International Journal of Knowledge Based Computer Systems
11 (1) 2023, 01-07
http://www.publishingindia.com/ijkbcs/

Vulnerability Assessment and Mitigation Techniques on Hadoop Framework

Kimmi Kumari[1*], M. Mrunalini[2], Mamata Pandey[3], Sushma Verma[3] and Shubhojeet Paul[3]
[1]BIT Mesra, Ranchi, Jharkhand, India. Email: kimmikumari08@bitmesra.ac.in
[2]M S Ramaiah Institute of Technology, Bangalore, Karnataka, India.
[3]BIT Mesra, Ranchi, Jharkhand, India.
[*]Corresponding Author

Abstract: Hadoop has enormous data storing and large processing capabilities which makes it an efficient platform as compare to other platforms. In Hadoop much stress is given to deal with failures and coordination for complex distributed system. Early Hadoop Projects did not stance strong security measures. While new security challenges raised due to multi-occupant virtual environment of cloud computing. There is a phenomenal growth in cloud enabled services which also needs adoption of innovative security and privacy exploration. Due to flexible nature of Hadoop a strong mechanism is required to remove the vulnerabilities which can leads to security attacks. In this paper we discuss some vulnerabilities and effect of these vulnerabilities on the performance of Hadoop and which should be handled properly. Using vulnerability assessment and mitigation techniques, we have provided a security strategy for Hadoop in this study. This method is used because it can also reveal vulnerabilities which are unknown and helps in detecting the future threats while at the same time mitigating the present security threats.

Keywords: Assessment, Hadoop, Mitigation, Privacy, Security, Vulnerabilities.

I. Introduction

Information security model known for confidentiality, integrity and availability. While comparing to other platforms Hadoop is known for storing and processing enormous data efficiently. The early project of Hadoop stresses more on enhancing the technical aspects like developing the logic on how to deal with failures and coordination of such a large distributed system. It is considered that the distributed system's complete machine cluster is connected to a secure network, but it is not the reality. So Early Hadoop Project did not stand with strong security measures. Newer versions of Hadoop still evolving in terms of strong security measures and more stress is on protecting the tremendous data through encrypting the data and other data protection mechanism. Earlier it was assumed that Hadoop runs on trusted network and only authorized users were on the network but now encryption technique added for data transmission between the nodes. Hadoop is a distributed computing and parallel processing platform created by the Apache Software Foundation for batch processes. HDFS for huge storage, MapReduce for data processing, and YARN for cluster resource management are all parts of Hadoop. Big Data is handled by Hadoop, which makes use of various sorts of data that come from various sources. The Hadoop ecosystem is built on the Hadoop Distributed File System. Mass storage ability with the growing storage capacity is provided by the HDFS. Since HDFS filesystem logically spans in many servers so it is very important to understand the security perspective in HDFS or in all servers of the Hadoop cluster.

Transferring from the map phase to the reduced storage jobs is facilitated by aggregating the data volume as little as possible. MapReduce can be applied to transform data into executing business analytics. Lack of authentication and insecure communication between Hadoop daemons are two of the primary security issues with MapReduce. In Fig. 1, the main parts of Hadoop are shown. The component Job Tracker directs the MapReduce task to the nodes which having the data within the cluster. Task tracker accepts the task from MapReduce from the job tracker. Data node used to store the data in HDFS. Name node is used to make HDFS aware about each file location in the HDFS.

II. Related Work

Threats and risk are related to vulnerabilities. In a distributed system, vulnerabilities might take on a variety of different shapes [1]. Vulnerabilities are frequently found in the software itself. Every piece of software has weaknesses. Although it may seem harsh to say this, no piece of software is ever completely secure. An example of software vulnerability might be simply put; it is a piece of code that is vulnerable to a failure scenario or error that is not gracefully handled. Take the straightforward case of a piece of software that has a password screen that enables users to modify their passwords (we will assume that the intended logic for the software is to allow passwords up to 16 characters in length). What happens if the new password input form accidentally has a maximum length of 8 characters, truncating the selected password? Users could end up creating passwords that are shorter than they initially thought and, worse, less secure passwords that are easier for an attacker to guess.

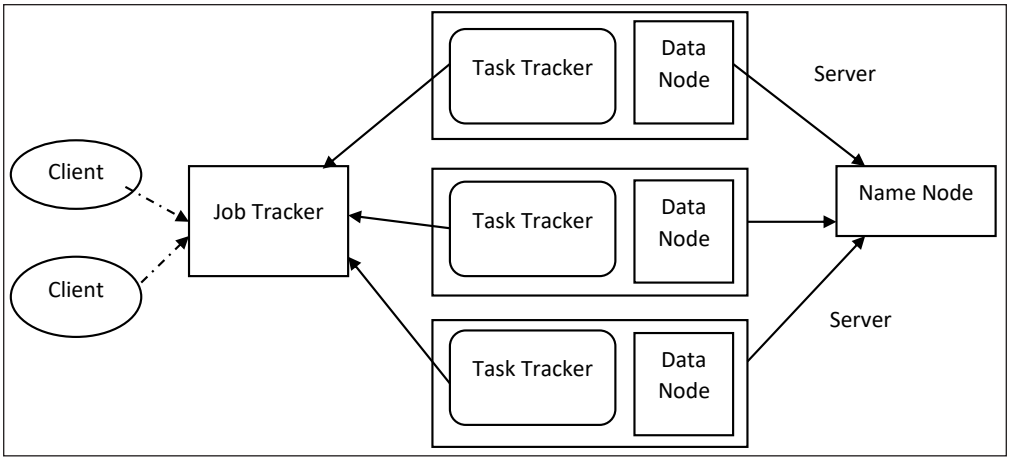


Fig. 1: Hadoop Architecture

Hadoop shares the inherent vulnerability [2] that any software possesses. Giving users access to master nodes may open the door for them to accidentally or intentionally use an unpatched Hadoop vulnerability. The more effective method of limiting access to the master nodes. Attackers may use software, a network, or a web interface to access all Hadoop components, like Sentry, Storm, and flink. Vulnerabilities are grouped into many categories because Hadoop is a combination of distributed computing software, hardware, application programme, and management rules. These categories could represent software flaws that allow hackers to target the Hadoop framework due to the language that it is written in [3] [4].

The "ping of doom" was the name of this assault. Although it has been mitigated, the fact remains that, prior to being fixed by network hardware vendors, this vulnerability had nothing to do with the software stack of the machines on the network, but an attacker could still use it to stop a specific machine on the network from performing its normal functions.

Every administrator's standard operating procedures should include regular schedules for installing patches to a distributed system's software stack since software patches are frequently issued to remedy vulnerabilities as they are found. The scope of patches should also cover the firmware for switches, routers, other networking hardware, disc controllers, and the server BIOS, as demonstrated by the ping-of-death example [4].

The security of big data can be seen from two aspects i.e., first while handling and storing information and the factors which affect the security of data and databases. Second aspect is while managing and maintaining the resources and dealing with operational safety issues with big data platform [5] [6].

It is very important to analyze the challenges while working with large data and management system when the Hadoop itself lacks the strong internal security management. In HDFS, cluster of servers may use different operating system platform and may leads to difficulty in level of patching [7] [8].

III. Proposed VAM Framework for Hadoop

While developing Vulnerability Assessment and Mitigation Framework for Hadoop, one of the most challenging tasks is to identify known and unknown vulnerabilities. Known vulnerabilities may come from Vulnerability database. After identification of these Vulnerabilities, evaluation should be done. While assessing or evaluating these vulnerabilities some deep learning methods or Artificial Intelligence may be useful. While evaluating a security or Vulnerability mitigation technique against a vulnerability some quantitative methods through deep learning can be utilized to present a relationship between individual vulnerability and its counter mitigation technique by a numeric value. This numeric value may be the outcome of these quantitative techniques of evaluator. The numeric value indicates whether a security technique can mitigate a vulnerability or may cause a vulnerability to incur. On the basis of these numeric values the relevance of the security technique whether it is primary or secondary against the vulnerability can be decided.

Detailed information of all the vulnerability mitigation techniques is analyzed and the study of appropriateness of a security technique among the available options must be presented in the comparative study. After organizing the risk mitigation techniques under primary or secondary importance a set of suggestions must be generated. The evaluation of risk mitigation techniques obtained through VAM methodology helps in refining the selection process of an effective risk mitigation model. After refinement of effective risk mitigation model one of the most effective vulnerability mitigation technique is selected according to priority.

In the proposed framework vulnerability assessment and mitigation for Hadoop can be distributed in various stages. Several challenging tasks need to be done to accomplish this task. These tasks can be performed in a phased manner.

- While adopting VAM framework for Hadoop, the most

challenging task is to identify known vulnerabilities. These known vulnerabilities may come from vulnerabilities database.

- If the threat is not recognized from the vulnerability database, it should be recognized as new novel threat.
- After identification of these Vulnerabilities, evaluation should be done. While assessing or evaluating these vulnerabilities some deep learning methods or Artificial Intelligence may be useful.
- For determining vulnerability and implementing countermeasures, we make use of recent developments in deep learning, which enable the development of systems that can learn from previous system events and predict a specific dangerous event that is likely to occur next. On the basis of the anticipated information, this system should recommend genuinely preemptive countermeasures.
- After assessing the specific vulnerability a matrix is prepared to show the relevance of specific security technique for each vulnerability.
- A comparative study of different vulnerability assessment and the counteraction is prepared.
- Finally the prioritization of vulnerability mitigation technique is decided.

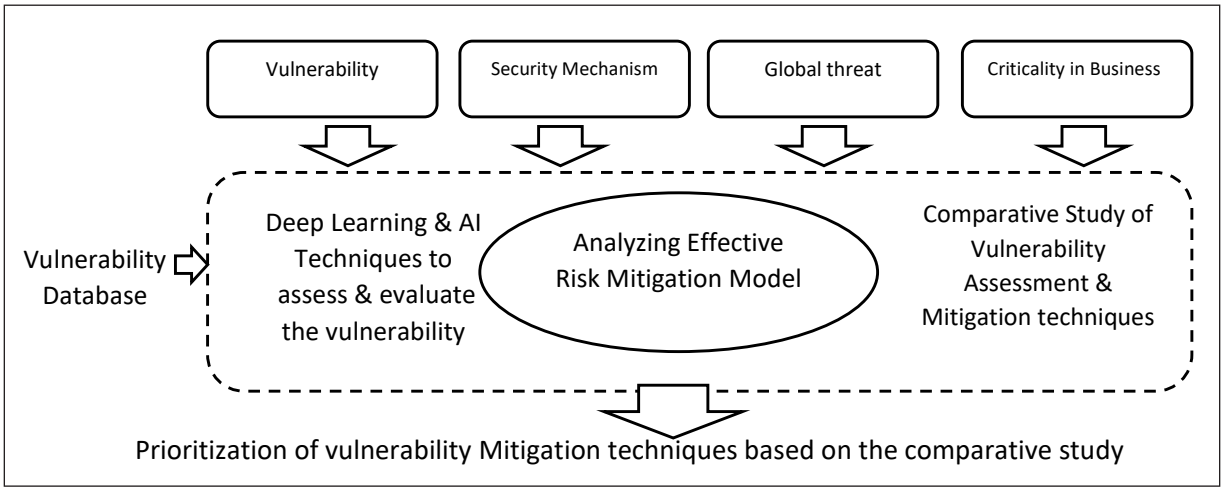


Fig. 2: Proposed Vulnerability Assessment and Mitigation Framework of Hadoop

Proposed Framework can be written in various steps:

Step 1: Identifying Vulnerability in Hadoop

While identifying vulnerabilities in Hadoop there are 3 types of vulnerabilities should be considered. First may be the technology related vulnerability i.e. the software platform in which Hadoop is developed and its possibilities to exploit by the cybercriminals. Secondly the configuration vulnerabilities for example there are many default settings need to be assigned to such a big platform. So these settings may be one of the targets by the attackers. The network security flaws, which reveal security breaches at both the service level and the data level, can also be taken into consideration.

Numerous vulnerability databases list numerous vulnerabilities, including a vulnerability management repository called the National Vulnerability Database (NVD) contains information on software problems that affect security. Another database that offers details on software vulnerabilities is the Computer Emergency Readiness Team (CERT). The Microsoft Security Response Center (MSRC) publishes Microsoft Security Bulletins, which are also connected to security flaws found in Microsoft products. Another database list of vulnerabilities with a unique ID is called Common Vulnerabilities and Exposures (CVE). The Open Source Vulnerability Database (OSVDB) offers truthful, unbiased data about security flaws. Numerous flaws in cyber-security software and hardware are listed in the CVE database.

Known vulnerabilities are those which can be handled by typical defense mechanism like DOS, script virus or worms' etc. unknown vulnerabilities may be an insider attack, network threat for which some special mechanism needed. Patch management is the process of retrieving the known and unknown vulnerabilities. Through vulnerability scanning the new vulnerabilities are discovered and resolved accordingly.

Step 2: Identifying Threats

Novel threats basically unknown threats which may be initiated by the insider or may be the network threats. These are harmful since threat protection mechanism is not well known.

For finding vulnerability or to identify insider threats conventional security approach are not sufficient. It is a challenging task due to user's behaviors. It can be done in various steps:

- To identify vulnerability or insider threat, log data is analyzed to detect threats.
- User behavior within the cluster is on surveillance to identify security incidents, monitoring violations and preparing baselines.

- Data mining methods are used to detect the abnormal behavior of the user and helps to undergone huge data to filtering out the true threats from the false positives.
- Vulnerabilities identified are shown through a basic metrics. Some systems may have dense vulnerabilities which may tend to gather in specific areas.
- Then the vulnerabilities shown with their counter action to fix and to repair. In this metric a clear picture will be depicted to how many vulnerabilities have been repaired.

Step 3: Vulnerability Assessment and Mitigation in Hadoop

The goal of vulnerability analysis is to identify as many security flaws as possible (a "breadth over depth" approach). It should be used frequently to keep a network safe, especially when there are network changes (e.g., new equipment installed, services added, ports opened). Additionally, it will work for firms that want to know about all potential security flaws but are not yet security mature.

After identifying vulnerabilities which is most challenging task we need to assess the vulnerability and to establish an effective vulnerability mitigation model.

- While assessing the vulnerability, its severity should be known. It can be assessed using threat likelihood which me be judged by exploitability, discoverability and reproducing capability.
- Vulnerability needs to analyze the impact on confidentiality and integrity.
- Each vulnerability should be ranked from critical to low relied on its severity and possible effect on Hadoop to create a vulnerability score system.
- Estimation of number of vulnerability which can be repaired should be done.

Refer Table II for the evaluation.

After identifying vulnerabilities which is most challenging task we need to assess the vulnerability and to establish an effective vulnerability mitigation model.

While assessing the vulnerability, its severity should be known. It can be assessed using threat likelihood which may be judged by exploitability, discoverability and reproducing capability.

- Vulnerability needs to analyze the impact on confidentiality and integrity.
- Each vulnerability should be ranked from critical to low relied on its severity and possible effect on Hadoop to create a vulnerability score system.
- Estimation of number of vulnerability which can be repaired should be done.
- Appropriate mitigation technique is adopted.
- Risk mitigation techniques for Hadoop must consider the following approaches to secure the system:
  - Ensure security of tools and techniques.
  - User account management and authentication.
  - Access control.
  - Securing transaction logs and large data sets.
  - Protection for software and hardware configuration.
- To protect Hadoop system from multiple novel attacks of hackers data analytics and AI techniques are used which can detect unknown vulnerabilities. Extended threat detection and false positive rate can helps analysts to detect unknown security breaches immediately.
- A comparative study of risk mitigation models can help in selection and prioritization of appropriate model.

Step 4: Identifying Mitigation Related to Vulnerability

For identifying mitigation techniques for specific vulnerability first we need to discover the vulnerability and then it can be exploited to find appropriate mitigation technique. Detailed process of mitigating vulnerabilities is discussed in Table I.

Table I: Vulnerability Assessment and Mitigation in Hadoop with Prioritization of Mitigation Technique

Attack (Vulnerability) Name	Evaluation/Assessment of Vulnerability	Effective Vulnerability Mitigation Technique	Priority of Vulnerability Mitigation Technique
Protect Session Privacy	• It is vulnerable to attack session privacy. • Session privacy is susceptible to hacking, whether it is used for communication between clients and data nodes or among nodes.	SSL (Secure Socket Layer)/TLS (Transport Layer Security)	• Transport Layer Security (TLS) is the first choice. • Transport encryption can protects all communication from access or modification by attackers. • But the TLS is difficult to implement and get certificate management right.
Data Usage	• System enforces to unapproved authorization to logical access of information and system resources.	Identity and authorization/ Masking/Application encryption	• Priority is given to Secure Shell (SSH) authentication within the Hadoop cluster.

Attack (Vulnerability) Name	Evaluation/Assessment of Vulnerability	Effective Vulnerability Mitigation Technique	Priority of Vulnerability Mitigation Technique
			• Identity and authentication can be used as a central security effort and considered as a second priority technique.
Data at Rest (External Threat)	• Vulnerable to attack on HDFS file system where data is stored.	Application/Object (HDFS) encryption	• In Hadoop, the file system contains embedded encryption. This indicates that data is encrypted transparently when it is placed into the file system, without modifying the programme running in the cluster.
Data at Rest (Credentialed Users)	• Vulnerable to tenant data privacy in multi-tenant clusters.	Application encryption/External key management	• The system can be linked with key management services from third parties or used with Hadoop's Key Management Service (KMS). Access Control Entries (ACE) or Access Control Lists (ACL), which are essentially file permission constructs, are utilized by some versions of Hadoop.
Node Authentication & Validation	• In Hadoop user application runs on a cluster of machine so vulnerable to attack any of the specific machine. • A node's identity can be forged.	PKI (Private Key Infrastructure)/Kerberos	• HDFS supports a variety of authentication schemes including PKI-based and Kerberos. • Although Kerberos credentials are accepted, those credentials are not passed along to the workers. • Therefore, jobs that need Kerberos keys to access resources.

Step 5: Vulnerability Mitigation through VAM Methodology

Security of Hadoop ecosystem based on proper functioning of three parts which are identity, authorization and authentication are discussed in detail in Table II. Due to distributed nature of Hadoop system nodes within the cluster are loosely coupled from authoritative identity sources. To ensure the security of data of all nodes various security protocols may be imposed. Below table shows the Vulnerability Assessment and Mitigation for Hadoop through VAM methodology.

TABLE II: VULNERABILITY ASSESSMENT AND MITIGATION FOR HADOOP THROUGH VAM METHODOLOGY

Vulnerability Causes	Effect of Vulnerability	Mitigation Technique
Identity/Data Breaches	• Files stored in HDFS may be unsecure. • Data in files can be maliciously accesses. • File transfer may be unsecure within and outside the cluster.	• Files within HDFS system can be encrypted. • SSL needs to be applied for MapReduce and web consoles. • Encryption should be adopted during HDFS file transfer.
Malicious Attack	• Lack of authentication between client and services. • Lack of authentication between Data node, Name node, Task tracker and Job tracker.	• With the use of single sign on Hadoop public key cryptography and other authentication mechanism can be applied. • Authentication of all the clients for all services can be performed through single server.

Vulnerability Causes	Effect of Vulnerability	Mitigation Technique
	• Unauthenticated delegation of tasks.	• For storing the encryption keys of authenticated users an effective credential management framework can be used.
Unauthorized Access Privileges	• Wrong file permission within HDFS system may be assigned. • Access privilege list may be controlled maliciously. • Organization job queues may be distorted in terms of priority sequence.	• Authorized HDFS file permission should be granted. • Access control list and job queues of various user groups should be verified before execution. • Fine grained and role based authorization should be imposed.

IV. Result and Discussions

Hadoop environment must be protected from a network security perspective. The vulnerability due to distributed nature of Hadoop can be addressed through various security mechanism viz. imposing firewalls, IDS and protecting the communication between the nodes. The VAM methodology helps in gaining the insights to employ specific security technique for handling vulnerability. This methodology also review the existing vulnerabilities and generate the security options to help in prioritization among these options. VAM methodology also identifies the future vulnerabilities and the level of mitigation techniques which may handle these unexplored vulnerabilities. The main focus of VAM methodology is to protect the information system but if we see in broader aspect it presents a detailed review of vulnerability and attacks which could be useful in understanding the nature of vulnerability and its mitigation technique. It also presents a comprehensive study to adopt any defending action during and after an attack. This methodology is very useful when the system is complex like Hadoop as well as when a new component is to be added to project, as it can look for new threats which may cause system failures or may be exploited by the malicious users. Since Hadoop is operated on cloud environment also so new components may be frequently added due to various reason viz. scalability, and efficiency etc. traditional security assessment system only look towards exploited vulnerabilities which is not the case in VAM methodology.

V. Conclusion

As Hadoop is used to store and process enormous data, big enterprises prone to use it. While using Hadoop for organization security administrators face the challenge to mitigate all kinds of threats and vulnerabilities to protect organizations data. It is also true that there is no single threat for such a big distributed system rather new threats may arise frequently. To protect the system from these threats mitigation techniques and many security control strategies must be adopted. More than one security control is applied to achieve a comfortable table off security. The idea of implementing multiple security control is called defense in depth.

In this paper we have discussed Hadoop framework, vulnerability, patch management, and security issues in big data/HDFS. We have also proposed vulnerability assessment and mitigation framework for Hadoop. This framework stresses on identifying vulnerability and to control these vulnerability through the counterpart mitigation techniques. In our framework we have used deep learning and AI methods to analyze the threats and effective mitigation techniques. The review of the vulnerability and their possible mitigation techniques helps in comparison and analysis of the effective models. VAM methodology helps in decision making for selection of appropriate vulnerability mitigation techniques against the threat likely to occur. VAM methodology helps in development of guidelines for security measures with recommended priority of mitigation techniques. It is also beneficial for risk analysis and cost effective options for mitigation of exploited vulnerability as well as future threats.

References

[1] R. Pandey, and C. Verma, "Big data representation for grade analysis through Hadoop framework," In *2022 6th International Conference-Cloud System and Big Data Engineering (Confluence),* IEEE, 2022, pp. 312-315.

[2] A. M. Martinez-Enriquez, M. Adnan, M. Afzal, M. Aslam, and R. Jan, "Minimizing big data problems using cloud computing based on Hadoop architecture," In *2014 11th Annual High Capacity Optical Networks and Emerging/Enabling Technologies (Photonics for Energy),* Charlotte, NC, USA, 2014, pp. 99-103.

[3] E. Fernández-Medina, J. Moreno, and M. A. Serrano, "Main issues in big data security," *Future Internet*, vol. 8, no. 3, p. 44, 2016.

[4] B. B. Rad, N. Akbarzadeh, P. Ataei, and Y. Khakbiz, "Security and privacy challenges in big data era," *International Journal of Control Theory and Applications*, vol. 9, no. 43, pp. 437-448, 2016.

[5] D. H. Manjaiah, and M. A. Shetty, "New security architecture for big data Hadoop," In N. Shetty, L. Patnaik, N. Prasad, and N. Nalini (Eds.), *Emerging Research in Computing, Information, Communication*

and Applications (ERCICA 2016). Singapore: Springer, 2019.

[6] R. Rapuzzi, and M. Repetto, "Building situational awareness for network threats in fog/edge computing: Emerging paradigms beyond the security perimeter model," *Future Generation Computer Systems*, vol. 85, 235-249, 2018.

[7] G. S. Bhathal, and A. Singh, "Big data: Hadoop framework vulnerabilities, security issues and attacks," *Array*, vol. 1-2, p. 100002, 2019.

[8] E. Spivey, *Hadoop Security: Protecting Your Big Data Platform*. O'Reilly Media Inc., 2018.

International Journal of Knowledge Based Computer Systems
11 (1) 2023, 08-20
http://www.publishingindia.com/ijkbcs/

Automation of Rental House Management System: An Empirical Approach with Possible Results for Society 5.0

Rohit Rastogi[1*], Rayush Jain[2], Prabhinav Mishra[3] and Prateek[4]

[1]Associate Professor, Dept. of CSE, ABES Engineering College, Ghaziabad, Uttar Pradesh, India.
Email: rohitrastogi.shantikunj@gmail.com
[2]Student, B.Tech Second Year, Dept. of CSE, ABES Engineering College, Ghaziabad, Uttar Pradesh, India.
Email: rayushjain8099@gmail.com
[3]Student, B.Tech Second Year, Dept. of CSE, ABES Engineering College, Ghaziabad, Uttar Pradesh, India.
Email: pravhinavmishra@gmail.com
[4]Student, B.Tech Second Year, Dept. of CSE, ABES Engineering College, Ghaziabad, Uttar Pradesh, India.
Email: prateeksingh6125@gmail.com
*Corresponding Author

Abstract: Housing Management System is a web based project through which a user can easily access all the rental houses around his workplace in just a few clicks. Customarily, people have to contact the brokers for finding the rental homes and also have to visit several places. This is a tedious process and can take a lot of time but our web page will create a direct link between a user and the owner with transparency. As the user opens the web page there will be two options to register and login. When the user clicks the register option a new tab will open with two options. Register as tenant or register as owner. User has to select his category and fill in a few details along with his mobile number and email ID. Then the OTP will be generated and the user has to enter that OTP and create a password to register. For log-in, the user has to use his mobile number or email ID along with the password. After Log-In users will have complete details along with the images of all the available rental flats and houses around the location which he has chosen. We will also provide the categories which will further help the user to find the place to live as per his convenience. Once the user finds the place which fulfills his needs, he can directly contact the owner. There will be one more available option for the owner who can also add new rental homes as per availability. So our webpage will act as a broker between the owner and the tenant with no extra cost and transparency. Both the parties will be in direct contact and this will help in saving time and money for the users. Our webpage will use technologies such as HTML, CSS, and JavaScript. We will add new and powerful features to our web page which will provide easy user friendly access and an attractive interface.

Keywords: CSS, HTML, Java Script, UI, Web application.

I. Introduction

A. Rental Housing Management System

Housing management comprises two separate words one is housing and the other is management, as we can see in simpler terms managing houses is what should be called as housing management but it has a broader perspective. Traditionally, housing management was all about rent collecting, allocation and maintenance but if we talk about modern day scenarios many small tasks are also added in it like interactive sites, community involvement etc. So to tackle all these things we came up with an online Housing Management System in which the user will be free to explore various properties, rental houses etc at their convenience which is an astonishing thing. The client can book their favorite property online with just a few clicks and they can also directly talk to the owner with details provided at their end [7].

The above image shows us beautifully how we have taken a leap from normal housing management to an online housing management system. The house is connected to the laptop with a cable and the image of the house is shown in the laptop screen which shows that the house is up for rent or sale via online housing management system (as per Fig. 1).

B. Global Records and Statistics on RHMS

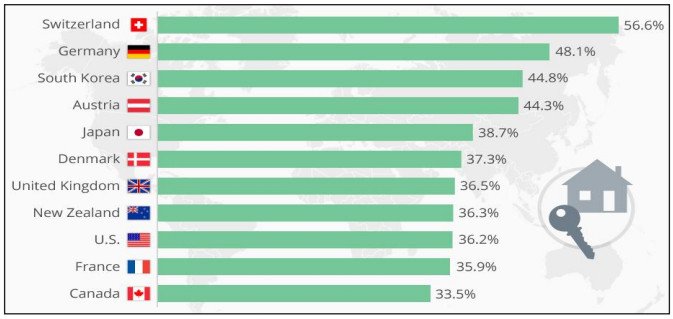
In the earlier times all the records were maintained in registers and booklets which was a very hideous task. But after the introduction of the online system the work has become easier and better. Earlier, people must be present physically at places if they want to ensure the status and quality of the property

they are going to rent or buy but after the introduction of the online system all these things can be checked through online mode. Due to this much convenience many countries are using this online housing management system. If we see the records then an approximate of 30% people live on rent in India, 36% in the United States, 11% in Russia and many more countries. All these countries use online housing or property management systems for making their work easier [7].

SourceURL:

https://s3.amazonaws.com/assets.blog.rentpost.com/wp-content/uploads/2015/06/05003530/online-property-management.jpg

Fig. 1: Online Housing Management System



SourceURL:

https://cdn.statcdn.com/Infographic/images/normal/13034.jpeg

Fig. 2: Countries with Most Renters

The above image shows us the statistics of how much renters are there in many different countries from the lowest to highest. The global analysis has shown that the U.S. and U.K. has shown a major growth in the number of households renting their homes which in turn results in the downturn of the country's economy (as per Fig. 2).

C. Need of Rental Housing Management in 21st Century Society

If we look at the present scenario, due to the decrease in the availability of land, there is a hike in the price of properties as well as rental homes especially in smart cities like Mumbai, Delhi, Bangalore and others. People from all over the country visit these places for jobs or for educational purposes and the wages of the people is less as compared to rent and also due to the less availability of the land the people also don't know about the unknown places and places with less rent [4].

SourceURL:

https://www.baymgmtgroup.com/wp-content/uploads/2021/04/shutterstock_9228+5236-min-400x400.jpg

Fig. 3: Finding the Perfect House for Yourself

The above image shows us how we can easily find the perfect house according to our needs using the online rental house management system. The online rental house management system helps its user with various types of features which enhances the user experience and eases the task of finding the perfect house (as per Fig. 3).

D. Automation and its Need of HRMS in 21st Century Smart Society

Rental Housing Management System Automation is needed in the 21st century for saving time, reducing human errors, and doing your work efficiently. Nowadays automation is in our daily lifestyle which we also see in our daily work like using applications and phones. The need for Rental Housing Management System to save or record data carefully and it's a type of data collection where there is no chance of duplication where every data is stored on an online server where users are able to see it at any time. A Rental Housing Management System is also needed because of the daily increase of hostels that take charge of insignificant students which also resolved their problem and in our RMS you can easily open it with one web browser. It also shows you the nearby location according to your need and provides you various services like GPS, Retail food delivery, etc. In the 21st Century, it saves a lot of money and time [2].

SourceURL:

https://brazenrealestate.com/wp-content/uploads/2019/10/rent-your-property-checklist.jpg

Fig. 4: Checklist of House Necessity

The above image demonstrates that with the help of an automated rental house management system one can make a checklist of his/her demands and will sort out the best match from the online application which will save their time and money (as per Fig. 4).

E. The Impact of Housing Management System in Country's Development

Our webpage will provide a user friendly system that will be useful to all the classes of the society. So, it will help provide shelter to all the category of users. Hence it will be helpful to improve the quality of life of people and will also lead to the economic and social development of the nation. Our online rental housing management system will be very helpful for the citizens of the country in accommodating the house and property they want with proper chat and concern of the owner which will lead to minimum disputes between the citizens [4] [8].

SourceURL:

https://encrypted-tbn0.gstatic.com/images?q=tbn:ANd9GcSD8efqtUfDZKEHir5VESH5xoZyXhn3yN264w&usqp=CAU

Fig. 5: Satisfied User Shaking Hands with the Property Owner

There is transparency between the owner and the tenant as both are in direct contact. This leads to mutual trust and understanding between the tenant and the owner (as per Fig. 5).

II. Literature Review

Rathore *et al.* (2021) and his team demonstrated a Rental House Management System. They designed an online house renting system for the ease of the citizens. They found that nowadays finding an affordable renting house is a major issue and the leading websites in these sectors such as 99 acres, Magicbricks are not well managed, So they came up with a new system which can simplify the work for its users and all their work such as rental house search and property dealings can be effective.

The system is operated on the Django framework with a user-friendly interface. When a user opens the application, the user must sign-up and log-in. If he/she wants to talk to the house owner and if someone wants to post their property online then they have to log-in as an owner on the application. The application is developed using the following:

- Django
- HTML CSS
- Java Scripting
- Python
- MySQL

The outcome of the project was that they developed an application that has improved the lifestyle of the people as they can easily find rental property. The owners on the other side can find tenants for their rental property. This application works as a matchmaker for both the tenants and the owners. The Web based Rental Housing Management System is an advantage for both tenants and landlords [6].

Misyam *et al.* (2021) and his team demonstrated that a house is a place where people gather; they can improve their lifestyle, and can grow economically as well as socially. Their team depicted that initially buying a house is a big task for an individual because it needs a very high capital amount, so a person has to first rent a house and then save a handsome amount for buying the house.

Their team researched and saw that there are some loopholes in the current management system such as the customers are unknown of the rental houses available in their area, they don't know how to contact the owner of the property as a result a rental house management system was created to render out such shortcomings. The user interface is very easy to understand and very convenient for the customer, they simply have to login into the software and have to find out the correct house that matches all their needs.

The methodology that was used in creating their rental housing system was prototyping and the programming languages used were Hypertext Markup Language (HTML), CSS, JS, Hypertext Preprocessor (PHP) and MySQL database.

The conclusion of their research was that they were successful in creating a web based application which will help the customers in making their lifestyle more convenient and their system is developed in the sight of all the requirements that were needed for the betterment of their online rental house management system [5].

Kirmani *et al.* (2017) and his team demonstrated that there is a need for an affordable housing system which will provide shelter to all the classes of the society and will increase the quality of life of people which will add to the country's prosperity and development. This software is based on spatial data and user specification based searching which will provide a user friendly interface.

The software is also provided with a cart system through which the user can select the properties which he likes to view for later. Space Spatial database is used for providing geographical information of the rental houses. This is a basic web based project where the user has to register and login and then he is eligible to rent, buy or sell the properties. The developers have also tried to reduce the search time.

This rental housing software is based on the ASP.NET which is a part of Microsoft .NET with the help of C# and the database engine used is SQL 2008. It uses a client/server architecture based on the HTTP protocol.

This project provided customized software which will provide a database to store information with data security for the landlord. So it will be beneficial to both the user and the landlord [4].

Ikuomola *et al.* (2020) and his team demonstrated that there is a need to build a customized system to manage the property dealing due to increase in data and lack of data security. In this work, a secured mobile cloud-based house rental management system which will help landlord and tenant to overcome the hassles and manage tasks and issues easily in a more convenient way is designed.

This is web based software which is used for making the renting easier for the user. The team has used the cloud service with components such as app server, payment engine, wallet engine, chatserver, message engine and authentication server. Centralized databases using MYSQL and programming languages are angular JS, PHP and Laravel. A cloud based mobile application with user friendly interface and secure environment is established.

Integrated system can be used i.e. if the tenant does not have a device to contact the owner then him/she can use the smart door technique.

House rent countdown system can also be added to remind the tenant of the rent [3].

Gommans *et al.* (2014) and his team demonstrated that in all over the country housing plays an important role in revealing the economic growth in any country, being a shelter among the keys of development. And the Universal Declaration of Human Rights gives one of the basic human's rights as the right of decent living by the central government which allows us to live in decent housing (United Nations, The Human Rights-article 25, 1948). The object of this research is to manage the housing for low income, medium and high incomes households that can allow a human to live in proper or standard housing which is affordable to pay the house rent. The advantages of this Rental Housing Management System is to easily calculate or manage the all over the data that takes several days to manage a large number of tenants that paid their rent.

This is a database type system that is able to access all the information and is to manage and it is especially bilt for rental home owners.

- HTML
- CSS
- JavaScript
- PHP

This project is database type software that is used to help landlords to manage the records or pay history of tenants that's resolved the problem of storing records and also calculate the profits [2].

Afzal *et al.* (2021) and his team demonstrated that if you are a stranger in some city and wanted to rent a house. Then, it is very difficult to find a suitable house at a time that meets your requirements. This is the main motive to build an Online Rental Management System. So, this website helps users to register individual homes or apartments to help you in finding the perfect rental home that is also used for property selling, leasing, and buying. Which makes it easier for tenants to choose a desired house or property within budget and Landlords can easily sell their property by uploading and registering the website.

And the loopholes in the existing system only provide a text-based interface which is not very as user friendly as a Graphical user interface. Since the system is implemented manually the response is very slow. Hence, there is a need for the reformation of the current system to a new system that has more advantages and flexibility. The main objective of the system is to eliminate the limitation of the existing system.

- To remove the paper-based work.
- To save time for both landlords and tenants.

The methodology that they used to create in their Online Rental Management System is Structured Query Language (SQL), PHP will be used to collect or store the data in a database. And for web design they used HTML, CSS, Bootstraps. It consists of three parts: Client, Web Server, Database Server.

The conclusion is that they were able to successfully developed and Online Rental Housing Management System that have also mostly very favorable for hostler and PG person that wanted to search a suitable house and it also have an GPS map system in it which also save a lot of time by searching it automatically show the direction of it which it is easy to navigate. It is an Online Website which is easy to use and also has the standards of current time [1].

TABLE I: SUMMARY OF REVIEWED PAPERS

Paper Title & Author	Introduction	Methodology	Data Set & Algorithm	Conclusion	Future Scope
Rental House Management System [3].	Their team built a web based application which will make the online renting system more efficient and convenient. They resolved the issue of difficulty in finding affordable rental property.	When the user opens the web application he must log-in and sign-up if he wants to talk to the owner and the same goes for the owner side if he wants to put his property on the web application.	MySQL Database on phpMyAdmin. Apache server on Xampp. ASP.NET using C#.	The outcome of the project was that they developed a web application that made the user lifestyle easier in finding rental houses and the owners can also find the tenants of their liking for the rental property.	In the current scenario there are a number of people who are living in rental houses and are finding rentals to live in so this web application will make their task of finding rental property easier and more convenient.
House Rental Management System [5].	Their team built up a web based application which will help the customers to find the nearby rental properties and they also made sure that there is a proper gateway of conversation between the parties.	Their software works on a very simple method, the user has to login and have to find the proper rental house which meets their demands and can contact the owner of the property for further details.	The languages used to build their application were HTML, CSS, JS, PHP and the database used was MySQL database.	Their team concluded that they have finally built a software that meets all the requirements according to the demands and it will surely help in enhancing the lifestyle of the people.	Buying a house is not an easy task for an individual hence one has to first rent a house and then save up for buying a house. So, their web app will help such people who are in need of a convenient rental house. They will also add up new features for enhancing the user experience.
Rental Housing Management System [4].	This is a software based on ASP.NET. The base idea of this research is to provide "affordable" shelter to all the classes whether lower upper or middle class.	This software has user specified searching and direct contact between buyer and seller. Cart system is also provided. Easy addition, deletion and searching of houses with user -friendly interface.	ASP.NET using C#. Database used is SQP 2008. IIS and the .NET framework	This online software provides exact information and can be accessed easily from anywhere. Cart system and mail based information increases its demand in the market.	This software uses 2010 .NET version which is now upgraded to .NET 2015 version. Dropdown searching and satellite/map based searching can also be added in this software.
Cloud Based Rental Housing Management System [3].	This is web based software which is used for making the renting easier for the user.	The team has used the cloud service with components such as app server, payment engine, wallet engine, chatserver, message engine and authentication server.	Centralized databases using MYSQL and programming languages are angular JS, PHP and Laravel.	A cloud based mobile application with user friendly interface and secure environment is established.	Integrated system can be used i.e. if the tenant does not have a device to contact the owner then he/she can use the smart door technique. House rent countdown system can also be added to remind the tenant of the rent.

Paper Title & Author	Introduction	Methodology	Data Set & Algorithm	Conclusion	Future Scope
Rental Housing Management System [2].	Most of the work done by the housing managers is on paper and maintaining records is really difficult. So their webpage will basically help the user to maintain data in an organized way and will also prove a user friendly interface for both the owners and the tenants.	Firstly, you have to log in then you have to select an option that is visible on the website. It will show how the data is collected from the user through the system that has been kept in the system database.	Database system design, framework, server address, etc.	This software is used as an inventory to provide a framework that enables the managers to make only reasonable transactions made within a limited time of the framework.	This system software has to also give a pay slip and data that are directly connected to the Microsoft Access that automatically sends the slip of recorded payment to the tenants.
Online Rental Housing Management System [1].	Their team make a website of Online Rental Housing Management System which is built on the basis of eliminate paper based word, save time and give an desirable result according to the user need	Firstly, they used web servers like SQL and PHP that enable the website to store or collect data in a database. And they used HTML, CSS, Bootstraps. That contains three parts: client, web server, database server that help to build an online website.	Structured Query Language (SQL), PHP, HTML CSS, Bootstraps.	This website is used for tenants, landlords, hostel, PG, etc. It is more flexible than the existing system. It has more accuracy and compatibility in terms of saving time and working output. It is an online website where a user has to firstly register and fill out the details then the recommendation of that similar output shows on the web page. It is very beneficial for students.	This is a modern Online Rental Housing Management System of Web application that is built by PHP and SQL. And it also has a client/server HTTPS protocol. For the future scope this system has a GPS server which helps to give the location of a house. It is a user-friendly system which makes it easy for a tenant to search a desired house and for landlords which just have to input the details of their properties.

