control function to a single-phase grid, without the reliance on an energy storage system. The reserve share of power is supplied to the real power regulation controller by the power regulation system. The appropriate fraction can be used to determine the quantity of reference power that can be extracted from the PV panels and fed to the grid. Using a PI Controller will result in slower system responsiveness due to the proportional and the combined derivative and integral controllers. Therefore, a fuzzy logic controller used instead of a PID controller to achieve faster response time. The power quality of these controllers is superior to that of traditional controllers, resulting in improved system performance. In the MATLAB/SIMULINK environment, the controller's performance on the system is modelled and simulated.**

Keywords— Active *power regulation, single-phase grid, reserve fraction, PV panel, fuzzy logic controller, maximum power point (MPP), power regulation*.

I. INTRODUCTION

Addressing environmental concerns associated with fossil fuels requires the utilization of renewable energy sources to meet the continuous escalating demand. Wind turbines and solar photovoltaics stand out as highly promising technologies in this regard. However, to optimize their performance across diverse operational conditions, effective power control strategies are indispensable, but in a variety of operational environments, effective power control strategies are crucial to maximizing their efficacy. Finding a way to regulate the frequency steadily in grid-connected, single-phase solar systems without using energy storage devices is the main objective of this research. Traditionally, a PI controller's ability to regulate real power injection is dependent on a reserve percentage that the plant operator supplies. However, alternative methods must be explored due to the PI controller's slow reaction time. As an extension strategy to improve system performance, this study investigates the implementation of a fuzzy logic controller (FLC).In grid integrated PV systems, the FLC enhances power quality and frequency management with its faster response times and better dynamic response compared to the PI controller. To validate the efficacy of the FLC, the controller is modelled and simulated and is compared with the conventional PID controller. The increasing importance of renewable energy sources in addressing environmental concerns makes solar photovoltaic (PV) system grid interconnection imperative. The functioning of a fuzzy logic controller is implemented on a single-phase grid-interfaced photovoltaic system with a power regulation scheme. The primary objective is to optimize the electricity output while maintaining reliable grid connectivity. Through simulated experiments, it is to evaluate how well the FLC performs in terms of improving power quality, grid stability, and energy efficiency when compared to traditional control approaches. Through enhanced control algorithms in grid-connected PV systems, this research advances sustainable energy solutions.

Analysis of grid-connected solar panels has been done by many authors. In [1-3] The study shows that the solar energy calibre of the Tibetan Plateau is enormous, even exceeding certain top-of-the-atmosphere estimations. Even with clouds and aerosols present, which can lower yearly radiation by 3–6% and up to 23% combined, the reduction is negligible when compared to other parts of the world. This emphasizes how well-suited the plateau is for the growth of solar energy, providing a viable path to supplying future energy demands while reducing CO2 emissions. Two different kinds of grid-connected solar panel power converters are examined in this study. In [4-5] They compare the efficiency of these converters with various control schemes. They discover that a particular kind, dubbed CSI7, requires an additional part but is superior at cutting down on wasted energy and enhancing performance. In [6-8] By adding virtual inertia, the PV-Virtual Synchronous Generator (PV-VSG) assists PV systems without energy storage in contributing to grid stability. They present the Active Power Reserve, which reserves power for frequency modulation using a variety of techniques. With the help of these techniques, PV stations can support grid stability without requiring intricate management. Simulation findings validate their efficacy. In [9-11] This study describes a transformer less dual-buck, semi-two-stage ground type grid-connected inverter for photovoltaic (PV) systems. Common mode leakage current is eliminated via the ground topology, which connects the PV's negative terminal to the neutral of the grid. This gets around the stray capacitance of the PV. The DC-link is shielded from short circuits by the dual-buck based architecture, which increases resilience. Multi-level features are provided by the semi-two-stage architecture, which lowers the forward voltage drop and reduces the dv/dt. Higher efficiency and reliability are achieved with this design in comparison to traditional two-stage inverters. High efficiency and low leakage currents are demonstrated by the experimental findings from a 1.5kW prototype. In [12-14] Through output reduction, the PV system keeps part of its power up/down capabilities. Nevertheless, energy is lost and the reserve power is not utilised effectively. In [16-18] Grid-connected PV systems should use MPPT to optimise energy production. Encouraging additional PV installations necessitates improving rules and power control mechanisms to prevent negative effects from PV systems, like burden on the grid. In [19-20] A novel method that offers simplicity, speed, and accuracy for precisely determining a solar array's maximum power point is presented in the work. The efficacy and feasibility of the REC-AE220 solar modules have been verified by simulation and experimental results. In [21-22] Using a fuzzy controller and a unique kind of converter, this research provides a novel approach to maximise power production from a photovoltaic (PV) system. It ensures that the PV system produces as much as possible of power by adjust the converter's parameters in response to sunlight. This technology tops typical approaches, as shown by experiments and simulations.

II. MATHEMATICAL MODELLING

A. *Active Power Regulation Scheme Proposal*

The control methodology employed to align the power generation of the solar photovoltaic lay out with the requirements of the electrical grid is commonly referred to as active power regulation. This strategy is specifically designed for solar PV systems that are interfaced to the grid. The stability and reliability of the grid can be maintained by managing the power supply, thereby enabling the solar photovoltaic (PV) system to achieve its MPP. An algorithmic control scheme, such as the Fuzzy Logic Controller (FLC), is employed in an real power regulation scheme to consistently monitor the operational attributes of the solar photovoltaic (PV) array and make necessary adjustments to the power output . The FLC considers various factors to ascertain the most feasible operating position for the PV array, including temperature, load demand, and solar radiation levels. The implementation of an active power regulation scheme has numerous benefits, including the enhancement of the complete functioning of the PV system. The technology can enhance the energy output of the solar PV system by allowing the array to run at its MPP. The maintenance of power supply dependability and the prevention of grid disruptions are contingent upon this factor.

Active power regulation is a control mechanism utilized to adjust the output of a grid-connected solar photovoltaic (PV) array in accordance with the requirements of the grid. The primary objective of active power regulation is to maintain equilibrium and dependability of the electrical grid, while simultaneously enabling the solar PV system to function at its maximum power point. In order to monitor the performance of the solar photovoltaic (PV) array and make necessary adjustments to the power output, real power regulation schemes frequently employ algorithmic control systems such as the Fuzzy Logic Controller (FLC) [23]. The Finite Loss Curve (FLC) considers multiple variables, including temperature, load demand, and solar radiation levels, to ascertain the most feasible operating position for the photovoltaic (PV) array. An active power regulation scheme offers a significant benefit by enhancing the overall performance of the solar photovoltaic (PV) system. By optimizing the power output of the solar photovoltaic (PV) array at its MPP, this methodology has the potential to enhance the energy yield of the array. In order to provide uninterrupted electricity supply and prevent grid failure, this is important. The Control scheme for a two-stage, single-phase solar system that is grid-connected is depicted in figure 1.

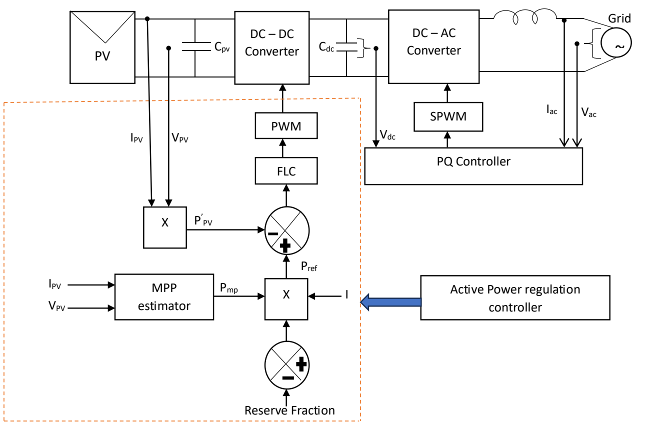


Fig. 1 Control scheme for a two-stage, single-phase solar grid-connected PV system

The principle of real power regulation is to follow the solar PV array's maximum power point by regulating the duty cycle of a DC-DC boost converter. To transfer power to the grid efficiently, the boost converter was used to adjust the output voltage of the PV array so that it matches the grid voltage. The following formula is used to determine the PV array's power output:

$$P_{out} = V_{mpp} \times I_{mpp} \tag{1}$$

The voltage and current at the maximum power point (MPP) are denoted as V_{mpp} and I_{mpp} respectively. An approximation of the voltage at the MPP is:

$$V_{mpp} \approx V_{oc} - I_{mpp} \times R_{pv} \tag{2}$$

PV array's open-circuit voltage is denoted by V_{oc}, and its equivalent series resistance is represented by R_{pv} . An approximation of the current at the MPP is:

$$I_{mpp} \approx \frac{V_{oc}}{R_{pv}} \tag{3}$$

The duty cycle of a boost converter D, is computed with the following formula:

$$D = \frac{V_{in} - V_{mpp}}{V_{in}} \tag{4}$$

where V_{in} is the converter's input voltage and the boost converter's power output as follows:

$$P_{boost} = V_{in} \times I_{boost} \tag{5}$$

where the converter's output current is represented by the expression I_{boost}. The inverter's power output, represented by P_{out}, can be computed as follows:

$$P_{out} = \frac{P_{boost}}{\eta_{inverter}} \tag{6}$$

where grid-tied inverter $\eta_{inverter}$ is the inverter's efficiency.

B. Maximum power point (MPP) estimation

In order to maintain optimal power generation in photovoltaic (PV) systems, Maximum Power Point Tracking (MPPT) is essential for continuously modifying the solar panels' operating point. The PV system functions most effectively at its maximum power point (MPP), which maximizes power output under specific circumstances including temperature and sunshine intensity. Algorithms are used to track the MPP in real-time during estimation. The Perturb and Observe (P&O) approach is one popular algorithm.

A one-diode model is the most popular type of PV stack. The output current and output voltage relationship can be expressed as [10]:

$$I = I_{ph} - I_s\left(exp\left(\frac{V + R_s I}{naV_t}\right) - 1\right) - \frac{V + R_s I}{R_P}. \tag{7}$$

The PV output voltage and current are denoted by I and V, respectively. A semiconductor's thermal voltage is represented by V_t, its p–n junction quality factor by a, its dark current by I_s, and its produced current by I_{ph}. The series and parallel resistances of the module are denoted by Rs and R_p, respectively. Lastly, the number of series-connected cells in the module is indicated by n.

At the MPP, there is no power derivation in relation to voltage. Consequently:

$$I_m = \frac{(V_m I_s/naV_t)exp((V_m + R_s I_m)/naV_t) + V_m / R_p}{1 + (R_s I_s/naV_t)exp((V_m + R_s I_m)/naV_t) + \frac{R_s}{R_p}} \tag{8}$$

Ignoring the impact of tiny and big resistances, R_s and R_p respectively, is one technique to get around this problem. Thus, (1) and (2) may be rephrased as:

$$I_1 = I_{ph} - I_s\left(exp\left(\frac{V_1}{naV_t}\right) - 1\right) \tag{9}$$

$$\boldsymbol{I_{1m}} = \frac{V_1 I_s}{naV_t}\, exp\left(\frac{V_{1m}}{naV_t}\right) \tag{10}$$

where I_1 and V_1 stand for the ideal current and voltage at the module terminals. The MPP voltage and current of the ideal model are, respectively, V_{1m} and $\boldsymbol{I_{1m}}$. Apparently, (9) and (10) are much simpler to solve. In order to identify an analytical fix, X is defined as:

$$X = \frac{V_{1m}}{naV_t} \tag{11}$$

The following equation is found by substituting $\boldsymbol{I_{1m}}$ in (8) from (9) while taking (11) into consideration:

$$(X + 1)exp(X) = \frac{I_{ph} + I_s}{I_s} \tag{12}$$

Since I_s is significantly less than I_{ph}, it can be disregarded in the nominator. Thus, X solution is:

$$X = W\left(\frac{I_{ph}\exp(1)}{I_s}\right) \tag{13}$$

where the following function is the inverse of the Lambert function, represented by the letter W:

$$y = xexp(x). \tag{14}$$

The simplified-module model's MPP can be found by using equations (13) and (14).

$$V_{1m} = naV_t\left(W\left(\frac{I_{ph}\exp(1)}{I_s}\right) - 1\right) \tag{15}$$

I_{ph} in (15) can be calculated using SCC in the following way:

$$I_{ph} = I_{sc}\frac{R_s + R_p}{R_p} \tag{16}$$

There is insufficient accuracy in (15). By accounting for the current flow in the parallel resistor and the voltage drop across the series resistor, (15) can be made much more accurate. Consequently, using the following formulas, the system MPP parameters may be determined:

$$I_m = I_{1m} - \frac{V_{1m}}{R_p} \tag{17}$$

$$V_m = V_{1m} - \left(I_{1m} - \frac{V_{1m}}{R_p}\right)R_s \tag{18}$$

To assess how successful the suggested approach is, η is defined as:

$$\eta = \frac{V_m I_m}{P_{m-real}} \times 100 \tag{19}$$

C. PQ Controller

Power quality controllers (PQ controllers) manage active and reactive power flow in power systems to meet specified criteria. They utilize proportional-integral (PI) controllers to adjust power converters and inverters' output based on error between measured and reference values. This ensures system voltage, frequency, and power factor stay within predefined bounds, improving energy efficiency and reliability while reducing voltage-related issues. A PI controller's transfer function is shown as follows:

$$G_c(S) = K_p + \frac{K_i}{s} \tag{20}$$

The following formulas can be used to represent the PQ controller's active and reactive power output:

$$V_g = \begin{bmatrix} V_\alpha \\ V_\beta \end{bmatrix} = \begin{bmatrix} V(\omega t) \\ V\left(\omega t + \left(\frac{\pi}{2}\right)\right) \end{bmatrix} \tag{21}$$

$$I_{load} = \begin{bmatrix} I_\alpha \\ I_\beta \end{bmatrix} = \begin{bmatrix} I(\omega t) \\ I\left(\omega t + \left(\frac{\pi}{2}\right)\right) \end{bmatrix} \tag{22}$$

The instantaneous real power ($P_{\alpha\beta}$) and reactive power ($Q_{\alpha\beta}$) are then defined by the pq theory in terms of the α–β components as:

$$P_{\alpha\beta} = V_\alpha I_\alpha + V_\beta I_\beta \tag{23}$$

$$Q_{\alpha\beta} = V_\beta I_\alpha - V_\alpha I_\beta \tag{24}$$

According to pq theory, the reference current in the α–β coordinate is given by:

$$I_{REF(\alpha,\beta)} = \frac{1}{V_{\alpha,\beta}^2}\begin{bmatrix} V_\alpha & -V_\beta \\ V_\beta & V_\alpha \end{bmatrix}\begin{bmatrix} P^* \\ Q^* \end{bmatrix} \tag{25}$$

In this $V_{\alpha,\beta}^2 = V_\alpha^2 + V_\beta^2$ where Q* is the reference reactive power and P* is the reference real power calculated by adding P_{mpp} and $P_{\alpha\beta}$.

D. Fuzzy Logic Controller

A precise mathematical model is not necessary for the operation of the fuzzy logic controller (FLC), and it is able to successfully manage nonlinearity and implicit inputs. Fuzzy also outperforms the norm for non-linear controllers. There are four components: the inference method, the rule basis, fuzzification, and defuzzification. In this inquiry, FLC is fed the error, E, and the change in error, CE, at sample time k, which are specified by equations (26) and (27). The duty cycle D, Equation (1,2), is the output of an FLC. The following two input factors are identified by:

$$E(K) = \frac{P(K) - P(K-1)}{V(K) - V(K-1)} \tag{26}$$

$$CE(K) = E(k) - E(K-1) \tag{27}$$

A PV module's voltage and power are denoted by V(k) and P(k), respectively.

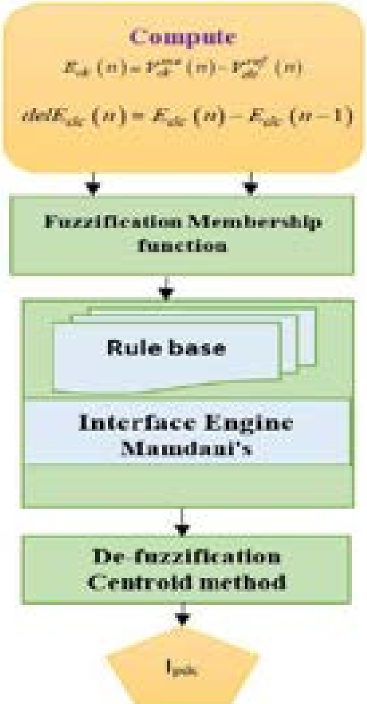


Fig 2.Fuzzy Logic Controller

There are three steps in the fuzzy controller:

1. *Fuzzification:*
   Fuzzification is a crucial stage in fuzzy logic controllers, allowing the interpretation of imprecise or ambiguous data by converting exact inputs into fuzzy sets, or degrees of membership. By handling inputs such as human language descriptors, this approach makes control systems more adaptive and makes reasoning with ambiguous data easier.

2. *Inference method:*
   Fuzzy logic controllers (FLCs) utilize fuzzy inference, with the Mamdani method emphasizing linguistic interpretability in its fuzzy outputs, while the Sugeno method prioritizes computational efficiency by producing crisp numerical results. The option between the two methods depends on the specific requirements of the control problem, balancing interpretability and computational speed.

TABLE1: FUZZY INFERENCE

E / CE	NB	ZE	PB
NB	PB	NS	NB
ZE	PS	ZE	NS
PB	NB	ZE	ZE

3. *Defuzzification:*
   Defuzzification is the process of turning a fuzzy inference system's fuzzy output into an exact numerical value for system control or decision-making, making sure fuzzy logic controllers can function properly in practical settings.

III. SOLAR PV MODULE OPERATION

In order to convert sunlight into energy for use in both home and commercial lighting systems, solar photovoltaic (PV) modules, which have wattages ranging from 50 to 350, are essential. They are made of semiconductor materials such as crystalline silicon and have 1-2 watt photovoltaic cells coupled to each other. By lowering dependency on fossil fuels, lessening their negative effects on the environment, and allowing excess energy to be reinjected into the system, these modules improve energy efficiency and advance a decentralized energy infrastructure. Solar cells have a special part called a p-n junction that separates charges, creating electricity. This electricity can power various things. Solar panels, made of many cells, are connected together to make more power and can be placed on buildings or structures.

Fig 3.solar pv panel

TABLE 2: PV MODULE SPECIFICATIONS

Parameters	Units	Values
OC Voltage	Volts	37.3
SC Current	Amps	8.66
Max Current	Amps	30.7
Max Voltage	Volts	8.15
Max Power	Watts	250.205
Resistance in series	Ohms	0.2372
Resistance in shunt	Ohms	224.18
Quality factor (D)	----	1.019

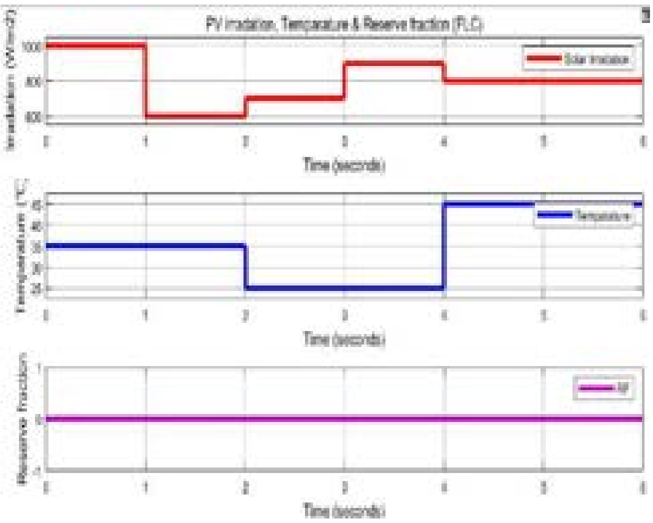


Fig 4. Variable Irradiance, variable Temperature and Reserve Fraction Zero

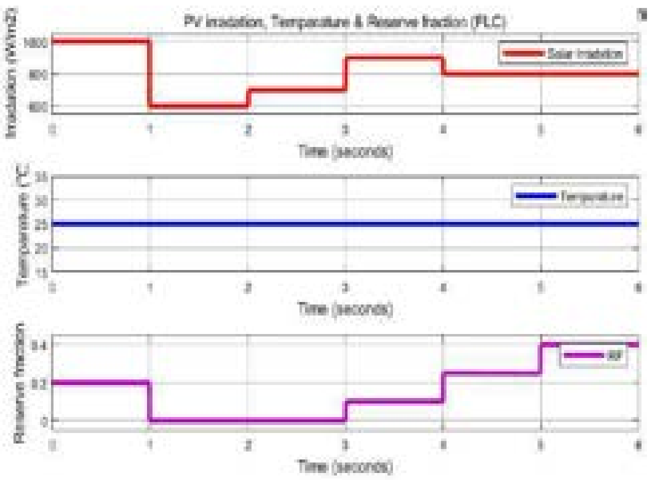


Fig 5. Variable Irradiance, Temperature constant and Reserve Fraction variable

IV. SIMULATION RESULTS

A 250 Wp solar PV system's proposed MPP estimator and controller is tested using a Matlab/Simulink model of the overall block diagram depicted in Fig. 1. In order to conduct the simulation study, the key system parameters are listed in Table 1. By running simulations of the three cases shown in the table, we can confirm that the proposed method works as intended.

.

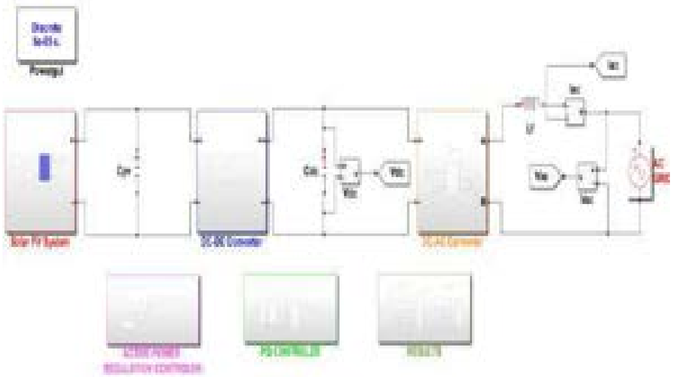


Fig 6. Two-stage grid-connected single-phase solar-PV system with control logic

First, the controller's transient reaction is tested in a range of environmental factors, including temperature and radiation. The actual grid-injected .Power was simulated under different radiation conditions with a fixed panel temperature of 25 °C and a 0% reserve fraction. As shown in Fig. 8, Following the proposed method, the inverter would inject the same amount of power into the AC grid as the predicted maximum power related to the irradiation.

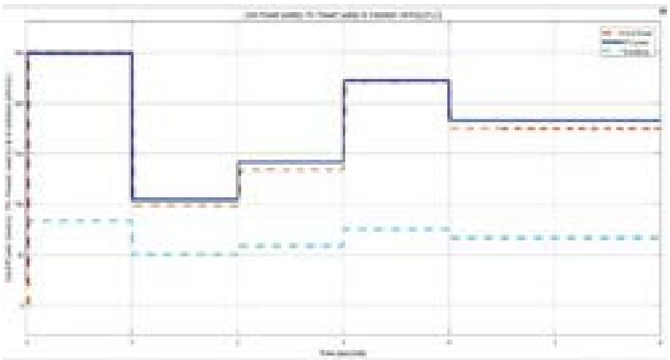


Fig 7. Variation in grid power and maximum power estimation under different radiation conditions at 25 °C with no reserve fraction

Changes in radiation levels and the zero reserve fraction cause variations in the MPP and grid-injected electricity. Figure 9 displays the temperature plot. When this occurs, the tracking speed of the controller is sufficient to determine the MPP and transmit the result to the grid.

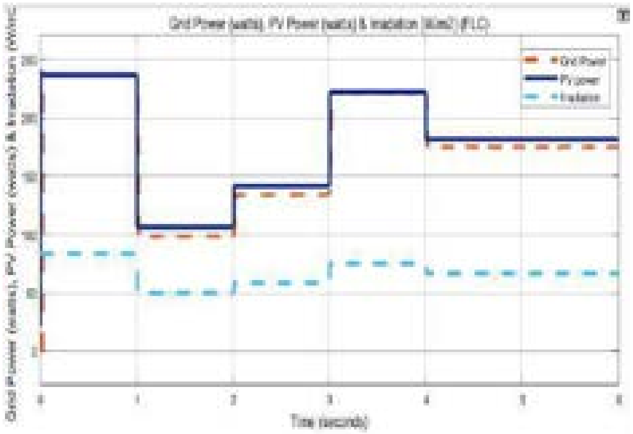


Fig 8. Variation in grid power and maximum power estimation at different temperatures and radiation levels with 0% reserve fraction

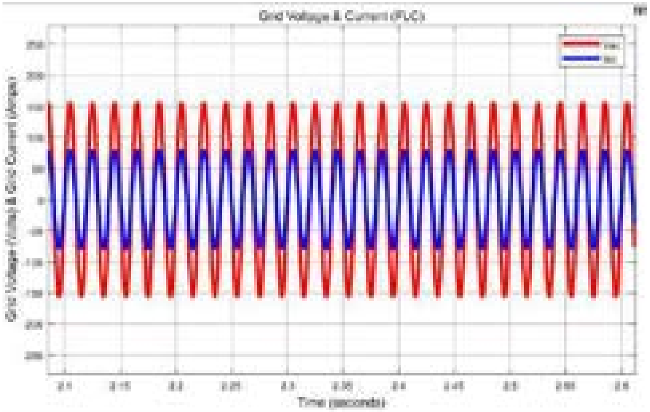


Fig 9. Voltage and current injected into the grid under various radiation circumstances at 25 °C without reserve fraction

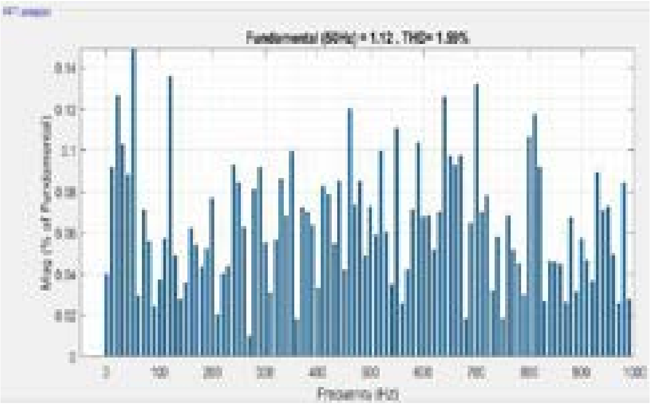


Fig 10.Total Hormonic Distortions in the system

V. CONCLUSION

Lastly, a power regulation scheme and a fuzzy logic controller (FLC) significantly improved the overall efficiency of the two-stage, single-phase, grid-connected solar PV system, according to the performance analysis. The FLC regulates power flow efficiently, which guarantees a seamless energy conversion process and grid integration. The two-stage setup increases control and durability over a wide range of weather conditions. Power quality, stability, and reduced fluctuations are all enhanced by the proposed approach, according to extensive simulations and analyses. The FLC's ability to adapt to changing solar conditions makes it a crucial component in maximizing energy extraction and grid integration. In conclusion, this study shows that FLC-based control techniques have promising potential for improving the performance of grid-connected solar PV systems.

REFERENCES

[1] N Gelsora, N Gelsorb, T Wangmob, et al. Solar energy on the Tibetan plateau: Atmospheric influences. Solar Energy, 2018, 173: 984-992.

[2] E Lorenzani, G Migliazza, F Immovilli, et al. CSI and CSI7 current source inverters for modular transformerless PV inverters. Chinese Journal of Electrical Engineering, 2019, 5(2): 32-42.

[3] X Zhang, Q Gao, Y Hu, et al. Active power reserve photovoltaic virtual synchronization control technology. Chinese Journal of Electrical Engineering, 2020.

[4] Thrinath, BV Sai, S. Prabhu, and B. Meghya Nayak. "Power quality improvement by using photovoltaic based shunt active harmonic filter with Z-source inverter converter." Electrical Engineering & Electromechanics 6 (2022): 35-41.

[5] A Sangwongwanich, Y Yang, F Blaabjerg. A sensorless power reserve control strategy for two-stage grid-connected PV systems. IEEE Transactions on Power Electronics, 2017, 32(11): 8559-8569.

[6] T. Stetz, F. Marten, and M. Braun, "Improved low voltage grid integration of photovoltaic systems in Germany," IEEE Trans. Sustain. Energy, vol. 4, no. 2, pp. 534–542, Apr. 2013.

[7] G Farivar, B Asaei, S Mehrnami. An analytical solution for tracking photovoltaic module MPP. IEEE J. Photovoltaics, 2013: 1053-1061.

[8] R. Rajkumar, N. Selvarani, B. B. Bright, S. Kaliappan, B. V. Sai Thrinath and A. X. A. Rebin, "Photovoltaic System based on Closed Loop DC-DC Converter with Fuzzy Logic Controller," 2023 4th International Conference on Electronics and Sustainable Communication Systems (ICESC), Coimbatore, India, 2023, pp. 113-119.

[9] Zafar, S.; Amin, M.A.; Javaid, B.; Khalid, H.A. On Design of DC-Link Voltage Controller and PQ Controller for Grid Connected VSC for Microgrid Application. In Proceedings of the 4th International Conference on Power Generation Systems and Renewable Energy Technologies, PGSRET 2018, Islamabad, Pakistan, 10–12 September 2018.

[10] Srilakshmi, Koganti, S. Poorna Chander Rao, G. Deepika, BV Sai Thrinath, Alapati Ramadevi, Kongari Dushanth Kumar, and Aravindhababu Palanivelu. "Performance Analysis of Artificial Intelligence Controller for PV and Battery Connected UPQC." International Journal of Renewable Energy Research (IJRER) 13, no. 1 (2023): 155-170.

[11] X Zhang, Q Gao, Y Hu, et al. Active power reserve photovoltaic virtual synchronization control technology. Chinese Journal of Electrical Engineering, 2020.

[12] C. Larbes, S.M. Aı ̈t Cheikh*, T. Obeidi, A. Zerguerras. Genetic algorithms optimized fuzzy logic control for the maximum power point tracking in photovoltaic system.

[13] X Li, H Wen, Y Zhu, et al. A novel sensor less photovoltaic power reserve control with simple real-time MPP estimation. IEEE Transaction on Power Electronics, 2019, 34(8): 7521-7531.

[14] S. B. J, S. Benvin, A. R, A. V, R. Maranan and B. V. S. Thrinath, "AI Based Electric Automobile Battery Drain Forecasting System," 2023 4th International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2023, pp. 756-761.

[15] E Batzelis, G Kampitsis, S Papathanassiou. Power reserves control for PV systems with real-time MPP estimation via curve fitting. IEEE Transactions on Sustainable Energy, 2017, 8(3): 1269 -1280.

[16] T W Zheng, L J Chen, W Liu, et al. Multi-mode operating control for photovoltaic virtual synchronous generator considering the dynamic characteristics of primary source. Proceedings of the CSEE, 2017, 37(2): 454-463.

[17] H. Xin, Y. Liu, Z. Wang, D. Gan, and T. C. Yang, "A new frequency regulation strategy for photovoltaic systems without energy storage, "IEEE Trans. Sustain. Energy, vol. 4, no. 4, pp. 985–993, 2013.

[18] B. V. S. Thrinath, P. S. Sathish, T. Manikanta, B. Harish, T. V. V. N. P. S. Harshath and T. Ramakrishna, "A Novel Dual Source Inverter for HES System Using SVM," 2023 Second International Conference on Trends in Electrical, Electronics, and Computer Engineering (TEECCON), Bangalore, India, 2023, pp. 69-74.

[19] Y. Tang, W. Yao, P. C. Loh, and F. Blaabjerg, "Highly reliable transformer less photovoltaic inverters with leakage current and pulsating power elimination," IEEE Trans. on Ind. Electron., vol. 63, no. 2, pp. 1016-1026, Sep. 2016.

[20] Farook, S., Sekhar, A. H., Reddy, K. B. N. K., Kumar, N. M. G., & Deepa, S. (2023, May). Fuzzy Based Intelligent Control Strategy to Mitigate Power Quality Issues in an Interconnected PV System. In *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)* (pp. 1602-1607). IEEE.

[21] Farook, S., Malligunta Kiran Kumar, A. Hema Sekhar, and K. Balaji Nanda Kumar Reddy. "A Fuzzy Logic-Based Control Strategy for Power Balance and Enhance Power Quality in a Grid-Tied PV System." In *2023 3rd International Conference on Intelligent Technologies (CONIT)*, pp. 1-5. IEEE, 2023.

[22] A. Woyte, V. V. Thong, R. Belmans, and J. Nijs, "Voltage fluctuations on distribution level introduced by photovoltaic systems," IEEE Trans.on Energy Convers., vol. 21, no. 1, pp. 202–209, Mar. 2006.A.

[23] Woyte, V. V. Thong, R. Belmans, and J. Nijs, "Voltage fluctuations on distribution level introduced by photovoltaic systems," IEEE Trans.on Energy Convers., vol. 21, no. 1, pp. 202–209, Mar. 2006.

Implementation of a Decentralised Platform for Digital File Verification and Sharing using Solana Blockchain and IPFS

Ayush Goel
Department of Networking and Communications
SRM Institute of Science and Technology
Chennai, India
ag7626@srmist.edu.in

Madhur Jain
Department of Networking and Communications
SRM Institute of Science and Technology
Chennai, India
mj2947@srmist.edu.in

R.Kayalvizhi
Department of Networking and Communications
SRM Institute of Science and Technology
Chennai, India
kayalvir@srmist.edu.in

***Abstract*— In recent years, the usage of digital files has increased and therefore data theft has also increased. Traditional centralized platforms often suffer from vulnerabilities such as single points of failure and lack of transparency. In response, decentralized technologies such as Interplanetary File System (IPFS) and blockchain have emerged as promising solutions. This research study presents a novel decentralized platform built upon the Solana blockchain and IPFS, designed to address the shortcomings of centralized systems. The proposed platform uses the transparency and immutability of the Solana blockchain network to provide a secure ledger for file verification. In addition to verification, the platform facilitates seamless sharing of digital files among users through IPFS, a distributed file system. By utilizing IPFS's decentralized architecture, data is stored across a chain of nodes, eliminating a single centralized server. It not only enhances data resilience and availability but also mitigates the risk of censorship and data manipulation. This paper proposes using Solana blockchain instead of traditional Ethereum blockchain for better scalability because Solana supports more than 2000 TPS (transaction per second) and reduced network fee.**

Keywords*— *Blockchain; File Sharing; Solana; Smart Contracts; Inter Planetary File Storage

I. INTRODUCTION

In today's digital era, the proliferation of data and the increasing need for secure and efficient file sharing mechanisms have propelled the exploration of decentralized technologies. The emergence of blockchain technology has revolutionized various industries by offering immutable, transparent, and decentralized solutions to long-standing challenges. Among these challenges, the verification and sharing of digital files stand out as crucial areas where traditional centralized systems often fall short in case of efficiency, trustworthiness and security.

The use of Solana network and IPFS presents a promising solution to address the challenges associated with centralized file verification and sharing systems. By decentralizing the storage and verification process, the platform eliminates single points of failure, enhances data integrity, and fosters trust among participants without the need for intermediaries. Furthermore, the use of smart contracts on the Solana blockchain enables the automation of file verification processes, ensuring transparency and immutability throughout the lifecycle of digital files.

A. History

File storage and sharing have evolved a lot since the advent of computers. Initially, files were stored on physical media such as floppy disks and later, hard drives. However, with the rise of the internet, the concept of file sharing underwent a drastic change. Peer-to-peer (P2P) file sharing emerged as a prominent concept, allowing users to share files directly between their computers. This was followed by the introduction of centralized file storage and sharing services like Napster, which revolutionized the music industry by enabling users to share MP3 files online. However, legal challenges led to the decline of Napster and the rise of decentralized alternatives like BitTorrent. The early 2000s saw the emergence of cloud storage services such as Dropbox, Google Drive, and Microsoft OneDrive. These platforms offered users the ability to store files remotely and access them from any device with an internet connection. As internet speeds improved and storage costs decreased, cloud storage became affordable, transforming the way individuals and businesses manage their data. Today, file storage and sharing continue to evolve with advancements in technology such as blockchain-based decentralized storage solutions and the integration of artificial intelligence for more efficient file organization and retrieval.

B. Motivation

The reason behind developing Solana powered file storage and sharing system is to develop a decentralized platform that leverages the power of Solana blockchain and IPFS for digital file verification and sharing. Solana blockchain is known for its high throughput and low transaction costs and it provides a robust infrastructure for maintaining a secure and scalable ledger of file transactions. On the other hand, IPFS offers a decentralized and distributed file system, enabling data to be stored and accessed from a decentralised network and also ensures redundancy.

Furthermore, the integration of Solana network and IPFS transcends the problems of traditional storage systems by developing an ecosystem that prioritizes user autonomy and data sovereignty. By empowering users with control over their files and eliminating single points of failure, the platform fosters a resilient digital infrastructure that is inherently resistant to censorship and manipulation. Conceptually, integrating blockchain into a file sharing system allows for several key benefits. First, blockchain provides a decentralized storage mechanism for metadata such as file ownership, access permissions, and transaction history. Each transaction (e.g., uploading a file, sharing with another user, granting access) can be recorded as a transaction on the blockchain, creating a transparent and auditable record of file-related activities. This transparency enhances trust among participants in the file sharing network. Additionally, blockchain's consensus mechanism ensures that the recorded transactions are tamper-resistant, providing a high level of security and integrity to the file sharing system. Furthermore, blockchain-based smart contracts can automate and enforce rules governing file sharing, ensuring that access permissions are managed efficiently and securely without relying on a central authority.

C. *Objectives*

The main objective of this project is to explore the development and implementation of a decentralized platform for digital file verification and sharing utilizing the Solana blockchain and IPFS (Interplanetary File System). The project elucidates the potential benefits of utilizing Solana blockchain and IPFS technologies in the context of digital file management.

D. *Scope And Application*

The scope of the Decentralized Platform for Digital File Verification and Sharing utilizing Solana Blockchain and IPFS is to enable secure and immutable verification of digital files, leveraging the speed and scalability of Solana Blockchain. Its application spans across sectors requiring transparent and tamper-proof data sharing such as healthcare, supply chain, and legal sectors. Through IPFS integration, it ensures efficient and decentralized file storage, improving censorship resistance and data integrity. This platform empowers users with trust-less verification mechanisms and seamless sharing capabilities, enhancing data reliability and accessibility in a decentralized ecosystem. The Platform considers various security issues such as Data Privacy and Confidentiality, immutability, Data Integrity.

II. TERMINOLOGY

A. *ReactJS*

React.js is an open-source solution for developing frontend interfaces that helps developers to make web applications efficiently.

B. *Smart Contract*

A smart contract is a self-executing contract with the terms of the agreement between two or more parties directly written into code. It operates on a blockchain, such as Solana, and automatically enforces, verifies, or facilitates the negotiation and performance of a contract.

C. *Transaction*

In web3, a transaction represents a digital interaction on a blockchain, involving the exchange of data or value between participants, often secured by cryptographic signatures and recorded on a public ledger.

D. *Proof of History Mechanism*

It is a novel consensus mechanism which uses timestamps and order of events for validation of transaction. It is very fast as compared to proof of work algorithm.

E. *Solana*

Solana is an innovative blockchain platform, it allows fast transactions as compared to other blockchains. It also have low fees and high efficiency which lowers wait time. It uses a consensus mechanism called Proof of History.

III. LITERATURE REVIEW

This section contains the prior research that has been done on Blockchain and File sharing.

A. *Related Work*

In research [1], the authors investigated the utilization of decentralised ledger technology for data storage, emphasizing its potential to offer secure and decentralized solutions. They highlighted blockchain's ability to reduce data leaks and unauthorised access, while concurrently making sure that availability of service is not affected. With the help of decentralised technology and its features, like immutable ledger and distributed network nodes, the study underscores its role in enhancing the security and reliability of file storage systems, thereby addressing critical concerns in data management and privacy.

In research [2], the authors investigated the potential applications of distributed ledger technology in decentralized data maintenance and retrieval. They proposed its capability to establish immutable methods for data transfer and retrieval. By leveraging blockchain's immutable ledger and decentralized nature, the study suggests a promising solution to enhance data security and integrity while facilitating efficient sharing across distributed networks.

In research [3], the authors investigated data management systems which can be used for decentralised ledger technology, analysing their features and limitations. They explored various use cases and discussed future directions for these systems, highlighting their potential impact on decentralized applications and distributed data storage.

In research [4], the authors introduced a proof of delivery protocol utilizing smart contracts and blockchain technology. They proposed a mechanism where participating entities, including the owner and file server, receive incentives as rewards upon successful content delivery to the customer. This protocol aims to enhance transparency and reliability in digital content distribution while incentivizing stakeholders for their contributions.

In research [5], the authors presented a novel approach to file sharing and access control utilizing blockchain technology and smart contracts. Their solution employed a private decentralised network to maintain and administer data, with access policies validated through contracts. By integrating blockchain's immutable ledger and smart contract capabilities, the system offers enhanced security and transparency in managing access to shared files, addressing critical concerns surrounding data privacy and authorization.

In research [7], the authors analyse how distributed ledger technology enhances affects the privacy of data being stored in the ledger. In research [8], the authors introduces a distributed data storage platform for cloud computing.

Some researchers have discussed about IPFS that aims to create a more resilient and censorship-resistant web infrastructure. It converts data into hash so that it can be shared across a decentralized network of nodes, reducing reliance on centralized servers. They also have addressed security issues in blockchain-based systems, including data confidentiality and authentication. Additionally, privacy-preserving techniques like ZKP has been used to improve user's data security in decentralized environments .efficiently

B. Challenges and Limitations

The Platforms which uses Ethereum faces several challenges and limitations. One significant challenge is scalability. Its current design struggles to handle high transaction volumes and processing speeds required for efficient file sharing applications. As a result, network congestion and increased transaction fees often affect the seamless sharing of files on the platform. Moreover, Its proof-of-work consensus mechanism contributes to high energy consumption, which raises concerns about sustainability and environmental impact. The centralized file sharing platforms also have their limitations. One limitation is data privacy and security. Centralized platforms store user data on centralized servers, making them vulnerable to hacking attacks and unauthorized access. Users have limited control over their data, relying heavily on the platform provider to safeguard their information.

IV. PROPOSED ARCHITECTURE

As shown in Fig.1, The architecture of the application is designed to provide a secure and transparent method for sharing files between parties. At the core of the architecture is the Solana Blockchain, which serves as the underlying network for storing the validation logic and maintaining a tamper-proof ledger of file-sharing transactions. Also, the Smart contracts deployed on the Solana Blockchain handle the validation of file-sharing requests, ensuring that only authorized users can upload and access files. These smart contracts enable access control rules and update the blockchain ledger with records of each transaction, providing an immutable and transparent history of file sharing activities. The frontend of the architecture is responsible for providing a user-friendly interface for the user to upload files. Built with React, the frontend allows users to initiate file-sharing requests and interact with the system seamlessly. When the user uploads a file through the frontend, a request is routed to the contract algorithm on the Solana network, including a hash of the file as a unique identifier. This hash serves as an identity and reference point for retrieving the file from IPFS, the decentralized storage network used in the architecture. IPFS, is leveraged for the storage of the uploaded data. Decentralized storage space refers to the

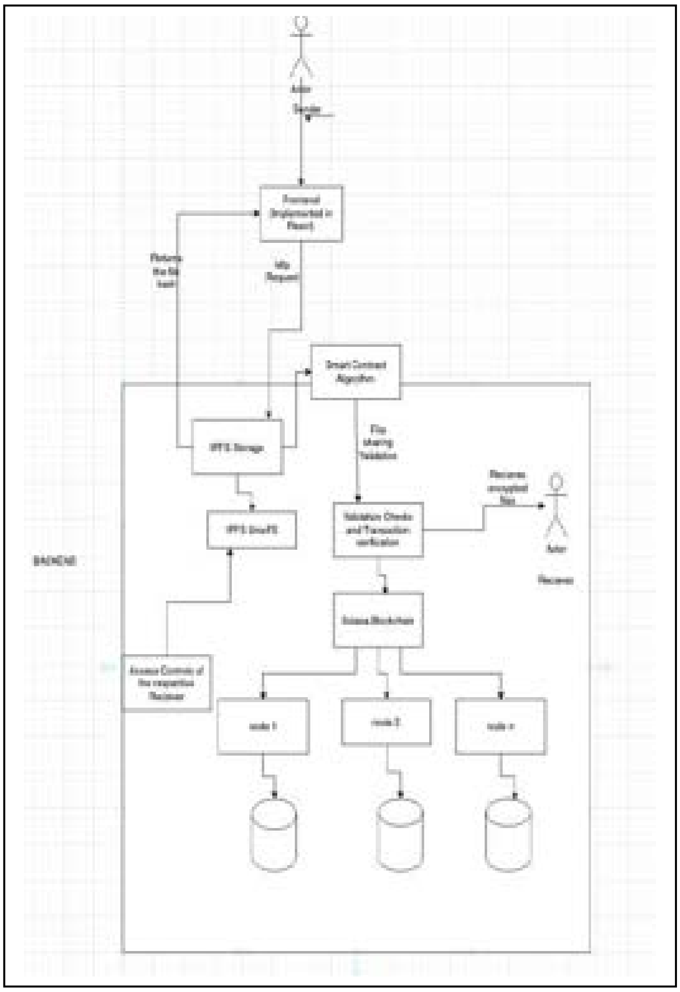


Fig 1. Proposed Architecture

storage of data across a network of nodes (computers) that collectively maintain and manage the data without relying on a central authority. Instead of using a traditional centralized server, data is stored in a decentralized fashion across multiple nodes.

When the smart contract validates the file-sharing request, the file is encrypted [10] and stored in IPFS, ensuring data security and integrity. IPFS uses a peer-to-peer network for storing and sharing data, eliminating the need for a central server, providing fault tolerance and less failure Additionally, user access rules are checked by the Solana contract which grants the Receiver access to the file based on predefined permissions. Overall, the architecture combines the immutability of the Solana Blockchain with the decentralized storage capabilities of IPFS to create a robust and transparent system for decentralized file sharing.

Solana's architecture is designed to process thousands of transactions per second (TPS). By utilizing a combination of PoH, a high-performance consensus mechanism (Proof of History), and parallel transaction processing, Solana achieves fast confirmation times. This rapid throughput significantly reduces transaction latency and improves the reliability of transaction finality. Solana maintains a decentralized network of validators (or nodes) that participate in the consensus process.

This distributed architecture enhances security and resilience against attacks. Validators work together to validate transactions and secure the network, ensuring the reliability of the blockchain. Solana leverages a technique called Transaction Processing Units (TPUs) to enable parallel transaction processing. The network is divided into multiple nodes, each responsible for processing a subset of transactions concurrently. This parallelism significantly increases the overall throughput of the network.

In an IPFS-based system, file sharing operates through a decentralized network where files are stored and accessed using a content-addressable method. When a user wants to share a file, the file is segmented into smaller chunks, and each chunk is assigned a unique cryptographic hash based on its content. These chunks are then distributed across multiple nodes (or peers) in the IPFS network. When another user requests the file, their IPFS client sends out requests to retrieve the specific chunks using their corresponding hashes. The IPFS network locates and retrieves these chunks from the nearest or most efficient nodes, taking advantage of the distributed nature of the system. Once all the chunks are retrieved, the file is reconstructed on the user's device. The metadata is uploaded to server and broken into number of fragments [9]. This method ensures efficient and resilient file sharing, as files are not dependent on a single server or location but are instead distributed across the network, leveraging the resources of participating nodes for storage and retrieval. This helps in securing Personal data from attacks [13].

V. MODULES

A. *Secure File Upload Module*

Users can upload the files and the file is encrypted into a hash by the IPFS storage and the files can only be decrypted by the user.

B. *Secure File Sharing Module*

Users can share the files with other user with the help of user name and this transaction is handled by smart contract.

C. *Access Control Module*

Sender can control who have the access to the file and allow only required people to access the file.

D. *Content ID Generation Module*

A content identifier is generated for all the user uploaded data which helps in access the data. The actual address is not shared instead a representational address is formed using the data information.

E. *User Registration Module*

Users can create their profile and a public key is also generated for the user for secure file sharing.

F. *Secure File Storage Module*

Users can store their files on the decentralised network, these files are stored securely using hashes and only the user with access can decrypt it.

G. *User Profile Module*

Users are able to see their unique public key and user ID. Also, Users can see uploaded and shared files on the dashboard

VI. IMPLEMENTATION

The proposed application combines front-end developed using React, and smart contract developed in Solidity for IPFS storage on the Solana network.

A. *Frontend*

The frontend implementation of the application is done using React, to provide a seamless and secure user experience. Using React, project have an intuitive user interface that allows user to upload files and share them with anyone. Axios client is used to communicate with the smart contract and wallet connection. Through a well-structured redux store, the application's state is managed efficiently. React hooks are used for state management in a single component.

Using react helps us in rapid application development, thus reducing the development time. React provides large set of libraries of pre-built modules which can be easily integrated in the application.

B. *Algorithm for integration with IPFS and Solana*

With the help of IPFS for decentralized storage, the application can hash files and store them across a distributed network, ensuring redundancy and resilience against data loss. Each file's unique hash acts as its identifier, facilitating efficient retrieval and verification. Also, integrating Solana Blockchain enables secure transaction processing and consensus mechanisms. Smart contracts are utilized to govern file verification and sharing protocols, ensuring transparency and immutability. With Solana's high throughput and low-latency capabilities, transactions are swiftly validated, enhancing the overall user experience.

As shown in Fig 2, the algorithm is designed to facilitate decentralized file sharing [11] on the Solana network with the help of managing URL storage and access control among users. It uses mappings and data structures to handle user ownership, access permissions, and users access data. Key functions are add, allow, disallow, display, and share-access. The add function enables users to append URLs to their personal storage, while allow and disallow functions manage access permissions by updating ownership mappings and access lists. The display function retrieves URLs of other users if access is granted, promoting controlled sharing, while share-access allows users to view their own access permissions. Together, these functions provide a framework for secure and transparent file sharing while maintaining user privacy and control over shared data.

Fig 2. Algorithm

C. Backend

The backend of the application is built on a Solidity and JavaScript. This provides data security, scalability, and seamless interactions. In this architecture, Solidity, a smart contract programming language, plays a pivotal role in managing access and by leveraging Solana Blockchain, the backend orchestrates the verification process, ensuring data integrity through its immutable ledger and enabling transparent and tamper-proof record-keeping. Concurrently, IPFS serves as the decentralized storage layer, facilitating efficient and secure file sharing by distributing files across a network of nodes, enhancing redundancy and accessibility.

VII. RESULTS

In this section, the results are shown for each implemented module of the proposed application.

A. User Registration Interface

Fig. 3 depicts the implemented interface, which prioritizes simplicity and clarity so the users scan seamlessly go through the registration process. The first field in the registration interface allows the entry of the username, an identification detail that enables easy and reliable identification within the platform. Users are allowed to select a unique and different username that represent their identity and aligns with any community guidelines or naming conventions established by the platform. Similarly, the password field prompts users to create a secure and memorable password for their account, enhancing the overall security of their user credentials. Users are allowed to create a new password that is made up of a strong combination of alphanumeric characters to minimize the risk of unauthorized access or account compromise.

Upon successfully submitting their registration details, users receive immediate notification confirming their registration and providing them with a unique hashed ID and token. The hashed ID serves as the user's unique identifier within the platform, allowing access to their account and enabling interaction with other users and platform features. Also, the generated token serves as the secure authentication token, granting users access to their account and verifying their identity during subsequent login attempts, further increasing the platform's security measures.

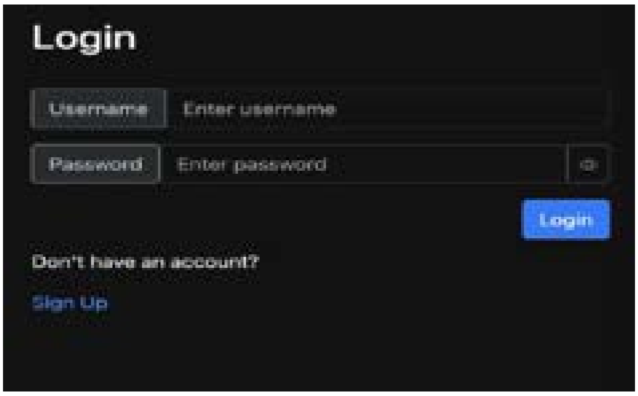


Fig 3. User Registration Interface

B. Content ID Interface

The Content ID (CID) generation as shown in Fig. 4 is generated in IPFS and it involves the creation of unique identifiers for digital content stored on their respective platforms. These CIDs serve as cryptographic hashes that uniquely identify each piece of content, allowing efficient retrieval, verification, and sharing of data within the decentralized network. In this application, the CIDs are generated with the help of smart contract functions. These functions also validate the generated CID and use it as decryption key for the file.

IPFS, or the Inter Planetary File System, uses a content-addressed system for data storage and retrieval. When a sender transfer data to IPFS, it generates a unique CID for that file based on its content. This process involves calculating the cryptographic hash of the file's content using a hashing algorithm. The resulting hash serves as the CID for the file, providing a secure and immutable identifier that uniquely identifies the file within the IPFS network.

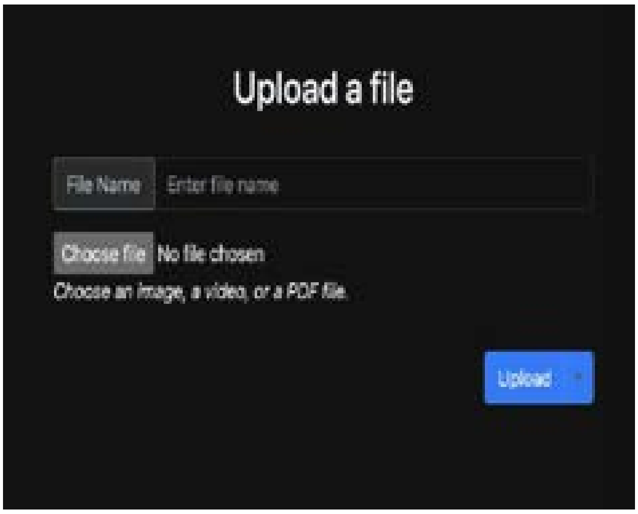


Fig 4. Content ID Generation Interface

C. *Secure File Sharing Interface*

As shown in Fig. 5 , users interact by selecting the file they wish to share from their device. The user interface includes a button for uploading files, allowing users to browse their local file system and select the desired file for sharing. Once the file is selected, the application provides feedback to confirm the successful upload, by displaying the file name and a thumbnail preview of the file. Next, users specify the recipients with whom they wish to share the file.