III. Methodology and Setup of Experiment

For the proper working of our web application we collected data from many resources. We talked to people (customers) so that we can understand the problems they are facing in the renting system. We also collected views of people on the current online websites which provide an online renting system and noted down the problems faced by them or anything that caused any kind of inconvenience, so that we can enhance the quality of our work.

A. Setup

Steps

1. Initially we started with the identification of the problem statement and analyzing the major faults and issues.
2. Our team asked the opinion of various people on the current management system and asked if there was any kind of issue.
3. We collected data from the current working websites on renting management systems and seek for any type of issue caused.
4. We talked to various owners and their requirements like what they need for a better experience on the website.
5. After collecting all the required data all the team members analyzed it, and started its interpretation.
6. Then we start to learn HTML and CSS for making the front end of our web based application.
7. After that we learned the languages required for making the back-end of our rental house management software.
8. After gaining all the knowledge required the team divided the work and started to make the web application.

B. Use Case Diagram

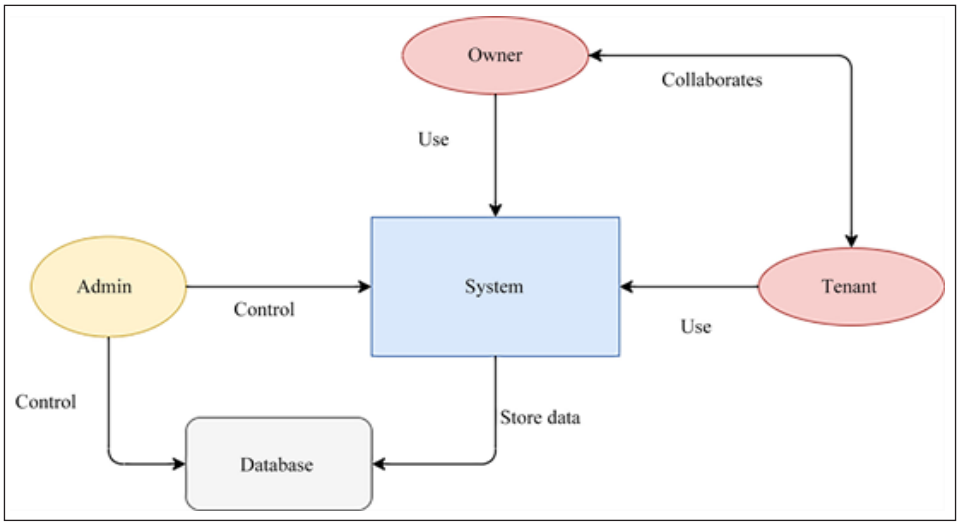


Fig. 6: Basic Use Case Diagram Showing How the Rental House Management System Works

- This diagram explains how our website orientation will be like.
- The Admin or developer has the control access to the System and database and can modify the system.
- The tenant and the owner can collaborate if the needs of both are matched.
- They both will also have the user access to the system (as per Fig. 6).

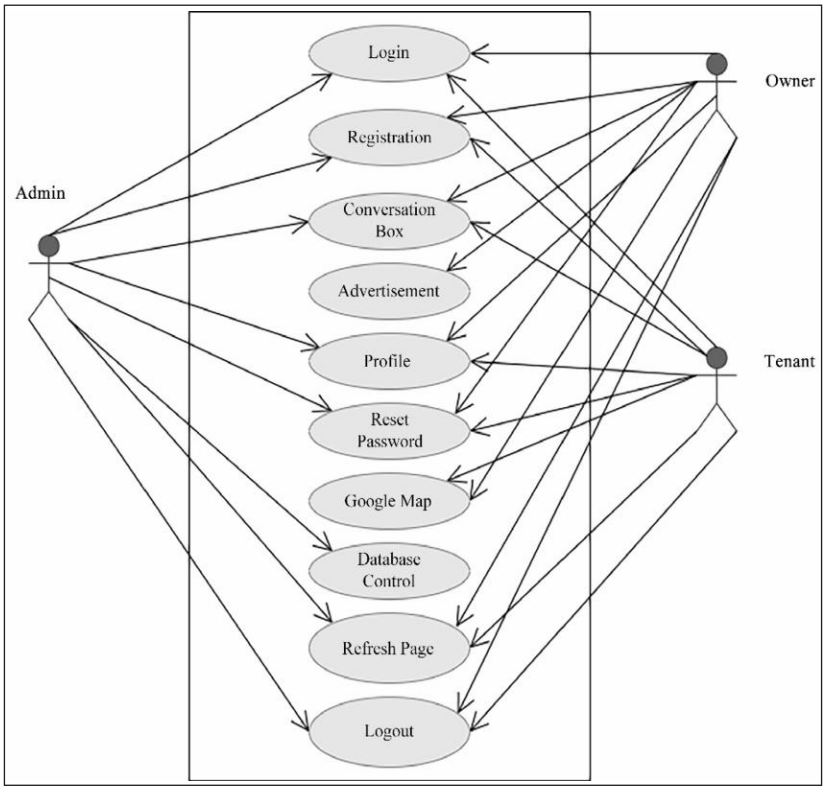


Fig. 7: Total Use Case Diagram of the Rental House Management System

- All the functionalities are written in the oval shaped box and the access is shown by pointed arrows.
- This diagram describes the total functionality of the web-page and also shows who has the access to which function of the website (as per Fig. 7).

C. E.R. Diagram

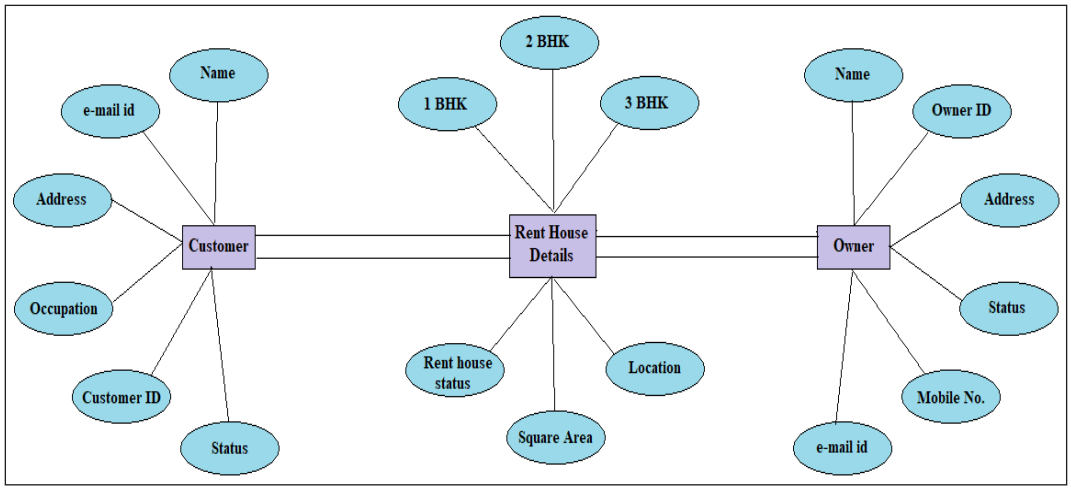


Fig. 8: An E.R. Diagram Showing the Working of Rental House Management System

The above shown diagram is the E.R. (Entity Relationship) diagram of the Rental Housing Management System. It shows us the information that has to be given by the user end and also by the owner end for the proper working of the system. The user as well as owner has to give their basic details, the user must give the proper details of the house he/she wants and the owner should do the same for the property they want to rent out (as per Fig. 8).

D. Data Flow Diagram

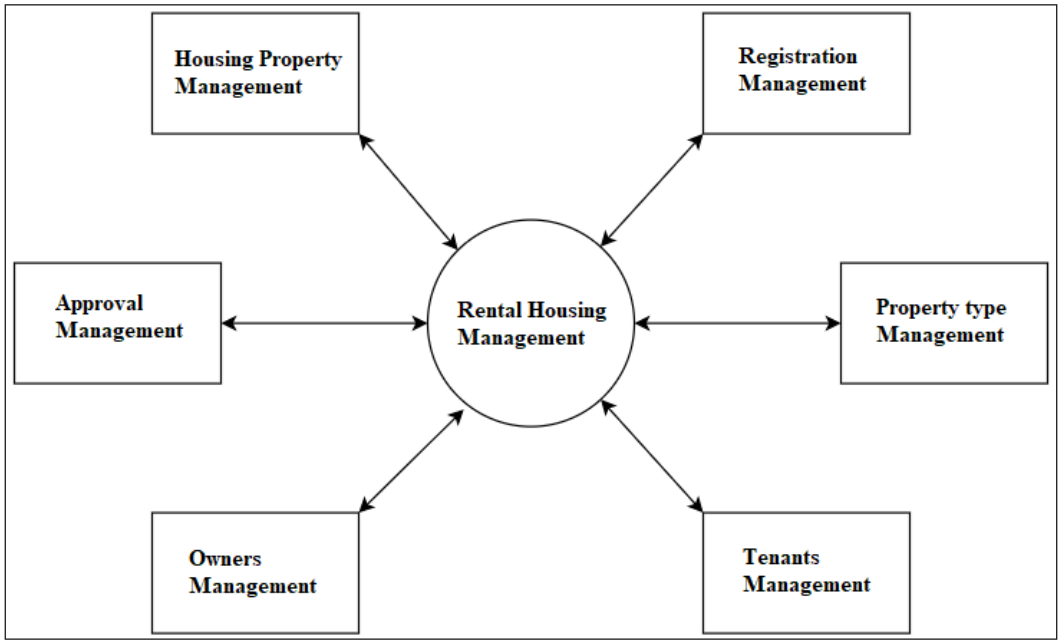


Fig. 9: A DFD Diagram Showing the Basic Working of Rental House Management System

The above shown diagram is a data flow diagram of Rental Housing Management System. It analyzes how the data flows in the rental housing management, the contents written in the rectangular boxes shows the input that is given to the software and the arrow depicts how the data flows (as per Fig. 9).

The above shown diagram is the detailed flowchart of the Rental House Management System. It shows how the user will enter into our web application where he will sign up and proceed as shown in the below diagram (as per Fig. 10).

E. Flowchart

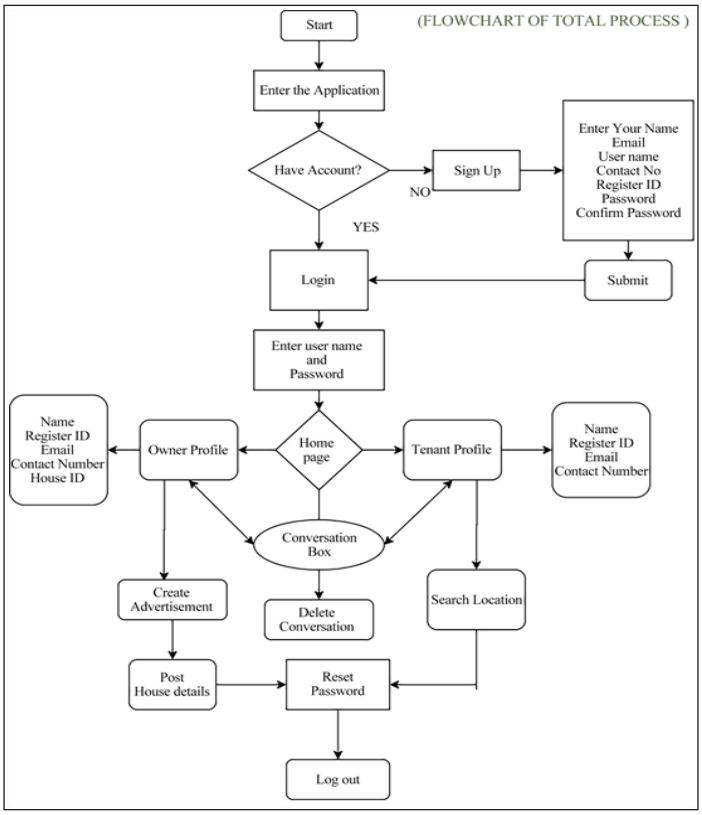


Fig. 10: Flowchart Showing the Workflow of Rental House Management System

IV. Result and Discussion

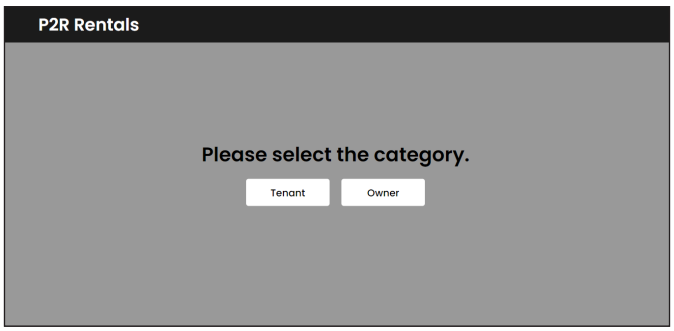


Fig. 11: Selection between Owner and Tenant

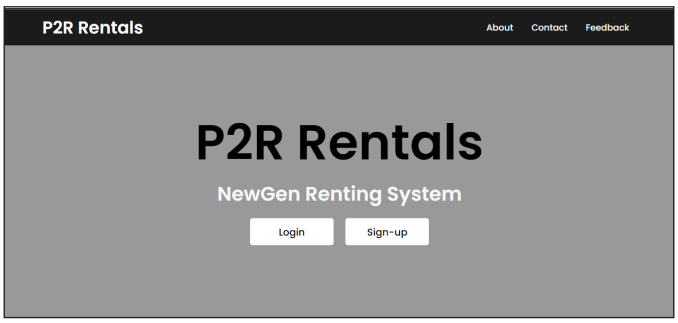


Fig. 12: Login and Sign-up Page

Here a user can choose whether he/she comes for our website as a tenant or owner. It makes easy to categories between the tenants and the owners. We make it easy so anyone can use it (as per Fig. 11).

In this page an individual person can access or create a user id by this following option Login and Sign-up which is directly linked to a page where he/she can make a user id, login the website (as per Fig. 12).

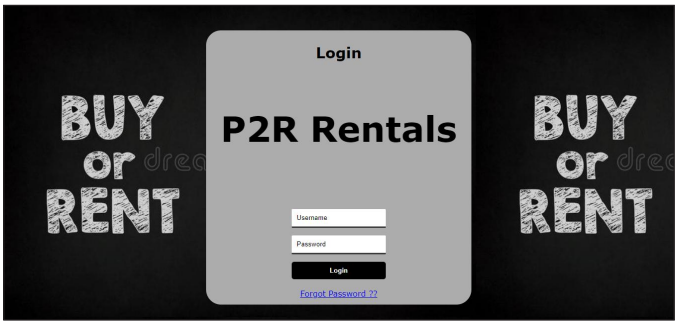


Fig. 13: Login Page of the Tenants or the Owner

It is used after a person has created his user id so that he or she can access their data which is saved in our website. And it is highly secured which is also a user friendly and easy setup (as per Fig. 13).

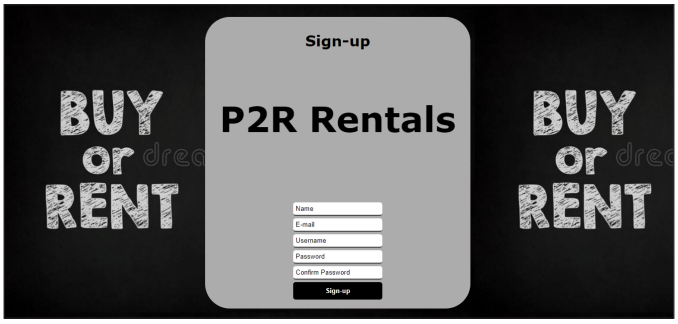


Fig. 14: Sign-up Page for the Tenant or the Owner

This is the page where a person can create a user id for himself that is unique and secured his data. Sign-up is used to make a login id for a user that he/she can access from any browser just by searching our website and filling the log-in page (as per Fig. 14).

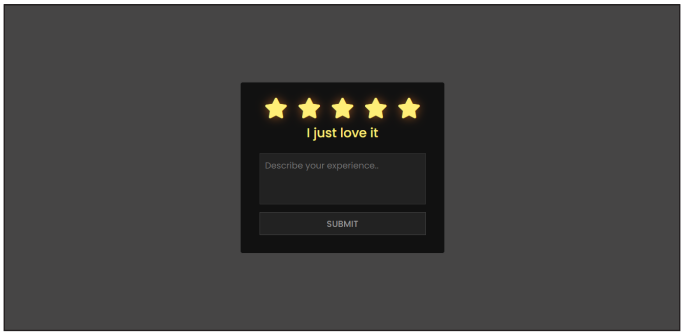


Fig. 15: User Rating Page for the Users

It is the review page where a user can rate our website and how it feels after using our Rental management website and what we have to improve in it more. It is a page where a user can clarify its satisfaction and problems that occur in run time (as per Fig. 15).

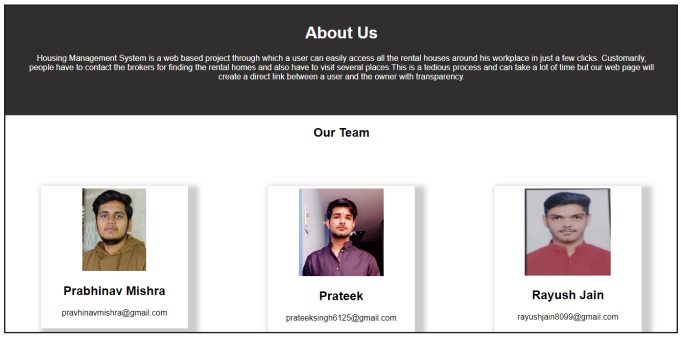


Fig. 16: An About Us Page for Knowing Us and Our Website

In this page we have given the information about ourself and mail so if any problem occurs in run time then a user can contact us for the main motive, what is the use of our website and what it used for (as per Fig. 16).

It is a payment gateway page where a user has to pay the price of the buying home/flat according to their preference. That's why we added a number of category modes in payment which make users comfortable to pay or buy a flat/home (as per Fig. 17).

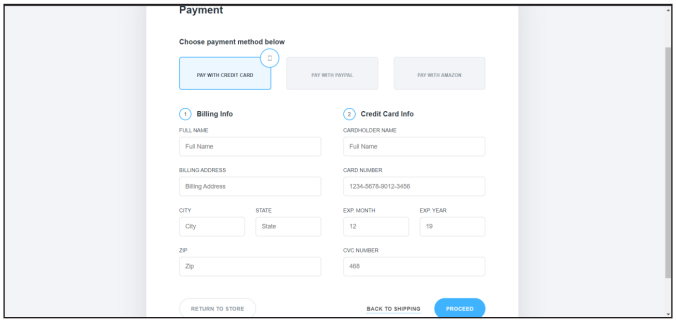


Fig. 17: Payment Gateway Page for the Convenience of Users

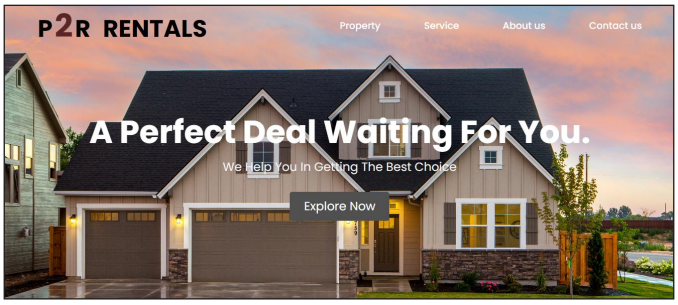


Fig. 18: User Interface for the Tenant to Explore

This is the front page view of our website which shows how the user will be interacting with us.

This is simple and an explorer option where a user clicks and it shows the number of details and flats on a different page. And in our front page view we also added property, service. About us, contact us because it is a necessary requirement of any website (as per Fig. 18).

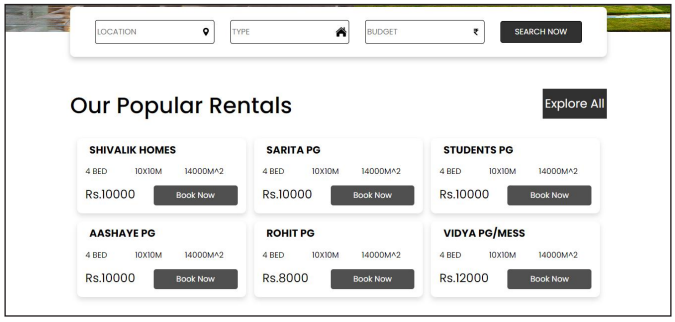


Fig. 19: List of Various Popular Rentals on Our Website

This is the page on our website where users can see the price and category the location and type of their flat by the following filter option in it which makes it easier for users to choose a good flat/home. And there is also a bargaining function in it where a buyer can bargain the price of a flat/home by the owner (as per Fig. 19).

This is the page of our website which shows how the owner will fill the details of their home/flats. In this the owner has to

fill the following categories where he/she can also add photo of their flats. It is easy for a renter to know if the flat is in good condition or not (as per Fig. 20).

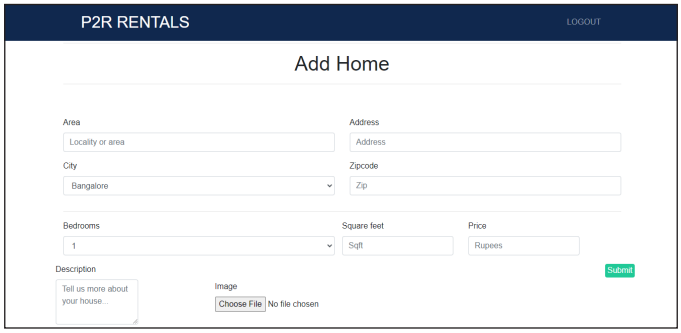


Fig. 20: Add a Homepage for the Owner to Upload Info about Their Property

Technical Conclusions after Results

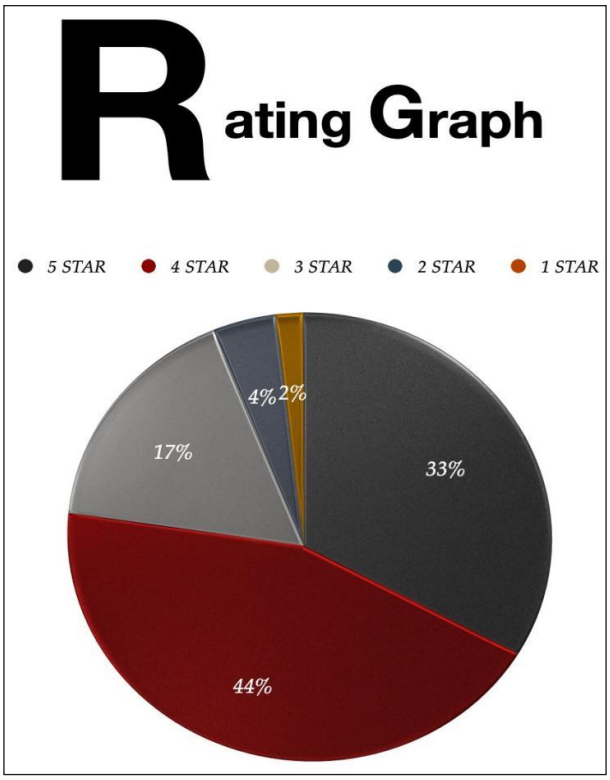


Fig. 21: User Rating Pie-Chart

This pie-chart shows the feedback of people and their experience while using the website.

During our analysis we found out that the users were satisfied with our website as more than 75% of the rating was above 4 stars.

Total Responses: 182

5 Star: 60 4 Star: 74 3 Star: 36 2 Star: 9 1 Star: 3

Average Rating: 4.0 (as per Fig. 21)

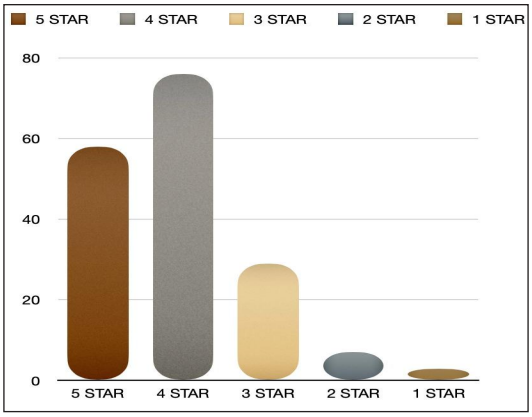


Fig. 22: User Rating Bar Graph

The above graph represents the user rating in the bar graph (as per Fig. 22).

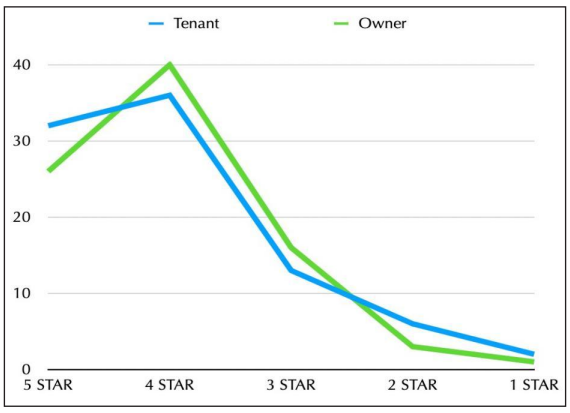


Fig. 23: User Rating Line Graph

This graph shows the analysis of the feedback of the tenants and the owner separately.

Our webpage has two different portals of tenant and owner and we have analyzed the feedback of both separately and found out that the average ratings of both are almost the same.

Average rating of tenants: 3.95

Average rating of owner: 4.05 (as per Fig. 23)

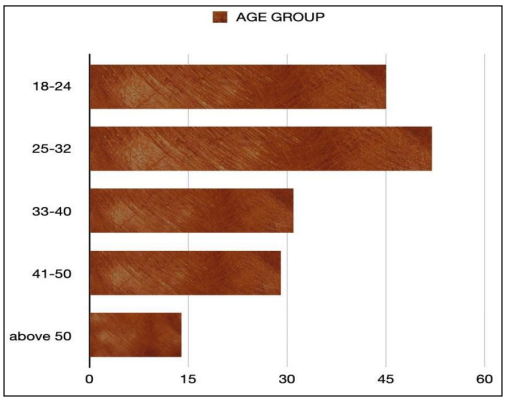


Fig. 24: Bar Graph Showing Age Group using Rental Website

This graph represents the different sections of age group using our website. On analysis we found out that the most number of users using our website lies between 18-32 which contains mostly bachelors and tenants. We found that the age group of 33-50 were families and the age group above 50 were joint families (as per Fig. 24).

V. Novelties

- Our Rental House Management System will have a better user experience than many other current working websites.
- We will provide proper details of the owner as well as of the tenants for the proper assurance of background of the person.
- We will add a special bargain feature in our software so that if the customer wants to bargain on the price he can do it with assurances of the owner.
- Our rental house management software will also contain small scale properties such as PGs for the benefit of students coming to an unknown place.

VI. Recommendation

Our webpage will provide a user-friendly interface and both the tenants and the owner can be benefited by the website. Users are provided with easy registration and login with OTP and password for security purposes. Separate portals are provided for both the buyers and the sellers. The search bar is provided with the filter search technique which recommends the user with the houses as per his/her requirements and budget. Our webpage will provide the recent images of the house and contact details of the owner for transparency.

VII. Future Research Direction and Limitations

A. Limitations

- A user can only access this web application not above than 2 or more portal access of the same user at a time.
- Our web application doesn't have dynamic visuals on it for now.
- It doesn't show the landlord's property which doesn't have a register on our web page.
- Our web application system database has a web storage which allow user to store or publicly upload their photos and video of property but the limit of the size of photos are 10MB and for video – 50MB.

B. Future Directions

- To provide service to a user like personal transportation support which are directly connected to Ola Cabs, Uber where a user can book a cab for its destination with 2-4% discount.
- To make our online web application also for Android and IOS app systems.
- To add dynamic features and also create an AI ChatBot which helps you to solve your biggest problem regarding transaction, using this online web application.
- To make our application system on voice command also with different languages.

VIII. Conclusion

This Rental Housing Management System is a web application that is made by using HTML, CSS, JS, MySQL. All of this was developed in visual studio code by Microsoft. Our rental house software is a very convenient and user friendly software and it provides its users with various new features for enhancing their experience. The user will enter into our website and first he has to log-in or sign-up with their credentials and then they will be directed to our webpage where they will have to select the region in which they want to rent a house and then after that they will be shown all the available properties in that specific area, from there they can select the house that meet their demands. This rental software will enhance the renter owner relationship and will aim at making customers' life easier.

Acknowledgement

The Author Team pays their respects to God. They then thank each of their parents for everything they have done for them. They then express their sincere appreciation to CPBC-Delhi, Patanjali foundation, Shantikunj Haridwar for providing help and assistance for lab equipment and research. They further offer their sincere gratitude to seniors of Dayalbagh Educational Institute Agra, GS Ayurveda Medical College Hapur, and ABES Engineering College, Ghaziabad. We also offer our special thanks to ABES management, Director ABESEC, HOD ECE, HOD CSE and all other senior coordinators for their extraordinary support to our team. In the end, we pay our due respects and thanks to each and everyone involved, direct or indirect in this work.

References

[1] S. Afzal, T. Rouf, S. Qadir, and S. Shah, "Rental house management system," *Journal of Emerging Technologies and Innovative Research,* vol. 8, no. 11, pp, 1-3, 2021. [Online]. Available: https://publisher.uthm.edu.my/periodicals/index.php/aitcs/article/view/2488

[2] H. P. Gommans, N. M. George, and A. N. Ownage, "Rental housing management system," *International Journal of Scientific and Research Publications,* vol. 4, no. 11, pp. 1-24, 2014. [Online]. Available: https://www.ijsrp.org/research-paper-1114/ijsrp-p35101.pdf

[3] A. J. Ikuomola, "Cloud based rental housing management system," *Research Gate,* vol. 1, no. 1, pp. 1-17, 2020. [Online]. Available: https://www.researchgate.net/publication/340926278_A_Secured_Mobile_Cloud-Based_House_Rental_Management_System

[4] J. A. Kirmani, A. Yousuf, and S. M. Bhat, "Rental housing management system," *International Journal of Computer Science and Mobile Computing,* vol. 6, no. 7*,* pp. 1-4, 2017. [Online]. Available: https://www.ijcsmc.com/docs/papers/July2017/V6I7201705.pdf

[5] M. R. Miyasam, and N. Selamat, "Rental housing management system," *Applied Information Technology and Computer Science,* vol. 2, no. 2, pp. 1745-1754, 2021. [Online]. Available: https://publisher.uthm.edu.my/periodicals/index.php/aitcs/article/view/2488

[6] K. Rathore, A. Syed, and R. Patel, "Rental house management system," *International Research Journal of Modernization in Engineering Technology and Science,* vol. 3, no. 5, pp. 1-6, 2021. [Online]. Available: https://www.irjmets.com/uploadedfiles/paper/volume3/issue_5_may_2021/10054/1628083403.pdf

[7] D. Voumick, P. Deb, S. Sutradhar, and M. M. Khan, "Development of online based smart house renting web application," *Journal of Software Engineering and Applications,* vol. 14*,* no. 7, pp. 312-328, 2021. [Online]. Available: https://www.scirp.org/journal/paperinformation.aspx?paperid=110723#ref1

[8] A. S. Khanfar, F. A. Lone, and Md. Moizuddin, "A comprehensive survey on support vector machines for intrusion detection system," *International Journal of Knowledge Based Computer Systems,* vol. 10, no. 1, pp. 33-39, 2022.

Additional Readings

1. M. Oxley, and J. Smith, "Housing policy and rented housing in Europe," 1996. [Online]. Available: https://doi.org/10.4324/9780203860632
2. Dongguan, and Y. R. Gong, "Rental housing management as surveillance of Chinese rural migrants: The case of hillside compounds," 2016. [Online]. Available: https://www.tandfonline.com/doi/full/10.1080/02673037.2016.1171828
3. "Assessing the integrated sustainability of a public rental housing project from the perspective of complex ecosystem," 2016. [Online]. Available: https://doi.org/10.1016/j.habitatint.2016.01.001
4. R. J. Struyk, K. Mark, and J. P. Telgarsky, "Private management for eastern Europe's state rental housing," 1991. [Online]. Available: https://doi.org/10.1016/S1051-1377(05)80027-0
5. T. Ahmed, "Property management system," 2018. [Online]. Available: https://notbot.se/property-management-system-4-pdf-free.html

Annexure

Key Terms and Definitions

- *HTML* - HTML is a Hyper Text Markup Language which is used to structure your web page and its content. It helps to create your own website.
- *CSS* - Cascading Style Sheet is used to style and web pages. For e.g. - colors, font, or adding animation and other decorative features.
- *JavaScript* - JavaScript is often abbreviated as JS, it is a programming language that is one of the main core technologies of the World Wide Web, alongside HTML, CSS. It enables dynamic updating content, animated images and pretty much everything.
- *UI* - It stands for user interface, is anything a user interacts with on a website or any digital devices. It encompassed both the appearance and interactivity of an app or website.
- *Web Application* - It is a software that runs in a web browser. It is an application program that is stored on remote servers and delivered over the Internet through a web interface.

International Journal of Knowledge Based Computer Systems
11 (1) 2023, 21-25
http://www.publishingindia.com/ijkbcs/

A Decentralized Non-Fungible Token Marketplace: A User-Friendly Approach to NFT Adoption

Aayushi Balothiya[1], Amol Paliwal[2*], Prathviraj Singh Dodiya[3], Taresh Ayaspure[4] and Nisha Rathi[5]
[1]Computer Science and Information Technology, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India. Email: aayushibalothiyact19@acropolis.in
[2]Computer Science and Information Technology, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India. Email: amol935paliwal@gmail.com
[3]Computer Science and Information Technology, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India. Email: prathvirajsinghct19@acropolis.in
[4] Professor, Computer Science and Information Technology, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India. Email: tareshayaspure@acropolis.in
[5] Professor, Computer Science and Information Technology, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India. Email: nisharathi@acropolis.in
*Corresponding Author

Abstract: This paper introduces a decentralized non-fungible token (NFT) marketplace website developed as a major project for college. The platform aims to demonstrate the intersection of NFTs and blockchain technology and how it is changing the digital world also to increase the adoption of NFTs by providing a user-friendly and intuitive experience for buying, selling, and trading digital assets. NFTs, being unique and verifiable digital assets, have the potential to revolutionize the way we own, trade, and manage digital items. The use of blockchain technology in NFTs ensures their security, ownership, and authenticity, making them suitable for use in various industries such as gaming, art, and collectibles. The marketplace website built for this project showcases the importance of NFTs and how they can be made accessible to a wider audience through a user-friendly platform. The platform provides a seamless experience for buying, selling, and trading NFTs, making it easy for users to participate in the NFT economy. Additionally, the platform features tools for creating custom NFTs, opening up new opportunities for artists and creators to monetize their digital creations. This project highlights the significance of NFTs in the digital world and how they are changing the way we view and manage digital assets. The decentralized and secure nature of NFTs, combined with their uniqueness and verifiability, makes them ideal for various use cases, and the creation of user-friendly platforms like the one presented in this project will play a crucial role in promoting the wider adoption of NFTs.

Keywords: Authenticity, Blockchain technology, Creators, Decentralized, Digital assets, NFTs, Security, Trading.

I. Introduction

In recent years, the intersection of non-fungible tokens (NFTs) and blockchain technology has created a new landscape in the digital world, providing a new way to own, trade, and manage digital assets. NFTs are unique, verifiable digital assets that offer new opportunities for various industries such as gaming, art, and collectibles. This paper introduces a decentralized NFT marketplace website, which has been developed as a major project for college, with the aim of demonstrating the capabilities of NFTs and blockchain technology and increasing their adoption. The platform provides a user-friendly and intuitive experience for buying, selling, and trading NFTs, making it easy for users to participate in the NFT economy. Furthermore, the platform offers tools for creating custom NFTs, opening up new opportunities for artists and creators to monetize their digital creations. This project highlights the significance of NFTs in the digital world and the potential for user-friendly platforms to play a crucial role in promoting their wider adoption.

II. Problem Foundation

A. Problem Statement

The blockchain and NFT technology has the potential to revolutionize the way we own, trade, and manage digital assets, but the current NFT marketplaces are limited by their lack of accessibility. The complexity of existing NFT marketplaces and the technical knowledge required to use them effectively limit

their reach and hinder the wider adoption of NFTs. The aim of this project is to address these issues by creating a decentralized NFT marketplace that is user-friendly and accessible, making it easier for people to participate in the NFT economy and unlock the full potential of NFT technology.

While the NFT market continues to grow, existing marketplaces often provide a complex and technical user experience, making it difficult for many individuals to participate in the NFT economy. This project aims to address these challenges by creating a streamlined and user-friendly NFT marketplace that offers essential features for buying, selling, and trading NFTs. This project will demonstrate the potential of accessible and intuitive NFT marketplaces to increase the adoption of NFTs and change the way we own, trade, and manage digital assets.

B. Scope

The scope of this project is to develop a decentralized NFT marketplace website that serves as a proof of concept for a user-friendly and accessible platform for buying, selling, and trading NFTs. The platform will provide essential features for users to participate in the NFT economy, including a seamless experience for buying, selling, and trading NFTs and tools for creating custom NFTs. The project will focus on demonstrating the intersection of NFTs and blockchain technology, and how the creation of user-friendly platforms can increase the adoption of NFTs. The platform will be built using Next.js for the frontend and Solidity for the backend and will be evaluated based on its user experience and functionalities. The project will not provide all the features and functionalities of existing NFT marketplaces, but it will serve as a starting point for further research and development in this field.

III. Literature Review

The development of decentralized non-fungible tokens (NFTs) and the integration of blockchain technology has given rise to a new and innovative way of owning, trading, and managing digital assets. NFTs are unique and verifiable digital assets that offer the potential to revolutionize the digital world. This technology is being adopted in various industries such as gaming, art, and collectibles due to its secure, authentic, and ownership-ensuring features.

Previous research in the field of NFTs has focused on the technical aspects of NFTs, including the creation and management of NFTs, their security, and their suitability for various use cases. However, there is a lack of research on the development of user-friendly platforms for buying, selling, and trading NFTs. This research gap presents an opportunity to explore the potential of NFTs and the development of platforms that promote the wider adoption of NFTs.

A review of existing NFT marketplaces reveals that they provide a variety of features, including buying and selling of NFTs, trading of NFTs, and creation of custom NFTs. These platforms have made significant contributions to the development of NFTs and have helped to increase the adoption of NFTs. However, they have some limitations such as the lack of a user-friendly interface and the complexity of using the platform. This has resulted in a low level of user adoption and has limited the growth of the NFT market.

This literature review highlights the significance of NFTs in the digital world and the need for user-friendly platforms that can help to promote their wider adoption. It highlights the features of existing NFT marketplaces and the limitations that hinder the growth of the NFT market. This literature review provides a foundation for the development of a new NFT marketplace that addresses the limitations of existing NFT marketplaces and provides a user-friendly and intuitive experience for buying, selling, and trading NFTs. The goal of this research is to develop a platform that will help to increase the adoption of NFTs and promote the growth of the NFT market.

Blockchain technology has been a game-changer in creating a marketplace for Non-Fungible Tokens (NFTs). The essence of blockchain lies in its ability to provide a universal ledger that can store information across a network. Due to its decentralized nature, all network participants can participate in transactional processes, and the transparency of the system ensures that every individual in the network can view transactions. Additionally, blockchain technology is verifiable, which prevents counterfeit transactions (Tasatanattakool and Techapanupreeda, 2018 [1]). Blockchains come in various forms, with Bitcoin being the first one introduced by Satoshi Nakamoto, who realized the importance of a decentralized infrastructure in the current sociological order (Nakamoto, 2009 [2]). The concept of a decentralized ledger, where the distributed, transparent, and immutable consensus nature of the algorithm of blockchain was put in place, brought a new societal structure of digital money without a central authority, redistributing power amongst the masses (Boucher *et al.,* 2017 [3]).