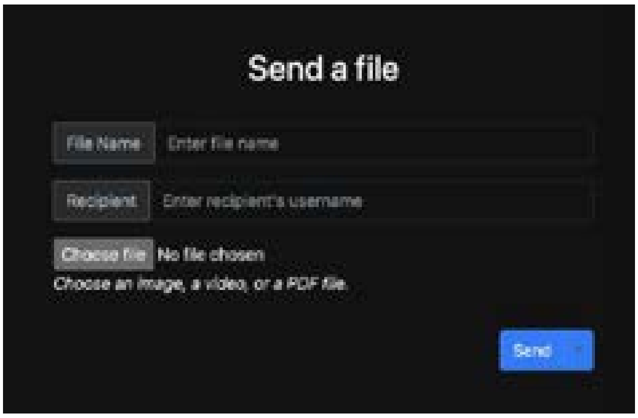


Fig 5. Secure File Sharing Interface

The interface communicates with the smart contract deployed on the Solana blockchain to execute the file sharing transaction. The smart contract verifies the authenticity of the transaction and implements any access control rules specified by the user, ensuring that only authorized recipients can access the shared file. In addition to facilitating file sharing transactions, the interface also include features for encrypting and decrypting shared files to enhance security and privacy. When a sender transfer data for sharing, the UI prompts the user to encrypt the file using their private key or a shared encryption key. The encrypted file is then securely transmitted to the smart contract for storage on the Solana blockchain.

D. *Secure File Storage Interface*

As shown in Fig . 5, the process of secure file storage starts with the user uploading the files on the profile which are then encrypted. This encryption ensures that files remain protected and inaccessible to unauthorized entities. The decentralized storage network IPFS, then distributes these encrypted and hashed files across a network of nodes, eliminating single points of failure.

Fig 6. Secure File Storage Interface

Smart contracts deployed on the Solana blockchain manage the storage and retrieval of file data transparently and securely, implementing access control policies to regulate file access and permissions. Authorized users can securely access files by their accessing their account. The files are shared by the sender to the username defined by the receiver in the application and then the receiver can download them locally using the necessary credentials, ensuring that only authorized individuals can access and modify the file data.

E. *User Profile Interface*

As shown in Figure 6, the user profile displayed on the dashboard exhibits critical information including the username, public key, and user ID. These elements serve as fundamental identifiers for each user within the system. The public key helps in creating a unique identity of the user on the platform facilitating secure transactions and interactions within the platform. Alongside the username, it forms a user ID ensuring the integrity and authenticity of user interactions. Additionally, the user ID serves as an additional layer of verification, enhancing the robustness of the identification process. Furthermore, the utilization of smart contracts on the Solana blockchain introduces a decentralized and tamper-resistant approach to access management.

Fig 7. User Profile Interface

Through the inherent features of blockchain, such as distributed consensus and cryptographic hashing, the system can mitigate potential security vulnerabilities and unauthorized access attempts effectively. Each user's public key serves as a cryptographic signature, ensuring that only authorized entities can interact with the system. As a result, the user profile dashboard becomes a gateway to secure and transparent user interactions, bolstered by the innovative capabilities of blockchain technology.

VIII. CONCLUSION AND FUTURE SCOPE

The proposed decentralized platform for digital file verification and sharing using Solana Blockchain and IPFS presents a solution to address issues of trust and security in file sharing. Also, this solution overcome the limitation posed by Ethereum as Solana is lightweight and allows more transactions per second. By leveraging blockchain's immutability and IPFS's decentralized storage, the platform ensures tamper-proof verification and efficient sharing. Furthermore, the scalability and speed offered by Solana Blockchain enhance the platform's capabilities to handle a vast array of transactions seamlessly. As the use case of the decentralised applications increases, this solution paves the way for a future where digital interactions are governed by transparency, security, and inclusivity, revolutionizing how we validate and share information in the digital realm. The performance developed platform is much better than existing solution because Solana's throughput can exceed 65,000 transactions per second (TPS), significantly surpassing Ethereum's TPS, which typically ranges from 15 to 30 TPS. Latency metrics also favor Solana, with block confirmation times averaging around 400 milliseconds, compared to Ethereum's block confirmation times that can range from 10 to 20 seconds. Additionally, Solana's proof-of-history (PoH) mechanism and Tower Byzantine Fault Tolerance (BFT) consensus algorithm contribute to its speed advantage by enabling parallel transaction processing and reducing the time needed to achieve network consensus. These quantitative metrics demonstrate platform' capability to handle a higher volume of transactions with lower latency compared to Ethereum, making it a compelling choice for applications requiring high-speed and scalable blockchain solutions. However, for the future scope, the platform can be integrated with IOT [6] and home automation devices to make the data transmission more secure. The Application can also be integrated with fintech solutions and EHR applications [12] to improve the adaptation and scale of the application

REFERENCES

[1] Z. Ning, L. Li, W. Liang, Y. Zhao, Q. Fu, and H. Chen, "A novel exploration for blockchain in distributed file storage," in Blockchain and Trustworthy Systems: First International Conference, BlockSys 2019, Guangzhou, China, December 7–8, 2019, Proceedings, 2020, pp. 740-746, Springer Singapore.

[2] B. N. Z. Benisi, M. Aminian, and B. Javadi, "Blockchain-based decentralized storage networks: A survey," Journal of Network and Computer Applications, vol. 162, p. 102656, Jul. 15, 2020.

[3] A. Ismail, M. Toohey, Y. C. Lee, Z. Dong, and A. Y. Zomaya, "Cost and performance analysis on decentralized file systems for blockchain-based applications: State-of-the-art report," in 2022 IEEE International Conference on Blockchain (Blockchain), Aug. 22, 2022, pp. 230-237, IEEE.

[4] H. R. Hasan and K. Salah, "Proof of delivery of digital assets using blockchain and smart contracts," IEEE Access, vol. 6, pp. 65439-65448, Oct. 21, 2018.

[5] I. Choi et al., "Blockchain-based secure file sharing and access control system," in International Conference on Big Data and Smart Computing, 2019.

[6] M. Ferrag, M. Derdour, M. Mukherjee, A. Derhab, L. Maglaras and H. Janicke, "Blockchain technologies for the internet of things: Research issues and challenges", IEEE Internet of Things Journal, vol. 6, no. 2, pp. 2188-2204, 2019.

[7] F. Casino, E. Politou, E. Alepis and C. Patsakis, "Immutability and decentralized storage: An analysis of emerging threats", IEEE Access, vol. 8, pp. 4737-4744, 2020.

[8] K. Bhosale, K. Akbarabbas, J. Deepak and A. Sankhe, "Blockchain based secure data storage", International Research Journal of Engineering and Technology (IRJET), vol. 6, no. 3, pp. 5058-5061, 2019.

[9] M. Naz, F. Al-zahrani, R. Khalid, N. Javaid, A. Qamar, M. Afzal, et al., "A secure data sharing platform using blockchain and interplanetary file system", Sustainability, vol. 11, 2019.

[10] I. Zhou, I. Makhdoom, M. Abolhasan, J. Lipman and N. Shariati, "A blockchain-based file-sharing system for academic paper review", 2019 13th International Conference on Signal Processing and Communication Systems (ICSPCS), 2019.

[11] J. Grabis, V. Stankovski and R. Zarins, "Blockchain enabled distributed storage and sharing of personal data assets", 2020 IEEE 36th International Conference on Data Engineering Workshops (ICDEW), 2020.

[12] H. Jin, Y. Luo, P. Li, J.Mathew. "A Review of Secure and Privacy Preserving Medical Data Sharing.", IEEE Access, Volume 7, 2019.

[13] G. Zyskind, O. Nathan, and A. S. Pentland. "Decentralizing privacy: Using Blockchain to protect personal data." in Proc. IEEE Secur. Privacy Workshops (SPW), pp. 180184, May 2015.

Hematological Image Analysis: Enhancing Blood Cell Segmentation through Advanced Image Processing Techniques

Mrs.B.Sandhiya
Information Technology
Easwari Engineering College
Chennai, India
sandhiya.b@eec.srmrmp.edu.in

Shahira Banu N
Information Technology
Easwari Engineering College
Chennai, India
nshahirabanu07@gmail.com

Vidya Annapoorna C
Information Technology
Easwari Engineering College
Chennai, India
cvidya2001@gmail.com

***Abstract*— This research study intends to solve the severe problem of poikilocytosis, a disorder that causes red blood cells (RBCs) to take on aberrant forms. This system analyzes blood smear images captured with a compound microscope by using sophisticated image processing algorithms. The method's main component is the application of the Faster R-CNN algorithm. It not only recognizes different RBC forms but also classifies them, providing information about possible related illnesses. The method significantly increases the dataset over existing methods while utilizing Faster R-CNN's better efficiency. The erythrocyte and BCCD databases data combines to create this expansion, which includes a wide range of RBC forms. The speed and precision of RBC shape detection is improved for poikilocytosis diagnosis by utilising Faster R-CNN. The larger dataset provides a deeper investigation of the wide variety of RBC forms, increasing the likelihood of finding connections to different blood cell-related illnesses. The main objective of this research study is to present a thorough report that improves patient outcomes, healthcare practices by advancing the diagnosis and knowledge of blood cell-related illnesses. This initiative paves way for better patient care and diagnostic accuracy by supporting continuing efforts to improve healthcare and medical diagnostics.**

Keywords—Blood cell, Poikilocytosis, Faster-rcnn, Image processing

I. INTRODUCTION

Poikilocytosis, the condition in which Red Blood Cells (RBCs) have non-normal forms, is still a crucial problem in the fields of haematology and clinical diagnostics. As possible markers for diverse underlying medical disorders, the unique and varied forms of RBCs linked to poikilocytosis have emerged. It emphasizes the urgent need for an accurate and timely diagnosis to enable successful treatment and improved patient outcomes. The study aims to use state-of-the-art image processing techniques to tackle this desperate problem. The novel method for identifying and classifying RBC shapes is demonstrated through the use of high-resolution compound microscope images of blood smears. Using the state-of-the-art Faster R-CNN algorithm forms the basis of this technique. This advanced method classifies RBCs in addition to identifying their distinctive forms. The main objective is to improve patient outcomes and streamline medical procedures by developing a comprehensive approach to identify and manage poikilocytosis. The target is to improve the accuracy and speed of RBC form recognition by utilizing state-of-the-art image processing techniques. It will help to ensure early diagnosis and effective management. Furthermore, the dataset have been expanded to include an expansive range of RBC types as part of objective. Through the integration of data from the erythrocyte and BCCD databases, a vast reservoir suitable for thorough examination is offered. This enhancement of the dataset allows us to investigate possible relationships between various RBC forms and blood cell-related diseases, improving the understanding of these conditions

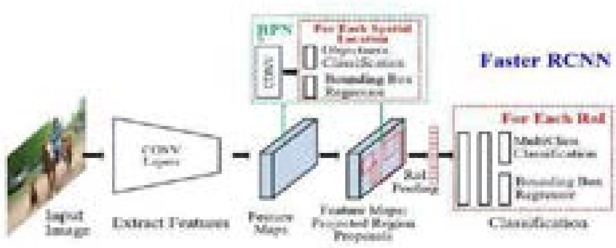


Figure 1. Faster RCNN Architecture

II. LITERATURE REVIEW

The past developments in automated blood counts and associated metrics. These include color normalization, pre-processing techniques, RBC segmentation, and the ability to distinguish between the background and foreground using a variety of segmentation methods based on the core ideas of image processing. To segment Red Blood Cells (RBCs) and deal with clumping, Berge devised morphological iterative threshold approaches. Their method created a Delaunay triangulation using real microscope pictures by using boundary curvatures. Cai et al. suggested a statistical model-based method for determining cell and cytoplasm borders in order to address segmentation issues. Once cells were separated from the cytoplasm, Bayesian classifiers became a popular option for classifying cells according to unique characteristics. Di Ruberto et al. used morphological operators and a watershed algorithm to handle clustered cells while utilizing image processing techniques to detect and identify RBCs infected with malaria. However, this approach frequently produced noisy RBC contours. The complicated properties and similar colors of blood cells made it difficult to separate RBCs from the cytoplasm.

III. METHODOLOGY

The faster R-CNN algorithm, which has been specially designed for better red blood cell (RBC) shape recognition and classification, is the fundamental component of this technique. The proposed system architecture is shown in Figure 1. The methodical process reveals how adaptable the algorithm is at effectively managing a large dataset. In order to capture characteristics indicative of various rbc types, key features are first extracted from pre-processed images. Annotations indicating the rbc forms in each of the pre-processed blood smear images are then carefully created to create a labelled dataset. To enable reliable model generalisation to unobserved data, it is critical to guarantee the representativeness and variety of the dataset. The next crucial step is to choose the right recognition model, which depends on the processing capacity and the difficulty of the task at hand. The selected model should be able to recognise complex patterns and differences in rbc morphology. After the selection of the model, the training phase begins, during which the model is trained to correlate retrieved features with corresponding rbc forms using iterative optimisation techniques such gradient descent. The model is evaluated using independent test data after training to determine how effective it is in real-world scenarios.

A. Blood Sample Acquisition

Blood Cell Count and Detection, or BCCD, is the dataset from which the blood smear images are acquired. For a thorough depiction of the analysis, the dataset offers a wide range of annotated images that include different red blood cell (RBC) morphologies linked to poikilocytosis. To ensure data integrity and relevance to the research's scope, careful selection of photographs is part of the acquisition procedure.

B. Image Processing

The procedures used to clean and preprocess the obtained blood smear images are done in the image processing section. First, a camera-equipped compound microscope is used to take high-quality pictures. To improve their quality, these photos go through preprocessing, which includes things like contrast and noise reduction. The Faster R-CNN algorithm, which was trained on a varied dataset assembled from the erythrocyte and BCCD databases, is then used to detect objects. This algorithm efficiently recognizes and classifies different types of red blood cells (RBCs), offering important information about possible related diseases. A more thorough study is made possible by expanding the dataset, which raises the possibility of finding links between abnormalities of the blood cells and RBC morphologies. After the data have been analyzed, a thorough report is produced that highlights the conclusions and highlights how they will affect patient outcomes and healthcare procedures. The continuing enhancement of patient care and diagnostic accuracy is facilitated by the cooperation of healthcare professionals and the constant improvement of algorithms. In the end, our image processing method improves patient outcomes and improves healthcare practices by contributing to the knowledge of blood cell-related disorders and helping to diagnose poikilocytosis. Steps are taken to improve image quality, fix artefacts, and prepare the dataset for further analysis.

- *Noise Removal*

  A variety of noise reduction techniques are used to remove unwanted artifacts and produce a clear depiction of blood smear photographs. These methods include deep neural networks and other machine learning techniques, together with sophisticated image processing algorithms including filtering and morphological operations. Several techniques are applied to reduce noise in bloodstain photos in order to get rid of unwanted artifacts and enhance clarity. These methods cover a wide range of complex approaches, such as machine learning, deep neural networks, and sophisticated image processing techniques including morphological operations and filtering. By combining these several techniques, the goal is to improve the way blood smear pictures are seen, which will enable more accurate analysis and detection of diseases such as poikilocytosis. This all-encompassing strategy demonstrates a commitment to use cutting-edge technologies to improve medical imaging and hematological diagnosis accuracy

- Median Filtering

  Image clarity is improved as noise is decreased by using median filtering. In order to eliminate undesired artefacts and create better images, noise reduction techniques like deep neural networks and machine learning algorithms are first used. Next, by replacing pixel values with median values from their surrounding neighbourhoods, edge preservation is ensured and noise is reduced through the use of median filtering. To compensate for colour differences, gray-world colour correction is applied. This allows for realistic portrayal of the various rbc types associated with poikilocytosis and consistent colour balance across photos. By fine-tuning the overall brightness and contrast, one may better see minute features that are essential for identifying and classifying rbc forms. Morphological identification methods such as erosion are used to inspect rbc structures and extract relevant characteristics that are necessary for next stages of analysis. Unsharp masking is the last technique used to improve photographs. The blood smear images are thoroughly cleaned through these pre-processing stages, improving their eligibility for precise analysis and diagnosis of blood cell-related disorders..

- Gray-World Color Correction

  Gray-World color correction is used to compensate for color variations in the images. This method ensures a more consistent color balance, allowing for accurate representation of different RBC types associated with poikilocytosis. To bring image colors closer to a neutral gray average, the gray-world color correcting approach is used. It first determines the average color of the entire image, and then it averages the red, green, and blue components to determine the gray world illumination. The color values of each pixel are then modified in accordance with the ratio of the lighting in the gray world to the initial average color. This procedure guarantees that the color representation is uniform and constant across the picture. Including this correction process prior to applying the Faster R-CNN algorithm to compound microscope blood smear image analysis will improve the accuracy of recognizing and categorizing various red blood cell types (RBCs). Gray-world color correction guarantees consistent color representation, which makes it easier to detect RBC shapes with greater accuracy and reliability. This enhances the diagnosis of disorders like poikilocytosis and other related conditions, ultimately improving patient care and hematology diagnostic procedures.

- Gamma Correction

The overall brightness and contrast of the images are finely tuned through the application of gamma correction. This process involves adjusting the gamma value to enhance the visibility of minute details, a critical factor in recognizing and categorizing red blood cell (RBC) forms. A precise adjustment of the gamma value can improve the visibility of minute features that are essential for a diagnosis by fine-tuning the overall brightness and contrast of the images. When performing gamma correction in this situation, it is crucial to carefully select the gamma value in order to avoid over-amplification that can mask significant features while still emphasising tiny changes in RBC morphology. To make sure that the pictures appropriately depict the fundamental properties of the blood cells, this adjustment is essential. When gamma correction is applied, the gamma correction formula is applied to each pixel in the image. By optimising the intensity levels throughout the picture, this method enhances the visibility of RBC structures and forms. To sum up, gamma correction is essential for improving blood smear picture quality, which in turn helps with better diagnostic results and the creation of more potent medical treatments for disorders involving blood cells.

- *Morphological Identification*

RBC structures and forms are examined using morphological identification techniques. Morphological processes like dilatation and erosion help to extract pertinent traits that are important for the later phases in analysis.

- *Unsharp Masking*

Images with details are enhanced by using unsharp masking, which highlights high-frequency components. By doing this, unsharp masking successfully sharpens the visual representation, which enhances the clarity and makes subtle details in the photos easier to see. This is especially useful when analyzing blood cells in detail for hematology.

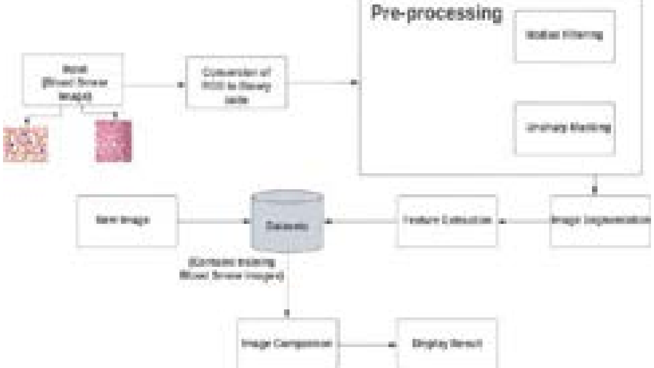


Figure 2. The proposed system architecture

C. *Image Segementation*

Several methods are used in the Image Segmentation stage to improve the analysis's granularity and separate distinct elements from the blood smear images. The aim is to enable a more thorough analysis of morphological anomalies in red blood cells (RBCs) linked to poikilocytosis.

First, the image is binary represented by applying thresholding techniques, which distinguish between pixels in the foreground and background. This first stage creates the framework for identifying and separating particular interest areas in the blood smear. Then, boundary and contour detection techniques are applied, which help to separate different cell structures and enable the recognition of asymmetric shapes.

Segmenting related regions within the image is possible by application of region-growing techniques, which group pixels with comparable qualities. The accuracy of this analysis is improved by this method, which is especially useful in identifying specific cells and structures in the blood scrape.

Furthermore, overlapping cells or structures in the image are addressed by using Watershed Segmentation. Watershed algorithms detect ridges and basins by treating pixel intensities as a topographic surface. This robust segmentation method successfully divides related components.

To further aid in a more thorough examination of the atypical forms connected to poikilocytosis, contour-based segmentation techniques are also used to outline and divide individual RBCs. This system can recognize and categorize different RBC types more accurately by extracting contours, which allows for a finer level of segmentation.

The segmented findings undergoes validation and post-processing procedures. In order to improve the segmented regions and guarantee the precision and dependability of the segmentation outcomes, this also entails noise reduction and validation checks. This stage lays the foundation for classifying the different cell shapes and detecting the anomalies suggestive of poikilocytosis, as well as for preparing the data for later phases of analysis.

D. *Feature Extraction*

To thoroughly analyze blood smear images for poikilocytosis identification, the Faster R-CNN architecture and the Feature Extraction phase collaborate well.

First, the Feature Extraction phase focuses on extracting pertinent information from segmented blood smear images. This stage uses several methods, including the extraction of morphological features, texture analysis, computation of statistical features, and shape descriptors. The combination of these methods produces a diverse range of characteristics that describe the deviations in the forms of red blood cells (RBCs) linked to poikilocytosis.

Next, the architecture of the Faster R-CNN is emphasized. The feature map, which is extracted from the Region Proposal Network (RPN) layer, is an essential element. The CNN's convolutional layers produce the feature map after it has been previously trained, and it captures the spatial properties of the input image. Using anchor boxes, the RPN provides region suggestions based on bounding box coordinates and objectness scores. This stage effectively identifies candidate locations that are likely to contain things of interest, including atypical red blood cells.

The feature map derived from the RPN is seamlessly included in the subsequent Fast R-CNN stage. This integration makes it feasible to precisely locate and classify things inside the identified zones, facilitating an easier detection of aberrant RBC morphologies associated with poikilocytosis.

On the whole, the rich feature set produced by the Feature Extraction stage enhances the visual information from segmented images. The Faster R-CNN architecture then uses the feature map which effectively identifies and categorizes items. All of these stages work together to improve the precision of identifying poikilocytosis in blood smear images.

E. Image Comparison

Region of Interest (ROI) Pooling is incorporated into the procedure to improve the comparison, and the extracted features from the Feature Extraction step are aligned. The ROI Pooling entails picking and combining characteristics from particular regions of interest in the processed image to provide a more targeted and insightful comparison with the dataset.

After alignment and ROI pooling, feature matching techniques are used to measure how similar related regions from various images are to one another. This stage makes sure that the RBC forms and anomalies connected to poikilocytosis are thoroughly compared.

Next, the dissimilarity between feature vectors is measured using distance metrics such as cosine similarity or Euclidean distance. These measures aid in the detection of trends and anomalies by providing information on the general dissimilarity of RBC shapes.

The outcomes of the Image Comparison stage help to clarify the differences in RBC shapes seen in the sample. Despite the comparison phase's main focus, the knowledge acquired from it might guide later stages of the analysis, such as classification or more focused research into particular abnormalities.

F. Classification and Prediction

During the Classification and Prediction phase, the focus moves from feature comparison to decision-making regarding the anomalies in red blood cell (RBC) shapes that have been observed. This crucial stage uses the knowledge gathered from the Image Comparison stage to categorize various morphologies, such as acanthocytes or degmacytes, and forecast possible illnesses linked to poikilocytosis.

The aligned and pooled features are used as inputs to a classification model after features are extracted and Region of Interest (ROI) Pooling is used. This model is crucial in classifying the blood smear images according to the detected abnormalities. It is frequently trained on a labelled dataset with annotations for various RBC forms and related disorders.

The Faster R-CNN model uses its object identification skills to recognize and categorize the blood smear images based on detected anomalies. The model has been trained on a labelled dataset with annotations for various RBC forms and related disorders. In doing so, particular labels—such as degmacyte or acanthocyte—are applied to the recognized RBC forms, and conjectures on possible illnesses are made.

Atlast, possible disorders are predicted using the categorized RBC forms. Given the anomalies seen in the blood smear images, the model is trained to identify patterns and relationships that help in forecasting the likelihood of particular diseases. Anaemia and other hematological diseases, for example, may be indicated by specific RBC shapes like sickle cell.

The results of this stage are integral for medical professionals to diagnose and comprehend the underlying medical issues of their patients. The system's ability to diagnose diseases is improved by the combination of classification and prediction, which offers important insights into possible illnesses linked to abnormalities in RBC forms.

To summarise, the Classification and Prediction stage expands on the knowledge obtained from Image Comparison by using a trained Classification Model to identify RBC shapes and forecast illnesses linked to poikilocytosis in blood smear pictures.

IV. EXPERIMENT AND RESULTS

A comprehensive evaluation is carried out to see how well the Faster R-CNN model performs in categorizing different shapes of red blood cells (RBCs) and predicting disorders linked to poikilocytosis. A sample image is shown in Figure 3.

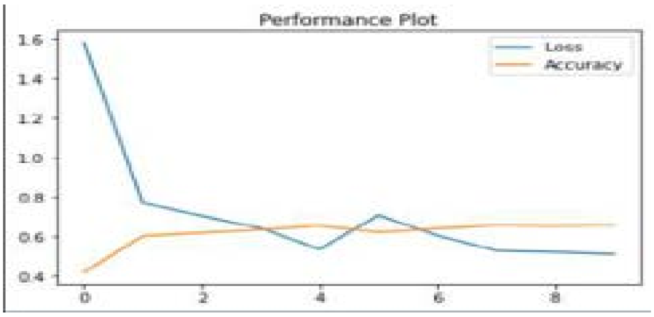
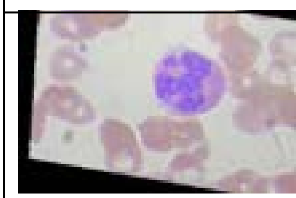
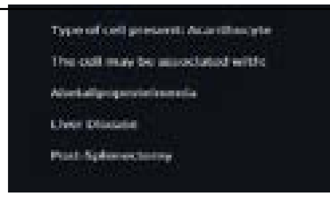


Figure 3. Performance Plot Performance Analysis of different techniques for blood cell detection

Several approaches were examined, including RCNN and CNN, ANFIS, ANN-YOLO, ResNet50, and Faster RCNN, in the Performance Analysis of various blood cell detection strategies. Table 1 shows the analysis of varying degrees of accuracy for each technique. The outcomes highlight the various blood cell detection capabilities of these methods and provide insightful information about their advantages and disadvantages.

Table 1. Result and Outcome of the system

Dataset	**Results**
	

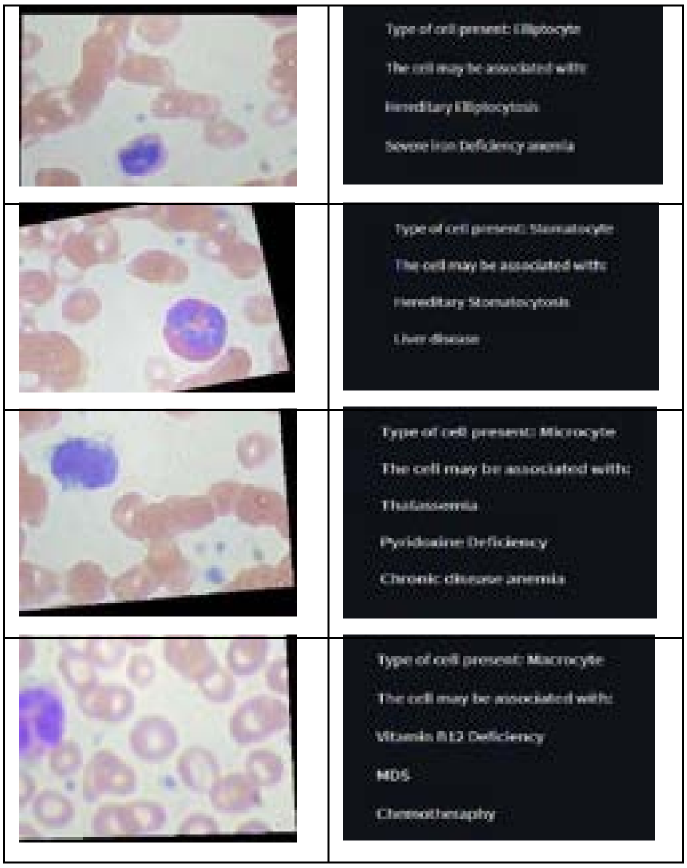


Table 2. Performance Analysis of different techniques for blood cell detection

Techniques	Total images	Accuracy
RCNN and CNN	313	72
ANFIS	21k	96.6
ANN-YOLO	364	96.09
ResNet50	3.4M	90
Faster RCNN	3k	75

ANFIS was able to achieve a 96.6% accuracy rate on a sizable dataset comprising 21,000 photos. Similarly, using 364 photos, ANN-YOLO achieved a high accuracy of 96.09%. ResNet50 demonstrated strong performance, achieving 90% accuracy on a large dataset consisting of 3.4 million photos. However, with 313 photos, RCNN and CNN demonstrated a lesser accuracy of 72%. The results emphasize the unique qualities and possible uses of different approaches while offering insightful information about the varying efficacy of these methods in blood cell detection. Among these, Faster RCNN was a standout performer, exhibiting encouraging outcomes that beat some other approaches. Based on these results, Faster RCNN may be a more powerful technique in some situations, highlighting its nuanced usefulness in blood cell detection.

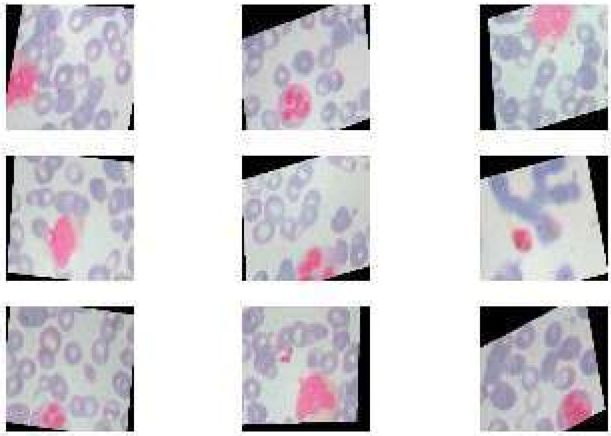


Figure 4. Blood smear images used for the system

A. Analysis of proposed system on different datasets

In the analysis of the suggested system on various datasets, the main goal was to evaluate the performance of datasets, on several different datasets, including BCCD, ErythrocytesIDB and Chula-RBC to understand the performance of the developed system. The system-attained accuracy percentages for each dataset, along with the image count, are presented in Table 3.

Table 3. Analysis of proposed system on different datasets

Dataset	Image count	Accuracy (%)
BCCD	4884	98.99
ErythrocytesIDB	626	97.14
Chula-RBC	1500	98.70
Our Dataset	3740	73.6

Four hematology datasets are evaluated in this study: Chula-RBC (1500 pictures, 98.70% accuracy), ErythrocytesIDB (626 photos, 97.14% accuracy), BCCD (4884 images, 98.99% accuracy), and Our Dataset (3740 images, 73.6% accuracy). Our diversity-focused dataset is now undergoing a comprehensive assessment. The work emphasizes how important dataset selection is to the best possible automated blood cell identification systems.

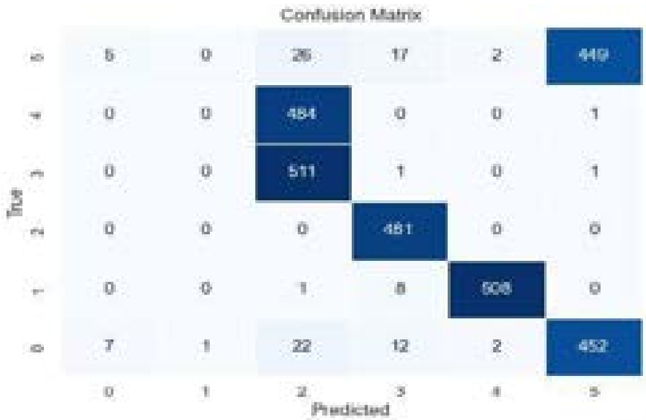


Figure 5. Confusion Matrix

B. *Performance Evaluation of target class using precision, accuracy and sensitivity*

The primary focus was on rigorously assessing the system's performance with specific attention to a designated target class. This study evaluates target class performance through precision, accuracy, and sensitivity metrics. Precision assesses correctness, accuracy gauges overall correctness, and sensitivity measures the ability to correctly identify positive instances. Table 4 encapsulates the precision, accuracy, and sensitivity metrics for this evaluation.

Table 4. Performance evaluation of target class using accuracy, precision and sensitivity

Target class	Precision (%)	Accuracy (%)	Sensitivity (%)
Microcyte	99.8	99.9	99.9
Elliptocyte	98.2	99.5	99.8
Macrocyte	98.4	99.8	99.4
Acanthocyte	99.3	99.7	99.3
Stomatocyte	97.2	99.6	99

Performance metrics for automated blood cell categorization are included in the table, with percentages for each target class for True Positives (TP), True Negatives (TN), False Positives (FP), False Negatives (FN), Precision, Accuracy, and Sensitivity. The efficiency of the categorization model is demonstrated by the high precision and accuracy that have been observed.

V. CONCLUSION

This research study offers a comprehensive solution to the challenging issue of poikilocytosis, a disorder characterized by abnormal shapes of red blood cells (RBCs). Leveraging sophisticated image processing algorithms, particularly the Faster R-CNN algorithm, the system analyses blood smear images captured with a compound microscope. By recognizing and classifying various RBC forms, the method not only enhances dataset size but also improves efficiency compared to existing methods. Through the combination of erythrocyte and BCCD databases, a diverse range of RBC forms is incorporated, facilitating a more extensive investigation into potential connections with blood cell-related diseases. The utilization of Faster R-CNN enhances both the speed and accuracy of RBC shape detection, thereby advancing the diagnosis of poikilocytosis and other related illnesses. Ultimately, this research study has improved patient outcomes and healthcare practices by enhancing diagnostic capabilities and expanding knowledge in the field of blood cell-related illnesses. This initiative represents a significant step forward in healthcare and medical diagnostics, supporting ongoing efforts to enhance patient care and diagnostic accuracy.

REFERENCES

[1] J. M. Sharif, M. F. Miswan, M. A. Ngadi, Md Sah Hj Salam, ”Red blood cell segmentation using masking and watershed algorithm: A preliminary study”, 2012 International Conference on Biomedical Engineering (ICoBE), Penang, Malaysia,27-28 February 2012, Publisher : IEEE.

[2] Siti Madihah Mazalan, Nasrul Humaimi Mahmood, Mohd Azhar Abdul Razak, “Automated Red Blood Cells Counting in Peripheral Blood Smear Image Using Circular Hough Transform”, 2013 IEEE First International Conference on Artificial Intelligence, Modelling & Simulation, Publisher : IEEE.

[3] Kareem, S. Morling, R. C. Kale I., “A Novel Method to count the red cells in thin blood films”, 2011 IEEE International Symposium of Circuits and Systems (ISCAS), Publisher : IEEE.

[4] M.Habibzadeh, A. Krzyz˙ak, T.Fevens, A.Sadr, “Counting of RBCs and WBCs in noisy normal blood smear microscopic images”, Medical Imaging 2011, Proc. of SPIE Vol. 7963, 79633-1.

[5] Razali Tomaria, Wan Nurshazwani, Wan Zakariaa, Muhammad Mahadi Abdul Jamila, Faridah Mohd Norb, Nik Farhan Nik Fuadb, ”Computer Aided System for Red Blood Cell Classification in Blood Smear Image”, International Conference on Robot PRIDE 2013-2014 - Medical and Rehabilitation Robotics and Instrumentation, ConfPRIDE 2013-2014, Publisher : Elsevier

[6] Sherif Abbas, ”Microscopic Images Dataset for Automation of RBCs Counting”, Middle East Technical University 2015, Publisher: Elsevier.

[7] Kwok, N. M., Wang D., Jia X., Chen S. Y., Fang G., Ha Q. P.,” Gray World based Color Cor- rection and Intensity Preservation for Image Enhancement”, 2011 4th International Congress on Image and Signal Processing, Publisher: IEEE

[8] Wilfried Kubinger, Markus Vincze, Minu Ayromlou, “The Role of Gamma Correction in Colour Image Processing”, 9th European Signal Processing Conference (EUSIPCO 1998), Publisher: IEEE

[9] Erik Reinhard, Michael Ashikhmin, Bruce Gooch, and Peter Shirley, University of Utah, “Colour transfer between Images”, IEEE 2001

[10] Derek Magee, Darren Treanor, Doreen Crellin, Mike Shires, Katherine Smith, Kevin Mohee, and Philip Quirke,” Colour Normalisation in Digital Histopathology Images”, Elsevier 2009.

[11] Complete Blood Count (CBC) Source: Wikipedia Human Blood and Its Components URL: https://www.hematology.org/Patients/Basics/

[12] D. Baswaraj, Dr. A. Govardhan, Dr. P. Premchand, “Active Contours and Image Segmen- tation: The Current State of the Art”, Global Journal of Computer Science and Technology Graphics & Vision Volume 12 Issue 11 Version 1.0 Year 2012

[13] Jinping Fan, Yonglin Zhang, Ruichun Wang, Shiguo Li, “A Separating Algorithm for Over- lapping Cell Images”, Department of Electronic Communication Technology, Shenzhen Institute of Information Technology, Shenzhen, China; Institute of Optoelectronic Engineering, Ji-nan University, Guangzhou, China

Blink Elimination in EEG signals using Two phase Variational Mode Decomposition

J. Sheshagiri Babu
Dept of ECE
Kakatiya Institute of Technology and Science
jsb.ece@kitsw.ac.in

B.Sai Vardhan
Dept of ECE
Kakatiya Institute of Technology and Science
bomminenisai69@gmail.com

K.Nikitha
Dept of ECE
Kakatiya Institute of Technology and Science
kalakolanikitha@gmail.com

R. Sai Nikhitha
Dept of ECE
Kakatiya Institute of Technology and Science
sainikhitharayarakula@gmail.com

Sai Dilip Suram
Dept of ECE
Kakatiya Institute of Technology and Science
suramsaidilipreddy@gmail.com

***Abstract*— The Electroencephalogram (EEG) are the data collected from human scalp to analyze functioning of the brain. These EEG signals usually contaminate with other signals like ocular artifacts (OA). There exist signal decomposition methods like wavelet decomposition, empirical mode decomposition (EMD) and Variational Mode Decomposition (VMD) to extract these artifacts. In this research work, a novel two-phase variational mode decomposition (TPVMD) is used to analyze EEG signals to extract OAs. In existing empirical mode decomposition (EMD) and multi variational empirical mode decomposition (MEMD), balancing parameter and regression is not applied. In this proposed method, Variational Mode Decomposition (VMD) is used to eliminate the OAs in two phases and in each phase, regression is applied. In this paper the EEG signal data is collected from semi-simulated Mendeley database. Blink artifacts are eliminated from corrupted EEG signal and the signal to artefact ration (SAR) for different balancing factor alpha(α). The simulation results show that the proposed method gives improved SNR and reduced difference in power spectral density (DPSD).**

Keywords—Electroencephalogram (EEG), Ocular Artifacts (OA), Variational Mode Decomposition (VMD), Regression, Signal to artifact Ratio (SAR), Difference In Power Spectral Density (DPSD).

I. Introduction

The scalp electrodes were used to record the electrogram, which came from the brain allows us to analyse the brain working and its characteristics. Significant advantages of this technique give rise to various techniques like noise elimination, feature extraction etc.. The most significant area of interest is noise elimination as it provides the data fidelity [1]. Data fidelity is a key aspect as EEG data is used in sensitive areas like diagnosis. The EEG signals are used in tracking epilepsy, dementia, head injuries, brain tumours, and sleep apnoea etc[2].

The process of recording EEG involves in several steps, firstly by cleaning the scalp and the several small sensors are spread across scalp which are termed as electrodes with the help of specific glue. These are all attached by wires to a EEG machine which records the signal electrogram of brain. Around 20 to 40 minutes of EEG recording with some preparation time concludes the session. The Results are given to trained doctors to analyse interpret the recorded electrogram and send the results. It is utilized not only for diagnostics but also for brain-computer interface devices, which give patients almost instantaneous control over external actuators. BCI applications based on EEG allow users to operate a computer or other device with just their thoughts, negating the need for software or other computer-related tools [3].

Table.1.EEG SIGNAL CHARACTERISTICS

EEG Type	**Characteristics**	
	Amplitude(μV)	***Frequency Range(Hz)***
δ	20-100	0.5-3.5
θ	10	3.5-7.5
α	2-100	7.5-12
β	5-10	12-30

The EEG recordings are risk free and no side effects are observed. The bandwidth of EEG spectrum ranges from 1 to 30 Hz with a amplitude ranging between 20 to 100 μV. Based on the frequency there is further sub classification into 4 groups: α (8-13 Hz), β (13-30 Hz), δ(0.5-4 Hz) and θ (4-7 Hz) as shown in Table 1. Each of these sub bands is associated with certain state of brain [4]. For example, Alpha waves are associated with the state of being relaxed wakefulness with high prominence of happening in parietal and occipital areas.

The 10-20 international system is a standard method for locating electrodes on the scalp to record EEG signals consistently across different recordings[5]. Each electrode site is labelled with a combination of letters indicating the general brain region it covers: To indicate pre-frontal, frontal, temporal, parietal, occipital, and central areas, use the letters Fp, F, T, O, and C. This systematic approach allows for accurate analysis of brain activity and function.

The main problem with EEG is noise which is picked by the other systems in the body such as circulatory system, sensory nervous system etc. The noise which is picked form the eye is termed as ocular artifacts [6][7]. As eyes are closer to the brain and hence the corruption by the eyes is more intense than other systems. These ocular artifacts are further classified into several bands based on frequency, morphology, duration, amplitude etc…The Table .II below shows different ocular artifacts.

Table.2. OA CHARACTERISTICS

OA Type	**Characteristics**		
	Amplitude(μV)	***Duration(ms)***	***Frequency Range(Hz)***
Blink	⩽800mvolt	200-400	<7.5
LEM	⩽800mvolt	200-4	<7.5
Flutter	⩽800mvolt		8-12

Mostly the EEG recorded from the frontal lobe is affected by these artifacts because of their neighbourhood with the eyes. The electrodes which are placed near to eye are affected by the eye blinks and eye movements[8]. As these electrogram from eye movements and eye blinks are inherent and cannot be differentiated manually or at the time of EEG, we need to remove them by using various noise elimination algorithms at the pre-processing stage[9].

There are several proposed algorithms on artifact elimination regression of EEG with a good reference EOG, Blind Source Separation methods such as ICA, and PCA. But because of the two-way corruption in the channels there is not much interest in multi-channel-based algorithms [10][11]. A different single-channel approach, the EMD method, was also put out; however, the algorithm did not follow any mathematical logic, and instead used techniques like cubic spline interpolation etc. make the job complex to deal with EMD based algorithms[12].

The method proposed by Chinmayee Dora et.al talks about the disadvantages of the multi-channel-based methods and EMD based methods and also provides a new method based on VMD (variational mode decomposition) to overcome those disadvantages and also successful in their approach[13][14]. The k value which decides the number of mode functions. Also there is no proper stand on their performance metric used which is DPSD(difference in power spectral density) with a controversial statement saying that more is DPSD more is the efficiency but the old papers are against the statement. Hence the metrics used in this paper are the one which are used from [15].

VMD based methods surpasses multi-channel-base methods, EMD based methods and Ensembled EMD (EEMD) regarding noise resilience, tone recognition, and separation. In Section 2, the paper presents the numerical method of VMD, linear regression, and minimized mean square error (mMSE0. In Section 3, the dataset, corrupt EEG segment detection, and the suggested process are addressed. Section 4 discusses the outcomes attained emphasizing the consequence and limitations of the proposed Method. Lastly, Section 5 draws conclusions and discussions.

II. Variational Mode Decomposition With Linear Regression And mMSE

A. VMD

The individual instruments can be recognized from any music while listening, have you ever thought of how could we able to separate the instruments from the music? This is because we inherently possess the ability to decompose the music signal and able to recognize them. The inherent ability is decomposition which separated all the different frequency components and from those components we can only concentrate on the desired instrument based on the characteristics (such as frequency) of the instrument that we are well aware. In the same way we use VMD to decompose a real time signal to decompose into several components and the identify and eliminate the undesired components using the known characteristics of those components from the contaminated EEG signal.

VMD is a technique which is used to decompose a non-stationary composite signal into several components which are band limited and qualify certain properties of VMD. The qualification mode components is based on the fact that the obtained signal should be a AM-FM signal and by adding the obtained all the obtained signals based on the qualification criteria should result in the same non-stationary composite signal. The mathematical model of VMD is shown in equation 1.

$$\min_{s_k,\omega_k}\{\sum_{m=1}^{K}||\partial_t\left[\left(\delta_t+\frac{j}{\pi t}\right)*s_k(t)\right]e^{-j\omega_k t}\;||_2^2\} \quad (1)$$

For the fact that

$$\sum_{k=1}^{K}s_k(t)=S(t) \quad (2)$$

Where (t) regarding k=1, 2, 3…. K are the BLIMFs of centre frequency derived via breakdown for a particular value of k and S(t) is the original signal.

To get around equation (1)'s constraint, a penalty term (α) and Lagrangian multiplier ( ) are employedExcellent convergence for limited weightsand strict constraint enforcement using the Lagrangian multiplier are advantages of the quadratic penalty term. Thus, we obtain the augmented Lagrangian equation as follows: where is the convolution operator and dot is the Dirac distribution.

$$L(\{s_k\},\{\omega_k\},\lambda)=\alpha\sum_{m=1}^{K}||\partial_t\left[\left(\delta_t+\frac{j}{\pi t}\right)*s_k(t)\right]e^{-j\omega_k t}\;||_2^2 + ||S(t)-\sum_{k=1}^{K}s_k(t)\,||_2^2+\langle\,\lambda(t),S(t)-\sum_{k=1}^{K}s_k(t)\rangle \quad (3)$$

The balance parameter of the data fidelity constraint is represented by α. Because of the integrated Wiener filter, the VMD approach is impervious to noise and sampling in this situation. With a convergence tolerance of €, the saddle point of Equation (2). It can find utilizing the following with the Alternate Direction Method of Multipliers (ADMM) convergence criterion and is identical to the solution of Equation (1):

$$\sum_k\frac{||\hat{s}_k^{n+1}-\hat{s}_k^{n}||_2^2}{||\hat{s}_k^{n}||_2^2}<€ \quad (4)$$

VMD-obtained BLIMFs exhibit less immediate frequency changes than EMD and EEMD. In contrast to EMD and EEMD, VMD performs better in terms of voice recognition, tonedeparture, noise robustness, and signal restoration.

B. Regression

By the regression model we can represented the contaminated ocular artifact in an EEG signal.. This model has the following mathematical expression:

$$s_i(t)=s_{true_i}(t)+a s_{false_i}(t) \quad (5)$$

where is the observed polluted EEG at the electrode locations, is the accurate or unpolluted EEG, and is the taint ocular artifact with a regressioncoefficient of a. Upon employing the least squares method of evaluation, the calculated is :

$$\hat{a}=\frac{\sum(s_i-\bar{s})(s_{false_i}-\overline{s_{false}})}{\sum(s_{false_i}-\overline{s_{false}})^2} \quad (6)$$

C. mMSE

Modified Multiscale Entropy(mMSE) is an extension of Multiscale Entropy (MSE) method, this is used to analyse the complexity of the time series data, especially bio signals like EEG, ECG etc… This method is based on the concept of entropy, this measures the entropy of the subject.

$$s_e(m, r, N) = -\ln\left(\frac{C(m+1, r)}{C(m, r)}\right) \tag{7}$$

Where:

- is the sample entropy of the signal
- m is the length of the template vector (embedding dimension)
- The similarity or tolerance criterion, denoted by r
- length of the time series denoted by N
- C (m, r) amount to the number of pairs of template vectors that match within a tolerance of r

Multiscale Entropy involves the computation of the at multiple scales by coarse-graining the time series data. This is typically achieved by dividing the time series into independent segments of different lengths and computing for each segment.

The formula for MSE at scale n is:

$$MSE(m, r, n) = \frac{s_e(m, r, N_n)}{s_e(m, r, N_{n+1})} \tag{8}$$

Where:

- is the length of the time series at scale n.

Now, mMSE involves some modification in the calculation of MSE to address specific limitations and to improve performance. Modifications may include variations in parameters m, r, or the way entropy is computed at each scale.

D. Turning point count based Artifact detection

When analysing EEG (Electroencephalogram) signals, the Turning Point Count (TPC) method is a technique used to find artifacts, which are undesired interruptions in the signal brought on by a variety of factors like muscle movement, eye blinking, or environmental interference.

The TPC approach analyses the EEG data by tracking variations in amplitude over time. The technique is specifically concerned with locating sudden shifts in the signal's direction, or "turning points." These pivotal moments signify noteworthy alterations in the EEG waveform and are frequently linked to artifacts.

The TPC method counts the number of EEG signal turning points that occur during a certain time window or segment. An increased number of turning points indicates a higher degree of artifacts in the signal.

The quality and dependability of the analysis can be increased by researchers by efficiently identifying, eliminating, or correcting artifacts from the EEG data by defining suitable thresholds for the turning point count. All things considered, the Turning Point Count technique offers a simple and quantitative method for identifying artifacts in EEG signals, which helps with the precise interpretation of brain activity data gathered from EEG recordings.

III. PROPOSED METHOD

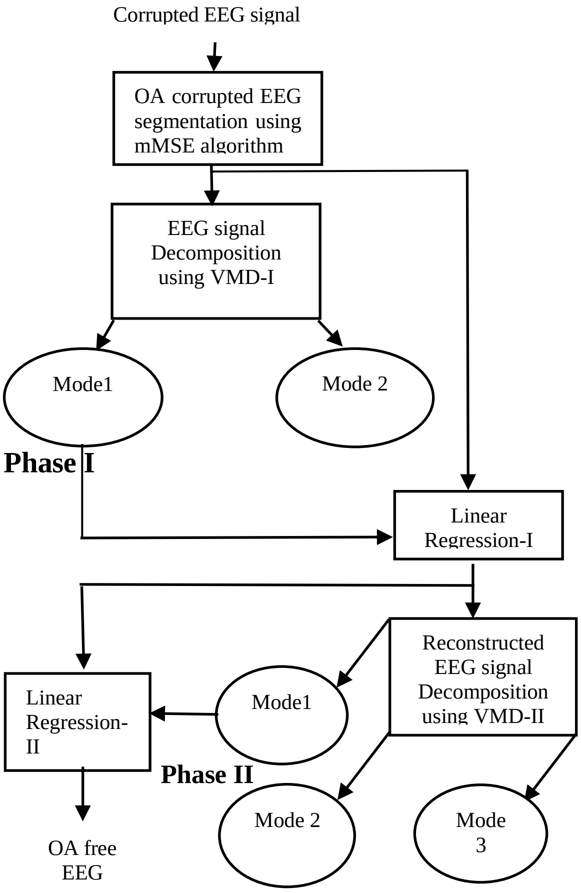


Fig. 1. *Illustrates the two phase algorithm for blink elimination.*

A. Phase – I

In the first step the EEG is applied with VMD-I with K value chosen as 2 and the balancing factor α chosen as 2000 for the eliminate of low frequency artifact components from the corrupted EEG signal ‘a’ in fig. 2. VMD-I decomposed the original signal into 2 signals ‘b’ and ‘c’ in fig. 2. out of which the signals having less than 0.5Hz is one of the mode functions.

The alpha(α) and K values are chosen specifically to attract the low frequency baseline components especially designed parameter from paper []. These values show a great potential in attracting the base line signal ‘b’ in fig. 2 form the contaminatedEEG signal.

After decomposing the corrupted EEG signal ‘a’ in fig. 2 into its mode components ‘b’ and ‘c’ in fig. 2 the low frequcy component ‘a’ in fig. 2 which is less than 0.5Hz is declined with the tainted EEG signal to get the reconstructed EEG singal ‘d’ in fig. 2 for further processing.

After decomposing the corrupted EEG signal ‘a’ in fig. 2 into its mode components ‘b’ and ‘c’ the low frequcy component ‘a’ which is less than 0.5Hz is declined with the conataminateEEG signal to get the reconstructed EEG singal ‘d’ in fig. 2 for further processing.

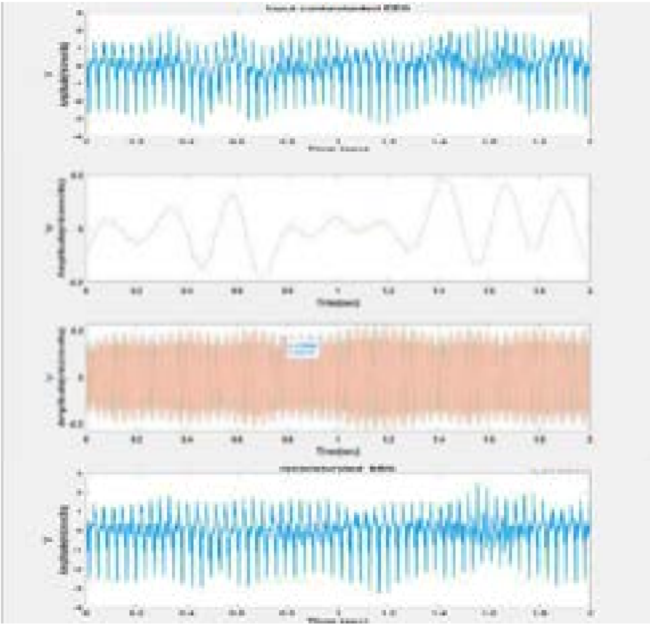


Fig. 2. *Illustrates the EEG Signal decomposition and reconstruction in Phase-I: (a) Original EEG signal of first subject in Mendeley database; (b), (c) Decomposed modes of VMD-I; (d) Reconstructed EEG signal after Linear Regression-I.*

B. Phase-II

In the second phase the EEG is treated with the α value chosen to be 1000 and K chosen to be 3 to extract the phase-I processed EEG input from the blink artifacts. After applying the VMD-Phase II with the chosen parameters form paper [] those obtained signals are then sent to turning point count algorithm which detects the corrupted segments from the three mode components. Then followed by regressing the selected mode components with the EEG signal, which will provide the ultimate pure EEG signal estimate.

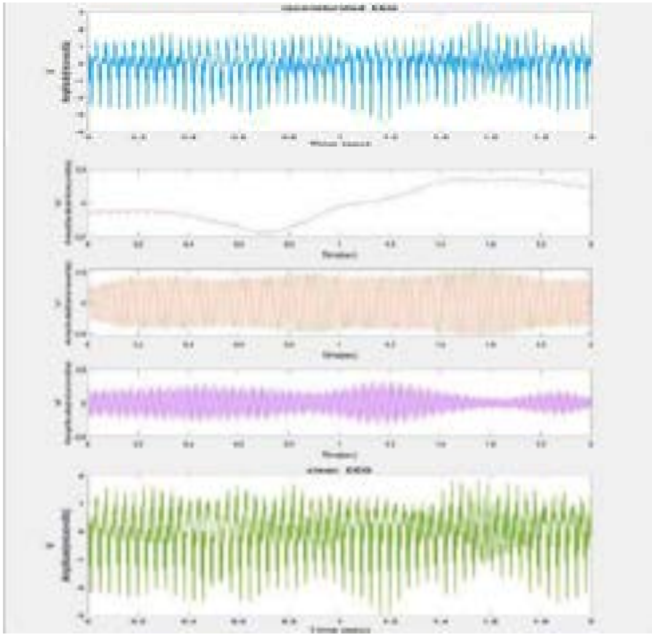


Fig. 3. *demonstrate the signal breakdown and restoration in phase-II : (a)processed EEG signal from phase-I;(b), (c), (d) are the mode functions of contaminated EEG;(e) is the clean EEG regressed from noisy mode function by turning point count.*

In phase-II the processed EEG signal ‘a’ in fig. 3 from phase-I is fed as input to the VMD-II which decomposes the signal into three mode components ‘b’, ‘c’ and ‘d’ in fig. 3. Among these three decomposed signals of the processed EEG signal one of the signal is a blink signal and it is identified using turning point count based method used in paper []. The blink component identified by the turning point method is ‘b’ in fig. 3. The processed EEG signal ‘a’ is regressed with the blink artifact component ‘b’. The resulting artifact free EEG signal after regression is ‘e’ which can be further used for analysis purpose.