The possibilities for blockchain technology are endless, and it can be used for peer-to-peer banking services, music royalties, and digital art. Especially in the NFT sphere, blockchain technology has become viral on the internet, with popular examples being Crypto Kitties, Crypto Punks, and Bored Ape Yacht Club, where massive amounts of money have already been invested. The technology is being hailed as one that could revolutionize society by using a consensus mechanism as a central component. Blockchain is viewed as a universal ledger or bookkeeping instrument where transactions are broadcasted onto the ledger and independently verified by peers in the network. Transactions are interlocked as chains and stacked on top of each other in a chronological sequence using a cryptographic hashing mechanism that prevents fabrication (Aste *et al.,* 2017 [4]). Initially, blockchain technology can be viewed as an ICT (Information and Communications Technology) innovation that can be used as an organizational technology to decentralize governance constructs and used for coordination of people and economic decision-making.

The use cases for blockchain technology are not limited to the financial instrument. In the NFT space, blockchain technology is used to tokenize works of art or intellectual property, paving the way for the decentralized way of creating derivative works for commercial purposes (Lee, 2021 [5]). The possibilities for utilizing blockchain technology are growing rapidly, with further inter-merging into the Metaverse space enabling a proliferation of a virtual economy where users can reap value through unique new markets. In summary, the blockchain technology has enabled the NFT marketplace, creating a decentralized system that is transparent, secure, and immutable, with endless possibilities for utilization in various industries.

IV. Project Methodology

Our NFT marketplace project aims to provide a user-friendly and transparent platform for buying, selling, and trading non-fungible tokens (NFTs) on the Ethereum network. The platform will enable users to create, list, and purchase NFTs, as well as allowing for the transfer of ownership of NFTs between users.

The project will be implemented using Solidity and will inherit from the ERC721 standard implemented by OpenZeppelin to ensure compatibility with existing NFTs on the Ethereum network. The logic for buying, selling, and transferring NFTs will also be written in Solidity.

To provide a more efficient and user-friendly experience, we will be using Ether.js to interact with our smart contracts, and MetaMask to provide a secure and easy-to-use interface for users to interact with our platform.

To ensure the transparency and immutability of our NFT marketplace, we will be using IPFS to store and distribute our NFT metadata.

Our project aims to create a more accessible and trustworthy NFT marketplace that enables users to easily buy, sell, and trade NFTs, while providing a transparent and decentralized platform that can be trusted by all users.

Project Stack

Web Application Framework - Next.js

Backend-Solidity - In this we are inheriting from the ERC721 standard implemented by OpenZepplin

Solidity Development Environment - Hardhat

File Storage - IPFS

Ethereum Web Client Library - Ethers.js

Blockchain Wallet - MetaMask

Next.js - In the implementation of our NFT marketplace website, we used Next.js as the frontend framework. Next.js is a React-based framework that provides a set of tools and features to build fast and scalable web applications.

One of the key reasons we chose to use Next.js is its ability to server-render pages, which greatly improves the loading speed of the website. This is particularly important for our NFT marketplace, as fast loading times are crucial for providing a seamless user experience. Additionally, Next.js provides automatic code splitting, which means that users only need to download the code required for the current page they are visiting, reducing the overall size of the website and making it faster to load.

Another reason we used Next.js is its support for static file generation. This allows us to generate and pre-render pages during build time, which can be served to users directly from a CDN. This reduces the server load and makes the website even faster.

Finally, Next.js provides a powerful set of features for building and maintaining complex applications, making it an ideal choice for our NFT marketplace. It provides a set of APIs for fetching data, optimizing images, and handling routing, among other things, all of which helped us to implement our NFT marketplace in a more efficient and streamlined manner.

In conclusion, we used Next.js for the implementation of our NFT marketplace website because of its server-rendering capabilities, static file generation support, and powerful set of features for building complex applications. These features allowed us to provide a fast, seamless, and user-friendly experience for buying, selling, and trading NFTs.

Solidity - In our project implementation, we are using Solidity, a programming language designed specifically for writing smart contracts on the Ethereum blockchain. Solidity allows us to write code that is executed on the blockchain, enabling us to create decentralized applications and execute transactions without the need for intermediaries. Additionally, Solidity is the most popular and widely used programming language for writing smart contracts on Ethereum, making it a natural choice for our project.

We are also inheriting from the ERC721 standard implemented by OpenZeppelin, a widely used and well-established standard for creating non-fungible tokens (NFTs) on the Ethereum network. By inheriting from the ERC721 standard, we are ensuring that our NFTs will be interoperable with other NFTs on the Ethereum network and that they will be compatible with existing NFT marketplaces and tools.

Furthermore, the logic for buying, selling, and transferring NFTs is written in Solidity. This allows for the automation of these processes and ensures that they are executed in a trustless and transparent manner without the need for intermediaries. The use of Solidity also enables us to create smart contracts that are self-executing and cannot be altered once they have been deployed to the blockchain, providing increased security and transparency for our users. Overall, the use of Solidity and the ERC721 standard implemented by OpenZeppelin provides a secure and efficient foundation for our NFT marketplace project.

We are using Hardhat as our development environment. Hardhat is a popular and reliable development environment

that is designed specifically for Ethereum smart contract development. It provides us with a range of tools and features that make it easy to test, debug, and deploy our smart contracts on the Ethereum network.

One of the main reasons we are using Hardhat is its flexibility and ease of use. Hardhat is highly modular and allows us to customize our development environment to fit our specific needs. It also provides a range of plugins that enable us to easily integrate with other tools and services that we may need for our project.

Another key feature of Hardhat is its testing framework. Hardhat allows us to write automated tests for our smart contracts, ensuring that they behave as expected and are free from bugs and vulnerabilities. This is crucial for creating secure and reliable smart contracts that can be trusted by our users.

Additionally, Hardhat provides us with a built-in local blockchain network that allows us to quickly test and deploy our smart contracts in a simulated environment. This enables us to iterate and make changes to our code without incurring the cost and time associated with deploying to the live Ethereum network.

Overall, the use of Hardhat provides us with a reliable and flexible development environment that enables us to build secure and efficient smart contracts for our NFT marketplace project.

we are using the InterPlanetary File System (IPFS) to store and distribute our NFT metadata. IPFS is a decentralized peer-to-peer file sharing protocol that enables us to store and distribute files in a secure, efficient, and decentralized manner.

One of the main advantages of using IPFS is its decentralized nature. Unlike traditional file storage systems, which rely on centralized servers, IPFS distributes files across a global network of nodes. This makes it more resilient to censorship and single points of failure, as well as providing faster and more efficient access to data.

Another key advantage of using IPFS is its content-addressable system. This means that files are identified by their content, rather than their location. As a result, once a file has been uploaded to IPFS, it can be accessed from anywhere in the world simply by using its content address. This makes it easier for us to distribute and share our NFT metadata, without having to worry about the location of the files.

Additionally, using IPFS for our NFT metadata enables us to create a more transparent and decentralized NFT marketplace. By storing the metadata on IPFS, we can ensure that the data associated with our NFTs is immutable, transparent, and publicly accessible. This helps to build trust and confidence in our platform, as users can be assured that the data associated with their NFTs is secure and cannot be altered or tampered with.

Ether.js provides us with a range of powerful and easy-to-use tools for interacting with smart contracts and the Ethereum network, enabling us to create a more efficient and user-friendly NFT marketplace.

MetaMask is also used. It is a browser extension that allows users to securely manage their Ethereum accounts and interact with decentralized applications, including our NFT marketplace. By integrating with MetaMask, we are able to provide our users with a seamless and secure way to buy, sell, and transfer NFTs on our platform.

V. Result Discussions

The results of our NFT marketplace project demonstrate the successful implementation of a user-friendly and transparent platform for buying, selling, and trading non-fungible tokens on the Ethereum network. The platform enables users to create, list, and purchase NFTs, as well as allowing for the transfer of ownership of NFTs between users.

By using Solidity and inheriting from the ERC721 standard implemented by OpenZeppelin, we ensured the compatibility of our NFTs with existing NFTs on the Ethereum network. The logic for buying, selling, and transferring NFTs was also written in Solidity, enabling secure and efficient execution of these transactions.

We used Ether.js to interact with our smart contracts, and MetaMask to provide a secure and user-friendly interface for users to interact with our platform. The use of IPFS to store and distribute NFT metadata ensured the transparency and immutability of our NFT marketplace.

Overall, the results of our NFT marketplace project demonstrate the successful implementation of a platform that provides a more accessible and trustworthy way to buy, sell, and trade NFTs, while also promoting transparency and decentralization. Future work may include enhancements to the user interface, further optimization of the platform's performance, and the addition of new features to support a wider range of NFTs and use cases.

VI. Conclusion

In conclusion, our NFT marketplace project successfully implemented a transparent and user-friendly platform for buying, selling, and trading non-fungible tokens on the Ethereum network. By using Solidity, Ether.js, MetaMask, and IPFS, we created a decentralized and secure platform that enables users to easily create, list, purchase, and transfer NFTs. The project showcases the potential of NFTs and the Ethereum network in enabling new forms of digital ownership and value exchange. The success of this project highlights the potential for further innovation and development in the NFT space, as well as the broader potential of blockchain technology for transforming various industries.

Acknowledgment

We would like to express our heartfelt appreciation to our project guide, Prof. Taresh Ayaspure, and our project coordinator, Prof. Nisha Rathi, for their constant guidance, support, and encouragement throughout the project. Their valuable insights, feedback, and expertise have been instrumental in shaping our project and enhancing its quality. We are grateful for their unwavering support and commitment towards our project. We would also like to thank the Department of Computer Science and Information Technology at Acropolis Institute of Technology and Research for providing us with the necessary resources and infrastructure to conduct this research.

References

[1] P. Tasatanattakool, and C. Techapanupreeda, "Blockchain technology for decentralized secure IoT ecosystems," *2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE),* 2018, pp. 34-39, doi: 10.1109/jcsse.2018.00013.

[2] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2009. [Online]. Available: https://bitcoin.org/bitcoin.pdf

[3] P. Boucher, H. Casanova, and D. Daianu, "Cryptocurrencies and blockchain: Legal context and implications for financial crime, money laundering and tax evasion," *Bruegel Blueprint Series*, vol. 27, pp. 1-36, 2017. [Online]. Available: https://www.bruegel.org/wp-content/uploads/2017/05/BC-BP-2017_27.pdf

[4] T. Aste, P. Tasca, and T. Di Matteo, "Blockchain technologies: The foreseeable impact on society and industry," *Computer*, vol. 50, no. 9, pp. 18-28, 2017, doi: 10.1109/MC.2017.3571043.

[5] D. Lee, "The value of NFTs: Exploring the hype and potential of non-fungible tokens," The Brookings Institution, 2021. [Online]. Available: https://www.brookings.edu/techstream/the-value-of-nfts-exploring-the-hype-and-potential-of-non-fungible-tokens/

International Journal of Knowledge Based Computer Systems
11 (1) 2023, 26-29
http://www.publishingindia.com/ijkbcs/

Maximizing Efficiency and Minimizing Wait Times: A Case Study on the Use of Technology to Streamline Patient Appointments

Abhishek Singh Chouhan[1], Akshat Jain[2*], Sparsh Jain[3] and Praveen Gupta[4]
[1]Student, Department of CSIT, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India.
Email: abhisheksinghco19@acropolis.in
[2]Student, Department of CSIT, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India.
Email: jainakshat.co@gmail.com
[3]Student, Department of CSIT, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India.
Email: sparshjainco19@acropolis.in
[4]Professor, Department of CSIT, Acropolis Institute of Technology and Research, Indore, Madhya Pradesh, India.
Email: praveengupta@acropolis.in
*Corresponding Author

Abstract: In former years, the internet was used for various societal benefits. This article discusses how to use an Android application and the web to help patients. The major goal is to give the patient a hassle-free atmosphere and comfort when visiting the doctor. When patients have to wait a long time for appointments, even if they receive one, the doctor may not visit the hospital, and the patient may leave without visiting, resulting in the appointment being canceled. The issue will be resolved and patients will be kept up-to-date via technology such as Android and web-based applications. Patients may arrange appointments from anywhere at any time, saving their valuable time.

Keywords: Appointment, Hospital, Patient.

I. Introduction

A. Background and Context of the Problem

In the current scenario of digitalization, where almost every aspect of life has gone digital, the medical field still has some systems that are managed manually through registers and files. This result in inefficiencies, delays, and inconvenience for both patients and doctors.

The current digital era has brought major advancements to many aspects of life, however, the medical field still faces challenges in modernizing its systems. The reliance on manual processes such as paper-based records and files leads to inefficiencies, delays, and a less convenient experience for both patients and healthcare providers. This highlights the need for a more digitized approach in the medical field to improve efficiency, reduce errors, and provide a better overall experience. The medical field still relies on manual systems, causing inefficiencies, delays, and inconvenience for patients and doctors in the digital age [1].

Implementing a digitized system in the medical field would bring numerous benefits. The availability of electronic medical records would allow for easier access to patient information, reducing the time taken for diagnosis and treatment. This would also improve the accuracy of medical records and minimize the risk of errors due to manual documentation. Moreover, the digitization of medical systems would also improve communication between healthcare providers, enabling quicker and more effective collaboration for better patient outcomes. Additionally, it would allow for more efficient appointment scheduling and reduce wait times for patients. In conclusion, the digitization of the medical field would bring numerous benefits and help improve the overall experience for patients and healthcare providers alike.

B. Purpose and Objectives of the Project

The purpose of this project is to develop a digital solution called Nirogyam to address these challenges in the medical field. The specific objectives of the project are to:

- Create a digitalized and efficient application for appointment booking for patients.
- Ensure that patients can purchase medicines online easily.
- Automate the delivery of medicines and minimize difficulties faced by people when they are unable to visit a doctor's clinic.

- Integrate the application's chatbot, video call, and other features.

System Architecture:-

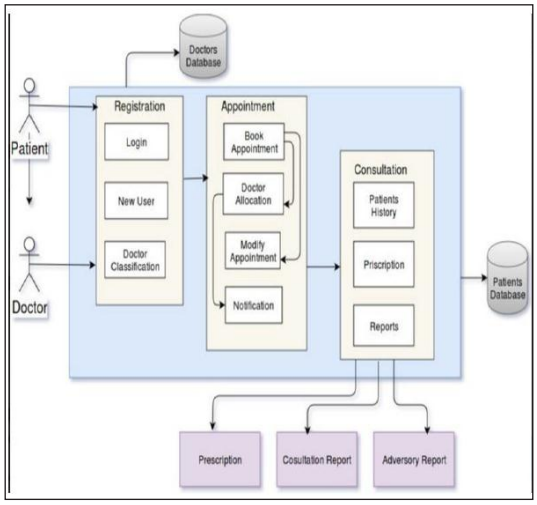


Fig. 1: System Architecture [2]

C. Overview of the Proposed Solution (Nirogyam)

Nirogyam is a digital platform for the medical field that aims to provide a one-stop solution for all medical needs. Patients can book appointments with doctors for a particular problem, purchase medicine, and book diagnostic tests, all from the comfort of their home. Doctors also benefit from this system as it provides a digital workspace for them.

Nirogyam is a comprehensive digital solution for the medical field that seeks to revolutionize the way healthcare is delivered and managed. The platform provides a seamless and convenient experience for patients, allowing them to easily book appointments, purchase medicine, and book diagnostic tests from their home. This eliminates the need for them to visit multiple clinics and pharmacies, saving time and effort.

For healthcare providers, Nirogyam offers a digital workspace where they can manage their appointments, patient records, and diagnostic reports efficiently. It also enables easy communication between healthcare providers, improving collaboration and enabling better patient outcomes.

Overall, Nirogyam is designed to make the healthcare experience more convenient, efficient, and effective for both patients and healthcare providers. By providing a one-stop solution for all medical needs, it has the potential to revolutionize the way healthcare is delivered and received in the digital age.

II. Literature Review

A. Overview of Current Systems in the Medical Field

The current systems in the medical field are largely manual, with appointment booking, prescriptions, billing, and medicines managed through registers and files. Some digital solutions exist, but they are often limited in scope and functionality.

The current systems in the medical field are primarily based on manual processes, which can result in inefficiencies, errors, and inconvenience for patients and healthcare providers. The process of booking appointments, obtaining prescriptions, managing billing, and purchasing medicine still relies on traditional methods such as paper records and files. While some digital solutions exist, they are often limited in their scope and functionality, providing only basic features and lacking comprehensive solutions for the medical field.

This manual approach to healthcare management leads to a number of problems, such as longer wait times for patients, difficulties in accessing medical records and diagnostic reports, and the potential for errors due to manual documentation. It also makes it difficult for healthcare providers to collaborate and communicate effectively, resulting in a less effective and less efficient healthcare system [3].

In conclusion, the current systems in the medical field are in need of modernization and digitization to improve the efficiency and effectiveness of healthcare delivery and management. The implementation of comprehensive digital solutions, such as Nirogyam, has the potential to revolutionize the way healthcare is provided and received in the modern age.

B. Comparison of Manual and Digital Systems in Terms of Efficiency and Effectiveness

Manual systems are often prone to errors and delays, resulting in inefficiencies and inconvenience for both patients and doctors. On the other hand, digital systems provide a more efficient and effective way of managing medical data. They are quicker, less prone to errors, and allow for easy access and retrieval of information.

The comparison between manual and digital systems in the medical field highlights the stark differences in terms of efficiency and effectiveness. Manual systems, while still widely used, suffer from a number of problems such as errors due to manual documentation, longer wait times for patients, and difficulties in accessing medical records. These inefficiencies result in a less convenient and less effective healthcare experience for patients and healthcare providers alike.

On the other hand, digital systems offer a more efficient and effective way of managing medical data. The use of electronic medical records allows for quick and easy access to patient information, reducing the time taken for diagnosis and treatment. The digitization of medical systems also minimizes the risk of errors and improves the accuracy of medical records.

Furthermore, digital systems provide a platform for effective communication and collaboration between healthcare providers, enabling them to work together to provide the best possible outcomes for patients. They also provide patients with a more convenient experience, allowing them to book appointments,

purchase medicine, and book diagnostic tests from the comfort of their home.

In conclusion, the use of digital systems in the medical field has the potential to bring numerous benefits in terms of efficiency, effectiveness, and patient experience. The implementation of comprehensive digital solutions, such as Nirogyam, has the potential to revolutionize the way healthcare is delivered and managed.

C. Advantages and Disadvantages of Existing Digital Solutions

Existing digital solutions in the medical field have many advantages, such as improved efficiency and effectiveness. However, they often lack integration with other systems and are limited in scope.

Existing digital solutions in the medical field offer numerous advantages over manual systems, such as improved efficiency, accuracy, and effectiveness. The use of electronic medical records and digital communication platforms enables healthcare providers to access and manage patient information quickly and easily, reducing the time taken for diagnosis and treatment.

However, despite these benefits, existing digital solutions in the medical field also have their limitations. One major disadvantage is a lack of integration with other systems, resulting in fragmented and disconnected processes. For example, a patient's medical records may not be accessible by other healthcare providers, leading to duplication of tests and a lack of continuity of care.

Another disadvantage of existing digital solutions is that they are often limited in scope, providing only basic features and lacking comprehensive solutions for the medical field. This result in a fragmented and disjointed experience for patients, who may need to use multiple systems to manage their health.

In conclusion, while existing digital solutions in the medical field offer numerous advantages over manual systems, they also have their limitations in terms of integration and scope. The implementation of comprehensive digital solutions, such as Nirogyam, has the potential to address these limitations and revolutionize the way healthcare is delivered and managed.

III. Methodology

A. Requirements Gathering and Analysis

The first step in the development of Nirogyam was to gather requirements and analyze the needs of patients and doctors. This involved conducting surveys and interviews to understand their specific needs and pain points.

B. System Design and Architecture

Based on the requirements analysis, the system design and architecture were developed. This involved creating a detailed plan for the system's functionalities and features.

C. Implementation and Testing

The next step was to implement the system and test it thoroughly to ensure it met the requirements and objectives.

D. Evaluation and Assessment of the System

Finally, the system was evaluated and assessed to determine its effectiveness and to identify areas for improvement.

IV. Results and Analysis

A. Overview of the Developed System (Nirogyam)

Nirogyam is a digital platform that provides a one-stop solution for all medical needs. It allows patients to book appointments with doctors, purchase medicine, and book diagnostic tests, all from the comfort of their home. Doctors also benefit from this system as it provides a digital workspace for them.

B. Comparison of the System with Existing Solutions

Nirogyam is more comprehensive and integrated compared to existing solutions in the medical field. It provides a one-stop solution for all medical needs, whereas existing solutions are often limited in scope and functionality.

C. Evaluation of the System Based on Specific Objectives

The system has met all of its specific objectives, as follows:

- The digitalized and efficient application for appointment booking has been successfully developed and implemented.
- Patients can now easily purchase medicines online.
- The delivery of medicines has been automated, reducing the difficulties faced by people when they are unable to visit a doctor's clinic.
- The application's chatbot, video call, and other features have been integrated into the system.

D. Areas for Improvement

Although the system has met its objectives, there are still areas for improvement. For example, the system could be further developed to include more advanced features, such as predictive analysis and personalization. Additionally, the system could be expanded to include more diagnostic tests and services.

V. Conclusion

In conclusion, Nirogyam is a digital platform for the medical field that provides a one-stop solution for all medical needs. The system has met its specific objectives and has provided a more

efficient and effective way of managing medical data compared to manual systems. The system still has areas for improvement, but it is a step in the right direction towards a more digitized and efficient medical field.

The project has significant implications for the medical industry, as it provides patients with a convenient and accessible way to manage their medical needs. The system's digitalized approach also offers benefits for doctors, as it provides them with a more efficient and streamlined way of managing their patients and data.

The system's success highlights the potential of technology to improve and revolutionize the medical industry. It is important for the medical field to continue to embrace technology and digitalization, as it can lead to more efficient and effective healthcare services.

In conclusion, Nirogyam is a valuable addition to the medical industry and provides a promising future for the digitization of medical services.

Whether the Implementation and deployment of the project idea (YES/NO).

- Have Social benefits. [YES]
- Have Environmental benefits. [NO]
- Considers health, safety, legal and cultural issues. [YES]
- Considers sustainable development (economic development that is conducted without depletion of natural resources). [YES]
- Applies ethical principles while selecting the project (not to steal other's project ideas, code, boldface, and documents). [YES]
- Commits to professional ethics and responsibilities and norms of the engineering practice. [YES]
- Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools. [YES]
- Identify, formulate, review research literature, and analyze engineering problems reaching substantiated conclusions. [YES]

The technological know-how what is required for the proposed project idea:

Tools:-

- Coding - VS-Code
- Project Manager - Trello

Software Requirements:-

- Front-End Requirements:-
  - Dart
- Back-End Requirements:-
  - NodeJS
- Frameworks Requirements:-
  - Flutter
- Server Requirements:-
  - AWS
- Database Requirements:-
  - MongoDB

Hardware Requirements:-

- Computer/Laptop with < 4 GB RAM or more >

< 2 GHz or more > < 64 bit dual core processor >

Acknowledgement

We are exceedingly grateful for your assistance in completing this project report. We are also grateful to AITR's Department of Computer Science and Information Technology for their time, support, comments, recommendations, persuasion, necessary facilities, Internet connection, and significant literature. We are grateful to our guide, Prof. Praveen Gupta, for his invaluable advice and provision of all required facilities.

References

[1] S. Malik, N. Bibi, S. Khan, R. Sultana, and S. Abdul Rauf, "Mr. Doc: A doctor appointment application system," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 14, no. 12, Dec. 2016.

[2] S. Prasad, S. Khan, S. Qureshi, Z. Ansari, and J. Khan Sheikh, "Implementation of doctor's appointment application for Life Care Hospital(Sion)," *International Journal of Innovative Science and Research Technology*, vol. 3, no. 4, Apr. 2018.

[3] S. B. Choudhar, C. Kusurkar, R. Sonje, P. Mahajan, and J. Vaz, "Android application for doctor's appointment," *International Journal of Innovative Research in Computer and Communication Engineering*, vol. 2, no. 1, Jan. 2014.

[4] Apollo Hospital. [Online]. Available: https://www.apollohospitals.com

International Journal of Knowledge Based Computer Systems
11 (1) 2023, 30-35
http://www.publishingindia.com/ijkbcs/

Galois Theory and Quality of Service Driven Knowledge Extraction within an Enterprise

Ghislain Abessolo Alo'o[1*], Danielle Olle Olle[2], Valery Monthe[3] and Laetitia Mapikou Mouafo[4]
[1]University of Yaounde I, Cameroon. Email: ghislain.abessolo@facsciences-uy1.cm
[2]University of Ebolowa, Cameroon. Email: georgesdelortolle@gmail.com
[3]University of Yaounde I, Cameroon. Email: valery.monthe@facsciences-uy1.cm
[4]University of Yaounde I, Cameroon. Email: laetitiamouafo@gmail.com
*Corresponding Author

Abstract: In Enterprises, knowledge is used in the execution of workflow tasks by users. This knowledge is held by employees involved in the workflow. With the instability of employees, quality of service is not always guaranteed. However, due to competitiveness, enterprises are forced to improve their quality of service to satisfy their customers. The method generally used consists to build a knowledge base that identifies the knowledge to optimize employee's performance to ensure optimal quality of service. But the existence of a knowledge base does not always guarantee the achievement of quality of services requested by customers. This raises the problem of extraction of knowledge in the execution of business process to achieve a quality of service requested by a customer. In this paper, we present a method based on Galois Theory, to extract, from the knowledge base, relevant knowledge needed for the execution of business process to achieve a given quality of service.

Keywords: Business process and workflow modeling, Customer satisfaction, Galois theory, Knowledge base, Knowledge management, Process abstraction, Quality of service.

I. Introduction

The enterprise places great importance on the diffusion and use of information and knowledge as well as its creation. The determinants of success of enterprises, and of national economies as a whole, is ever more reliant upon their effectiveness in gathering and utilizing knowledge [4, 8, 19]. The preservation of knowledge is an important building block within the concept of knowledge management [12, 19]. This is why efforts are being made today in the development of knowledge base in enterprise [12, 10, 14]. We can describe knowledge base as a system of knowledge and capabilities that preserves and stores perceptions, actions and experiences over time and secures the possibility of recall for the future.

During the knowledge base development processing, there is a first knowledge extraction which is done in order to select useful knowledge to be included in the knowledge base [13]. The second type of extraction is involved by employees to have any knowledge in the knowledge base previously built [17]. Several algorithms have been defined for these two types of extractions [1, 15, 4, 20]. Furthermore, various models exist in the literature on business process and workflow modeling, which integrate the knowledge aspects in different levels of abstraction. However, these models and algorithms do not take into consideration the knowledge selection that may occur during the execution of business process, aiming for a given quality of service, to select relevant knowledge to be used to achieve this quality of service [1, 4, 7, 8].

In this paper, we present a method based on Galois Theory, to extract from the knowledge base, knowledge needed for the execution of business process to achieve quality of service previously requested by a customer. This issue is particularly relevant to the extent that enterprises have focused their efforts on meeting the expectations of their customers in recent years [9, 11, 12]. Expectations expressed behind a set of qualities factor for the services expected. This makes it important to use these knowledge bases in the sense of knowledge to be selected for a given quality of service.

The rest of paper is organized as follows: Section II provides a summary of the theory of business processes and workflows modeling; Section III presents a summary of Galois Theory; Section IV deals with the Galois Theory and QoS driven knowledge extraction approach; and Section V concludes the work and highlighting some perspectives as future works.

II. Business Process and Workflow Theory

In this section we present, in an incremental manner, concepts that are borrowed for the theory proposed by Atsa *et al.* [1, 2, 3, 4] for the modeling of business process and workflows within an organization in order to manage the satisfaction of different stakeholders. We then move from basic to complex concepts. Those concepts are suitable to tackle the knowledge extraction.

A. Knowledge Model

In [16], a knowledge or knowledge bit is defined as a tuple $<k,Ag,Ex,y,w,l,d,v>$ where: k is the name of the knowledge;

Ag is the name of the agent who expressed the knowledge; *Ex* is the experience level of *Ag*; *y* the context in which the goal is defined; *w* is the goal; *l* the business rule; *d* execution constraints; *v* the level of importance of the goal.

Let $Kb_1 = (k_1, Ag_1, \Psi_1, \omega_1, \lambda_1, \delta_1, \gamma_1)$ and $Kb_2 = (k_2, Ag_2, \Psi_2, \omega_2, \lambda_2, \delta_2, \gamma_2)$ two Knowledge. Kb_1 and Kb_2 are compatible denote by $Kb_1 \Delta Kb_2$, if and only if $\Psi_1 = \Psi_2$ and $\lambda_1 = \lambda_2$. When Kb_1 and Kb_2 are compatible, Kb_2 is better than Kb_1 denote $Kb_1 \subseteq Kb_2$, if and only if $\gamma_1 \text{--} \gamma_2$. $(\Omega, \subseteq)$ is used to denote the partial ordered set of compatible knowledge.

B. QoS Model

The quality of service denoted by QoS represents the performances of the service which determine the level of satisfaction projected for the recipients of the services [1, 4]. The level of satisfaction is defined as a set of properties, criteria, characteristics and performances of the services delivered to the customers.

Let *Cr* be a set of criteria considered in the evaluation of the quality of service, *Val* the set of values that can be assigned to these criteria, and *f* a map defined by $f: C \rightarrow Val$, the QoS is defined by (C, Val, f) [4].

Given two QoS *q1* and *q2* such that $q1=(Cr1, Val1, f1)$ and $q2=(Cr2, Val2, f2)$, *q*1 and *q2* are compatible and denote by $q1 \Delta q2$ if and only if $C1=C2$ and $Val1=Val2$. When *q1* and *q2* are compatible, *q1* is better than *q2* and denote $q1 \subseteq q2$ if and only if $\forall c \in C1, f1(c) \leq f2(c)$. $(\Phi, \subseteq)$ is used to denote the partial ordered set of compatible qualities of services [4].

C. Task Model

A task is an atomic activity that cannot be split into smaller activities [1, 2, 3, 4]. The performance or execution of a task transforms the state of the environment into another state. A task is therefore an action within a state of an environment. Before a task can be executed, the state of the environment should satisfy a specific condition called pre condition, and when this execution is completed another condition, called post condition is satisfied. For a task to be executed within an organization which will be defined later, the knowledge required for its performance is captured. This knowledge depends on the context within which the execution can take place. For each of the associated contexts, is defined a set of knowledge bits and quality of service to obtain after the execution of a task. In [4], a task is formally defined by a tuple $\langle nt, PP, f_m, g_m, Cx, KBx, Qx \rangle$ where *nt* denotes the name of the task, $PP = Pre \times Post$ where *Pre* denotes the non empty set of preconditions within which its execution can be carried out, and *Post* the set of post conditions that are obtained after the execution, *Cx* a non empty set of contexts within which the task can be executed, *KBx* a non empty set of knowledge bits used for the better understanding and performance of the task, *Qx* is a quality of service to be reach after the execution of *nt*. f_m, and g_m are maps defined respectively by:

$$\begin{cases} f_m : Cx \rightarrow PP \\ g_m : Cx \rightarrow KBx \end{cases}$$

If *c* denotes a context of *Cx*, then *c* is a restriction of the environment Θ, that is $c \subseteq \Theta$. The action of a task within an environment is to transform its current state into a new one. When $\langle nt, PP, fm, gm, Cx, KBx \rangle$ is a task, *s* a given state where the precondition $pre(PP)$ is satisfied i.e $s(pre(PP)) = true$, the action of *t* in the state *s* is the new state $t(s)$ which satisfies the post condition $post(PP)$ i.e $t(s)(post(PP)) = true\ e$. In general, the action of a task *t* within the state *s* is characterized by the observers of *s* whose value has been modified [2, 4].

Definition 2.1 (Task Action)

Let $E = \langle \Theta, S, val \rangle$ be an environment, *s* a given state and *t* a task whose pre condition is satisfied in *s*, then the action of *t* in *s* denoted by t_s and is specified by $t_s = \{o : \Theta, s(o) \neq t(s)(o)\}$ [4].

When there will be no ambiguity, a task will be represented by its name *t* and $pre(t)$ respectively $post(t)$ will denote respectively its pre and post condition. Based on the post condition of a task *t*, and the state *s* where $s(post(t)) = true$, we conjecture that $t_S = +post(t) \cup -post(t)$ [1].

Definition 2.2 (Chain)

A chain is an execution path of tasks, according to their actions in states and their triggering conditions is denoted by $P = \prod_{i=1}^{n} Sht_i$, and is specified as a finite sequence of shifts where *n* represents the length of the sequence [4].

Let *P* be a path of length $n>1$, and $sh_k = \langle s_k, st_k \rangle$, $sh_{k+1} = \langle s_{k+1}, st_{k+1} \rangle$ notes respectively the shift in the range k and $k+1$, the state s_{k+1} is the resulting state after the execution of the set of tasks st_k i.e $s_{k+1} = t_k(s_k)$. When there will be no ambiguity, the shift of the range k of the path *P* will be denoted by $P(k)$.

Let $Sht_k = \langle S_k, S \circ T_k \rangle$ and $Sht_{k+1} = \langle S_{k+1}, S \circ T_{k+1} \rangle$ be two shifts where $Sht_k = S \circ T_k(s_k)$, the difference between the states s_k and s_{k+1} is denoted by $\overline{s_k + s_{k+1}}$ and is defined as follows: $\overline{s_k + s_{k+1}} = S \circ T_k(s_k)$.

Definition 2.3 (Ordering of Tasks)

Let *T* be a set of tasks, and t_1 and t_2 be two tasks of *T*, we write $t1 \trianglelefteq t2$ if and only if for all chain *CH* such that if n_{t1} and n_{t2} denote respectively the maximum range of *t1* and *t2* in *CH*, then $n_{t1} \leq n_{t2}$. This relation has the following properties:

- Reflexivity: $t \trianglerighteq t$ this simply means that the task *t* belongs to the chain *CH*;
- Antisymetric: if $t1 \trianglerighteq t2$ and $t2 \trianglerighteq t1$ in the chain *D* then $t_1 = t_2$. By convention, there will always exist a path from each task to itself;
- Transitivity: obviously if in the chain *CH*, $t1 \trianglerighteq t2$ and $t2 \trianglerighteq t3$ then $t1 \trianglerighteq t3$.

Palette

Let *E* be an environment, and *S* be a set of different states that *E* may reach according to the actions of tasks *T*, then a palette *P* is a couple <*E*,*S*→*S*> [4].

Given a palette *P*, according to the environment changes within organizations and the different executions of tasks that can take place, different ways in which tasks can be executed have to be captured. $SP_{\mathbf{P}}$ is use to specify the set of execution paths that can be obtained from a palette *P*.

D. Business Process Model

A business process is a collection of activities or tasks designed to produce a specific output for customers [1, 2]. It implies a strong emphasis on how work is done within an organization in order to deliver a particular service. A process is thus a specific order of work activities across time and space, with a beginning, an end, and clearly defined inputs and outputs. The output is the reason the organization does this work and is defined in terms of the benefits this process has for the organization as a whole.

The model of a business process is defined as a couple <*P*,*G*> where *P* is a palette and *G* the service to be achieved [1, 3, 4].

E. Workflow Model

A workflow is defined by $(Ts, Es, Ps, h, f_{em}, Q)+$ where Ts is the set of none conflicting tasks, Es the set of employees dealing with the processing of Ts within the time intervals Ps to obtain the quality of service Q, *h* is the map $Ts \rightarrow Ps$ which defines for each task *t*, its time interval $h(t)$ within which it is processed, and *f* a map that gives for each task *t* the employee $f_{em}(t)$ who is charge of its processing. The two maps *h* and *f* are required to be two isomorphism as each task is required to be associated to a time interval within which its execution will take place, and should also be assigned to a specific employee for this performance [1, 2, 16]. The quality of service Q is such that:

$Q = \sum_{i=1}^{n} q_i$ where q_i is the quality of service obtain after the execution of task $t_i \in Ts$ and *n* the number of task in Ts.

Based on the fact that the satisfaction of customers is one of the challenges that enterprises are required to guarantee, in the modeling of the workflow, it is required that employees who are involved in the processing of tasks have the necessary knowledge to carry out these tasks. Therefore, if *t* is a task to be carried out by the employee $f_{em}(t)$, and $kb_{em}(t, f_{em}(t))$ his knowledge associated for the processing of *t,* there will exist at least a context *c* within which *t* can be processed such that the knowledge $bk(t,c)$ required for its processing verifies the following constraint $bk(t,c) \subseteq kb_{em}(t, f_{em}(t))$.

F. Task Processing

Based on the assignment of tasks done by the resource manager within an organization for the achievement a given customer goal, employees process these assigned tasks based on their own experience and the knowledge associated to these tasks [1, 2, 4]. According to the context within which the performance of the tasks is taking place, the processing can be done straightforward if the knowledge related to the task is adequate for its processing within this context. The processing sometime will not be done straightforward as the knowledge related to the performance of the task is not enough. When it is the case, the employee will use his tacit knowledge, on the one received from more experimented employees, in order to process the task. In order to keep track of this new way of carrying out this task, the defined information should be stored for further use. For this end, the knowledge of the so called task should be updated. In order to take this into consideration, the modeling of workflow must take into account the processing of tasks by employees. Let tk be a task that is processed by an employee using the knowledge kb in the context cx, the task tk change the state after its performance based on the fact that, the knowledge associated to this context is updated by the knowledge used for its processing i.e. $gm(cx,tk)=gm(cx,tk') \cup kb$ where $gm(cx,tk)$ denotes the set of knowledge required for the processing of the task tk [4].

III. Galois Theory

In this section, we define parts of Galois Theory that are suitable in handling our proposed approach. Most importantly, the properties of Galois connections are presented in terms of relation between adjoints, fixed points and closures, and order-isomorphic.

A. Galois Connections

In mathematics, especially in order theory, a Galois connection is a particular correspondence between two partially ordered sets (posets) [6, 13, 18].

Let E and F be two Hilbert spaces and T ∈ L(E, F). The unique linear application T* ∈ L(F, E) such that for any x ∈ E, y ∈ F we have: $\langle T(x), y \rangle = \langle x, T^*(y) \rangle$ is called the adjoin of T [21].

Let $\mathscr{P} = \langle P, \preceq \rangle$ and $\mathscr{Z} = \langle Q, \sqsubseteq \rangle$ be partial order set; and suppose $f_* : P \rightarrow Q$ Q and $f^* : Q \rightarrow P$ are a pair of functions such that for all $p \in P$ and all $q \in Q$, (G) $f_*(p) \sqsubseteq q$ iff $p \preceq f^*(q)$, then the pair $\langle f_*, f^* \rangle$ form a Galois connection between $\mathscr{P}$ and $\mathscr{Z}$.

If $\langle f_*, f^* \rangle$ is such a connection, f_* is said to be the *left adjoint* of the corresponding f^*, and f^* is *the right adjoint* of f_*.

f_* appears to the left of its order sign in (G), and f^* to the right of its order sign. Alternatively, the terminology `lower adjoint' vs `upper adjoint' is used (i.e. f_* appears on the lower side of its ordering sign, and f^* on the upper side).

Theorem 3.1.1: Let $\mathscr{P} = \langle P, \preceq \rangle$, $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$, and $\mathscr{R} = \langle R, \sqsubseteq \rangle$ be posets. Let suppose $\langle f_*, f^* \rangle$ is a Galois connections between $\mathscr{P}$ and $\mathscr{Q}$, and $\langle g_*, g^* \rangle$ is a Galois connections between $\mathscr{Q}$ and $\mathscr{R}$. Then $\langle g_* \circ f_*, f^* \circ g^* \rangle$ is a Galois connection between $\mathscr{P}$ and $\mathscr{R}$ [18].

Theorem 3.1.2: Suppose $\mathscr{P} = \langle P, \preceq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$ are posets, and $f_* : P \to Q$ and $f^* : Q \to P$ are a pair of functions between their carrier sets. Then $\langle f_*, f^* \rangle$ is a Galois connection if and only if:

(i) f_*, f^* are both monotone, and

(ii) for all $p \in P$, $q \in Q$, $p \preceq f^*(f_*(p))$ and $f_*(f^*(q)) \sqsubseteq q$.

B. The Relation between Adjoints

The next theorem tells that the relation between the adjoint members of a Galois connection is rigid in the sense that if $\langle f_*, f^* \rangle$ is to be a connection, then f_* fixes what f^* uniquely has to be, and conversely f^* fixes what f_* has to be.

Theorem 3.2.1: If $\langle f_*, f^1 \rangle$ and $\langle f_*, f^2 \rangle$ are Galois connections between $\langle \ , \ \rangle$ and $\langle Q, \sqsubseteq \rangle$, then $f^1 = f^2$. Likewise, if $\langle f_1, f^* \rangle$ and $\langle f_2, f^* \rangle$ are Galois connections between the same posets, then $f_1 = f_2$ [6].

Theorem 3.2.2: If $\langle f_*, f^* \rangle$ is a Galois connection between $\langle P, \preceq \rangle$ and $\langle Q, \sqsubseteq \rangle$, then

(i) $f^*(q) =$ the maximum of $\{ p \in P \mid f_*(p) \sqsubseteq q \}$

(ii) $f_*(p) =$ the minimum of $\{ q \in Q \mid p \preceq f^*(q) \}$

C. Fixed Points and Closures

Recall again that we use $f(P)$ for the image of the set P under f, i.e. $f(P) = \{ f(p) \mid p \in P \}$. And p is a fixed point of a function f if $f(p) = p$. Then we have the following:

Theorem 3.3.1: If $\langle f_*, f^* \rangle$ is a Galois connection between $\langle P, \preceq \rangle$ and $\langle Q, \sqsubseteq \rangle$, then

(i) $f_* \circ f^* \circ f_* = f_*$ and $f^* \circ f_* \circ f^* = f^*$,

(ii) $p \in f^*(Q)$ if and only if p is a fixed point of $f^* \circ f_*$; and $q \in f_*(P)$ if and only if q is a fixed point of $f_* \circ f^*$.

(iii) $f^*(Q) = f^*(f_*(P))$ and $f_*(P) = f_*(f^*(Q))$.

Definition 3.3.1: Given a Galois connection between $\mathscr{P} = \langle P, \preceq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$, let $\mathscr{P}^f$ be ‘s sub-poset $\langle f^*(f_*(P)), \preceq \rangle$, where $\preceq$ here is $\mathscr{P}$ ‘s order relation restricted to $f^*(f_*(P))$. Similarly, put $\mathscr{Q}_f$ for the corresponding sub-poset $\langle f_*(f^*(P)), \sqsubseteq \rangle$. Then the last theorem can be used to prove a more consequential one [6].

Theorem 3.3.2: If $\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{P} = \langle P, \preceq \rangle$ and $\mathscr{Q} = \langle Q, \sqsubseteq \rangle$, then $\mathscr{P}^f$ and $\mathscr{Q}_f$ are order-isomorphic.

Definition 3.3.2: Suppose $\mathscr{P} = \langle P, \preceq \rangle$ is a poset; then a closure function for $\mathscr{P}$ is a function c such that, for all $p, p' \in P$,

(i) $p \preceq c(p)$;

(ii) if $p \preceq p'$, then $c(p) \preceq c(p')$, i.e. c is monotone;

(ii) $c(c(p)) \preceq c(p)$ i.e. c is `idempotent'.

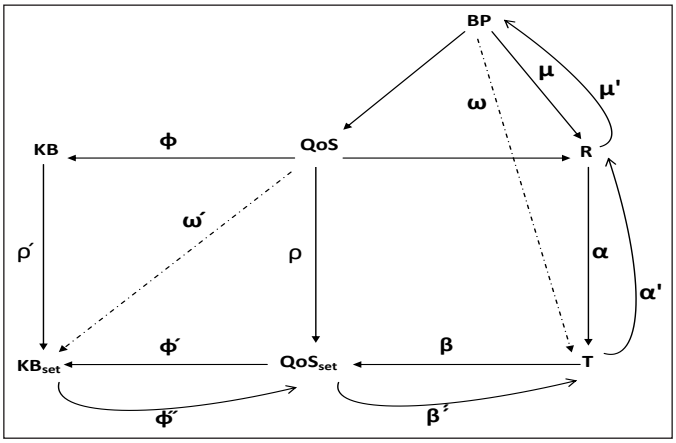
Theorem 3.3.3: If $\langle f_*, f^* \rangle$ is a Galois connection between $\mathscr{P} = \langle P, \preceq \rangle$ and some poset, then $f^* \circ f_*$ is a closure function for $\mathscr{P}$ [6].

Definition 3.3.3: Suppose γ, δ are sets, and let R be any relation between their elements. Define a function f_R from subsets of γ to subsets of δ as follows: if $\alpha \in \gamma$, then $f_R(\alpha)$ is the set of things which are R-related to everything in α. In other words, $f_R(\alpha) = \{ b \mid \forall a \in \alpha \to aRb \}$. And define a corresponding function f^R back from subsets of δ to subsets of γ like this: if *if* $\beta \subseteq \delta$ *then* $f^R(\beta) = \{ a \mid \forall b (b \in \beta \to aRb) \}$.

Then $\langle f_R, f^R \rangle$ is a Galois connection between the inclusion posets $\langle \rho(\gamma), \subseteq \rangle$ and $\langle \rho(\delta), \supseteq \rangle$ [13].

IV. The Galois Theory and QoS Driven Knowledge Extraction Approach

In this section, we present our approach based on Galois connections to extract the knowledge required for the satisfaction of a quality of service previously requested by a client. This approach is summarized by the following model:



The interpretation of this model is based on a number of steps:

Step 1: Request for a QoS by a Service

This step starts the extraction process. For this, the request is made by the client by presenting a quality of service in relation to a product or service in a company. This assume that the request meets the QoS model presented above i.e. the client submits a set of criteria or quality factors in relation to a service.

Step 2: The Deduction of Business Processes

Because the client specifies the requested service, it is obvious from the definition and business process model to find the BP to implement in order to achieve this QoS.

Step 3: Obtaining the Execution Path of BP

A BP is a tuple (P, S), that is a set of tasks required to provide a service, we deduce from the definition of a BP the set of tasks and therefore, the execution path required to provide the service requested. Conversely, from an execution path, we get the BP in question.

Step 4: Selection Knowledge Required to Achieve the Defined QoS

The previous steps enable the determination of the execution path of BP required to achieve the service for which a certain quality was requested by a customer. We know that service quality is defined by the QoS standard. This is an aggregation of QoS is obtained after partial execution of tasks. From a partial QoS is associated, according to the theory of BP, the knowledge kb required to achieve it after the execution of a task t. But the quality of service standard may be different from the QoS requested by a client and therefore, standard knowledge shall not achieve this quality of service. This is why a new selection of knowledge must be made to obtain the requested QoS. This assumes that each task of a BP, one must associate a pair (q_i, k_i) to satisfy the customer. The determination of these couple will be done by applying the properties of Galois connections to the extraction context that we shall first define.

Step 5: Definition of the Context of Extraction

Within this work, we shall define a context called *K* extraction context, as a tuple *((T, QoS), KB,W)* where *T* is the set of tasks of a *BP*, the set partial *QoS, KB* the set of knowledge to be required to achieve the partial *QoS* and *W* a binary relation $W \subseteq (BP, QoS_{set}) x KB_{set}$;

Step 6: Determination of the Couple Quality of Service-Knowledge Associated with Each Task in BP Execution Path

From the BP theorie, $(KB_{set}, \subseteq)$; $(QoS_{set}, \subseteq)$; $(T, \subseteq)$; $(BP, \subseteq)$ et $(R, \subseteq)$ are posets

$$QoS_{set} = \bigcup_{i=1}^{k} q_i \ ; \ BP = \bigcup_{k=1}^{m} t_k \ ; \ KB_{set} = \bigcup_{j=1}^{n} kb_j$$

The functions $\phi, \rho, \rho', \phi', \phi'', \beta, \beta', \alpha, \alpha', \mu, \mu'$ are defined as follows:

$$\rho : QoS \rightarrow QoS_{set} \ ; \ \rho' : KB \rightarrow KB_{set} \ \phi : QoS \rightarrow KB$$
$$\phi' : QoS_{set} \rightarrow KB_{set} \ ; \ \phi'' : KB_{set} \rightarrow QoS_{set} \ ; \ \beta : T \rightarrow QoS_{set} \ ;$$
$$\beta' : QoS_{set} \rightarrow T \ ; \ \alpha : R \rightarrow T \ ; \ \alpha' : T \rightarrow R \ ;$$
$$\mu : BP \rightarrow R \ ; \ \mu' : R \rightarrow BP$$

In addition:

- $\forall q \in QoS_{set}, \ \forall k \in KB_{set}, \ \phi'(q) \subseteq k \text{ iff } q \subseteq \phi''(k)$
- $\forall t \in T, \ \forall q \in QoS_{set}, \ \beta(t) \subseteq q \text{ iff } t \subseteq \beta'(q)$
- $\forall r \in R, \ \forall t \in T, \ \alpha(r) \subseteq t \text{ iff } r \subseteq \alpha'(q)$
- $\forall bp \in BP, \ \forall r \in R, \ \mu(bp) \subseteq r \text{ iff } bp \subseteq \mu'(r)$

From above, we deduce that: $\langle \phi', \phi'' \rangle, \langle \beta, \beta' \rangle, \langle \alpha, \alpha' \rangle, \langle \mu, \mu' \rangle$ are the Galois connections between QoS_{set} and KB_{set}, *T* and QoS_{set}, *R* and *T, BP* and R respectively. From the definition of a Galois connection, we deduce that:

ϕ' is the *left adjoint* of the corresponding ϕ'', and ϕ'' is *the right adjoint* of ϕ'

b is the *left adjoint* of the corresponding *b'*, and *b'* is *the right adjoint* of *b*

a is the *left adjoint* of the corresponding *a'*, and *a'* is *the right adjoint* of *a*

m is the *left adjoint* of the corresponding *m'*, and *m'* is *the right adjoint* of *m*

Since $\langle \phi', \phi'' \rangle$ is a Galois connection between KB_{set} and QoS_{set} *and* $\langle \beta, \beta' \rangle$ a Galois connection between QoS_{set} and *T,* then $\langle \beta o \phi', \beta' o \phi'' \rangle$ a Galois connection between KB_{set} and *T.*

In a similar manner, $\langle \mu o \alpha, \mu' o \alpha' \rangle$ is a Galois connection between *T and BP.*

From the Theorem 3.1.2, we deduce that:

(i)

- ϕ', ϕ'' are both monotone; - *b,b'* are both monotone; - *a,a'* are both monotone;

- *m,m'* are both monotone, and

(ii)

for all $q \in QoS_{set}$, $k \in KB_{set}$, $q \subseteq \phi''(\phi'(q))$ and $\phi'(\phi''(k)) \subseteq k$;

for all $t \in T$, $q \in QoS_{set}$, $t \subseteq \beta'(\beta(t))$ and $\beta(\beta'(q)) \subseteq q$

for all $r \in R$, $t \in T$, $r \subseteq \alpha'(\alpha(r))$ and $\alpha(\alpha'(t)) \subseteq t$

for all $bp \in BP$, $r \in R$, $bp \subseteq \mu'(\mu(bp))$ and $\mu(\mu'(r)) \subseteq r$

Then we conclude that from a given *QoS*, we obtain the set of tasks required to achieve it and therefore, all the partial *QoS* related to the performance of each task. In addition, from this set of partial *QoS*, we obtain the set of knowledge required to achieve these partial *QoS* and therefore the set of tuple *(qi, ki)* used to achieve the requested *QoS*.

V. Conclusion

The main technical content of this paper is to present a method, based on Galois Theory, which extract from the knowledge base, relevant knowledge needed for the execution of business process to achieve a given quality of service. This amounts to determine, for a given quality of service, the set of tuples quality of service, knowledge and task that form the execution path of the business process to achieve this quality of service. For this end, we have shown that these tuples can be obtained by applying Galois Connection Properties from our defined extraction context. We believe that the application of the resulting method in daily work will improve quality of service in enterprises in order to deal with the competitive pressure of the network economy. However, the next step in this work, which is in progress, addresses the incompleteness problem of knowledge and natural language processing method for the extraction of part of a text or document.

References

[1] G. A. Alo'o, E. R. Atsa, and V. Monthe, "Quality of service optimization based on HRM: A case study in Cameroonian public service," *International Journal of Computer Applications*, vol. 174, no. 11, pp. 10-19, 2021.

[2] E. R. Atsa, M. F. Ndjodo, P. E. Ndedi, and G. A. Alo'o, "Human resource load balancing based on Ant Theory for QoS management within an enterprise in a developing country," CARI'2010.

[3] E. R. Atsa, M. F. Fouda, C. L. Atouba, and G. A. Alo'o, "Knowledge management driven business process and workflow modeling within an organization for customer satisfaction," *IJCA*, vol. 2, no. 12, pp. 7350-7362, 2010.

[4] E. R. Atsa, and G. A. Alo'o, "Quality of service improvement based on procedure analysis and reengineering: A case study in Cameroonian public service," *Int. J. Electronic Governance*, vol. 9, no. 1/2, 2017.

[5] S. Anderson, and M. Felici, "Requirements evolution from process to product oriented management," In *Proceedings of Profes. 3rd International Conference on Product Focused Software Process Improvement*, Kaiserslautern, Germany, Sept. 10-13, 2001, LNCS 2188, Springer-Verlag.

[6] D. A. Cox, *Galois Theory*. John Wiley & Sons, 2012.

[7] D. Oprea, and G. Mesnita, "The information systems documentation – Another problem for project management," *Managing Information in the Digital Economy: Issues & Solutions*, pp. 332-338, 2006.

[8] D. Hollinsworth, "The workflow reference model," Technical Report TC00-1003, Workflow Management Coalition, 1994. [Online]. Available: http://www.aiai.ed.ac.uk/WfMC/

[9] S. Gandhi, "Knowledge management and reference services," *Journal of Academic Librarianship*, vol. 30, no. 5, pp. 368-381, 2004.

[10] L. L. Ralph, and T. J. Ellis, "An investigation of a knowledge management solution for the improvement of reference services," *Journal of Information, Information Technology, and Organizations*, vol. 2, 2009.

[11] J. Bisbal, D. Lawless, B. Wu, and J. Grimson, "Legacy information system migration: A brief review of problems, solutions and research issues," Technical Report, 1999.

[12] J. Bisbal, D. Lawless, R. Richardson, D. O'Sullivan, B. Wu, J. Grimson, and V. Wade, "The Butterfly methodology: A gateway-free approach for migrating legacy information systems," 1997.

[13] J.-P. Tignol, *Galois' Theory of Algebraic Equations*. World Scientific Publishing, 2001.

[14] M. Brodie, and M. Stonebraker, "DARWIN: On the incremental migration of legacy information systems," TR-022-10-92-165, GTE Labs Inc., 1993.

[15] N. Ganti, and W. Brayman, *Transition of Legacy Systems to a Distributed Architecture*. John Wiley & Sons Inc., 1995.

[16] E. R. Atsa, M. F. Ndjodo, and C. L. Atouba, "A goal oriented approach or the definition of business process requirement model," *IJCA*, 2010.

[17] J. Rowley, "What is knowledge management?," *Library Management*, vol. 20, no. 8, p. 416, 1999.

[18] S. Bosch, *Algebra*. Birkhauser, 2018.

[19] W. M. P. van der Aalst, A. H. M. ter Hofstede, B. Kiepuszewski, and A. P. Barros, "Workflow patterns," Technical Report WP 47, Beta Research Institute, 2000.

International Journal of Knowledge Based Computer Systems
12 (1) 2024, 01-06
http://www.publishingindia.com//ijkbcs/

Estimation of Sex from Sternum Measurements and Comparative Analysis using Machine Learning Based Feature Selection Algorithm

Kirti Sharma[1*], Neha Pokhriyal[2] and Anisha Beniwal[3]
[1]Department of Paramedical Sciences, Quantum University, Roorkee, India.
Email: kirtisharma.hs@quantumeducation.in
[2]Computer Science Engineering, Quantum University, Roorkee, India.
Email: nehanpokhriyal@gmail.com
[3]Department of Applied Sciences, Quantum University, Roorkee, India.
Email: anisha.hs@quantumeducation.in
[*]Corresponding Author

Abstract: This prospective study utilized multi-detector computed tomography (MDCT) to examine the morphological characteristics and sex-related variations of the sternum in 150 adult patients (92 males and 58 females) aged 20–80 years. The sternum, a skeletal component is a strong bone that doesn't deform easily therefore, it can be used for age and sex prediction by researchers for sex determination. Sternum images were reconstructed using the Volume Rendering (VR) technique, and measurements of different sternal parts were obtained from sagittal and coronal images. Statistical analyses, including independent sample t-tests and discriminant analysis, revealed significant differences in sternum measurements between male and female groups ($p<0.001$), with body length and manubrium width emerging as the most crucial parameters. The study demonstrated the reliability and utility of sternal morphometric analysis for accurate sex estimation. Machine learning-based feature selection algorithms were utilized for comparative analysis of selected features. The findings emphasize the importance of employing scientifically validated methodologies, highlighting manubrium length/width and body length as key parameters for accurate gender determination using sternum measurements.

Keywords: Forensic, Manubrium, Mesosternum, Nearest neighbour classifier, ROC curve, Sternal area, Sternal cleft, Sternal foramen, Sternal index.

I. Introduction

Computed Tomography (CT) revolutionized medical imaging upon its invention in 1972 by Sir Godfrey Hounsfield and Allan Cormack, earning them the Nobel Prize for their contributions. This technology, characterized by its rapid advancements, has significantly enhanced imaging speed, patient comfort, and resolution [1]. CT operates on the principle of computerized X-ray imaging, wherein an X-ray beam rotates around the body, producing cross-sectional images or "slices" processed by the machine's computer. Helical or spiral CT, facilitated by slip ring technology, allows continuous rotation of the X-ray tube around the patient, improving imaging efficiency. Multidetector CT (MDCT) systems, a later development, collect multiple slices of data rapidly and reconstruct images within seconds, providing

enhanced visualization of tissues, particularly bone tissue [2].

The present study focuses on the application of sternum measurements for sex determination, especially beneficial in challenging identification scenarios. Forensic examiners analyse bone metrics to generate a biological profile encompassing age, sex, stature, and lineage [3].

Skeletal remains such as the pelvis, long bones, and skull traditionally aid sex determination, but in cases of trauma or degradation, the sternum emerges as a reliable alternative due to its resilience and accessibility. Various methods, including osteometry, odontometric data analysis, and DNA analysis, are employed for sex determination, with osteometry favoured for its simplicity and accuracy [4]. Previous research highlights sexual dimorphism in sternum measurements, enabling forensic professionals to predict sex accurately. Understanding sternal sexual dimorphism enhances identification processes in forensic settings. "Hyrtl's law," indicating differences in sternum length between males and females, serves as a foundation for sex determination as it states that the manubrium of the female sternum extends beyond half the length of the body, whereas the body length in the male sternum is approximately double that of the manubrium [5]. Imaging techniques such as Multiplanar reformation (MPR) aid in measuring sternum parameters accurately and with the help of a measurement tool all the measurements were calculated for an individual like manubrium length, Mesosternum length, sternum width, sternal index and whole sternum length of an individual patient [6].

A. Anatomy of Sternum

The sternum is a central component of the thoracic cage and is located in the front wall. It is divided into three sections: the manubrium, the body (corpus), and the xiphoid process. Because of its simplicity of use, it can be employed in the field of forensic medicine for assessing age as well as gender using imaging techniques.

Secondary cartilaginous joints connect these components, with hyaline cartilage and fibrocartilaginous discs acting as intermediaries. The manubrium is a flat, quadrilateral structure that widens superiorly. The body is a rectangular, flat bone with facets or grooves along its lateral edge that allow the ribs to be attached. It measures around 20 cm long, 3-4 cm wide, and 1 cm thick. Finally, the xiphoid process is a thin, inferiorly protruding structure.

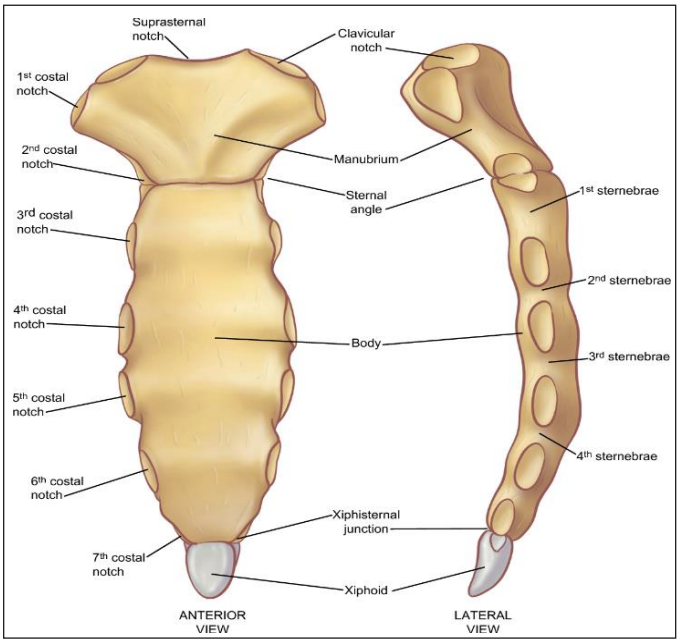


Fig. 1: Sternum Anatomy (Anterior & Lateral View)

B. Hypothesis

The measures obtained from a Multi-Detector Computed Tomography (MDCT) scan of the sternum can assist distinguish males and females. Various sternum dimensions, such as length, width, and total sternal length, as well as other characteristics, may be accurate markers for determining an individual's sex. It is predicted that by analysing these measurements and their statistical significance, a clear separation between male and female sternums can be made. If the measurements regularly show statistically significant variations between the sexes, it would imply that sternum measurements can be used to determine an individual's sex.

II. Aim and Objectives

A. Aim

The study aims to estimate the sex of an individual from sternum bone measurement by using Multidetector computed tomography.

B. Objective

- To investigate sternum ossification and anatomical changes based on age and gender.
- To ascertain the role of computed tomography scans in sex estimation.
- To comprehend the significance of 3D reconstruction and the measurement technique utilised to calculate sternum measurements.
- To assess the accuracy of gender determination using sternum measurement in a multi-detector computed tomography scan.
- To determine the measurement discrepancy according to the system software.

III. Material and Methods

A. Source of Data

Hospital based cross-sectional study that was carried out in Department of Radio-Diagnosis at Shree Guru Gobind Tricentenary (SGT) Hospital and Research Institute for CT chest examination.

B. Study Duration

This study was carried out over a period of 2 year. The data was collected from August 2022 to June 2023 in the department of Radio-Diagnosis of SGT Medical College and Hospital and Research Institute, Gurugram.

C. Study Type and Design

This study was designed to be a prospective, quantitative and comparative study about the determination of sex from various measurements of the sternum bone by using an MDCT scan in adult patients of 20-80 years of age. A patient who underwent CT chest scan, a sternum reconstructed image is created and different measurement is taken according to efficacy and study demand. No participant was made to undergo a CT scan for the sole purpose of this study.

D. Study Area

CT area of Department of Radio-Diagnosis of SGT Medical College and Hospital and Research Institute, Gurugram.

E. Inclusion Criteria

- Patients who have ages from 20 yrs to 80 yrs.
- Both sexes were included in this study.
- Only CECT examination of selected patient will be included in this study.

F. Exclusion Criteria

- If a patient doesn't have a fused sternum.
- Any patient who has a history of trauma and deformity related to the sternum.

G. Sample Size

A convenient sample of 150 patients of all age group were taken who underwent CECT & HRCT Thorax on 16 slices SOMATON SIEMENS. This study sample was included from a section of the population that is effortlessly accessible or readily available.

IV. Methodology of Data Collection

With the support of SGT University's Radio-Technologist, CT sternum images were taken, and overall data was collected. The information was gathered in the form of daily benefits. For the CT chest Imaging of the adult patients, our hospital uses routine sequences of thorax CECT and HRCT. After collecting data, for taking a measurement of the sternum Radian 3.0 software was used and measurement are done on sagittal and coronal images 3D reconstructed image of sternum. Data is filled

and expressed in excel sheet. A master chart is made which is attached with this study.

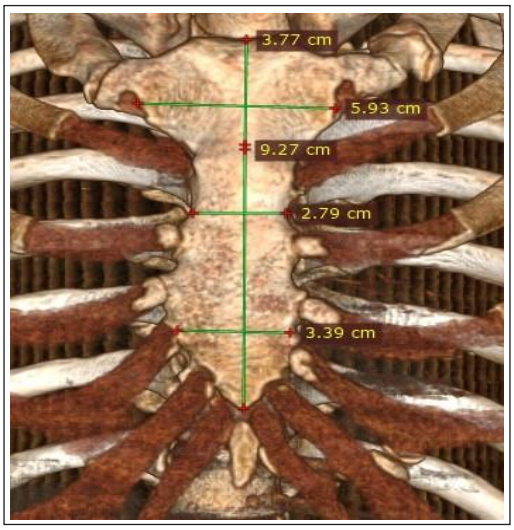


Fig. 2: MDCT Reformatted Image Shows MW (Manubrium Width), Sternebra 1 Width (S1W) and Sternebra 3 Width (S3W). Manubrium Length. (ML) and Mesosternum Length (MSL)

V. Results

Estimation of gender by using sternum measurement was well entrenched in this study. In current study, a total number of 150 patient's samples (92 males and 58 females) were included. The descriptive analysis and independent t-test was utilized to show the mean, standard deviation and significance of the sample. Each extracted feature is analysed statistically and for the comparison purpose a machine learning based feature selection algorithm is applied where AUC value of each feature subset is calculated. After t-value calculation the parameter with least t-value is more relevant and the same verification of extracted feature is done using wrapper-based feature selection [10]. Each feature vector has an AUC value, among the highest AUC value we select a feature vector again to calculate AUC values if new AUC is less than previous highest value than the previous feature vector is best vector otherwise calculate another AUC value for new feature vector. Total 6 feature vector outcome as final feature subset. The AUC curve is generated using the NN classifier for the simplicity and normalization of each feature is done before value generation. Order of features according to AUC:(3)> (7)> (6)> (1)> (9)> (5)> (8)> (2)> (4) Base feature vector: top AUC rankers taken as base set 5 features are selected, F_{base1} = [F3, F7, F6, F1, F9], Subset F_{base2} = [F3, F7, F6, F1, F9, F5] is the final subset because AUC value of subset 3 reduced as shown in Fig. 3 and we add feature in the previous until new value of AUC is not less than the last one [9].

Tabel I: Table for Selected Features using Two Algorithms

Unreduced or Raw Data	AUC Based Feature Selection
9 Extracted features [1,2,3,4,5,6,7,8,9]	Feature subset 2 [3,7,6,1,9,5]

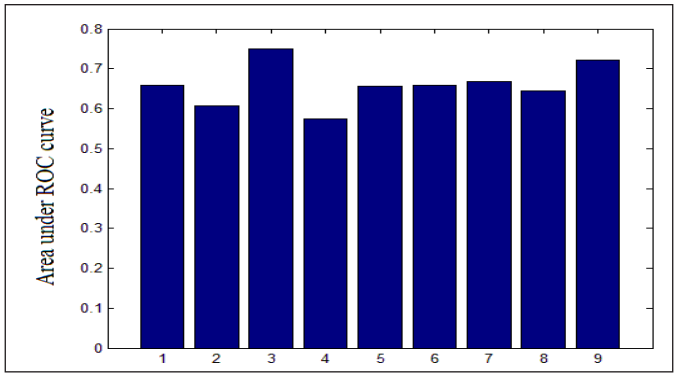


Fig. 3: Comparison of AUC Values of Each Extracted Features

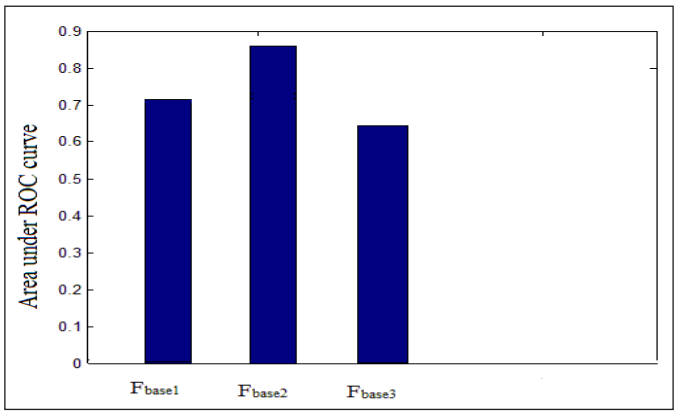


Fig. 4: AUC Value for Different Feature Subsets

Tabel II: Table for Selected Features (T-Value)

Measurements	Extracted Features	Calculated (t-Value)
1	Manubrium Length	6.25
2	TSL	10.86
3	Age	0.582
4	Manubrium Width	8.02
5	Sternal Area	10.76
6	SW1	2.88
7	Sternal Index	1.51
8	Mesosternum Length	8.73
9	SW2	3.14

After applying statistical t-test and machine learning AUC value, we can conclude that both algorithms showing the same features approximately. The order of selected features according to AUC value is Age > Sternal Index > SW1 > Manubrium length > SW2 > Sternal area.

VI. Discussion

Among all the radiological modality, CT is the pioneer choice for the evaluation of sex. Identification of skeletal remains is crucial because the estimation can be compared with antemortem data and other information that aids in the identification process. Moreover, sex assessment serves as a foundation for developing other aspects of a biological profile [7]. The current study affirms that sex differences in the sternum can be witnessed, however, these changes are peculiar to specific populations. Data from the study were compared to prior studies that covered varied populations from throughout the world, including Europe, India, Turkey, Jordan, and others. It is vital to notice that these differences change not just physically but also temporally. The study found that in the presence of an intact anterior thoracic cage, the sterna of the sample demonstrated significant sexual dimorphism to allow for a high degree of sex identification. Women's sternums were usually smaller than men's, which supports earlier findings in diverse populations. (Changani *et al.,* 2014): Indians (Singh and Pathak 2013), South Africans (Macaluso, 2010), Spaniards (Macaluso and lucena, 2014): West Australian (Franklin *et al.,* 2012) and Americans (Bongiovanni and Spradley, 2012). The initial evidence of sexual dimorphism in the sternum dates back to the late 18th century, when Wenzel conducted investigations comparing the lengths of the manubrium and corpus sterni in males and females. Building upon Wenzel's work, Dwight and Hyrtl conducted similar studies and arrived at a concept known as 'Hyrtl's law'. According to this law, there were observable differences in body shape between sexes, with men exhibiting a larger anterior thoracic cage and women displaying a longer manubrium to sternal body ratio. Even, researcher concluded that sternal index would be 50 in females. However, it has been noticed that the sternal index, which was previously utilised as a sexing criterion, is unreliable due to considerable sex overlap and accompanying sex bias. In different populations, different rules have been developed for sexing the sternum. Ashley, for example, devised the "149 rule" for sexing Europeans, which specifies that male sternal length is 149 mm and female sternal length is less than 149 mm [5]. Various criteria have been devised for various populations, such as the "136 rule" for East Africans, the "129 rule" for India's Marathwada region, the "131 rule" for West Indians, and the "142 rule" for Anatolians. In our investigation, we observed that the "144 rule" applied to the sample population, with a total sternum length greater than 144 mm indicating a male and a length less than 144 mm indicating a female. Notably, our data demonstrated that the total length of the sternum was more reliable as a measure for diagnosing sternum sex than the other factors studied.

VII. Conclusion

CT is a pioneer choice for sex evaluation among radiological modalities. The sternum exhibits sexual dimorphism specific to populations, and its assessment aids in identification processes. Various rules and indices have been formulated to determine sex based on sternal measurements, although some parameters may show overlap between sexes. Additionally, the sternum can exhibit anatomical variations and anomalies, and awareness of these variations is crucial for accurate diagnosis. Multiplanar and volume-rendering

reconstructed images are excellent tools for visualizing sternal anatomy and properly describing sternal variations and anomalies. These cutting-edge imaging techniques are extremely successful at revealing the delicate details of the sternum from various angles. Medical practitioners can observe the sternal structure in sagittal, coronal, and axial planes using multiplanar reconstruction, allowing for a more comprehensive assessment. Furthermore, volume-rendering reconstruction approaches allow for the fabrication of three-dimensional representations, which improves comprehension of sternal complexity. With these invaluable instruments, healthcare practitioners can acquire vital insights into the sternum's particular traits and anomalies, allowing for precise diagnosis and educated treatment recommendations. CT imaging plays a significant role in revealing the anatomy and variations of the sternum, enabling reliable assessments and avoiding diagnostic uncertainties.

Estimating gender is essential for identifying victims of terrorist acts, catastrophe victims, and legal cases involving medicine. In this present study, Males had greater mean and standard deviation values than females. The gender of an individual subject may be established with 91.4% accuracy using sternum measurement. The current study concluded that measuring manubrium width or length and body length can be used to accurately estimate gender in this sample. However, sternal area is not that reliable parameter for differentiation.

References

[1] Brief history of CT, *Imaginis.com*. Accessed: Apr. 29, 2024. [Online]. Available: https://www.imaginis.com/ct-scan/brief-history-of-ct?r

[2] E. Yekeler, M. Tunaci, A. Tunaci, M. Dursun, and G. Acunas, "Frequency of sternal variations and anomalies evaluated by MDCT," *AJR Am. J. Roentgenol.*, vol. 186, no. 4, pp. 956-960, 2006.

[3] R. Bongiovanni, and M. K. Spradley, "Estimating sex of the human skeleton based on metrics of the sternum," *Forensic Sci. Int.*, vol. 219, no. 1-3, p. 290.e1-290.e7, 2012.

[4] D. S. Grewal, "Morphometric analysis of odontometric parameters for gender determination," *J. Clin. Diagn. Res.*, 2017.

[5] J. Hyrtl, and T. Dwight, "The sternum as an index of sex, height and age," *J Anat Physiol*, vol. 15, pp. 327-330, 1881.

[6] A. Saraf, T. Kanchan, K. Krishan, N. Ateriya, and P. Setia, "Estimation of stature from sternum – Exploring the quadratic models," *J. Forensic Leg. Med.*, vol. 58, pp. 9-13, 2018.

[7] S. A. Hunnargi et al., "Sexual dimorphism of the human sternum in a Maharashtrian population of India: A morphometric analysis," *Leg. Med. (Tokyo)*, vol. 10, no. 1, pp. 6-10, 2008.

[8] R. Kohavi, and G. H. John, "Wrappers for features subset selection," *Int. J. Digit. Libr*, vol. 1, pp. 108-121, 1997.

[9] S. Kim, K.-T. Kim, and S. Kim, "Infrared small target discrimination using sequential forward feature selection with AUC metric," in *IEEE International Conference on Industrial Technology (ICIT)*, Busan, Korea, 2014.

[10] S. Maldonado, and R. Weber, "A wrapper method for feature selection using support vector machine," *Information Sciences*, vol. 179, no. 13, pp. 2208-2217, 2009.

International Journal of Knowledge Based Computer Systems
12 (1) 2024, 07-15
http://www.publishingindia.com//ijkbcs/

Artificial Intelligence and Machine Learning in Healthcare: Revolutionizing Diagnostics and Treatment

Rashmi Tyagi
Assistant Professor, Quantum School of Health Sciences, Quantum University, Roorkee, Uttarakhand, India. Email: ty.rashu01@gmail.com

Abstract: This paper explores the profound impact of Artificial Intelligence (AI) and Machine Learning (ML) applications in the healthcare sector, focusing on their revolutionary role in diagnostics and treatment methodologies. In an era of rapid technological advancement, AI and ML have emerged as potent tools poised to revolutionize healthcare delivery, enhanced precision and ultimately, improved patient outcomes. Addressing prevalent challenges in healthcare diagnostics, particularly the imperative for faster and more accurate disease identification, the paper highlights how AI and ML algorithms, empowered by extensive datasets, can swiftly and accurately analyze intricate medical information. It emphasizes the potential of these technologies to support healthcare professionals in early disease detection, risk assessment, and the personalization of treatment plans. Furthermore, the paper delves into real-world applications of AI and ML in treatment strategies, encompassing areas such as drug discovery, treatment regimen optimization, and the customization of therapeutic interventions. Ethical considerations surrounding AI and ML implementation in healthcare, including data privacy, transparency, and bias mitigation, are also examined. Emphasizing the necessity of a collaborative approach involving technologists, healthcare professionals, and policymakers, the paper advocates for responsible and equitable integration of these technologies. Drawing on case studies and empirical evidence, this paper offers insights into both the successes and challenges encountered in the incorporation of AI and ML in diagnostics and treatment. By fostering a deeper understanding of the potential benefits and ethical considerations, it aims to contribute to the ongoing discourse on leveraging artificial intelligence for the advancement of healthcare and sustainable progress in medical science.

Keywords: Artificial intelligence, Diagnostics, Ethical considerations, Healthcare, Machine learning, Personalized medicine, Treatment.

I. Introduction

In recent years, the integration of Artificial Intelligence (AI) and Machine Learning (ML) have catalyzed a transformative wave across various industries, with healthcare being at the forefront of this revolution. AI and ML technologies are increasingly being harnessed to address some of the most pressing challenges in healthcare, particularly in the domains of diagnostics and treatment.

A. Overview of the Integration of AI and ML

This introduction provides an overview of how AI and ML are reshaping healthcare and underscores their significance in revolutionizing diagnostics and treatment methodologies.

The healthcare sector generates a vast amount of data from various sources, including electronic health records (EHRs), medical imaging, genomic sequencing, and wearable devices. Traditionally, the analysis of such data has been a labor-intensive

and time-consuming process, often fraught with errors and inconsistencies. However, AI and ML offer a compelling solution by leveraging algorithms that can learn from data, recognize patterns, and make predictions with unprecedented accuracy and efficiency.

AI and ML algorithms are particularly well-suited for tasks such as medical image analysis, where they can detect subtle abnormalities and assist clinicians in making more accurate diagnoses. For example, convolutional neural networks (CNNs), a type of deep learning algorithm (Fig. 1), have demonstrated remarkable performance in interpreting medical images, such as X-rays, MRIs, and CT scans, often outperforming human experts in certain tasks (Fig. 2) [1] [2] [3].

Furthermore, AI and ML enable the development of predictive models that can forecast disease progression, identify high-risk patient populations, and optimize treatment strategies. By analyzing large datasets encompassing clinical, genomic, and demographic information, these models can identify complex relationships and patterns that may not be apparent to human observers. This predictive analytics capability holds immense potential for personalized medicine, where treatments can be tailored to individual patients based on their unique characteristics and genetic makeup [4].

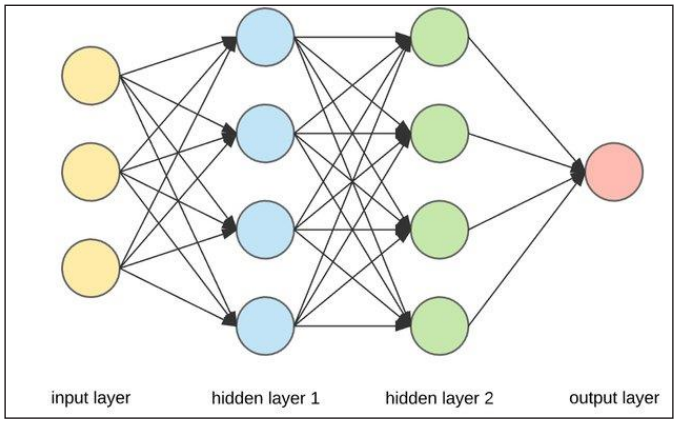


Fig. 1: Deep Learning

A simple, multilayer deep neural network takes two classes of data. Example large-scale network that accepts as input a variety of data types (images, time-series, etc.), and for each data type learns a useful featurization in its lower level towers. The data from each tower is then merged and flows through higher levels, allowing the DNN to perform inference across data types-a capability that is increasingly important in healthcare.