The final outcome is a pure EEG signal which is further established in the results section. In this phase the artifact components are removed by varying the VMD parameters to obtain the characteristics of the artifacts in a heuristic way. The other mode components obtained are the factors of the processed EEG signal which are not our course of interest. Out the three mode function obtained the first mode function is our artifact component extracted from the processed EEG signal.

IV. Results and Discussion

The proposed two-phase VMD algorithm is validated in this section using suitable performance metrics by different EEG signals collected from many different databases articles. This section gives the idea of datasets used and performance metrics and validation of the proposed algorithm using them.

Table.3. Phase-wise comparative analysis

Phase	**Parameters**			
	K(no.of modes)	***α(balancing parameter)***	***DPSD (Difference in Power Spectral desity)***	***SAR(Singnla to artifact ratio)***
Phase-I	2	2000	0.04	7.85
Phase-II	3	100000	0.018	25.18

A. Databases

The data bases operate to verifythe proposed technique are the EEG signal collected from EEG information which include the semi-simulated Mendeley database, EEGMAT database and Polysomnographic database.

B. Performance metrics

The performance metrics used in this database are Signal to Artifact Ratio (SAR), Difference in power spectral density (DPSD) and Normalized Cross Correlation (NCC) for both frequency, power, variation based metric analysis for denoising of the given EEG signal. The performance metrics are calculated for the actual EEG signal and the estimated EEG signal are given by:

- $SAR\,(in\ dB) = 10 * \log\left(\frac{Signal\ Pow}{Noise\ Pow}\right.$ (9)

$DPSD = PSD(Original\ Signal) - PSD(Cleaned\ signal)$

where PSD is the strength spectrum density of the communication.

C. Performance evaluation

The evaluation of the algorithm phase wise based on the VMD parameters is performed initially to showcase the phase wise performance of the proposed algorithm. The phase wise performance is shown in Table. 3. The variation in the VMD parameters have yielded good results (SAR) in both the phases as we can see in Table. 3.

The SAR ratio is good enough to yield a EEG signal as pure as possible. The signal strength is greater than the noise signal, as indicated by the SAR. Additional parameter Power spectral density variation.

DPSD is defined by calculating the difference in power spectral density between the actual EEG signal and the signal processed with the blink elimination algorithm, you can quantify the reduction in power associated with blink artifacts. A greater reduction in power at frequencies and time intervals corresponding to blink events indicates more effective blink artifact removal.

A lower DPSD signifies that the processed signal more closely resembles the original EEG signal, with fewer artifacts remaining after blink elimination. This indicates higher signal quality and reliability for subsequent analysis tasks such as frequency domain analysis. Fig. 4 shows the DPSD plot for both phases. The plot 'a' shows the DPSD values in phase 1 and plot 'b' shows the DPSD values in phase 2.

(a)

(b)

Fig. 4. *Illustrates the Difference in power spectral desity plots in both phases: (a) . is the power spectral desity in phase 1, (b) is the power spectral density in phase 2.*

The results from Fig. 4 conclude that the DPSD values are sufficiently low which resembles the elimination of power frequency components of noise efficiently in both the phases. The phase wise elimination of noise yields better results in terms of Signal to Artifact Ration (SAR) and Difference in Power Spectral Density (DPSD).

The proposed technique has also performed well in comparision with the other VMD based algorithms. Table. 4 shows the results of the other VMD based algorithms

V. CONCLUSION

This study addressed blink extraction in the EEG signal using a single channel framework based on Two-Phase VMD. The proposed architecture consists of two stages: first, VMD breaks down the signal into several components, and then, using OA features, the contaminated EEG signal is regressed to produce noise-free EEG. The four components of the proposed framework are as follows: VMD-I for decomposing the EEG signal into two modes; VMD-II for processing the EEG signal into three modes; and regression of mode with ocular artifacts (OAs) derived frombased on threshold criteria based on turning point counts. The Mendeley database, one of the publicly accessible databases, is used to collect EEG signals with a variety of OAs present, each with a different magnitude and shape, aiming toverify the efficacy of the suggested structures. The findings show that the suggested framework can remove OAs from all EEG local rhythms and the reconstructed EEG signal with little loss of important clinical aspects. Additionally, it performs better in both objective and subjective assessments than the few applied current OAs removal strategies.

VI. REFERENCES

[1] D. Chinmayee and B. Pradyut Kumar, "An improved algorithm for efficient ocular artifact suppression from frontal EEG electrodes using VMD," in *Dept. of Electronics & Telecommunication Engineering, International Institute of Information Technology*, Bhubaneswar, India, 2020.

[2] M. Saini, Payal and U. Satija, "An Effective and Robust Framework for Ocular Artifact Removal From Single-Channel EEG Signal Based on Variational Mode Decomposition," in *IEEE Sensors Journal,vol. 20, no. 1, pp. 369-376*, 2020.

[3] R. Patel, M. Janawadkar, S. Sengottuvel, K. Gireesan and T. Radhakrishnan, "Suppression of eye-blink associated artifact using single channel EEG data by combining crosscorrelation with empirical mode decomposition," in *IEEE Sens Journal, 16:6947–54*, 2016.

[4] M. Soomro, N. Badruddin, M. Yusoff and M. Jatoi, "Automatic eye-blink artifact removal method based on EMD-CCA," in *ICME International Conference on Complex Medical Engineering (CME);pp. 186–90.*, 2013 .

[5] K. Sweeney, S. McLoone and T. Ward, "The use of ensemble empirical mode decomposition with canonical correlation analysis as a novel artifact removal technique," in *IEEE Trans Biomed Eng 2013; 60:97–105.*.

[6] X. Chen, X. Xu, A. Liu, M. McKeown and Z. Wang, "The use of multivariate EMD and CCA for denoising muscle artifacts from few-channel EEG recordings," IEEE Trans Instrum Meas 2018; 67:359–70..

[7] M. Mannan, S. Kim, M. Jeong and M. Kamran, "Hybrid EEG Eye tracker:automatic identification and removal of eye movement and blink artifacts from electroencephalaographic signal," in *Sensors 2016;16:241.*.

[8] A. Jafarifarmand, M.-A. Badamchizadeh, S. Khanmohanmmadi, M. Nazari and B. Tazehkand, "Real time ocular artifacts removal of EEG data using a hybrid ICA-ANC approach," in *Biosignal process Control 2017;31:199-210.*

[9] B. Somers, T. Francart and A. Bertrand, "A generic EEG artifact removal," in *J Neural Eng ; 15:036007.*, 2018

[10] F. P. S. ,. Tamburro G, J. Haueisen and S. Comani, "A new ICA-based fingerprint method for the automatic removal of physiological artifacts from EEG recordings," in *PeerJ;6:e4380*, 2018.

[11] T. Dutta, U. Satija, B. Ramkumar, M. S. Manikandan, "A novel method for automatic modulation classification under non-Gaussian noise based on variational mode decomposition," 22nd Nat. Conf. Commun., 2016.

[12] Q. Zhao et al., "Automatic identification and removal of ocular artifacts in EEG-improved adaptive predictor filtering for portable applications," IEEE Trans. NanoBioscience, vol. 13, no. 2, pp. 109-117, June 2014.

[13] MA Klados and PD Bamidis, "A semi-simulated EEG/EOG dataset for the comparison of EOG artifact rejection techniques." Data in brief, vol.8, pp. 1004-1006, 2016.

[14] A. Goldberger, L. Amaral, L. Glass, J. Hausdorff, P. Ivanov, R.Mark, J. Mietus, G. Moody, C.-K. Peng, and H. Stanley, "Physiobank,physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," Circulation, vol. 101, no. 23, pp. e215 e220, 2000.

[15] M. S. Manikandan and S. Dandapat, "Wavelet energy based diagnostic distortion measure for ECG," Biomedical Signal Processing and Control, vol. 2, no. 2, pp. 80-96, 2007.

CMOS VLSI Implementation of Implicit Pulsed Dual Edge Triggered Flip Flop using Pass Transistor Logic for Power Efficient Applications

P.Nagarajan[1]
Associate Professor / ECE
SRM Institute of Science and Technology, Vadapalani, Chennai,
Tamilnadu, India.
nagarajan.pandiyan@gmail.com

N.Ashokkumar[2]
Professor/ECE
Mohanbabu University (Erstwhile Sree Vidyanikethan Engineering College)
Tirupati-AndraPradesh.
ashoknoc@gmail.com

Kavitha Thandapani[3]
Professor/ECE
Vel Tech Rangarajan Dr Sagunthala R&D Institute of Science and Technology Chennai, TamilNadu India
kavithaecephd@gmail.com

B.Manimaran[4]
Assistant Professor / ECE
Rajalakshmi Institute of Technology, Kuthambakkam,
Chennai, Tamilnadu, India.
manimaran.b@ritchennai.edu.in

Krishnamurthy Mayathevar [5]
Assistant Professor / ECE
PSNA College of Engineering and Technology, Dindigul,
Tamilnadu, India.
krishnamurthy184@gmail.com

P.Kabilamani[6]
Assistant Professor / ECE
SRM Institute of Science and Technology, Vadapalani, Chennai,
Tamilnadu, India.
kabilamanipandy08@gmail.com

Abstract— In general, the flipflops are acting as a key circuit and major power intaking element in several digital systems modeling. In this work, a new power efficient flipflop topology titled Pass Transistor Logic based Implicit Pulsed – Dual Edge Triggered Flipflop (PTIP-DETFF) is offered, by fusing the power minimization techniques such as minimum counts of device utilization, implicit pulsed clocking, dual edge triggering and parallel paradigm schemes. The implicit clock pulsing scheme eludes the exterior usage of clock generation network which intakes extra power. The twin edge triggering scheme halves the clock frequency which minimizes the dynamic power intake of flipflop as half. The parallel paradigm style diminishes the undesirable intake of power by the circuitry section which is not significant to result the preferred output. The offered flip flop topology is simulated using 0.12µm CMOS process technology and assessed in view of overall count of transistors, count of clock enabled loads, Layout area, Data in to output delay, entire power intake and also in terms of power delay product, Energy delay product and power energy products. The anticipated topology intakes the entire power of 5.265µW and achieves the power efficiency ranging from 12% to 63.30%, which conveys this anticipated flipflop will be appropriate for power efficient VLSI applications.

Keywords— flip flops, low power, CMOS technology, dual edge triggering

I. INTRODUCTION

In present era, the Power intake is the primary issue while designing the circuits and systems for low power applications [1, 2]. In general, the flipflops are acting as a key circuit and major power intaking element in various digital systems design [3]. In the global power intake of digital system, the flipflops accounts the power intake of 30% to 60% which is extreme value in the overall power intake [3, 4]. To drop the overall power intake of digital system design, the crucial flipflop element should be low power intaking element. The lessening of power intake of flipflops will have deep impression on the entire power intake of the digital systems.

The power efficient flipflops which are realized by Master-slave fashion and twin edge stimulating are used as major circuitry sections in contemporary low power processors and embedded systems. Typically master slave pattern are designed with a collection of latches in which one is transparent high and another one is opaque low [5, 6]. The master- slave style may be realized by using the parallel paradigm fashion where one latching section is replicated into manifold times as per the topology desires. Edge triggering flipflops diminish multiple stages into single stage by ensuing the positive and negative set up time. The pulse triggering flip flops are categorized into explicit and implicit style which duplicates the latching sections into several copies and aligned as parallel fashion [5, 6].
The Entire Power intake can be articulated as

P= Static (Ps) + Dynamic (Pd) + Short circuit (Psc)

Ps- designates static power utilization which happens due to leakage current flow. Pd - indicates dynamic power dissipation which arises because of undesirable on off activities. Psc- designates short circuit power dissipation which ensues due to a pathway exists between power supply (VDD) and ground (VSS) if both pull up and pull down net conducts simultaneously.

Power utilization of the flipflops can be diminished by ensuing various power efficient strategies as follows. [7, 8]

- Twin edge triggering

- Low swing supply voltage
- Pathway Splitting scheme
- Lessening capacity of clocking transistors
- Diminishing the overall count of loads

In recent years, various flipflops are designed by using different power efficient strategies and listed as CTS-DETFF [9], DDNET-D2 [10], S-TCRFF [11], RTSPCFF [12], DCSFF [13], MS-PTDFF [14] in comparison table-1 which are taken into account to evaluate the significance of an anticipated flipflop topology.

II. Proposed Pass Transistor Logic Based Implicit Pulsed – Dual Edge Triggered Flipflop (PTIP-DETFF) Design

A. Construction of anticipated flipflop topology

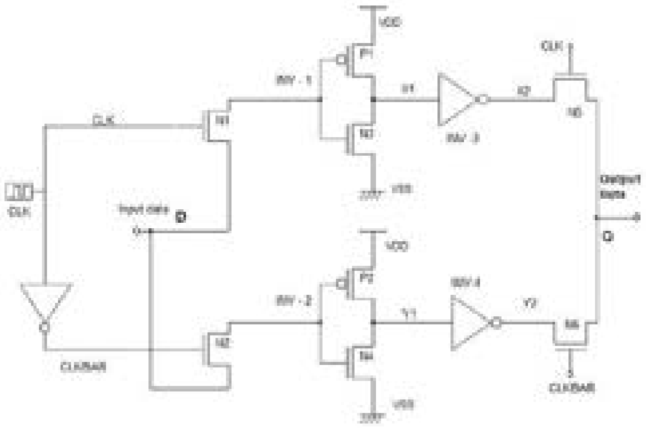


Fig. 1. Circuit Topology of Proposed Pass Transistor Logic based Implicit Pulsed – Dual Edge Triggered Flipflop (PTIP-DETFF), Overall 14 counts of transistor loads including 6 counts of clock enabled loads

The anticipated flipflop topology is framed by three dynamic sections titled Upper Latch (UL), Lower Latch (LL), the data selection network and entirely using 14 counts of transistor loads comprising 6 counts of clock enabled Transistor loads. An upper latch is realized by 5 numbers of elements N-type transistors N1, N3 and P-type transistor P1with an inverter named INV-3. The lower latch is realized by 5 numbers of devices N-type N2, N4 and P-type P2 with an inverter named INV-4 shown in Fig.1. By following the power efficient technique named minimum counts of device utilization, both the UL and LL latching sections are topologied by only 5 counts of devices each. Due to the fewer count of element usage in the latching sections, the power utilization by the latches becomes less. In this anticipated flipflop topology the clock triggering signal not generated by an external clock network. Instead, the clock is applied as an implicit manner to the latching sections directly, without any separate clock distribution network. This implicit clock feeding scheme obviously leads to lessening of power wastage. The data selection network of this proposed design is framed by using pass transistor logic which consists of N-type transistors Pair N5 and N6. Instead the complex data selection logic, this proposed design using the simple dual parallel connected pass transistors which minimizes the power usage of the data selection network which are highlighted in Fig.2.

The upper latch and lower latch sections are coupled in parallel style with data selection logic by following standard parallel paradigm approach. To trigger the complete flipflop topology, the clock signal CLK is directly fed to an upper latch and data selection pass transistor N5. An inverted clock signal CLKBAR drives the lower latch and data selection pass transistor N6.

The input data D is commonly fed to both upper latch and lower latch through the source terminal of the data feeding transistors N1, N2 Respectively. An upper latch is exclusively responsible to capture the input D while the clock triggering signal CLK is logically high. The UL results the data out at node X2 and the same fed to pass transistor N5 in the data selection logic. The device N5 is completely under the control of clock triggering signal CLK which passes the output data Q if the CLK is logical high. In this scenario the LL is completely in deactivate mode and logically isolated from the entire circuit.

The lower Latch is exclusively responsible to capture the input D while the clock triggering signal CLK is logical low (i.e.) CLKBAR becomes logically high value. The LL results the data out at Y2 node and the same fed to pass transistor N6 in the data selection network. The element N6 is completely under the control of CLKBAR signal which passes the output data Q if the CLKBAR is logically high. In this state the UL is completely in deactivate mode and logically isolated from the whole circuit.

B. Functionality of the proposed topology

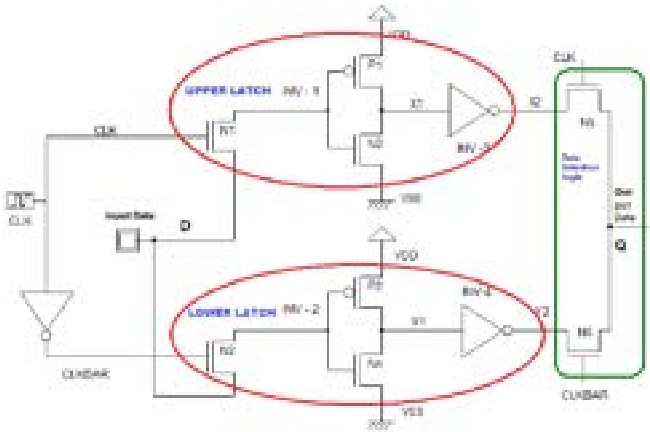


Fig. 2. schematic of PTIP-DETFF with its highlighted dynamic sections

The functionality of the proposed PTIP-DETFF is elaborated as 4 set of operations with right data set of input Data D and output Q as follows:

Operation (1): When the input Data D=0 (LOW) and the dropping edge of the clock arrives CLK=0 (LOW), the device N1 turns off. The off state N1 is incapable to drive both an elements N3 and P1 which formulates an inverter I1. Presently an Inverter-1 and an Inverter-3 is momentarily cutoff from the circuitry data path and no data is served as an input to the pass transistor N5. Due to CLKBAR=1 (HIGH), the transistors N2 gets ON. The on state N2 permits the input D=0 value to an Inverter I2. The transistors P2 and N4 employed in an Invetter-2 gets ON and OFF correspondingly. The on state P2 makes the Y1 node to high & an Inverter-4 results logical low and the same value served as input to pass transistor N6. Due to the control signal CLKBAR =1, the device N6, got enabled and passes the logic low value to the output data path. The output data Q results logic low as an output i.e. Q = 0 (LOW).

Operation (2): when the input Data D=0 (LOW) and the growing edge of the clock arrives CLK=1 (HIGH), the device N1 turns on. The on state N1, drives the elements P1 and N3 which results the node X1 = high. Due to X1= high, an inverter INV-3 results X2=low and the same served as input for the pass transistor N5 in the data selection logic. Due to CLK=1. The Pass transistor got excited and directly passes the X2=low value as low output. i.e Q=0 (LOW). In this scenario due to CLK=1 & CLKBAR =0, the lower latch and the pass transistor N6 are momentarily detached from that circuitry part.

Operation (3): when the input Data D=1 (HIGH) and the dropping edge of the clock arrives CLK=0 (LOW), the device N1 gets off. The off state N1 is incapable to drive an inverter I1. Presently an Inverter-1 and an Inverter-3 is momentarily detached from the circuitry data path and no data is served as an input to the pass transistor N5.Due to CLKBAR=1 (HIGH), the device N2 gets ON. The on state N2 passes the input D=1 value to an Inverter I2 and results the Y1 node to low and an Inverter-4 results logical high. Due to the control signal CLKBAR =1, the device N6 got enabled and passes the logic high value in node Y2 to the output data path. The output data Q results logic high as an output i.e. Q = 1 (HIGH).

Operation (4): when the input Data D=1 (HIGH) and the growing edge of the clock arrives CLK=1(HIGH) the device N1 turns on. The on state N1, drives the elements P1 and N3 which results the node X1 = low. Due to X1= low, an inverter INV-3 results X2=high and the same served as an input for the pass transistor N5 in the data selection logic. Due to CLK=1. The Pass transistor N5 got energized and directly passes the X2=high value as low output. i.e Q=1 (HIGH). In this scenario due to CLK=1 & CLKBAR =0, the lower latch and the pass transistor N6 are momentarily detached from that circuitry part.

In this this proposed topology, by following the parallel paradigm approach, the upper latch section samples the data input while clock triggering signal becomes HIGH i.e CLK=1 . Same the lower latch section samples the data input while clock triggering signal becomes LOW i.e CLK=0. Due to this parallel paradigm approach while the proposed topology is under working condition, one latch section will be in active mode and another one latch section will be temporarily in ideal mode. So the power utilization and the Data D – Q out delay will be minimized for the proposed flipflop design.

III. Results and Discussion

The proposed PTIP-DETFF is implemented in digital schematic DSCH and simulated at the layout level with 0.12 µm CMOS VLSI technology in Microwind EDA environment. The significance of this proposed topology is assessed in terms of performance factors such as overall count of transistor loads, count of clocked loads, Data D – Q out delay, triggering mode, Layout area occupation and Entire power intake and optimization factors such as Power & Delay Product (PDP), Energy & Delay Product (EDP) and Power & Energy Product (PEP).

The anticipated design results the response at both edge of the clock triggering signal and Summarized as follows.

Operation -1: Datain D=0, CLK=0, CLKBAR=1, Qoutput =0

Operation -2: Datain D =0, CLK=1, CLKBAR=0, Qoutput=0

Operation -3: Datain D=1, CLK=0, CLKBAR=1, Qoutput=1

Operation -4: Datain D=1, CLK=1, CLKBAR=0, Qoutput=1

The functionalities of proposed PTIP-DETFF topology are analyzed by exiting the all possible combinations of input data D (D=0 & D=1) and Clock (CLK=0 & CLK=1) as given above operation -1, 2, 3 and 4. The simulation results demonstrates the correct functioning of proposed design for respective Datain D & CLK and shown in Fig. 3,4,5 and 6.

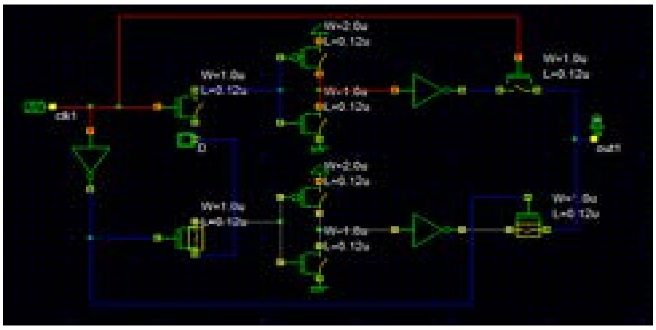


Fig. 3. Operation -1: Datain D=0, CLK=0, CLKBAR=1, Qoutput =0

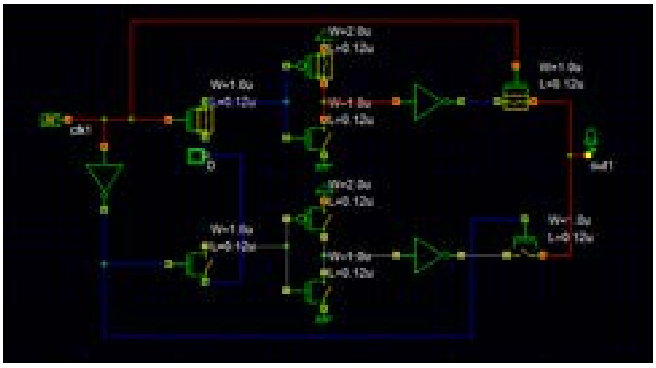


Fig.4.Operation -2: Datain D =0, CLK=1, CLKBAR=0, Qoutput=0

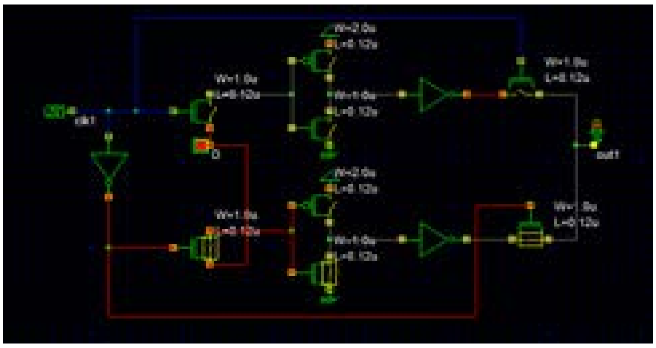


Fig.5. Operation -3: Datain D=1, CLK=0, CLKBAR=1, Qoutput=1

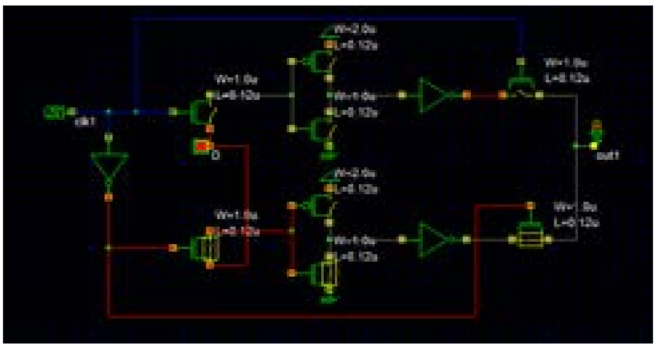


Fig.6. Operation -4: Datain D=1, CLK=1, CLKBAR=0, Qoutput=1

The performance and optimization factors for existing and anticipated flipflop topology are tabulated in the Table-1 and Table -2 respectively.

In point of overall count of Transistors, the anticipated design uses 14 count of elements which is considerably less compared with existing design. Due to usage of power efficient technique named minimum count of transistors results less power intake.

Principally the clock enabled devices are highly dynamic and intakes extra power compared with nonclocked devices [9, 14]. In sense of clock enabled transistor count, the anticipated design consist 4 elements, which is considerably fewer compared with existing topologies. This effects fewer dynamic power intake and leads to minimum entire power intake.

In aspect of entire power intake, the anticipated topology, consumes the power of 5.265µW as shown in Fig.7, which is very less compared to existing flipflop topologies. The anticipated topology achieves the power efficiency ranging from 12% to 63.30% compared to the existing topologies listed in Table-1. This minimum power intake makes the proposed PTIP-DETFF as greatly appropriate for power efficient VLSI applications.

In vision of Datain D – Q out delay, the anticipated design takes the delay of 130 ps which is minimum value compared with existing topologies in Table-1 except the design DDNET-D2. This lowest delay makes the proposed topology as also fit for high speed applications.

In note of Physical Layout area utilization, the anticipated PTIP-DETFF occupies an Layout area of 247µm2 which is demonstrated in Fig.8. This is considerably less compared to the topologies listed in table-1, except DDNET-D2, S-TCRFF and DCSFF. The Comparison chats Fig.9 and 10 clearly demonstrates the Comparison of overall Count of transistors& clock enabled loads and Comparison of Entire power intake.

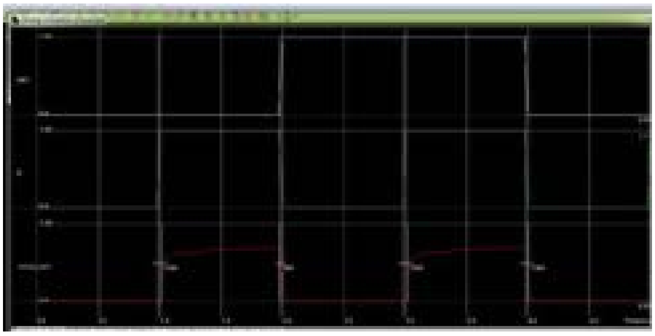


Fig.7.Entire power intake of proposed PTIP-DETFF (Power = 5.265 µW)

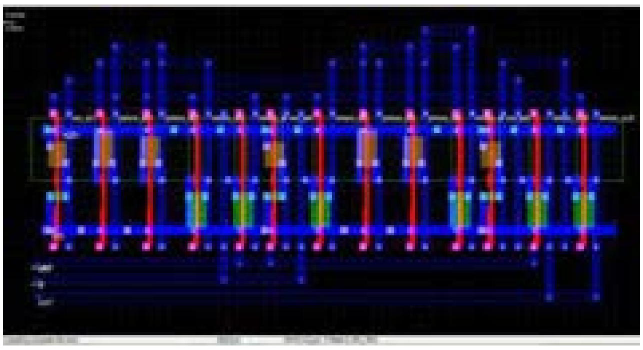


Fig.8 Physical Layout area of proposed PTIP-DETFF (Area = 247 μm^2)

In account of power delay product, the anticipated topology possess the value of 0.684 fj which is less compared to the existing topologies except S-TCRFF. In view of Energy Delay Product, the proposed PTIP-DETFF intakes the cost of 0.088*10-24 which is considerable rate compared with existing topologies. Principally Power Energy Product (Power * Power * Delay) is the metric which gives more weightage to power consumption compared with delay [15]. In account of

TABLE I. ASSESSMENT OF PERFORMANCE FACTORS

Flipflop Designs	Performance factors					
	Overall count of transistor Loads	Count of clock enabled loads	Triggering mode	Layout area (µ m^2)	Data D –Q Out delay (ps)	Entire Power Intake (µw)
CTS-DETFF [9]	23	8	Dual	424	179	8.347
DDNET-D2 [10]	14	2	Dual	176	91	7.610
S-TCRFF [11]	16	12	Dual	201	86.31	7.310
RTSPCFF [12]	26	5	Single	616	77.31	8.598
DCSFF [13]	24	7	Single	180	112.7	7.40
MS-PTDFF [14]	16	6	Dual	252	206	5.945
Proposed PTIP-DETFF	**14**	**6**	**Dual**	**247**	**130**	**5.265**

TABLE II. ASSESSMENT OF OPTIMIZATION FACTORS

Flipflop Design	Optimization Factors		
	PDP (fj)	EDP ($*10^{-24}$)	PEP ($*10^{-20}$)
CTS-DETFF	1.503	0.2690	1.263
DDNET-D2	0.692	0.0630	0.527
S-TCRFF	0.630	0.0545	0.461
RTSPCFF	0.664	0.0513	0.571
DCSFF	0.833	0.093	0.617
MS-PTDFF	1.224	0.252	0.728
Proposed PTIP-DETFF	**0.684**	**0.088**	**0.360**

Power Energy Product, the PTIP-DETFF circuit achieves very less value of 0.360*10-20 and reported in Fig.11.For the proposed PTIP-DETFF, the PEP is improved in the range of 28.05% to 71.49 % which states the anticipated design is more appropriate for power efficient VLSI applications.

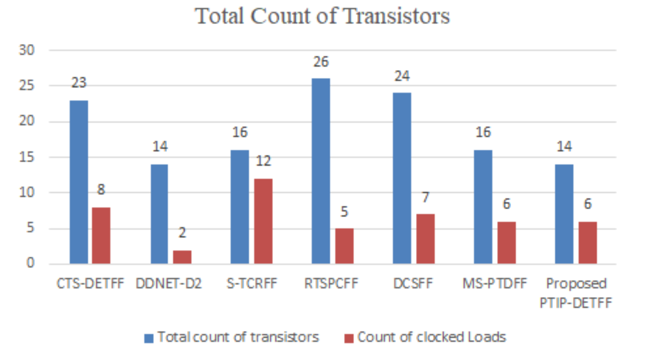


Fig.9 Comparison of overall Count of transistors and clock enabled loads

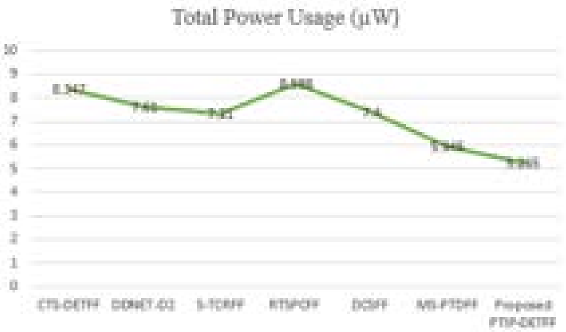


Fig.10 Comparison of Entire power intake

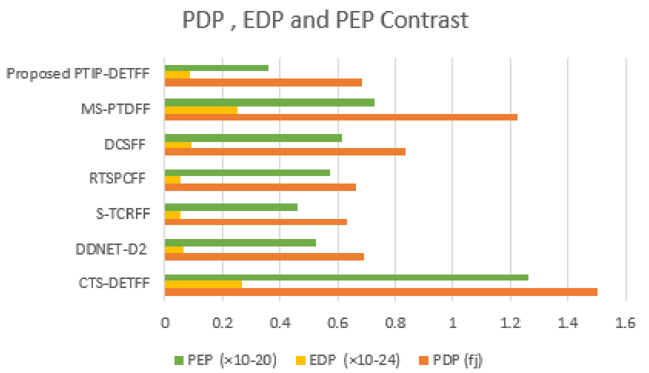


Fig.11 Comparison of PDP, EDP and EDP

IV. Conclusion

This PTIP-DETFF topology is realized by ensuing the power effective practices such as minimum counts of device utilization, implicit pulsed clocking, twin edge triggering and parallel paradigm schema. The implicit clock pulsing scheme evades the external usage of clock generation network which intakes surplus power. The parallel paradigm approach reduces the unwanted intake of power by the circuitry section which is not vital to result the desired output. The anticipated topology intakes the power of 5.265µW , attains the power efficiency ranging from 12% to 63.30% and reaches PEP enrichment in the range of 28.05% to 71.49 % which states that the anticipated design is well appropriate for power efficient VLSI applications.

References

[1] Neil H.E.Weste and Kamaran Eshraghian.,Princinles of CMOS VLSI Design: A System perspective, (2003).

[2] Peiyi Zhao, Jason Mc Neely, Weidong kuang, Nanwang and Zhongfeng Wang, "Design of sequential elements for low power clocking system", IEEE Transaction on VLSI systems. vol.19 ,pp.914 – 918, 2011.

[3] Gary Yeap, Practical Low Power Digital VLSI Design, 1998,Kluwer Academic Publishers Boston/ Dordrecht/London.

[4] P.Nagarajan , G.Mohanbabu and R.Prasanna , "Design of power gated double edge triggered flip flop using sleep transistor technique (PG-STDETFF)", The patent office journal, Government of India , pp. 18111,2017.

[5] Kunwar Singh, Satish Chandra Tiwari and Maneesha Gupta, "A Modified Implementation of Tristate Inverter Based Static MasterSlave Flip-Flop with Improved Power-Delay-Area Product", Hindawi Publishing Corporation Scientific World Journal ,2014, Article ID 453675.

[6] G. Gerosa, "A 2.2 W, 80 MHz superscalar RISC microprocessor",IEEE J.Solid-State Circuits, vol.29, pp.1440-1454, 2005.

[7] C.K.Teh, M.Hamada, T.Fujita, H.Hara, N.Ikumi and Y.oowaki, "Conditional data mapping flip-flops for low power and high performance systems," IEEE Trans. VLSI syst., vol.14,1379-1383, 2006.

[8] Kalarikkal Absel, Lijo Manuel & Kavitha, RK, "Low-Power Dual Dynamic Node Pulsed Hybrid Flip-Flop Featuring Efficient Embedded Logic", IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 21, 2013.

[9] P.Nagarajan, R.Saravanan, and P.Thirumurugan, "Design of Register Element for Low Power Clocking System", Information an interdisciplinary research, ISSN 1343-4500, vol.17, pp.2903-2913, 2014.

[10] P.Nagarajan , T.Kavitha and S.Shiyamala , "Efficient timing element design featuring low power VLSI applications", International Journal of Engineering and Technology , vol. 8,pp. 1696-1705, 2016.

[11] Neethu Anna Sabu and K. Batri, " Design and Analysis of Power Efficient TG based Dual Edge Triggered Flip-Flops with Stacking Technique", Journal of Circuits, Systems and Computers, Vol..29,No.8, ,2020.

[12] Heng you, zenghui Yu and shushan qiao " Low Power Rentive True Single Phase Clocked Flop-flop with redundant precharge free operation", IEEE Transactions on VLSI Syatems, Vol.29, No.5, pp.1022-1032, 2021.

[13] Jun-Young Park, Minhyun Jin, Soo-Youn Kim and Minkyu Song 2022," Design of a Dual Change-Sensing 24T Flip-Flop in 65 nm CMOS Technology for Ultra Low-Power System Chips", Electronics, pp 01-10.

[14] Pandian Nagarajan, Thandabani Kavitha et al "Power energy and power area product simulation analysis of master-slave flip-flop. (2023). Revue roumaine des sciences techniques — série électrotechnique et énergétique, ISSN (online): 00354066, December 2023.

[15] Dipanjan Sengupta and Resve saleh, "Power – Delay metrics Revisited for 90nm CMOS Technology", Proc. of the sixth international symp. on ISQED, 2005.

Blockchain-powered Federated Learning: A Review on Current Practices

Dr. Neda Fatima
School for Computer Applications, Manav Rachna International Institute of Research and Studies in Faridabad, India
drnedafatima@gmail.com

Dr. Mukesh Prasad
School of Computer Science, UTS Australia
Mukesh.Prasad@uts.edu.au

Dr. Usha Batra
1. Visiting Postdoctoral Fellow, University of Technology, Sydney, Australia
2. Professor, Computer Science Engineering Skill Department of CS and IT, ShriVishwakarma Skill University, India.
dr.ushabatra@gmail.com

Hasim Ali
School for Computer Applications, Manav Rachna International Institute of Research and Studies in Faridabad, India.
hasimali8682@gmail.com

***Abstract*—With improvements in machine learning technology, viable methods for processing the massive amounts of data created in real life are now available. However, data protection and scalability concerns will slow down the further development of machine learning algorithms. Federated learning (FL) prevents data leaks by distributing training jobs to numerous clients and thus isolating the central server from the local devices. Nevertheless, FL still has issues like malicious data and single points of failure. Blockchain Technology's debut offers a secure and practical solution for FL deployment. This research study presents a comprehensive evaluation of the blockchain-based FL. First, the concept of Federated Learning is extensively dwelled upon. Then, the limitations of Federated Learning were analyzed, which led to the adoption of BC-FL. Some real-world uses of BCFL are also analyzed. Subsequently, the integration of BC with FL has been studied. Finally, the architecture of BC-FL has been developed and described. BC-FL is an upcoming research field that is comparatively unexplored. The paper offers insights into its development and how the technology can be implemented by addressing data privacy and security issues.**

Keywords—Blockchain, Federated Learning, Machine Learning, Privacy, Security.

I. INTRODUCTION

The introduction effectively outlines the challenges of data protection and scalability in machine learning. It highlights the issues of data transmission costs and privacy risks associated with uploading raw data to central servers. These challenges hinder the efficiency and security of traditional machine-learning architectures. To address these issues, the paper introduces federated learning (FL), a unique ML architecture proposed by Google, which aims to preserve user privacy while allowing distributed devices to train models collectively. However, traditional FL architectures face limitations such as speed bottlenecks, unreliable data, and privacy breaches.

To achieve privacy in federated learning (FL), particularly when merging it with blockchain (BCFL), several key strategies are employed:

Decentralization: FL inherently operates decentralized, allowing training to occur on local devices without uploading raw data to a central server. This decentralization minimizes the risk of privacy breaches associated with data transmission.

Encryption: Utilizing encryption techniques such as homomorphic encryption and differential privacy ensures that sensitive data remains private even during model training and communication between devices and servers.

Anonymization: Before data is shared or used for training, it can be anonymized to remove personally identifiable information, reducing the risk of privacy leakage.

Blockchain Immutable Ledger: Integrating blockchain technology provides an immutable ledger for recording transactions and interactions within the FL system. This ensures transparency and trust in the system's operations while maintaining participants' privacy.

Smart Contracts: Smart contracts can enforce rules and permissions regarding data access and model training, ensuring that only authorized parties can participate and access specific data, thus preserving privacy.

Consensus Mechanisms: Blockchain's consensus mechanisms ensure that network participants agree upon updates to the FL model, maintaining the model's integrity while preserving privacy.

Machine learning (ML) is used in every industry and has dramatically impacted people's lives. To train machine learning (ML) models, data generated daily from a sizable number of end users may be obtained. These models serve us by delivering improved services that enhance our quality of life. However, there are two issues with the current ML architecture, which frequently requires end devices to send data they have collected to a central server for model training. For starters, data transmission might cost a significant quantity of communication resources. Second, uploading raw data increases the risk of privacy leakage, discouraging data owners from doing so out of security concerns.

To solve the challenges above, Google suggested federated learning (FL). This unique ML architecture may successfully preserve user privacy while collectively allowing many end devices to train an ML model. A cutting-edge deep learning strategy called federated learning makes learning easier from a remote environment. A new artificial intelligence system called Google created Federated Learning in 2016 with the primary goal of solving the problem of the local update model for Android mobile device consumers. The design aims to perform fast machine learning with several parties or compute nodes while ensuring legal compliance. It also aims to secure terminal data and personal data privacy through massive data sharing and information security. The global data learning model maintained on tens of millions of remote client devices is the foundation for the traditional federated learning problem. The client device must often be connected to the central server during training. It causes speed bottlenecks, unreliable data, privacy breaches, and other issues. Figure 1 depicts the menthodology of Federated Learning (FL).

Fig. 1. Federated Learning

Blockchain is the world's most recent windfall, mainly in banking and finance. Blockchain is also employed in the healthcare management system to effectively maintain electronic health and medical records. Security, privacy, and immutability are all guaranteed by the technology. Blockchain distributed trust offers a fresh approach to designing the FL architecture and paradigm. FL could be applicable in the blockchain setting since distributed consensus and transaction records allow for safe communication in an unreliable environment. Concurrently, incentive and consensus-based systems are inherently suitable for creating a market economy that may successfully promote participation and information flow in FL. This paper presents a framework for providing a personalized recommendation system by merging blockchain with Federated Deep Learning. Figure 2 shows the integration of blockchain and Federated Learning (FL).

Fig. 2. Integration of Blockchain with Federated Learning

The research article is structured as follows: Section II begins with an overview of the literature. Section III examines the history and origins of the ideas of Federated Learning and Blockchain. Section IV elaborates on the limitations of federated learning. Section V examines the relevance of the amalgamation of the two technologies of Federated Learning and Blockchain BCFL. Section VI describes the BCFL architecture in detail. The paper comes to a close with Section VII.

II. Literature Review

Federated learning is a relatively novel notion in machine learning, and recent research has focused on how the blockchain could be used in this scenario. Biscotti is described in [1]. It is a fully decentralized peer-to-peer (P2P) approach to multi-participant machine learning that employs cryptographic primitives and blockchain to coordinate an ML process while protecting peering clients' anonymity. According to the author, Biscotti is extendable, error-resistant, and resistant to known dangers. When 30% of the adversaries are present, Biscotti can ensure the privacy of each client's update while maintaining the global model's performance at scale. The review research in [2] undertakes a detailed investigation of the privacy and security debate in the context of BC-FL approaches using empirical data to provide an impartial road map of the problem state. The study's findings indicate that the usage of blockchain-based FL has expanded dramatically over the previous five years. Blockchain technology is rapidly being utilized to tackle difficulties with hospital administration, hospital data on various diseases, CT scans and X-ray records, medical equipment, and other issues.

In the study reported in [3], smart devices based on the Internet of Things (IoT) use federated learning to use distributed clinical data most. This technique is based on blockchain technology and is aimed at intelligent healthcare. To protect privacy, an adaptive differential privacy approach is presented, combined with a gradient-verification-based consensus mechanism to identify poisoning attempts. Using a diabetic dataset from the real world, the method is contrasted with two equivalent techniques for performance metrics. The test’s findings indicate that the recommended method may produce highly accurate models in a manageable length of time while also consuming less privacy budget and thwarting attempts at poisoning.

In [4], the article thoroughly analyzes BCFL and considers the effects of such a unique paradigm. It begins by briefly introducing FL technology and discussing the issues that it faces. The Blockchain ecosystem is then summarised. Following that, it highlights BCFL's structural design and platform. Furthermore, improvement in FL performance using Blockchain and many integrated uses of FL incentive systems are demonstrated. Finally, BCFL's industrial application possibilities are described. The review paper [5] uses in-depth literature analysis in this work to analyze existing FL issues and synthesize knowledge on blockchain-based FL (BFL). The differences between these gathered BFL designs are discussed, and BFL applications are categorized and examined. Finally, several suggestions are made for BFL's potential development and use.

The release of this paper [6] gives a comprehensive review of the use and implementation of frameworks for Blockchain-Enabled Federated Learning for the Internet of Vehicles. Also emphasized are possible problems, challenges, issues, solutions, and future research goals for BC-enabled FL frameworks for IoVs. The survey results might then be utilized to develop BC-enabled FL solutions that address various data privacy issues and IoV scenarios. The paper [7] explores FL and looks at the main evolutionary pathway for challenges during the FL development process. This study will examine current industrial engineering applications to identify potential future landing purposes. Furthermore, our research suggests six study topics to address the FL literature and increase our understanding of FL for potential future optimization. This work contributes to the conclusion of applications in computer science and industrial engineering while also providing an overview of FL applications. In [8], Besides providing an overview of current approaches, the article looks at the distinct characteristics and restrictions of federated learning. It also identifies several potential areas for future study that might benefit many different academic communities. The study provides an overview of federated learning, a learning paradigm in which statistical models are taught at the network's edge. It has spoken about federated learning that differs from standard distributed data center computing and the issues that come with it. Using classic privacy-protecting learning, there is a thorough explanation of traditional findings and current federation work. [9] presents an architecture for distributed computing defense for a sustainable society using blockchain technology and federated learning. The recommended model contains a method for overcoming the issues of little training data to obtain high accuracy without using a reason-specific model. The dataset is used to investigate how the proposed framework performs in various circumstances to determine how successful the model is.

To preserve client data privacy in the IIoT, the study [10] suggests a BC-FL device failure detection approach. The initial step was creating a BC-FL architecture architecture for IIoT failure detection. The design allows you to verify the correctness of consumer data. According to the architecture, each client routinely creates a Merkle tree on the blockchain, with each leaf node representing a client data record. Furthermore, a distinctive centroid distance-weighted federated averaging technique was used to solve the issue of data heterogeneity in IIoT failure detection.

This method compensates for the difference between positive and negative classifications in each client dataset. The work in [11] creates a blockchain-enabled safe data sharing for the multiple stakeholders participating in this inquiry. Data privacy is safeguarded by supplying the data model rather than real data. Last but not least, federated learning is integrated into the permissioned blockchain consensus process so that federated training may be done using the same computer resources as consensus.

In [12], the paper looks at how data and security problems have emerged in bright edge computing devices and how to fix them. The proposed strategy combines machine learning methods like classifiers and optimization algorithms with Blockchain consensus to secure data. In addition, the authors used the tools available in an edge computing environment to send different data batches to various customers. The use of Blockchain servers resulted in the preservation of client anonymity. Furthermore, the authors developed a federated learning technique to segment the client data into learnable batches. The study's findings reveal how to deploy a Blockchain-based training model in a computing environment with an advantage. The article [12] examines how smart edge computing devices are developing data and security flaws, as well as how to correct them. The suggested method combines machine learning methods like classifiers and optimization approaches with data security. The authors also used the proposed methods in an edge computing setting, sending various clients distinct batches of data. The utilization of Blockchain servers ensured the protection of client anonymity. The authors separated the client data into batches, which were then taught via a federated learning approach. The findings of this study demonstrate how to implement a Blockchain-based training model in an edge-computing environment. The paper [13] described an FL with an incentive-aware MD powered by blockchain. It contributes in two important ways: first, it uses competition to encourage employees to maximize their benefits, and second, the design requires employees to obey the protocol without using cryptographic techniques such as homomorphic encryption or zero-knowledge proof. Second, several criteria for the protocol design's functionality have been quantitatively identified using contest theory. These theoretical proof results are optimistic for implementing the blockchain-enabled FL.

In [14], the research presents a special FL-Block (federated learning with blockchain) approach. With FL-Block, end-device local learning updates may be shared with a blockchain-based, miner-verified global learning model. Based on this, FL-Block uses the blockchain's PoW consensus mechanism to allow autonomous machine learning to maintain the worldwide model and coordinate without a centralized authority. Additionally, the performance of the FL-latency Block has been investigated, and the appropriate block production rate has been estimated while considering communication, consensus delays, and computational costs. Numerous assessment results demonstrate FL-Superior Block's efficiency, resistance to poisoning assaults, and privacy protection.

The research [15] gives an unbiased road map of this problem's current status by fully investigating the privacy and security argument in the context of blockchain-based FL methods on scientific datasets. The results of this study show that BC-FL has expanded dramatically over the last five years, and blockchain technology is being used more frequently to address issues with patient healthcare records, image retrieval, cancer datasets, industrial equipment, and economic data in the context of IoT applications and intelligent environments. The literature review in [16] suggests blockchain-based federated learning (FL) as a decentralized, efficient, and trustworthy federated learning solution. The concept has three major components: customers, blockchain technology, and machine learning engineers. The benefits of the new federated learning system built on the blockchain are also highlighted, as are several techniques for protecting model privacy on the blockchain.

III. FEDERATED LEARNING

Standard machine-learning algorithms require that training data be stored on a single computer or in a data center. Additionally, Google has constructed one of the most dependable and secure cloud infrastructures to improve services. It introduces Federated Learning, a fresh method for building models users interact with on mobile devices.

Federated learning reduces the requirement for machine learning to save data on the cloud by allowing mobile devices to participate in the creation of a shared predictive model, with all training data kept locally. Extending model training to the device, goes beyond utilizing local models to make predictions on mobile devices (such as the Mobile Vision API and On-Device Smart Reply).

The approach follows: after receiving the current model, the device updates it using the phone's data and then summarizes the changes in a brief, focused update. Only this model modification is safely transferred to the cloud, where it is quickly averaged with the changes of other users to enhance the overall model. All training data is kept on the device; no individual updates are saved on the cloud.

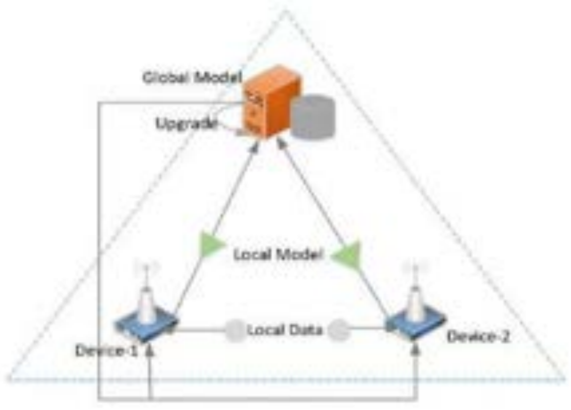


Fig. 3. Federated Learning

A distributed machine learning strategy for protecting data privacy is called federated learning. Federated learning systems rely on a central server to coordinate model training over several distributed client devices that store data. No data is transferred outside the client devices during the model training, which takes place locally on the clients. Decentralized methods can also be used for federated learning. Image, structured, and text data types are the most frequently utilized, with picture categorization being the most popular application. Federated learning is widely used by applications that infer personal data from mobile devices, such as images, financial or medical information, and text.

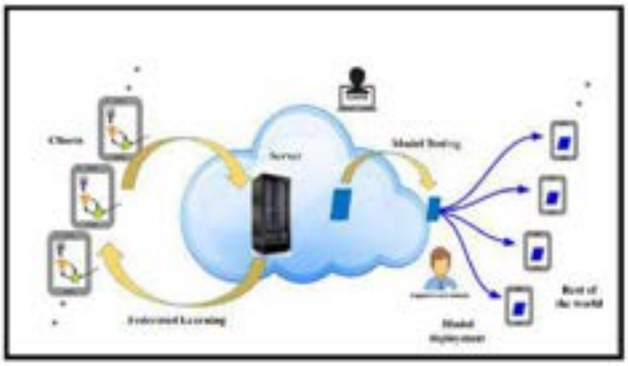


Fig. 4. Integration of Blockchain with Federated Learning

IV. LIMITATIONS OF FEDERATED LEARNING

FL relies on a central server susceptible to malicious activity, resulting in erroneous global model updates. All local model updates are distorted, reducing the precision of subsequent local model updates. Furthermore, due to network speed limits, sending many models at once may overburden the central server. One of the most common issues with federated learning is the single point of failure.

Fraudulent participants can cause training samples or models to fail, invalidating machine learning predictions, yet FL cannot audit malevolent trainers. Malicious clients can, therefore, tamper with the local model, upload erroneous mask gradients, and publish unsponsored shares to the center server, which can all damage the global model.

FL assumes that each local device will give out all of its data resources however this needs to be more practical. Because participants must use their data and computing resources, there is minimal incentive for participants to do model training. Mobile devices that are self-centered will be hesitant to engage in model learning unless well compensated. As a result, FL has a loss of motivation for model training.

The privacy of training data is compromised by FL frameworks even when the training resource is stored locally. Some researchers have created attacks that extract considerable information from intermediate gradients. Thus, there is a reasonable *threat to privacy* in FL.

1. Communication Overhead and Latency: FL relies on communication between local devices and the central server. However, transmitting model updates over

potentially slow or unreliable networks can introduce significant delays and increase communication overhead. This can lead to synchronization issues and hinder the efficiency of the learning process, especially in scenarios where real-time model updates are crucial.

2. Heterogeneity of Devices and Data: In a federated learning setting, devices contributing to the training process may vary widely in terms of hardware capabilities, network bandwidth, and the quality of their data. This heterogeneity can introduce challenges in effectively aggregating model updates. Models trained on data from diverse sources may generalize poorly to unseen data or suffer from biases inherent in the training data distribution.

3. Data Distribution Shifts: Data distribution across participating devices may evolve due to various factors such as user behavior changes, device upgrades, or changes in application usage patterns. These changes can lead to distributional shifts in the training data, potentially causing degradation in model performance if not adequately addressed through techniques like continual learning or adaptive federated learning algorithms.

4. Privacy Risks and Regulatory Compliance: FL raises significant privacy concerns as sensitive data remains on user devices and is only shared as model updates. However, even with this decentralized approach, FL frameworks are susceptible to privacy attacks such as membership inference and model inversion attacks. Ensuring compliance with privacy regulations like GDPR becomes challenging, mainly when dealing with sensitive data such as healthcare records or financial transactions.

5. Model Robustness and Security: Federated learning models may be vulnerable to adversarial attacks, where malicious actors attempt to manipulate the learning process by injecting carefully crafted data or perturbations into the model updates. Ensuring the robustness and security of federated learning models against such attacks requires the development of robust aggregation mechanisms and defenses against adversarial manipulation.

6. Scalability and Resource Management: As the number of participating devices or clients in a federated learning system grows, managing resources such as computation, storage, and bandwidth becomes increasingly challenging. Scalability issues may arise, impacting the ability to efficiently coordinate model updates and maintain system performance under heavy loads.

V. BC-FL INTEGRATION

The blockchain is a distributed, decentralized database. The system generates interactive records (transactions), organised into blocks and saved in each time segment. Any node in the system may carry out safe peer-to-peer transactions because each transaction is encrypted and secured by Proof of Work (PoW) mechanisms that cannot be manipulated or falsified. A block, as seen in Fig. 5, comprises many transaction records and a block header containing information. The limited definition of Blockchain connects these blocks with the hash reference in the block header to form an entire ledger. The blockchain comprises five layers: data, incentive system, consensus, network, and application.

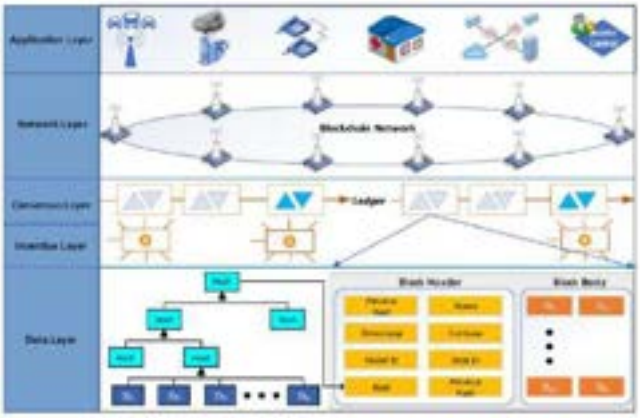


Fig. 5. Blockchain Architecture

Based on application scenarios and constructed technologies, blockchain is divided into three groups: public chain, consortium chain, and private chain. Typically, several Blockchain variants are selected depending on the requirements of distinct business scenarios. However, the public chain may generally only satisfy the initial design goal of the Blockchain.

Distributed blockchain nodes can be the FL center server to solve single-point failure. Miner nodes can transmit model updates for neighboring devices. Miner nodes are chosen independently or at random.

Each local device performs computations and sends local model changes to the linked miner. Miners can trade after completing Proof of Work, validate uploaded model modifications, and record authorized local model updates (PoW). The aggregated model adjustments are supplied, and the blocks holding those modifications are uploaded to the blockchain. The newly constructed block may also provide global model updates to any local device. The confirmed model may be aggregated, and chunks generated by local devices can be employed as miners.

Blockchain may also encourage collaboration and serve as a disincentive to misbehavior to advance the interests of the consumers. The local device may receive incentives if the local model updates are determined to be valid. If improper updates are uploaded, the local device will also be penalized. Blockchain consensus algorithms may verify improper changes to prevent poison attacks. Researchers have created the Multi-Krum defense, which checks local model changes using a Byzantine-tolerant aggregation approach. To protect players against poison assaults, a unique selection method based on a scoring system has been developed. An alternative consensus technique based on proof of verification has also been developed. Preset miners can remove hazardous and deadly model parameters by ensuring accurate model updates. The validity of is decided by the validator.

The findings show that while the use of blockchain technology eliminates most of the difficulties associated with conventional FL, certain drawbacks remain.

While blockchain consensus algorithms have certain security concerns, several researchers have used them to prevent poison attacks. Misbehaving miners use most of the system's computational power in PoW, which can result in blockchain splits.

Furthermore, PBFT has the potential to accept roughly one-third of the error nodes. As a result, various research has developed novel consensus algorithms such as PoK, PoQ, PoF, and so on. They can lower power usage but need to increase consensus mechanism security. The findings show that while blockchain technology eliminates most of the difficulties associated with conventional FL, there are still certain drawbacks.

Several research employed blockchain consensus algorithms to avoid poison attacks. However, blockchain consensus algorithms have several security problems. Misbehaving miners use most of the system's computational power in PoW, which can result in blockchain splits.

VI. BC- FL Architecture

For the completely decentralized, privacy-protected FL system, the Blockchain primarily serves as a central database. As a result, the primary goal is to recompense clients based on their contributions' value while protecting the underlying dataset's privacy and avoiding malicious assaults.

A platform must be used to support the Blockchain layer's functionality in BCFL to operate. Different Blockchain platforms have various characteristics. For instance, consortium networks offer robust security, public chains provide dependable performance, and private chains offer more customization options. Four platforms, Ethereum, Hyperledger Fabric, EOS, and Custom Blockchain, are

Ethereum: Ethereum is a public blockchain platform known for its smart contract functionality and wide adoption in various decentralized applications (DApps). It offers robust security and transparency due to its decentralized nature, making it suitable for BCFL projects where openness and accessibility are paramount. Ethereum's large developer community and ecosystem also provide extensive support for building and deploying FL solutions.

Hyperledger Fabric: Hyperledger Fabric is a permissioned blockchain framework designed for enterprise use cases. It offers advanced privacy features, scalability, and modular architecture, making it well-suited for BCFL projects that require fine-grained control over access permissions and data privacy. Hyperledger Fabric's focus on consortium networks aligns with the collaborative nature of federated learning, where multiple stakeholders need to securely share and process data.

EOS: EOS is a public blockchain platform known for its high throughput and low latency, making it suitable for applications requiring fast transaction processing. In BCFL projects, EOS may be preferred for real-time or latency-sensitive applications where quick model updates are essential. Its delegated proof-of-stake (DPoS) consensus mechanism enables efficient block production and transaction validation, enhancing the overall performance of FL systems.

Custom Blockchain: Custom Blockchain refers to bespoke blockchain solutions tailored to specific project requirements. While less standardized compared to established platforms like Ethereum or Hyperledger Fabric, custom blockchains offer flexibility and customization options tailored precisely to the needs of BCFL projects. These custom solutions may incorporate unique consensus mechanisms, privacy enhancements, or interoperability features optimized for federated learning environments.

These are predominantly used by the current BCFL, according to a thorough assessment of the literature.

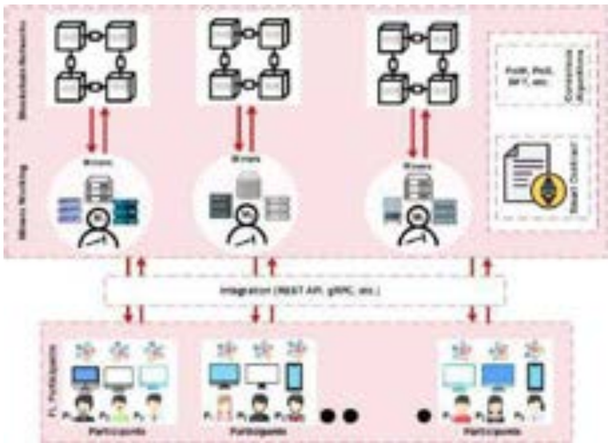


Fig. 6. FC-BL Architecture

Players in federated learning: Just like in a typical FL environment, participants take on the roles of entities or devices. Participating FL members train models and provide local model updates for the following stage of verification and aggregation. All FL system participants are first given access to the core model. Then, FL participants update local models using their unprocessed datasets. Direct communication occurs between miners and FL members.

The Blockchain and FL integration acts as a go-between for FL players and the blockchain. Users linked to the Hyperledger Fabric blockchain using the REST-API to record and reward gradient uploads. The API also enables data transmission between FL clients and the Ethereum blockchain network by utilizing Google's remote procedure calls.

If the mining program is easily downloadable, miners can utilize their personal computers, standby servers, or cloud-based nodes. The FL participants are now sending local model modifications to the miners. Each FL participant/data holder is immediately connected to the miner, enabling continuous communication. Miners are responsible for obtaining local model updates from FL devices or participants. Furthermore, aggregation is carried out utilizing the consensus mechanism, and a block is then put into the blockchain.

The blockchain system's Smart Contract (SC) component opens up new possibilities for decentralized applications by dynamically executing program logic in response to preset conditions. Before starting the FL model training process, participating FL clients will agree that all criteria are clear and immutable. SC also gives customers the option to establish agreements without the help of a reliable third party. Researchers used smart contact in several ways, including signing up participants, organizing model training, compiling local model updates, rating participant contributions, and providing incentives. A smart contract assignment between FL players and miners is shown in Figure 5.

Transaction verification is based on the consensus algorithm, which is the blockchain network's backbone. On a blockchain network, everyone agrees on how to generate, verify, and approve new blocks. When miners reach a consensus mechanism, such as Proof of Work (PoW), Proof of Stake (PoS), or Byzantine Fault Tolerance (BFT), a new block is added to the blockchain. When blockchain technology and federated learning combine, the system becomes more adaptable. FL participants will enter a new FL training phase, and miners will learn how to create a completely converged global model using a consensus technique. Once the consensus method is completed, the block is added to the blockchain network.

BCFL ARCHITECTURE AND COMPONENTS

Finally, the freshly produced and validated blocks are uploaded. The FL model strategy is repeated until the desired learning rate is achieved. Following that, FL customers or other parties may request the global model for their usage. Finally, miners may download global models, which FL players can purchase.

1. Federated Learning Participants:
- Operate similarly to a typical federated learning environment.
- Entities or devices participate by training models, providing local updates, and engaging in verification and aggregation processes.
2. Blockchain and FL Integration:
- Acts as an intermediary between FL participants and the blockchain.
- Hyperledger Fabric and Ethereum are utilized for integration, facilitating the recording and rewarding of gradient uploads through REST-API, and Google's remote procedure calls for data transmission.
3. Miners' Role:
- Miners, utilizing personal computers, standby servers, or cloud-based nodes, receive local model updates from FL participants.
- Continuous communication is established, and aggregation is performed using a consensus mechanism.
- Aggregated updates are then added to the blockchain as a new block.
4. Smart Contracts:
- Smart Contracts (SC) play a crucial role in automating processes in response to preset conditions.
- SC facilitates agreements among FL participants and miners without relying on a third party.
- Used for participant registration, organizing model training, compiling local model updates, rating contributions, and providing incentives.
5. Consensus Mechanism:
- Transaction verification is based on a consensus algorithm, such as Proof of Work (PoW), Proof of Stake (PoS), or Byzantine Fault Tolerance (BFT).
- The consensus mechanism ensures agreement on generating, verifying, and approving new blocks, contributing to the security and integrity of the BC-FL system.
6. Blockchain Network:
- Newly produced and validated blocks are uploaded to the blockchain network.
- FL model training phases continue until the desired learning rate is achieved.

Global models, created through consensus techniques, are made available for FL customers or other parties to request and use.

VII. FEDERATED LEARNING AND BLOCKCHAIN

3.1 Overview of Federated Learning

Federated Learning, as proposed by McMahan et al. (2017), serves as a centralized training method with a primary focus on protecting user privacy through unique data distribution properties. FL participants, representing consumers (FL users), contribute training data from their private local datasets to the FL server, enabling the creation of a global model. The FL server consolidates these local model updates, forming a globally converged model that users can download.

Equation (1) defines the FL server's communication with selected FL participants, conveying the initial model changes (mt) for each training round (t). Following this, FL clients utilize an initial model to train local data, updating their local model. These local model changes are then submitted to the central FL server, which compiles them. The process iterates until Gt represents the global model as of the tenth iteration, and Gt+1 signifies a globally converged model.

Performing federated learning involves a series of steps to train a global model collaboratively across multiple devices while preserving data privacy. The process typically follows these general steps:

1. Initialization: The FL server initializes the global model and communicates it to selected FL participants.
2. Training Rounds: Each training round begins with the FL server sending the current global model (mt) to selected FL participants. Participants then use their local datasets to train the model locally, resulting in local model updates.
3. Model Aggregation: FL participants send their local model updates back to the FL server, which aggregates them to form a globally converged model. This

aggregation process typically employs algorithms like Federated Average (FedAvg) or its variants to combine local updates while addressing system heterogeneity.
4. Iteration: Steps 2 and 3 are iterated for multiple rounds (t) until the global model (Gt) converges to a satisfactory level of performance.

Evaluation: The final converged model (Gt+1) is evaluated for performance on validation data or testing datasets to assess its effectiveness.

$$G_{t+1} = G_t + \frac{1}{m_t}\sum_{k=1}^{m_t} l_t^k,$$

The foundation of FL lies in the aggregation algorithm, specifically the vanilla Federated Average (FedAvg), allowing the accumulation of local model updates. To address system heterogeneity, FedProx, a generalization of FedAvg, is introduced. Various improvements to the aggregation approach, such as Federated Matched Averaging (FedMa) and Federated Optimisation (FedOpt), aim to overcome specific FL challenges. Additionally, three types of FL are identified, including Horizontal Federated Learning (HFL).

Real-world Applications of Blockchain-based Federated Learning (BC-FL):
1. Healthcare Data Sharing:
- Scenario: Collaborative improvement of a global disease prediction model without sharing sensitive patient data.
- BC-FL Implementation: BC-FL enables secure participation of multiple hospitals, keeping data private and sharing local model updates securely through blockchain. Smart contracts ensure fair compensation.

2. Finance and Fraud Detection:
- Scenario: Collaborative enhancement of fraud detection models without revealing individual transaction details.
- BC-FL Implementation: Blockchain ensures the integrity and transparency of financial transactions. FL participants from different banks securely share model updates through the blockchain, with smart contracts automating reward distribution.

3. Smart Grid Optimization:
- Scenario: Collective improvement of predictive models for energy consumption to optimize smart grids.
- BC-FL Implementation: BC-FL allows decentralized collaboration among energy providers, securing local data on energy consumption patterns through blockchain. The consensus mechanism ensures accuracy and reliability.

4. Supply Chain Management:
- Scenario: Optimization of inventory management and logistics in a supply chain without exposing proprietary data.
- BC-FL Implementation: Blockchain ensures transparency and traceability. Federated learning participants contribute local insights through blockchain, and smart contracts automate compensation based on contributions' value.

5. Autonomous Vehicles Collaboration:
- Scenario: Collaboration among autonomous vehicle manufacturers to improve traffic prediction models without sharing real-time location data.
- BC-FL Implementation: BC-FL allows secure sharing of model updates through blockchain, preserving the privacy of individual vehicle data. Smart contracts facilitate fair compensation.
6. Cross-Border Research Collaboration:
- Scenario: Researchers from different countries collaborate on a scientific project requiring analysis of diverse datasets without compromising data privacy.
- BC-FL Implementation: BC-FL ensures secure collaboration through blockchain, allowing researchers to contribute local model updates while maintaining privacy. Smart contracts automate reward distribution based on the significance of contributions.

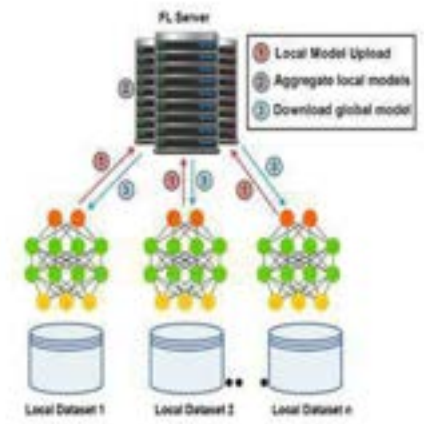


Fig. 7. Federated learning architecture

VIII. OVERVIEW OF BLOCKCHAIN

Blockchain technology has emerged as a revolutionary force in Federated Learning (FL), with the ability to solve key privacy, security, and cooperation issues. This section provides a detailed explanation of how Blockchain is being incorporated into the fabric of Federated Learning, altering the landscape of decentralized machine learning. At its foundation, Blockchain functions as a decentralized and distributed ledger, dramatically disrupting the previous paradigm of centralized control. In the context of Federated Learning, this invention provides a safe and transparent framework for collaborative model training over a network of different devices without compromising individual datasets' privacy.

Blockchain's design is based on a succession of blocks, each containing transactional data, a timestamp, and the previous block's cryptographic hash. This structure guarantees an immutable and chronological record, which protects transaction integrity. Applied to Federated Learning, these blocks encapsulate crucial information about model updates, preserving the lineage of collaborative efforts while maintaining data privacy. Timestamps and cryptographic hashes not only secure the transactions but also establish a secure and tamper-resistant link between blocks, fortifying the reliability of the transaction history.

The decentralized nature of Blockchain is well-aligned with the principles of Federated Learning. In traditional machine learning models, data is often centralized, posing significant privacy risks. Blockchain, through its peer-to-peer (P2P) network, redistributes control and information, allowing participants, often individual devices or nodes, to engage in collaborative model training. Each member of the blockchain network maintains a synchronized copy of the transaction history. This increases transparency and reduces the dangers associated with a single point of failure.

Federated Learning categorizes blockchain into three types: public or permissionless blockchain, private or commissioned blockchain, and consortium blockchain. In the case of a public blockchain, no single entity dominates, and participants can join or exit the network at will, fostering decentralization and transparency. In contrast, private blockchains have a centralized structure in which a single party has power over transaction validation, increasing efficiency and speed. Consortium blockchains strike a balance, granting specific privileges to selected participants and necessitating consensus before implementing transactions, thus combining aspects of both public and private configurations.

Central to Blockchain's functionality are consensus algorithms, determining the rules governing the acceptance of transactions. The choice of consensus algorithm is crucial in the context of Federated Learning, impacting the trade-off between security, efficiency, and scalability. Algorithms like Proof of Work (PoW) and Proof of Stake (PoS) play pivotal roles in securing the integrity of the Blockchain network.

Smart contracts, autonomous and self-executing pieces of code deployed on the Blockchain, further enrich the landscape of Federated Learning. In the context of FL, smart contracts serve as digital agreements between participating nodes, executing predefined functions to facilitate seamless and automated interactions. These contracts, distributed across all nodes on the Blockchain, ensure transaction traceability and enforce irreversibility, contributing to the overall integrity of the collaborative process.

Achieving security in blockchain-based federated learning involves implementing several key measures. Firstly, encryption techniques such as homomorphic encryption ensure that sensitive data remains encrypted during transmission and computation, thus safeguarding it from unauthorized access. Secondly, decentralized consensus mechanisms, inherent to blockchain technology, validate transactions and updates to the federated learning model, ensuring integrity and preventing tampering. Additionally, access controls and smart contracts enforce permissions, dictating who can participate in model training and access the federated learning system, thereby mitigating potential security threats. Finally, regular audits and monitoring of the blockchain network help identify and address security vulnerabilities, ensuring the overall robustness of the system. Through these integrated security measures, blockchain-based federated learning systems can provide a high level of security and trustworthiness for collaborative model training across distributed devices

In conclusion, the integration of Blockchain in Federated Learning represents a paradigm shift, offering innovative solutions to the challenges posed by traditional machine learning approaches. This overview elucidates the core principles of Blockchain technology and its transformative potential in revolutionizing Federated Learning. As the intersection of Blockchain and FL continues to evolve, this synergistic relationship holds the promise of fostering a secure, transparent, and collaborative future for decentralized machine learning.

IX. Blockchain-based Federated Learning (BFL)

When performing a systematic review of Blockchain-based Federated Learning (BFL), a thorough and precise approach is used to examine the current body of research, providing significant insights into the junction of these two cutting-edge technologies. The research selection criteria will be purposefully designed to prioritise papers that directly address the combination of Blockchain and Federated Learning, resulting in a targeted and relevant evaluation. Inclusion criteria will include works published in respectable peer-reviewed journals, conference proceedings, and recognised research databases. To capture the latest developments, the review will concentrate on literature within the last ten years, allowing for an up-to-date exploration of the field.

A structured data extraction form will be used to obtain relevant information from the selected research systematically. This includes details on authors, methodologies, key findings, Blockchain technologies utilized, and implications for Federated Learning. The collected data will undergo thematic analysis, allowing for the categorization and synthesis of findings. The focus will be on critical aspects such as integration challenges, benefits, and the impact on user perspectives, providing a holistic perspective.

Quality assessment will be a critical component of this systematic review, ensuring the reliability and robustness of the included studies. The evaluation will consider the methodologies employed in the research, offering a qualitative lens to interpret the findings and draw meaningful conclusions. Adhering to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, the reporting phase will present a clear and transparent depiction of the study selection process through a flow diagram, and a narrative synthesis will be crafted to communicate the key findings cohesively.

This meticulously crafted methodology emphasizes systematic search strategies, rigorous study selection, and comprehensive data extraction. The goal is to give a complete assessment of the environment of Blockchain-based Federated Learning. Through this approach, the systematic review contributes substantively to the understanding of the integration of Blockchain in the evolving field of Federated Learning, shedding light on challenges, benefits, and potential avenues for future research.

X. CHALLENGES AND OPPORTUNITY

The intersection of Blockchain technology with Federated Learning (FL) creates a complex interplay of obstacles and possibilities that shapes the trajectory of this novel paradigm. Scalability is one of the most challenging obstacles to overcome when integrating Blockchain with Federated Learning. While effective in ensuring transparency and immutability, the classic Blockchain structure faces challenges when the number of participants and transactions grows. The sheer amount of computations necessary for consensus algorithms such as Proof of Work (PoW) or Proof of Stake (PoS) creates a scalability barrier that may impede the smooth execution of collaborative machine learning projects. Striking a balance between maintaining decentralization, a cornerstone of Blockchain, and meeting the scalability demands of Federated Learning remains a persistent challenge.

The inherent commitment to privacy in Federated Learning introduces a layer of intricacy when coalescing with the transparent nature of Blockchain. Privacy-preserving techniques, fundamental to the ethos of Federated Learning, demand meticulous attention when navigating the decentralized Blockchain ecosystem. As the assurance of individual data privacy clashes with the necessity for a transparent and auditable transaction history, reconciling these seemingly opposing principles emerges as a nuanced challenge. This intricate dance requires innovative solutions to preserve the privacy of individual contributions while upholding the integrity and transparency promised by Blockchain.

Another challenge emanates from the heterogeneity of devices within a federated environment. The diverse range of computational capabilities and network bandwidth across participating nodes introduces a potential stumbling block. Disparities in node processing speeds can impede collaborative learning efforts' synchronization and efficiency. Addressing this challenge necessitates designing mechanisms that accommodate the variability in device capabilities, ensuring an inclusive and equitable participation framework for all nodes.

Despite these challenges, the fusion of Blockchain and Federated Learning unveils promising opportunities that could reshape the landscape of collaborative machine learning. Blockchain's immutable and transparent nature provides a foundation for enhanced accountability and traceability in Federated Learning transactions. This feature addresses the challenge of maintaining a trustworthy ecosystem and builds a foundation of trust among participants. The decentralized design of Blockchain reduces the risks associated with a single point of failure, resulting in a more resilient and secure environment for collaborative learning.

Moreover, the integration of smart contracts within the Blockchain framework introduces a layer of automation to the coordination and execution of tasks in Federated Learning. Smart contracts can streamline and automate the negotiation of terms and conditions among participating nodes, thereby enhancing the overall efficiency of collaborative learning initiatives. This automation holds the potential to minimize manual intervention, reducing friction in the collaborative process and fostering a more seamless and efficient workflow.

An additional opportunity arises with the prospect of tokenization within Blockchain-based Federated Learning. Implementing token-based incentive mechanisms can stimulate active participation and contribution from nodes. The introduction of tokens aligns with the principles of decentralized incentive structures, creating a dynamic and engaging collaborative environment. Token-based rewards systems have the potential to incentivize participation, fostering a sense of ownership and motivation among nodes within the federated ecosystem.

In conclusion, the challenges and opportunities entwined in Blockchain-based Federated Learning delineate a dynamic and intricate landscape. While scalability, consensus mechanisms, and privacy concerns present formidable challenges, the inherent advantages of Blockchain offer novel solutions to reshape collaborative machine learning. Navigating these hurdles and capitalizing on possibilities will help define the future of Blockchain-based Federated Learning, opening the path for a safe, transparent, and efficient collaborative machine learning paradigm.

XI. FUTURE DIRECTIONS

As the combination of Blockchain technology with Federated Learning (FL) continues to mature, a fascinating array of future paths emerges, opening the door for significant breakthroughs in collaborative machine learning. One prominent avenue for exploration revolves around refining consensus mechanisms tailored specifically for the nuances of Federated Learning within a Blockchain environment. While traditional Proof of Work (PoW) and Proof of Stake (PoS) procedures have shown to be effective in typical Blockchain contexts, they may not be ideal for FL's dynamic and decentralized nature. Future research should delve into developing consensus algorithms that strike a delicate balance between ensuring data integrity and minimizing computational overhead, thus addressing one of the primary challenges currently faced. Furthermore, privacy-preserving strategies in the context of Blockchain-based Federated Learning need for improved methodology. Future research endeavors could focus on pioneering cryptographic solutions or innovative privacy-preserving algorithms that bolster the seamless integration of individual privacy with the transparent nature of Blockchain. This would involve not only protecting sensitive information during collaborative machine learning but also ensuring that the integrity and traceability promises of Blockchain are upheld. Exploring cutting-edge cryptographic approaches such as homomorphic encryption or secure multi-party computation might be the key to unlocking new dimensions in privacy and security. The heterogeneity of devices participating in federated environments presents a rich area for future research. Tailoring Federated Learning algorithms to accommodate the diversity in computational capabilities and network bandwidth across nodes becomes imperative. Future directions may involve the development of adaptive and dynamic federated learning strategies that adjust according to the capabilities of individual nodes. Ensuring an inclusive and equitable participation framework for all nodes, irrespective of their computational prowess, will be essential for the continued success and scalability of Blockchain-based Federated Learning.

Furthermore, the concept of tokenization within Blockchain-based Federated Learning introduces novel opportunities. Future research may delve into the design and implementation of sophisticated token-based incentive mechanisms. Exploring the dynamics of token economies within federated ecosystems could yield insights into how incentivizing active participation and collaboration enhances the overall efficiency and engagement. This involves not only the creation of tokens as rewards but also the exploration of governance models and decentralized decision-making frameworks fueled by these tokens.

In the realm of practical applications, future research could focus on developing use case-specific frameworks for Blockchain-based Federated Learning. Tailoring the integration to suit the requirements of specific industries, such as healthcare, finance, or manufacturing, would involve understanding the unique challenges and opportunities within these sectors. This industry-specific approach could pave the way for more targeted and impactful implementations of federated learning in real-world scenarios, aligning with the broader trend of responsible and contextual AI deployment.

In conclusion, the future directions of Blockchain-based Federated Learning are rich with possibilities, beckoning researchers to delve into innovative solutions and advancements. Addressing the intricacies of consensus mechanisms, refining privacy-preserving techniques, accommodating device heterogeneity, integrating explainability, exploring token economies, and tailoring frameworks to specific industries all contribute to the burgeoning landscape of collaborative machine learning. As research unfolds, these future directions hold the potential to reshape the trajectory of Blockchain-based Federated Learning, propelling it into a realm of increased efficiency, transparency, and real-world applicability.

XII. CONCLUSION

In conclusion, this systematic review has unveiled the complexities between Blockchain and Federated Learning (FL), shedding light on the transformative potential and challenges of their integration. The examination of the literature points to a scenario in which Blockchain's transparency and decentralisation might successfully address FL's inherent issues with data security and privacy. While challenges such as scalability and consensus mechanisms persist, the amalgamation of these technologies presents a groundbreaking avenue for collaborative machine learning. The transparent and immutable nature of Blockchain addresses trust and accountability concerns in FL, fostering a more secure environment for shared data. The automation introduced through smart contracts and tokenization mechanisms not only streamlines collaboration but also incentivizes active participation. Looking forward, it is imperative for researchers to focus on refining consensus mechanisms and privacy-preserving techniques, balancing the decentralized principles of Blockchain with the scalability demands of FL. In order to fully realise the potential of Blockchain-based Federated Learning and provide the groundwork for a future in which collaborative machine learning is founded on the values of decentralisation, security, and transparency, this comprehensive evaluation serves as a guide.

REFERENCES

[1] M. Shayan, C. Fung, C. J. M. Yoon, and I. Beschastnikh, "Biscotti: A Blockchain System for Private and Secure Federated Learning," *IEEE Trans. Parallel Distrib. Syst.*, vol. 32, no. 7, pp. 1513–1525, 2021, doi: 10.1109/TPDS.2020.3044223.

[2] D. Li, Z. Luo, and B. Cao, "Blockchain-based federated learning methodologies in smart environments," *Cluster Comput.*, vol. 25, no. 4, pp. 2585–2599, 2022, doi: 10.1007/s10586-021-03424-y.

[3] Y. Chang, C. Fang, and W. Sun, "A blockchain-based federated learning method for smart healthcare," *Comput. Intell. Neurosci.*, vol. 2021, 2021, doi: 10.1155/2021/4376418.

[4] D. Li *et al.*, "Blockchain for federated learning toward secure distributed machine learning systems: a systemic survey," *Soft Comput.*, vol. 26, no. 9, pp. 4423–4440, 2022, doi: 10.1007/s00500-021-06496-5.

[5] D. Hou, J. Zhang, K. L. Man, J. Ma, and Z. Peng, "A Systematic Literature Review of Blockchain-based Federated Learning: Architectures, Applications and Issues," *2021 2nd Inf. Commun. Technol. Conf. ICTC 2021*, no. May, pp. 302–307, 2021, doi: 10.1109/ICTC51749.2021.9441499.

[6] M. Billah, S. T. Mehedi, A. Anwar, Z. Rahman, and R. Islam, "A Systematic Literature Review on Blockchain Enabled Federated Learning Framework for Internet of Vehicles," vol. 4, pp. 1–21, 2022, [Online]. Available: http://arxiv.org/abs/2203.05192.

[7] L. Li, Y. Fan, M. Tse, and K.-Y. Lin, "A review of applications in federated learning," Computers & Industrial Engineering, p. 106854, 2020.

[8] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, "Federated learning: Challenges, methods, and future directions," IEEE Signal Processing Magazine, vol. 37, no. 3, pp. 50–60, 2020.

[9] P. Kumar, J. Hyuk, and K. Cho, "Blockchain and federated learning-based distributed computing defence framework for sustainable society," *Sustain. Cities Soc.*, vol. 59, no. March, p. 102220, 2020, doi: 10.1016/j.scs.2020.10222.

[10] W. Zhang, Q. Lu, Q. Yu, Z. Li, Y. Liu, S. K. Lo, S. Chen, X. Xu, and L. Zhu, "Blockchain-based federated learning for device failure detection in industrial iot," IEEE Internet of Things Journal, pp. 1–1, 2020.

[11] Y. Lu, X. Huang, Y. Dai, S. Maharjan, and Y. Zhang, "Blockchain and federated learning for privacy-preserved data sharing in industrial iot," IEEE Transactions on Industrial Informatics, vol. 16, no. 6, pp. 4177– 4186, 2020.

[12] L. Cui, X. Su, Z. Ming, Z. Chen, S. Yang, Y. Zhou, and W. Xiao, "Creat: Blockchain-assisted compression algorithm of federated learning for content caching in edge computing," IEEE Internet of Things Journal, pp. 1–1, 2020.

[13] K. Toyoda, J. Zhao, A. N. S. Zhang, and P. T. Mathiopoulos, "Blockchain-enabled federated learning with mechanism design," IEEE Access, vol. 8, pp. 219744–219756, 2020.

[14] Y. Qu, L. Gao, T. H. Luan, Y. Xiang, S. Yu, B. Li, and G. Zheng, "Decentralized privacy using blockchain-enabled federated learning

in fog computing," IEEE Internet of Things Journal, vol. 7, no. 6, pp. 5171– 5183, 2020.

[15] I. Martinez, S. Francis, and A. S. Hafid, "Record and reward federated learning contributions with blockchain," in 2019 International Conference on Cyber-Enabled Distributed Computing and Knowledge Discovery (CyberC), pp. 50–57, 2019.

[16] K. Toyoda and A. N. Zhang, "Mechanism design for an incentiveaware blockchain-enabled federated learning platform," in 2019 IEEE International Conference on Big Data (Big Data), pp. 395–403, 2019.

[17] Qu, Youyang & Uddin, Md Palash & Gan, Chenquan & Xiang, Yong & Gao, Longxiang & Yearwood, John. (2022). Blockchain-Enabled Federated Learning: A Survey. ACM Computing Surveys. 55. 10.1145/3524104.

[18] Wang Zhibo, Song Mengkai, Zhang Zhifei, Song Yang, Wang Qian, and Qi Hairong. 2019. Beyond inferring class representatives: User-level privacy leakage from federated learning. In IEEE Conference on Computer Communications (IEEE INFOCOM'19). IEEE, 2512–2520

[19] EU. 2006. Regulation (EU) 2016/679 of the european parliament and of the council. https://op.europa.eu/en/publication-detail/-/publication/3e485e15-11bd-11e6-ba9a-01aa75ed71a1/.

[20] Yuan Jiawei and Yu Shucheng. 2014. Privacy preserving back-propagation neural network learning made practical with cloud computing. IEEE Transactions on Parallel and Distributed Systems 25, 1 (2014), 212–221.

[21] McMahan H. Brendan, Moore Eider, Ramage Daniel, Hampson Seth, and Arcas Blaise Agüera y. 2017. Communication-efficient learning of deep networks from decentralized data. In 20th International Conference on Artificial Intelligence and Statistics (AISTATS'17). 1273–1282.

[22] Roh Yuji, Heo Geon, and Whang Steven Euijong. 2019. A survey on data collection for machine learning: A big data-AI integration perspective. IEEE Transactions on Knowledge and Data Engineering 33 (2019), 1328–1347.

[23] Jordan Michael I. and Mitchell Tom M.. 2015. Machine learning: Trends, perspectives, and prospects. Science 349, 6245 (2015), 255–260.

[24] Konecny Jakub, McMahan H. Brendan, and Richtarik Peter. 2016. Federated optimization: Distributed machine learning for on-device intelligence. Computing Research Repository (CoRR) (2016), 38.

[25] Abadi Martin, Chu Andy, Goodfellow Ian, McMahan H. Brendan, Mironov Ilya, Talwar Kunal, and Zhang Li. 2016. Deep learning with differential privacy. In ACM SIGSAC Conference on Computer and Communications Security (CCS). ACM, Vienna, Austria, 308–318.

[26] Aloqaily Moayad, Ridhawi Ismaeel Al, and Guizani Mohsen. 2022. Energy-aware blockchain and federated learning-supported vehicular networks. IEEE Trans. Intell. Transp. Syst. 23, 11 (2022), 22641–22652. DOI: DOI: https://doi.org/10.1109/TITS.2021.3103645

[27] Sarwate Anand D. and Chaudhuri Kamalika. 2013. Signal processing and machine learning with differential privacy: Algorithms and challenges for continuous data. IEEE Signal Processing Magazine 30, 5 (Sept. 2013), 86–94. DOI:

[28] Abadi Martin, Chu Andy, Goodfellow Ian, McMahan H. Brendan, Mironov Ilya, Talwar Kunal, and Zhang Li. 2016. Deep learning with differential privacy. In ACM SIGSAC Conference on Computer and Communications Security (CCS '16). 308–318. DOI:

[29] O. El Rifai, M. Biotteau, X. de Boissezon, I. Megdiche, F. Ravat, andO. Teste, "Blockchain-based federated learning in medicine," in ArtificialIntelligence in Medicine (M. Michalowski and R. Moskovitch, eds.),(Cham), pp. 214–224, Springer International Publishing, 2020.

[30] S. M. Halim, L. Khan, and B. Thuraisingham, "Next - location predictionusing federated learning on a blockchain," in 2020 IEEE SecondInternational Conference on Cognitive Machine Intelligence (CogMI),pp. 244–250, 2020.

GPS based IoT Module for Vehicle Safety in Epileptic Seizure Detection and Alcohol Monitoring

Sushma S.Jagtap
Assistant Professor(SS)
Department of Electronics and Communication Engineering
Rajalakshmi Engineering College
Chennai,India
sushmasjagtap@rajalakshmi.edu.in

Saranya G
Assistant Professor(SS)
Department of Electronics and Communication Engineering
Rajalakshmi Engineering College
Chennai,India
saranya.g@rajalakshmi.edu.in

Dr Sathish M
Professor
Department of Electronics and Communication Engineering
Rajalakshmi Engineering College
Chennai,India
sathish.m@rajalakshmi.edu.in

Noor Jesima G
UG Student
Department of Electronics and Communication Engineering
Rajalakshmi Engineering College
Chennai,India
200801144@rajalakshmi.edu.in

Priyadharshini S
UG Student
Department of Electronics and Communication Engineering
Rajalakshmi Engineering College
Chennai,India
200801156@rajalakshmi.edu.in

Purnima
UG Student
Department of Electronics and Communication Engineering
Rajalakshmi Engineering College
Chennai,India
200801158@rajalakshmi.edu.in

***Abstract*—This research presents a comprehensive Internet of Things-enabled safety system designed for individuals with epilepsy. The system integrates epileptic seizure detection, alcohol monitoring, and vehicle engine control mechanisms to enhance the overall safety of individuals prone to seizures. Utilizing MQ3 and EEG sensors, the system continuously monitors physiological indicators to detect and analyse epileptic events. Additionally, an alcohol detection module ensures that individuals under the influence are identified. The incorporation of GPS technology enables real-time location tracking, crucial for prompt intervention during seizures. The system's IoT architecture facilitates seamless communication with cloud platforms, allowing for data storage, analysis, and remote monitoring. Furthermore, in instances of detected alcohol consumption, the system employs relay control to prevent engine start up, minimizing the risk of impaired driving. This integrated solution aims to provide a robust safety net for individuals with epilepsy, promoting independence while prioritizing their well-being in various scenarios.**

Keywords— Internet of Things; Global Positioning System; Seizure; MQ3; Electroencephalogram.

I. INTRODUCTION

Epilepsy presents substantial problems to the day-today lives of those who are affected by it, with the unexpected nature of seizures offering inherent hazards. In the management of epilepsy, one of the most important aspects is making sure that persons who are prone to seizures are safe. This is especially important in situations where impairments, such as drinking alcohol, make these risks even more substantial. The findings of this study present a novel safety system that is based on the Internet of Things (IoT). This system combines the detection of epileptic seizures, the monitoring of alcohol consumption, and the management of car engines to produce a holistic solution for people who suffer from epilepsy. By utilizing cutting-edge technology, the primary objective of this integrated system is to improve the safety and freedom of individuals who are diagnosed with epilepsy.

The system is able to enable real-time detection and analysis of epileptic occurrences because it makes use of sophisticated sensors that are able to monitor physiological indications that are connected with seizures. This makes it possible to intervene at the appropriate time and improves the overall management of hazards associated with epilepsy. One of the most important aspects of this safety system is the Internet of Things (IoT), which enables cloud platforms to communicate with one another in a smooth manner. This connectivity makes it possible to perform remote monitoring, data storage, and analysis.It aims to provide a full safety net to the individuals who has absent seizure that last for less than 15 seconds.

II. RELATED WORKS

[1] Ugochi A. Okengwu, et.al, 2022, “Design and Implementation of In-Vehicle Alcohol Detection and Speed Control System”, Research Gate, studied to address the issue of road accidents caused by driver’s

alcohol consumption in Nigeria. Despite existing measures, accidents persist due to speeding, rash driving, and drunk driving. The researchers designed an In-Vehicle Alcohol Detection and Speed Control System using Arduino Nano. The system met engine start standards, and if the driver's BAC exceeds 0.5 mg/mL, authorities are notified via GSM. Experimental results show rapid alcohol detection and prolonged operation of the system.

[2] Anuja Khodaskar, et.al, 2011, "IOT based DD Detection System", IJCRT, implemented an IoT based Drunk and Drive Detection (DDD) system aimed at early identification of alcohol-impaired drivers. Drunk driving is a significant contributor to traffic accidents globally, leading to numerous fatalities. The proposed system monitors alcohol levels using a sensor and, if the driver is found to be intoxicated, instantly halts the vehicle's ignition through the processor. Additionally, the system utilizes GPS and GSM technology to relay the vehicle's location to pre-programmed contacts. The implementation of this system aims to reduce accidents caused by alcohol consumption, contributing to improved road safety and advancements in the automotive industry.

[3] Gorlagunta Latha, et.al, 2023, "IOT Based Automatic Speed Control And In-Vehicle Alcohol Detection System Forsmart Vehicles", Academia, introduced an IoT-based Automatic Speed Control and In-Vehicle Alcohol Detection System for smart vehicles. Utilizing the Arduino UNO microcontroller board, the system integrates an alcohol sensor, control switch, RF Module, Wifi Module, and speed control modules. The alcohol sensor continuously monitors the driver's blood alcohol level through breath analysis, and in the event of alcohol detection, the system automatically shuts off the vehicle's ignition. This technology presents a promising approach to reducing accidents caused by impaired driving in the realm of intelligent smart vehicle systems.

[4] Chaitanya V Shembekar, et.al, 2020, "IoT based Alcohol and Driver Drowsiness Detection and Prevention System", Research Gate, implemented a system which utilizes visual data and artificial intelligence to detect drowsiness by analysing the driver's face and eyes, measuring the level of eye closure (PERCLOS) with Softmax for neural network processing. Additionally, the system incorporates alcohol and heart rate detection to assess the driver's condition. The aim is to address fatigue-related accidents, particularly in drivers of large vehicles with extended driving periods and challenging conditions.

[5] Diwakaran.S, et.al, 2021, "Android Based Alcohol Detection System Using an IOT", IJCRT, developed a road transportation safety system for the those living in cities to enjoy the availability of Alcohol detection system. It works with the integration of IoT. I t traces the content of alcohol and denies driving.It provides awareness to the public surrounded by making loud noise.

[6] Arun Francis, et.al, 2019, "Health monitoring with alcohol detection and ignition control system using IOT", Research Gate, proposed a solution which involves monitoring alcohol consumption and heart rate. If the driver is identified as drunk, the vehicle's ignition system is disabled, preventing movement and potential accidents. Additionally, abnormal changes in heart rate trigger the transmission of the driver's status to friends via IoT. Due to practical constraints, the system is implemented using a DC motor, with Node MCU serving as the controller.

[7] Stanley Uzairue, et.al, 2018, "IoT-Enabled Alcohol Detection System for Road Transportation Safety in Smart City", core.ac.uk, discussed the development of an IoT-enabled alcohol detection system for road transportation safety in smart cities. It utilizes a microcontroller to monitor two Blood Alcohol Content (BAC) thresholds. Upon reaching the first threshold, the system transmits the driver's BAC level and vehicle coordinates to a central monitoring unit. Upon reaching the second threshold, the system shuts down the vehicle's engine, triggers an alarm, and activates a warning light. The prototype includes a DC motor as the vehicle's engine and a push button as its ignition system. The system's efficiency is tested, aiming to reduce drunk driving-related accidents in smart cities.

[8] Sobhana Jahan, et.al, 2023, "AI-Based Epileptic Seizure Detection and Prediction in Internet of Healthcare Things: A Systematic Review", IEEE, conducted a systematic review titled "AI-Based Epileptic Seizure Detection and Prediction in Internet of Healthcare Things" published in IEEE. The study focuses on epilepsy, a neurological condition affecting 50 million individuals globally, recognized as a hypersensitive disease.
Electroencephalography (EEG) is widely used for its stable and universal characteristics in detecting and predicting seizures. The researchers reviewed 56 articles from diverse academic databases, exploring EEG technology and its applications, including Machine Learning, Deep Learning, and the Internet of Things. The study provides a comprehensive analysis of classification algorithms' performance and discusses open issues in the field, offering insights for potential future research. The study focuses on epilepsy, a neurological condition

affecting 50 million individuals globally, recognized as a hypersensitive disease .

[9] Robertas Damaševičius, et.al, 2023, “From Sensors to Safety: Internet of Emergency Services (IoES) for Emergency Response and Disaster Management”, MDPI, proposed a system using the Internet of Emergency Services (IoES) and its impact on emergency response and disaster management. IoES involves integrating internet connected devices and systems to enhance real-time data collection and coordination among emergency services. The study emphasizes the role of sensors and IoT devices in providing timely information to responders. It discusses the potential benefits of IoES, such as improved speed and efficiency in emergency response, as well as enhanced public safety. However, the paper also highlights the need to address challenges and risks associated with IoES implementation to ensure responsible use. Overall, the paper provides a comprehensive understanding of IoES and its implications for emergency management.

[10] Rohith Yerrapalem, et.al, 2023, “The IoT Based Health Monitoring with Alcohol Detection and Control System”, IJECS, proposed a health monitoring system based of IoT with alcohol detection. It is designed to reduce the accidents caused due to drunken driving by including a Alcohol Sensor, Heart beat rate sensor, Global System for Mobile communication, Wi-Fi modem, are utilized. Once it reaches above the threshold level,the engine is made to stop.

III. EXISTING SYSTEM

An integrated approach that includes alcohol monitoring, real-time intervention, and seizure detection is currently missing from the landscape of care for individuals with epilepsy. There isn't a complete system out there that incorporates all of the features of the proposed system, even though different technologies handle different parts of it. Wearable devices are a common component of current seizure detection systems, which aim to track the physiological changes that occur during seizures. One way to deal with the problem of drunk driving is by using alcohol detection technology, which includes breathalyzers and ignition interlock devices. Internet of Things (IoT) healthcare platforms and GPS tracking devices also aid in remote monitoring and data analysis. When it comes to situations where people with epilepsy are driving while under the influence of alcohol, there is a significant lack of comprehensive safety measures due to the lack of an integrated approach. To close this gap and to enable them to live an independent life, we have come up with an idea to create a unified system that safeguards people dealing with epilepsy by taking into account its ever-changing nature. Though previous systems focus only on detection of alcohol, there is no idea over how to implement it in driving for those with absent seizures.

IV. PROPOSED SYSTEM

The proposed system envisions a holistic and integrated IoT-based safety solution designed to address the multifaceted challenges faced by individuals with epilepsy. At its core, the system will deploy advanced sensors and algorithms for accurate and real-time detection of epileptic seizures, minimizing the shortcomings observed in existing technologies by observing the voltage which may vary from 10μV to 100μV in amplitude for a normal human being when measured from scalp. And for a epileptic patient. Simultaneously, a user-friendly and continuous alcohol monitoring system which can be seamlessly integrated into wearable devices is developed as a prototype, ensuring real-time insights into alcohol levels. The incorporation of precise GPS tracking technology will provide immediate location data, enhancing overall safety by facilitating prompt responses during seizure events. To ensure scalability, security, and efficient data management, the system will embrace a cloud-based IoT architecture, enabling seamless communication between components and facilitating remote monitoring, data analysis, and customization of parameters by healthcare professionals. A pivotal feature includes a relay control mechanism, preventing vehicle engines from starting upon detecting alcohol consumption, thereby curbing the dangers associated with impaired driving. The system will prioritize user customization, boasting an intuitive interface and customizable parameters to cater to individual profiles and medical conditions. Data security measures will be paramount, ensuring compliance with privacy standards and protecting sensitive health and location information. Through interdisciplinary collaboration, involving technology developers, healthcare professionals, and end-users, the proposed system aspires to offer a transformative solution that not only fills existing gaps but also prioritizes the safety, independence, and well-being of individuals with epilepsy in various daily-life scenarios. Additionally, the system is equipped with a module for monitoring alcohol intake, which allows it to identify instances of alcohol use in addition to seizure detection. The use of global positioning system technology guarantees continuous location tracking, which is an essential component for ensuring a rapid response to seizures and improving the individual's overall awareness of their surroundings. This system combines the detection of epileptic seizures, the monitoring of alcohol consumption, and the management of car engines to produce a holistic solution for people who suffer from epilepsy. By utilizing cutting-edge technology, the primary objective of this integrated system is to improve the safety and freedom of individuals who are diagnosed with epilepsy.

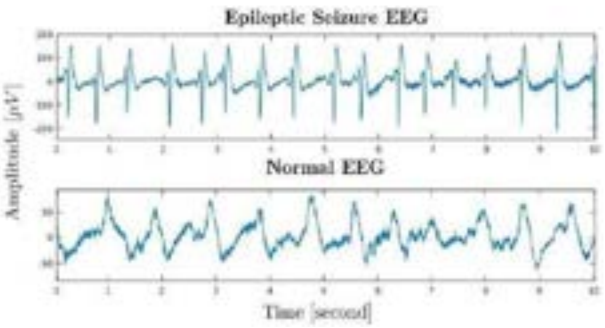


Fig 1: Difference of EEG wave

V. SYSTEM ARCHITECTURE

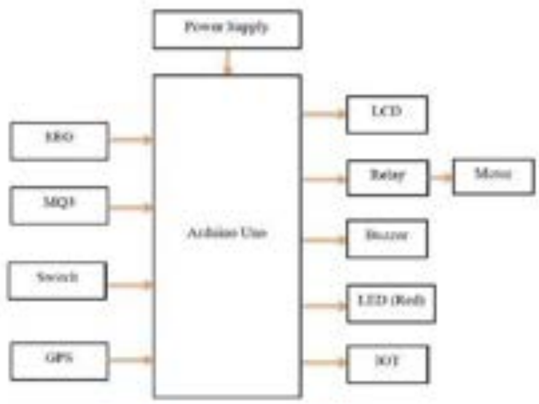


Fig 2:Block diagram of GPS based IOT module for Vehicle Safety in Epileptic Seizure Detection and Alcohol Monitoring

VI. METHODOLOGY

A. ESP8266

This module uses an ESP8266 microcontroller which has an inbuilt Wi-Fi connectivity to control the system through IOT and is programmable using the Arduino IDE, where the data linked to the latitude and longitude of the person involved in the accident is sent to be displayed on the LCD and to be stored in the webpage. The connection is established between IOT module and the Arduino.

Fig 3: ESP8266

B.GPS Module

This is the core component that receives signals from GPS satellites and processes them to determine the device's location. There are various types and models available in the market, such as those based on chipsets from companies like u-blox, MediaTek, and Broadcom. This indicates the Latitude and Longitude of the person undergoing seizure.

Fig 4:GPS Module

C. Alcohol sensor and EEG sensor

By using this technology, we can provide a faster and hands-free experience to the user. This whole system comes with an EEG sensor with three leads, each one to be placed on the scalp.Once the voltage beings to vary drastically, spikes in the EEG wave also varies.

Fig 5: MQ3 Alcohol sensor and EEG Sensor

MQ3 sensor is a gas sensor which is sensitive to alcohol vapours which is the target. It provides an analog output voltage proportional to concentration of alcohol percent in air. It detects the alcohol level of the person driving and if it exceeds the denoted threshold level, a LED light incorporated in the system would be blinking and the location is sent through IOT module. Instead of directly connecting a DC motor to stop the vehicle, it is interfaced with a relay of 5V for halting the vehicle on receiving warnings in the sense of blinking of LED. Table I shows the overview of the integrated features.

Table 1:Overview of features integrated

Feature	Epileptic Seizure Detection	Alcohol Monitoring
GPS Integration	Yes	Yes
Real-time Tracking	Yes	Yes
Seizure Detection Sensors	EEG sensor	-
Alcohol Monitoring Sensors	-	MQ3
Alerts/Notifications	Yes	Yes
Data storage	Yes	Yes
Indicator	Buzzer	LED

VII. RESULTS AND DISCUSSION

The use of MQ3 and EEG sensors has shown the potential for continuous monitoring of physiological signs associated with epileptic events. The ability to detect and analyse epileptic seizures in real time facilitates immediate help and intervention, improving the overall safety of people in these crucial moments. We strive to prove our worth through accurate, real-time measurements of operating systems expressed in millivolts. Normal brain voltage value for an individual is between 70 mV; if the attacks are distorted, the deviation from normal values is about 0.01 mV to 0.1 mV. In addition, with this prototype, we can analyse and detect scene variations and alert the cloud module. The ability of the alcohol detection module to correctly identify intoxication has been demonstrated.

Fig 6: Display of Alcohol level and EEG voltage in LCD display

The system's ability to recognize alcohol consumption is important to prevent drunk driving, as it uses relay control to prevent the engine from starting in such situations .The real-time location tracking enabled by the integration of GPS technology is crucial for fast response and intervention during seizures. The GPS function improves the overall efficiency of the system by ensuring that help can be immediately sent to the exact location of the person in need. The IoT architecture enables seamless communication with cloud platforms, storing, analysing and monitoring data remotely. Cloud integration provides nurses, medical professionals and other stakeholders with access to historical and real-time data, enabling a thorough understanding of a patient's condition and potential triggers. The comprehensive solutions of the study associated with both seizures and the potential dangers of alcohol consumption, providing a strong safety net for people with epilepsy. The system promotes independence by allowing people to prioritize their own safety in different driving situations and participate in the driving of vehicles. A webpage which stores and shows the latitude and longitude of the person has been created and the results are shown in Fig.6 and Fig.7.Only the people with specific username and password can access it. The same will be displayed on the LCD and indicates the alcohol level and the voltage level in microvolts.

Fig 7:Data observed sent to webpage through IOT module

The data observed through the alcohol sensor and EEG sensor is sent through the IOT module to the webpage developed to store the same. It acts like a vehicle blackbox except it is used for detection of seizure and drunk driving purposes. The latitude and longitude will be sent to the webpage admin so as to monitor the patient and ensure a user-friendly lifestyle.

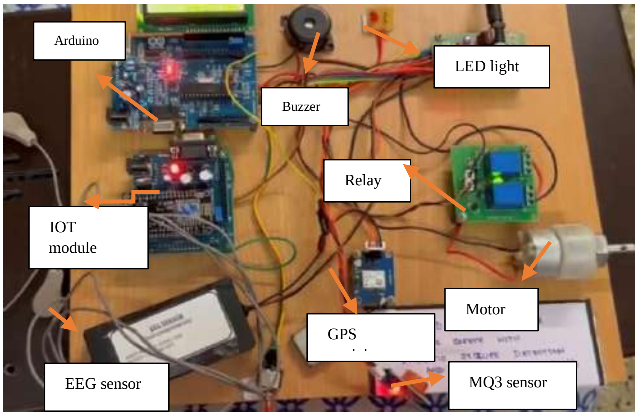


Fig 8: Developed Hardware Prototype

Fig 9 : Webpage to store and display alcohol level, Seizure detection and location of the person

Search:

	Logdate	LogTime
	01/27/2024	12:56:10
	01/27/2024	12:57:17
	01/27/2024	12:57:28
	01/27/2024	12:58:26
	01/27/2024	12:59:02

Fig 10:Webpage displaying the time of detection

VIII. CONCLUSION AND FUTURE WORKS

The proposed integrated IoT-based safety system is a game-changing innovation that fills in important holes in current technology for people with epilepsy. This system provides an all-encompassing method of controlling risks associated with epilepsy by integrating features such as the detection of epileptic seizures, the monitoring of alcohol levels in real time, GPS tracking, and cloud connectivity. A compassing safety net that allows people with epilepsy to live more independently is made possible by the benefits of real-time intervention, prevention of impaired driving, and adjustable parameters. Secure your sensitive health information with confidence thanks to the user-friendly UI and strong data security features. By bringing together experts from different fields, the system is better able to meet real-world requirements and win over users. In the end, the suggested system is a revolutionary leap forward in improving overall the security, independence, and general health of people with epilepsy in many everyday situations, thanks to its basic features of scalability and adaptability.

REFERENCES

[1] D. Torse, V. Desai, and R. Khanai, ‘‘A review on seizure detection systems with emphasis on multi-domain feature extraction and classification using machine learning,’’ BRAIN. Broad Res. Artif. Intell. Neurosci., vol. 8, no. 4, pp. 109–129, 2017.

[2]K. Abualsaud, A. Mohamed, T. Khattab, E. Yaacoub, M. Hasna, and M. Guizani, ‘‘Classification for imperfect EEG epileptic seizure in IoT applications: A comparative study,’’ in Proc. 14th Int. Wireless Commun. Mobile Comput. Conf. (IWCMC), Jun. 2018, pp. 364–369.

[3]M. Iftikhar, S. A. Khan, and A. Hassan, ‘‘A survey of deep learning and traditional approaches for EEG signal processing and classification,’’ in Proc.IEEE 9th Annu. Inf. Technol., Electron. Mobile Commun. Conf. (IEMCON),Nov. 2018, pp. 395–400.

[4] L. De Clerck, A. Nica, and A. Biraben, ‘‘Clinical aspects of seizures in the elderly,’’ Geriatrie et Psychologie Neuropsychiatrie du Vieillissement, vol. 17, no. 1, pp. 7–12, 2019.