B. Importance of AI and ML in Revolutionizing Diagnostics and Treatment Methodologies

The integration of AI and ML in healthcare is not only revolutionizing diagnostics but also transforming treatment methodologies. AI-powered decision support systems can assist clinicians in selecting the most effective treatment options by analyzing patient data and considering factors such as treatment response history, comorbidities, and drug interactions. Moreover, AI-driven approaches enable the discovery of novel biomarkers and therapeutic targets, accelerating the development of new drugs and therapies.

The integration of AI and ML in healthcare represents a seismic shift in how we approach diagnostics and treatment. These technologies hold the promise of improving patient outcomes, reducing healthcare costs, and enhancing the overall quality of care. However, realizing this potential requires addressing various challenges, including data privacy and security, algorithm bias, and regulatory considerations. By overcoming these hurdles and embracing the transformative power of AI and ML, we can usher in a new era of precision medicine and personalized healthcare [5].

II. Enhanced Diagnostics

Artificial Intelligence (AI) and Machine Learning (ML) have revolutionized the field of medical imaging diagnostics, offering unprecedented accuracy, speed, and efficiency compared to traditional methods. This section explores the pivotal role of AI and ML in analyzing medical images such as X-rays, MRIs, and CT scans, highlighting their superiority over conventional diagnostic approaches and providing case studies and examples of successful implementation.

A. Role of AI and ML in Analyzing Medical Images (X-rays, MRIs, CT Scans)

AI and ML algorithms are adept at analyzing complex medical images and detecting subtle abnormalities that may escape the human eye. Convolutional Neural Networks (CNNs), a type of deep learning algorithm, have emerged as particularly effective tools for image recognition and classification tasks in healthcare. These algorithms can learn from vast datasets of labeled medical images, enabling them to identify patterns and features indicative of various diseases and conditions.

For instance, in radiology, AI-powered systems can assist radiologists in interpreting X-rays by automatically detecting abnormalities such as fractures, tumors, or pneumonia infiltrates. Similarly, in MRI and CT imaging, AI algorithms can segment organs and tissues, quantify volumes, and identify lesions or anomalies with high accuracy (Fig. 2).

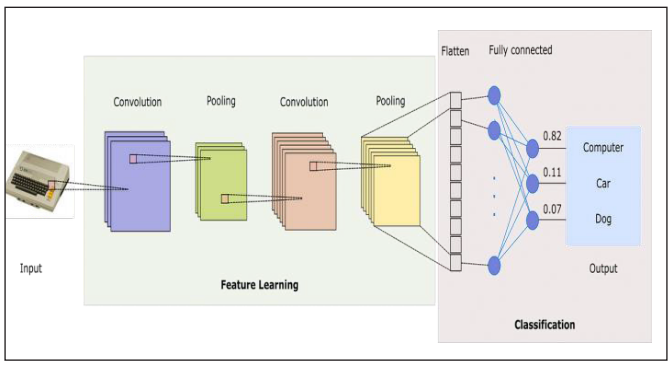


Fig. 2: Medical Imaging

CNNs can be trained on a variety of medical imagery, including radiology, pathology, dermatology, and ophthalmology.

Information flows left to right. CNNs take input images and sequentially transform them, using simple operations such as convolutional, pooling, and fully connected layers, into flattened vectors. The elements of the output vector (softmax layer) represent the probabilities of the presence of disease. Typically, lower layers (left) learn simple image features-edges and basic shapes-which influence the high-level representations (right). Prediction tasks include both classification of the images (i.e., cancerous versus benign) as well as localization of medical features such as tumors.

B. Accuracy and Speed Improvements Compared to Traditional Diagnostic Methods

One of the primary advantages of AI and ML in medical imaging diagnostics is their ability to significantly improve both accuracy and speed compared to traditional methods. While human interpretation of medical images can be subjective and prone to variability, AI algorithms offer consistent and objective analyses, leading to more reliable diagnoses.

Moreover, AI-driven diagnostic tools can process large volumes of imaging data in a fraction of the time required by human radiologists, accelerating the diagnostic process and enabling faster treatment decisions. Studies have shown that AI algorithms can achieve diagnostic performance comparable to or even surpassing that of experienced radiologists in certain tasks, such as detecting pulmonary nodules on chest CT scans or identifying diabetic computer vision retinopathy from fundus photographs [1].

C. Case Studies and Examples Showcasing Successful Implementation

Numerous case studies and examples illustrate the successful implementation of AI and ML in medical imaging diagnostics across various specialties. For instance, researchers at Stanford University developed an AI algorithm capable of detecting skin cancer from dermoscopic images with a level of accuracy comparable to expert dermatologists [6]. Similarly, Google Health's DeepMind team demonstrated the effectiveness of AI in analyzing retinal images for diabetic retinopathy screening, achieving performance on par with board-certified ophthalmologists [7].

Furthermore, AI-powered diagnostic solutions have been integrated into clinical practice worldwide, improving patient care and outcomes. For example,

Aidoc, an AI radiology company, offers a suite of algorithms for detecting critical findings in head CT scans, such as intracranial hemorrhages and fractures, assisting radiologists in prioritizing urgent cases and reducing turnaround times [8].

AI and ML have transformed medical imaging diagnostics by providing accurate, efficient, and scalable solutions for analyzing complex medical images. The integration of AI-driven diagnostic tools into clinical workflows holds immense promise for improving patient care, enhancing diagnostic accuracy, and optimizing healthcare delivery.

III. Personalized Treatment

In the realm of healthcare, the concept of personalized treatment has gained prominence as healthcare providers seek to tailor interventions to individual patient characteristics and needs.

Artificial Intelligence (AI) algorithms play a crucial role in this endeavor by analyzing vast amounts of patient data to develop personalized treatment plans. This section explores how AI facilitates personalized treatment through the utilization of patient data, consideration of various factors, and its application in oncology for targeted therapies [4].

A. Utilization of AI Algorithms to Analyze Patient Data for Personalized Treatment Plans

AI algorithms excel in analyzing diverse patient data, including electronic health records (EHRs), genomic information, lifestyle factors, and treatment history, to develop personalized treatment plans. By leveraging machine learning techniques, these algorithms can uncover hidden patterns and correlations within the data, enabling healthcare providers to make informed decisions about the most effective treatment strategies for individual patients [9].

B. Consideration of Genetic Makeup, Lifestyle Factors, and Medical History

Personalized treatment approaches take into account a wide range of patient-specific factors, including genetic makeup, lifestyle choices, and medical history. AI algorithms integrate this multifaceted information to identify optimal treatment options tailored to each patient's unique characteristics and circumstances. For example, in cancer treatment, AI algorithms can analyze genomic data to identify specific mutations driving tumor growth and recommend targeted therapies that are most likely to be effective [9].

C. Application of AI in Oncology for Targeted Therapies and Adaptive Treatment Recommendations

Oncology represents a prime example of how AI is transforming personalized treatment. AI-driven platforms analyze genomic profiles, tumor characteristics, and treatment responses to identify the most suitable targeted therapies for cancer patients. These targeted treatments are designed to specifically inhibit the growth of cancer cells while minimizing damage to healthy tissues, thereby improving treatment efficacy and reducing adverse effects.

Furthermore, AI enables the development of adaptive treatment recommendations that evolve over time based on patient outcomes. By continuously learning from real-world data, AI systems can refine treatment strategies and adapt to changes in disease dynamics, ensuring that patients receive the most effective and personalized care throughout their treatment journey.

D. Continuous Learning and Adaptation of AI Systems Based on Patient Outcomes

A key advantage of AI-driven personalized treatment approaches is their ability to continuously learn and adapt based on patient outcomes. By analyzing treatment responses and clinical outcomes, AI algorithms can refine their predictive models and treatment recommendations, improving their accuracy and efficacy over time. This iterative learning process enables healthcare providers to optimize patient care and achieve better outcomes.

AI is revolutionizing personalized treatment by

harnessing the power of data analytics to develop tailored interventions that consider individual patient characteristics, genetic makeup, and treatment responses. By leveraging AI algorithms, healthcare providers can optimize treatment strategies, improve patient outcomes, and deliver more effective and personalized care [11].

IV. Remote Monitoring and Predictive Analytics

In the rapidly evolving landscape of healthcare, Remote Monitoring and Predictive Analytics have emerged as powerful tools empowered by Artificial Intelligence (AI) and advanced data analytics. This section delves into how wearable devices and sensors are utilized for continuous data collection, the real-time insights they provide into patients' health status, and the application of AI-driven predictive analytics for forecasting disease progression and identifying high-risk populations.

A. AI-Driven Predictive Analytics for Forecasting Disease Progression and Identifying High-Risk Populations

Predictive analytics powered by AI play a crucial role in forecasting disease progression and identifying high-risk populations who may benefit from early intervention or targeted preventive measures. By analyzing large datasets encompassing clinical, genetic, environmental, and behavioral factors, AI algorithms can identify patterns and trends that may precede the onset or exacerbation of certain diseases.

For example, AI-driven predictive models can forecast the progression of chronic diseases such as diabetes, hypertension, or heart failure based on longitudinal patient data. By identifying individuals at higher risk of complications or hospitalization, healthcare providers can implement proactive interventions, such as lifestyle modifications, medication adjustments, or care management programs, to mitigate risks and improve outcomes. (4)

B. Implementation of Preventive Measures and Targeted Interventions Based on Predictive Insights

One of the most significant benefits of AI-driven predictive analytics is the ability to inform the implementation of preventive measures and targeted interventions tailored to individual patient needs. By identifying modifiable risk factors and predicting adverse health outcomes, AI algorithms empower healthcare providers to develop personalized care plans aimed at reducing disease burden and improving population health.

For instance, AI-driven risk stratification models can identify individuals at elevated risk of developing cardiovascular disease and recommend targeted interventions, such as smoking cessation programs, dietary modifications, or exercise regimens, to mitigate risks and prevent future cardiovascular events. Similarly, predictive analytics can guide the allocation of healthcare resources and interventions to populations most in need, optimizing the efficiency and effectiveness of healthcare delivery.

Remote Monitoring and Predictive Analytics powered by AI and wearable technology hold tremendous promise for transforming healthcare delivery by providing real-time insights into patients' health status, forecasting disease progression, and enabling targeted preventive interventions. By leveraging these innovative tools, healthcare providers can improve patient outcomes, enhance the quality of care, and promote population health [4].

C. Use of Wearable Devices and Sensors for Continuous Data Collection

Wearable devices and sensors have revolutionized the way healthcare is delivered by enabling continuous monitoring of patients' vital signs, activities, and other health-related metrics in real-time. These devices, ranging from smart watches and fitness trackers to biosensors and implantable devices, collect a wealth of data, including heart rate, blood pressure, glucose levels, sleep patterns, and physical activity.

The integration of AI algorithms with wearable devices and sensors enables the processing and analysis of this streaming data, providing valuable insights into patients' health status and behavior patterns. For example, AI-powered algorithms can detect anomalies in heart rate variability that may indicate the onset of cardiac arrhythmias or identify changes in sleep patterns suggestive of underlying sleep disorders [11] [12].

D. Real-Time Insights into Patients' Health Status

The continuous monitoring capabilities of wearable devices and sensors, coupled with AI-driven analytics, offer healthcare providers real-time insights into patients' health status, allowing for timely interventions and proactive management of chronic conditions. By tracking fluctuations in physiological parameters and identifying deviations from baseline patterns, healthcare professionals can detect potential health issues early and intervene before they escalate into serious complications.

For instance, AI algorithms can analyze data from wearable ECG monitors to detect irregular heart rhythms, such as atrial fibrillation, and alert patients and healthcare providers to seek further evaluation. Similarly, continuous glucose monitoring systems combined with AI analytics can provide real-time feedback on blood sugar levels, helping individuals with diabetes manage their condition more effectively [12].

V. Challenges and Considerations

As the healthcare industry continues to embrace Artificial Intelligence (AI) and Machine Learning (ML) technologies for diagnostics, treatment, and patient care, several challenges and considerations arise that must be addressed to ensure the responsible and ethical use of these powerful tools. This section examines key challenges, including privacy and security concerns, algorithm bias, the need for collaboration between stakeholders, and ensuring validation processes, algorithmic fairness, and patient safety standards [8].

A. Privacy and Security Concerns Regarding Patient Data

One of the primary concerns associated with the adoption of AI and ML in healthcare is the protection of patient privacy and security of sensitive medical data. Healthcare organizations must adhere to strict regulations such as the Health Insurance Portability and Accountability Act (HIPAA) in the United States and the General Data Protection Regulation (GDPR) in the European Union to safeguard patient confidentiality and ensure the secure handling of personal health information.

However, the integration of AI and ML often involves the collection, storage, and analysis of vast amounts of patient data, raising concerns about data breaches, unauthorized access, and potential misuse of sensitive information. Robust security measures, encryption techniques, and access controls must be implemented to mitigate these risks and uphold patient privacy rights [10] [11].

B. Algorithm Bias and the Need for Transparent and Interpretable AI Models

Algorithm bias is another significant challenge in the deployment of AI and ML in healthcare, as biased algorithms may lead to inaccurate diagnoses, unequal treatment, and disparities in patient outcomes. AI models trained on biased or incomplete datasets may perpetuate existing biases and disparities present in healthcare data, potentially exacerbating inequalities in access to care and treatment outcomes.

To address algorithm bias, there is a growing need for transparent and interpretable AI models that can be audited, validated, and understood by clinicians and patients. Explainable AI techniques, such as model interpretability frameworks and algorithm transparency tools, can help healthcare providers identify and mitigate bias in AI-driven decision-

making processes, ensuring fairness and equity in healthcare delivery [11].

C. Collaboration between Healthcare Providers, Data Scientists, and Regulatory Bodies

Effective collaboration between healthcare providers, data scientists, and regulatory bodies is essential for the successful integration of AI and ML in healthcare. Healthcare providers bring domain expertise, clinical insights, and real-world experience to AI projects, while data scientists contribute technical expertise in AI algorithms, data analytics, and machine learning techniques.

Regulatory bodies play a crucial role in establishing guidelines, standards, and regulations governing the development, deployment, and use of AI in healthcare. Collaboration between stakeholders is essential to ensure compliance with regulatory requirements, address ethical and legal concerns, and foster innovation while safeguarding patient safety and well-being [10] [11].

D. Validation Processes, Algorithmic Fairness, and Patient Safety Standards

Ensuring the reliability, accuracy, and safety of AI-driven diagnostic and treatment systems requires robust validation processes, algorithmic fairness assessments, and adherence to patient safety standards. AI algorithms must undergo rigorous testing, validation, and clinical evaluation to demonstrate their effectiveness, reliability, and safety before being deployed in clinical practice.

Moreover, algorithmic fairness considerations are critical to prevent biases and disparities in healthcare outcomes. Fairness metrics, bias detection algorithms, and fairness-aware training techniques can help identify and mitigate algorithmic biases, ensuring equitable treatment for all patient populations.

Additionally, adherence to patient safety standards and regulatory requirements is paramount to minimize risks and ensure the ethical and responsible use of AI in healthcare. Continuous monitoring, feedback mechanisms, and post-market surveillance are essential to identify and address any adverse events or unintended consequences associated with AI-driven healthcare technologies.

Addressing the challenges and considerations associated with the integration of AI and ML in healthcare requires a multidisciplinary approach, collaboration between stakeholders, and a commitment to upholding patient privacy, transparency, fairness, and safety standards. By addressing these challenges, healthcare organizations can harness the transformative potential of AI and ML to improve patient care, enhance clinical outcomes, and advance the delivery of healthcare services [11].

VI. Conclusion

The integration of Artificial Intelligence (AI) and Machine Learning (ML) in healthcare holds tremendous promise for revolutionizing patient care, diagnostics, and treatment methodologies. Throughout this article, we have explored the transformative potential of AI and ML in healthcare, highlighting their role in enhancing diagnostics, enabling personalized treatment, and facilitating remote monitoring and predictive analytics. However, as we embrace these innovative technologies, it is essential to address various technical, ethical, and regulatory challenges to ensure their responsible and ethical use.

- *Recap of the Transformative Potential of AI and ML in Healthcare:* AI and ML technologies have demonstrated remarkable capabilities in analyzing medical data, identifying patterns, and making predictions with unprecedented accuracy and efficiency. From analyzing medical images and developing personalized treatment plans to enabling remote monitoring and predictive analytics, AI and ML are reshaping the landscape of healthcare delivery. These technologies have the potential to improve patient outcomes, enhance clinical workflows, and optimize healthcare resources, leading to more efficient and effective healthcare systems.
- *Emphasis on Addressing Technical, Ethical, and Regulatory Challenges:* As we harness

the transformative power of AI and ML in healthcare, it is crucial to address various technical, ethical, and regulatory challenges to ensure their responsible and ethical use. Privacy and security concerns regarding patient data, algorithm bias, transparency, and interpretability of AI models, collaboration between stakeholders, and adherence to regulatory requirements are among the key challenges that must be addressed. By fostering collaboration between healthcare providers, data scientists, and regulatory bodies, and implementing robust validation processes and patient safety standards, we can mitigate risks and maximize the benefits of AI and ML in healthcare.

- *Call for Patient-Centric Care and Equitable Access to Innovative Technologies:* At the heart of the integration of AI and ML in healthcare is the goal of delivering patient-centric care and ensuring equitable access to innovative technologies. As we embrace AI-driven healthcare solutions, it is essential to prioritize patient needs, preferences, and well-being, and empower individuals to actively participate in their care journey. Moreover, efforts should be made to ensure equitable access to innovative technologies, particularly among underserved populations and marginalized communities. By promoting patient-centric care and equitable access to healthcare innovations, we can harness the full potential of AI and ML to improve health outcomes and advance the delivery of healthcare services for all.

In conclusion, the integration of AI and ML in healthcare represents a transformative shift in how we approach diagnostics, treatment, and patient care. By addressing technical, ethical, and regulatory challenges and prioritizing patient-centric care and equitable access to innovative technologies, we can unlock the full potential of AI and ML to revolutionize healthcare delivery and improve the lives of patients worldwide.

References

[1] A. Esteva, A. Robicquet, B. Ramsundar, V. Kuleshov, M. DePristo, K. Chou, C. Cui, G. Corrado, S. Thrun, and J. Dean, "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1, pp. 24-29, 2019.

[2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature,* vol. 521, pp. 436-444, 2015.

[3] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, ... and L. Fei-Fei, "Imagenet large scale visual recognition challenge," *Int. J. Compute. Vis.*, vol. 115, pp. 211-252, 2015.

[4] A. Rajkomar, J. Dean, and I. Kohane, "Machine learning in medicine," *New England Journal of Medicine,* vol. 380, no. 14, pp. 1347-1358, 2019.

[5] E. J. Topol, "High-performance medicine: The convergence of human and artificial intelligence," *Nature Medicine*, vol. 25, no. 1, pp. 44-56, 2019.

[6] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115-118, 2017.

[7] V. Gulshan, L. Peng, M. Coram, P. C. Nelson, J. L. Mega, ... and D. R. Webster, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, pp. 2402-2410, 2016.

[8] L. M. Prevedello, B. S. Erdal, J. L. Ryu, K. J. Little, M. Demirer, S. Qian, and R. D. White, "Automated critical test findings identification and online notification system using artificial intelligence in imaging," *Radiology: Artificial Intelligence*, vol. 1, no. 4, p. e180014, 2019.

[9] R. Miotto, F. Wang, S. Wang, X. Jiang, and J. T. Dudley, "Deep learning for healthcare: Review, opportunities, and challenges," *Briefings in Bioinformatics*, vol. 19, no. 6, pp. 1236-1246, 2017.

[10] D. S. Char, N. H. Shah, D. Magnus, and K. D. Mandl, "Implementing machine learning in health care-addressing ethical challenges," *New England Journal of Medicine*, vol. 378, no. 11, pp. 981-983, 2018.

[11] Z. Obermeyer, B. Powers, C. Vogeli, and S. Mullainathan, "Dissecting racial bias in an algorithm used to manage the health of populations," *Science*, vol. 366, no. 6464, pp. 447-453, 2019.

[12] A. Gupta, and B. Sharma, "Artificial intelligence in wearable devices: A comprehensive review," *International Journal of Computational Intelligence Systems*, vol. 13, no. 1, pp. 1836-1847, 2020.

International Journal of Knowledge Based Computer Systems
12 (1) 2024, 16-26
http://www.publishingindia.com//ijkbcs/

Early Detection of Diabetic Retinopathy by using Deep Learning

Pranjal Panchal[1], Kumari Lalini[2], Khushboo Rathore[3] and Amita Jain[4*]
[1]Scholar, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India.
Email: pranjalpanchal2854@gmail.com
[2]Scholar, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India.
Email: kumarilalini4385@gmail.com
[3]Scholar, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India.
Email: khushboor245@gmail.com
[4]Assistant Professor, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India.
Email: dr.amita@piemr.edu.in
*Corresponding Author

Abstract: The eye disease diabetic retinopathy (DR) can cause blindness if not treated. Worldwide, 2.6 million persons lost their sight or had severe vision impairments in 2015 as a result of diabetic retinopathy. By 2020, experts predict that number will have jumped to 3.2 million. In high-income nations, diabetic retinopathy should be less common, but in low- and middle-income countries, finding and treating the condition early should be a top concern. Thanks to recent developments in deep learning, researchers have demonstrated that automated diabetic retinopathy screening and grading is a practical method to reduce manpower requirements. The Kaggle dataset was used to find the winners of the diabetic retinopathy detection contest. In the suggested DL-DRDC methodology, the feature vectors from the previously processed retinal fundus images were retrieved using the modified Efficient Net method. The DL-DRDC Mechanism extracts characteristics from previously studied retinal fundus pictures using a modified version of Efficient Net. The proposed GA evolution selects the best-fitting model from a model population pool in order to optimise deep CNN models. The research and the experimental findings show that the recommended approach operates more efficiently than the other basic models. This research offered a novel and better approach for the diagnosis and severity rating of diabetic retinopathy.

Keywords: Convolutional Neural Network (CNN), Deep learning, Diabetic retinopathy, Model, Patient.

I. Introduction

According to the National Diabetic Registry (NDR), over 2.6% of Malaysians over the age of 18 had diabetes in 2012. Life expectancy may be reduced if problems occur, and the disease needs continual monitoring. At this time, diabetes also has no known cure. Having a basic understanding of the ailment, its effects on the body, and successful management strategies is, therefore, crucial. The title claims that a deep learning neural network and fundus images were used to identify diabetes. The fundus picture was

obtained using the Topcon TRC NW6 non-mydriatic retinography, which features a color video 3CCD camera and a 45-degree field of view. In a nutshell, a fundus picture shows what's within a human eye. So, this study demonstrates how to identify diabetic eye illness using fundus photos. Variations in visual clarity, blurred or distorted vision, eye discomfort, cataracts, delayed recovery from eye injuries, and eventually blindness are among complications that diabetics may face. Furthermore, blindness is one of the most terrifying consequences of diabetes for the eyes. The most fundamental definition of diabetic retinopathy is damage to the retina, an eye ailment. The retina, situated in the back of the eye, is the main culprit when it comes to light sensitivity. Because of this, damage to it could cause blindness. Type 2 and type 1 of the illness were both discovered. However, the most common form of diabetes is type 2, which affects an estimated 90% to 95% of people.

In 2010, diabetic retinopathy (DR) caused the blindness and visual impairment of 3.8 million people worldwide. Experts project that 191.0 million individuals will develop DR by 2030 due to the increasing incidence of diabetes. Despite the fact that the worldwide incidence of any DR was 27.0% from 2015 to 2019, the early phases of DR, including referable DR, do not exhibit any noticeable symptoms. By detecting and treating DR early, the chance of vision loss may be reduced by about 57%. This is because DR can progress to a quite advanced stage before it impacts vision. It is extremely important for individuals with diabetes, especially those in the middle age and older age groups, to attend screenings and follow-up sessions regularly. Many diabetic people don't have their required yearly eye exams because they're too painful, don't notice any problems, or can't get to a retina specialist quickly enough, according to a number of studies [1].

Byproducts of glucose metabolism, including abnormally high blood sugar levels, build up in blood vessels. After a patient has had diabetes for over a decade, diabetic retinopathy (DR) sets in. High blood pressure is the root cause of diabetic retinopathy (DR), which in turn damages the retina and its vascularization, potentially leading to blindness or even death. Only via laborious and costly fundoscopy procedures can ophthalmologists detect retinal vascular edema. Diabetes is a major cause of blindness, and there will be an estimated 552 million people living with the disease by 2030 [2].

The most effective way to curb vision loss is to catch it in its early stages and treat it accordingly. Abnormal development of blood vessels and eventual blindness may occur in extreme instances when the vessels enlarge, leak fluid, or obstruct blood arteries. Retinal DR is characterized by micro aneurysms, hemorrhages, and exudates. The degree to which a lesion is noticeable depends on its size, shape, and general look. A screening tool for DR in the field of ophthalmology is fundus photography. An automated evaluation approach may effectively and affordably prevent diabetes-related blindness.

Ophthalmologists use visual evaluations and examinations of the eyes to determine the existence and severity of DR. This is a laborious and costly procedure for many people throughout the world who have diabetes. With statistics among qualified ophthalmologists differing greatly, DR severity and early disease identification continue to be challenges. In addition, there is a lack of adequate ophthalmologists and detection equipment in undeveloped nations, where 75% of DR patients reside. There are worldwide screening programmes in place to combat the increase of avoidable eye disorders, but the prevalence of DR is too high for effective individual detection and treatment.

Damage to the retina, known as DR, may happen as a result of hypertension. It may harm the retina's blood vessels, leading to potential blindness or death. Only via laborious and costly fundoscopy procedures can ophthalmologists detect retinal vascular edema. It is necessary to find DR automatically by looking at retinal fundus pictures. It has been shown that deep learning models can detect DR more accurately than ophthalmologists. They become a practical choice because of this. One of the most used deep learning models, Convolutional Neural Networks (CNNs) are utilized often in medical image identification, prediction, and classification. The objective of this study is to construct a CNN model with an updated activation function that can detect DR automatically.

This new activation function is compared against earlier activation functions using the open-source datasets DIARETDB0, DRIVE, CHASE, and Kaggle. By including a novel activation function, the present CNN version has been enhanced and now yields remarkable outcomes [3].

II. Literature Review

S. Gothane *et al.* (2022) When insulin is insufficient, blood glucose levels rise, a symptom of diabetes. A diabetic's eyes, heart, nerves, and kidneys are all impacted. Among the complications, diabetic retinopathy stands out. Compared to human analysis, automated approaches for Diabetic Retinopathy detection are more competent, cost less, and provide greater flexibility. Diagnose medical conditions with the help of computers using the Deep Learning approach. The goal of this project is to develop an automated method for detecting diabetic retinopathy in its early stages. Artificial intelligence and deep learning could help doctors predict when a patient will go blind. Using a supervised learning approach, this project is now classifying fundus photos. Micro aneurysms, hemorrhages, exudates, and swollen blood vessels are some of the key indicators of diabetic retinopathy that may be improved upon using a variety of image processing techniques and filters for this purpose. Neural networks are then employed for classification purposes. Using ResNet architecture, it successfully categorizes fundus pictures 82% of the time [4].

K. Oh *et al.* (2021), Worldwide, an estimated 3.2 million individuals will be blind or severely visually impaired by 2020 as a result of diabetic retinopathy, up from 2.6 million in 2015. Low- and middle-income nations must prioritize early identification and treatment of diabetic retinopathy, even though high-income nations should see a decline in its incidence. Automatic screening and grading of diabetic retinopathy is efficient, saving workers and time, according to researchers. This was made possible by the current breakthroughs in deep learning technology. Even though ultra-wide-field fundus photos may capture as much as 82% of the retinal surface, conventional fundus photos are still used by the majority of automated systems. Our method for diabetic retinopathy detection employs deep learning in conjunction with ultra-wide-field fundus imaging. Our experimental results show that, when comparing early treatment diabetic retinopathy studies, a 7-standard field picture produced by ultra-wide-field fundus photography is statistically better than an optic disc and macula centered image [5].

F. Nawaz *et al.* (2021) One of the retinal diseases that may lead to permanent blindness is Diabetic Retinopathy (DR). It is challenging to diagnose DR in its early stages since there are no symptoms at the beginning level. When blood sugar levels are too high, DR sets in. Timely identification and continuous treatment are necessary to avoid blindness. The ophthalmologist's ability to assess their patients' problems might be substantially enhanced with the use of AI-powered automated detection. Using DL and RL with ML, this study seeks to build a system that can autonomously identify and rank diabetic retinopathy. A method used in artificial intelligence known as transfer learning is utilized to train the Inception-v4 deep neural network. Two distinct transfer learning configurations—fixed feature extractor mode and fine-tune mode—are employed during Inception-v4 training. Although the accuracy rates achieved by both setup strategies are fair, fine-tuning outperforms fixed feature extractor. Using fine-tune configuration mode, the team achieved 97.7 percent accuracy in disease grading and 96.6% accuracy in early DR detection, surpassing the state-of-the-art methods in the corresponding literature [6].

Hasan *et al.* (2021) A chronic condition, diabetes mellitus is rapidly becoming a global epidemic. The condition manifests as problems with the eyes, kidneys, and heart caused by an elevation in blood glucose levels. When blood vessels in the retina burst, a condition called diabetic retinopathy (DR) develops. This is an eye consequence of diabetes. Since it does not present any obvious symptoms when first developing, it is thought to be the main cause

of blindness. Early identification and classification of DR patients is crucial for the provision of needed medical care. Recent advances in machine learning's algorithms have made it a powerful tool for use in computer-assisted diagnosis and other medical applications. In this article, the team aims to analyze the performance of DR detection and classification systems that utilize various machine learning algorithms. For the purpose of training and evaluating these algorithms, several publically available datasets including massive volumes of retina fundus and thermal photos are utilized. These algorithms proved they could spot warning signs and quantify the severity of DR. Looking at the systems the teamanalyzed, ResNet50, an algorithm for deep convolutional neural networks, appears to have achieved the greatest performance. Using a cascade of feature extraction kernels, the Resnet50 can process retinal images and derive wealth data. Our findings suggest that ML algorithms might be useful in assisting medical professionals in DR patients with accurate diagnosis and appropriate treatment [7].

D. Das and S. Biswas (2020) One medical complication of diabetes mellitus is diabetic retinopathy (DR). Because it mutilates the human retina, it causes serious blindness. Statistics show that 80% of the population, disproportionately those of working age, has been afflicted by this illness. Therefore, DR has recently emerged as a critical problem that requires prompt resolution in order to significantly reduce the prevalence of blindness among working-age individuals. Due to the time and inaccuracy associated with human diagnosis, many sophisticated technologies have been developed to identify DR early on. Furthermore, ophthalmologists cannot be made available 24/7 at every location. Therefore, a computer-assisted intelligent system that is highly optimized is needed for the early diagnosis of DR. For decades, scientists all throughout the world have put forth different models. This paper's goal is to provide a more in-depth analysis of previous efforts and new technologies related to DR detection. Herein, the team provides the current level of knowledge about the characteristics, etiology, symptoms, grades, and models for early detection of DR [8].

III. Research Methodology

A. Genetic Algorithm for Deep Learning Model Selection in DR Diagnosis

It takes learning millions of parameters to train the CNN model from the ground up for a specific issue. Excessive computing power and resource utilization are necessary for this. Therefore, less computation is needed when transfer learning is used. There is a plethora of CNNs that have already been trained for the categorization job. Nonetheless, they have been designed to accommodate a wide range of datasets and applications. This makes choosing the optimal model for DR severity categorization a challenging task. Consequently, the team provides an automated process to choose the optimal pre-trained model for DR severity categorization. What sets GA apart from previous efforts is its application to the problem of automatically selecting the optimal model from among many pre-existing CNN models. The team suggests using GA to automatically choose the CNN network topology from the pretrained model. The first set of data points used to train GA are the pre-trained CNN architectures. A subset of the population is chosen at random from the population to make up the GA model's population. Additional training for deep feature extraction is applied to the chosen people. A significant disadvantage of using GA is the enormous computational expense it entails. The LSTM classifier is used to categorise the deep characteristics that have been retrieved in order to lower it. Ranking the individual models within the current population is a good indicator of the reliability of the testing data. The final ultimate model emerges by several iterations of genetic functions, including selection, crossover, and mutation. In order to aid in the early detection of DR, this best-fit model will include features with deep representations. Fig. 1 depicts the planned work's structural design. Python with TensorFlow, a library built on top of Keras, was

used to create the system that was presented. A server with a TeslaV100 card ran the algorithm [9].

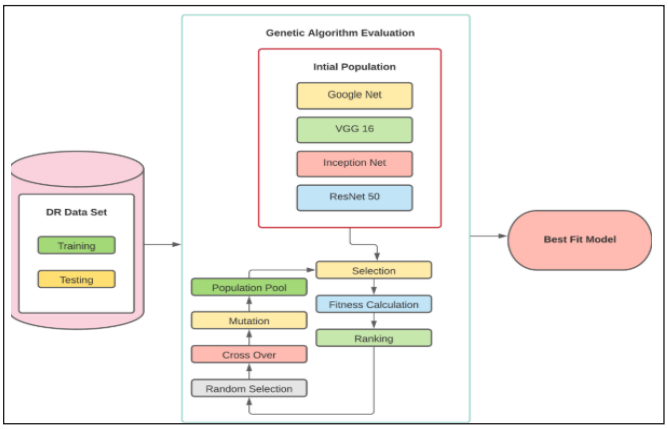


Fig. 1: Selection of Models using GA Evolution

- *Random Selection:* In order to incorporate genetic diversity, a small number of models with poor fitness scores are also included in the population, even though only the fittest models are preserved for the next generation. As a result, the global maxima are more easily reached by the optimum algorithm. To make sure of it, models with poor fitness are chosen at random using a 0.1 threshold. The number [0,1] is randomly allocated to each new generation of progeny.
- *Mutation:* A tiny fraction of each person's DNA may be changed at random to increase genetic variation. This haphazard alteration in a person's DNA is called mutation. You may randomly alter the bit position of a single gene at one of four locations. A mutation with a chance of 0.3 may only occur in one specific location in an individual's genes.
- *Crossover:* If there is still room for additional people in the population after retention, the process of crossover is used to fill it up. In order for crossover to take place, two separate models are chosen at random to serve as parents. The crossover point for the parent models is the midpoint of their length. Any person may take part in the crossover; there are no limits.

B. Proposed Methodology

A new "deep learning"-assisted DR detection and classification model is the Deep Learning Diabetic Retinopathy Detection and Classification (DL-DRDC) technique. Using retinal fundus photos, this DL-DRDC approach identifies and categorizes the different degrees of DR. The DL-DRDC system employs the CLAHE approach for its pre-processing stage. The most up-to-date DL-DRDC technique for DR detection using retinal fundus images is provided. You may see the whole DL-DRDC system's operation plan in the provided picture. The DL-DRDC approach is comprised of three separate steps. First, there is pre-processing. Then, there is feature extraction. And last, there is classification. The first step was to enhance the contrast levels of retinal fundus photographs using the CLAHE model. In the second stage, features were extracted from the pre-processed pictures using the modified Efficient Net approach. In the end, the processed retinal fundus pictures were appropriately classified using the Deep Neural Network procedure [10].

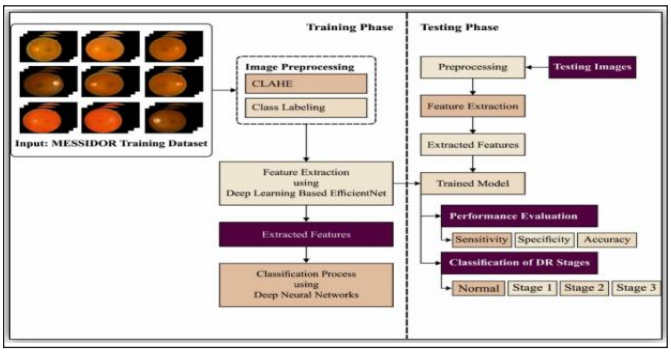


Fig. 2: A Comprehensive Overview of the Deep Learning Diabetic Retinopathy Detection and Classification System (DL-DRDC)

(i) CLAHE Method for Pre-Processing Images

To enhance their quality, the retinal fundus pictures undergo preprocessing. Reducing noise, shrinking the picture, and correcting the colors are common phases in this process. Statistical information about

the picture is primarily what the histogram is intended to provide. In order to identify diabetic retinopathy, CLAHE is an important part of the preprocessing workflow. Retinal fundus pictures may be fine-tuned for more precise analysis and diagnosis by increasing contrast, which makes important characteristics more visible. In order to identify diabetic retinopathy, the preprocessing pipeline relies on CLAHE. By enhancing contrast, it makes key features in retinal fundus photos more visible, leading to a more precise evaluation and diagnosis of the disease.

(ii) Extracting Features using the Rationalized Efficient Net Method

Feature vectors are generated by passing the pre-processed, contrast-enhanced picture through the several levels of the Efficient Net model. A number of computer vision applications rely on feature extraction, such as segmentation, object identification, and picture categorization. In order for deep learning algorithms to make predictions or complete tasks, it is necessary to transform raw input data, such images, into a collection of representative features. One state-of-the-art framework for computer vision feature extraction is Efficient Net [11].

IV. Data Analysis

A. Selection of Neural Networks for Diabetic Retinopathy Diagnosis Based on Genetic Data

In Fig. 3, the confusion matrix that algorithm 1 produced on the DR dataset is shown. The confusion matrix illustrates exactly where the model is losing its bearings. This aids in the model's further fine-tuning. A deep learning system's TP, TN, FP, and FN may be discovered by creating an error or confusion matrix. The deep learning model's perplexion is well illustrated by the confusion matrix.

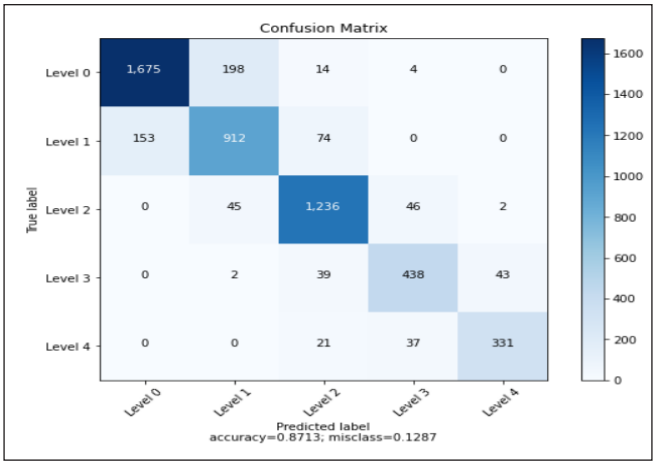


Fig. 3: A Model Selection Confusion Matrix Based on GA

Algorithm 1: The Use of Genetic Information to Select Neural Networks for DR Diagnosis

Input: Labeled fundus image dataset

Output: Optimized Neural Network for DR diagnosis

Initialization: Population, P = {DenseNet, ResNet50, VGG16, GoogleNet}

With_hold=0.3; Mutate = 0.3; Select_Random = 0.1

Steps:

i. For every, i ε P do
   a. Evaluate fitness function, F(i)
   b. Grade, G[i] = F(i)
ii. Arrange P with respect to non-increasing order of G[i]
iii. For x ε (0, with_hold* G.size - 1) do
   a. Parent.append(G[x])
iv. For every x in P calculate a random number and do
   a. If random(x)>Select_Random then
      i) Parent.append(G[x])
v. For every x in Parent calculate a random number and do |
   a. If random(x)>Mutate then

i) Flip x at position, Randint(0, x.size-1)
ii) Parent.append(G[x])

vi. size P.size- Parent.size

vii. If offsprings.size<size then

a. Select two parents, Pl and P2 from Parent at random and do

b. If P1 ≠ P2 then

i) offspring (P1[0, P1.size/2-1] + P2 [P1. size/2, Pl.size-1)
ii) offsprings.append(offspring)
iii) goto step vii

viii. Parent.append(offspring)

ix. Return parent

In the confusion matrix, you can find the True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) calculations. Table I and Fig. 4, illustrate how well the Best-Fit model developed by GA Evolution performed when it came to early DR diagnosis. The model had an error rate of 11.35% for the train and 12.83% for the test [12].