[5]N. van Klink, A. Mooij, G. Huiskamp, C. Ferrier, K. Braun, A. Hillebrand, and M. Zijlmans, ''Simultaneous MEG and EEG to detect ripples in people with focal epilepsy,'' Clin. Neurophysiol., vol. 130, no. 7, pp. 1175–1183, Jul. 2019.

[6] N. Pianou and S. Chatziioannou, ''Imaging with PET/CT in patients with epilepsy,'' in Epilepsy Surgery Intrinsic Brain Tumor Surgery. Berlin, Germany: Springer, 2019, pp. 45–50.

[7]G. Zazzaro, S. Cuomo, A. Martone, R. V. Montaquila, G. Toraldo, and L. Pavone, ''EEG signal analysis for epileptic seizures detection by applying data mining techniques,'' Internet Things, vol. 14, Jun. 2021, Art. no. 100048.

[8] S. Cheruvu, A. Kumar, N. Smith, and D. M. Wheeler, ''IoT frameworks and complexity,'' in Demystifying Internet of Things Security. Berlin, Germany: Springer, 2020, pp. 23–148.

[9]M. Soufineyestani, D. Dowling, and A. Khan, ''Electroencephalography (EEG) technology applications and available devices,'' Appl. Sci., vol. 10, no. 21, p. 7453, Oct. 2020.

[10]M. Carbó-Carreté, C. Cañete-Massé, M. D. Figueroa-Jiménez, M. Peró-Cebollero, and J. G. Rdia-Olmos, ''Relationship between quality of life and the complexity of default mode network in resting state functional magnetic resonance image in down syndrome,'' Int. J. Environ. Res. Public Health, vol.17, no. 19, p. 7127, Sep. 2020.

[11]P. Gelisse, P. Genton, A. Crespel, and P. H. Lefevre, ''Will MRI replace the EEG for the diagnosis of nonconvulsive status epilepticus, especially focal?'' Revue Neurologique, vol. 177, no. 4, pp. 359–369, Apr. 2021.

[12] A. Biondi, V. Santoro, P. F. Viana, P. Laiou, D. K. Pal, E. Bruno, and M. P. Richardson, ''Noninvasive mobile EEG as a tool for seizure monitoring and management: A systematic review,'' Epilepsia, vol. 63, no. 5, pp. 1041–1063, May 2022.

Modeling and Control of Solar Powered BLDC Motor System with PID and Sheppard Taylor Converter

S. Karthikeyan
Department of Electrical and Electronics Engineering
Sri Krishna College of Engineering and Technology
Coimbatore, India
karthitamilsona@gmail.com

K. Harish
Department of Electrical and Electronics Engineering
Sri Krishna College of Engineering and Technology
Coimbatore, India
harishkrishnan571@gmail.com

S. Dhinesh Balaji
Department of Electrical and Electronics Engineering
Sri Krishna College of Engineering and Technology
Coimbatore, India
dhineshbalaji004@gmail.com

R. Harikrishnan
Department of Electrical and Electronics Engineering
Sri Krishna College of Engineering and Technology
Coimbatore, India
harihk075@gmail.com

C. Anush
Department of Electrical and Electronics Engineering
Sri Krishna College of Engineering and Technology
Coimbatore, India
anushchinnasamy@gmail.com

***Abstract*—The proposed system presents a comprehensive approach for designing, simulating, and implementing a solar-powered air-purification system. This system incorporates a Sheppard-Taylor converter for voltage regulation, a Brushless DC (BLDC) motor, and an MPPT (Maximum Power Point Tracking) algorithm. For analysis and simulation tasks involve MATLAB application, ensuring the system's performance under diverse conditions. Sheppard-Taylor converter optimizes voltage regulation, while the BLDC motor drives the air purification unit efficiently. The MPPT algorithm maximizes power extraction from the solar panels. This integrated solution offers a sustainable approach to air purification, with MATLAB simulations validating its effectiveness for implementation. Suitable for off-grid and environmentally sensitive areas. Additionally, this system contributes to cleaner air and energy sustainability.**

***Keywords*—BLDC Motor, MATLAB Simulation, Insulated-Gate Bipolar Transistor (IGBT) Switch, Sheppard Taylor Converter, Maximum Power Point Tracking (MPPT) Algorithm.**

I. INTRODUCTION

Our paper is focused on exploring the use of solar-powered BLDC motor systems, in conjunction with the innovative Sheppard-Taylor Converter for voltage boosting and control. These advancements are crucial in addressing energy needs sustainably. By harnessing abundant solar energy, these systems offer a promising shift towards cleaner and greener energy solutions, particularly in vital sectors such as agriculture, water management, and rural electrification.

BLDC motors are widely recognized for their efficiency and precise control, making them the backbone of these systems. Their adaptability and reliability make them ideal for a wide range of tasks, from powering water pumps in remote areas to ventilating urban spaces sustainably. The Sheppard-Taylor Converter emerges as a game-changer, providing a cost-effective alternative to traditional boost converters. Its ability to efficiently regulate voltage and energy consumption, even amidst fluctuations in sunlight intensity, ensures optimal performance and reliability. Moreover, its simplicity in operation and maintenance makes it accessible to a wider audience, empowering communities to embrace solar energy with confidence. Furthermore, the Sheppard-Taylor Converter's capability to amplify voltage at minimal costs adds a significant advantage over conventional converters, making it an attractive choice for solar-powered setups, especially in resource-constrained environments. To maximize the impact of our project, we undertake a comprehensive analysis of MPPT algorithms. These algorithms are very useful in optimizing the efficiency and performance of solar-powered BLDC motor systems, ensuring that every ray of sunlight is utilized effectively.

II. SYSTEM DESCRIPTION

A. Solar Panel

The solar panel, a vital part of our system, contains many PV cells linked together to produce electricity from sunlight. These cells are usually made of silicon which creates electricity when sunlight hits them. The electricity generated is then collected and used to power electronic devices. Essentially, the solar panel acts as the main energy source for converting solar energy into usable electrical power for electronic devices.

TABLE I. SOLAR PANEL SPECIFICATION

Parameter	Magnitude
No load Voltage (Voc)	82.1V
Peak Power level Voltage (Vmp)	75V
Max. Current (Isc)	19.55A
Peak Power level Current (Imp)	18.7A
Peak Power (W)	1402.5W
No. of. Cells (Ncell)	34

V-Voltage, A-ampere, W-watts

B. BLDC System with PID Controller

The approach developed for BLDC motor speed management assisted by PID control technique involves several important steps. Firstly, we detect the motor's position and back electromotive force (BEMF) using hall sensors. The BEMF usually looks like a trapezoidal wave. We then use this information to figure out when to activate each phase of the motor. Next, we carefully create sequences of gate pulses, which control the power transistors in the motor's circuit. We compare the desired motor speed with its actual speed, creating an error signal. This signal goes into a PID controller, which adjusts the output based on how big the error is. The improved signal is combined with the gate pulse sequences, changing when or how long they're active to control the motor's speed better. This method keeps checking and adjusting the speed continuously, making sure it stays just right no matter what. Overall, using hall sensors, decoder blocks, PID controllers, and gate pulse sequences together helps us manage the speed of BLDC motors accurately and flexibly.

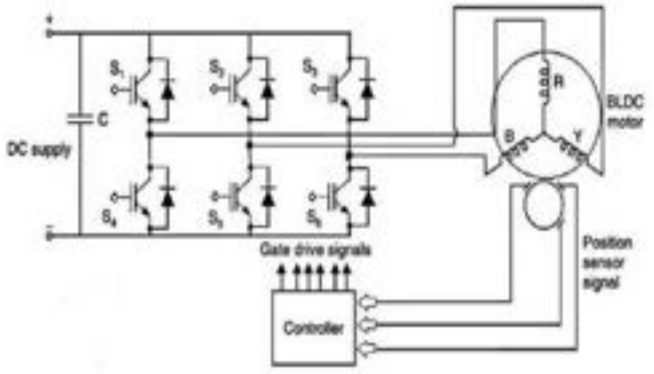


Fig 1. Speed control of BLDC motor

TABLE II. BLDC MOTOR SPECIFICATION

Parameter	Magnitude
Rated Power	630 W
Rated torque	0.21 Nm
Rated speed	3000 rpm
Rated current per phase	1.2 A
Rated Voltage	320 V
No of Pole pairs	2

W-watts, Nm-Newton meter, rpm-Revolutions per minute, V-Voltage, A-Ampere

In Table III the Logic table for the Inverter switching sequence, explaining how the transistors in the inverter switch for each step in controlling a BLDC motor is shown.

TABLE III. LOGIC TABLE FOR INVERTER

Gate Pulse Logic Table								
ε_a	ε_b	ε_b	P_1	P_2	P_3	P_4	P_5	P_6
0	0	0	F	F	F	F	F	F
0	-1	+1	F	F	F	T	T	F
-1	+1	0	F	T	T	F	F	F
-1	0	+1	F	T	F	F	T	F
+1	0	-1	T	F	F	F	F	T
+1	-1	0	T	F	F	T	F	F
0	+1	-1	F	F	T	F	F	T
0	0	0	F	F	F	F	F	F

In the PID Controller, the Kp improves responsiveness, Ki eliminates steady-state error by integrating past errors, and Kd minimizes overshoot by anticipating future error trends, collectively enhancing system stability and response. PID parameter values employed in our design are exhibited in Table IV

TABLE IV. VALUES OF PID CONTROLLER

PID VALUES			
K_p	K_i	K_d	Filter coefficient (N)
3.13	12.53	-0.05	7.56

C. Working principle of ST converter

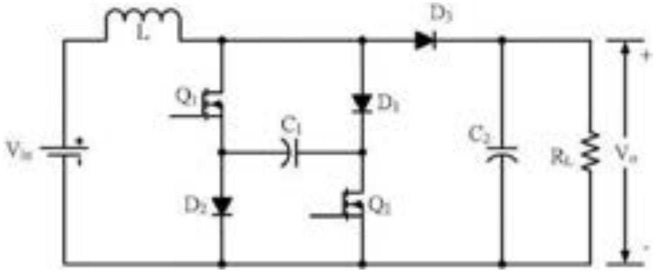


Fig. 2. ST Boost Converter Circuit

The ST converter, which helps pump up voltage, is made up of two IGBTs, three diodes, one inductor, and capacitor. These components are picked thoughtfully to keep energy steady, preventing any wobbles. It ensures a smooth flow of power, making sure everything runs smoothly and efficiently. The converter functions through three distinct stages. Initially, as the input voltage is introduced, the inductor begins to accumulate energy while the capacitor expels energy via the activation of Q1 and Q2 switches. At the outset, the voltage contained within the capacitor represents the discrepancy between the input voltage and that of the inductor. In the second stage, we halt the gate pulse to turn off switches Q1 and Q2. This action forward-biases diodes D1 and D2, allowing the capacitor to begin storing energy from both the source and the inductor. Meanwhile, diode D3 remains reverse-biased since the capacitor voltage hasn't reached a sufficient level to forward-bias it yet.. Over time, as the capacitor gains energy and its voltage exceed the output voltage, diode D3 becomes forward-biased, delivering the boosted voltage to the load with support from the inductor and capacitor.

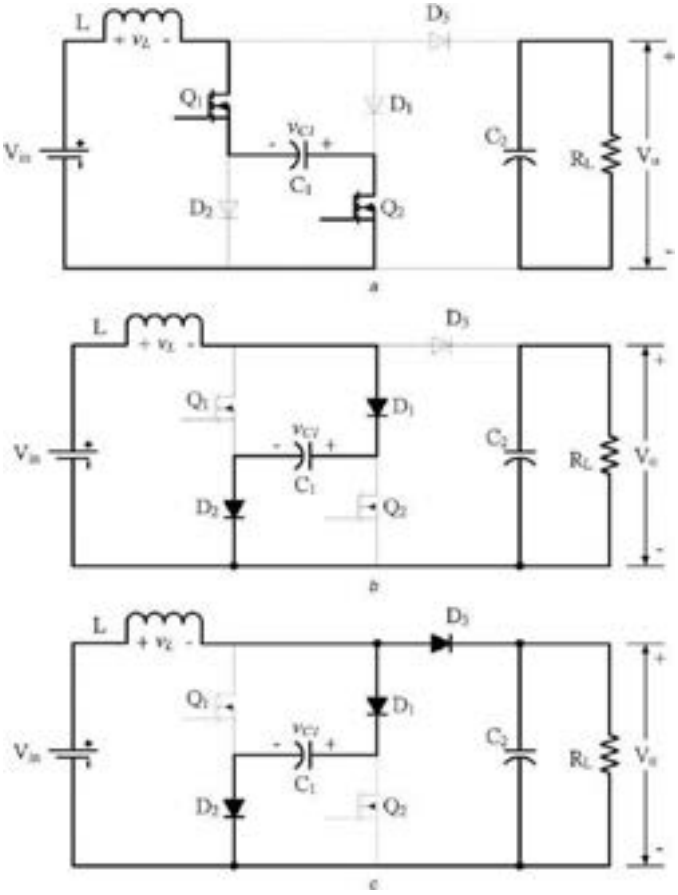


Fig. 3. ST Converter States of Working

TABLE III. ST CONVERTER SPECIFICATION

Component/Parameter	Value/Type
Input Inductor L_1	4.4 mH
Input Capacitor C_1	100 µF
Flying Capacitor C	47 µF
Output Capacitor C_0	56 µF

mH – millihenry, µF-microfarad

D. *MPPT Algorithms:*

In the Perturb and Observer (P&O) algorithm, the duty cycle of the converter is adjusted bit by bit. Initially, a small change is made, and then we check how much power the solar panel produces. If the power goes up, we keep moving in the same direction; if it goes down, we change direction. This process helps us find the best point where the solar panel produces the most power, which we call the MPP. By changing the duty cycle, the algorithm makes sure the system works at this best point, making the most power produced from the solar panel.

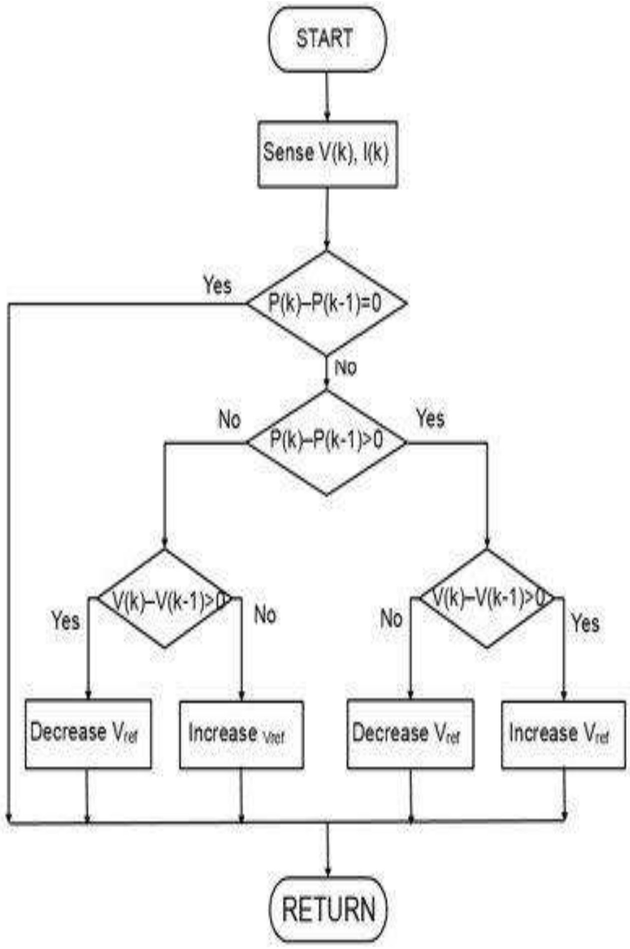


Fig. 4. PO Algorithm Process Diagram

In the Incremental Conductance algorithm, the duty cycle adjusts based on power differences between two points. If power increases, the duty cycle shifts in one direction; if it decreases, the shift is reversed. This process repeats until the algorithm identifies the maximum power point (MPP), where the power difference becomes zero. Following this method, the algorithm guarantees the system operates at the MPP, optimizing power generation from the solar panel.

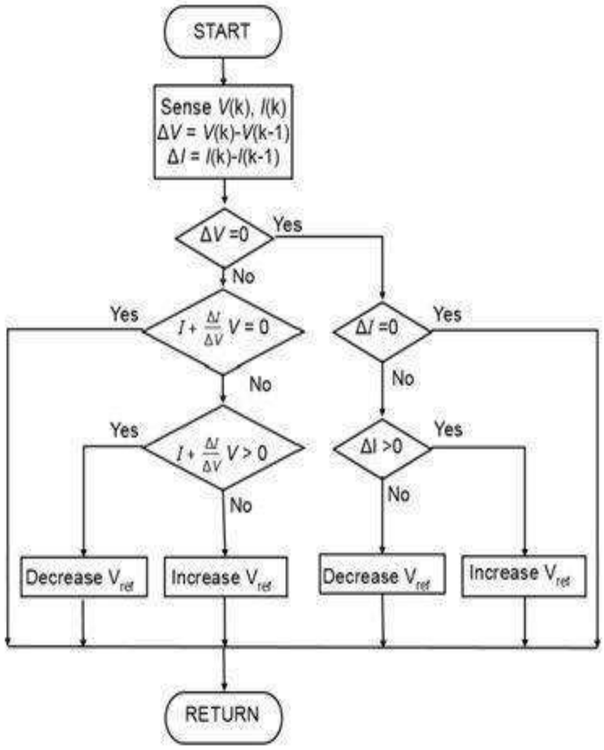


Fig. 5. IC Algorithm Process Diagram

III. PROPOSED SYSTEM

Our system proposal is designed to efficiently utilize solar energy. It utilizes a photovoltaic (PV) panel as its main power source, employing sophisticated MPPT techniques like Perturb and Observer (PO) and Incremental Conductance (INC). These algorithms optimize the PV panel's output, ensuring maximum power extraction even in different sunlight conditions. The voltage output from the PV panel undergoes regulation and amplification using an advanced ST boost converter. This converter plays a crucial role in maintaining a stable and elevated voltage level suitable for powering the load. Additionally, a 3Φ inverter is employed to transform converter voltage to an inverted state, with its output controlled by a precise Proportional-Integral-Derivative (PID) controller. This controller adjusts the input voltage to the Brushless DC (BLDC) motor based on a reference speed, thereby ensuring efficient speed regulation. Each component is intricately integrated, contributing to the overall efficiency and reliability of the system.

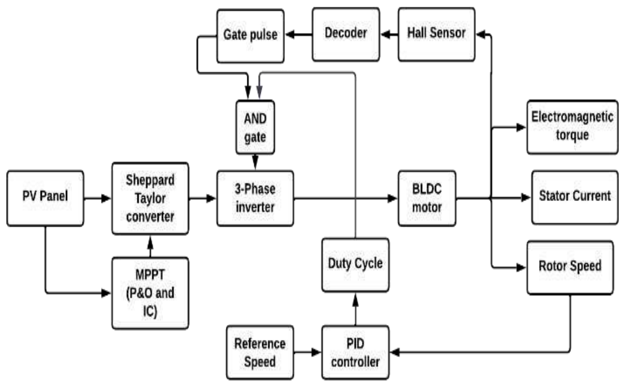


Fig.6.Proposed Block Diagram

IV. SOFTWARE SIMULATION

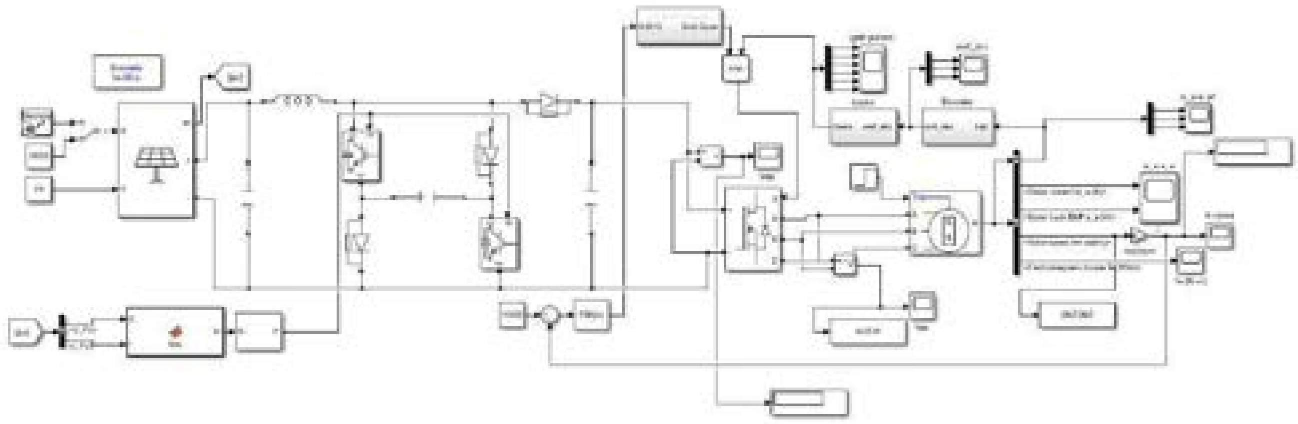


Fig. 7. MATLAB Simulation of Solar based PID controlled BLDC system with ST converter using Perturbation and Observation algorithm

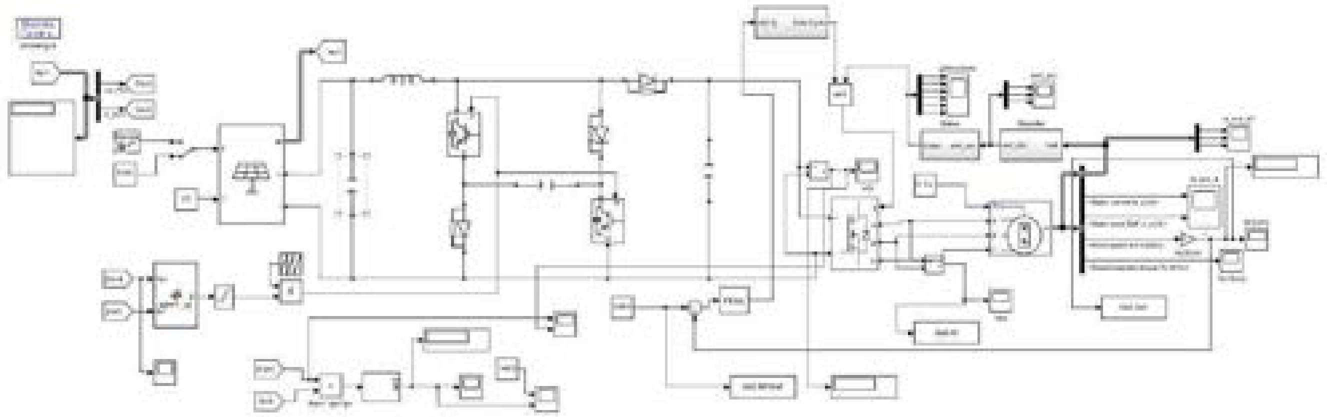


Fig. 8. MATLAB Simulation of Solar based PID controlled BLDC system with ST converter using Incremental Conductance algorithm

The complete system design is implemented as a Simulink model utilizing MATLAB R2023a software. Additionally, the Scope tool is utilized to capture and visualize the system's output graph, providing valuable data for analysis and discussion.

V. SOFTWARE SIMULATION

The simulations were conducted using a MATLAB Simulink model with specified parameters and the resultant voltage-time characteristics of both input voltage and ST Converter graph are depicted in Fig.9. Fig.10. and Fig.11.

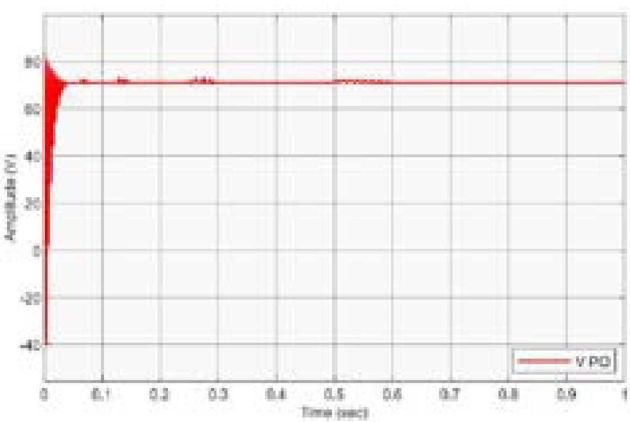


Fig 9. Voltage-Time graph of P&O algorithm

From Fig.9 and Fig 10, a close examination brings to light the unique behaviors of the P&O and IC algorithms. The P&O algorithm takes around 0.17 seconds to gradually approach the MPPT voltage, consistently hovering around this optimal level.

On the other hand, the IC algorithm demonstrates a swifter response, reaching the MPPT voltage more promptly and with steady accuracy. This detailed analysis provides deeper insights into the intricacies of MPPT algorithms and their role in enhancing solar energy conversion

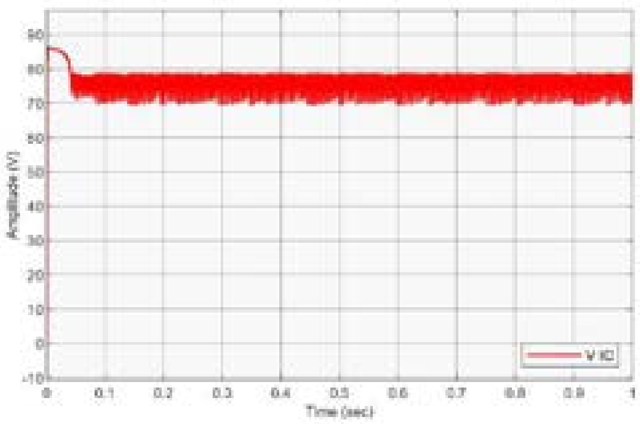


Fig 10. Voltage-Time graph of IC algorithm

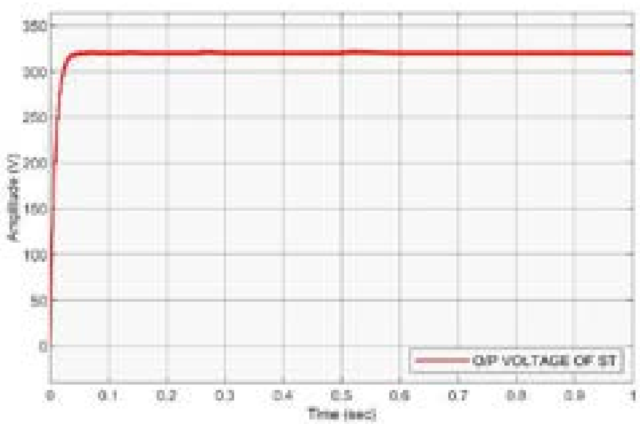


Fig 11. Resultant output voltage graph of ST converter

In Fig. 11, we can see how the ST Converter quickly ramps up the DC bus voltage to 320 V in just 0.18 seconds. Looking closely at the output voltage graph, it's clear that the converter handles this voltage increase swiftly and efficiently. This suggests that the converter can effectively regulate voltage changes, which is crucial for keeping power systems stable.

Its fast response time shows that it's suitable for tasks where voltage adjustments are needed quickly. This detailed analysis gives us valuable insights into how the ST Converter improves the reliability and efficiency of power systems, which can be useful for future studies and practical applications.

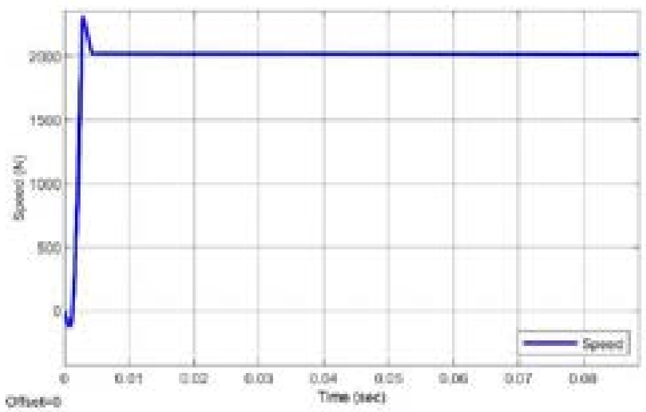


Fig. 12. Speed-time graph of BLDC motor

Fig 12 illustrates its speed-time profile. Notably, the motor achieves a target speed of 2000 RPM within just 0.18 seconds, though it experiences a peak overshoot of 350 RPM. These results, based on the constants Kp, Ki, and Kd set in the PID controller as mentioned before, clearly highlight the BLDC motor's quick reaction and dynamic behavior. This shows how suitable it is for tasks needing precise speed control, making it a great choice for different applications.

VI. CONCLUSION

This study has evaluated the efficiency and energy management of sustainable energy using MATLAB simulations. The results demonstrate the smooth operation of the ST converter alongside the PV system and BLDC system with the PID controller. Notably, the performance of the Incremental Conductance algorithm surpasses that of the Perturb and Observe algorithm in terms of accuracy, stability, and adaptability to varying irradiance levels. The effectiveness of the ST converter and BLDC system underscores their suitability for a range of industrial applications like air purifiers, blower fans, etc. The simulation output provides insights into optimizing this system and promoting the cost-effective operation of the ST converter. Overall, the system's significant impact lies in energy management and promoting clean and green technology.

VII. REFERENCE

[1] Ibrahim, Mohammad Haziq, Swee Peng Ang, Muhammad Norfauzi Dani, Mohammad Ishlah Rahman, Rafidah Petra, and Sheik Mohammed Sulthan. "Optimizing Step-Size of Perturb & Observe and Incremental Conductance MPPT Techniques Using PSO for Grid-Tied PV System." IEEE Access 11 (2023): 13079-13090.

[2] Sayyed Aadil, K., R. Kuwar Rohit, R. Kolhe Shubham, A. Ganesh Pandore, and H. R. Aher. "SOLAR OPERATED INDOOR AIR PURIFIER."

[3] Karthika, M., and K. C. R. Nisha. "Torque Ripple Reduction Control Strategies of Sensor and Sensorless BLDC Motor: A Review." In *ICDSMLA 2021: Proceedings of the 3rd International Conference on Data Science, Machine Learning and Applications*, pp. 681-692. Singapore: Springer Nature Singapore, 2023.

[4] Padilla-García, Erick Axel, Héctor Cervantes-Culebro, Alejandro Rodriguez-Angeles, and Carlos Alberto Cruz-Villar. "Selection/control concurrent optimization of BLDC motors for industrial robots." *Plos one* 18, no. 8 (2023): e0289717..

[5] Qaisar, Muhammad Waqas, Hassan Mujtaba, Muhammad Tanveer Riaz, Muhammad Shahid, Ahmad Abdul Ghani, Muhammad Abbas Khan, and Kashif Hussain. "Comparative Analysis of Fuzzy and Neural Controller for a Boost Converter." In *Proceedings of the 27th International Conference on Evaluation and Assessment in Software Engineering*, pp. 414-422. 2023..

[6] Islam, Sheikh Aminul, Md Osiur Rahaman, Md Faisal Hossain, Md Shaon Mridha, and Mohammad Jahidul Islam. "design, fabrication, and performance analysis of a solar powered air purifier." PhD diss., Sonargoan University (SU), 2023.

[7] Pathak, Pawan Kumar, Sanjeevikumar Padmanaban, Anil Kumar Yadav, P. A. Alvi, and Baseem Khan. "Modified incremental conductance MPPT algorithm for SPV-based grid-tied and stand-alone systems." *IET Generation, Transmission & Distribution* 16, no. 4 (2022): 776-791.

[8] Stephen, Akinyemi Ayodeji, Kabeya Musasa, and Innocent Ewean Davidson. "Modelling of solar PV under varying condition with an improved incremental conductance and integral regulator." *Energies* 15, no. 7 (2022): 2405.

[9] Sadick, Abubakari. "Maximum power point tracking simulation for photovoltaic systems using perturb and observe algorithm." (2023).

[10] Somasundaram, A., and A. Ravi. "Optimal Power Flow in Grid Connected EV Charging System Using High Gain Sheppard-Taylor Converter." *Electric Power Components and Systems* (2024): 1-21.

[11] Rinwa, Gouri, Nidhi Agarwal, Shaurya Dubey, and Punit Ratnani. "A Novel Technique for the Implementation of Dynamical Modeling of Brushless DC Motor using MATLAB/Simulink." In *2021 2nd Global Conference for Advancement in Technology (GCAT)*, pp. 1-4. IEEE, 2021.

[12] Lodhi, Poonam, Hitesh Joshi, and Pankaj Kumar. "Matlab Simulation of Photovoltaic System for Speed Control of BLDC Motor."

[13] Robert, Reenamol, and P. Sandhya. "Speed Control of Bldc Motor Using Pv and Sepic Converter." *International Journal of Innovations in Engineering Research and Technology* 8, no. 05: 34-41.

[14] Dinakaran, S. Amose, A. Bhuvanesh, A. S. Kamaraja, P. Anitha, K. Karthik Kumar, and P. Nirmal Kumar. "Modelling and performance analysis of improved incremental conductance MPPT technique for water pumping system." *Measurement: Sensors* 30 (2023): 100895.

[15] Anshory, Izza, Jamaaluddin Jamaaluddin, Arief Wisaksono, Indah Sulistiyowati, Bagus Setya Rintyarna, Ahmad Fudholi, Yuli Asmi Rahman, and Kamaruzzaman Sopian. "Optimization DC-DC boost converter of BLDC motor drive by solar panel using PID and firefly algorithm." *Results in Engineering* 21 (2024): 101727

[16] Sen, Aryadip, and Bhim Singh. "Parametric Observer Controlled Motor Invariant Electronic Commutation of Photovoltaic powered BLDCM without Position Sensor." *IEEE Transactions on Power Electronics* (2023).

Detection of Real Time Pothole System using Edge Detection

Ranjitha R
Assistant Professor, Dept. of EEE, *CMR Institute of Technology,*
Bengaluru-560037
ranjitha.maheshms@gmail.com

Archana K
Assistant Professor, Dept. of EEE,
CIT, Bengaluru *-560036*
archana.k@cit.com

Sankalp, Dept. of EEE,
CMR Institute of Technology,
Bengaluru *-560037*
Salp18ee@cmrit.ac.in

Sandesh B K, Dept. of EEE,
CMR Institute of Technology,
Bengaluru *-560037*
Sads18ee@cmrit.ac.in

Shashank, Dept. of EEE,
CMR Institute of Technology,
Bengaluru *-560037*
Shah18ee@cmrit.ac.in

***Abstract*— The purpose of streets is to facilitate safe vehicle movement at designated speeds. Surveillance is particularly crucial at intersections, sharp curves, congested areas, accident-prone zones, and other critical points on the road. Various measures are implemented to enhance safety, including mandatory speed limit signs, flashing signals for alerts, and road markings. While speed bumps may be necessary in certain situations, their widespread use is not considered ideal due to potential vehicle damage and risk of injury. Additionally, traffic signs can sometimes distract drivers, especially in low-light conditions, making it challenging to anticipate speed bumps. This research study intends to address these concerns by employing Machine Learning (ML) techniques, specifically edge detection, to identify potholes and bumps on roads. The proposed model achieves an accuracy of 98% in detecting the road hazards. An image-processing system is utilized for pothole detection, while ultrasonic sensors are employed to measure the depth of potholes, providing drivers with crucial information to avoid accidents or vehicle damage.**

Keywords—Potholes, Accidents, Machine Learning, Edge Computing

1. Introduction

A recent NCRB report on road accidents in India states that approximately 1.5 lakh people are injured from road accidents because of the bad infrastructure of the road, climate conditions, and Over-speed or overtaking. The majority of drivers today know what potholes are because they have experienced them at some point while driving down a city street, in their neighbor- hood, on the interstate or other highway, or any other road. These experiences are typically not positive memories for any driver. Potholes can cause serious damage to any vehicle and can create dangerous situations for drivers as well when trying to avoid them or by losing control after hitting them. Potholes are places or spots on roads where the structural integrity of an area has failed or been compromised, leaving an uneven surface or "hole" in the pavement or roadway. Potholes come in all shapes, sizes, and depths. These variations in size, depth, and location help determine the kind of pothole that it is, which also affects the method of re- pair that will be required to correct the dam- age in the most effective manner potholes most commonly appear in urban areas on city streets or in other areas where the road, pavement, or street receives regular traffic from motor vehicles. Although potholes are most common on city streets, they can appear on just about any road, street, highway, or other roadway. There are many reasons that a pothole may exist but here are the few most common reasons that potholes form due to, Improper roadway construction or compaction, Excessive or frequent use over a problematic area, Poor maintenance,Improper road or pothole repairs,Shifting in the earth, ground or surface area, Extreme or unnatural weather conditions, Material composition of the roadway or street. In India, potholes are a known fact and in every kilometre a potholes can be seen. A pothole is a type of depression on any plain ground or any surface. Potholes are the surface areas of road which have cracked, eroded, and eventually results in a form of hole. Initially they start as cracks of smaller size. If it is not fixed immediately,which results in the damage of the complete road and sometimes the complete area. The pothole may results from the capability of water absorption. The formation occurs when a rock surface is worn away by a circular water stream that carries small stones and debris. The force of water and the sediment it carries is greater than the resistance of the rock.After it starts, the process carries on, and the rock there keeps eroding. Potholes have to be taken seriously and are capable of causing serious or even fatal accidents[1-3] .

Common damage to vehicles inflicted by potholes includes flat tires (or other tire damage), damage to rims, suspension, or steering, and (in some cases) damage to a car's exterior. Hitting a pothole at any sort of speed can also affect the wheel alignment of a vehicle which can have a significant impact on how quickly tires wear and perform. Potholes are detected along with their severity using a built-in accelerometer. On the road, presence of variations are detected with the help of an accelerometer along with an ultrasonic sensor. Potholes are also detected by using the lasers which scan and measure the depth of the pothole. This paper presents a real-time building a prototype for the detection and reveal the presence of potholes by examining the road with the help of a camera. Potholes are identified based on the edge detection algorithms. It collects information about the potholes location coordinates with the help of a GPS module and stores the location and the images in the server of a database for future reference using the Wi-Fi module added with the microcontroller. This study aims at providing smart detection using image processing and simultaneously reporting the detected potholes using a web server. This process ensures a stable Wi-Fi connection. Smoothing filters such as Median blur are applied to remove the noise and simultaneously preserve the edges of the image. Threshold techniques are applied to adjust the color variants. Morphological techniques such as opening are applied to remove the small edges and thicken the required images. The canny edge algorithm, along with arc length techniques, is used to identify the pothole using image processing. Water has a way of eating away at road surfaces in any climate[4- 5]. Even without a freeze-and-thaw cycle, water can still get underneath the pavement and create divots by eroding the materials, such as loose gravel, that form the roadbed. With the proposed system an attempt has been made to endorse drivers to ward off the accidents causeddue to potholes.

Accidents are pretty common on the roads, as the cars move very close to each other, which can be prevented by assisting the driver by informing him that another vehicle is very close to him. In Emergency help and accident cases, sending the real-time location to registered numbers can save lives. Potholes and speed ramps are among the most common obstacles on most road networks. These road obstacles and disruptions in the surface condition of roads are among the causative factors of road accidents, carnages, and fatalities[6]. This project examines the prospects and challenges of the Internet of Things GPS Tracker in detecting, reporting, and management of pot-holes and other obstacles on our road networks fortunately, road maintenance crews can fix pot holes easily. Fixing a pothole requires a crew to clean all the loose rock and dirt out of the pot- hole before filling it with a mixture of hot or cold asphalt (depending upon the time of year) to create a patch in the roadway surface, according to the report of the National Crime Record Bureau(NCRB).

2.LITERATURE REVIEW

Road infrastructure repair has accrued much knowledge over the last few days. Poor road conditions often cause problems and harm to automobiles and passengers that cannot be ignored. Some of the big factors that are causing harm are potholes, speed-breakers, rough patches, etc. The existence of such abnormalities also affects the consistency of the ride and the protection of the passengers. A YOLO algorithm is proposed to identify the potholes, which evaluates the distortion rate, restoration rate, and model validity [7]. A collective data of potholes were collected using the built-in application from the mobile smartphone and which features out from the frequency of different domains. The features were extracted with the different types of machine learning classifiers. Among the various classifiers, the random forest algorithm gave a very fruitful result with an accuracy of 88.5%[8]. The paper explains the low-cost method that uses a smartphone and an OBD-II module to find and locate road pot-holes[9]. A Generative adversarial network is used to generate a pseudo image and combine it with the real image to detect the road dam- age which was used as a new set of data and was trained. this helps to improve the F measure by training the synthesized data images of road[10-11].

3.METHODOLOGY

The intention of the proposed work are design and development of the Pothole and hump detection System,which help for the road maintenance, aids the economic growth, and results in avoidance of accidents through a real-time alerts with the help of GSM. To exhibit the pothole location with latitude and longitude, in the Google Maps with the link is also showed in the TCP/UDP test tool.

Potholes are the road surface area that have cracked down, worn apart, and have formed huge holes varying from a few centimeters to meters. The most dominant reason for the formation of these holes is water. However, spillages and emissions by vehicles, accidents, and other dam- ages may also be the various causes. Moisture in the roads and surfaces freezes during the winter season leading to a decrease in the material density due to a rise in the volume. These holes further expand due to overlooking the issue and lack of proper care/maintenance[12]. Potholes can create a traffic congestion which leads to gross wastage of fuel. This ultimately adds to air pollution causing various, serious health problems for people. Besides, many accidents are caused due to the presence of these potholes. These accidents may be severe enough to cause damage to the body and the life of individuals or may simply affect the functioning of vehicles and tyres. The main criteria of this work is to plan the design and develop a intelligent model based on data-gathering system which can be fitted in any vehicle or car to detect road or highway conditions. The prototype develops an information-gathering system that informs the authorities about the conditions of road. By the gathered information, the authorities could take measures to repair the roads which in turn results to avoid accidents caused due to non-technical humps and hence lower the percentage of damaged roads and avoid road accidents[13- 16]. An Artificial Neural Network(ANN) can be implemented to detect the path holes and by training the models more accuracy can be implemented[17-19].

Figure 1: Block Diagram

The proposed system block diagram is shown as in fig.1. This research work uses Raspberry Pi 4, which is based on microprocessor and interfaced with codes. A system based on real-time uses the camera to detect the potholes of road and an information alert along with the accuracy of it being the pothole. The device must be trained with a sample amount of pictures of potholes to prepare it for the real-time scenario.Using a software-based method, the Raspberry pi is utilized to feed the programs intended for hardware implementation. The Pi camera records the video and sends the feed to the board of Raspberry Pi. The video in the input feed is converted into grayscale images and fits the dimension set up for the device. The image is then checked for the pothole and if there is detected, the alert will be sent to the user along with the probability of the pothole being detected.

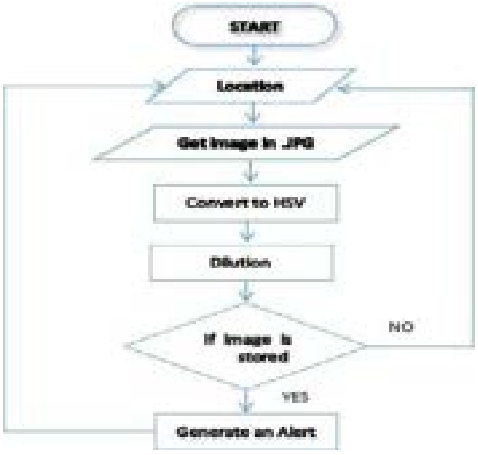


Figure 2: Flow Chart

4. Results & Discussion

The components are interfaced with Raspberry Pi 4. It is larger than the average micro- controller, uses more power, has a clock speed of between 700MHz-1.5GHz, and has a RAM capacity between 1GB to 8GB. Also, the Rasp- berry Pi, unlike microcontrollers, can run both 32-bit and 64-bit. A real-time detection project and it makes use of the camera to detect the potholes present on the road and sends an alert along with the accuracy of it being the pothole. The device must be trained with a less number of pothole images to prepare it for the real-time scenario. In this type of analysis a data sets of two types are used: the first type of database consists of 2000 distinct photos with potholes and may be accessed online. The second set of datasets will have 1800 photos gathered from various online sources and 344 photos taken from the Labanese roadways which indicated potholes. Most of the images are from videos that were recorded via cameras of dashboard of people driving their cars.This results into a 2087 different images, with more than 3000 potholes are used in this analysis. We used an image annotation program to name the thousands of pothole photos that we have gathered from the internet. Both commercial and open-source picture tagging solutions are available with some good features. We have made use of Microsoft's VoTT, which is accessible on GitHub under an MIT license and is open source.with the help of Python, deep learning models were built using the canny edge algorithm, and the model was built to detect the real path hole. The accuracy was developed with 68% with the first trial. Reusing the previous model and training it more vigorously we came up with the new trained model with an accuracy of 98%.

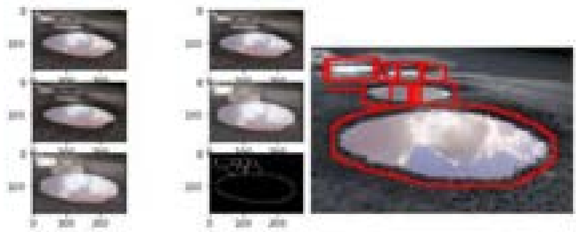


Figure 3: Image Recognition

Figure 5: Results obtained for real path holedetected images.

5. Conclusion

This research work presents a smart system for the identification of pothole based on Machine learning algorithms. The system starts by acquiring pothole images as data and pre-processing the data by creating image annotation and image augmentation on the pothole images. This study has carried out a real-time pothole identification based on multiple images of pothole. This work can further be extended by detection on a live-streaming video using the Open CV library.Then by using this Machine Learning code, we implement it using Raspberry Pi 4 Micro-processor and pie cam we take real-time video which allows us to implement it in any Automobile. Our model is based on machine learning and was developed completely with the help of image processing. The barriers of social to be over-come by the autonomous vehicle and is based on the t technology advancement. Nowadays for these cars, new rule will create a new opportunities.

The proposed work will be helpful for driver-less cars with the introduction of advanced automobiles to enable pothole detection. This work can be extended by considering the datasets, which consists of more than 2000 images, for training purposes of the network which may contain images of potholes from different roads, with different severities and several lighting and weather conditions. Manholes images can also be used for our system training. Potholes and manholes will have a similar characteristic and this desired improvement have to be done in our present system for segregate the difference between manholes and potholes. The present system can also be implemented into several cars to identify the road condition in real-time with the help of GPS to get the maintenance of potholes dimensions.

REFERENCES

[1] Anand, Sukhad, et al. "Crack-pot: Au- tonomous road crack and pothole de- tection." 2018 Digital Image Computing:Techniques and Applications (DICTA). IEEE, 2018.

[2] Baek, J.-W.; Chung, K. Pothole Classi- fication Model Using Edge Detection in Road Image. Appl. Sci. 2020, 10, 6662. https://doi.org/10.3390/app10196662

[3] Wu, C.; Wang, Z.; Hu, S.; Lepine, J.; Na, X.; Ainalis, D.; Stettler, M. An Au- tomated Machine-Learning Approach for Road Pothole Detection Using Smartphone Sensor Data. Sensors 2020, 20, 5564. https://doi.org/10.3390/s20195564

[4] Guerrieri, Marco, et al. "Asphalt Pavement Damage Detection through Deep Learning Technique and Cost-Effective Equipment: A Case Study in Urban Roads Crossed by Tramway Lines." Infrastructures 9.2 (2024): 34.

[5] Chugh, Gunjan, Divya Bansal, and San- jeev Sofat. "Road condition detection us- ing smartphone sensors: A survey." Inter- national Journal of Electronic and Electri- cal Engineering 7.6 (2014): 595-602.

[6] Jin Lin and Yayu Liu Ajit Danti, Jyoti Y. Kulkarni, and P. S. Hiremath, "An Image Processing Approach to Detect Lanes, Pot- holes and recognize road Signs in Indian Roads", International Journal of Modelingand Optimization, Vol. 2, No. 6, Decem- ber2012.

[7] Kiran Kumar Vupparaboina and Roopak

R. Tamboli "Overtaking and receding ve- hicle detection for driver assistance and naturalistic driving studies." In Intelli- gent Transportation Systems (ITSC), 2014 IEEE 17th International Conference on, pp.697-702. IEEE, 2014.

[8] Zhajian Li I. Kolmanovsky, E. Atkins, J. Lu, D. Filev, and J. Michelini, "Cloud aided safety- based route planning," in 2014 IEEE International Conference on Systems,Man, and Cybernetics, pp. 2525–2530, Octo- ber2014.

[9] Rui Fan H. Youquan, W. Jian, Q. Hanx- ing, Z. Wei, and X. Jianfang, "A research of pavement potholes detection based on three-dimensional projection transforma- tion," in Proc. 4th Int.Congr. ImageSignal- Process (CISP),Oct. 2011,pp. 1805–1808

[10] I. Moazzam, K. Kamal, S. Mathavan, S. Us- man, and M. Rahman, "Metrology and vi- sualization of potholes using the microsoft Kinect sensor," in Proc. 16th Int. IEEE Conf. Intell.Transp. Syst., Oct. 2013,pp. 1284–1291 .

[11] Sachin Bharadwaj Sundra Murthy and Golla Varaprasad, "Automatic Detection and Notification of Potholes and Humps on Roads to Aid Drivers", IEEE Sensors J., Vol. 15, No. 8, August 2015.

[12] Zhaojian LiI. Kolmanovsky, E. Atkins, J. Lu, D. Filev, and J. Michelini, "Cloud aided safety- based route planning," in 2014 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2525–2530, Oc- tober2014.

[13] Vijaya Kumar Samyak Kathane, Vaibhav Kambli, Tanil Patel and Rohan Kapadia,"Real Time Potholes Detection and Vehicle Accident Detection and Reporting System and Antitheft (Wireless)", IJETT, Vol. 21, No. 4, March 2015.

[14] Van Khang Nguyen and ´ Eric Renault, Taehyeong Kim, Seung-Ki Ryu, "A Guide- line for Pothole Classification", Interna- tional Journal of Enginering and Technol- ogy (IJET), Vol. 4, No. 10, October 2014.

[15] Madli, R., Hebbar, S., Pattar, P., Golla, V. (2015). Automatic detection and noti- fication of potholes and humps on roads to aid drivers. IEEE Sensors Journal, 15(8), 4313–4318,https://doi.org/10.1109/JSEN.2015.2417579.

[16] Rauber, T. W., da Silva Loca, A. L., de Boldt, F. de A., Rodrigues, A. L., Varejao, ˜ F. M. (2021). An experimen- tal methodology to evaluate machine learning methods for fault diagno- sis based on vibration signals. Expert Systems with Applications, 167, 114022, https://doi.org/10.1016/j.eswa.2020.114022.

[17] Ranjitha, R. Agalya, V. Archana, K.. (2022). Diabetes Prediction by Artificial Neural Network. 10.1007/978-981-16-5529- 6$_7$6.

[18] Savitha, S., and N. Sreedevi. "A Study on Accident Detection Systems Using Machine Learning." *International Conference on Innovations in Computer Science and Engineering*. Singapore: Springer Nature Singapore, 2022.

[19] Behera, Sagarik, et al. "Chat Application Using Homomorphic Encryption." *ITM Web of Conferences*. Vol. 50. EDP Sciences, 2022.

Detection and Classification of Multi Cancers using Transfer Learning

1[st] Reddy Hemanth
Networking and Communication
SRM Institute of Science and Technology
Chennai, Tamil Nadu
rq0665@srmist.edu.in

2[nd] G Punith Sai Reddy
Networking and Communication
SRM Institute of Science and Technology
Chennai, Tamil Nadu
gr3062@srmist.edu.in

3[rd] Dr. Fancy C
Networking and Communication
SRM Institute of Science and Technology
Chennai, Tamil Nadu
fancyc@srmist.edu.in

***Abstract*—Worldwide, cancer is regarded as the leading cause of death; brain and lung cancers are especially difficult to identify and categorize in their early stages. For effective cancer treatment and better patient outcomes, early cancer detection and classification are essential. However, because these cancers are complex and don't always exhibit distinct symptoms, conventional diagnostic techniques frequently fall short of identifying them in their early stages. Deep learning has demonstrated significant impact in the identification and categorization of a number of illnesses, including cancer. But there is still space for improvement in the precision of current deep learning models for the identification and categorization of early-stage cancers, especially when it comes to malignancies like brain and lung cancers. The process of creating precise methodologies is fraught with difficulties. This study suggests using deep learning based transfer learning models to classify images of various cancer types, such as lung, brain, Acute Lymphoblastic Leukemia and oral cancer. In this work, Transfer learning model is evaluated against the classification of images exhibiting malignant features. Pre-trained CNN models include ResNet50, MobileNet, DenseNet and VGG, use the knowledge acquired from the ImageNet dataset to identify various types of cancer. The utilization of the ResNet50 model within this research has yielded notably superior results. These outcomes are crucial for making precise decisions in the medical domain.**

***Index Terms*—Convolutional Neural Network, ResNet50, SDG, ADAM, Transfer Learning**

I. INTRODUCTION

Cancer is a complicated and widespread illness defined by abnormal cells in the body growing and spreading out of control. It can impact different tissues and organs, causing serious health risks and frequently require extensive medical care. Numerous risk factors, including genetics, environmental exposures, lifestyle decisions, and others, can contribute to the development of cancer. Every year, millions of new cases of cancer are diagnosed worldwide, raising serious concerns about public health. Tobacco use, poor diets, alcohol consumption, physical inactivity, exposure to carcinogens like chemicals and genetic predisposition and UV radiation, are the main causes of cancer. These factors contribute to the development of various types of cancer, leading to significant morbidity and mortality rates.

Cancer is responsible for a substantial number of deaths worldwide. The disease can progress rapidly if not detected and treated early, leading to poor prognosis and decreased survival rates. Effective cancer prevention, early detection, and timely treatment are critical in reducing mortality rates and improving patient outcomes. Some of the most common types of cancer worldwide include lung cancer, breast cancer, oral cancer, brain cancer, skin cancer (melanoma), and cervical cancer. These cancers vary in their risk factors, symptoms, diagnostic approaches, and treatment modalities, highlighting the need for personalized and comprehensive cancer care strategies.

The medical industry faces several challenges in combating cancer effectively. These challenges include limited access to healthcare services, inadequate screening and diagnostic tools, high treatment costs, disparities in healthcare delivery, and insufficient public awareness about cancer prevention and early detection. Additionally, the rapid advancements in cancer research and treatment require continuous education and training for healthcare professionals.

Early prediction and detection of cancer play a vital role in improving therapy outcomes and patient survival rates. Key steps to facilitate early prediction include promoting healthy lifestyles and behaviors, regular screening and diagnostic tests for high-risk individuals, implementing advanced imaging and molecular technologies for early detection, and enhancing public education and awareness campaigns about cancer signs and symptoms.

In the field of medicine, the application of deep learning in cancer detection models has had a revolutionary effect. These cutting-edge technologies have greatly improved cancer diagnosis speed, accuracy, and efficiency, enabling early identification and individualized treatment plans. Medical professionals can make well-informed decisions and improve patient care pathways by using ML models to predict cancer risk factors, treatment responses, and prognosis. Convolutional neural networks (CNNs), one type of deep learning algorithm that is especially good at image recognition, allow for accurate and automated analysis of medical images for the purpose of classifying and detecting tumours. This is resulting in better patient outcomes, lower healthcare costs, and faster advancements in research.

II. Related Works

Various conventional methods for tumor detection and classification are both straightforward and intricate, demanding extensive time and continuous labor. Over time, there has been a shift towards employing deep learning models for cancer classification. This segment provides an information of recent endeavors in utilizing deep learning models for the detection and classification of a diverse range of cancers.

Subramanian et al.'s research on multiple types of cancer classification using VGG16, VGG19, DenseNet, and MobileNet, presents a significant contribution to the field[1]. The utilization of Bayesian optimization for hyperparameter tuning and the comparison of accuracy among various CNN models offer valuable insights into the performance of different architectures. However, the study also highlights certain limitations, such as the requirement for a supercomputer with high RAM and CPU capabilities to process the dataset efficiently. Additionally, the achieved accuracy of around 80percent indicates room for improvement in future endeavors to enhance the precision and reliability of cancer classification systems based on deep learning models.

B. H. Sai et al.'s research on deep learning classifier and Multi-Layer Perceptron (MLP) for lung cancer stage prediction offers valuable insights into the potential of advanced machine learning techniques in healthcare. The utilization of SVM and Random Forest models alongside deep learning capabilities demonstrates a comprehensive approach [2] to accurate staging, which could significantly aid in treatment decision-making processes. However, the study also acknowledges the challenges posed by the need for large datasets and substantial computational power for efficient training, highlighting areas for future research to address scalability issues and further improve the reliability and accessibility of predictive models in clinical settings.

Ravuri, Viswanadham et al.'s work on multi-cancer early detection and classification SVM, Decision Tree, and K-nearest neighbor, represents a significant effort in leveraging data- driven methods for cancer detection[3]. The positive aspects include the utilization of a large dataset for classification purposes, which is crucial for training robust models. However, the study also identifies limitations such as the reliance on traditional machine learning models, which may not capture the complexity of cancer data as effectively as deep learning models. Additionally, the accuracy achieved may not meet the desired standards, suggesting avenues for improvement through the exploration of more advanced techniques and model architectures in future research endeavors.

R. Cao et al.'s research on predicting microsatellite instability in colorectal cancer through deep learning and multi- scale pathological pictures offers a promising avenue for improving diagnostic capabilities[4]. The inclusion of multiple diseased picture scales in the processing approach enhances the accuracy and reliability of microsatellite instability pre- diction. However, potential drawbacks such as the necessity for specialized knowledge in image processing may present challenges in implementing the method widely across health-care settings. Future research could focus on developing user-friendly tools or platforms that integrate advanced image processing techniques seamlessly, bridging the gap between sophisticated technology and practical clinical application for improved cancer diagnostics.

Through a multi-net architecture, E. M. Roopa Devi et al. use CNN to solve unbalanced breast cancer prediction[5] by using vgg19 model. It is probable that the process includes correcting the dataset's class imbalance. The advantage is that unbalanced data may be handled for more precise predictions; nevertheless, selecting suitable multi-net framework designs could provide some difficulties.

For the purpose of predicting pancreatic cancer, X. Chen et al. provide a combined spiral transformation and model-driven multi-modal deep learning scheme[6] using regression and random forest. Model-driven techniques and spiral trans-formation are included in the process. Potential advancements in mutation prediction are one of the merits; yet, the model's interpretability may provide difficulties.

Using histopathology pictures, J. Moranguinho et al. present an attention-based lung cancer prediction[7] using Grad-Cam algorithm. It is possible that the approach makes use of attention processes to capture pertinent characteristics. Improved feature representation is a benefit, but there may be drawbacks, such as the requirement for a large labeled training data.

Al-Shouka et al., explored the potential of transfer learning for lung cancer detection using convolutional neural network (CNN) architectures like ResNet, VGG16, Xception, and MobileNetV2. Their approach achieved a promising accuracy of 91 percent, demonstrating the effectiveness of transfer learning in this domain. However, the approach has limitations. It necessitates large datasets for training, and the computational demands can be significant. Future research efforts could address these limitations by exploring techniques for efficient training with smaller datasets or by investigating more lightweight CNN architectures.

Welikala et al.'s research on detection and classification of oral cancer using deep learning, specifically ResNet-101-based object detection frameworks, marks a significant step towards early detection of oral cancer. Using ResNet-101 as an image classification model demonstrates the potential of deep learning in improving the accuracy and efficiency of oral lesion detection. However, the study also acknowledges limitations such as the restricted dataset size and challenges associated with object detection in oral images, which may affect the model's generalizability and robustness. Future endeavors could focus on addressing these limitations through the acquisition of larger datasets and the development of more robust object detection techniques tailored for oral cancer detection, thereby enhancing the clinical utility and impact of deep learning-based approaches in oral cancer diagnostics.

III. Proposed Methodology

A. Dataset

As part of cancer research and diagnosis, datasets containing medical images play a crucial role in training ML models for accurate classification. One such dataset, sourced from Kaggle, comprises MRI/CT images related to seven distinct classes of cancer. These classes represent a diverse range of cancer types, each with its unique characteristics and imaging features. To facilitate robust model, the dataset is split into 2 subsets: training and testing. Training set forms 80 percent of the data, providing a substantial volume of examples for the model to learn from. This large training set allows the model, particularly a deep learning architecture like ResNet50, to capture intricate patterns and nuances associated with each cancer class. The remaining 20 percent of the data forms the testing set, serving as an independent benchmark to assess the model's performance and generalization capabilities.

Cancer	Classes	Images
Acute Lymphoblastic Leukemia	4	20000
Brain Cancer	3	15000
Lung and Colon Cancer	5	25000
Oral Cancer	2	10000

TABLE I
INFORMATION OF DATASET

B. Image Processing

The image preprocessing step in the training process is essential for enhancing the performance and robustness of deep learning models, such as the ResNet50 algorithm in this case. In the provided code snippet, the tf.keras. preprocessing. image. ImageDataGenerator class is utilized with various parameters to augment and preprocess the input images before feeding them into the model.

Firstly, the zca epsilon parameter controls the intensity of ZCA whitening, which helps in reducing redundancy and improving the model's ability to learn discriminative features. The rotation range, width shift range, heightshift range, shear range, and zoom range parameters introduce random transformations to the images during training, such as rotation, shifting, shearing, and zooming. These augmentations help in making the model more robust to variations in input data and improve its generalization ability. Additionally, fill model handles how pixel values are filled in case of transformations, while horizontal flip and vertical flip perform horizontal and vertical flips, respectively, to further diversify the training data. The preprocessing function parameter allows applying custom preprocessing functions, such as preprocess input, which can normalize pixel values and prepare the images for optimal model training. Finally, the validation split parameter splits the dataset into training and validation sets, with 5 percent of the data reserved for validation to monitor the model's performance during training and prevent overfitting.

C. Model

In this study, used resnet50 CNN model to segregate the varieties of cancer by taking MRI and CT images as input.

CNNs leverage filters (kernels) to scan images, extracting key features. Through stacked convolutional layers, they progressively build a deeper understanding of the image content. This architecture, capturing spatial relationships between pixels, makes CNNs ideal for image analysis tasks.

Model Building:

ResNet-50 is a CNN model widely used in transfer learning for various deep learning tasks. ResNet-50 is better because it can learn hierarchical features well because it has a deep design with 50 layers that include convolutional layers, pooling layers, and residual blocks with shortcut connections. In transfer learning, the already trained ResNet-50 model's acquired characteristics are kept by turning off the top layers. Putting custom layers like MaxPooling2D, Flatten, and Dense layers (fully linked layers) also made a big difference in getting these excellent outcomes. Because this method cuts down on training time and computer resources by a lot, ResNet-50 is often used for transfer learning tasks like object detection, semantic segmentation and image classification where labeled data is scarce but high model accuracy is needed. In this work, the ResNet50 model leverages Adaptive Moment Estimation (ADAM) to dynamically adjust learning rates for individual parameters during the training process. This iterative optimization technique utilizes estimates of gradient moments to accelerate convergence and enhance training performance. ADAM effectively addresses challenges like vanishing gradients and the need for parameter-specific learning rates.

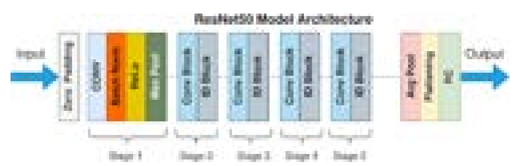


Fig. 1. ResNet50 Architecture

IV. System Architecture

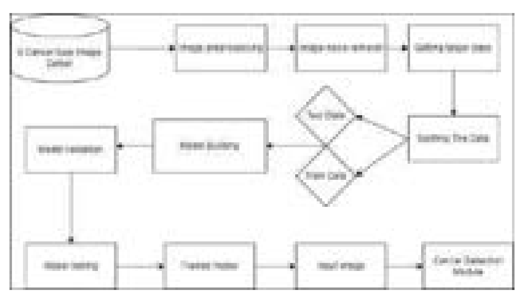


Fig. 2. System Architecture

Utilized a dataset comprising MRI/CT images representing eight distinct types of cancers, each further categorized into various mutations. Standard preprocessing techniques are applied to normalize pixel values and resize the images to a standardized input size of 224x224 pixels, aligning with the requirements of the ResNet50 model. The ResNet50 architecture, a deep CNN pre-trained on the ImageNet dataset, serves as the base model. By disable its top layers, which include fully connected layers, and add additional layers such as MaxPooling2D, Flatten, and Dense layers, along with batch normalization, to adapt the model for specific classification task and the use of Adaptive Moment Estimation (ADAM) iteratively adjusts learning rates for individual parameters in a neural network. After incorporating all the necessary preprocessing layers, the model will be compiled and then subjected to training with 50 epochs. This signifies that the algorithm will undergo 50 cycles of training.

V. Results and Discussion

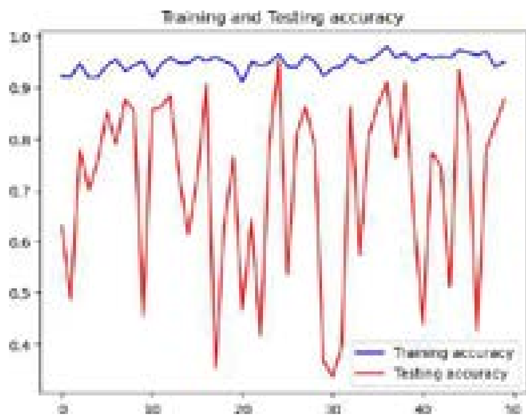


Fig. 3. Accuracy Graph of Brain Cancer

In Figure 3, the accuracy graph of Trained ResNet50 model for Brain Cancer reveal significant progress in the model's performance over the course of training. By the end of 50 epoch, the model have achieved 95 percent of Training accuracy and 87.67 percent of Testing accuracy. Figure 4, the accuracy graph of Trained ResNet50 model for Oral Cancer reveal significant progress in the model's performance over the course of training. By the end of 50 epoch, the model have achieved 87.18 percent Training accuracy and of 86.71 percent of Testing accuracy.

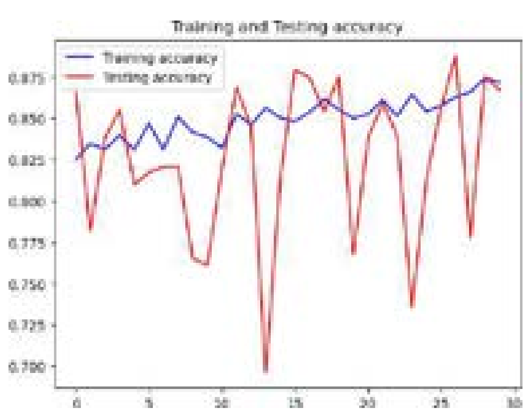


Fig. 4. Accuracy Graph of Oral Cancer

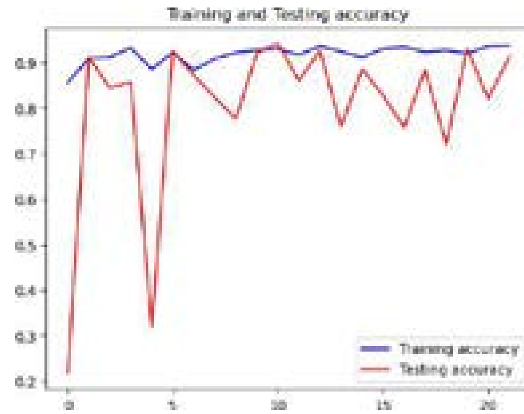


Fig. 5. Accuracy Graph of Lung and Colon Cancer

The accuracy curve of the trained ResNet50 algorithm for Lung and Colon Cancer in Figure 5 shows that the model's results over the course of training. By the end of 50 epoch, the model has achieved 93.7% of Training accuracy and 91.27 percent of Testing accuracy.

Classifier	Train Accuracy	Test Accuracy
Resnet50	93.77	91.27
VGG19	80.93	76.34
DenseNet	84.78	79.82
MobileNet	83.04	81.67

TABLE II
Comparison of Accuracy within models

The comparative analysis of different deep learning models, namely ResNet50, VGG19, DenseNet, and MobileNet, reveals that the proposed ResNet50 model achieves the highest accuracy of 91 percent. This signifies that ResNet50 outperforms the other models in terms of its predictive capabilities and ability to accurately classify data. The superior performance of ResNet50 can be attributed to its deeper architecture, which allows for more complex feature extraction and representation learning, leading to enhanced accuracy in various tasks such as image classification. Overall, these findings highlight the effectiveness of ResNet50 as a robust and reliable deep learning model for achieving high accuracy in classification tasks when compared to VGG19, DenseNet, and MobileNet.

The ResNet50 model was trained on dataset contains of different classes, achieving an aggregated accuracy of approximately 91% shown in fig 6. The dataset encompasses various types of cancers such as brain cancer, oral cancer, and lung cancer, among others. During training and evaluation, the model demonstrated robust performance across these diverse classes, showcasing its efficacy in accurately identifying and classifying different types of cancer. In addition to accuracy, other metrics such as precision, recall, and F1 score were also evaluated to assess the model's performance comprehensively. These results suggest that the ResNet50 model holds promise as a valuable tool in cancer diagnosis and classification tasks, offering high accuracy and reliable performance across multiple cancer types.

Classifier	Cancer Type	Class Labels	Train Accuracy	Test Accuracy
ResNet50	Acute Lymphoblastic Leukemia	all_benign	98.64	98.69
		all_early		
		all_pre		
		all_pro		
	Brain Cancer	brain_glioma	95.00	87.67
		brain_menin		
		brain_tumor		
	Lung and Colon Cancer	colon_aca	93.77	91.27
		colon_bnt		
		lung_aca		
		lung_bnt		
		lung_scc		
	Oral Cancer	oral_normal	87.18	86.71
		oral_scc		

Fig. 6. Model accuracy for multi cancer

VI. Conclusion and Future Works

In conclusion, this research stands out by providing a comprehensive platform for the classification of various types of cancers, offering a unified solution compared to other papers that focus on predicting or classifying individual cancer types. By leveraging the ResNet50 model, which boasts a superior architecture with 50 layers and achieving an accuracy of 98.69 percent accuracy in identifying Acute Lymphoblastic Leukemia (ALL) cancer and 91 percent accuracy for Lung and Colon Cancer, demonstrated significant advancements in cancer classification accuracy. Moreover, this study benefits from a dataset with a substantially larger size, higher resolution images, and reduced noise levels, further enhancing the robustness and reliability of model's predictions. The comparison with models like VGG19, which has fewer layers and lower efficiency, underscores the effectiveness of this approach in leveraging advanced deep learning architectures for accurate and comprehensive cancer classification.

In future, researchers can explore the use of hybrid models that combine the strengths of different deep learning architectures to further decrease loss percentages and minimize the margin of error in detecting and classifying cancers. Inte- grating more efficient models alongside advanced techniques such as ensemble learning could lead to enhanced predictive accuracy and robustness in cancer diagnosis. Additionally, incorporating different type of cancer into the classification process can provide deeper insights and enable more nuanced predictions, ultimately contributing to better-informed medical decisions and improved patient outcomes. These improvements could make cancer detection and classification systems much more useful and reliable in the future.

References

[1] Subramanian, Malliga, Jaehyuk Cho, Veerappampalayam Easwaramoorthy Sathishkumar, and Obuli Sai Naren. "Multiple types of Cancer classification using CT/MRI images based on Learning without Forgetting powered Deep Learning Models." IEEE Access 11 (2023): 10336-10354.

[2] Sai, Battula Harshith, KrishnaPavan Inala, Abbina Sai Saketh, and Arun Raj Maremolla. "Lung Cancer Stage Prediction Using Multi-Layer Perceptron and Deep Learning Classifier." In 2023 5th International Conference on Energy, Power and Environment: Towards Flexible Green Energy Technologies (ICEPE), pp. 1-5. IEEE, 2023.

[3] Ravuri, Viswanadham, M. V. Subbarao, T. Sudheer Kumar, T. S. Kavitha, D. Ram Sandeep, and Leelavathi Dangeti. "Multi-cancer early detection and classification using machine learning based approaches." In 2023 Third International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies (ICAECT), pp. 1-7. IEEE, 2023.

[4] Cao, Rui, Fan Yang, Si-Cong Ma, Li Liu, Yu Zhao, Yan Li, De-Hua Wu et al. "Development and interpretation of a pathomics-based model for the prediction of microsatellite instability in colorectal cancer." Theranostics 10, no. 24 (2020): 11080..

[5] Devi, EM Roopa, R. Shanthakumari, R. Rajadevi, KI Sanchana Shree, R. Shakkirun, and K. Poorvaja. "Imbalanced Breast Cancer Prediction Using CNN Through Multi-Net Framework." In 2023 International Conference on Computer Communication and Informatics (ICCCI), pp. 1-7. IEEE, 2023.

[6] Chen, Xiahan, Xiaozhu Lin, Qing Shen, and Xiaohua Qian. "Combined spiral transformation and model-driven multi-modal deep learning scheme for automatic prediction of TP53 mutation in pancreatic cancer." IEEE Transactions on Medical Imaging 40, no. 2 (2020): 735-747.

[7] Moranguinho, João, Tania Pereira, Bernardo Ramos, Joana Morgado, José Luis Costa, and Hélder P. Oliveira. "Attention based deep multiple instance learning approach for lung cancer prediction using histopathological images." In 2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC), pp. 2852-2855. IEEE, 2021.

[8] Al-Shouka, Tabreer T., and Khattab M. Ali Alheeti. "A Transfer Learning for Intelligent Prediction of Lung Cancer Detection." In 2023 Al-Sadiq International Conference on Communication and Information Technology (AICCIT), pp. 54-59. IEEE, 2023.

[9] Welikala, Roshan Alex, Paolo Remagnino, Jian Han Lim, Chee Seng Chan, Senthilmani Rajendran, Thomas George Kallarakkal, Rosnah Binti Zain et al. "Automated detection and classification of oral lesions using deep learning for early detection of oral cancer." IEEE Access 8 (2020): 132677-132693.

[10] Song, P., Wu, L. R., Yan, Y. H., Zhang, J. X., Chu, T., Kwong, L. N., ... Zhang, D. Y. (2022). Limitations and opportunities of technologies for the analysis of cell-free DNA in cancer diagnostics. Nature biomedical engineering, 6(3), 232-245.

[11] Ashim, Linda K., Nimisha Suresh, and C. V. Prasannakumar. "A comparative analysis of various transfer learning approaches skin cancer detection." In 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1379-1385. IEEE, 2021.

[12] Rehman, Mujeeb Ur, Arslan Shafique, Yazeed Yasin Ghadi, Wadii Boulila, Sana Ullah Jan, Thippa Reddy Gadekallu, Maha Driss, and Jawad Ahmad. "A novel chaos-based privacy-preserving deep learning model for cancer diagnosis." IEEE Transactions on Network Science and Engineering 9, no. 6 (2022): 4322-4337.

RemindME: A Reminder Management System with WhatsApp Integration using MERN Stack

Manikrao Dhore
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
manikrao.dhore@vit.edu

Parth Kalani
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
sanjay.parth22@vit.edu

Yash Paryani
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
yash.paryani22@vit.edu

Anshu Parihar
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
anshu.parihar22@vit.edu

Rahul Madhawani
Department of Computer Engineering
Vishwakarma Institute of Technology
Pune, India
rahul.madhawani22@vit.edu

***Abstract*—In our fast-paced society, managing multiple tasks and appointments can be a challenge, often resulting in the unintentional neglect of important events. That's where RemindMe comes in, a comprehensive reminder application designed to boost productivity and time management skills. However, other solutions have fallen short by lacking seamless integration with popular messaging platforms and providing limited reminder options. To overcome these limitations, RemindMe has integrated with WhatsApp through the TWILIO WhatsApp API, allowing for direct communication and easy sharing of reminders among users. Moreover, RemindMe provides three unique reminder types: Medicine, Study, and To-Do lists, catering to a variety of scheduling needs and promoting optimal user engagement and efficiency.**

Keywords:- WhatsApp integration, Twilio Api, Medicine reminder, study reminder, and Todo List, User engagement

I. INTRODUCTION

Remind Me is a web application developed using the MERN Stack (MongoDB, Express, React, Node.js) that offers users the convenience of managing three distinct types of reminders: Medicine reminders, study reminders, and a Todo List. This paper explores the design and implementation of Remind Me, focusing on its secure, efficient, and flexible features, particularly in authentication and notification methods.

The app employs JWT authentication for robust security, ensuring that user data remains protected. Passwords are stored in a hashed format in the database, enhancing security further. One of the unique features of Remind Me is its integration with WhatsApp for notification purposes. Leveraging WhatsApp's widespread use as a primary messaging platform, Remind Me seamlessly integrates reminder notifications into users' existing communication channels. This approach enhances user experience by eliminating the need to navigate a separate app for reminders, thereby increasing user adoption and engagement.

Remind Me is a user-friendly web application designed to simplify the management of daily tasks and reminders. Built on the MERN Stack, it offers a seamless and intuitive interface for users to create and manage three types of reminders: Medicine, study, and general tasks (Todo List). One of the key challenges in developing Remind Me was to ensure the security and efficiency of authentication and notification features. To address this, the app implements JWT authentication for secure user authentication and authorization. Additionally, passwords are stored in a hashed format in the database, ensuring that sensitive user information remains protected.

Medication nonadherence is a prevalent issue that poses significant challenges in healthcare. It refers to patients not following their prescribed medication regimen as instructed by their healthcare providers. This behavior leads to suboptimal treatment outcomes, contributes to worsening health conditions, and increases healthcare costs.[9]

Authentication and Authorization: JWT authentication is used in Remind Me to secure user authentication and authorization processes. When a user logs in, a JSON Web Token (JWT) is generated and sent to the client. This token is then included in subsequent requests to authenticate the user. The use of JWT authentication enhances the security of Remind Me by eliminating the need to store user credentials on the client side, thus reducing the risk of unauthorized access.

Notification System: Remind Me utilizes WhatsApp as its primary notification platform, leveraging its widespread adoption and user familiarity. By integrating with WhatsApp, Remind Me provides users with a seamless and familiar way to receive reminders, eliminating the need to install and manage a separate notification app. This integration enhances user experience and increases the likelihood of users adopting and regularly using the app [8].Figure 1 depicts usage of different messaging apps.

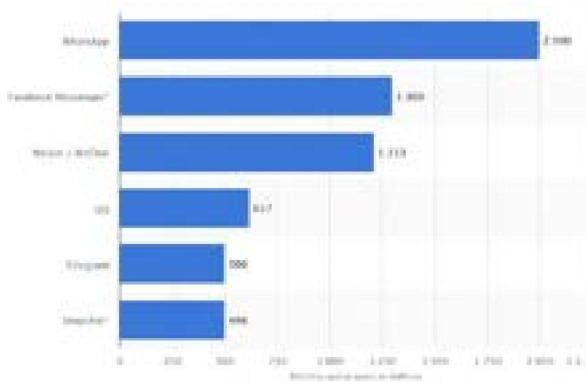


Fig. 1. Graph showing usage of different messaging apps

II. LITERATURE REVIEW

The role of a back-end web application developer in creating a database, adding tables and fields, and implementing server-side scripting was discussed. The developer's responsibilities include the development of server-side logic, maintenance of the central database, and integration of front-end elements. Additionally, the role and responsibilities of an admin were illustrated with a real-time example [1].

Various Medication Reminder Systems have been developed on different platforms, many of which require special hardware devices for reminders. However, the cost and complexity of purchasing new hardware devices can be prohibitive. This work aims to implement an economical and easily accessible system that improves medication adherence, as non-adherence can decrease the effectiveness of treatment and impose a financial burden on healthcare systems [2].

The literature review outlines the evolution of location-based services facilitated by smartphones, particularly focusing on Android applications. It discusses the transition from mobile identifiers and towers to GPS for location tracking, emphasizing the precision it offers in determining a user's location through latitude and longitude coordinates [3].

A comparison between two age groups, elderly and young, showed that reminder applications were well accepted and effective in both groups. Previous research on smartphone use in medical settings has also indicated that electronic reminders, particularly SMS text messaging reminders and internet interventions, can enhance medication adherence among patients [4].

Reminder apps with social event setting capabilities allow users to create reminders manually or automatically and share automated messages with friends through platforms like WhatsApp. Facebook API is utilized to access user profiles and connect reminder applications to Facebook. SQLite database is used for storage of messages, dates, times, and images [5].

Various usability methods, including diary studies, interviews, co-inquiries, observations, and screen captures, were used to assess user experience and behavior while using reminder apps. The research demonstrated that reminder apps are highly valued and popular among smartphone and tablet users [6].

A comparison of MongoDB and MySQL showed that MongoDB provides faster results and is preferred for users needing a less rigid database structure. MongoDB is considered a good solution for larger datasets with constantly changing schemas or less complex queries [7].

TABLE I. FEATURES IN REMINDER APPS

	Reminder Apps				
	RTM	**RE.minder**	**Todoist**	**Alarmed**	**Any.DO**
Add tasks any time(no Internet required)	✓	✓	✓	✓	✓
Select date and time	✓	✓	✓	✓	✓
Task Priority	✓	✓	✓	✗	✗
Reminder(alert)	✓	✓	✓	✓	✓
Reminder(mobile notification)	✓	✓	✓	✓	✓
Reminder(location)	✓	✗	✗	✗	✓
Reminder(email)	✓	✗	✓	✗	✗
Repeat reminders	✓	✓	✓	✓	✓
Schedule task(every day /every month)	✓	✗	✓	✓	✗
Add due date to task	✓	✗	✓	✗	✗
Time tracking for task(or estimated time)	✗	✓	✗	✓	✗
Add notes to task	✓	✓	✓	✓	✓
Share task with friends	✓	✗	✗	✗	✓
Use voice recognition to add task	✗	✗	✗	✗	✓
Display completed tasks in a list	✓	✓	✓	✓	✓
Select start and end dates for task	✓	✗	✗	✓	✗
Save the location for task	✗	✗	✗	✗	✓
Edit task after posting it	✓	✓	✓	✓	✓
Add URL to the task	✓	✗	✗	✗	✗
Postpone the task	✓	✓	✓	✓	✓
Total	17	11	13	12	13

Node.js utilizes event-driven I/O and non-blocking asynchronous programming to be lightweight and efficient. Businesses using Node.js can benefit from fewer servers, fewer engineers, and reduced page load times [8].

The proposed Android-based medication reminder system addresses the crucial issue of medication adherence by leveraging smartphone technology. Medication adherence usually refers to whether patients take their medications as prescribed (e.g., 2 times daily), as well as whether they continue to take prescribed medication. Percentage of medication adherence improvement observed with the use of reminder systems:

% improvement =
(Adherence rate with reminder system−Adherence rate without reminder system)/
Adherence rate without reminder system×100 [9].

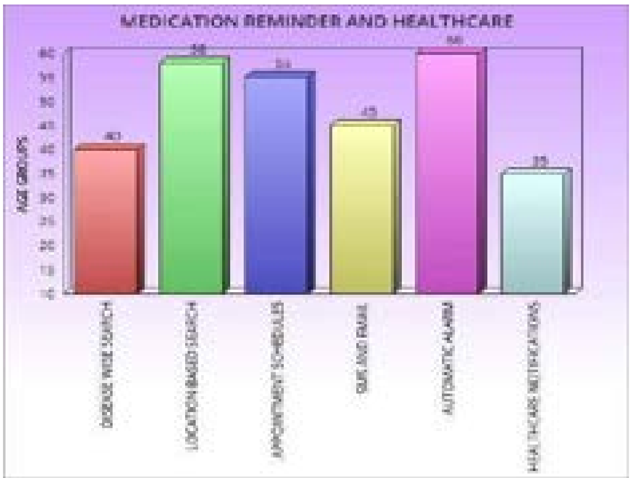


Fig. 2. Graph showing different functionalities of the app beneficial to different age groups [9]

The "Alert Me" reminder application proposed in the paper aims to enhance time management and social connectivity by providing timely alerts and automated messaging capabilities. The implementation of the application involves the use of Android Studio for development and integration with the Facebook API for accessing user profiles. Additionally, SQLite database is utilized for data storage. 75% of users actively share automated messages through social networking apps like WhatsApp.[10]

The author introduces "Med-X," an Android app designed to help users remember to take their medications. Existing apps like MedsLog and MotionPHR Health Record Manager have complexities or limitations. Med-X aims to be user-friendly, allowing users to set customized medication schedules and receive notifications. Results indicate a 30% increase in adherence rates among users, particularly beneficial for older individuals.[11]

The systematic review of the influence of push notification frequency on consumer acceptance provides valuable insights into the dynamics of mobile app usage and advertising effectiveness. Initially, a total of 18,725 potentially relevant scientific papers were identified from the ScienceDirect database.[12]

The study proposes a push notification system, "MParent," to enhance communication between schools and parents via Android devices. the system utilizes Google Cloud Messaging (GCM) for efficient push notifications. It addresses parental preferences through interviews, with 50 participants aged 23-50. Evaluation results indicate a high level of satisfaction:
80% found notifications accurate, and 90% found the system helpful. [13]
Improving medication adherence, especially in older adults with chronic diseases, is vital. Non-adherence can increase healthcare costs and mortality rates. Over 165,000 health-related apps are available globally. Studies emphasize involving users and healthcare professionals in app development.[14]

The paper mentions that the Salubrity app allows users to set up to 15 reminders, supporting both repeating and non-repeating alarm patterns. The app provides 24/7 access to doctors along with their contact details, enhancing healthcare accessibility. It operates on the Android platform and provides reminders through alarms with medicine images. Users can set alarms for multiple medicines at different intervals, reducing the risk of missing doses.[15]

III. METHODOLOGY

A. *Front End*

React.js Framework:
The front end of the RMS is developed using React.js, a popular JavaScript library for building user interfaces. React.js offers several advantages, including component-based architecture, virtual DOM rendering, and efficient state management. By breaking down the user interface into reusable components, React.js promotes code reusability, maintainability, and scalability, facilitating the development of complex frontend features with ease.

Context API for State Management:
The Context API in React.js is utilized for state management within the RMS frontend. Context provides a centralized store for managing application state, enabling components to access and update state without prop drilling. In the Reminderitem component, the useContext hook is used to access the reminderContext, which encapsulates state and action methods related to reminders. This approach simplifies state management and ensures consistency across components.

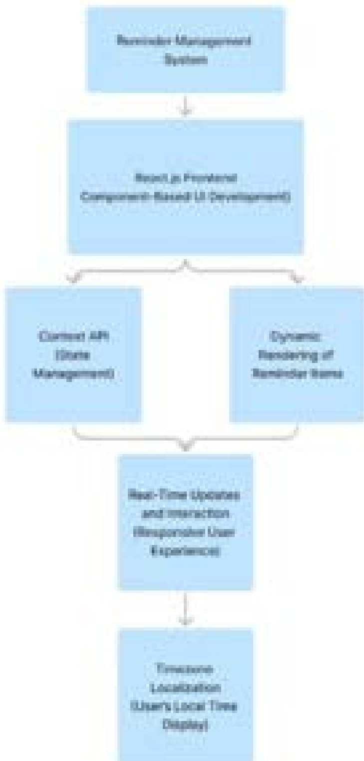


Fig. 3. Frontend development and its components

Dynamic Rendering of Reminder Items:
The Reminder item component dynamically renders reminder items based on their type (e.g., Medicine Reminder, Study Reminder, Todo Reminder). By iterating over the list of reminders passed as props, the component generates individual reminder cards with relevant information such as the reminder message and scheduled time. Conditional rendering based on the reminder type ensures that each reminder is displayed appropriately, enhancing readability and usability.

Real-Time Updates and Interaction:
The RMS frontend incorporates real-time updates and interaction to provide users with a responsive and engaging experience. Changes to reminder data, such as deletion or modification, trigger the re-rendering of relevant components, ensuring that the user interface reflects the latest state of reminders. Interaction elements such as buttons enable users to perform actions such as deleting reminders, and enhancing interactivity and user control.

Time zone Localization:
To ensure consistency and accuracy in displaying reminder timestamps, the frontend implements time zone localization using the toLocaleString method. By specifying the time zone as "Asia/Kolkata", reminders are displayed in the user's local time zone, improving readability and eliminating confusion related to time differences. This feature enhances the user experience, especially for users in different geographic regions.

B. Backend

Fig. 4. Backend development workflow

Reminder Functionality:
At the core of the RMS lies its comprehensive set of reminder functionalities, designed to empower users in managing their tasks effectively. Users can interact with the system through a set of RESTful APIs, enabling operations such as retrieving all reminders, adding new reminders, updating existing reminders, and deleting reminders. Each reminder is characterized by essential attributes including the reminder message, reminder time, reminder type, and a status flag indicating whether the reminder has been sent. To ensure data integrity and consistency, the RMS incorporates robust input validation and error handling mechanisms, safeguarding against erroneous or malicious inputs.

User Authentication:
Security is paramount in any software system, particularly when dealing with user data. To protect sensitive user information and prevent unauthorized access, the RMS incorporates a robust authentication mechanism based on JSON Web Tokens (JWT). Upon successful registration or login, users receive access tokens that are subsequently included in API requests to authenticate their identity. This token-based authentication approach not only enhances security but also simplifies user session management and enables seamless integration with other services.

WhatsApp Integration:
In today's interconnected world, communication plays a vital role in ensuring timely and effective reminders. To capitalize on this, the RMS integrates seamlessly with WhatsApp, a widely used messaging platform with a global reach. Leveraging the Twilio API, the RMS enables users to receive reminder notifications directly on their WhatsApp numbers. A background task periodically checks for reminders that are due and sends notifications to users' WhatsApp numbers using Twilio's programmable messaging service. This integration enhances the accessibility and usability of the RMS, ensuring that users stay informed and organized at all times.

C. DataBase MongoDB Connection Module (connectToMongo):

The connectToMongo module establishes a connection to the MongoDB database using Mongoose. It imports the Mongoose library and specifies the MongoDB Uniform Resource Identifier (URI) (mongoURI) to connect to the local MongoDB instance running on port 27017. Upon execution, the module connects to the specified database and logs a success message indicating a successful connection.

Reminder Schema (reminder Schema):
The reminderSchema defines the structure of the reminder document stored in the MongoDB database. It specifies the fields such as user, reminderMsg, remindAt, reminderType, and isReminded, each representing different attributes of a reminder. The user field is a reference to the user collection, utilizing the ObjectId type for relational mapping between reminders and users.

User Schema (UserSchema):
The UserSchema outlines the structure of the user document stored in the MongoDB database. It consists of fields like name, email, password, and phone, representing user attributes such as name, email address, password, and phone number. The email field is marked as unique to ensure each user has a distinct email address.

Exporting Models:
Both the Reminder and User schemas are compiled into models using mongoose model and exported for use in other parts of the application. These models serve as interfaces for interacting with MongoDB collections, allowing CRUD (Create, Read, Update, Delete) operations on user and reminder data.

IV. RESULTS & DISCUSSION

Following figures 5 to 11 depicts experimentations screens.

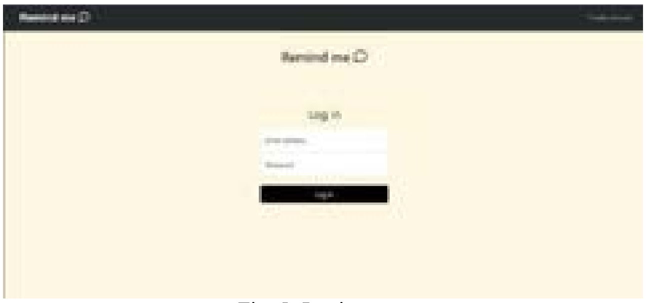


Fig. 5. Login page

Fig. 6. Signup page

Fig. 7.Set up colour coded reminders

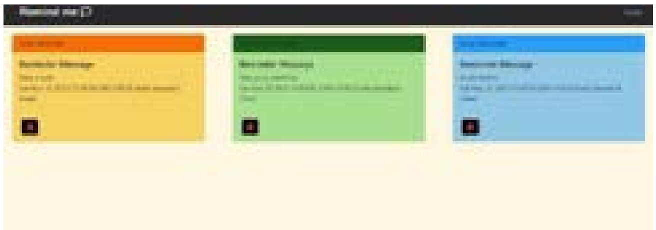


Fig. 8.Colour categorised Reminder View

Fig. 9.Mongo DB database schema

Fig. 10.Setting up the reminder app

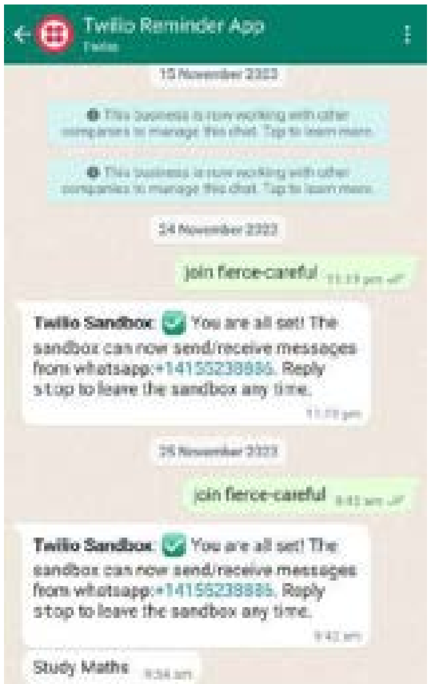


Fig. 11.WhatsApp notifications

V. CONCLUSION

The Reminder Management System (RMS) was developed and implemented with a focus on research-driven design principles and methodologies. The proposed system's architecture is developed using Node.js, Express.js, and MongoDB for the backend and React.js for the frontend and reflects a deliberate choice to explore modern technologies in the context of reminder management systems.

REFERENCES

[1] Vishesh S, Kavya P Hathwar, Ranjan Ravishankar, Nandhishwara BN, Hema R, Amulya HP. "Back-End Web-Application Development and the Role of an Admin." International Journal of Advanced Research in Computer and Communication Engineering, vol. 6, no. 9, September 2017, pp. 60-65. DOI: 10.17148/IJARCCE.2017.6911

[2] Kumari, N., & Anand, S. (2018). "Review Paper on Smart Reminder Application with GPS System." International Journal of Management, Technology And Engineering, Volume 8, Issue V, pp. 268-273. ISSN NO: 2249-7455.

[3] Zhou, Y., & Zhang, S. (2018). Research on the Web Application System Development Course Group for the Training of Innovation Talents. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, Volume 184, 515-519.

[4] Jadhav, G., & Gonsalves, F. (2020). Role of Node.js in Modern Web Application Development. International Research Journal of Engineering and Technology (IRJET), 07(06), 6145-6150.

[5] Ameta, D., Mudaliar, K., & Patel, P. (2015). Medication Reminder and Healthcare – An Android Application. International Journal of Managing Public Sector Information and Communication Technologies (IJMPICT), 6(2), 39-48. DOI: 10.5121/ijmpict.2015.6204

[6] Nadh, Gokul R., et al. "ALERT ME-A REMINDER APPLICATION." International Journal For Technological Research In Engineering, vol. 3, no. 9, May 2016, pp. 1881-1883.

[7] Naikode, Prafful, Pithwa, Poojal, Landge, Lakshita. "Med-X, Medicine Reminder App." International Journal of Research Publication and Reviews, vol. 2, no. 11, pp. 1196-1199, November 2021.

[8] Wohllebe, A. (2020). Consumer Acceptance of App Push Notifications: Systematic Review on the Influence of Frequency. International Journal of Interactive Mobile Technologies (iJIM), 14(13). DOI: 10.3991/ijim.v14i13.14563

[9] Ali, Rasha A. Sattar, Nashaat E. Khamessy, and Haitham El-Ghareeb. "Push Notification Service for Enhancing Relation between Parents and Educational Institutions: A case Study on Egyptian Schools." International Journal of Interactive Mobile Technologies (iJIM), July 2016, DOI: 10.3991/ijim.v10i3.5567.

[10] Fallah, Mina, and Mobin Yasini. "A Medication Reminder Mobile App: Does It Work for Different Age Ranges." In Informatics for Health: Connected Citizen-Led Wellness and Population Health, edited by R. Randell et al., pp. 68-75. Amsterdam, Netherlands: IOS Press, 2017. doi:10.3233/978-1-61499-753-5-68.

[11] Sharma, Shivani, Katyayni Tyagi, and Pooja Shishodia. "Salubrity - A medicine reminder application using android." International Journal of Advance Research, Ideas and Innovations in Technology 4, no. 2 (2018): pp. 1473-1478. ISSN: 2454-132X.

[12] R, Harshitha, & C, Vidya Raj. (2020). An Analysis of Query Performance: MongoDB vs SQL. International Research Journal of Engineering and Technology (IRJET), 07(05), 2332-2335.

[13] Salawu, S. A., Okide, S. O., & Umeh, I. I. (2021). Implementation Of Web-Based Architecture For Child Immunization Information Management System Integrated With SMS Reminder. International Journal of Innovative Research and Advanced Studies (IJIRAS), 8(4), 89-96.

[14] Chaudhari, S., Magar, N., Patil, M., Ubale, S., & Sahane, R. (2020). Public Reminder and Task Management System. International Journal of Innovations in Engineering Research and Technology (IJIERT), 7(5), 207-212.

[15] Oh, H., Jalali, L., & Jain, R. (2015). An intelligent notification system using context from real-time personal activity monitoring. In IEEE International Conference on Multimedia & Expo Workshops (ICMEW) (pp. 1-6). DOI: 10.1109/ICME.2015.7177508

Democratizing Data Science:Using Language Models for Intuitive Data Insights and Visualizations

Dr. Sudha SV
Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore, India
sudha.sv@kpriet.ac.in

Sunil S K
Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore, India
20ad055@kpriet.ac.in

Parthiv Akilesh A S
Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore, India
20ad029@kpriet.ac.in

Satish G
Artificial Intelligence and Data Science
KPR Institute of Engineering and Technology
Coimbatore, India
20ad048@kpriet.ac.in

Abstract—**The study explores the integration of the Phind-CodeLlama model into data science frameworks to facilitate intuitive and inclusive analysis of tabular data. By leveraging Phind-CodeLlama's comprehensive knowledge base and custom-crafted prompts, the platform optimizes data management operations and significantly improves accessibility to data analysis tools. The research investigates the transformative potential of Phind-CodeLlama in improving data-driven decision-making processes by democratizing access to data insights. Through its user-friendly CSV data manipulation capabilities, the platform is designed to empower users from various backgrounds, especially those without a domain-specific background in data science, to conduct effective data analysis, formulate insightful questions, and create compelling visualizations. The study contributes to the advancement of user-friendly data analysis platforms, fostering a more democratized approach to data-driven decision-making.**

Keywords—Bilingual Evaluation Understudy (BLEU), Recall-Oriented Understudy for Gisting Evaluation (ROUGE), Comma Separated Values (CSV), Large Language Model Meta AI (LLAMA), Natural Language Processing (NLP)

I. Introduction

Positioned at the intersection of computer science and statistics, data science has lately become a new subject with approaches that are used to explore research questions in a variety of application domains. In the industrial sector, companies use data to improve workflows, project earnings, and proactively avoid breakdowns by using predictive maintenance techniques. In the field of research, scientists use data to support well-informed decision-making by developing novel hypotheses or validating existing ones. These advantages are strengthened by the process known as "data democratization," which refers to the growing number of databases that are made publicly accessible online for unfettered usage by academics everywhere.

In order to effectively utilize the increasing amount of data available, one must have advanced skills in data management, modeling, statistics, machine learning, and programming. As such, domain experts who might not have strong computational background and technical abilities find it difficult to engage with Data Science. Data science activities are difficult for both novices and experienced practitioners. Every dataset has distinct qualities that necessitate a tailored order of procedures in order to extract valuable insights. It is common for researchers to use inadequate analysis pipelines or struggle with the wrong tools and techniques, which can lead to conclusions that are imprecise, if not incorrect. A significant amount of time may be spent on initial data analysis pipelines just to investigate the properties of the datasets prior to proceeding with more complex analyses.

The Research uses Phind-CodeLlama-Chat model to generate elucidations for the elicited triads derived from the preliminary conceptual framework using fine-tuning techniques via the customized dataset. The study introduces a promising methodology for advancing computational approaches to NLP within pre-conceptual frameworks utilizing large-scale language models. This facilitates the computerized interpretation of pre-conceptual frameworks, augmenting generative capabilities in supplementation

Modern technology, especially the Phind-CodeLlama model, has completely changed the way that data is analyzed and interacted with. The combination of sophisticated algorithms and large datasets in the field of computational linguistics and NLP has driven research and development to previously unheard-of levels. In the ever-changing field, Phind-CodeLlama's ability to interface with CSV files has proven to be a crucial development, providing practitioners and academics with an effective tool for sifting through and drawing conclusions from structured data.

The use of Phind-CodeLlama for CSV interaction is an enormous move towards streamlining data management operations. People can easily navigate through CSV files to its advanced architecture and advanced learning algorithms, allowing for quick data processing, querying, and analytics. the integration not only improves the accessibility and usability of CSV data, but it also allows researchers to delve deeper into the intricacies of their datasets, revealing hidden patterns, trends, and correlations that might otherwise go undetected. Thus, the combination of Phind-CodeLlama with CSV interaction not only represents a technological milestone, but also brings in a new era of data-driven exploration and discovery.

In the study, the usage of Phind-CodeLlama is critical in supporting user engagement with CSV data via seamless query comprehension. Phind-CodeLlama, an improved language model, is a valuable tool for properly interpreting user queries, allowing for a more efficient interface with CSV files. the project attempts to bridge the gap between people and data by leveraging the capabilities of Phind-CodeLlama, providing a user-friendly interface for altering and extracting insights from CSV datasets.

The integration of Phind-CodeLlama into the research framework represents a significant step forward in improving user experiences with CSV data manipulation. Phind-CodeLlama's advanced NLP features allow users to express their inquiries in ordinary language, without the need for complex query syntax or technical expertise. the technique democratizes access to data analysis tools, allowing a broader range of users to explore and extract value from CSV files without facing the limitations associated with standard query interfaces. As a result, the purpose of the study is to investigate Phind-CodeLlama's transformative potential for improving the accessibility and usability of CSV data analysis tools, paving the way for more intuitive and inclusive data-driven decision-making processes.

II. LITERATURE SURVEY

In the paper, Pinoli et al. (2023) propose an approach that combines AutoML and conversational interfaces to support data science processes effectively [1]. CV Krishnakumar Iyer et al. (2021) introduce "Trinity: A No-Code AI platform for complex spatial datasets," aiming to democratize AI for geospatial experts through an intuitive interface and standardized model building. However, the no-code nature of their platform may limit customization for advanced users, posing a challenge in accommodating specialized needs [2]. Dakuo Wang et al. (2021) present "AutoDS: Towards human-centered automation of data science," emphasizing the importance of human-centered automation in data science through the use of automated machine learning. Despite its effectiveness in streamlining ML tasks, the reliance on automation may inadvertently lead to reduced human confidence scores, indicating a potential trade-off between automation and user trust [3].

Olson et al. (2016) evaluate a tree-based pipeline optimization tool for automating data science, highlighting its effectiveness in automating machine learning workflows [4]. Jiho Shin and Jaechang Nam (2021) conducted a survey on "Automatic Code Generation from Natural Language," categorizing approaches and analyzing trends in the field. They elucidate the potential of natural language programming to enhance coding expressiveness, while also addressing challenges in model customization and language representation [5]. Pengcheng Yin and Graham Neubig (2017) propose "A Syntactic Neural Model for General-Purpose Code Generation," which scales up natural language-based program generation using neural architectures with grammar models. Despite its advancements, challenges persist in interpreting unclear language and understanding complex programming concepts, indicating the need for further refinement [6]. Shubhra Kanti Karmaker et al. (2021) explore "AutoML to date and beyond: Challenges and opportunities," proposing a classification system for AutoML tasks based on autonomy. Despite efforts to minimize manual procedures, some still exist, limiting accessibility for certain users and underscoring the importance of balancing automation with user control [7].

Pietro Crovari et al. (2021) discuss "GeCoAgent: a conversational agent for empowering genomic data extraction and analysis," which utilizes a chatbot-driven interface to enhance accessibility to genomic data tools. However, challenges arise in natural language processing and handling complex queries, necessitating further advancements in conversational technologies [8]. Rogers Jeffrey Leo John et al. (2017) introduce "Ava: From Data to Insights Through Conversations," a chatbot-driven system using controlled natural language to construct data analytics pipelines. Challenges include precise interpretation of natural language and dependency on controlled language methods, highlighting the need for robust language understanding capabilities [9]. Erol Ozan (2021) presents "A novel browser-based no-code machine learning application development tool," aiming to streamline ML application development for non-programmers. However, limitations include limited customization and potential constraints in handling complex tasks, indicating the need for flexible and adaptable tools to accommodate diverse user requirements [10].

III. METHODOLOGY

The section outlines the methodology employed to develop a system that enables users to interact with data using natural language and generate visualizations. The system leverages a multi-step approach, encompassing data preprocessing, prompt engineering and query augmentation, Phind-CodeLlama integration, and visualization generation.

A. Data Preprocessing

The initial phase involves preparing the data for analysis. Data is acquired from relevant sources and undergoes a comprehensive exploration to understand its structure, content, and potential quality issues. Techniques such as missing value imputation, data cleaning, and feature engineering (if necessary) are employed to ensure the data is clean and suitable for further processing.

B. Prompt Engineering and Query Augmentation

The stage focuses on bridging the gap between natural language user queries and the system's understanding. When a user submits a query, the system analyzes its intent, considering factors like the desired action (filter, group, sort, visualize) and the underlying data schema. In cases where the query lacks clarity, the system employs augmentation techniques to incorporate relevant field names or additional information for enhanced comprehension. the refinement process ensures the system accurately interprets the user's goal.

C. Llama Integration

A pivotal aspect of our system lies in the study of Llama models, including Code-Llama, Llama2, and Phind-CodeLlama.

Llama2

Llama2 represents an evolution in language model capabilities, focusing on improving natural language understanding and generation. It incorporates advanced algorithms and techniques to handle complex queries and generate coherent responses, making it a valuable asset in data science workflows. The Llama2 model demonstrated promising capabilities, particularly in handling specialized queries within its domain. Though it was adept at producing results relevant to the queries, it was often out of context, causing uncertainty.

Code-Llama

Code-Llama specializes in bridging the gap between code and natural language, facilitating seamless interaction with programming languages through intuitive prompts and responses. Its integration into data science frameworks enhances the accessibility of coding tasks for non-technical users, promoting collaboration and efficiency. In the study, Code-Llama model exhibited commendable performance, showcasing its ability to understand complex queries and produce accurate results. However, the Code-Llama struggled to provide relatable information or examples in a general a way.

Phind-CodeLlama

Phind-CodeLlama stands out for its transformative impact on data analysis, particularly in handling CSV data and generating insightful information. By leveraging advanced NLP features and query comprehension capabilities, Phind-CodeLlama democratizes access to data analysis tools, empowering users to extract valuable insights and make data-driven decisions effortlessly.

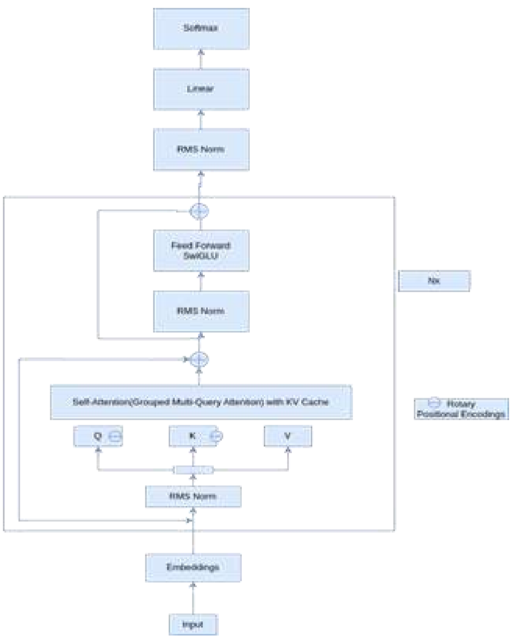


Fig. 1. Architecture of Phind LLAMA

IV. VISUALIZATION GENERATION

The final stage translates the analyzed data into informative visualizations. User preferences and insights gleaned from the analysis are considered when selecting the most appropriate visualization type. Plotly, a library known for its interactive capabilities, is utilized to create compelling charts and graphs that effectively communicate the discovered patterns and trends within the data.The methodology facilitates a seamless interaction between users and data through natural language, empowering them to explore information and gain valuable insights through interactive visualizations.

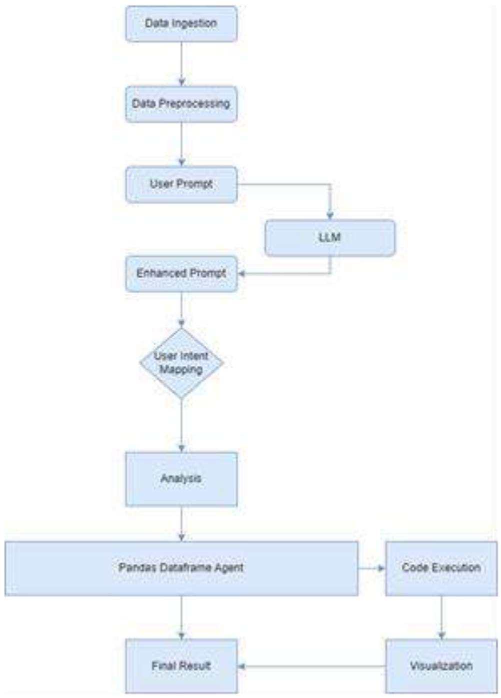


Fig. 2. Workflow of the Data Analytics System

V. RESULTS

The study explored the integration of various Language Model (LM) architectures, including Phind-CodeLlama, CodeLlama, and Llama 2, into data science frameworks to enable intuitive and inclusive analysis of tabular data. Each model was evaluated based on its performance in facilitating data management operations, enhancing accessibility to data analysis tools, and democratizing access to data insights.