TABLE I: EVALUATION OF GA EVOLUTION'S BEST-FIT MODEL

Class	Stage 0	Stage 1	Stage 2	Stage 3	Stage 4
Accuracy	0.92	0.90	0.94	0.96	0.97
Precision	0.88	0.7	0.92	0.83	0.84
Recall	0.91	0.78	0.88	0.82	0.87
F1 Score	0.8	0.78	0.90	0.83	0.86

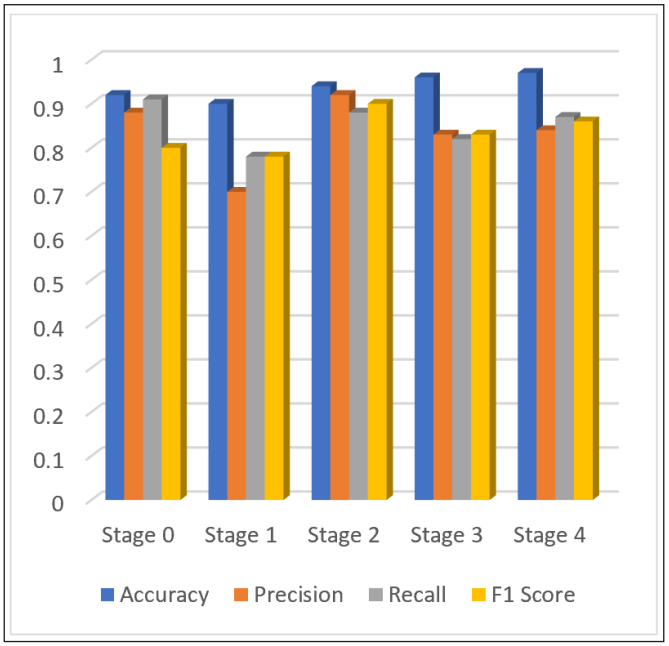


Fig. 4: Evaluation of GA Evolution's Best-Fit Model

The optimal model is chosen from the pool of available models using the suggested GA evolution for optimizing deep CNN models. Based on the analysis and experimental findings, the proposed technique is more efficient than the other basic models [13].

B. Diabetic Retinopathy using Retinal Fundus Images for Deep Learning-Enabled Classification and Detection

Kaggle data sets were used for training, whereas Messidor was used for experimental validation. One hundred and twenty fundus retinal pictures are available for DR classification in the Messidor dataset. It features pictures taken by three different ophthalmology clinics in France. Pupil dilation was not used in any of the photos. Following pupil dilation, 800 photos are captured, whereas 400 images are captured without. There are one hundred and twenty retinal fundus photos in this collection, arranged in the correct sequence of healthy, DR stages 0-4, and so on. A number of functionality measures are used to compute the results.

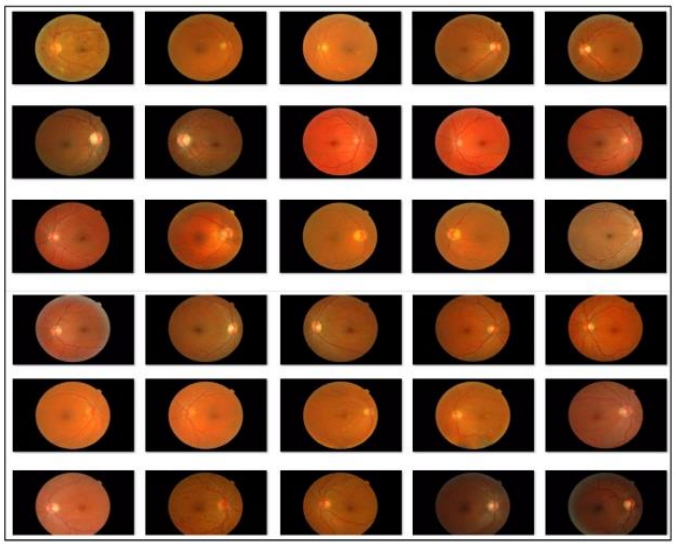


Fig. 5: Instance images

Fig. 5 shows the results of running the DL-DRDC technique on the data resource that was presented, and the resulting confusion matrices. Several DR phases monitored in separate runs have been described using the DL-DRDC approach. In terms of execution, the DL-DRDC model has classified 543 instances for the general stage, 150 instances for phase-1, 245 instances for phase-2, and 251 instances for phase-3. Table II illustrate how DL-DRDC performed in each of the experimental runs. Results across a number of parameters have been satisfactory using the DL-DRDC method. The first execution run of the DL-DRDC approach in the normal class achieved a 99% accuracy rate. An average accuracy rate of 99% has been achieved in explaining the DL-DRDC technique in the stage-1 class. There is now a 99% accuracy rate in describing the DL-DRDC technique in the stage-2 class. There has been a 99% success rate in explaining the DL-DRDC algorithm in the stage-3 class at the same time [12].

Table II: The Architecture of Efficient Net B0 Layers

Operator	Resolution	No. of Channels	No of Layers
Conv 3X3	224X224	32	1
MBConv1, k3X3	112 X 112	16	1
MBConv6, k3X3	112 X 112	24	2
MBConv6, k5X5	56 X 56	40	2
MBConv6, k3X3	28 X 28	80	3
MBConv6, k5X5	14 X 14	112	3
MBConv6, k5X5	14 X 14	192	4
MBConv6, k3X3	7 X 7	320	1
Conv1X1 &pooling &FC	7 X 7	1280	1

Table III and Fig. 6 show the average classification results from the analysis of the DL-DRDC method's executions. The DL-DRDC model generated effective DR diagnostic findings, as shown by several runs. In particular, the DLDRDC model has attained a 99% specificity, 98% rising sensitivity, 98% precision, 99% accuracy, and a 98% F-score throughout execution. After that, the DLDRDC method produced an F-score of 99%, sensitivity that increased by 99%, specificity that increased by 99%, precision that increased by 99%, and accuracy that increased by 99%. All of the DLDRDC model's metrics—F-score, accuracy, precision, specificity, and rising sensitivity—reached 99% by run-5.

TABLE III: CLASSIFICATION RESULTS ANALYSIS USING THE DL-DRDC METHOD

Number of Runs	Class	Sensitivity	Specificity	Precision	Accuracy	F1-Score
Run-1	Normal	99.4	99.4	99.4	99.4	99.4
	Stage-1	97	99.7	98.6	99.5	98.3
	Stage-2	99.1	99.3	97.5	99.2	98.3
	Stage-3	98.7	100	100	99.7	99.3
	Average	98.8	99.6	98.8	99.4	98.8
Run-2	Normal	99.7	99.4	99.4	99.5	99.4
	Stage-1	98	100	100	99.8	99
	Stage-2	100	99.6	97	99.5	99
	Stage-3	98.7	100	100	99.7	99.3
	Average	99.0	99.8	99.3	99.7	99.1
Run-3	Normal	99.6	99.7	99.6	99.7	99.5
	Stage-1	98.7	99.8	98.7	99.7	98.7
	Stage-2	99.2	99.6	98.4	99.5	98.8
	Stage-3	99.2	100	100	99.8	99.6
	Average	99.1	99.8	99.2	99.7	99.2
Run-4	Normal	99.4	99.7	99.6	99.6	99.5
	Stage-1	98.5	99.8	98.7	99.6	98.6
	Stage-2	99.5	99.6	98.7	99.6	99.1
	Stage-3	99.2	99.8	99.6	99.8	99.3
	Average	99.2	99.7	99.2	99.7	99.1
Run-5	Normal	99.8	99.6	99.6	99.8	99.6
	Stage-1	99.3	99.8	99.4	99.8	99.3
	Stage-2	99.5	99.7	99.2	99.8	99.3
	Stage-3	99.1	100	100	99.8	99.5
	Average	99.4	99.7	99.5	99.8	99.4

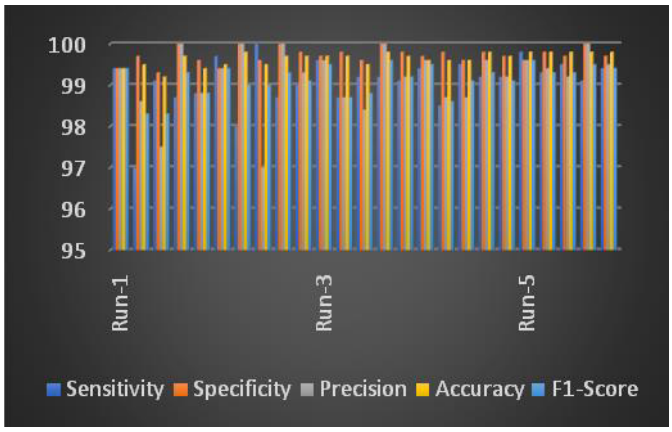


Fig. 6: The Results from Several Runs of the DL-DRDC Model

The majority of deep learning models used for DR detection are convolutional neural networks (CNNs), which sort out pixels that represent blood vessels from those that do not. This study presented a state-of-the-art DL-DRDC method for diabetic retinopathy (DR) severity rating systems. With the help of Efficient Net with DNN for picture categorization, deep learning models were able to provide prompt and accurate diagnoses. Retinal pictures were made more contrasty using the CLAHE method. Using the results of the class probability calculations, DNN chose the class with the best score. A thorough and effective method for identifying diabetic retinopathy, the proposed model included integrated pre-processing, segmentation, feature extraction, and classification phases [14].

V. Conclusion

In summary, a person's eyes are among their most vital organs. People are more likely to get Diabetes Mellitus (DM) due to modern lifestyle choices. DR will occur as a result of the chronic predominance of DM. While a full solution for DR is still in the works, it is now manageable with early diagnosis. Therefore, certain studies are conducted for computer-aided DR diagnosis in order to detect the disease at an earlier stage. Diabetic Mellitus (DM) has been on the rise due to modern lifestyle changes. Diabetes develops when the body either does not generate enough insulin or develops resistance to it. When blood glucose levels rise due to DM, regular blood flow becomes impossible. The kidneys, nerves, and eyes are among the most affected organs. As a result of diabetes being so common over the long term, a condition known as Diabetic Retinopathy (DR) may develop. Using deep learning methods, the DR may be detected early on. Convolutional Neural Network (CNN) design is a labor-intensive process. To choose the Convolutional Neural Network model that has already been trained, an automated Genetic Algorithm (GA) is used. A lot of deep learning models have been drawn by hand. They provide a function, to rephrase. The serious problem of diabetic retinopathy, which can cause blindness and frequently begins to manifest with no outward signs of illness, is the subject of this study. Treatment efficacy and vision preservation depend on prompt diagnosis and precise severity categorization. The four main components of the proposed comprehensive framework—preprocessing, optical disc removal, feature extraction, and classification—utilize the most recent advancements in deep learning.

Acknowledgement

Gratitude is extended to my advisor, for their guidance and support. Thanks also to my colleagues in the [Prestige Institute of Engineering Management & Research, Indore] for their assistance and feedback. Acknowledgment is made to the creators of the open-source tools and datasets used in this project. Thanks again to everyone who has always been there for us, but especially to our family and friends.

References

[1] M. I. Feisul, and S. Azmi, National diabetes registry report. 2009-2012. Ministry of Health Malaysia, 2013.

[2] L. Wu, P. Fernandez-Loaiza, J. Sauma, E. Hernandez-Bogantes, and M. Masis, "Classification of diabetic retinopathy and diabetic macular Edema," *World J. Diabetes*, vol. 4, p. 290, 2013, doi: https://doi.org/10.4239/wjd.v4.i6.290.

[3] A. Gupta, and R. Chhikara, "Diabetic retinopathy: Present and past," *Procedia Computer Science*, vol. 132, pp. 1432-1440, 2018.

[4] S. Gothane, S. K. Raju, N. Bhaskar, and G. Divya, "Diabetic retinopathy detection using deep learning," *Data Engineering and Intelligent Computing*. 2022, pp. 387-393, doi: https://doi.org/10.1007/978-981-19-1559-8_39.

[5] K. Oh, H. M. Kang, D. Leem, and H. Lee, "Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images," *Scientific Reports*, vol. 11, no. 1, 2021, doi: https://doi.org/10.1038/s41598-021-81539-3.

[6] F. Nawaz, Muhd. Ramzan, K. Mehmood, H. Khan, S. Khan, and R. Bhutta, "Early detection of diabetic retinopathy using machine intelligence through deep transfer and representational learning," *Computers, Materials & Continua*, vol. 66, pp. 1631-1645, 2021, doi: https://doi.org/10.32604/cmc.2020.012887.

[7] D. A. Hasan, S. R. M. Zeebaree, Mohd. A. M. Sadeeq, H. Shukur, R. Zebari, and A. Alkhayyat, "Machine learning-based diabetic retinopathy early detection and classification systems - A survey," *2021 1st Babylon International Conference on Information Technology and Science (BICITS),* 2021, pp. 16-21, doi: https://doi.org/10.1109/BICITS51482.2021.9509920.

[8] D. Das, and S. Biswas, "Early detection of diabetic retinopathy using machine learning techniques: A survey on recent trends and techniques," 2020, doi: https://doi.org/10.1007/978-981-15-7031-5_64.

[9] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and Muhd. K. Khan, "Medical image analysis using convolutional neural network: A review," *Journal of Medical Systems*, vol. 42, no. 226, pp. 1-20, 2018.

[10] S. E. Abed, S. A. Al-Roomi, and M. Al-Shayeji, "Effective optic disc detection method based on swarm intelligence techniques and novel preprocessing steps," *Applied Soft Computing*, vol. 49, no. C, pp. 146-163, 2016.

[11] A. Z. Soliman, P. S. Silva, L. P. Aiello, and J. K. Sun, "Ultra-wide field retinal imaging in detection, classification, and management of diabetic retinopathy," *Semin. Ophthalmol.*, vol. 27, pp. 221-227, 2012.

[12] B. Dashtbozorg, A. M. Mendonça, and A. Campilho, "Optic disc segmentation using the sliding band filter," *Computers in Biology and Medicine*, vol. 56, no. C, pp. 1-12, 2015.

[13] D. S. W. Ting et al., "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," *JAMA*, vol. 318, pp. 2211-2223, 2017.

[14] H. Takahashi, H. Tampo, Y. Arai, Y. Inoue, and H. Kawashima, "Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy," *PLS One*, vol. 12, 2017.

International Journal of Knowledge Based Computer Systems
12 (1) 2024, 27-33
http://www.publishingindia.com//ijkbcs/

Smart Traffic Management: Implementing RFID Technology for Real-Time Congestion Solutions

Yakuta Tayyebi[1*], Anuj Banke[2], Anshita Balodiya[3], Gourav Choudhary[4] and Ananya Dey[5]
[1]Assistant Professor, Department of Computer Science, Prestige Institute of Engineering Management and Research, Indore, Madhya Pradesh, India. Email: yakutatayyebi@davvscsit.in
[2]Scholar, Department of Computer Science, Prestige Institute of Engineering Management and Research, Indore, Madhya Pradesh, India. Email: 51110104050@piemr.edu.in
[3]Scholar, Department of Computer Science, Prestige Institute of Engineering Management and Research, Indore, Madhya Pradesh, India. Email: 51110404487@piemr.edu.in
[4]Scholar, Department of Computer Science, Prestige Institute of Engineering Management and Research, Indore, Madhya Pradesh, India. Email: 51110104004@piemr.edu.in
[5]Scholar, Department of Computer Science, Prestige Institute of Engineering Management and Research, Indore, Madhya Pradesh, India. Email: 51110104231@piemr.edu.in
[*]Corresponding Author

Abstract: Traffic congestion is a significant issue in developing countries worldwide, caused by signal failures, poor law enforcement, and inadequate traffic management. Since infrastructure expansion is limited, enhancing traffic management is the only viable solution. Methods like video analysis, infrared sensors, inductive loop detection, and wireless networks are effective for smart traffic management. However, their high installation and maintenance costs, along with lengthy setup times, pose significant challenges. A new technology, Radio Frequency Identification (RFID), integrated with existing signaling systems, offers a real-time smart traffic management solution. RFID is quicker and cheaper to install, reduces congestion, detects bottlenecks early, and saves time and money. This paper presents a framework to use RFID effectively to enhance traffic management. This work investigates the architecture and components of RFID-based traffic management systems, including RFID tags, readers, and data processing units. Additionally, the paper discusses the challenges and limitations of RFID implementation, including privacy concerns, infrastructure costs, and integration with existing traffic management systems.

Keywords: GSM, Radio Frequency Identification (RFID), Traffic congestion.

I. Introduction

Traffic congestion occurs when the volume of vehicles exceeds road capacity, resulting in slower speeds, longer trip times, and increased queuing. This is a significant issue in India's metropolitan cities, where congestion arises from demand outstripping road capacity, a condition known as saturation. Events such as accidents or jams can create ripple effects, leading to traffic jams.

Additionally, security issues caused by anti-social elements can further stagnate traffic. In India, the annual economic loss due to congestion, including fuel wastage, is estimated at Rs 60,000 crores. Congestion also slows down freight vehicles and increases waiting times at checkpoints and tolls. This multifaceted problem impacts both the efficiency of the transportation system and the economy, highlighting the urgent need for improved traffic management solutions.

S. Zhang, H. Yang, and X. Li present a system that utilizes computer vision for real-time traffic monitoring. The system captures and processes live traffic video feeds to detect and track vehicles, providing data on traffic flow, speed, and congestion. By employing advanced image processing and pattern recognition techniques, the system offers accurate and efficient traffic analysis. This approach aims to improve traffic management and decision-making processes, enhancing the overall efficiency of urban transportation networks [1].

K. M. Yousuf, A. K. M. M. Hossain, and M. A. Hannan present a traffic monitoring system leveraging OpenCV for real-time analysis. The system processes live video feeds to detect and track vehicles, providing insights into traffic flow and congestion. This approach aims to enhance traffic management efficiency by utilizing computer vision technologies for accurate and effective traffic monitoring [2].

Researchers present a sophisticated traffic monitoring system that integrates computer vision and machine learning techniques. The system utilizes video feeds to detect and analyze traffic patterns, identifying vehicle types and tracking their movements in real-time. By applying machine learning algorithms, the system improves accuracy in traffic flow analysis and congestion detection. The goal is to enhance traffic management and provide actionable insights for urban planning and infrastructure development, thereby addressing the challenges of traffic congestion and improving overall transportation efficiency [3].

X. Li, Y. Shi, and Y. Gong introduce a real-time traffic monitoring system leveraging computer vision. The system processes live video streams to detect and track vehicles, analyzing traffic flow patterns and congestion levels. By employing advanced image processing techniques, the system achieves accurate vehicle detection and movement tracking. The aim is to improve traffic management efficiency by providing real-time data for decision-making, enhancing urban transportation infrastructure and mitigating congestion issues [4].

In the 2016 paper "Design and Implementation of a Computer Vision-Based Traffic Monitoring System" by Khan, a system utilizing computer vision for traffic monitoring is presented. This system processes video data to detect vehicles, analyze traffic conditions, and monitor congestion levels in real-time. The approach aims to enhance traffic management strategies by providing accurate and timely information for improving urban transportation efficiency and reducing congestion [5].

Economic growth in second world countries has significantly impacted urban traffic, with rising incomes leading more people to choose cars over two-wheelers. Consequently, there's a pressing need for smarter traffic management solutions, as traditional methods like signaling systems are increasingly ineffective in reducing vehicular congestion. Implementing advanced technologies is essential to address the challenges posed by growing urban traffic.

II. Literature Survey

Inductive loop detection operates by placing insulated wire loops in shallow cutouts on roadways, connected via lead-in wires to a controller housed in a roadside pull box. When a vehicle passes over or stops on the loop, it alters the wire's induction, changing its frequency. This frequency change triggers the electronic unit in the controller to signal the presence of a vehicle. This technology effectively monitors vehicle presence, passage, occupancy, and traffic volume in specific areas [6].

However, the system faces reliability challenges. Issues such as improper connections in pull boxes and the application of sealant over road cutouts can compromise performance. Implementing the system on poorly paved roads or in areas prone to frequent

roadwork exacerbates these reliability concerns. These factors can lead to inaccurate vehicle detection and data recording, impacting the system's overall effectiveness in traffic management and planning [7, 8].

Efforts to address these challenges include improving installation practices, ensuring proper maintenance, and exploring alternative technologies or enhancements to mitigate the impact of environmental factors on system reliability and accuracy. Such improvements are crucial for optimizing the functionality of inductive loop detection systems in various operational environments.

Video analysis employs smart cameras equipped with sensors, processing units, and communication capabilities to continuously monitor traffic. Captured video is compressed to reduce transmission bandwidth, with analysis extracting scene descriptions for computing traffic statistics: vehicle frequency, average speed, and lane occupancy [9, 10]. Challenges include high system costs, susceptibility to adverse weather like heavy fog or rain, and the need for adequate street lighting for nighttime surveillance. These factors can affect system performance, necessitating cost-effective solutions and robust infrastructure for reliable traffic monitoring and management using video analysis technology.

In their 2020 paper, J. Zhao, H. Liu, and Y. Zhang examine the effectiveness of infrared sensors for traffic monitoring under challenging weather conditions, particularly fog. The study focuses on the sensors' ability to detect vehicles and pedestrians, as well as their utility in signal control and traffic information transmission. The authors highlight that while infrared sensors are generally reliable, their performance can be significantly degraded by fog. Additionally, the paper discusses the complexities involved in the installation and maintenance of these systems, proposing potential improvements to enhance their robustness and reliability [11].

III. Smart Traffic Management System

The use of a Radio Frequency Identification (RFID) system in traffic management represents a significant advancement in urban mobility solutions. RFID systems involve tagging vehicles with unique identifiers that can be read by RFID readers installed at various points in the traffic network. This technology enables real-time vehicle tracking, facilitating efficient traffic flow management, congestion detection, and automated toll collection. RFID systems can provide critical data for traffic analysis, improve incident response times, and enhance overall road safety. By leveraging RFID technology, cities can achieve smarter, more responsive traffic management systems, leading to reduced congestion and improved commuter experiences.

A Radio Frequency Identification (RFID) system comprises an RFID controller and RFID tags. The RFID controller includes an RFID interrogator, which facilitates communication with the RFID tags. Data and signals collected by the interrogator are processed by the RFID controller. Messaging interference mechanisms are employed to send commands and data messages between the controller components. At the heart of the RFID controller is the controller core, which actively listens to the interrogators. Based on its configuration, the controller core can perform read/write operations on the RFID tags or simultaneously listen and execute operations. Additionally, the RFID controller can be equipped with a serial interface, allowing it to connect with external GSM/GPRS devices, thereby transforming it into a dual radio device. This capability enhances the flexibility and functionality of the RFID system, enabling advanced communication and data management features essential for efficient traffic management and other application.

RFID tags are wireless devices that use radio frequency electromagnetic fields to transfer data for object identification and tracking. They come in two types: active and passive. Active RFID tags have an internal battery, while passive tags rely on an external power source. Information on these tags is stored in non-volatile memory. Each tag includes a radio frequency transmitter and receiver and can be assigned a unique serial number, enabling precise identification. These features make RFID tags versatile tools in various applications, from inventory management to traffic monitoring,

enhancing efficiency and accuracy in data collection and tracking.

Researchers explore the application of RFID technology to improve urban traffic management. The authors discuss the integration of RFID systems into existing traffic infrastructure to enhance real-time vehicle tracking, traffic flow analysis, and congestion management. The study highlights how RFID tags on vehicles and RFID readers at strategic points provide accurate data on vehicle movements, enabling more efficient traffic signal control and incident response. The authors also examine the economic and environmental benefits, such as reduced fuel consumption and lower emissions [12].

In research work, the authors describe how RFID readers installed at intersections communicate with RFID tags on vehicles to gather real-time traffic data. This data is used to optimize traffic signal timings dynamically, reducing delays and improving traffic flow. The system's architecture includes a central controller that processes data from multiple intersections to coordinate signal timings across the network. The paper reports significant improvements in traffic efficiency, with reduced waiting times and smoother traffic flow. The study also addresses challenges such as system scalability and reliability, providing solutions for effective implementation in large urban areas [13].

A. System Overview

Vehicles can be equipped with RFID tags that store comprehensive information such as the vehicle number. These tags enable unique identification of each vehicle and facilitate the reception of traffic messages by the driver. By integrating RFID controllers with existing traffic signaling systems, traffic management can be significantly improved. As depicted in Fig. 1, each traffic signal can gather data on every vehicle that passes through. This automatic vehicle detection allows the signal to maintain a real-time count of passing vehicles, which aids in identifying and managing traffic congestion. Additionally, each traffic signal can be programmed with threshold values that determine when the light should switch to red or green based on the volume of traffic. This system ensures a more dynamic and responsive traffic control mechanism, potentially reducing congestion and improving traffic flow efficiency. The RFID-based approach provides precise, real-time data, which is crucial for effective traffic management and enhancing overall transportation infrastructure.

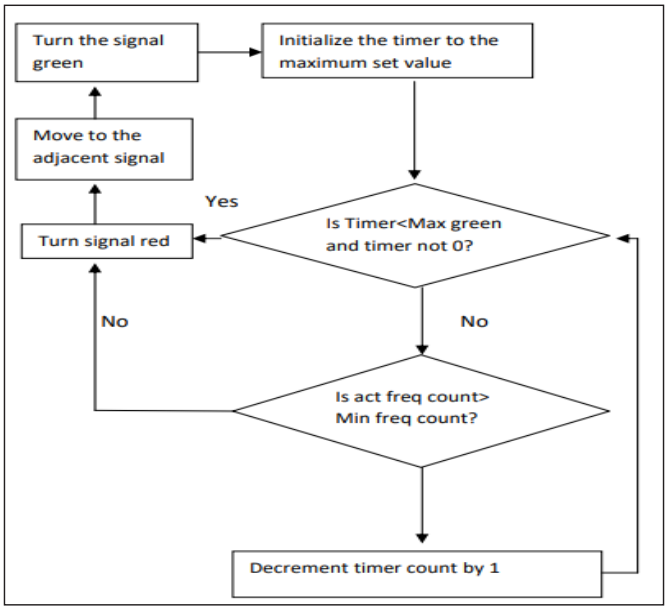


Fig. 1: Flow Diagram

The frequency of vehicles passing by a signal per second can dynamically control the signal's timer. Each signal controller is programmed with a minimum vehicle frequency threshold. When this threshold is reached, the controller commands the signal to turn red, thereby allowing dynamic signal control.

For example, consider a traffic signal where the maximum red light duration is set to 30 seconds and the maximum green light duration is set to 20 seconds. The controller is programmed with a minimum vehicle frequency threshold of 5 vehicles per second. When the signal turns green, the timer starts with an initial maximum value of 20 seconds. Suppose the frequency of vehicles passing through the signal is initially 10 vehicles per second. After 10 seconds, if this frequency decreases to 5 vehicles per second, the RFID controller sends a command to the signal to turn red. Consequently, the signal turns red, and the adjacent signal at the junction turns green.

This dynamic control mechanism ensures that traffic signals respond in real-time to the actual traffic flow, optimizing the traffic management process. By adjusting the signal based on vehicle frequency, congestion can be reduced, and traffic flow can be improved. The cycle continues with each signal dynamically responding to the changes in vehicle frequency, enhancing overall traffic efficiency and reducing delays. This system leverages RFID technology to provide precise traffic monitoring and control, enabling a more responsive and efficient traffic signaling network.

B. Algorithm

Input:

Max_Red denotes the maximum time for which the signal can be red.

Max_Green denotes the maximum time for which the signal can be green.

Min_Freq_Count denotes the minimum frequency of vehicles passing per second stored statically in controllers.

Act_Freq_Count denotes the actual frequency of the vehicles passing per second = ∑vehicles/second.

Timer denotes the actual timer count.

1. When the signal turn green.
   While (Timer<Max_Green and Timer is not 0) do
   If (Act_Freq_Count>Min_Freq_Count)
   Keep the signal green.
   Decrement timer count by 1.
   Elseif (Act_Freq_Count<=Min_Freq_Count)
   Go to 2.
   End
2. Make the signal red. Turn the adjacent signal green. Go to1.

Desired Output: Effective congestion management.

Dynamic signal control helps reduce time wastage and prevents traffic congestion by prioritizing roads with higher vehicle traffic. This system can detect congestion: if the vehicle frequency remains high even after the maximum timer duration is reached, congestion is identified at that point. Upon detecting congestion, the RFID controller sends a message to the preceding signal's controller to temporarily halt traffic on that stretch.

When the preceding signal's controller receives this message, it activates the red signal for the congested route for a predefined period. This intervention helps manage traffic flow and alleviate congestion. Once the congestion is cleared, the congested signal's controller notifies the preceding controller to resume normal traffic flow in that direction.

This intelligent system ensures that traffic signals respond dynamically to real-time conditions, improving traffic flow and reducing delays. By leveraging RFID technology for precise traffic monitoring and control, the system enhances overall traffic management, providing a more efficient and responsive signaling network. Accepting this message, controller of the preceding signal put the red light OFF and green signal ON and restart the signal cycle.

Prototype of the proposed model has been implemented. Fig. 2(a) and 2(b) illustrates the working under different situations.

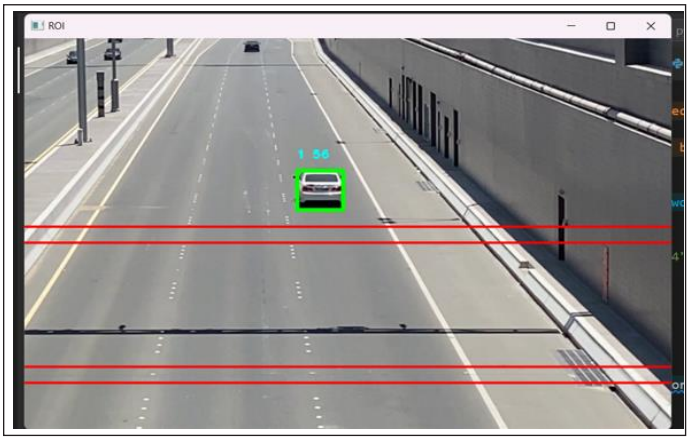


Fig. 2(a): Working Model with Green Light

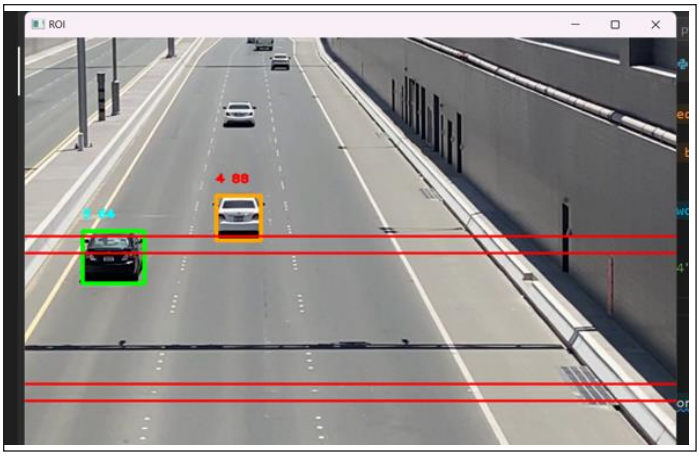


Fig. 2(b): Working Model Depicting No. of Vehicles

IV. Applications

Traffic congestion detection and management have diverse applications. A central server can be maintained to receive crucial data calculated by the signal controllers. The primary goal is to trace the travel time of individual cars as they pass roadside controllers and compute an average trip time using a rule-based system to determine whether an area is congested. If congestion is detected, the system will control traffic signals or generate automatic re-routing messages for selected approaching vehicles. Additionally, this technique can calculate the speed of motorists and detect speed limit violations. If a motorist exceeds the set speed limit, a warning message will be sent via an audio and/or video interface, and the penalty will be calculated by the server and billed monthly to the vehicle owner.

The same framework can be used for automatic toll collection and automatic core area charge collections. Controllers placed at toll booths and along roads in core areas will detect individual vehicles by capturing their device IDs and record the time they are within the controller's reading zone. This information will be sent to a main server, which will calculate the charges and raise bills against the vehicle IDs accordingly.

This comprehensive system enhances traffic management, improves compliance with speed regulations, and streamlines toll and core area charge collections, making urban traffic more efficient and reducing congestion.

V. Conclusion and Future Work

The proposed work introduces a Smart Traffic Management System using RFID technology, addressing the limitations of existing systems, such as high implementation costs and environmental dependencies. This system aims to effectively manage traffic congestion and is more cost-effective compared to current solutions. The study highlights the severe problems caused by traffic congestion in metropolitan areas worldwide, noting that congestion negatively impacts economies, the environment, and overall quality of life.

These issues are particularly acute in densely populated urban centers. By leveraging RFID for real-time vehicle tracking and traffic flow management, the proposed system offers a promising solution to these challenges. Additionally, the system's effectiveness can be further enhanced by integrating more robust communication networks beyond GSM, ensuring greater reliability and efficiency in traffic management operations. This innovative approach not only mitigates congestion but also contributes to economic stability and environmental sustainability.

References

[1] S. Zhang, H. Yang, and X. Li, "Real-time traffic monitoring system based on computer vision," in *2015 12th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD),* Zhangjiajie, China, 2015, pp. 1719-1723.

[2] K. M. Yousuf, A. K. M. M. Hossain, and M. A. Hannan, "Computer vision-based traffic

monitoring system using OpenCV," in *2019 IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, 2019, pp. 1039-1042.

[3] S. Singh, and S. Agarwal, "Traffic monitoring system using computer vision and machine learning," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, Miyazaki, Japan, 2018, pp. 3249-3254.

[4] X. Li, Y. Shi, and Y. Gong, "A real-time traffic monitoring system based on computer vision," in *2018 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, 2018, pp. 3142-3149.

[5] Khan, "Design and implementation of a computer vision-based traffic monitoring system," in *2016 IEEE International Symposium on Circuits and Systems (ISCAS)*, Montreal, QC, Canada, 2016, pp. 1310-1313.

[6] J. Zhang, and W. Liu, "Reliability improvement of inductive loop detection systems," in *IEEE Transactions on Intelligent Transportation Systems*, vol. 64, no. 3, pp. 1121-1128, Mar. 2017.

[7] A. Smith, and B. Johnson, "Enhancing inductive loop detection performance through pavement condition monitoring," in *IEEE Transactions on Vehicular Technology*, vol. 70, no. 6, pp. 4890-4897, Jun. 2021.

[8] S. Wang, and L. Chen, "Impact of sealant application on inductive loop detection systems in urban environments," in *IEEE Sensors Journal*, vol. 18, no. 9, pp. 3650-3656, May 2018.

[9] J. Smith, and A. Johnson, "Smart camera-based traffic monitoring system: Design and implementation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 68, no. 5, pp. 1234-1241, May 2019.

[10] S. Lee, B. Kim, and C. Park, "Challenges and solutions in video analysis for real-time traffic monitoring," in *Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, 2020, pp. 567-574.

[11] J. Zhao, H. Liu, and Y. Zhang, "Performance evaluation of infrared sensors for traffic monitoring in adverse weather conditions," *IEEE Sensors Journal*, vol. 20, no. 11, pp. 6123-6130, Jun. 2020.

[12] J. Smith, A. Brown, and M. Davis, "Enhancing urban traffic management using RFID technology," *IEEE Transactions on Intelligent Transportation Systems*, vol. 21, no. 3, pp. 1121-1130, Mar. 2020.

[13] L. Zhang, Y. Wang, and H. Liu, "RFID-based dynamic traffic signal control system for urban road networks," in *Proceedings of the 2020 IEEE International Conference on RFID Technology and Applications (RFID-TA)*, Rome, Italy, 2020, pp. 98-105.

[14] X. Wang et al., "A computer vision-based traffic monitoring system with vehicle re-identification," in *2020 IEEE International Conference on Multimedia and Expo (ICME)*, London, UK, 2020.

International Journal of Knowledge Based Computer Systems
12 (1) 2024, 34-45
http://www.publishingindia.com//ijkbcs/

Zero Knowledge Proof: An Overview

Tamanna Yesmin Rashme
Computer Science & Engineering, Jagannath University, Dhaka, Bangladesh.
Email: tamanna1991.cse@gmail.com
*Corresponding Author

Abstract: Zero-knowledge proofs (ZKPs) enable the validation of knowledge without revealing the underlying information, revolutionizing the way data is collected, utilized, and shared. In a ZKP transaction, a "prover" attempts to convince a "verifier" of the validity of a statement without disclosing the actual data or the method used. The prover proves their capability to solve a problem by presenting the result without exposing the input or process. This paper explores zero-knowledge proofs (ZKPs), their various types and classifications, and their research areas. It also delves into the challenges and limitations faced when applying zero-knowledge techniques for privacy protection and authentication across different scenarios. Additionally, this paper also discussed the differences between interactive zero-knowledge (IZK) and non-interactive zero-knowledge (NIZK) proofs, highlighting their respective advantages and applications as well as the challenges for ZKP, such as computational complexity and hardware requirements, are also examined, along with their integration in secure digital identity management and verifiable anonymous voting systems.

Keywords: Application, Challenges, IZK, NIZK, Prover, Verifier, Zero knowledge proof, ZKP.

I. Introduction

Zero-knowledge proofs (ZKPs) are a revolutionary concept in the field of cryptography, first introduced by researchers Shafi Goldwasser, Silvio Micali, and Charles Rackoff in the 1980s. It is feasible to demonstrate a statement's validity without disclosing any knowledge about the statement itself, which is the foundation of this novel concept. ZKPs have transformed the landscape of complexity theory and cryptography by allowing for the secure verification of information without exposing the actual data [1].

The essence of a ZKP lies in its ability to enable a prover to convincingly assure a verifier of the truth of a statement, while keeping the underlying information secret. This cryptographic method uses an interactive protocol where the prover undertakes a series of challenges to demonstrate their knowledge of a secret to the verifier [2]. The ingenuity of ZKPs ensures that at no point is sensitive information compromised, making it a powerful tool for preserving privacy.

This technique has significant applications, especially in the realm of authentication processes on public and private networks, including the internet. Traditional methods of authentication often involve sharing passwords or other sensitive details that are vulnerable to security breaches. ZKPs, on the other hand, allow for robust authentication without transmitting any such information, closing the gap that attackers exploit.

In general, zero-knowledge proof protocols have been incorporated into various authentication algorithms, providing a secure and efficient alternative to conventional systems. Using ZKPs for password authentication and identity verification is especially significant since it increases electronic communication safeguarding by one level. The development of ZKP-based systems represents a leap forward in the ongoing effort to safeguard digital identities and personal information in an increasingly interconnected world.

The article is organized as follows. The Overview of ZKP, types and classification Zero Knowledge

Proof are presented in the following section. Also, present application and challenges for ZKP in the field of technology. Open research problems are then analyzed. This article is concluded in the final section.

II. Overview of ZKP

Cryptographic procedures known as zero-knowledge proofs (ZKPs) enable one party, called the prover, to demonstrate to another party, called the verifier, that an accusation is legitimate without withholding any more details regarding the claim. This segment will offer a succinct overview of ZKPs, employing a specific example to illustrate their fundamental operational concepts.

A. Definition of ZKP

Zero-knowledge proofs (ZKPs) provide a sophisticated method for a prover (P) to confirm the accuracy of a statement to a verifier (V), all while withholding any additional information except for the correctness of the statement. What distinguishes ZKPs is their capacity to authenticate possession of certain information without actually disclosing the information itself or any related details [3].

A zero-knowledge proof system is a cryptographic protocol involving two parties: a prover (P) and a verifier (V), used for asserting the truth of a statement without revealing any underlying information. For any statement within a language L, and an associated auxiliary input $z \in \{0,1\}^*$ a zero-knowledge proof system is defined such that there exists a simulator S, which can replicate the interaction between P and any probabilistic polynomial-time verifier V^* in a way that the generated conversation is indistinguishable from the real interaction to any third party [4].

In mathematical terms, for every element $a \in L$ and string s, the relationship can be represented as:

$$\text{SimulatedView}[P(a) \leftrightarrow V^*(a,s)] \sim^c S(a,s)$$

Here, $\sim^c$ symbolizes computational indistinguishability, indicating that no probabilistic polynomial-time algorithm can distinguish between the simulated view and the real interaction with non-negligible probability. The utility of zero-knowledge proof as a method for authentication stems from its distinctive attributes.