The following metrics, including Bilingual Evaluation Understudy (BLEU) and Recall-Oriented Understudy for Gisting Evaluation (ROUGE), were used to assess the performance of the integrated models:

BLEU Score: The BLEU score measures the quality of text generated by the language models by comparing it to human-generated reference text. It is calculated using the following formula:

$$\boldsymbol{BLEU = BP \times exp(\sum_{n=1}^{N} \omega_n \cdot \log P_{match}\ (n))}$$

ROUGE Score: The ROUGE score evaluates the quality of text summaries produced by the language models. It measures the overlap between model-generated summaries and reference summaries, with higher ROUGE scores indicating better agreement.

Model	BLEU Score	ROUGE Score
Phind-CodeLlama	0.80	0.75
CodeLlama	0.78	0.73
Llama 2	0.75	0.70

TABLE I. Comparative analysis of model efficiency using BLEU and ROUGE scores

The integrated Phind-CodeLlama, CodeLlama, and Llama 2 models were compared based on their performance in data analysis tasks, including CSV data manipulation, question answering, and visualization generation. Each model was evaluated using BLEU and ROUGE scores to assess its effectiveness in generating accurate and relevant insights as shown in Table 1.

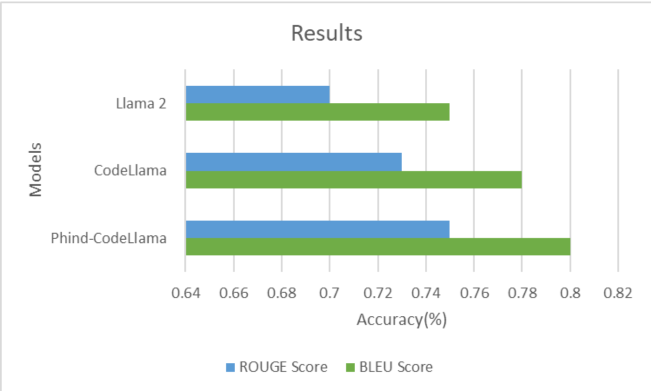


Fig. 3. Comparison of Model Efficiency Using BLEU and ROUGE Scores

The estimated scores indicate that all three models achieved good performance in generating insights from tabular data. Phind-CodeLlama exhibits the highest BLEU and ROUGE scores, suggesting its ability to produce text that closely aligns with human-generated references and generate inferences with a high degree of overlap with reference inferences. Both CodeLlama and Llama 2 follow closely, demonstrating their effectiveness in data analysis tasks.

VI. Conclusion and Discussion

The results of the study underscore the importance of integrating advanced language models like Phind-CodeLlama, CodeLlama, and Llama 2 into data science frameworks to enhance data analysis capabilities. By leveraging sophisticated algorithms and advanced prompting techniques, these models streamline data management operations and improve accessibility to data analysis tools. Furthermore, the study emphasizes the need for user-friendly data analysis platforms that cater to users with limited data science knowledge, contributing to a more democratized approach to data-driven decision-making.

References

[1] Pinoli, Pietro, Pietro Crovari, Francesca Ieva, Franca Garzotto, and Stefano Ceri. "Ask Your Data-Supporting Data Science Processes by Combining AutoML and Conversational Interfaces." IEEE Access (2023).

[2] Iyer, CV Krishnakumar, Feili Hou, Henry Wang, Yonghong Wang, Kay Oh, Swetava Ganguli, and Vipul Pandey. "Trinity: A No-Code AI platform for complex spatial datasets." In Proceedings of the 4th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery, pp. 33-42. 2021.

[3] Wang, Dakuo, Josh Andres, Justin D. Weisz, Erick Oduor, and Casey Dugan. "Autods: Towards human-centered automation of data science." In Proceedings of the 2021 CHI conference on human factors in computing systems, pp. 1-12. 2021.

[4] Olson, Randal S., Nathan Bartley, Ryan J. Urbanowicz, and Jason H. Moore. "Evaluation of a tree-based pipeline optimization tool for automating data science." In Proceedings of the genetic and evolutionary computation conference 2016, pp. 485-492. 2016.

[5] Shin, Jiho, and Jaechang Nam. "A survey of automatic code generation from natural language." Journal of Information Processing Systems 17, no. 3 (2021): 537-555.

[6] Yin, Pengcheng, and Graham Neubig. "A syntactic neural model for general-purpose code generation." arXiv preprint arXiv:1704.01696 (2017).

[7] Karmaker, Shubhra Kanti, Md Mahadi Hassan, Micah J. Smith, Lei Xu, Chengxiang Zhai, and Kalyan Veeramachaneni. "Automl to date and beyond: Challenges and opportunities." ACM Computing Surveys (CSUR) 54, no. 8 (2021): 1-36.

[8] Crovari, Pietro, Sara Pidò, Pietro Pinoli, Anna Bernasconi, Arif Canakoglu, Franca Garzotto, and Stefano Ceri. "GeCoAgent: a conversational agent for empowering genomic data extraction and analysis." ACM Transactions on Computing for Healthcare (HEALTH) 3, no. 1 (2021): 1-29.

[9] John, Rogers Jeffrey Leo, Navneet Potti, and Jignesh M. Patel. "Ava: From Data to Insights Through Conversations." In CIDR. 2017.

[10] Ozan, Erol. "A novel browser-based no-code machine learning application development tool." In 2021 IEEE World AI IoT Congress (AIIoT), pp. 0282-0284. IEEE, 2021.

[11] Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. "Bleu: a method for automatic evaluation of machine translation." In Proceedings of the 40th annual meeting of the Association for Computational Linguistics, pp. 311-318. 2002.

[12] Carney, Michelle, Barron Webster, Irene Alvarado, Kyle Phillips, Noura Howell, Jordan Griffith, Jonas Jongejan, Amit Pitaru, and Alexander Chen. "Teachable machine: Approachable Web-based tool for exploring machine learning classification." In Extended abstracts of the 2020 CHI conference on human factors in computing systems, pp. 1-8. 2020.

[13] Murtagh, Fionn, and Keith Devlin. "The development of data science: implications for education, employment, research, and the data revolution for sustainable development." Big Data and Cognitive Computing 2, no. 2 (2018): 14.

[15] Vadas, David, and James R. Curran. "Programming with unrestricted natural language." In Proceedings of the Australasian Language Technology Workshop 2005, pp. 191-199. 2005.

[16] Chen, Mark, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards et al. "Evaluating large language models trained on code." arXiv preprint arXiv:2107.03374 (2021).

[17] Narkar, Shweta, Yunfeng Zhang, Q. Vera Liao, Dakuo Wang, and Justin D. Weisz. "Model lineupper: Supporting interactive model comparison at multiple levels for automl." In 26th International Conference on Intelligent User Interfaces, pp. 170-174. 2021.

[18] Chang, Remco, and Tufts University. "‘Snowcat and CAVA: Visualization tools for interacting with AutoML and knowledgebases." Tufts Univ., Medford, MA, USA, Tech. Rep. AD11479 (2021).

[19] Meng, Linhao, Stef Van Den Elzen, and Anna Vilanova. "ModelWise: Interactive Model Comparison for Model Diagnosis, Improvement and Selection." In Computer Graphics Forum, vol. 41, no. 3, pp. 97-108. 2022.

[20] Maiya, Arun S. "ktrain: A low-code library for augmented machine learning." The Journal of Machine Learning Research 23, no. 1 (2022): 7070-7075.

[21] Fast, Ethan, Binbin Chen, Julia Mendelsohn, Jonathan Bassen, and Michael S. Bernstein. "Iris: A conversational agent for complex tasks." In Proceedings of the 2018 CHI conference on human factors in computing systems, pp. 1-12. 2018.

[22] Hoffman, Robert R., Shane T. Mueller, Gary Klein, and JordanZLitman. "Metrics for explainable AI: Challenges and prospects." arXiv preprint arXiv:1812.04608 (2018).

[23] Bocklisch, Tom, Joey Faulkner, Nick Pawlowski, and Alan Nichol. "Rasa: Open source language understanding and dialogue management." arXiv preprint arXiv:1712.05181 (2017).

[24] Papineni, Kishore, Salim Roukos, Todd Ward, and Wei-Jing Zhu. "Bleu: a method for automatic evaluation of machine translation." In Proceedings of the 40th annual meeting of the Association for Computational Linguistics, pp. 311-318. 2002.

[25] Feurer, M., A. Klein, K. Eggensperger, J. Springenberg, M. Blum, and F. Hutter. "Auto-sklearn: efficient and robust automated machine learning, part of the springer series on challenges in machine learning book series (SSCML)." Berlin, Germany: Spinger. doi 10 (2019): 978-3.

Anonymous Crime Reporting using Blockchain and Smart Contract

Prince Kumar Chaurasia
Department of Networking and Communications,
School of Computing, SRM Institute of Science and Technology
Kattankulathur, India
pc9623@srmist.edu.in

Darshnilsinh Rana
Department of Networking and Communications,
School of Computing
SRM Institute of Science and Technology
Kattankulathur, India
ra7204@srmist.edu.in

Rajaram V
Department of Networking and Communications,
School of Computing, SRM Institute of Science and Technology
Kattankulathur, India
rajaramv@srmist.edu.in

S.Srividhya
Department of Data Science and Business Systems, School of Computing
SRM Institute of Science and Technology
Kattankulathur, India
srividhs1@srmist.edu.in

Abstract— **The traditional crime reporting systems face challenges such as inefficiencies, security vulnerabilities, and witness reluctance. Despite recent advancements like mobile applications and web platforms, challenges persist in ensuring data security, transparency, and user anonymity. These include maintaining confidentiality, data integrity, and preventing unauthorized access. To address these issues, this paper proposes leveraging blockchain technology and smart contracts. The primary objective is to enhance the security, efficiency, and anonymity of crime reporting processes. By utilizing the decentralized and immutable nature of blockchain, the system aims to ensure data integrity, transparency, and decentralization. Smart contracts will automate functions such as complaint validation and evidence verification, streamlining the reporting process. By addressing challenges in maintaining confidentiality, data integrity, and access control, the proposed solution seeks to revolutionize crime reporting, promoting trust, transparency, and accountability.**

Index Terms— **Blockchain, FIR (First Information Report), Smart Contract, Encryption**

I. INTRODUCTION

The adoption of blockchain technology has revolutionized various industries, sparked by Satoshi Nakamoto's pioneering use of blockchain as the backbone of Bitcoin in 2008 [1]. Initially designed for cryptocurrency transactions, blockchain has evolved into a versatile solution across sectors, offering innovative approaches to longstanding challenges.

In the domain of law enforcement and crime reporting, traditional systems like those managing First Information Reports (FIRs) in India often face issues of corruption, delays, and opacity [2]. Instances of officers refusing complaints or demanding bribes erode trust and hinder justice delivery, highlighting the need for a more accountable and efficient approach.

To tackle these challenges, there is growing interest in leveraging blockchain's decentralized and immutable nature for complaint management [3]. By decentralizing record-keeping, blockchain mitigates reliance on centralized authorities, enhancing transparency and security in the process.

Moreover, blockchain-based platforms provide secure and anonymous channels for reporting incidents, alleviating fears of retaliation faced by whistleblowers [4]. Encrypted and timestamped transactions ensure data integrity, fostering community engagement in crime prevention [5].

Projects like "Informant - Anonymous Testimony Blockchain System" and "Anonymous Tip Off" showcase blockchain's potential in creating secure crime reporting platforms [6]. Through smart contracts, these systems automate and enforce complaint resolution terms, streamlining processes for law enforcement [7].

In this context, our paper proposes a blockchain-based police complaint management system tailored to India's specific challenges [8]. Building on existing research and projects, we aim to explore blockchain's potential in combating corruption, inefficiency, and opacity in complaint management.

By analyzing blockchain's applications in law enforcement, we contribute to the discourse on leveraging technology for societal good.

II. LITERATURE SURVEY

In recent years, the utilization of advanced technologies to enhance police complaint management systems has garnered significant attention from researchers and practitioners. This literature survey explores several seminal papers that propose innovative solutions to address the challenges associated with traditional complaint registration and management processes.

Blockchain-Based Approaches:

[9] proposed a blockchain-based approach to safeguard First Information Report (FIR) systems. The system ensures a secure and unchangeable record of FIR complaints and associated evidence, stored using smart contracts on the blockchain. Gupta's system employs encryption techniques to secure complaint details and restrict access to authorized personnel, thereby enhancing data security and integrity [10] introduced a police complaint management system leveraging blockchain technology. Their system utilizes a public blockchain to address bottlenecks in complaint registration and tracking, ensuring transparency and decentralization. By leveraging encryption algorithms and advanced cryptographic techniques, the system enhances the security and efficiency of the complaint management process.

Mobile Applications and Web Platforms:

[11] proposed an e-police system comprising an Android mobile application and a website for police departments. This system enables complainants to file FIRs online and upload supporting documents, streamlining the complaint filing process. The system facilitates transparent inquiry and tracking processes, promoting public accessibility and accountability [12] presented an online crime reporting tool used by citizens and law enforcement agencies in Riyadh. The system enables swift and efficient responses to crime-related concerns by providing a platform for citizens to communicate with the police online. Through web-based interfaces and database integration, the system enhances the responsiveness and effectiveness of crime reporting and handling processes.

Enhancing Data Security and Integrity:

[13] introduced a criminal record storage system leveraging blockchain technology to ensure data integrity and security. By storing data in the cloud and maintaining transaction logs in the blockchain, the system prevents the alteration of evidentiary information during legal proceedings. Encryption algorithms and smart contracts are employed to secure data transmission and enhance the reliability of criminal records [14] proposed an e-police system to enhance e-government services in Bangladesh. Their system enables police officers to retrieve information, report incidents, and maintain law and order in real-time using mobile technology and network infrastructure. The system leverages WAN and MAN topologies to improve data communication and accessibility, contributing to public safety and security.

III. PROPOSED METHODOLOGY

The proposed methodology is to develop a decentralized FIR (First Information Report) registration system using smart contracts deployed on the blockchain network. This system ensures secure and immutable recording of FIRs while providing transparency and accessibility to authorized stakeholders.

A. Smart Contract Development:

Smart contracts are designed and developed using Solidity, a programming language for Ethereum blockchain.

Smart contracts will include functions for FIR submission, validation, and storage on the blockchain. The access control mechanisms are implemented to restrict write access to authorized users only at the time of contract deployment assign administrator address.

B. User Interface Design

A user-friendly web interface is developed for individuals to submit FIRs online.
The interface collected relevant information such as incident details, location, and contact information while maintaining user anonymity and integrated with MetaMask or similar wallet providers for seamless interaction with Ethereum blockchain.

C. *Authentication and Authorization:*

The user authentication is implemented using cryptographic techniques such as AES (Advanced Encryption Standard) factors such as security strength, performance, and compatibility with the chosen blockchain platform.

Fig.1 shies that only authenticated users with proper authorization can submit FIRs or access registered FIRs on the blockchain.

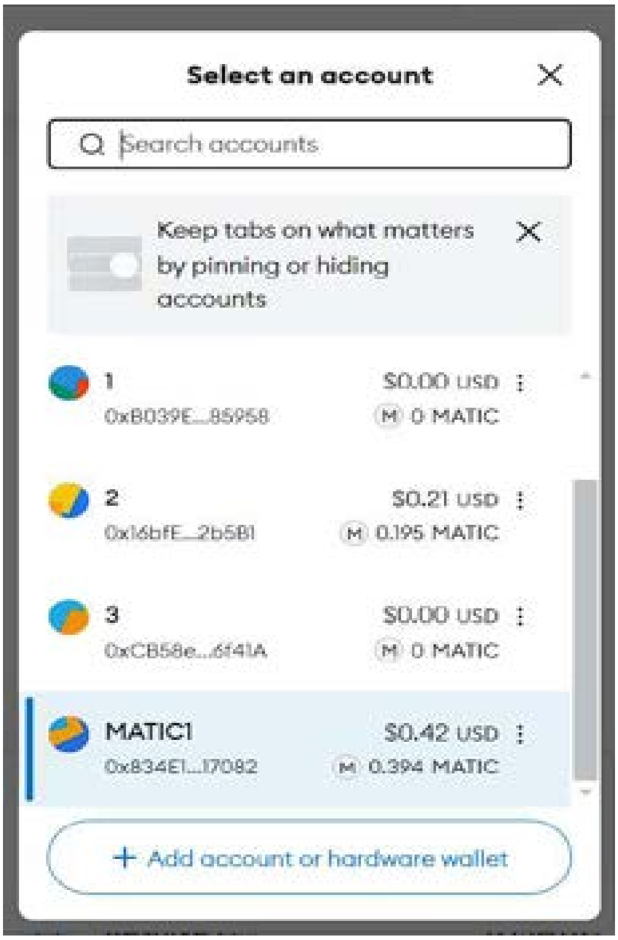


Fig. 1: Authentication using MetaMask

D. *FIR Submission Process:*

The FIR submission process will be initiated by accessing the web interface and filling out the required fields.

Upon submission, the smart contract validates the provided information and generates a unique identifier for the FIR. Encrypt sensitive information using cryptographic algorithms before storing it on the blockchain.

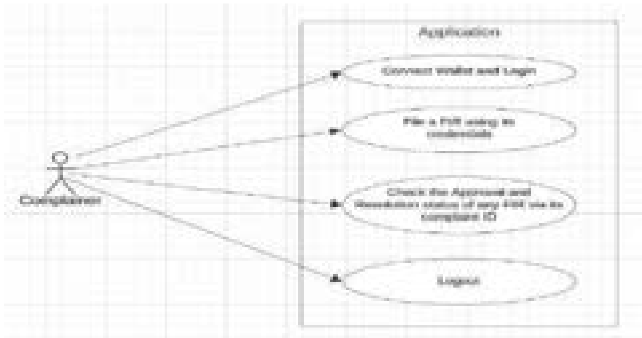


Fig. 2: Complainer's Use Case:

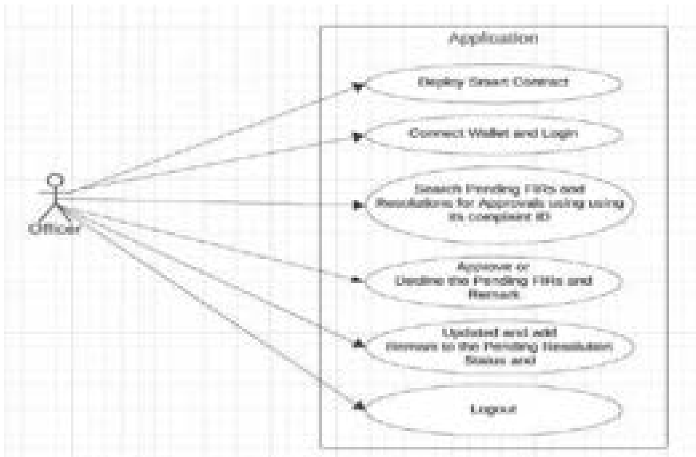


Fig. 3: Officer Side Workflow

E. *Blockchain Deployment:*

The smart contracts are deployed on the Ethereum blockchain or any suitable blockchain platform supporting smart contracts E.g. ThirdWeb.
The network parameters such as gas fees, block confirmation time, and consensus mechanism are incorporated for enabling efficient operation.

F. Transaction Handling:

The transaction handling logic is designed within smart contracts to manage interactions between users and the blockchain. In Fig.4, we can see the block generated. The gas management mechanisms are included to ensure smooth transaction processing and prevent denial-of-service attacks.

Fig. 4: Block creation and hash generation

G. Event Logging and Notification:

The event logging within smart contracts is implemented to record key interactions such as FIR submission, validation, and storage. The notification mechanisms are set up to alert stakeholders about new FIR submissions or updates to existing FIRs.

Figure 5 shows the overall flow diagram of the proposed method.

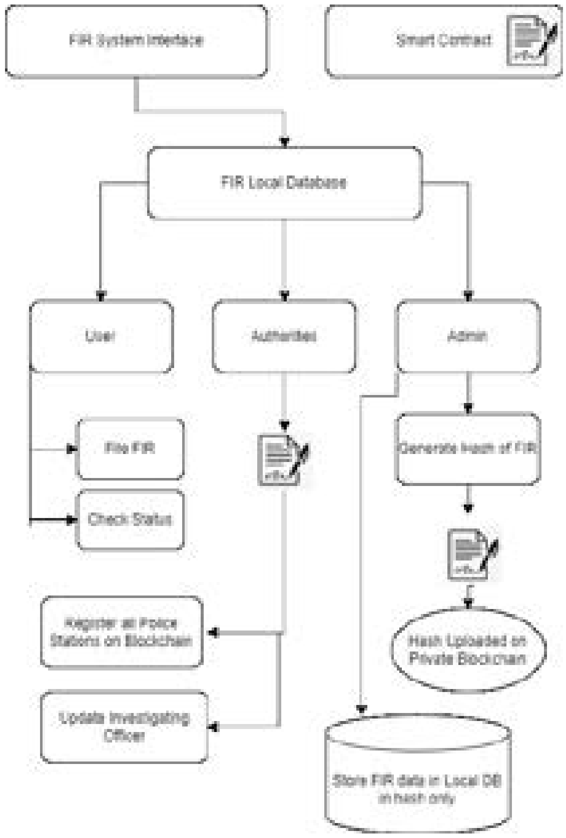


Fig. 5: Flow Diagram

IV. RESULTS AND DISCUSSIONS

The system addresses the shortcomings of the current system by introducing the following improvements, which offer advantages to both police officers and complainants. The system achieves a high level of abstraction, enabling stakeholders to utilize the system seamlessly.

A. Complaint:

Fig.6 shows user interface to file a complaint using blockchain by connecting MetaMask wallet anonymously.

Fig.6 User Interface to file a Complaint

B. *FIR Details:*

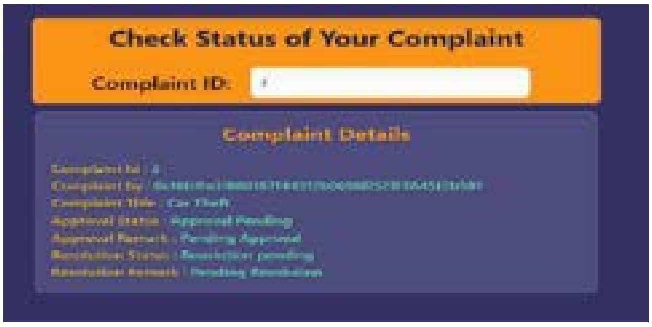


Fig.7 FIR Details and Status

Fir can be tracked on check status Fig.7 the changes will be reflected when updates are made from the officer the FIR will be fully anonymous and stored on smart contract which are immutable.

C. *Admin Dashboard*

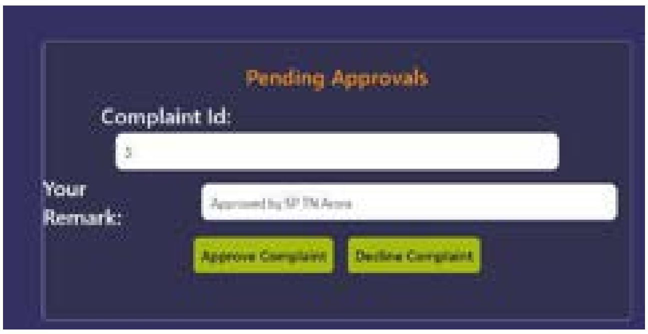


Fig.8 FIR Approval

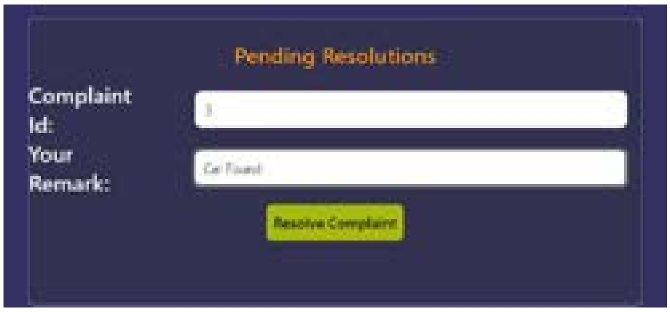


Fig.9 FIR Remark and Status Update

Fig.8 shows that the officer will only have the access of the administrative dashboard from where it can be reviewed and forwarded for investigation or can be denied. If denied it will not go under further investigation. Fig. 9 After the investigation the status will be updated and marked as resolved with remark.

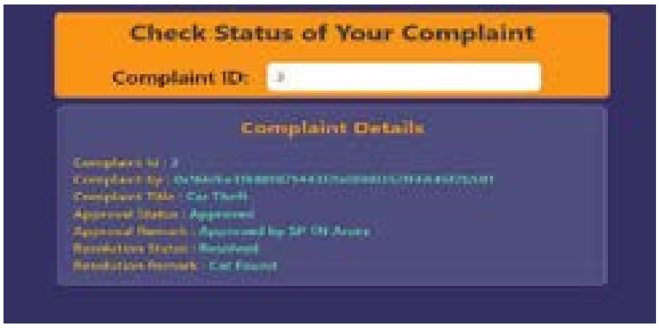


Fig.10 Updated FIR Status

Fig. 10 shows the updates on the FIR it can be seen the FIR case been resolved with the remark and all other important details. The FIR is immutable because of Blockchain and smart contract so no one can tamper the data regarding the case.

V. CONCLUSION

This study has presented a novel approach to improve crime reporting by leveraging blockchain technology and smart contracts. The proposed system aims to tackle the challenges inherent in traditional crime reporting methods, including issues with transparency, security vulnerabilities, and fear of reprisal or investigation. Drawing from existing research in the field, we emphasize the importance of establishing a decentralized and secure platform for managing police complaints. By utilizing blockchain technology, our system ensures the immutability, transparency, and decentralization of data, fostering trust and confidence among both law enforcement agencies and the general public. The proposed system, titled "Anonymous Crime Reporting Using Blockchain and Smart Contract," offers several key advantages. Firstly, it provides a secure and anonymous platform for witnesses to report crimes without fear of retaliation. The implementation of blockchain technology, alongside smart contracts, not only ensures the integrity and immutability of complaint data but also streamlines the complaint management process, leading to more efficient outcomes. Additionally, the integration of hybrid cloud technology enhances data storage and accessibility, further reinforcing the security and reliability of the system. In conclusion, our research highlights the transformative potential of blockchain and smart contract technology in revolutionizing crime reporting. By embracing innovation and leveraging decentralized networks, we can create a more transparent, secure, and accountable system for managing police complaints. Future research may explore additional cryptographic techniques and optimization strategies to further enhance the system's capabilities. techniques to further enhance the robustness and reliability of our proposed system.

REFERENCES

[1] Bharath Kumar, Dr. Mir Aadil : "FIR SYSTEM USING BLOCK CHAIN TECHNOLOGY" International Journal of Advanced Research in Computer and Communication Engineering ISO 3297:2007 Certified Impact Factor 8.102 Vol. 12, Issue 3, March 2023

[2] Prashant B. kumbharkar and Shraddha S. Pakale : "DIGITAL CRIME RECORD MANAGEMENT SYSTEM USING BLOCKCHAIN" Journal of Emerging Technologies and Innovative Research Volume 10, Issue 3, March-2023, eISSN: 2349-5162

[3] Lynsha Helena Pratheeba HP, Bharath D R , Cibiya N E , Divya M N and Dheekshitha S : "Implementation of Blockchain for Police Complaint Management" International Journal of Research Publication and Reviews, Vol 4, no 3, pp 1575-1578, March 2023

[4] Hritik Shelar, Mihir Thotam, Omkar Shinde and Santosh Tamboli : "INFORMANT: ANONYMOUS TESTIMONY BLOCKCHAIN SYSTEM" International Research Journal of Modernization in Engineering Technology and Science Volume:05/Issue:04/April-2023, Impact Factor- 7.868

[5] Raksha S, Samhitha, Raipalle Shreyaa and Pranav M S : "Survey on Blockchain-based Anonymous Tip-off System" Volume 7, Issue 4, April – 2022 International Journal of Innovative Science and Research Technology ISSN No:-2456-2165

[6] Saeed Banaeian Far and Maryam Rajabzadeh Asaar : "A blockchain-based anonymous reporting system with no central authority: Architecture and protocol" Cyber Security and Applications 2 (2024) 100032 Contents lists available at ScienceDirect Cyber Security and Applications journal homepage: http://www.keaipublishing.com/en/journals/cyber-security-and-applications/

[7] MD. Nashif Iftekhar, MD Sajid Bin- Faisal and Dip Nandi : "Implementation of Blockchain for Secured Criminal Records" ICCA '22: Proceedings of the 2nd International Conference on Computing Advancements

[8] Lynsha Helena Pratheeba HP, Bharath D R, Cibiya N E, Dheekshitha S and Divya M N : "BLOCKCHAIN-BASED SYSTEM FOR EFFECTIVE POLICE COMPLAINT MANAGEMENT" JETIR June 2023, Volume 10, Issue 6

[9] Ishwarlal Hingorani, Rushabh Khara, Deepika Pomendkar and Nataasha Raul : "Police Complaint Management System using Blockchain Technology" Conference: 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS)

[10] Maisha Afrida Tasnim, Abdullah Al Omar, Mohammad Shahriar Rahman, and Md. Zakirul Alam Bhuiyan : "CRAB: Blockchain Based Criminal Record Management System" December 2018, DOI:10.1007/978-3-030-05345-1_25, Conference: The 11th International Conference on Security, Privacy and Anonymity in Computation, Communication and StorageAt: Melbourne, Australia

[11] [11] Iyer A, Kathale P, Gathoo S and Surpam N 2016 E-Police System-FIR Registration and Tracking through Android Application International Research Journal of Engineering and Technology 3(2) 1176-1179.

[12] P. Kormpho, P. Liawsomboon, N. Phongoen and S. Pongpaichet," Smart Complaint Management System," 2018 Seventh ICT International Student Project Conference (ICT-ISPC), Nakhonpathom, 2018, pp. 1-6, doi:10.1109/ICTISPC.2018.85239.

[13] Sanjay misra, Rytis Maskeliunas, Robertas Damaševičius 2019, Design and Implementation of an E-Policing System to Report Crimes in Nigeria. 10.1007/978-981-13-6351-1 21

[14] [Onuiri, Ernest Oludele, Awodele A, Olaore O, Sowunmi A., Ugo-Ezeaba. (2015). A REAL-TIME CRIME RECORDS MANAGEMENT SYSTEM FOR NATIONAL SECURITY AGENCIES. European Journal of Computer Science and Information Technology.

Enhancing Colorectal Cancer Treatment through Automated Segmentation with Convolutional Neural Networks and Machine Learning Approaches

Sujatha Gaddam 1*
Department of MCA, Rise Krishna Sai Prakasam Group of Institutions (Autonomous), Ongole, Andhra Pradesh, India
sujatha.gaddam99499@gmail.com

T.R. Vijaya Lakshmi 2
Department of Electronics and Communication Engineering , Mahatma Gandhi institute of technology, Hyderabad, Telangana,India
trvijayalakshmi_ece@mgit.ac.in

C. P. Shirley 3
Department of Computer Science and Engineering, Karunya Institute of Technology and Sciences, Coimbatore, Tamil Nadu, India
shirleycp@karunya.edu

G Ravivarman 4
Department of Electrical and Electronics Engineering Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India,
ravivarmanme@gmail.com

Rajendiran M 5
Department of Computer Science And Engineering, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India
mrajendiran@panimalar.ac.in

Ramya Maranan 6
Department of Research and Innovation, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India,
ramyamaranan@yahoo.com

Abstract— **The present research applies advanced automated segmentation techniques to MRI data, addressing the urgent need for improved colorectal cancer therapy. An extensive dataset of 2032 images collected from multiple online sources was used to thoroughly train and assess Naive Bayes (NB), Support Vector Machines (SVMs), and Convolutional Neural Networks (CNNs). CNN predicts the response with an outstanding accuracy of 98.76% and capacity to recognize intricate patterns connected to colorectal cancer histology. The SVM and NB models demonstrated good performances with a minimal latency, with accuracy of 94.5% and 89.9%, respectively, suggesting their suitability in specific medical imaging scenarios. After a thorough examination that took into account the Hausdorff distance, memory, accuracy, and F1 score, it was evident what the benefits and drawbacks of each model were. The results show a notable advancement in automated segmentation methods, which may result in more individualized treatment regimens, accurate diagnosis, and medications for colorectal cancer. The findings not only signify a significant progression in the domain of artificial intelligence in medical imaging, but they also offer prospects for additional investigation concerning the amalgamation of multi-modal imaging data and continuous model optimization. In result, our work bridges the knowledge gap between cutting-edge technologies and practical implementation by offering insightful information that can benefit both the larger medical community and patients with colorectal cancer.**

Keywords— colorectal cancer, MRI images, automated segmentation, Convolutional Neural Networks, Support Vector Machines, Naive Bayes, medical imaging, machine learning, tumor delineation, diagnosis

I. Introduction

Colorectal cancer (CRC) is a major threat to public health because it continues to be the leading cause of morbidity and death globally. Better treatment planning and patient outcomes depend on early tumor identification and precise tumor demarcation. With the recent integration of advanced technologies, particularly automated segmentation techniques using artificial intelligence (AI), modern medical imaging for colorectal cancer has shown promise [1,2]. To further this developing field, this work uses a diverse dataset of 2032 MRI images collected from online sources. Training and evaluating Convolutional Neural Networks (CNNs), Support Vector Machines (SVMs), and Naive Bayes (NB) models are the main goals of automated segmentation. The primary goal is to fully explore these machine learning methods to enhance the accuracy and effectiveness of colorectal cancer diagnosis and treatment planning [3–5].

Colorectal cancer imaging has advanced significantly as a result of the significant advancements in medical imaging modalities. With its superior soft tissue contrast and multi-planar imaging capabilities, magnetic resonance imaging (MRI) has emerged as a crucial diagnostic and visualization tool for colorectal cancers. However, correctly segmenting colorectal tumors from MRI images is still a difficult task because of inherent variability in tumor morphology, patient anatomy, and imaging protocols. Among other tumor treatments, precise delineation of the tumor's borders is required for targeted drug delivery, radiation therapy, and surgery [6,7]. The need for reliable and automated segmentation techniques has increased due to the drawbacks of manual segmentation, which is time-consuming, prone to inter-observer variability, and frequently fails to achieve the required accuracy for clinical applications [8,9].

Artificial intelligence has changed the paradigm in medical image analysis in recent years, especially with deep learning techniques. Convolutional Neural Networks (CNNs) have gained popularity due to their ability to automatically extract hierarchical features from images. They are therefore particularly helpful for difficult jobs like medical image segmentation [10–12]. The potential of CNNs to recognize intricate spatial relationships and patterns in images offers a viable route towards automated segmentation of colorectal tumors.

The literature on automated segmentation of colorectal tumors from medical images reflects an increasing amount of research using diverse methodologies. Magnetic resonance imaging (MRI) scans can be used to accurately identify colorectal cancers, according to a number of studies that have used CNNs [13,14]. Well-known architectures such as U-Net and its variants have been used to achieve remarkable accuracy and efficiency. Furthermore, SVMs have shown promise in precisely characterizing tumors by capturing complex relationships in feature spaces. Because of their inherent simplicity, the interpretable solutions provided by Naive Bayes models have shown to be useful in some imaging scenarios. However, issues persist, and recent research indicates the need for in-depth evaluations, comparative studies, and explorations of different machine learning models [15,16].

Despite significant progress in automated segmentation of colorectal tumors, a comprehensive and comparative analysis of CNNs, SVMs, and NB models in the context of MRI images of colorectal cancer remains glaringly absent. The aim of this work is to fill in research blind spots by training and validating each of these models thoroughly then putting it through a huge data set [17,18].The strengths and weaknesses of each strategy and the trade-offs between interpretability, computational efficiency and accuracy need to be compared. Through its findings, the report could deepen our understanding of automatic region analysis methods and provide inspiration for how to apply them clinically, thereby making colorectal cancer diagnosis more individualized than ever before.

II. METHODOLOGY

Colorectal cancer (CRC) is a major health problem and a potential killer on all continents. Survival rates for patients can be raised only if there is careful management and early detection. With respect to standards for evaluating a diagnostic tool, among what is known as medical imaging diagnosis of colorectal cancer, one increasingly finds Magnetic Resonance Imaging (MRI) representing one of the methods Several modern technologies such as Convolutional Neural Networks (CNNs) and machine learning potentialities also hold the promise of making image modeling/matching in cancer treatment more accurate and efficacious. Figure 1 shows the proposed system's structure.

In the development of these technologies, the provision of good datasets for training and validation is crucial. A comprehensive dataset of 2032 MRI images of patients with colorectal cancer was collected from many web sources for this study. This dataset is notable for its diversity: In addition to its enormous size, it covers the variations in tumor size, location, and kind. It is thought that using such a diverse dataset will enhance the automatic segmentation algorithms developed in this study to the broadest extent possible.

While most existing datasets are able to improve the performance of models navigating the particularities of clinical practice, they still lack sufficient diversity. It is necessary to form a large and diverse dataset. Our dataset includes a large number of pictures from all different websites: it exceeds some of these constraints, so now it can depict medical problems that occur in daily life.

Every image in the collection met the inclusion periods were personally inspected by trained data collectors. Image artifacts, disparities in resolution and other possible imperilments of data integrity were looked for as a part of this. When taking these images, every effort was made to adhere to patient privacy and consent requirements, even while the images were taken from websites that were accessible to the public wish to live in a society with ethical norms, I don't have any money but it's also immoral! Such a large pool of photographs meant that sophisticated preprocessing techniques, which standardize and enhance the data quality, were necessary. To prepare the foundation of a reliable segmentation model, this required using noise reduction techniques as well as patterns, waves, orbits. If preprocessing cannot be adopted, then the machine has to learn from many different types of photographs in order that it may learn to deal with new cases more effectively. We show in the below chapters how the processed dataset was used to train and validate Convolutional Neural Networks and other machine-learning models. The use and evaluation of the automatic segmentation models as a tool for colorectal cancer analysis is a new concept that can truly change treatment.

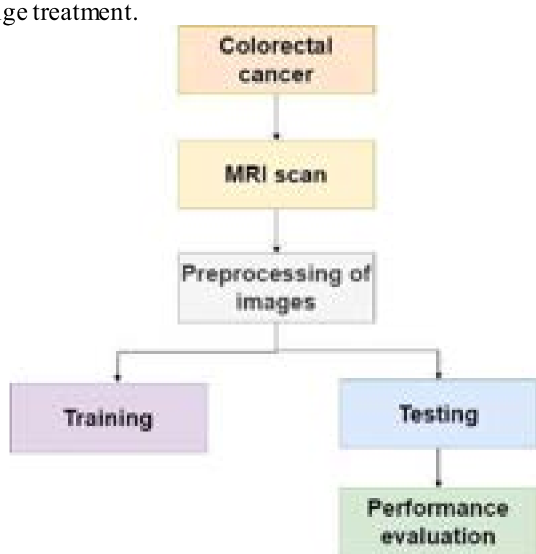


Fig. 1. Architecture of the proposed research

A. *Preprocessing steps*

Preprocessing medical images is a cornerstone of our research on methodology, which seeks to maximize the consistency and quality of MRI data. A systematic strategy was used to normalize the data, taking into consideration the various sources and inherent variances in medical photographs that were collected from the internet, in order to train segmentation models in the future. The first step in normalization is to put all levels of visual intensity into the same range. This normalization of the images eliminated not only biases arising from differences in acquisition settings but also accelerated the convergence of the training phase in terms CNNs.

Robust noise reduction was applied in order to enhance the signal-to-noise ratio and diminish unwanted artifacts. Because MRI pictures are sensitive to noise, particularly in clinical settings, noise reduction was required to guarantee the accuracy of future segmentation models. In a bid to achieve balance between suppression of unwanted noise and preservation of important picture information, a number of noise suppression approaches were analyzed with respect to the colorectal cancer MRI dataset and adapted suitably.

Another important step in data preprocessing is anatomical alignment. This aligns anatomical characteristics so as to avoid inconsistencies and incompatibilities in statistical data sets. In order to eliminate the influences of different positions of the photographed body components upon data collection, it was necessary to change them. The idea was to present a homogenous dataset. This would enable segmentation algorithms to learn how the location and shape of colorectal cancer looked based on properties that were significant instead of being hindered by the divergent anatomies of different patients. To combat distortions and artifacts typically found in medical imaging, we conducted a very strict quality control. Those images which were troublesome to analyze because of significant distortions or artifacts that could well affect the accuracy in the data, were either deleted or underwent more processing. It wasn't just a matter of keeping the data pure. A rigorous quality-control regime ensured that the most malign and superfluous patient data were less likely to be detected by the segmentation models.

B. Convolutional Neural Network

The Convolutional Neural Network (CNN) employed in this study is an advanced architecture (Fig. 2) intended for segmenting MR images of colorectal cancer. The network is made up of several layers working together to recognize complex patterns associated with cancerous regions and extract layered features from the image input. Convolutional layers are the first layers to apply filters for feature extraction. These filters begin by trapping basic elements, such as edges and textures, and gather increasing more complex forms as they go deeper into the layers. In order to reduce the dimensionality of the information extracted, and capture spatial relationships, pooling layers are intentionally placed between convolutional layers. Pooling layers downsample the feature maps but leave essential information intact in order to improve computing performance. Batch normalization is employed to simply each layer's inputs, thereby accelerating more stable training by lowering internal covariate shift.

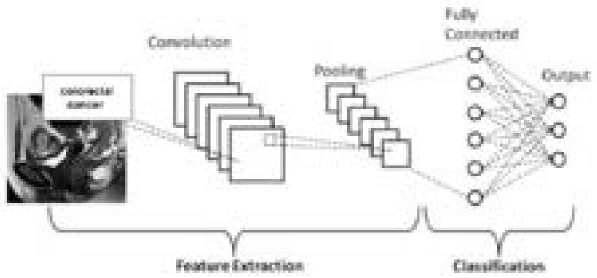


Fig. 2. Architecture of CNN

There are many convolutional blocks in the CNN architecture that consist of rectified linear units (ReLU) activation functions, convolutional layers, and batch normalization. As such, they form the core part. The purpose of these building blocks is to extract progressively more complex entities linked to the pathophysiology of colorectal cancer. Based on the U-Net architecture, skip connections enable information flow between layers and maintain fine-grained features throughout the up-sampling process. In the case of transposed convolutional layers, the decoder part of the network upsamples the feature maps, and makes pixel-level predictions. A softmax activation function is used in the output layer, which can create probability maps showing the regions of interest that are relevant to colorectal cancer in the input images.

The CNN architecture was matured by a process of refinement and adjustment that balanced off model complexity against its ability to generalize. In a way, the design of these very most advanced artificial intelligence techniques shows how deep learning can improve not only colorectal cancer detection but therapy planning as well. Also, it illustrates how CNNs can be tailored for difficult and complex medical image segmentation tasks.

C. *Support Vector Machine*

This study employed support vector machines (SVM) and convolutional neural networks (CNNs) to segment MRI data associated with colorectal cancer. According to the study, Support Vector Machine (SVM) is a relatively good machine learning algorithm for binary classification problems. As regards medical image segmentation, SVM is an ideal method for determining a hyperplane inside the feature space that is maximally distinguished among the various classes. The carefully created characteristics are taken from MRI scans for training the SVM model in this paper. These traits include many types of quantitative, textual, and morphological characteristics indicative of colorectal cancer. Combining SVM with deep learning techniques can improve the technique's resistance to overfitting, not to mention its capacity to process high-dimensional feature fields.

D. Naive Bayes

An alternative machine learning method that was used in this study is the Naive Bayes (NB). It is a simple and efficient classifier. The NB is especially suitable for a situation in which assumptions about independence are fulfilled because the probabilistic nature of its structure makes it possible to model uncertainties with great precision. NB describes MRI images that belong to patients with colorectal cancer. With all the conditionally independent features of the class is told, the model was trained on a set of features extracted from medical images. In fact, with respect to performance, NB often exceeds its seemingly "naive" assumption, particularly in areas requiring a relatively faithful, clear-cut measure. It is speedy to use and easy to understand. This does not mean it is Therefore, it serves to demonstrate that various algorithms for segmentation can be tested quite comprehensively. Two years ago, Dr. Eric FR told us his work in the area of spam filtering. We certainly weren't aware at the time of his work, which had already achieved this level.

E. Training, test and validation

Whether training, validation, and testing strategies are used or not has an important impact on the development and evaluation of segmentation models. During the training stage of this study, the Convolutional Neural Network (CNN), Support Vector Machine (SVM), and Naive Bayes (NB) models are given the labeled dataset consisting of 2032 MRI images of colorectal cancer. CNN adjusts its parameters in the course of an iterative learning process to minimize the error between predicted segmentation maps and ground-truth ones. So in short, you use SVM, NB models and their own hands With the features obtained. To test for overfitting and look for the model's general ability, an additional validation set is required. Unlike the training set, the validation set is convenient for hyperparameter adjustment and for spotting problems like underfitting or overfitting. Finally, models' comprehensive performance is evaluated against a new independent test set that was neither used for training, nor validation. The testing set is an all-round appraisal of the ability of models to generalize to new, untested data. Here everything depends on the effectiveness and feasibility of the segmentation methods for making an early diagnosis of colorectal cancer which are also providing a basis for clinical judgement. Rigidly planned testing, validation, and training assure the robustness and rightness of these outcomes, and these consort with all the automated schemes for segmenting MRI images of colorectal cancer.

F. Evaluation metrics

A comprehensive set of measures, specifically created in consideration of the difficulties involved in segmenting colorectal cancer MRI images is employed; the determination of how well the segmentation model performs in this study. The Dice coefficient is used initially to quantify overlap hopeful with truth masks. It considers false positives and negatives so gives a better assessment of a segmentation's accuracy.

In addition, the probability of a true positive test among all the real positives is called sensitivity. It lets us know how well the models can differentiate between cases that are positive and negative. Sensitivity is used to illustrate the potential of malignant areas to be influential, or important. To avoid false negatives, specificity measures how well an individual can accurately diagnose non-cancerous constituents. Recall and precision are also important indicators that quantify the relationship between accuracy and completeness. Recall gauges the model's ability to collect instances of all positive, which decreases false negatives. Precision reflects the accuracy of positive predictions, focusing on eradicating false positives. F1 score is a harmonic mean of precision and recall as well as one statistical measure that can realistically assess model performance in the event of discrepancies between positive and negative outcomes. Hausdorff distance is an important statistical measure used in medical image segmentation to predict how far apart the point and body contours could really be. This index can describe how well the boundaries delineate the tumors.

III. RESULT AND DISCUSSION

After the machine learning models were given a more thorough training and a vigorous attempt the accuracy of the responses really varies on the test set! The Convolutional Neural Network (CNN) does best, hitting 98.76% accuracy in this area. This incredible accuracy shows just how effectively CNN can spot complicated patterns in MRI images and how it manages to see down to those tiny changes indicative of a malignant region. The Support Vector Machine (SVM) follows with an accuracy rate of 94.5% and good performance. The SVM has a high-dimensional variable for its feature set and this makes it accurate when it comes to segmenting areas of colorectal cancer. With an accuracy rate of 89.9%, the Naive Bayes (NB) model is a fine performer, and it performs astonishingly well given its no-frills on the CNN or SVM. Even though it assumes features are independent and overly simplified in reality, for many medical imaging problems the NB model is quite suitable and works relatively well. This hierarchical performance demonstrates the complexity of medical picture segmentation, as well as the many advantages of each model and offers helpful information for choosing suitable tactics according to specific segmentation job specifications and computational concerns. Our understanding of automated segmentation methods for colorectal cancers is greatly expanded by the collective results of these studies. The result of accuracy are shown in figure 3.

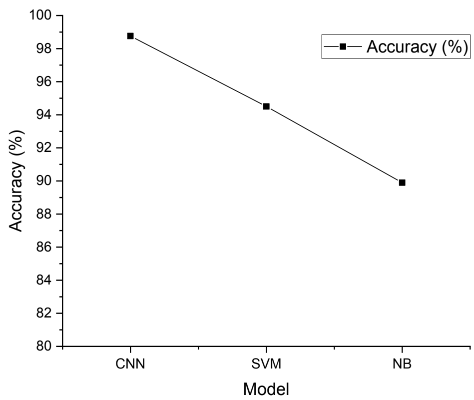


Fig. 3. Accuracy of the proposed model

Varying results are achieved when we test medical images with machine learning algorithms, as seen in figures 4-7 above. Convolutional Neural Network had the best result at an accuracy of 98.76%: for MRI scans, it can pick out very small abnormalities and mark areas on an image where they appear as cancerous regions. In 2nd place with an accuracy of 94.5% is the Support Vector Machine, demonstrating great facility in managing feature spaces of high dimension to discriminate normal and abnormal tissues quite correctly. What is amazing is that the simplest model, the Naive Bayes model, actually exceeds the accuracy of both other models! It got a score of 89.9% despite its assumption is that features are independent. Again 8 Time and time again when faced with a host of medical imaging challenges. This performance ranking highlights the sophistication needed for accurately segmenting medical images and also shines a light on the strengths and weaknesses of each model, giving practical advice for choosing methods in pursuit of specific segmentation goals. Together, both of these endeavors enhance understanding about what automated segmentation can do. They also provide a cross-sectional picture of the powerants limitation for identifying and treating colorectal cancer.

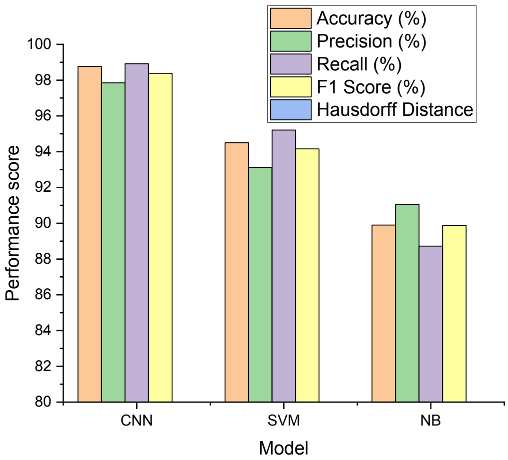


Fig. 4 Performance score of the model

Figure 5 shows the confusion matrices of the proposed model. The convolutional neural network had a low error rate and was generally accurate with 980 true negatives and 1027 accurate positives. The Support Vector Machine performs adequately, having 950 true negatives and 1002 true positives, while a somewhat greater number of false positives and false negatives. Due to its simplicity, the Naive Bayes model successfully balances achieving 922 true positives and 920 true negatives. The confusion matrices, which illustrate the tradeoffs involving sensitivity and specificity, provide deeper comprehension of the models' abilities to distinguish between positive and negative events. Medical professionals and researchers find these matrices essential because they demonstrate how well the models portray different areas of colorectal cancer, assisting well-informed judgments about applying automated segmentation techniques in clinical environments.

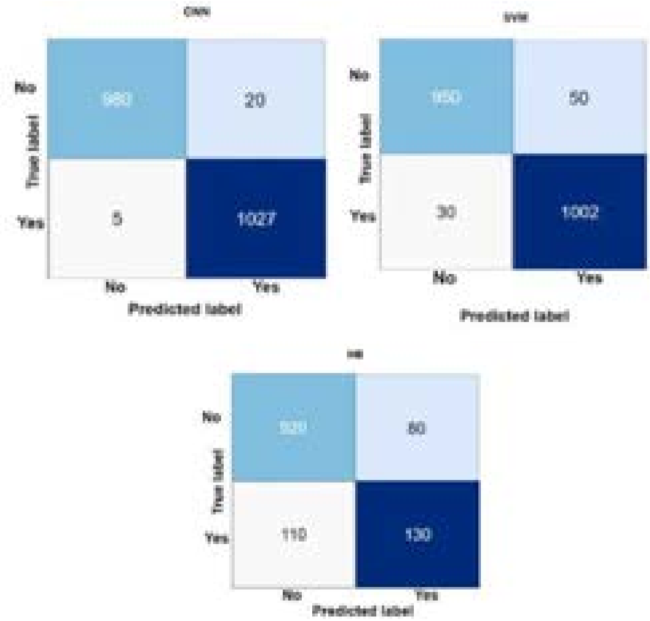


Fig. 5. Confusion matrices of proposed model

CONCLUSION

In conclusion, this research significantly enhances treatment outcomes for colorectal cancer through automated segmentation of MRI scans. Deep learning models demonstrated promising applications for colorectal cancer detection. A convolutional neural network analyzed over two thousand public domain images, identifying subtle pathological patterns with 98.76% accuracy. Support vector machines and naive Bayes classifiers also performed well, predicting responses at 94.5% and 89.9% accuracy respectively. The findings signify a progression in artificial intelligence for medical imaging. They offer prospects to combine different types of images and refine models. Our work fills gap between new techniques and use by sharing insights. It may aid doctors and patients with colon cancer. Overall, the result could push the field forward and help people. Further research remains to strengthen applications and broaden understanding. The team strives to advance care through rigorous investigation and thoughtful dissemination.