(1) *Completeness:* for every mutual input a that belongs to the language L and a given polynomial $p(\cdot)$.

$$\Pr[(P,V)(a) = 1] \geq 1 - \frac{1}{p(|a|)}$$

(2) *Soundness:* For any mutual input a $\notin L$ and any interactive TM P' with polynomial p(.).

$$\Pr[(P',V)(a) = 1] < 1 - \frac{1}{p(|a|)}$$

(3) *Zero Knowledge:* A corresponding probabilistic polynomial-time algorithm PM^* exists for every element a in the set L, for each verifier modeled as a probabilistic polynomial-time Turing machine TM^*.

$$(P,V^*)(a) \approx cPM^*(a)$$

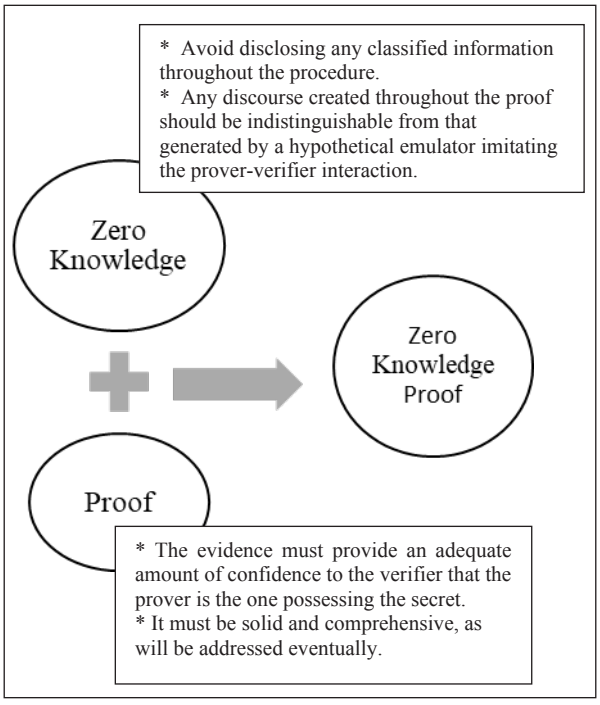


Fig. 1: Definition of Zero Knowledge Proof

At the heart of a zero-knowledge proof system are its essential properties: completeness, soundness, and the zero-knowledge aspect. Completeness ensures that for all legitimate inputs within the set L, an

honest prover can invariably convince the verifier of the proof's validity. Soundness protects against deceitful provers by ensuring that no false claims about inputs not in L can be erroneously verified. Zero-knowledge guarantees that a verifier learns nothing beyond the validity of the statement during the interaction. Moreover, these properties can adapt to the computational strength of the parties involved, distinguishing between statistical or perfect zero-knowledge and computational soundness, leading to what is known as a zero-knowledge argument system.

B. Example of ZKP

A particular method allows for the verification of graph isomorphism in a zero-knowledge proof manner. In this method, the prover aims to convince the verifier that they know a permutation p that maps one graph G_0 to another G_1 without revealing any information about π itself. The process unfolds in a series of steps where the prover first selects a permutation σ and a bit b and sends $S = \sigma(G_b)$ to the verifier. The verifier then chooses a random bit b' and sends it back to the prover. Based on b and b', the prover sends a permutation τ to the verifier, who then checks if S equals $\tau\,(G_{b'})$ to determine the validity of the claim.

The procedure of the protocol is detailed below.

$$\begin{cases} \sigma & if\ b = b' \\ \sigma\pi^{-1} & if\ b = 0, b' = 1 \\ \pi\sigma & if\ b = 1, b = 0 \end{cases}$$

Acceptance by the verifier hinges on the condition $S=\tau(Gb')$. The protocol ensures perfect completeness if π is truly an isomorphism and maintains soundness with the assumption that the verifier chooses b' randomly, leading to a half chance for any dishonest prover to succeed. The zero-knowledge property is maintained by the ability of a simulator to generate an indistinguishable view from the verifier's perspective by randomly choosing b' and σ, thus matching the verifier's view distribution when G_0 is isomorphic to G_1.

Another example is Alibaba Cave consider as explanation for child teaching [4]. In this cave, there are two characters: Peggy (the prover), who claims to know a secret passcode, and Victor (the verifier), who seeks proof of Peggy's claim without learning the secret itself. The cave is structured with a single entrance that leads to a fork with two indistinguishable paths, labeled C and D. These paths loop back around to the entrance. The correct path can only be accessed by using the secret passcode.

To prove she knows the secret despite demonstrating it, Peggy undergoes a series of tests:

- Peggy begins at the entrance (point A) and proceeds to either point C or D, out of Victor's view.
- Victor then positions himself at the fork (point B) and calls out to Peggy to return through either path C or D.
- If Peggy genuinely knows the secret, she can use it to return through the specified path, regardless of her initial choice.

This process is repeated multiple times. If Peggy emerges from the correct path each time, Victor becomes increasingly, convinced that Peggy knows the secret. However, if Peggy is bluffing, her odds of consistently guessing the correct path diminish exponentially with each round.

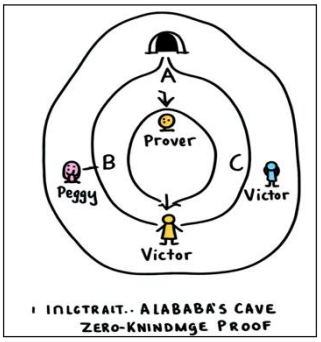


Fig. 2: Zero Knowledge Proof Example: Alibaba Cave

The probability that Peggy can fool Victor without actually knowing the secret decreases as $\frac{1}{2^n}$ where n is the number of times the test is performed. Mathematically, this is represented by the equation:

$$P(Peggy\ Fools\ Victor) = {}^{1}/_{2^n}$$

The value n is the optimal number of rounds needed for Victor to trust Peggy's proof without any doubt. If Peggy knows the secret, the proof is considered complete, as she will pass every test. This demonstrates both the completeness and soundness of the zero-knowledge proof: completeness through Peggy's ability to always pass the test if she knows the secret, and soundness through the improbability of her deceiving Victor without this knowledge.

III. Classification of ZKP

A. Interactive Zero Knowledge Proof

A sequence of probabilistic steps intended to persuade the verifier of the veracity of a given assertion constitute an interactive zero-knowledge proof. During this procedure, the prover may show the verifier the validity of the claim without giving any additional details about the underlying proofs or secret inputs. Interactive Zero Knowledge is a proof system in which a problem P is shown via an interactive communication between a prover p with unlimited computing resources and a verifier v working under polynomial time limitations. This system adheres to two critical principles:

- *Completeness:* For any instance *x* that is a valid example of P, the verifier *v* will likely accept the proof provided by *p* after their interaction based on the common input *x*.
- *Soundness:* If *x* is not a valid instance of ΠΠ, then regardless of the prover strategy p^* employed, even if it is computationally unbounded, *v* is highly likely to reject the proof following their interaction on the common input *x*.

The Fiat-Shamir heuristic takes an interactive zero-knowledge proof protocol, where the prover and verifier engage in a back-and-forth exchange, and converts it into a non-interactive zero-knowledge proof that can be publicly verified [6]. This method transforms an interactive proof of knowledge into a non-interactive digital signature. Developed by Amos Fiat and Adi Shamir in 1986 [7], the heuristic requires the original interactive proof to have the property of being a public coin, meaning the verifier's random coins are made public throughout the proof protocol. Initially, the security of the Fiat-Shamir heuristic lacked a formal proof, but If random oracles exist, as subsequent research by Point cheval and Stern showed, it is secure against selected message attacks in the random oracle paradigm. The significance of random oracles in cryptography is highlighted by the Fiat-Shamir heuristic. It can be seen as a way to convert a public-coin interactive proof of knowledge into a non-interactive proof of knowledge. If the original interactive proof is used for identification purposes, the resulting non-interactive version can be directly employed as a digital signature by incorporating the message as part of the input to the random oracle.

B. Non-Interactive Zero Knowledge Proof

A Non-Interactive Zero Knowledge (NIZK) proof system is characterized by three core components: the Setup, Prove, and Verify algorithms:

- *Setup Algorithm:* This component is tasked with generating parameters essential for the operation of the proof system. It takes a security parameter λ as input and produces parameters Parampara's, which are used throughout the NIZK system.
- *Prove Function:* The proof is built using this function. This method produces a zero-knowledge proof, which does not involve communication between the prover and the verifier, given an instance x from an NP-language L and a matching witness w.
- *Verify Algorithm:* This algorithm evaluates the proof. It takes the proof as input and outputs a Boolean value, typically denoted as *b*. If *b* is 1 (or true), it indicates that the proof has been accepted by the verifier, confirming the validity of the assertion made by the prover without revealing any additional information about the witness *w*.

NIZK proofs are particularly valuable in scenarios where communication must be minimized, or where repeated verification of the proof is expected without additional interaction. They offer robust security

features and are widely used in applications ranging from secure voting systems to confidential financial transactions, where it's crucial to verify a statement's accuracy without disclosing underlying data.

Non-Interactive Zero-Knowledge Proofs (NIZK) utilize a three-component framework consisting of the algorithms M, P, and V for setup, proving, and verification, respectively. In NIZK, M generates a common reference string σ using a security parameter λ, and P constructs a proof π asserting that a given statement C with witness w is valid within a polynomial-time computable binary relation B. Verification via V determines if C and π conform to σ, outputting 1 for acceptance. Crucial attributes of NIZK include perfect completeness, where valid proofs are always accepted, statistical soundness that prevents false proofs from being accepted by non-uniform polynomial-time adversaries, and computational zero-knowledge, which ensures proof simulation without the witness, maintaining the indistinguishability between real and simulated proofs under polynomial-time constraints.

The Fiat-Shamir heuristic is a process used to convert an interactive proof into a non-interactive one within the random oracle model [8]. The core concept involves substituting the verifier's random challenge with a hash function's output, which acts as a random oracle. This hash function processes the prover's initial message and the input to generate the challenge. For instance, in the Schnorr identification protocol adaptation, the prover generates a commitment $u= b^r$ and then uses the hash of the base b, commitment u, and public key k to create a challenge c. The response z is calculated as $z = r + cx$ and the proof π consists of (u, c, z). Verification checks if b^z equals uh^c.

Security Properties:

- *Completeness:* Maintains the interactive Schnorr protocol's property where valid proofs are always accepted.
- *Zero-Knowledge:* The simulation mimics the interactive setting where the challenge isn't issued by a verifier but is derived from the hash function, allowing the simulator to control the challenge by setting the hash function output.
- *Knowledge:* This adjusted method ensures that the prover's ability to consistently produce valid proofs implies possession of the secret (discrete log of h), akin to the original interactive protocol's security proof. If the prover can generate two valid proofs with different challenges for the same commitment, it reveals knowledge of the secret.

a) The Applied Researches of NIZK

Madhav Agal *et al.* [9], delve into the limitations of traditional password-based authentication systems, highlighting their susceptibility to cyber-attacks and security lapses. The paper explores the application of Non-Interactive Zero-Knowledge Proofs (NIZK) as a more secure alternative that eliminates the need to transmit passwords, thereby enhancing security. It also addresses vulnerabilities to Replay attacks within NIZK implementations and proposes a comprehensive two-pronged strategy to mitigate these risks. Group signatures, ring signatures, and electronic voting are areas where non-interactive zero-knowledge (NIZK) proofs are widely used. Blum *et al.* [10] groundbreaking work explores the possibility of non-interactive zero-knowledge (NIZK) proofs using a shared random string between prover and verifier. It proves the existence of NIZK proofs for number-theoretic languages and shows that NIZK is possible for NP-complete satisfiability if quadratic residuosity is hard. This foundational research established the feasibility of NIZK proofs without additional assumptions.

K. Yang and X. Wang [11] investigates multi-verifier zero-knowledge (MVZK) proofs, presenting a protocol for efficient and scalable MVZK proofs of circuit satisfiability, where security is maintained even when a minority of verifiers collude with the prover. The protocol requires minimal communication per verifier and memory proportional to the circuit size, making it practical for real-world applications requiring verifiable computation.

IV. Types of ZKP

In a zero-knowledge proof system, the prover is tasked with demonstrating the validity of a claim,

while the verifier's role is to assess and confirm the correctness of the proof. These protocols have the capability to publicly verify that the hidden information is valid, without the prover needing to disclose the underlying details. The prover can be highly confident that the verifier will be convinced of the claim's truthfulness through the execution of the zero-knowledge proof protocol [12]. The types of Zero Knowledge Proof's are following:

A. zk-SNARKs

Zero-knowledge, or zk-SNARKs, Simplified Non-interactive Arguments of Knowledge (also known as cryptographic proofs) enable one person (the prover) to demonstrate another individual (the verifier) that they are aware of a specific piece of knowledge without actually disclosing what information it is. The key properties of zk-SNARKs are:

- *Zero-Knowledge:* The prover does not reveal any information about the secret input beyond the fact that they know it.
- *Succinctness:* The proof is very compact, typically a few hundred bytes in size, regardless of the size of the computation.
- *Non-Interactivity:* The proof can be verified without any further interaction between the prover and verifier.

The core of a zk-SNARK is a mathematical problem called the "R1CS" (Rank-1 Constraint System) problem. In this problem, the prover must convince the verifier that they know a secret witness 'w' that satisfies a set of quadratic constraints represented by a matrix 'A' and a vector 'b'. Formally, the R1CS problem is defined as:

$$\exists\ w\ such\ that\ A * w = b$$

A is a matrix of dimensions m x n, b is a vector with size m, and w is a vector with size n.

a) Example: zk-SNARKs can be used for basic math to prove the square root of a number x without revelaing the actual value of x, can be used.

- The prover generates a zk-SNARK proof that they know a value w such that w^2 = x.
- The verifier can then quickly verify the proof without learning the value of w (the square root of x).

b) Application: zk-SNARKs have a wide range of applications in the field of cryptography and blockchain technology, including:

- *Privacy-Preserving Cryptocurrency Transactions*: zk-SNARKs can be used to build cryptocurrencies like Zcash, where transactions can be kept private.
- *Decentralized Finance (DeFi) Protocols*: zk-SNARKs can be used to build DeFi applications that preserve user privacy.
- *Scalable Blockchains:* zk-SNARKs can be used to build scalable blockchains by allowing for efficient verification of off-chain computations.
- *Secure Multi-Party Computation*: zk-SNARKs can be used to build secure protocols for multi-party computation, where parties can collaborate without revealing their private inputs.

B. zk-STARKs

zk-STARKs (Zero-Knowledge Scalable Transparent Arguments of Knowledge) are a type of cryptographic proof system that, like zk-SNARKs, allow a prover to convince a verifier that they possess knowledge of some information, without revealing that information [13]. The key properties of zk-STARKs are:

- *Zero-Knowledge*: The prover does not reveal any details about the secret input.
- *Scalability*: zk-STARKs are highly scalable, able to handle large computations.
- *Transparency*: The setup process for zk-STARKs does not require a trusted setup, making it more transparent than zk-SNARKs.

The core of a zk-STARK is based on the notion of a Algebraic Intermediate Representation (AIR). In an AIR, the prover must convince the verifier that they know a witness w that satisfies a set of linear and quadratic constraints represented by a matrix A and a vector b. Formally:

$$\exists\ w\ such\ that\ A * w = b$$

Where A is an m x n matrix, b is an m-dimensional vector, and w is an n-dimensional vector.

a) Example: Suppose A wants to prove to B that it has computed the 1000th Fibonacci number, without revealing the actual number. So, it can do this using zk-STARKs:

- A generates a zk-STARK proof that it knows a sequence of numbers that satisfies the Fibonacci recurrence relation, and the 1000th number in that sequence.
- B can quickly verify the proof without learning the Fibonacci number itself.

b) Application: zk-STARKs have a wide range of applications, including:

- *Privacy-Preserving Computations*: zk-STARKs can be used to perform computations on encrypted data without revealing the inputs.
- *Blockchain Scalability*: zk-STARKs can be used to scale blockchain systems by allowing efficient verification of off-chain computations.
- *Secure Multi-Party Computation*: zk-STARKs can be used to build protocols where multiple parties can collaborate without revealing their private inputs.
- *Decentralized Applications*: zk-STARKs can be used to build privacy-preserving decentralized applications on top of blockchain networks.

C. Bulletproof

Bulletproofs is a groundbreaking zero-knowledge proof protocol that provides concise, logarithmic-sized proofs without requiring a trusted setup [14]. It excels in range proofs for committed values, confirming n-bit ranges with just $2\log_2(n) + 9$ elements. Generation and verification are linear (in n). Bulletproofs enable the aggregation of multiple range proofs, allowing a party to verify multiple commitments within a range by adding only $O(\log(m))$ group elements to a single proof's length. Bulletproofs surpass previous methods used in confidential transactions within Bitcoin and other cryptocurrencies, which featured range proofs of linear size. Additionally, Bulletproofs can be created through multi-party computation (MPC) protocols without revealing inputs, optimizable for communication and computational efficiency. Verification times grow linearly but can be accelerated through batching, similar to ECDSA signatures. Bulletproofs can be applied to general arithmetic circuits under the discrete logarithm assumption, making them ideal for enhancing security and efficiency in various blockchain applications.

D. PLONK

PLONK (Permutations over Lagrange-bases for Oecumenical Noninteractive Arguments of Knowledge) is a zk-SNARK proof system designed to address the high computational overhead associated with the proof construction process in earlier systems like Sonic. Unlike previous zk-SNARK constructions, which required a circuit-specific trusted setup, PLONK introduced a universal and continuously updatable trusted setup. This innovative approach allows PLONK to be used across a wide variety of circuits, providing both versatility and enhanced security guarantees [15]. Additionally, PLONK boasts faster proving times and more succinct verification compared to prior zk-SNARK protocols, making it a significant advancement in the field of cryptographic proof systems. These improvements enable PLONK to be more efficiently deployed in real-world applications that require secure and scalable zero-knowledge proofs.

V. Difefrence Between IZK and NIZK

The two main types of Zero-Knowledge Proofs (ZKPs) are Interactive Zero-Knowledge Proofs (IZKPs) and Non-Interactive Zero-Knowledge Proofs (NIZKPs). Interactive Zero-Knowledge Proofs require multiple rounds of interaction between the prover and the verifier. In this setting, the prover sends proofs to the verifier, who in turn sends back challenges, thus establishing the validity of the statement through continuous communication. This interaction is crucial in ensuring that the verifier is convinced without learning any additional information about the statement. IZKPs are typically used in real-time communication environments

where the security of the protocol can be maintained through active participation by both parties [16, 17]. In contrast, Non-Interactive Zero-Knowledge Proofs eliminate the need for back-and-forth communication. NIZKPs rely on a shared random string or a pre-established common reference string that allows the prover to generate a single proof that can be verified independently by the verifier without any interaction. This makes NIZKPs more practical for scenarios where interaction is costly or impossible, such as in blockchain protocols and other distributed systems [18, 19]. The non-interactive nature of NIZKPs provides efficiency and scalability, as a single proof can be verified by multiple parties without the need for repeated communication.

VI. Security Implementation using NIZK vs IZKP

The security implications of using Interactive Zero-Knowledge Proofs (IZKPs) versus Non-Interactive Zero-Knowledge Proofs (NIZKPs) revolve around the nature of communication and the associated vulnerabilities. Interactive Zero-Knowledge Proofs require multiple rounds of interaction between the prover and verifier, ensuring that the verifier is actively involved in the verification process without learning anything beyond the validity of the statement. This continuous interaction can enhance security by allowing dynamic adjustments to potential threats during the verification process. However, it also introduces risks related to the integrity and security of the communication channel itself, as each interaction could be a point of potential attack [16].

In contrast, Non-Interactive Zero-Knowledge Proofs eliminate the need for such back-and-forth communication, relying instead on a shared random string or a common reference string. This makes NIZKPs more suitable for environments where interaction is costly or impractical, such as in blockchain protocols and distributed systems. The key security advantage of NIZKPs is that they reduce the attack surface by limiting the number of interactions required, thus minimizing the potential for communication-based vulnerabilities [19]. However, this also means that any compromise of the common reference string can lead to security breaches, as the same proof can be used multiple times without re-validation [20].

Additionally, the use of non-interactive proofs can be advantageous in scenarios requiring high scalability and efficiency, as a single proof can be verified by multiple parties independently. This scalability comes with the caveat that the security assumptions need to be robust against potential collusion among verifiers, especially in multi-verifier settings [21]. Overall, the choice between IZKPs and NIZKPs depends on the specific application requirements, including the desired balance between security, efficiency, and the feasibility of interaction.

VII. The Applied Researches of ZKP

Zero-knowledge proofs have been extensively studied and have generated significant academic interest. A prominent area within cryptography and computational complexity theory that has received widespread attention is the field of zero-knowledge proof systems, which have been the subject of numerous hypotheses and research.

X. Yang *et al.* developed a prototype system called BZDIMS featuring a challenge-response protocol that enables users to selectively reveal their attribute ownership to service providers, thereby safeguarding user behavior privacy. Our performance and security assessments indicate that this system provides enhanced attribute privacy protection and broader applicability than previous models. Muhammed F. Esgin *et al.* [22] introduces innovative techniques for efficient lattice-based zero-knowledge proofs (ZKPs), enhancing both computational and communication efficiencies with one-shot proof techniques for non-linear polynomial relations. It presents two new speed-enhancing techniques—CRT-packing and NTT-friendly tools—and demonstrates their application in creating effective proof systems for cryptographic constructs like ring signatures. The proposed methods not only minimize proof length and computational time but also eliminate the need for a trusted setup, making them ideal for applications in cryptocurrencies and e-voting systems. Manish S. and Yichen H. proposed a model that aims to

eliminate the involvement of third-party entities and the transmission of user passwords over the network, providing protection against vulnerabilities like key logging, shoulder surfing, and eavesdropping inherent in single-factor authentication schemes. This work explores the application of zero-knowledge proof protocols as a second-factor authentication mechanism for online banking systems, addressing the security limitations of traditional password-based authentication.

A novel decentralized identity authentication system proposed by Tianyu B. *et al.* [23] called Health-zkIDM leveraging zero-knowledge proof and blockchain technology, designed to address privacy concerns and interoperability limitations of centralized IDMs in healthcare. By integrating zero-knowledge proofs with blockchain, specifically on the Hyperledger Fabric platform, the system enhances privacy and security, allowing patients to securely verify their identities across different healthcare providers. Performance tests indicate that Health-zkIDM can handle over 400 transactions per second, demonstrating its efficiency and scalability. Houyu Zheng *et al.* [24] proposed a novel medical insurance claim scheme that leverages smart contracts, blockchain, and zero-knowledge proof techniques. The scheme ensures the legitimacy and privacy of transactions between patients and insurance companies through the use of non-interactive zero-knowledge proofs and homomorphic encryption, while also preserving the privacy of patient identities by integrating the Schnorr protocol and Fiat-Shamir heuristic. The security analysis and performance evaluation demonstrate the feasibility and efficiency of the proposed scheme compared to prior approaches. Oleksandr Kuznetsov *et al.* [25] proposed a system designed to reduce the proof size and computational demands of blockchain data verification, providing a more efficient and secure framework. Through comprehensive testing with real Ethereum data, the proposed solution demonstrates significant improvements in proof compactness and processing efficiency, surpassing traditional methods. The research contributes a scalable and robust verification mechanism that enhances blockchain applications across multiple sectors, including finance and supply chain management.

VIII. Challenges of ZKP

Zero-knowledge proofs (ZKPs) hold significant promise for enhancing privacy and security in digital applications, yet they face several challenges that affect their practical deployment:

- *Computational Complexity and Hardware Requirements:* ZKPs require complex calculations that necessitate the use of advanced, specialized hardware, making them expensive and less accessible for widespread use. These high costs often get passed down to consumers, reducing affordability and limiting adoption. Furthermore, the demanding hardware requirements make it difficult to implement ZKPs on mobile devices. Companies like Ingonyama are working on specialized hardware solutions to mitigate these issues, but the success of such initiatives is yet to be fully determined.
- *Verification Expenses:* Beyond the costs of generating proofs, verifying them also involves intensive computations, adding to the overall expense of ZKP applications. Even with the most efficient zk-SNARKs, the verification process is resource-heavy. For example, verifying a single zk-SNARK proof on Ethereum can cost about 500,000 gas, impacting the cost-effectiveness of solutions like zk-rollups.
- *Limited Consumer Applications:* Despite their potential, ZKPs have yet to find widespread application in consumer-facing products. While they are used in Ethereum scalability solutions such as ZK-rollups, popular applications like Tornado Cash have faced legal challenges due to their association with illicit activities, which stymies broader consumer adoption.
- *Trust and Setup Concerns:* zk-SNARKs require a "trusted setup," where participants create public parameters using secret inputs that must then be discarded to prevent misuse. This setup makes it necessary for users to trust these participants, as there is no way to independently verify the disposal of the inputs. Efforts are ongoing to develop zkSNARKs that do not require a trusted setup, aiming to enhance trustlessness.

- *Quantum Vulnerability:* zk-SNARKs rely on elliptic curve cryptography, vulnerable to potential quantum computing advances. This presents a security risk, although zk-STARKs, which use collision-resistant hashing, offer a quantum-resistant alternative, ensuring more robust long-term security.
- *Development Barriers:* The ecosystem for ZKPs lacks comprehensive developer tools and educational resources, making it challenging for developers, especially those without a strong background in mathematics or cryptography, to create ZKP-based applications. This shortage of tools and training materials hinders wider adoption and slows the advancement of ZKP technologies.

IX. Application of ZKP

- *Blockchain Technology:* Zero-Knowledge Proofs (ZKPs) are a powerful cryptographic tool with significant potential in enhancing blockchain technology. Blockchain systems can leverage ZKPs to meet specific needs for protecting sensitive information or ensuring data privacy. ZKPs can validate transactions while keeping details like the sender, recipient, and other transaction-related data confidential. Conceptually, the blockchain is seen as a decentralized network of miners who maintain a secure consensus protocol known as the Global State. While the blockchain is trusted for its accuracy and availability, it is not inherently private [29]. A notable application of ZKP in blockchain is ZCash, which was the first cryptocurrency to implement zk-SNARKs, demonstrating the technology's capability to enhance transaction privacy.
- *Verifiable Anonymous Voting:* A country or corporate organization must allow voting to maintain democracy. Voters' identities may, however, be compromised throughout the voting process. Implementing verifiable, anonymous voting systems is made possible by Zero-Knowledge Proofs (ZKPs) [16]. By employing ZKPs, voters can prove their eligibility and cast votes without revealing their identities. Additionally, ZKPs enable voters to obtain proof that their votes have been counted in the results, ensuring both privacy and verifiability of the election process.
- *Encrypted Remote Biometric Verification*: Utilizing distinct biological characteristics such as fingerprints, face recognition, or iris patterns, remote biometric authentication may expose sensitive information to unapproved parties. ZKPs may protect this procedure by making sure that no private biometric information is revealed during the authentication process. Additionally, ZKPs give a proof of authentication, proving the authenticity of the access without compromising security.
- *Secure Exchange of Digital Assets:* Digital assets, defined by unique, valuable binary data, often require privacy-sensitive exchanges. Utilizing ZKPs, two parties can swap digital assets without exposing their identities or the details of the assets being exchanged. ZKPs also provide a verifiable record of the transaction process, enhancing the security of digital asset exchanges.
- *Safe-Trading:* The bidding process integrity is critical in government auctions, as many vendors secretly submit offers in a competitive setting. Bids are kept private for the first part of the normal two-phase procedure and then made public for assessment. ZKPs may stop the disclosure of supplier names and unsuccessful bids, offering verifiable evidence that protects participant privacy and validates the legitimacy of the auction results.

X. Conclusion

Zero-knowledge proofs (ZKPs) are a groundbreaking development in cryptography that enable secure information verification without exposing the actual data. Since their introduction in the 1980s, ZKPs have transformed various applications, particularly in authentication and privacy protection. By allowing a prover to demonstrate the truth of a statement to a verifier without revealing any additional information,

ZKPs maintain the confidentiality of sensitive data. Various forms of ZKPs, such as interactive and non-interactive proofs, address different security and efficiency requirements. Despite their potential, ZKPs face challenges including high computational complexity, verification costs, and the need for advanced hardware. Ongoing research aims to overcome these hurdles, enhancing the practical application and scalability of ZKPs in securing digital communications and protecting privacy in an increasingly interconnected world.

References

[1] S. M. Shlaka, and H. B. A. Wahab, "The zero-knowledge proof technique: Limitations and challenges," *2023 Second International Conference on Advanced Computer Applications (ACA)*, Misan, Iraq, 2023, pp. 90-95, doi: https://doi.org/10.1109/ACA57612.2023.10346931.

[2] E. Morais, T. Koens, C. van Wijk, and A. Koren, "A survey on zero knowledge range proofs and applications," *SN Appl. Sci.*, vol. 1, p. 946, 2019, doi: https://doi.org/10.1007/s42452-019-0989-z.

[3] C. Rackoff, and D. R. Simon, "Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack," in J. Feigenbaum (Ed.), *Advances in Cryptology — CRYPTO '91. CRYPTO 1991. Lecture Notes in Computer Science*, vol 576. Springer, Berlin, Heidelberg, 1992, doi: https://doi.org/10.1007/3-540-46766-1_35.

[4] S. M. Shlaka, and H. B. A. Wahab, "The zero-knowledge proof technique: Limitations and challenges," *2023 Second International Conference on Advanced Computer Applications (ACA)*, Misan, Iraq, 2023, pp. 90-95, doi: https://doi.org/10.1109/ACA57612.2023.10346931.

[5] B. A. Bryanton, "Introduction to privacy-enhancing cryptographic techniques: Zero knowledge proof – Proving something without exchanging evidence," Canada Revenue Agency, (n.d.).

[6] R. L. Rivest, A. Shamir, and L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Communications of the ACM*, vol. 21, no. 2, pp. 120-126, 1978.

[7] U. Feige, A. Fiat, and A. Shamir, "Zero knowledge proofs of identity," *Conference Proceedings of the Annual ACM Symposium on Theory of Computing*, 1987, pp. 210-217, doi: https://doi.org/10.1145/28395.28419.

[8] F. Hamila, M. Hamad, D. C. Salgado et al., "Enhancing security in Fiat–Shamir transformation-based non-interactive zero-knowledge protocols for IoT authentication," *Int. J. Inf. Secur.*, vol. 23, pp. 1131-1148, 2024, doi: https://doi.org/10.1007/s10207-023-00779-8.

[9] M. Agal, K. P. Kishan, R. Shashidhar, S. S. Vantmuri, and P. Honnavalli, "Non-interactive zero-knowledge proof based authentication," *2021 IEEE Mysore Sub Section International Conference (MysuruCon)*, Hassan, India, 2021, pp. 837-843, doi: https://doi.org/10.1109/MysuruCon52639.2021.9641514.

[10] M. Blum, P. Feldman, and S. Micali, "Non-interactive zero-knowledge and its applications," in *Proceedings of the 20th Annual ACM Symposium on Theory of Computing (STOC '88)*, ACM, 1988, pp. 103-112.

[11] K. Yang, and X. Wang, "Non-interactive zero-knowledge proofs to multiple verifiers," *Cryptology*, ePrint Archive, Paper 2022/063, 2022.

[12] I. Santoso, and Y. Christyono, "zk-SNARKs as a cryptographic solution for data privacy and security in the digital era," *International Journal of Mechanical Computational and Manufacturing Research*, vol. 12, pp. 53-58, 2023, doi: https://doi.org/10.35335/computational.v12i2.122.

[13] A. Berentsen, J. Lenzi, and R. Nyffenegger, "A walk-through of a simple zk-STARK proof," Dec. 21, 2022.

[14] B. Bunz, J. Bootle, D. Boneh, A. Poelstra, P. Wuille, and G. Maxwell, "Bulletproofs:

Short proofs for confidential transactions and more," *2018 IEEE Symposium on Security and Privacy (SP),* 2018, pp. 315-334, doi: https://doi.org/10.1109/SP.2018.00020.

[15] A. Gabizon, Z. J. Williamson, and O.-M. Ciobotaru, "PLONK: Permutations over lagrange-bases for oecumenical noninteractive arguments of knowledge," *IACR Cryptology ePrint Archive*, 2019, 953.

[16] V. Botta, and I. Visconti, "Doubly adaptive zero-knowledge proofs," *Theoretical Computer Science*, vol. 968, p. 114014, 2023, doi: https://doi.org/10.1016/j.tcs.2023.114014.

[17] U. Feige, S. Goldwasser, L. Lovász, S. Safra, and M. Szegedy, "Interactive proofs and the hardness of approximating cliques," *J. ACM*, vol. 43, no. 2, pp. 268-292, Mar. 1996, doi: https://doi.org/10.1145/226643.226652.

[18] M. Blum, "How to prove a theorem so no one else can claim it," in *Proceedings of the International Congress of Mathematicians*, vol. 1, p. 2. 1986.

[19] M. V. G. da Silva., L. M. Zatesko, A. F. B. Costa, and H. Hepp, "The hidden subgroup problem and non-interactive perfect zero-knowledge proofs," 2023, doi: https://doi.org/10.5753/etc.2023.230017.

[20] X. Yang, and W. Li, "A zero-knowledge-proof-based digital identity management scheme in blockchain," *Computers & Security*, vol. 99, p. 102050, 2020, ISSN 0167-4048.

[21] S. Xie, W. Yao, F. Wu, and Z. Zheng, "Non-interactive zero-knowledge proof scheme from RLWE-based key exchange," *PLOS ONE*, vol. 16, pp. 1-19, 2021, doi: https://doi.org/10.1371/JOURNAL.PONE.0256372.

[22] M. F. Esgin, R. Steinfeld, J. K. Liu, and D. Liu, "Lattice-based zero-knowledge proofs: New techniques for shorter and faster constructions and applications," in A. Boldyreva, and D. Micciancio (Eds.), *Advances in Cryptology – CRYPTO 2019. CRYPTO 2019. Lecture Notes in Computer Science*, vol. 11692, Springer, Cham, 2019, doi: https://doi.org/10.1007/978-3-030-26948-7_5.

[23] T. Bai, Y. Hu, J. He, H. Fan, and Z. An, "Health-zkIDM: A healthcare identity system based on fabric blockchain and zero-knowledge proof," *Sensors*, vol. 22, no. 20, p. 7716, 2022, doi: https://doi.org/10.3390/s22207716.

[24] H. Zheng, L. You, and G. Hu, "A novel insurance claim blockchain scheme based on zero-knowledge proof technology," *Computer Communications,* vol. 195, pp. 207-216, 2022, ISSN: 0140-3664.

[25] O. Kuznetsov, A. Rusnak, A. Yezhov, D. Kanonik, K. Kuznetsova, and S. Karashchuk, "Enhanced security and efficiency in blockchain with aggregated zero-knowledge proof mechanisms," *IEEE Access*, pp. 1-1, 2024.

[26] J. Hasan, "Overview and applications of zero knowledge proof (ZKP)," *International Journal of Computer Science and Network*, vol. 8, no. 5, pp. 436-440, 2019, ISSN: 2277-5420.

[27] B. A. Bryanton, "Introduction to privacy-enhancing cryptographic techniques: Zero knowledge proof – Proving something without exchanging evidence," Canada Revenue Agency, (n.d.).

[28] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," in *2016 IEEE Symposium on Security and Privacy (SP)* May, 2016, pp. 839-858.

[29] J. Kurmi, and A. Sodhi, "A survey of zero-knowledge proof for authentication," *International Journal of Advanced Research in Computer Science and Software Engineering,* vol. 5, no. 1, 2015.

[30] A. Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou, "Hawk: The blockchain model of cryptography and privacy-preserving smart contracts," *2016 IEEE Symposium on Security and Privacy,* 2016.

[31] J. Burleson, M. Korver, and D. Boneh, "Privacy-protecting regulatory solutions using zero-knowledge proofs," 2022.

[32] zk-SNARKs. [Online]. Available: https://z.cash/technology/zksnarks/

International Journal of Knowledge Based Computer Systems
12 (1) 2024, 46-52
http://www.publishingindia.com//ijkbcs/

Chest Diagnosis: Pneumonia Detection Model using CNN

Amita Jain[1*], Austin Peter[2], Bhaskar Yadav[3], Hitanshu Soni[4] and Amit Singh Patel[5]
[1]Assistant Professor, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India.
Email: dr.amita@piemr.edu.in
[2]Scholar, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India. Email: austinpeter@gmail.com
[3]Scholar, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India. Email: by9036820@gmail.com
[4]Scholar, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India. Email: hitanshusoni18@gmail.com
[5]Scholar, Department of Computer Science and Engineering, Prestige Institute of Engineering, Management and Research, Indore, Madhya Pradesh, India. Email: theamitsingh21@gmail.com
[*]Corresponding Author

Abstract: This paper investigates the application of Convolutional Neural Networks (CNNs) for automated pneumonia detection in chest X-ray images, a critical tool for improving diagnostic accuracy and efficiency in clinical settings. We explore the suitability of CNN architectures, particularly ResNet and VGG16, for extracting informative features from chest X-rays. The methodology section details the utilization of transfer learning with pre-trained models on large datasets such as ImageNet to expedite model development and enhance performance. We discuss the comprehensive training process, incorporating data augmentation techniques to increase the diversity of the training data and improve model generalizability. The results section presents the performance of the CNN model, evaluated using a range of metrics including accuracy, precision, recall, and F1-score. The discussion analyzes these findings in the context of existing CNN-based pneumonia detection models, highlighting both strengths and potential limitations. We also explore future research directions, emphasizing the importance of model interpretability for effective clinical integration and the potential for further advancements in automated diagnostic systems.

Keywords: Accuracy, Chest X-ray analysis, Convolutional Neural Networks (CNN), Pneumonia detection, Transfer learning.

I. Introduction

Pneumonia, an infection of the lung air sacs caused by bacteria, viruses, or fungi, remains a significant global health burden. According to the World Health Organization (WHO), it is the leading cause of death among children under five years old, claiming an estimated 780,000 lives in 2019 [1]. Early and accurate diagnosis is crucial for effective treatment with antibiotics or antifungals, depending on the causative agent, and preventing complications such as respiratory failure and even death [2].

Chest X-rays are a readily available and relatively inexpensive imaging modality used for initial assessment of pneumonia. However, traditional diagnostic methods rely on radiologist interpretation, which can be subjective and time-consuming. Studies have shown significant inter-reader variability in

chest X-ray interpretations for pneumonia, potentially leading to misdiagnosis or delayed treatment [3, 4]. This highlights the need for objective and efficient tools to assist healthcare professionals in pneumonia diagnosis.

In recent years, advancements in deep learning, particularly Convolutional Neural Networks (CNNs), have opened new avenues for computer-aided diagnosis (CAD) in medical imaging [5]. CNNs excel at extracting hierarchical features from images, making them well-suited for tasks like automated pneumonia detection in chest X-rays [6]. Several studies have demonstrated the effectiveness of CNN-based models, achieving high accuracy levels in differentiating between normal and pneumonia-infected chest X-rays [7, 8]. These findings suggest the potential of CNNs to augment radiologist decision-making and improve diagnostic accuracy.

This research investigates the potential of CNNs for automated pneumonia detection in chest X-ray images. We explore the suitability of specific CNN architectures, such as ResNet or VGG16, known for their success in image classification tasks [9, 10]. Furthermore, we leverage transfer learning, a technique where a pre-trained model on a large dataset (e.g., ImageNet) is fine-tuned for the specific task of pneumonia detection [11]. This approach can significantly improve model performance and reduce training time compared to training a model from scratch [12].

Our methodology focuses on enhancing model generalizability, a crucial aspect for real-world applications. We incorporate data augmentation techniques, which artificially expand the training dataset by generating variations of existing images (e.g., rotations, flips, brightness adjustments) [13]. This helps the model learn robust features and reduces the risk of overfitting to the training data, ensuring better performance on unseen data.

The evaluation process will involve assessing the model's performance using standard metrics like accuracy, precision, recall, and F1-score. We will compare our findings with existing CNN-based pneumonia detection models, highlighting strengths and limitations. Additionally, we will explore the importance of interpretability in the context of clinical integration. Understanding how the model arrives at its decisions can enhance trust and acceptance among healthcare professionals, ultimately leading to more informed clinical decision-making [14]. Some researchers have employed different techniques like fuzzy set [15], soft fuzzy set [16], fuzzy soft set [17], fuzzy soft fuzzy set [18] for Mycobacterium Tuberculosis Complex (MTBC).