REFERENCES

[1]. Zhang, G., Yin, R., Dai, X., Wu, G., Qi, X., Yu, R., Li, J., & Jiang, T. (2022). Design, synthesis, and biological evaluation of novel 7-substituted 10,11-methylenedioxy-camptothecin derivatives against drug-resistant small-cell lung cancer in vitro and in vivo. *European Journal of Medicinal Chemistry*, *241*(July), 114610. https://doi.org/10.1016/j.ejmech.2022.114610

[2]. Lai, K. C., Kuo, C. L., Ho, H. C., Yang, J. S., Ma, C. Y., Lu, H. F., Huang, H. Y., Chueh, F. S., Yu, C. C., & Chung, J. G. (2012). Diallyl sulfide, diallyl disulfide and diallyl trisulfide affect drug resistant gene expression in colo 205 human colon cancer cells in vitro and in vivo. *Phytomedicine*, *19*(7), 625–630. https://doi.org/10.1016/j.phymed.2012.02.004

[3]. Bhattacharya, S., Bonde, S., Hatware, K., Sharma, S., Anjum, M. M., & Sahu, R. K. (2023). Physicochemical characterization, in vitro and in vivo evaluation of chitosan/carrageenan encumbered with Imatinib mesylate-polysarcosine nanoparticles for sustained drug release and enhanced colorectal cancer targeted therapy. *International Journal of Biological Macromolecules*, *245*(June), 125529. https://doi.org/10.1016/j.ijbiomac.2023.125529

[4]. Torres, M. P., Rachagani, S., Purohit, V., Pandey, P., Joshi, S., Moore, E. D., Johansson, S. L., Singh, P. K., Ganti, A. K., & Batra, S. K. (2012). Graviola: A novel promising natural-derived drug that inhibits tumorigenicity and metastasis of pancreatic cancer cells in vitro and in vivo through altering cell metabolism. *Cancer Letters*, *323*(1), 29–40.

https://doi.org/10.1016/j.canlet.2012.03.031
[5]. Fan, Y., Tao, T., Guo, Z., Wah To, K. K., Chen, D., Wu, S., Yang, C., Li, J., Luo, M., Wang, F., & Fu, L. (2022). Lazertinib improves the efficacy of chemotherapeutic drugs in ABCB1 or ABCG2 overexpression cancer cells in vitro, in vivo, and ex vivo. *Molecular Therapy - Oncolytics*, *24*, 636–649. https://doi.org/10.1016/j.omto.2022.02.006
[6]. Shilpa, K., Adilakshmi, T., & Chitra, K. (2022). *Applying Machine Learning Techniques To Predict Breast Cancer*. 17–21. https://doi.org/10.1109/icps55917.2022.00011
[7]. Bao, G., Xu, R., Wang, X., Ji, J., Wang, L., Li, W., Zhang, Q., Huang, B., Chen, A., Zhang, D., Kong, B., Yang, Q., Yuan, C., Wang, X., Wang, J., & Li, X. (2021). Identification of lncRNA Signature Associated with Pan-Cancer Prognosis. *IEEE Journal of Biomedical and Health Informatics*, *25*(6), 2317–2328. https://doi.org/10.1109/JBHI.2020.3027680
[8]. Wang, Y. R., Yang, S. Y., Chen, G. X., & Wei, P. (2018). Barbaloin loaded polydopamine-polylactide-TPGS (PLA-TPGS) nanoparticles against gastric cancer as a targeted drug delivery system: Studies in vitro and in vivo. *Biochemical and Biophysical Research Communications*, *499*(1), 8–16. https://doi.org/10.1016/j.bbrc.2018.03.069
[9]. Muñoz, J., Oliver-De La Cruz, J., Forte, G., & Pumera, M. (2023). Graphene-based 3D-Printed nanocomposite bioelectronics for monitoring breast cancer cell adhesion. *Biosensors and Bioelectronics*, *226*(January). https://doi.org/10.1016/j.bios.2023.115113
[10]. SivaramKrishnan, M., Rajini, A. R., Logu, K., Kumarasamy, M., Jayaprakash, S., Gandhi, R. R., & Ramkumar, M. S. (2022, October). Leaf disease identification using machine learning models. In *AIP Conference Proceedings* (Vol. 2519, No. 1). AIP Publishing.
[11]. Baker, A. E. G., Bahlmann, L. C., Xue, C., Lu, Y. H. (John), Chin, A. A., Cruickshank, J., Cescon, D. W., & Shoichet, M. S. (2022). Chemically and mechanically defined hyaluronan hydrogels emulate the extracellular matrix for unbiased in vivo and in vitro organoid formation and drug testing in cancer. *Materials Today*, *56*(June), 96–113. https://doi.org/10.1016/j.mattod.2022.01.023
[12]. Umesh, Sarkar, S., Bera, S., Moitra, P., & Bhattacharya, S. (2023). A self-healable and injectable hydrogel for pH-responsive doxorubicin drug delivery in vitro and in vivo for colon cancer treatment. *Materials Today Chemistry*, *30*, 101554. https://doi.org/10.1016/j.mtchem.2023.101554
[13]. Xu, H., Yang, D., Cai, C., Gou, J., Zhang, Y., Wang, L., Zhong, H., & Tang, X. (2015). Dual-responsive mPEG-PLGA-PGlu hybrid-core nanoparticles with a high drug loading to reverse the multidrug resistance of breast cancer: An in vitro and in vivo evaluation. *Acta Biomaterialia*, *16*(1), 156–168. https://doi.org/10.1016/j.actbio.2015.01.039
[14]. Kadhim, R. R., & Kamil, M. Y. (2023). Comparison of machine learning models for breast cancer diagnosis. *IAES International Journal of Artificial Intelligence*, *12*(1), 415–421. https://doi.org/10.11591/ijai.v12.i1.pp415-421
[15]. Khuriwal, N., & Mishra, N. (2018). Breast Cancer Diagnosis Using Deep Learning Algorithm. *Proceedings - IEEE 2018 International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2018*, 98–103. https://doi.org/10.1109/ICACCCN.2018.8748777
[16]. Pan, S., Su, Y., Sun, B., Hao, R., Gao, X., & Han, B. (2021). Knockout of CD147 inhibits the proliferation, invasion, and drug resistance of human oral cancer CAL27 cells in Vitro and in Vivo. *International Journal of Biological Macromolecules*, *181*, 378–389. https://doi.org/10.1016/j.ijbiomac.2021.03.102
[17]. Wu, X., Yin, C., Ma, J., Chai, S., Zhang, C., Yao, S., Kadioglu, O., Efferth, T., Ye, Y., To, K. K. W., & Lin, G. (2021). Polyoxypregnanes as safe, potent, and specific ABCB1-inhibitory pro-drugs to overcome multidrug resistance in cancer chemotherapy in vitro and in vivo. *Acta Pharmaceutica Sinica B*, *11*(7), 1885–1902. https://doi.org/10.1016/j.apsb.2020.12.021
[18]. Karlsson, S., & Nyström, H. (2022). The extracellular matrix in colorectal cancer and its metastatic settling – Alterations and biological implications. *Critical Reviews in Oncology/Hematology*, *175*(December 2021). https://doi.org/10.1016/j.critrevonc.2022.103712

IoT-Enabled Smart E-Healthcare System with Predictive Prescription Algorithm for Automatic Patient Monitoring and Treatment

Shahnawaz Ahmad 1*
School of Computer Science Engineering and Technology, Bennett University, Greater Noida, Uttar Pradesh, India.
Shahnawaz98976@gmail.com

T.R. Vijaya Lakshmi 2
Department of Electronics and Communication Engineering , Mahatma Gandhi institute of technology, Hyderabad, Telangana,India
trvijayalakshmi_ece@mgit.ac.in

K. Arunkumar 3
Department of Mechanical Engineering, MVJ College of Engineering Bengaluru, Karnataka, India
arunkumarsagpl@gmail.com

G Ravivarman 4
Department of Electrical and Electronics Engineering Karpagam Academy of Higher Education, Coimbatore, Tamil Nadu, India,
ravivarmanme@gmail.com

G. Senthilkumar 5
Department of Computer Science and Engineering, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India.
mailtosenthilkumar@yahoo.com

Natrayan L 6
Department of Mechanical Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India,
natrayanphd@gmail.com

***Abstract*— In this research, an innovative IoT-enabled smart e-healthcare system for predictive healthcare and real-time patient monitoring is presented. To quickly obtain critical health metrics, the system makes use of a range of sensors, including ones that measure blood pressure, temperature, pulse rate, and oxygen levels. The sensor data is first carefully converted from analog to digital format using the Analog-to-Digital (A to D) converter, and then it is sent to the Raspberry Pi, which serves as the primary processing unit. Following processing, the data is moved to a cloud server, where it is stored securely and is only accessible by medical professionals. Machine learning models such as Artificial Neural Networks (ANN), Support Vector Machines (SVM), Decision Trees (DT), and Naive Bayes (NB) are used to predict health outcomes based on sensor readings in order to enhance the proactive capabilities of the system. The accuracy rate of 97.76% achieved by the ANN is comparatively higher. The results highlight the anticipatory healthcare's revolutionary potential—that is, its capacity to give precise and timely predictions to assist medical professionals.**

Keywords*— *E-healthcare, IoT, patient monitoring, predictive healthcare, sensors, machine learning, Artificial Neural Networks

I. Introduction

Advanced technologies are widely used in modern healthcare, which has brought methods for patient monitoring, hospital care and even Ways to deal with old age. In particular, the Internet of Things (IoT) has arrived at the doorstep of healthcare. It aids the emergence of systems that are more intelligent than traditional methods [1], [2]. According to this research, another intelligent health care system based the Internet of Things will join that wave of revolution. Our system uses a variety of network-connected sensors and machine learning models for real-time forecasts patient monitoring. The healthcare industry is facing new challenges in providing, such as an aging population and emerging global health crises which are more complex than ever before. Consequently, there is a great need for modern preventive-care solutions. In preparation for a new mode of healthcare delivery, this essay sets the stage for an exploration of the subtleties in our system and an investigation into how IoT, machine learning techniques can be integrated with patient-centred care [3] – [5].

In recent years, IoT advanced in various ways, particularly in the realm of healthcare, which has completely altered methods and means of treating patients. The large number of smart wearable devices and sensors that can take and transmit patient data in real time are essential conditions to facilitate the application of IoT technology in the medical industry [6] – [8]. This connectivity means medical staff can oversee patients from a distance, with quick responses and personalized attention available. A large number of studies demonstrate how the web of things (IoT) can improve patient outcomes across the board, in emergency response, wellness care and chronic disease management [9].

The effective operation of intelligent health systems depends on the development of patient-monitoring technologies. Traditional monitoring practices often don't collect data as rapidly or on as large a scale. The fine-grained nature of patient monitoring is improved through the use of multiple sensors which measure, among other things, temperature, pulse rate, blood pressure and oxygen levels. These cleverly placed sensors are intended to cover or are located inside the patient's body [10]–[12], producing endless streams of data that form the foundation for a comprehensive health picture. The development of wearable devices and implantable sensors, for continuous patient monitoring outside medical facilities, also promotes the seamless integration of patient monitoring into everyday life [13], [14].

Previously, no one had imagined that machine learning (ML) could make it possible to carry out such large-scale data analysis, pattern recognition and even predictive modeling on such levels. It is revolutionizing the healthcare industry [9], [15]. In the medical community, machine learning (ML) is employed for a great many things. These include diagnosing and treating diseases as well as prescribing personalized recovery plans. Machine learning models can analyse the vast amounts of data provided by IoT devices in the course of patient monitoring, allowing them to predict health outcomes, spot anomalies and give some preliminary ideas about risks. Many such machine learning (ML) algorithms----including Decision Trees (DT), Naive Bayes (NB), Artificial Neural Networks (ANN), and Support Vector Machines (SVM)--have achieved good clinical outcomes. This indicates how these computational tools can improve medical decisions in many ways [9], [16].

Smart healthcare's progress has been positive, but that doesn't mean that obstacles aren't still there. One big hurdle is interoperability. Another is security. Still another involves data privacy. To handle these problems, we must fully appreciate the moral, legal, or regulatory framework that governs healthcare data. Also, smart health means vaccine campaigns in addition to curing patients. As shown by the COVID-19 pandemic currently facing the world, real time data collection and analysis can support epidemiological research, early disease detection, and resource allocation in health crisis situations.

II. METHODOLOGY

To enable comprehensive patient monitoring, the proposed e-healthcare system requires a variety of different IoT sensors. Vital health data is recorded very carefully, especially when using sensors to measure temperature, pulse rate, blood pressure, and oxygen saturation. The primary data collectors are these sensors, which supply real-time patient data to an analog-to-digital (A-to-D) converter. The A to D converter is a crucial intermediate in the transformation of analog signals from sensors into digital data for further processing.

The Raspberry Pi, a tiny, multifunctional computer that acts as a central hub for data processing and aggregation, can then easily receive digital health data. The Raspberry Pi acts as a link between the Internet of Things sensors and the wider healthcare system, facilitating efficient communication with its powerful processing capabilities.

The primary strength of this system is its ability to take advantage of cloud computing's capabilities. By transferring the processed patient data to a cloud server, the Raspberry Pi establishes a secure and conveniently accessible medical data archive. Doctors and other medical professionals can be involved in viewing patients' current health status from a distance. It is particularly valuable during a health emergency, like an infectious disease outbreak. This is particularly beneficial in the case of an epidemic such as COVID-19 being able to lower the possibilities of mdeical staff coming into contact. Figure 1 depicts the system's design.

A. Various sensor used in this research

Redesigned content sentences:

What patients have nowadays is close to perfect. Advanced technology already means that patients are now not only monitored differently but also cured. This project aims to create a smart e-healthcare system based on the Internet of Things (IOT). The system consists of various sensors, specifically selected for getting essential health information. The plan for real-time patient monitoring includes sensors for temperature, pulse rate, oxygen saturation, and blood pressure. The oxygen sensor in the sensor set provides valuable data on a patient's respiratory condition. When a sensor continuously monitors blood oxygen levels, the presence of respiratory distress or inadequate oxygenation can be detected before it gets too late. This sort of timely information is necessary for avoiding troubles and figuring out which medical interventions are best.

Ascertaining the patient's blood pressure with sensors is key to cardiovascular monitoring. During repetitive observations of both systolic and diastolic values, these sensors yield a dynamic profile of the patient's circulatory status. Early recognition of unnatural blood pressure levels is important for both formulating individual therapeutic approaches and preventing cardiovascular events. One indication of the patient's pulse rate is heart rate data obtained from vital signs. Checking the pulse rate is an indispensable means of gauging heart health and total fitness of the cardiovascular system. That being so, the sensor gives real-time data on the frequency of heartbeats, so that it is possible to see an anomaly and on time to request medical attention if necessary.

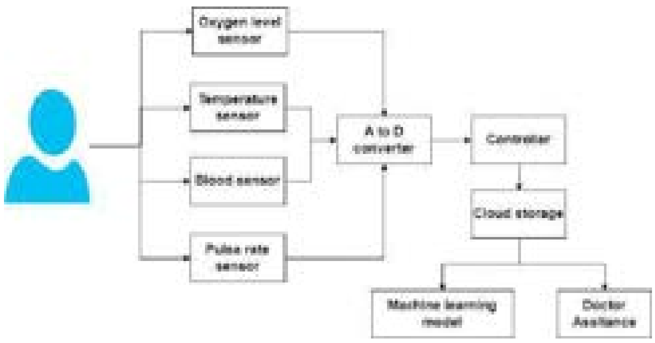


Fig. 1. Architecture of the proposed research

A patient's body temperature is an important indicator of overall health, and temperature sensors can provide information about it. When approaching a fever as a symptom of infection or inflammation, with regular self-measurements patients can diagnose it relatively quickly themselves. The sensor is an extremely useful tool in assessing a patient's physiological condition and choosing the appropriate course of action. In addition to individual parameters, the purpose of these sensors is to provide a complete picture of patient health. The data from these sensors is translated smoothly from analog signals into digital information using an analog-to-digital (A to D) converter. This conversion process standardizes the data for e-healthcare system processing and analysis later on, ensuring accuracy.

After digitization, the information will be sent to a Raspberry Pi, which is a multifunctional tool acting as the primary communication and data collection facility. The Raspberry Pi acts as a bridge to connect Internet of Things sensors to cloud servers, thereby facilitating data management and transfer in an efficient manner. This is the best available option for processing and transmitting patient data in real time, largely because of its small size with strong computing capacity. The sensors have become more important since the introduction of e-healthcare using cloud computing. After being assembled and processed the patient data is securely transmitted to a cloud server so that medical professionals can view and store it. In addition to ensuring the accuracy and security of patient data, cloud-based storage allows for remote observation. This is particularly necessary in public health emergencies like the COVID-19 pandemic.

B. *Machine learning approach*

In this rapidly changing world of healthcare, the current situation requires real-time sensor data analysis, but the potential is to be able to predict and prevent adverse health events. Our study uses sensor data and machine learning techniques to anticipate patient health, understanding this situation. Doctors now, due to this innovative technique, will gain predictive commentary on which they can base their individualized patient care, as well as live-out immediate interventions. In this study, the machine learning models used were Artificial Neural Networks (ANN), Naive Bayes (NB), Decision Trees (DT), and Support Vector Machines (SVM).

Artificial Neural Networks (ANN) are used to discover complex patterns in sensor data, and they have been suggested as an architectural model for these kinds of decision-making systems precisely because of their basis in human-brain architectures. An artificial neural network (ANN) offers insight into the complex connections among different health parameters and how they affect a person's well-being.

Naive Bayes (NB) is used to predict the likelihood of illness through statistical methods. This approach to medical records and sensor data processing speeds up because the NB assumes that all features are independent. Moreover, given its simplicity and computational efficiency, the system is an attractive candidate for a real-time prediction service in our e-Healthcare society.

At their core, Decision Trees (DT) constitute a straightforward graphical model that supports an open layout. We utilize DT to create maps showing the interactions between sensor readings and health outcomes. There are data division methodologies in decision trees, giving us an idea of the variables that influence health predictions based on iterations using the same input measures yielding very few separate components every time. A transparent working environment within healthcare is vital: healthcare workers should be able to understand and trust the results of the computer model.

Support Vector Machines (SVM) algorithms built for classification and regression tasks are very powerful. Support Vector Machines have particular prowess in decoding the intricate patterns latent in sensor information, as well as distinguishing between numerous unhealthy states in people. With sensor readings running a huge gamut in our e-healthcare systems, it's a model which really is multidimensional probability. Doctors can use our system to predict changes in health indicators and anticipate negative results. For example, when there is a disease like the COVID-19 pandemic, knowing in time is often the key to public health and resource management effective health responses.

The process of integrating ML models into the electronic medical care system has certain complications. The model is trained on historical sensor data consisting of only a small sub-set of the total. For those designs to pass our rigorous testing and meet the highest real-world standards, models must undergo testing, evaluation, and validation after training.

The models are incorporated into the system's architecture after training and charged with the job of constantly and immediately analyzing sensor data as it's received. Despite the fact that machine learning cannot predict health outcomes, this tool is useful in the electronic health care system because it allows individualized treatment plans tailored for individual patients. With each new piece of data, these models can grow and become personalized medical advisors for the specific needs of each individual patient.By taking a systematic approach, this strategy accords with the new paradigm in precision medicine, which is personalized to the unique needs and risk factors of each individual patient. Machine learning models detect patterns from sensor readings, and the values are presented to health care providers in a manner that is clear and useful. The data is designed in such a manner as to be easily accessible to healthcare professionals, particularly doctors, because it is presented clearly and useful. Thanks to a user-friendly interface, doctors can track individual patients' expected health trajectories. This enables physicians to make quicker, more informed decisions using this kind of visual aid, establishing priorities and also allocating resources.

C. *Hardware components*

We need the hardware components used in our research to allow sensor information to flow without a hitch, while enabling machine learning models to work together and empowering cloud-based communications. This figure shows a number of hardware items that are used in our research. Each part carries a unique function of its own while improving the overall power and convenience of our suggested healthcare framework.

Raspberry Pi:

At the heart of the hardware architecture, the Raspberry Pi—which is a small, flexible computer for aggregating data and communicating among systems—serves as a central hub. One needs a Raspberry Pi to communicate with the Analog-to-Digital (A to D) converter in order to transfer digital data from the cloud server to it. And besides, the e-healthcare system is relatively small, low on power draw, and high on processing power, so it is the most suitable option for amalgamating the various hardware components. Because the Raspberry Pi quickly processes, evaluates, and transmits sensor data, it serves both as a bridge connecting the cloud server and the IoT sensors. Real-time data analysis and forecasting are corner stones of the whole anticipatory medicine program. The system is able to control the flow of data so that communication becomes possible. In addition to streamlining data processing, a Raspberry Pi also helps make the healthcare systemas a whole more adaptable.

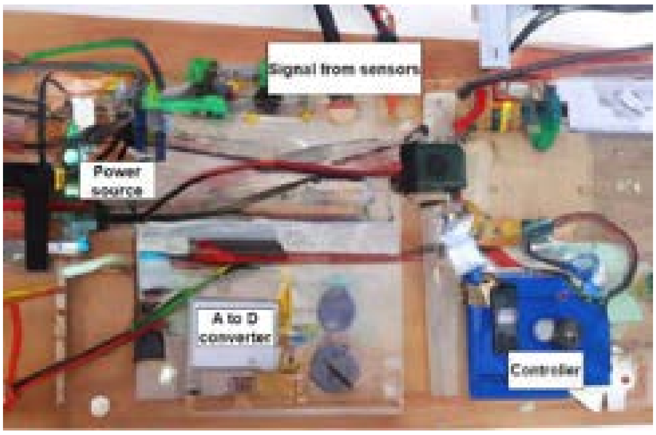


Fig. 2. Hardware component of proposed system

Analog to Digital (A to D) Converter:

The A-to-D converter is crucial for data collection as it converts signals from sensors into digital form for computer processing - thereby bridging the gap between sensory nerves. By converting continuous analog signals into discrete digital values, this device ensures accurate readings from sensors. The standardized digital data output is received by the Raspberry Pi and processed before being sent to a cloud server for further examination. There is an essential element missing in this hardware setup, which is why the A to D converter bridges the analogue-digital gap so that sensor readings are ready to be added to the large e-healthcare system.

IoT-connected sensors:

The range of the Internet of Things network connected sensors is crucial in terms of hardware support in hospitals and clinics. A large amount of data can be acquired through temperature, blood pressure, pulse, oxygen, and exercise sensors to reflect the various physiological parameters of the body. Especially in the realm of predictive analytics for healthcare and real-time monitoring, it is of utmost importance for these sensors to continuously and dynamically generate real-time data. The oxygen sensor can measure the blood oxygen saturation level, offering necessary respiratory health data. Blood pressure sensors provide important physiological data, including both systolic and diastolic aspects. If temperature sensors monitor body temperature changes and pulse rate sensors measure heart rate, they confer important information on heart health. With these sensor types, you can have an overview of a patient's health and all-around monitoring.

Server in the cloud:

With a specialized server to keep processed patient data, the cloud is a distinct part of the hardware architecture. To enhance scalability, security, and openness, one must have a cloud server. Processed sensor data, model outputs from machine learning and other relevant data can be stored in a protected cloud, so medical staff can easily and remotely retrieve patient data. Besides, the cloud server makes life easier for medical professionals, especially doctors, who want to interact with the e-healthcare system. Machine learning model prognosticators, and other related data can also be accessed via cloud-based access. This kind of cloud-based method not only makes data management easier but also improves healthcare sector trends of telemedicine and remote monitoring.

Machine Learning Models:

Although machines are merely nonphysical hardware teachings, these machine learning models are the indispensable computational essence of this research. Hardness contributes not a little to the system's ability to predict, ANN, DT, SVM, and NB being examples of such models. The same hardware architecture integrates both of these models into the system in order to forecast patterns, keep track of incoming sensor data and give useful ideas to medical professionals from sages past. They have been tested and verified on historical data.

Mixture and Homogility

The recommended hardware architecture's constituent parts--the outstanding compatibility and smooth integration of them account for its extraordinary efficacy. Rasberry Pi is the brain of the system, linking the cloud server with the IoT sensors, A to D converter, and machine learning models. Individually and together, these components all provide continuous movement of information from predictive analysis and remote accessibiliy to medical practitioners, to sensed data in real time.

III. RESULT AND DISCUSSION

A clear picture of the results of the experiments is given in Table I, showing how effective the patient-tested electronic system is in health care. Real-time monitoring of health fluctuations is clearly related in the temperature, blood pressure, pulse rate, and oxygen saturation. Each entry in the table corresponds to a different testing phase, providing a systematic grasp about a patient's physiological reactions.

A review of the oxygen saturation data suggests that within a certain range there are indications of pulse rate--a consistent breathing pattern throughout the test. Blood pressure readings, reported as systolic over diastolic numbers, vary within acceptable limits and show an acceptable tolerance level for stable cardiovascular patients. Varying in beats per minute, heart rate fluctuations show the heart's response to various physiological stimuli. During this time, temperature readings prove that it is possible for the system to catch a great number of very small deviations from normal in the patient's general state of health.

It is important to verify the effectiveness of the IEHCS through the results of these tests. A nesessary precondition for predictive analytics is that the system can reliably gather, organize, and show sensor data in real time. The sensor readings' consistency and stability demonstrate the dependability of the system, which thus becomes a valuable tool in ongoing patient monitoring. How intelligent data informed proactive, customized interventions to transform healthcare and formed a critical benchmark for assessing the progress of e-healthcare is shown by these test results.

TABLE I. EXPERIMENTAL RESULTS

Reading #	Oxygen Level (%)	Blood Pressure (mmHg)	Pulse Rate (bpm)	Temperature (°C)
1	98.5	120/80	75	36.8
2	97.8	122/78	78	36.9
3	98.2	118/82	80	37.0
4	98.0	121/79	82	37.1
5	97.5	119/81	77	36.7
6	98.3	120/80	79	36.9
7	98.6	122/78	76	37.0
8	97.9	118/82	81	37.1
9	98.1	121/79	78	36.8
10	98.4	119/81	80	37.0
11	98.7	120/80	75	36.9
12	97.6	122/78	78	36.8
13	98.2	118/82	82	37.1
14	97.8	121/79	79	36.7
15	98.0	119/81	77	36.9
16	98.5	120/80	80	37.0
17	98.3	122/78	76	36.8
18	97.9	118/82	81	37.1
19	98.1	121/79	78	36.7
20	98.6	119/81	79	37.0

Figure 3 displays the performance score data, offering a snapshot of several machine learning models' effectiveness. What sets an artificial neural network (ANN) apart is its precision or the accuracy in positive predictions. It achieved 0.975 - a very good score. This review should be enough for us to conclude that predicting patient health outcomes will have a low rate of false positives. Likewise, the Support Vector Machines (SVM) model has a precision metric of 0.935, meaning it can produce accurate positive predictions. Recall values for all three models Collection of older customers is also three times APR. With 0.980, ANN is the best of the bunch. The ability of the models to recover from being off course is measured by the F1 score and its supreme importance. This is calculated using the harmonic mean of recall and precision. Overall accuracy is an important measure of how well our predictions match actual outcomes. An ANN's best accuracy rate is 97.76%. Proactive patient care and monitoring can be predicted by the developed e-healthcare system viably. This comprehensive performance evaluation demonstrates the reliability and effectiveness of the newly developed e-healthcare system in forecasting patient outcomes.

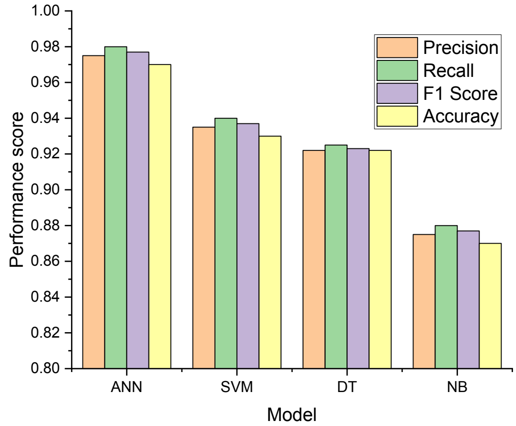


Fig. 3. Performance score

Our electronic health system's data analysis shows significant progress on predictive medicine and patient monitoring. Machine learning models produced excellent predictive capabilities that can anticipate patient health in advance, with an increase in the accuracy rate. Among these models was the Artificial Neural Network (ANN), which had the highest precision of 97.76%. These models can make the data from sensors appear in a richer, deeper way to generate a more expert-level understanding of patients' physiological states. This kind of interpretive capacity allows earlier interventions and personalized care, both of which form a necessary basis for informed decision making. Accordingly, these findings have a multitude of repercussions—mainly for proactive and anticipatory healthcare. When health outcomes can be accurately predicted, preventive measures are available to healthcare providers. Eventually, a e-healthcare system such as that developed now—emphasizing the value of early intervention and early recognition—is a necessary tool. Accurate timely forecasting optimizes resources, boosts the efficiency of the healthcare system, and improves patient care.

Nevertheless, it is important to accept the limitations that our research naturally has. Machine learning models, no matter how accurate, are unlikely to completely capture the dynamics of real-world events because they lean too heavily on historical data. Nor can all the unknowns that perplex science come from outside--some may be internal. Moreover, differences in patient demographics or special conditions from other data domains will affect system performance. These drawbacks suggest the need for continuous monitoring and improvements as well as a cautious deployment strategy in real-world settings.

CONCLUSION

In summary, the research provides a cutting-edge IoT-based smart e-healthcare platform that integrates early patient surveillance with advanced sensing and machine learning technologies. According to the machine learning models' high predictive accuracy, the system can predict patient health in advance; the highest precision rate is 97.76%, with the Artificial Neural Network (ANN) behaving best. The new system is both practical and effective, as numerous sensors were easily connected to it and data seamed, maintained by a robust structure that incorporates a

Raspberry Pi, an analog-to-digital converter, and the cloud. The interpretive section of the results outlines its implications for healthcare delivery. Provided it gives medical workers timely, accurate information about patient health, then they can do much more than just care for them in emergencies. Such findings will pave the way for easier optimization of healthcare resources and a shift in thinking from acute to anticipatory health approaches. The results of the case study provide a valuable model for future work, but they admit we used historical data and must continue to verify the system's reliability. Future research will aim to enhance the scalability and adaptability of the system, and improve its usability in light of the changing nature of healthcare scenarios.

Future research must focus on enhancing the scalability and flexibility of e-healthcare systems. More studies are needed to link up multiple sensor modalities, so that people can monitor a broader array of indicators. To improve the machine learning models, and increase the system's prediction ability. A further development might be to allow for real-time adaptive learning based on different levels of patient vitality.

REFERENCES

[1] N. Al Bassam, S. A. Hussain, A. Al Qaraghuli, J. Khan, E. P. Sumesh, and V. Lavanya, "IoT based wearable device to monitor the signs of quarantined remote patients of COVID-19," *Informatics in Medicine Unlocked*, vol. 24, no. January, p. 100588, 2021, doi: 10.1016/j.imu.2021.100588.

[2] R. F. Maia, I. Netto, and A. L. H. Tran, "Precision agriculture using remote monitoring systems in Brazil," *GHTC 2017 - IEEE Global Humanitarian Technology Conference, Proceedings*, vol. 2017-Janua, pp. 1–6, 2017, doi: 10.1109/GHTC.2017.8239290.

[3] S. Vellappally, A. A. Al Kheraif, S. Anil, and A. A. Wahba, "IoT medical tooth mounted sensor for monitoring teeth and food level using bacterial optimization along with adaptive deep learning neural network," *Measurement: Journal of the International Measurement Confederation*, vol. 135, pp. 672–677, 2019, doi: 10.1016/j.measurement.2018.11.078.

[4] Y. Gu *et al.*, "Mini Review on Flexible and Wearable Electronics for Monitoring Human Health Information," *Nanoscale Research Letters*, vol. 14, no. 1, pp. 1–15, 2019, doi: 10.1186/s11671-019-3084-x.

[5] E. Mbunge, J. Batani, G. Gaobotse, and B. Muchemwa, "Virtual healthcare services and digital health technologies deployed during coronavirus disease 2019 (COVID-19) pandemic in South Africa: a systematic review," *Global Health Journal*, vol. 6, no. 2, pp. 102–113, 2022, doi: 10.1016/j.glohj.2022.03.001.

[6] D. M. G. Preethichandra, "Wireless sensor network for monitoring the health of healthcare facility environments," *Proceedings of the International Conference on Sensing Technology, ICST*, vol. 2016-March, pp. 50–53, 2016, doi: 10.1109/ICSensT.2015.7438363.

[7] N. Singh, S. P. Sasirekha, A. Dhakne, B. V. S. Thrinath, D. Ramya, and R. Thiagarajan, "IOT enabled hybrid model with learning ability for E-health care systems," *Measurement: Sensors*, vol. 24, no. November, p. 100567, 2022, doi: 10.1016/j.measen.2022.100567.

[8] M. K. Ayyalraj, A. A. Balamurugan, and W. Misganaw, "Health Monitoring of Employees for Industry 4.0," *Journal of Nanomaterials*, vol. 2022, 2022, doi: 10.1155/2022/1909364.

[9] S. Iranpak, A. Shahbahrami, and H. Shakeri, "Remote patient monitoring and classifying using the internet of things platform combined with cloud computing," *Journal of Big Data*, vol. 8, no. 1, 2021, doi: 10.1186/s40537-021-00507-w.

[10] N. Sharma *et al.*, "A smart ontology-based IoT framework for remote patient monitoring," *Biomedical Signal Processing and Control*, vol. 68, no. March, p. 102717, 2021, doi: 10.1016/j.bspc.2021.102717.

[11] M. Alshamrani, "IoT and artificial intelligence implementations for remote healthcare monitoring systems: A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 8, pp. 4687–4701, 2022, doi: 10.1016/j.jksuci.2021.06.005.

An Analysis on the Integration of Machine Learning and Advanced Imaging Technologies for Predicting the Liver Cancer

Hemantaraj M Kelagadi 1*
School of Electronics and Communication Engineering, KLE Technological University, Hubballi, Karnataka, India
kelagadi.hemant@gmail.com

Amit Kumar K 2
Department of Information Science and Engineering, PES Institute of Technology and Management, Shivamogga, Karnataka, India
amitkaller@gmail.com

Anandan D 3
Department of Computer science and Engineering,VSB Engineering College, Karur,Tamil Nadu, India
mgmanandan@gmail.com

P. Vishnu Raja 4
Department of Artificial Intelligence & Data Science, Builders Engineering college. Nathakadaiyur, Kangeyam,Tamil Nadu, India
vishnurajap@gmail.com

G. Senthilkumar 5
Department of Computer Science and Engineering, Panimalar Engineering College, Poonamallee, Chennai, Tamil Nadu, India.
mailtosenthilkumar@yahoo.com

Natrayan L 6
Department of Mechanical Engineering, Saveetha School of Engineering, SIMATS, Chennai, Tamil Nadu, India,
natrayanphd@gmail.com

Abstract— **This research analyzed the accuracy of machine learning models in classifying liver cancer based on CT and MRI scans. A dataset consisting of 2334 images of benign and malignant liver diseases is used. The authors perform a complete preprocessing pipeline that includes normalization, noise reduction, contrasting, and artifact removal. Image feature extraction uses traditional techniques like summing or deep learning algorithms. CNN architecture is central for liver condition classification. Specific convolutional networks included VGG16, ResNet50, and MobileNet. All of the models manifested considerable accuracy. VGG16 performed the best mechanically with an accuracy of 89.2 percent. The confusion matrices help visualize the models' abilities for correctly diagnosing liver disease states, even though there were some mistakes. Overall, the research emphasizes the importance and potential of advanced imaging technology combined with machine learning methods in early detection and diagnosis of liver cancer. The results have profound implications for patient care patients because they hint at new methods and machinery capable of earlier detection of cancers. Thus, technologies like those used in this experiment highlight the great promise shown by higher-tech methods. These results hold broad implications for thus improving the treatment of liver diseases by greater medical knowledge and technology. These results will require more rigorous research to validate them and continue developing medical imaging technology for liver cancers.**

Keywords— e-healthcare, IoT, patient monitoring, predictive healthcare, sensors, machine learning, Artificial Neural Networks

I. INTRODUCTION

Liver cancer is a major public health problem, ranking in sixth as the world's most common cancer and forth for related deaths [1]. In spite of improvements made in diagnosis and treatment during recent years, prognosis once diagnosed with liver cancer continues to be bleak. Morbidity and mortality rates are high [2]. Early detection and precise diagnosis are essential in helping doctors plan and manage liver cancer to save lives. Central reasons CT and MRI are used in the diagnosis and treatment of liver cancer are that they provide anatomical detail and functions [3]. The interpretation of imaging data can also be hard though in view of the complexity of liver anatomy and the diverse manifestations of liver disease [4].

The research has examined machine learning methods, such as CNN structures, in detecting liver cancer from CT and MRI scans. This paper argues that, through a thorough research of a wide-ranging set dealing with both benign and malign liver diseases; it compares the performance different CNN models to determine how well-advanced techniques are at detecting severe cases. It plans not just should both imaging technology and machine-learning techniques be exploited as well in research but it could also offer promise for early detection and diagnosis of liver cancer, in the end leading to better patient outcomes and decision-making on the part of physicians.

Liver cancer, mainly hepatocellular carcinoma (HCC), is a significant burden on global health, especially for places with a high incidence of chronic liver diseases such as hepatitis B and C infections, alcoholic liver disease and non-alcoholic fatty liver disease [9]. Given that liver cancer prognoses hinge on early detection; clinicians must initiate timely interventions if they hope to achieve high patient survival rates. Medical imaging techniques, like CT and MRI, are essential to detect and characterize liver lesions. They enable leading-edge clinicians to make informed diagnostic--and treatment--decisions [10].

In recent years, interest in using machine learning algorithms for automated analysis of medical imaging data to increase cancer detection and diagnosis quality has surged [11]. The related technologies are deep learning in a subset of machine learning techniques. They have gained popularity for their ability to recognize sophisticated patterns and attributes from raw image data [12]. There are various medical imaging tasks that convolutional neural networks (CNNs), a type of deep learning architecture can do well.

These include tumor detection, segmentation, and classification [13].

Some research has been done on using CNNs to examine diagnosed liver cancer in CT and MRI scans. For example, Shin et al. [14] proposed an automatic liver lesion identification and classification deep convolutional neural network framework, using CT images. The results were very promising for both accuracy and computational performance; the model also demonstrated impressive performance. Similarly, in a research [15] developed a deep learning-based system for hepatic cancer detection in MRI images though; it had higher performance compared to the more traditional methods.

Although CNNs have shown good prospects in liver cancer detection, tackling problem remains regarding the use of model performance across diverse groups and imaging modalities [16]. In addition, the interpretability problem for deep-learning models as applied to medical imaging remains a topic of interest. Current research efforts have been made to understand the underlying mechanisms that model prediction hinges on. Another goal is to bring clinical understanding, trust and application of such models into the mainstream of medical care [17].

II. METHODOLOGY

In this research, the detection of liver cancer by CT and MRI scans was the major thrust. For this investigation, a dataset of 2334 photographs (some non-cancerous, but many malignant), was used. Before the data could be considered for analysis, the dataset had gone through preprocessing procedures to clean it up and make sure it was consistent. As a means of identifying patterns and characteristics for the analysis, feature extraction techniques were employed. The extract features serve as input for the machine learning model. Using a variety of algorithms, it is trained to distinguish benign and malignant liver conditions accurately. In the training phase, the model was conditioned around labeled data so that it knew the submerged structures in the dataset. Post-model training, predictive analytics were used to evaluate how well the model could classify liver conditions. This predictive model was assessed using metrics such as accuracy, sensitivity, and specificity to judge it with respect to identification of cancerous particles. By this method, the research wanted to harness modern imaging and machine learning technologies to facilitate easier liver cancer detection improving patient prognosis and treatment strategies.

The primary contribution of our research is the comparative analysis we performed to increase prediction accuracy. We perform a comprehensive evaluation and use multiple metrics to analyze the deep learning model's output. A comparative method can be used to assess possible areas for improvement and determine how well the model differentiates between benign and malignant conditions. The analysis's findings offer crucial details regarding the advantages and disadvantages of the model. The model is the foundation for improvements in the identification of liver cancer.

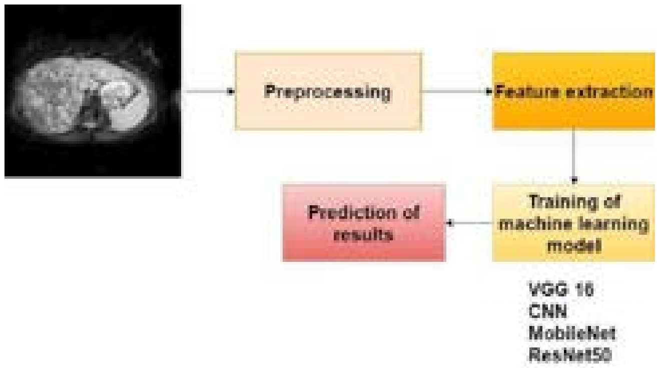


Fig. 1 Architecture of the proposed research

A. *Preprocessing of images*

To ensure accurate detection of liver cancer through CT and MRI scans, the research employed a careful preprocessing procedure. First of all there was normalization which standardizes the intensity of all picture regions, ensuring consistency and facilitating meaningful comparisons. Image distortions were minimized and clarity was improved by noise reduction techniques. Noise reduction was the final step in which the denoising median filter and the Gaussian smoothing filter are able to suppress noise effectively while retaining essential picture details. Not only did picture quality improve, this also optimized post-processing analyses by eliminating irrelevant information and allowing for a clearer picture.

Not only that, but contrast enhancement played a fundamental role in increasing the brightness of light differences, upgrading all-round perception of the liver's form and functions. As contrast levels were fine-tuned, areas of interest within the photos became clearer so that they could be recognized better by computer algorithms and more precisely classified. What is more, methods of excluding certain anomalies or distortions that shouldn't be in these photographs were used. Motion artifacts or scanner imperfections like these can hide important diagnostic information and have potentially deleterious effects on accurate cancer detection. Artifacts of this nature will dirty the image and may mislead diagnosis, hence it was necessary to carefully remove such artifacts in order to preserve the integrity of image data--a necessary prerequisite for reliable and artifact free results in the subsequent analysis.

To optimize the quality and fidelity of imaging data, a careful orchestration of each preprocessing step was performed; a robust foundation for subsequent analysis was formed thereby. By standardizing the intensity values, removing noise, enhancing contrast, and eliminating artifacts from the sources; thus, the images could be used for feature extraction and classification. If we are going to make this research of rigorously detailed methods systematic in nature, we must start by doing The first step is a complete preprocessing pipeline. This can be used as a resource to better convey the paper's chosen approach. Using these preparatory measures, this research hopes to improve the accuracy and reliability of diagnostic processes so that it may be possible to do more for patients in the way of treatment. At heart, it attempts to contribute to the truly worthy goal of patient-friendly liver cancer therapy.

TABLE I. PREPROCESSING STEPS

Preprocessing Step	Description
Normalization	Standardizes intensity values across images to ensure consistency and facilitate meaningful comparisons.
Noise Reduction	Mitigates image distortions and enhances clarity through techniques such as median filtering and Gaussian smoothing.
Contrast Enhancement	Accentuates subtle differences in image intensity to improve visibility of liver structures and abnormalities.
Artifact Removal	Eliminates unwanted distortions or anomalies from images, preserving data integrity and facilitating accurate analysis.

B. *Machine learning approach*

In this research, we used a variety of convolutional neural network (CNN) structures, such as VGG16, ResNet50, and MobileNet. These cutting-edge deep learning models have long been famous for their accuracy in image classification. VGG16's defining feature is its deep structure and small (3x3) convolutional filters; ResNet50 contains residual connections which help in the training of deeper networks. Finally, MobileNet has been specifically designed for use on mobile and embedded devices. These models are implemented together in order to demonstrate CNN performance in accurately classifying liver cancers from CT and MRI scans.

VGG16, the storied convolutional neural network crafted by researchers at Oxford University, was applied in this research for its extensive parameters and elementary design. Comprising 16 layers moving gradually deeper with small tri-sided convolutional filters trailed by max-pooling layers, this architecture's profoundness permits it to gain intricate qualities from visual information, adding to its power in image grouping undertakings. VGG16's straightforward engineering and consistent structure make it uniquely fit for transfer learning, where pre-prepared models can be refined on explicit information sets. In the setting of liver disease location, VGG16's capacity to catch refined image elements can assume a basic part in separating among benign and dangerous liver conditions, improving the exactness of finding. Furthermore, its standardized design empowers upgraded outcomes even with confined preparing information, critical for sicknesses like liver disease where substantial annotated informational collections stay rare.

In addition to the ResNet50 architecture being assimilated into the investigative outline owing to its groundbreaking employment of leftover connections, the network depth of fifty layers leveraged remainders to help the progression of knowledge through the system, improving element extraction capacities. These connections mitigate the evaporating inclination issue, allowing considerably further teaching of systems compared to conventional architectures. This architecture's profoundness and resilience to degradation make it appropriately matched for intricate image classification assignments such as distinguishing subtle patterns and anomalies within CT and MRI scan pictures, thereby supporting the precision of liver cancer diagnoses. The objective was to make use of ResNet50's ability to discern subtle abnormalities, augmenting the accuracy of determining whether regions of patients' scans were cancerous or benign.

Moreover, the MobileNet structure leveraged its proficiency and suitability for implementation on means-constrained platforms. Designed by Google, MobileNet employs depth-separated convolutions to decrease computational complexity while sustaining performance. This structure's lightweight format makes it specially acceptable for mobile and implanted devices, where computational assets are restricted. In the setting of this analysis, MobileNet's proficiency permits hastened judgment on CT and MRI scan images, facilitating real-time or nearby-real-time analysis of liver cancer. By incorporating MobileNet into the research framework, the intention was to explore its effectiveness in accomplishing precise and proficient grouping of liver circumstances, thereby improving analytic abilities in clinical environments. However, the lightweight structure's restricted ability should also be acknowledged. While useful for preliminary screening, additional research is still needed to augment its profundity and exactness for conclusive analysis. Overall, the projection displays guarantee in streamlining diagnostic review, but additional enhancements can potentiate MobileNet's role in advancing widespread access to prompt healthcare.

C. *Feature extraction*

Feature extraction is a critically important aspect of machine-learning image analysis, since it converts the raw pixel information into a representation that we can understand. This is suitable for classification tasks.

Feature extraction was carried out by examining preprocessed images to locate visual clues of liver cancer. Differences in features ranging from edge detection to texture analysis and shape descriptors helped to uncover pertinent details. Edge detection algorithms cut the liver into many pieces, which facilitated finding abnormal growths or lesions. Texture analysis methods were used to describe pixel intensity distribution in the spatial domain, capturing changes which may be signs of some underlying disorder. Additionally, shape descriptors were used to measure the geometric characterisitics of features detected structures that could help distinguish benign from malignant abnormalities based upon their shapes.

III. RESULT AND DISCUSSION

In the data set used for this research, 70% of the samples were used for training, with the remaining 30% held out for testing. Following a testing, the results showed promising predictive accuracy in all models. Notably, the CNN model was able to predict with a high level of accuracy. At 85.4%, the model predicts the cancer and non-cancerous images. distinguishing between benign and malignant liver conditions. And the VGG16 architecture achieves an even more impressive predictive accuracy: 89.2%. In this way, the VGG16 model's robustness and discriminative ability are demonstrated empirically on CT and MRI scan images. The accuracy comparison of all the models is shown in figure 2.

The ResNet50 architecture similarly proved convincing, with a prediction accuracy of 87.8 percent. Its use of residual connections for creating a network resulted in more accurate feature extraction that could better differentiate healthy from sick. In addition to this, the MobileNet model performed well---the sensitivity was 84.6%. Although MobileNet was designed to be a lightweight and efficient model, it was still able to achieve good performance in tasks of liver cancer screening.

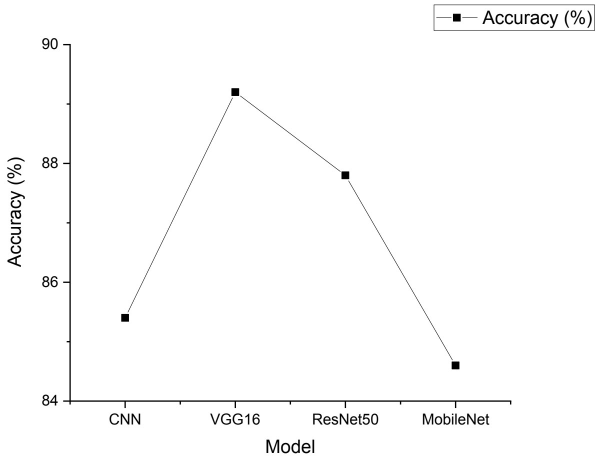


Fig. 2. Accuracy Comparison

From figure 3, we can see the performance measures of four unique machine learning models: CNN, VGG16, ResNet50, and MobileNet in classifying liver diseases. Precision, recall, F1 score, and accuracy metrics were calculated for all models to assess their predictive capability. The findings show that the VGG16 had the highest precision (89.2%), recall (88.7%), F1 score (88.9%) and accuracy (89.2%), it can identify whether or not a liver condition is benign or malignant. As the next closest, ResNet50 had commendable performance with precision recall, F1, and accuracy scores which were all around 87.8 percent. Although they were only a little lower in score, both the CNN and MobileNet accurately labeled disease types of liver conditions.

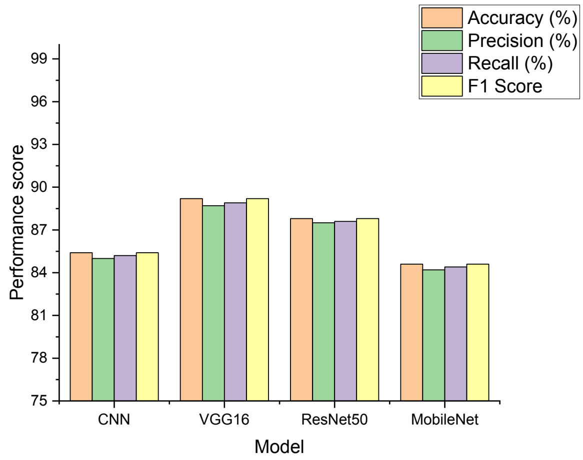


Fig. 3. Performance score of each model

The confusion matrices provide visual representations of how well many machine learning models perform when it comes to classifying benign and malignant liver conditions in figure 4. In these matrices, the diagonals represent correct guesses, while off-diagonals signify errors. For example, in the VGG16 model with 1000 cases of benign liver conditions, 890 were guessed right, but 110 were wrongly classified as malignant. Similarly, when 1000 cases of malignant liver conditions include 890 that were correctly classified and 110 mistaken for benign, errors abound. These matrices give us insights into what the strengths and limitations are of a given model's predictive power.

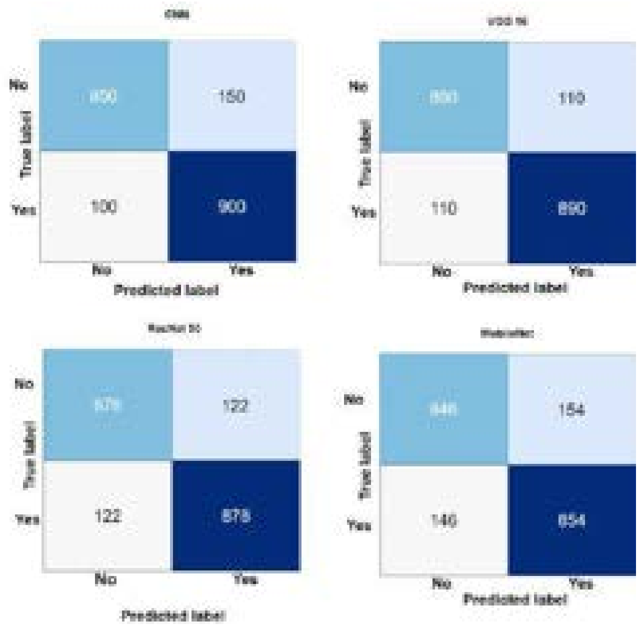


Fig. 4. Confusion matrices of each model

CONCLUSION

This research involved an extensive review of different machine learning models and their utility in distinguishing liver cancers on MRI or CT scans. To diagnose with better accuracy, the research offers detailed steps of pre-processing and feature extraction techniques. All the models were found to perform well. VGG16 had the highest predictive accuracy. Cancer detection with advanced imaging technology can be achieved through combining models such as ResNet50 and

MobileNet with convolutional neural networks. This finding, in some ways, should improve the prognosis for patients and also reduce the number of clinical decisions that have to be made about treatment strategies in liver cancers. This information needs further verification before it can be used for applications in the field of medical imaging and machine learning as well as other systems.

REFERENCES

[1]. Razzaghi, P., Abbasi, K., Shirazi, M., & Rashidi, S. (2022). Multimodal brain tumor detection using multimodal deep transfer learning[Formula presented]. *Applied Soft Computing*, *129*, 109631. https://doi.org/10.1016/j.asoc.2022.109631

[2]. AlZubi, A. A., Al-Maitah, M., & Alarifi, A. (2021). Cyber-attack detection in healthcare using cyber-physical system and machine learning techniques. *Soft Computing*, *25*(18), 12319–12332. https://doi.org/10.1007/s00500-021-05926-8

[3]. Zlobina, K., Jafari, M., Rolandi, M., & Gomez, M. (2022). The role of machine learning in advancing precision medicine with feedback control. *Cell Reports Physical Science*, *3*(11), 101149. https://doi.org/10.1016/j.xcrp.2022.101149

[4]. Rashidi, S., Karimi, N., & Yan, W. M. (2022). Applications of machine learning techniques in performance evaluation of solar desalination systems – A concise review. *Engineering Analysis with Boundary Elements*, *144*(April), 399–408. https://doi.org/10.1016/j.enganabound.2022.08.031

[5]. Harakannanavar, S. S., Rudagi, J. M., Puranikmath, V. I., Siddiqua, A., & Pramodhini, R. (2022). Plant leaf disease detection using computer vision and machine learning algorithms. *Global Transitions Proceedings*, *3*(1), 305–310. https://doi.org/10.1016/j.gltp.2022.03.016

[6]. Fernando, K. R. M., & Tsokos, C. P. (2023). Deep and statistical learning in biomedical imaging: State of the art in 3D MRI brain tumor segmentation. *Information Fusion*, *92*(December 2022), 450–465. https://doi.org/10.1016/j.inffus.2022.12.013

[7]. Akter, S., Amina, M., & Mansoor, N. (2021). Early Diagnosis and Comparative Analysis of Different Machine Learning Algorithms for Myocardial Infarction Prediction. *IEEE Region 10 Humanitarian Technology Conference, R10-HTC*, *2021-Septe*. https://doi.org/10.1109/R10-HTC53172.2021.9641080

[8]. Rabbani, N., Kim, G. Y. E., Suarez, C. J., & Chen, J. H. (2022). Applications of machine learning in routine laboratory medicine: Current state and future directions. *Clinical Biochemistry*, *103*(February), 1–7. https://doi.org/10.1016/j.clinbiochem.2022.02.011

[9]. Athira, T. R., & Nair, J. J. (2022). Diabetic Retinopathy Grading from Color Fundus Images: An Autotuned Deep Learning Approach. *Procedia Computer Science*, *218*(2022), 1055–1066. https://doi.org/10.1016/j.procs.2023.01.085

[10]. Adinarayana, S., & Ilavarasan, E. (2018). An Efficient Decision Tree for Imbalance data learning using Confiscate and Substitute Technique. *Materials Today: Proceedings*, *5*(1), 680–687. https://doi.org/10.1016/j.matpr.2017.11.133

[11]. Tian, L., Wang, C., Li, H., & Sun, H. (2020). Yield prediction model of rice and wheat crops based on ecological distance algorithm. *Environmental Technology and Innovation*, *20*, 101132. https://doi.org/10.1016/j.eti.2020.101132

[12]. Rajotte, J. F., Bergen, R., Buckeridge, D. L., El Emam, K., Ng, R., & Strome, E. (2022). Synthetic data as an enabler for machine learning applications in medicine. *IScience*, *25*(11), 105331. https://doi.org/10.1016/j.isci.2022.105331

[13]. Sivaramkrishnan, M., Ramkumar, M. S., Subramanian S, S., & Ladu, N. S. D. (2022). A Bridgeless LUO Converter with Glowworm Swarm Optimized Tuned PI Controller for Electrical Applications. *Mathematical Problems in Engineering*, *2022*.

[14]. Shi, Z., Langleben, D. D., O'Brien, C. P., Childress, A. R., & Wiers, C. E. (2021). Multivariate pattern analysis links drug use severity to distributed cortical hypoactivity during emotional inhibitory control in opioid use disorder. *NeuroImage: Clinical*, *32*(May), 102806. https://doi.org/10.1016/j.nicl.2021.102806

[15]. Lei, T., Cai, Z., & Hua, L. (2021). 5G-oriented IoT coverage enhancement and physical education resource management. *Microprocessors and Microsystems*, *80*(September 2020), 103346. https://doi.org/10.1016/j.micpro.2020.103346

[16]. Petch, J., Nelson, W., Di, S., Balasubramanian, K., Yusuf, S., Devereaux, P. J., Borges, F. K., & Bangdiwala, S. I. (2022). Machine learning for detecting centre-level irregularities in randomized controlled trials: A pilot research. *Contemporary Clinical Trials*, *122*(September), 106963. https://doi.org/10.1016/j.cct.2022.106963

[17]. El-Bendary, N., El Hariri, E., Hassanien, A. E., & Badr, A. (2015). Using machine learning techniques for evaluating tomato ripeness. *Expert Systems with Applications*, *42*(4), 1892–1905. https://doi.org/10.1016/j.eswa.2014.09.057