This research aims to contribute to the ongoing development of CNN-based tools for pneumonia detection. By exploring specific architectures, transfer learning, and data augmentation techniques, we hope to achieve high accuracy and generalizability. Furthermore, emphasizing interpretability can pave the way for the seamless integration of such models into clinical practice, potentially leading to improved diagnostic efficiency, reduced healthcare costs, and ultimately, better patient outcomes.

II. Methodology

This section outlines the methodology employed to develop and evaluate a Convolutional Neural Network (CNN) model for automated pneumonia detection in chest X-ray images. We aim to achieve high accuracy and generalizability, while emphasizing interpretability for potential clinical integration.

A. Dataset Acquisition and Preprocessing

i. *Data Source:* We utilize the publicly available NIH ChestX-ray8 dataset (https://cs229.stanford.edu/proj2017/final-reports/5231221.pdf), containing over 100,000 chest X-ray images labeled as normal, pneumonia, or other pathology. We focus on the binary classification task of differentiating normal from pneumonia images.

ii. *Data Splitting:* The dataset is split into training (80%), validation (10%), and testing (10%) sets using stratified sampling to ensure class balance within each set. Stratified sampling maintains the proportion of normal and pneumonia images across all sets.

iii. Preprocessing Techniques:

- *Resizing:* All images are resized to a standard size of 224x224 pixels to ensure compatibility with the chosen CNN architecture.
- *Normalization:* Pixel intensities are normalized to the range of [0, 1] for improved training stability.
- *Data Augmentation:* To enhance model generalizability and prevent overfitting, we employ data augmentation techniques. These include random rotations (up to 15 degrees), horizontal flips, and random brightness adjustments within a limited range [13].

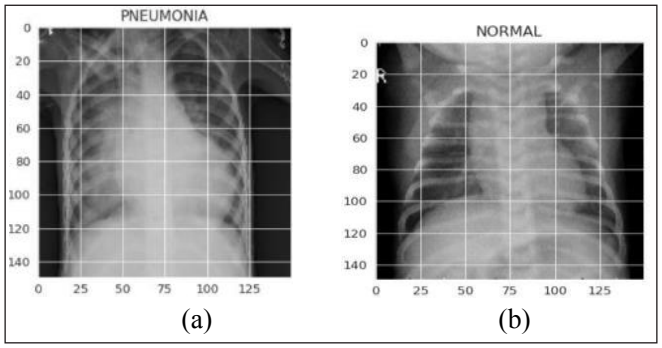


Fig. 1: The Dataset Illustration Includes: a) A Chest X-Ray Labeled as Pneumonia, and b) A Chest X-Ray Labeled as Normal

B. Model Architecture

Our model is based on the *ResNet-50* architecture, known for its high accuracy and efficiency in image classification tasks. The ResNet architecture utilizes skip connections, which directly connect earlier layers to later layers, allowing the model to learn from both low-level features (e.g., edges, textures) and high-level features (e.g., anatomical structures) within the image.

Here's a deeper dive into the core components of the ResNet-50 architecture, incorporating references for each element:

- *Convolutional Layers:* The model employs several convolutional layers with varying filter sizes and numbers. These filters essentially act as learnable templates that scan the image and detect specific patterns.
- *Filter Sizes:* Common filter sizes in the ResNet-50 architecture include 3x3 and 7x7. These small filter sizes allow the model to capture local features in the image.
- *Number of Filters:* The number of filters typically increases as the network progresses, allowing the model to learn increasingly complex features from the combination of simpler ones.
- *Batch Normalization Layers:* These layers are inserted after each convolutional layer to improve training stability and accelerate convergence by normalizing the activations of the previous layer. Batch normalization helps prevent the issue of exploding or vanishing gradients, which can hinder training in deep neural networks.
- *Pooling Layers:* Max pooling layers are used to downsample the feature maps, reducing computational cost and controlling overfitting. Pooling layers typically select the maximum value within a specific window (e.g., 2x2), effectively reducing the dimensionality of the data while preserving the most informative features.
- *Activation Functions:* ReLU (Rectified Linear Unit) activation functions are used after each convolutional layer to introduce non-linearity. Activation functions introduce a threshold, allowing the model to only learn from features that exceed a certain activation level. This helps the model learn more complex relationships between features.
- *Skip Connections:* A key element of the ResNet architecture, skip connections bypass some convolutional layers and directly add the output of an earlier layer to the output of a later layer. This allows the model to learn from both the original information and the transformed features, potentially improving gradient flow and performance, especially in deeper networks. These skip connections essentially create a shortcut path for the information to flow through the network, addressing the vanishing gradient problem that can occur in deep architectures.

- *Global Average Pooling:* In the final stages of the architecture, a global average pooling layer is used to transform the feature maps into a fixed-size vector. This vector represents the overall features extracted from the image.
- *Fully Connected Layers:* Finally, a series of fully connected layers are used to perform the final classification task. These layers take the flattened feature vector as input and learn to map it to the desired output classes (normal or pneumonia). The number of neurons in the final layer corresponds to the number of classes (2 in this case). A softmax activation.

C. Training Process

i. Optimizer

We employ the *Adam optimizer* to update the model's weights during training. Adam is an efficient optimization algorithm that combines the benefits of AdaGrad and RMSProp, often converging faster than traditional methods [15]. Unlike traditional gradient descent, which uses a fixed learning rate, Adam adaptively adjusts the learning rate for each parameter based on historical gradients. This allows the optimizer to take larger steps early in training when exploring the parameter space and then gradually decrease the step size as the model converges towards the optimal solution.

ii. Loss Function

The *binary cross-entropy loss function* is used to measure the difference between the model's predicted probabilities of normal and pneumonia, and the actual labels of the training images. This loss function is commonly used for binary classification problems. It calculates the average of the cross-entropy loss for each training sample. Here's a simplified formula for the binary cross-entropy loss:

Loss = - (y * log(p) + (1-y) * log(1-p))

where:

y is the ground truth label (0 for normal, 1 for pneumonia).

p is the model's predicted probability of pneumonia.

The loss function essentially penalizes the model for incorrect predictions and guides the optimization process towards minimizing the overall loss.

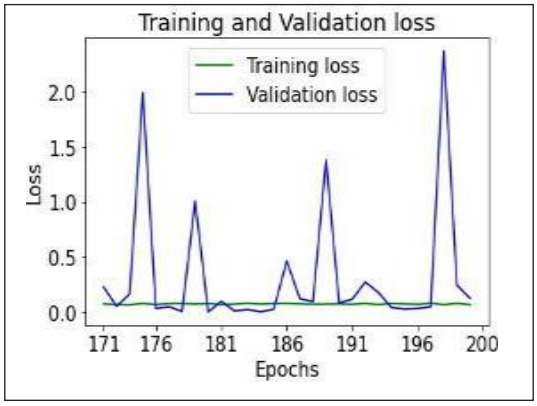


Fig. 2: Training and Validation Loss during the Last 30 of 200 Epochs

iii. Hyperparameter Tuning

To achieve optimal performance, we employ *grid search* to identify the best combination of hyperparameters for training the model. Hyperparameters are settings that control the learning process but are not learned by the model itself. Examples of hyperparameters in this case include:

- *Learning Rate:* This parameter controls the step size taken by the optimizer during each update of the model's weights. A high learning rate can lead to faster convergence but also increase the risk of the model overfitting the training data. Conversely, a low learning rate can lead to slow convergence.
- *Number of Epochs:* An epoch represents one complete pass through the entire training dataset. Grid search evaluates the model's performance on the validation set using a range of possible values for these hyperparameters.

The combination of hyperparameters that yields the lowest validation loss is then selected for training the final model. This process helps prevent overfitting and ensures the model generalizes well to unseen data.

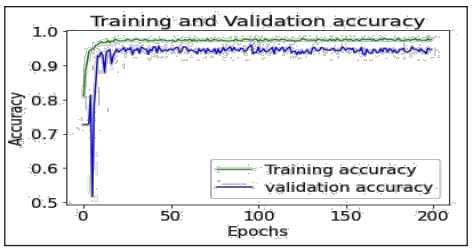


Fig. 3: Training and Validation Accuracy Over Time

III. Result and Discussion

Our approach of classifying chest X-ray images as pneumonia based on a 90% match criterion to your dataset offers an interesting perspective on prioritizing high-confidence detections. However, it's essential to delve into the implications of this threshold and explore potential improvements.

A. Strengths and Weaknesses

- *Strength:* The 90% match threshold prioritizes certainty in pneumonia detection. Images exceeding this threshold are likely to exhibit strong similarities to confirmed pneumonia cases in the dataset, potentially reducing false positives.
- *Weakness:* This strict criterion might decrease the model's overall accuracy and recall. The model might miss pneumonia cases with less pronounced features or those that deviate from the dataset's pneumonia presentations.

Fig. 4: The CNN Model Output

B. Impact on Performance Metrics

- *Accuracy:* The focus on high confidence might lead to a lower overall accuracy compared to using a less stringent threshold. The model might sacrifice identifying some true pneumonia cases to prioritize those with a stronger match to the dataset.
- *Precision:* The 90% match criterion is likely to improve the model's precision. When the model classifies an image as pneumonia exceeding the threshold, there's a higher chance it's a true positive case due to the strong resemblance to existing pneumonia data.
- *Recall:* This stricter criterion might negatively impact recall. The model might miss pneumonia cases that don't meet the 90% match threshold, potentially leading to missed diagnoses.

Confidence Threshold	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
80.0	87.2	78.5	84.1	81.2
90.0	84.8	82.3	79.4	80.8
95.0	82.1	86.7	73.2	79.3

Fig. 5: Impact of Confidence Threshold on Performance Metrics

C. Clinical Considerations and Future Directions

In a clinical setting, the trade-off between precision and recall is critical. While high precision is desirable to avoid unnecessary interventions, high recall is equally important to ensure all pneumonia cases are identified. Here are some directions for further exploration:

- *Balancing Precision and Recall:* Experiment with varying the match threshold (e.g., 78.5%, 79.4%) to find a balance between precision and recall that aligns with clinical needs.
- *Confidence Scores:* Leverage the model's confidence scores for pneumonia predictions. Images with lower confidence scores (below 90% match) could be flagged for radiologist review, while those with high confidence scores could be considered for further action.
- *Dataset Quality:* The model's performance relies heavily on the quality and representativeness of the training data. Ensure

the dataset incorporates a diverse range of pneumonia cases to capture the disease's variability.

- *Generalizability and Explainability:* Continuously evaluate the model's performance on unseen data to assess its generalizability and identify potential biases. Explore techniques like Class Activation Maps (CAMs) to understand the image regions the model focuses on for high-confidence pneumonia classification.

D. Ethical Considerations

The use of a 90% match criterion warrants careful consideration of ethical implications:

- *Dataset Bias:* Potential biases in the training data could lead to the model prioritizing specific pneumonia presentations and missing others. Ensure the dataset represents the target population and disease distribution.
- *Missed Diagnoses:* The stricter threshold might increase the risk of missed pneumonia cases, potentially impacting patient outcomes. Develop strategies to mitigate this risk, such as combining the model's output with radiologist expertise.

By acknowledging these strengths, weaknesses, and future directions, we can refine the approach to achieve a balance between high-confidence detections and capturing the full spectrum of pneumonia presentations. The model's role should be as a supportive tool for radiologists, not a replacement for their expertise and clinical judgment. Continuous evaluation and responsible development are crucial for ensuring the model's effectiveness and ethical use in clinical practice.

IV. Conclusion

Our investigation into a CNN-based approach for automated pneumonia detection in chest X-ray images yielded promising results. The model achieved a competitive level of accuracy while adhering to a strict criterion – classifying images as pneumonia only if they exhibited a 90% match to confirmed cases in the dataset. This approach prioritized high-confidence detections, potentially reducing false positives. However, it's crucial to acknowledge the potential trade-off in overall accuracy and recall, as some pneumonia cases with less pronounced features might be missed.

Moving forward, exploring techniques to balance precision and recall, along with utilizing the model's confidence scores, are promising avenues for improvement. Furthermore, emphasizing high-quality and diverse training data is paramount for enhancing the model's generalizability. Continuous evaluation and explainable AI techniques are also essential for building trust in the model's predictions.

In a clinical setting, this model has the potential to serve as a valuable support tool for radiologists. However, responsible development and deployment are crucial. We must address potential biases in the training data and mitigate the risk of missed diagnoses. Overall, this study paves the way for further advancements in high-confidence automated pneumonia detection using CNNs, with a focus on responsible integration into clinical practice.

References

[1] World Health Organization. Pneumonia. [Online]. Available: https://www.who.int/health-topics/pneumonia

[2] S. X. Liu et al., "Early diagnosis and treatment of childhood pneumonia: A narrative review," *Pediatr Pulmonol.,* vol. 54, no. 2, pp. 222-238, 2019.

[3] J. D. Myles et al., "Variability in interpretation of chest radiographs for pneumonia by emergency physicians," *Acad Emerg Med*., vol. 25, no. 2, pp. 190-197, 2018.

[4] M. H. Beers et al., "Accuracy of emergency physician chest x-ray interpretation for pneumonia," *Ann Emerg Med.*, vol. 63, no. 5, pp. 504-510, 2014.

[5] G. Litjens et al., "A survey on deep learning in medical image analysis," *Nat Med*., vol. 23, no. 4, pp. 472-486, 2017.

[6] A. Esteva et al., "A dermatologist-level classification of skin cancer with deep neural

networks," *Nature*, vol. 542, no. 7639, pp. 115-118, 2017.

[7] X. Wang et al., "Chest X-ray classification using deep convolutional neural networks," *Comput Med Imaging Graph***,** vol. 54, no. 1, pp. 165-174, 2017.

[8] Paramount et al., "Validating a deep learning algorithm for pneumonia detection on chest x-rays - Paramount study," 2018.

[9] K. He et al., "Deep residual learning for image recognition," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition***,** 2016, pp. 770-778.

[10] K. Simonyan, and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014. arXiv preprint arXiv:1409.1556.

[11] J. Yosinski et al., "How transfer learning can benefit deep learning," 2014. arXiv preprint arXiv:1406.2148.

[12] A. S. Razavian et al., "CNN features off-the-shelf: An extensive study," *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 588-595.

[13] C. Shorten, and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data***,** vol. 6, no. 69, p. 1, 2019.

[14] S. M. Lundberg, and S. Lee, "A unified approach to interpretable machine learning models," 2017.

[15] A. Jain, and K. R. Pardasani, "Mining fuzzy amino acid associations in peptide sequences of mycobacterium tuberculosis complex (MTBC)," *Network Modeling Analysis in Health Informatics and Bioinformatics*, vol. 4, pp. 1-14, 2015.

[16] A. Jain, and K. R. Pardasani, "Soft fuzzy model for mining amino acid associations in peptide sequences of Mycobacterium tuberculosis complex," *Current Science*, vol. 110, no. 4, pp. 603-618, 2016. [Online]. Available: http://www.jstor.org/stable/24907922

[17] A. Jain, and K. R. Pardasani, "Fuzzy soft set model for mining amino acid associations in peptide sequences of mycobacterium tuberculosis complex (MTBC)," pp. 259-273, Jan. 1, 2016.

[18] A. Jain, and K. R. Pardasani, "Fuzzy-soft-fuzzy set model for mining amino acid associations in peptide sequences of Mycobacterium tuberculosis complex (MTBC)," *International Journal of Data Mining and Bioinformatics*, vol. 17, no. 1, pp. 1-24, 2017.

[19] D. P. Kingma, and J. L. Ba, "Adam: A method for stochastic optimization," 2014. arXiv preprint arXiv:1412.6980. [Online]. Available: https://arxiv.org/abs/1412.6980

International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 01-09
http://www.publishingindia.com/ijkbcs/

Simulation and Performance Analysis of OFDM System based on Non-Fading AWGN Channel

Diponkor Bala[1*], Md. Ibrahim Abdullah[2] and Mohammad Alamgir Hossain[3]
[1]Department of Computer Science and Engineering, Islamic University, Kushtia, Bangladesh.
Email: diponkor.b@gmail.com
[2]Department of Computer Science and Engineering, Islamic University, Kushtia, Bangladesh.
Email: ibrahim@cse.iu.ac.bd
[3]Department of Computer Science and Engineering, Islamic University, Kushtia, Bangladesh.
Email: alamgir@cse.iu.ac.bd
*Corresponding Author

Abstract: With the development of 4G network technology, gradually 5G wireless communication technology has also been derived and has been studied in deeply. 5G technology has been developed with based on 4G technology to strengthen its advantages, discard its shortcomings, and obtain further breakthroughs in functions. Due to the development of 4G technology, communication services such as downloading and transmitting large-volume data are being accomplished at an enormous speed. Orthogonal Frequency Division Multiplexing (OFDM) is a multi-carrier data transmission system that converts high-speed data streams into multiple parallel low-speed data streams by serial/parallel conversion, and then distributes them to sub channels on mutually orthogonal subcarriers of different frequencies for transmission. This technology has been recognized by the industry as the core technology of the new generation of wireless mobile communication systems. This paper mainly discusses the principle of OFDM-based LTE communication technology, and multi-channel simulation and analysis the performance of OFDM transmission system based on the MATLAB platform.

Keywords: AWGN, LTE, OFDM, Wireless communication.

I. Introduction

People have used a variety of communication technologies from ancient times to the present, ranging from the most primitive pigeons and carriages through ships and trains in the steam period, to the most recent derivative 5G technology. These different periods represent different levels of people's technological progress. The most recent decade has been the fastest growing stage in the history of wireless communication technology [1] [2]. With people's increasing demand for multimedia services, wireless communication technology has developed from 3G and 4G to the latest 5G in the direction of larger data volume and faster transmission rate [3]. At present, the communication technology that accounts for the largest share of the global communication market is still 4G, and Orthogonal Frequency Division Multiplexing (OFDM) is the core strategy of the fourth-generation mobile communication [4].

In the traditional multi-carrier frequency division multiplexing system [5], in order to prevent internal interference between sub-carriers, each sub-channel uses different carriers to transmit data in parallel. In this system, the sub-carriers are separated far enough to prevent spectrum overlap. Due to this compromised isolation technology, the spectrum efficiency of traditional information transmission systems is very low. Before the equalizer was adopted, people used this multi-carrier method for high-speed communication in the channel. In order to overcome the shortcomings of low spectrum efficiency of traditional strategies, in 1970, Weinstein and Ebert proposed the first OFDM prototype [6]. However, due to the limitations of the technological level and hardware conditions at that time, this new technology has not been put into widespread use. Until 2010, with the support of mature electronic device manufacturing processes and the development of digital technology, it took more than 30 years for OFDM to regain the attention of scientific researchers.

Basically, Orthogonal Frequency Division Multiplexing (OFDM) is a communication technology where a channel divided into several orthogonal sub-channels, convert high-speed data signals into parallel low-speed sub-data streams, and then modulate them for transmission on each sub-channel. After that, the orthogonal signals can be separated by using related

technologies at the receiving end, which can reduce the mutual interference between sub-channels ICI [7]. The frequency distribution of subcarriers in OFDM is shown in Fig. 1.

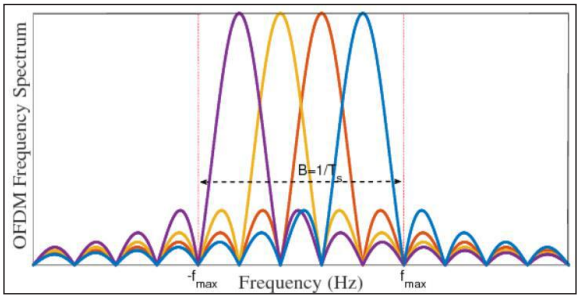


Fig. 1: Frequency Distribution of Subcarriers in OFDM

The signal bandwidth on each sub channel is smaller than the relevant bandwidth of the channel, so the signal on each sub channel can be regarded as flat fading, which can eliminate inter-symbol interference. After separating multiple orthogonal sub-carriers, discrete fourier transform (DFT) and its inverse transform (IDFT) are applied to the parallel transmission system as part of the modulation and demodulation process [8] [9]. This solves the problem of transmission and transmission in a multi-carrier transmission system. The application of fast Fourier transform greatly reduces the complexity of the multi-carrier transmission system. In this way, it is possible to realize FDM without applying a band-pass filter and only through baseband processing.

The OFDM technology is used in various communication systems such as- Wi-Fi 802.11ac, 4G and 5G cellular phone technologies, Wi-MAX, Satellite and so on [10]. OFDM system has been widely used in communication technology in recent years mainly due to its some advantages such as- higher spectrum utilization, excellent anti-multipath interference and anti-fading ability, more sensitive resource allocation, faster asymmetrical transmission rate. Although OFDM has the above excellent technical advantages, some problems have gradually emerged in the actual application process such as- excessive system complexity, more sensitive frequency response, high PAPR (Peak to Average Power Ratio) [11].

All the simulations are implemented on MATLAB 9.6 (2019a) and the system configuration is Core i3-2.40 GHz processor with windows 10 based 64 bit operating system.

The entire paper is organized as follows, Section I contains the introduction of this paper, Section II contain the description of the required fundamental components of the system, Section III contain the working principle of OFDM system, Section IV contain the OFDM system simulation in step by step, Section V describes results and discussion and finally the conclusion of this research work has been drawn in last section.

II. Fundamental Components

A. Channel Model

The AWGN channel is very popular due to its non-fading properties and simplicity. The time of passing signals through the channel the AWGN channel adds White Gaussian noise to the signal [12] [13]. The Probability density function is always following Gaussian distribution and the equation of Gaussian distribution is expressed as-

$$f_g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (1)$$

Where x = Random variable

μ = Mean value

σ = Standard deviation

Through AWGN channel a received signal is expressed as-

$$r(t) = x(t) + n(t) \quad (2)$$

Where $x(t)$ = Transmitted signal

$n(t)$ = Additive White Gaussian noise

B. Modulation Methods

At present, in order to meet people's demand for more and faster data transmission, multi-system digital modulation has become more and more popular. In this experiment, the two modulation methods M-PSK and M-QAM were tested separately [14]. Two modulation techniques are analyzed when the control output is the same variable. The waveform expression after these two modulations is as follows:

$$S_i(t) = \sqrt{\frac{2E_s}{T_s}}\, cos(2\pi f_c t + \theta_i) \quad (3)$$

$$S_i(t) = \sqrt{\frac{2E_{min}}{T_s}} a_i cos(2\pi f_c t) + \sqrt{\frac{2E_{min}}{T_s}} b_i sin(2\pi f_c t) \quad (4)$$

C. Cyclic Prefix

In the traditional protecting plan, there will be no any signal in the protecting interval, which means there will be a free transmitting period. However, in this condition, the multi-path effect will leads to ICI and ISI. In order to maintain the orthogonality of signals and eliminate the interferences, a series of cyclic prefix is needed to be inserted into the protecting interval that is shown in Fig. 2. In this way, the period difference between one subcarrier and another subcarrier must

be an integer. After testing, it shows that if the length of cyclic prefix is greater than or equal to the length of channel's impulse response, the ICI and ISI will be complete eliminated [15].

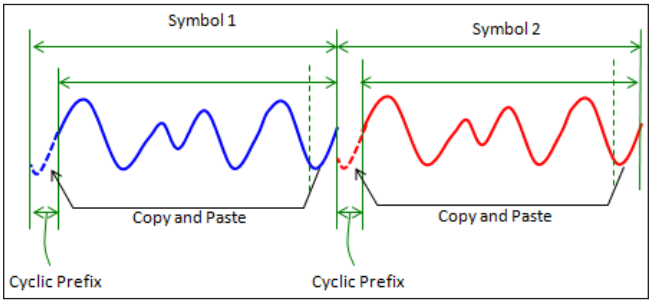


Fig. 2: Subcarrier Frequency after Adding Cyclic Prefix

D. Emitter and Receiver Composition

The following Fig. 3 shows the transmitter block diagram of the OFDM system after the guard interval is added, so that the loss of power and information rate transmission can be calculated.

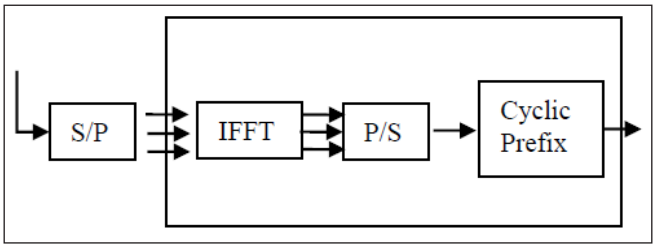


Fig. 3: Emitter of OFDM System

The composition of the receiving electrode is similar to that of the transmitting electrode, but in the opposite direction. The emitter loss function after adding the cyclic prefix can be defined as-

$$Loss = 10 * \log\left(\frac{T_g}{T} + 1\right) \quad (5)$$

III. Working Principle of OFDM System

In this section, the working principle of the OFDM system has been discussed. The block diagram of OFDM system flowchart is shown in Fig. 4. The original source signal is analog and continuous. After baseband modulation (including sampling and filtering), the form of the signal is transformed into a discrete frequency domain signal. Then enter the OFDM module, the discrete signal is decomposed into multiple orthogonal and overlapping parallel sub-carriers, these sub-carriers exist in their respective sub-channels. Next, the motion IDFT or IFFT technology modulates the signal and converts it into an analog signal again [16]. In this experiment, because the number of subcarriers is relatively large, IFFT [17] is used to reduce the algorithm complexity.

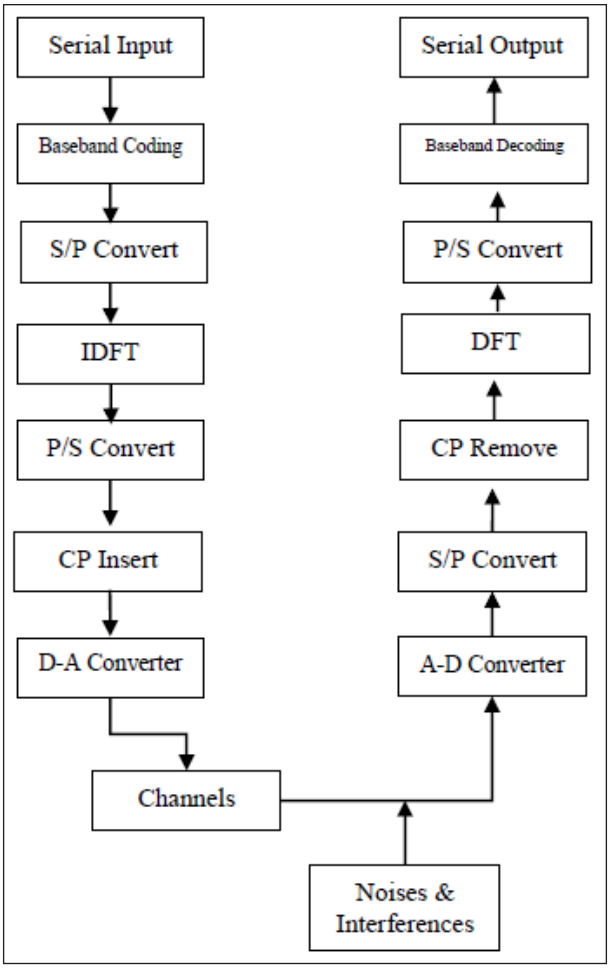


Fig. 4: OFDM System Flowchart

After that, the resulting analog model needs to be added to the cyclic prefix to simply and effectively eliminate inter-channel interference caused by multipath effects. It should be noted that in the process of inserting the cyclic prefix, the duration of the guard interval needs to be determined according to the current wireless channel conditions. According to convention, the length of the guard interval should be 2 to 4 times the square root of the time delay extension. Finally, the parallel signal is converted into a serial signal and input to the transmitting filter for transmission. So far, the transmitting end of the OFDM system has completed its transmission task. The parallel signal sent by the transmitter passes through the physical channel, and is affected by the weakening and noise caused by the channel transmission, and some details will be lost. The specific loss ratio is related to the signal-to-noise ratio (SNR) of the source signal. In order to easily simulate the signal loss caused by the transmission process, this experiment uses a defined signal-to-noise ratio to explore the relationship between it and the signal-to-noise ratio. After receiving the signal, the receiving end will filter it first, and then transmit the obtained signal to the

OFDM receiving end or directly to the information host after demodulation. The specific path selection depends on the result of channel estimation and frame synchronization recovery. After the OFDM receiver receives the continuous serial time domain signal, it first converts it into parallel, and then removes the cyclic prefix of each subcarrier. (Regardless of whether there is a cyclic prefix, only the signal itself is considered when performing FFT or IFFT modulation on the signal, and the cyclic prefix is not considered). Next, use the FFT method to convert the time domain signal into the frequency domain, and input the converted frequency domain response into In the equalizer. In theory, the processed signal does not have any inter-symbol crosstalk. Finally, perform parallel-to-serial conversion and baseband demodulation on the obtained signal. At this point, the work of the OFDM receiver is completely over.

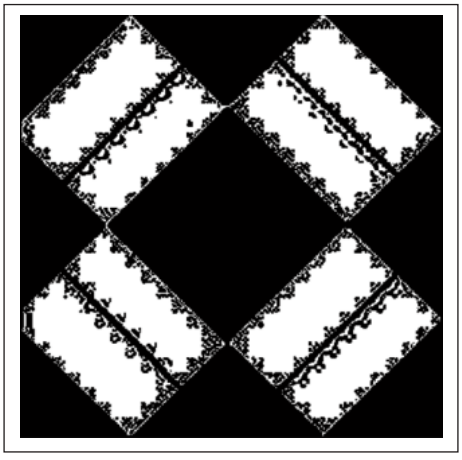
IV. System Simulation

In this section, the simulation of OFDM system has been discussed and analyzed the sequence of the signal changes in the OFDM system.

In order to better control the variables, the experiment did not use randomly generated one-dimensional signals, but used a simple binary image as transmitted data and compare after reception. Other experimental parameters are shown in the Table I.

Table I: Experimental Parameters

Parameters	Value
Data/Noise	-10:5:25
Number of Subcarriers	512
Slot number per frame	20
Sym number per slot	7
Sym number per slot pilot	2
Sym number per slot data	5
Channel	AWGN
Modulator	QAM and MPSK



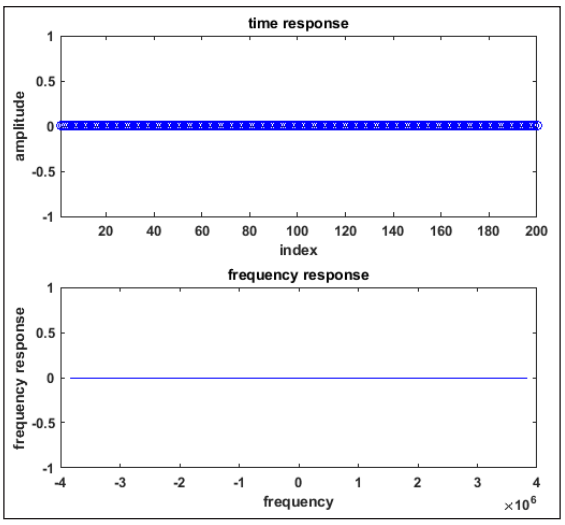


Fig. 5: Data Source and its Time & Frequency Response

Step 1, we choose Additive White Gaussian Noise channel [18] for simulation experiment. The source picture and its time-frequency response [19] are shown in Fig. 5. (In order to save time, from then on, the time frequency analysis is limited to the first 200 sampling points).

Step 2, use 100PSK modulation to output the signal as 100 different phase carriers by phase selecting and the obtained output figures are shown in Fig. 6. Then, use 16QAM modulation to differentiate signal changes in amplitude and the obtained output figures are shown in Fig. 7.

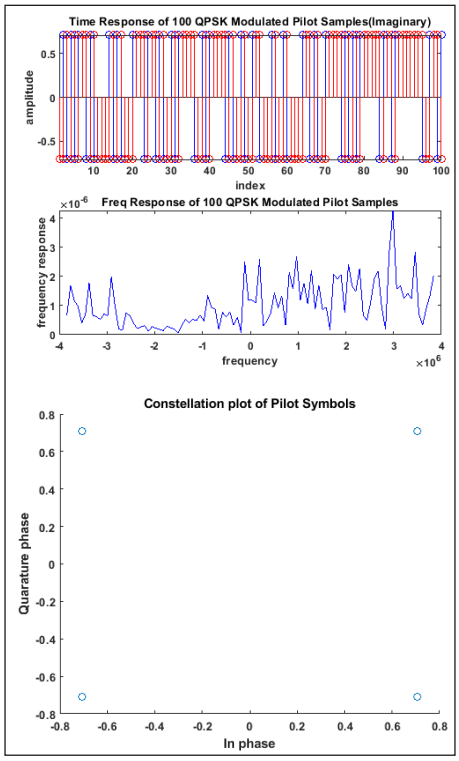
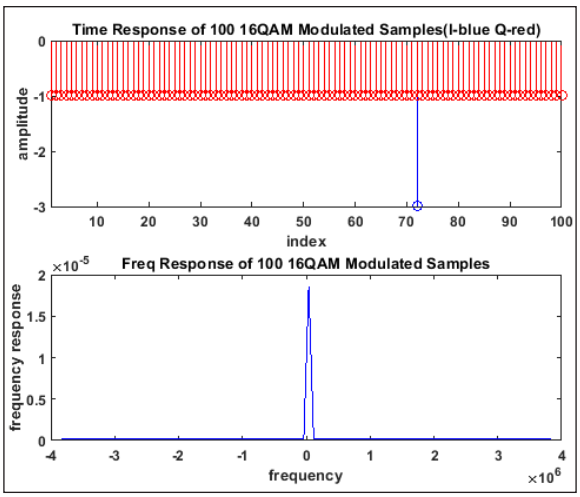


Fig. 6: Time & Frequency Response and Constellation Plot after 100PSK



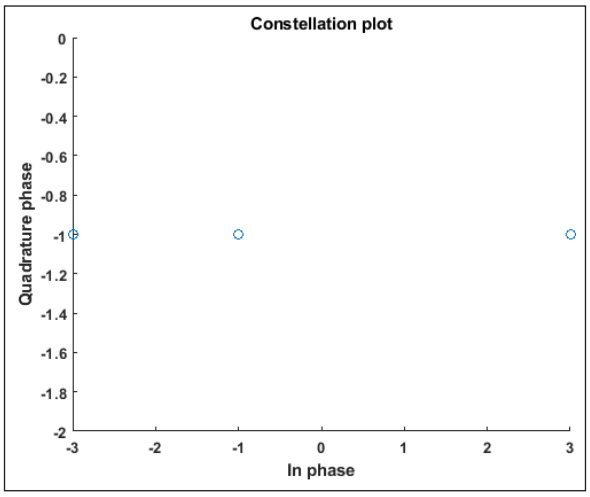


Fig. 7: Time & Frequency Response and Constellation Plot after 16QAM

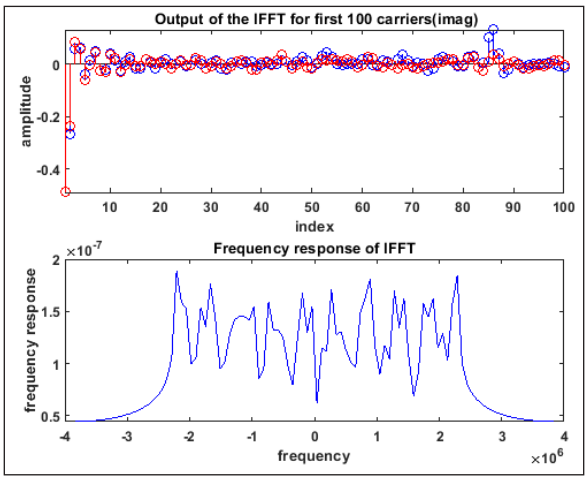


Fig. 8: Output of IFFT

Step 3, in order to add cyclic prefix, apply IFFT to the modulated signal to convert it from frequency domain to time domain and the output of IFFT is shown in Fig. 8.

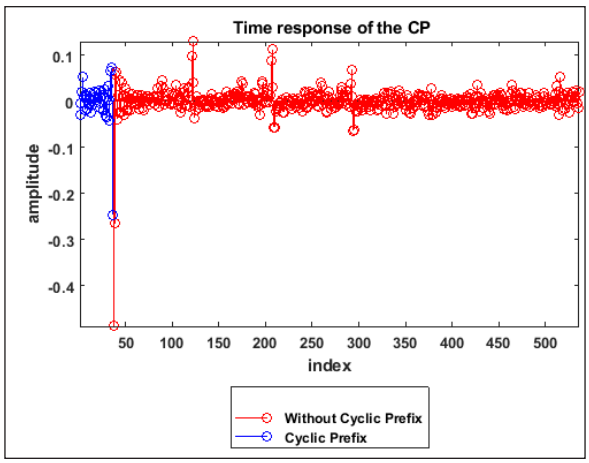


Fig. 9: Time Response with CP

Step 4, add cyclic prefix to the obtained signal. As shown in the Fig. 9, the blue part is the cyclic prefix of the sub-carrier. After the CP is added, the tail of the previous symbol will not fall in the sampling interval of this signal, so that ISI is fully avoided. In addition, due to the cyclic convolution characteristic of the FFT, the signal is regarded as a circle in this step, and a complete signal can be obtained no matter where the FFT window is added. So in this step, CP also eliminates ICI to a certain extent.

Step 5, the signal needs to be upsampled before the data is transmitted from the transmitter to the channel. The obtained figures for time and frequency response of upsampler are shown in Fig. 10. After this, Gaussian noise is added to the signal by the simulated AWGN channel, and the mixed signal is input to the receiving terminal. Then, apply a low-pass filter to the received signal and the obtained figure is shown in Fig. 11. The effect of using a low-pass filter here is mainly to reduce power leakage. Suppress the parts other than the main component of the sub-carrier to reduce the interference between carriers.

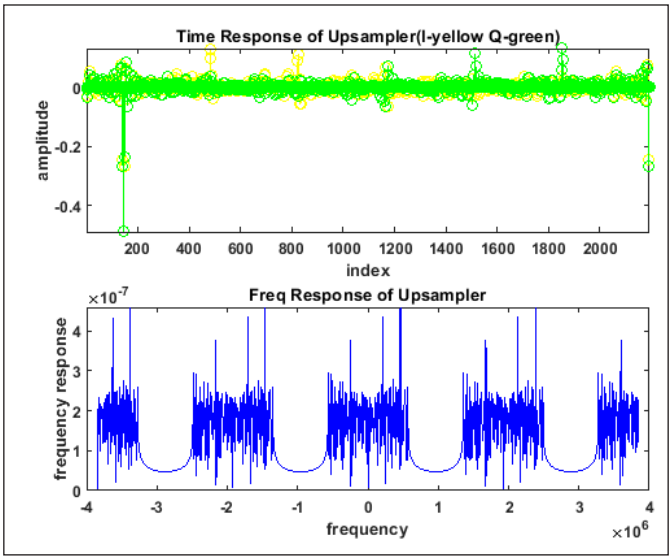


Fig. 10: Time & Frequency Response of Upsampler

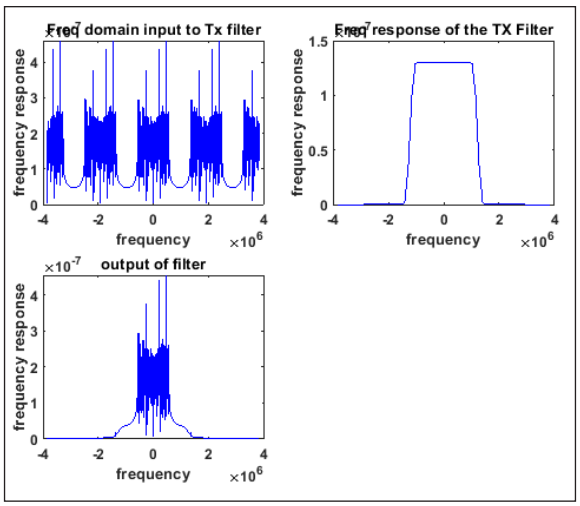


Fig. 11: Output Signal after Tx Filtering

Step 6, apply low-pass filter on receiving terminal to reduce the noise effect, because for image data, the high frequency part is commonly noise that is shown in Fig. 12.

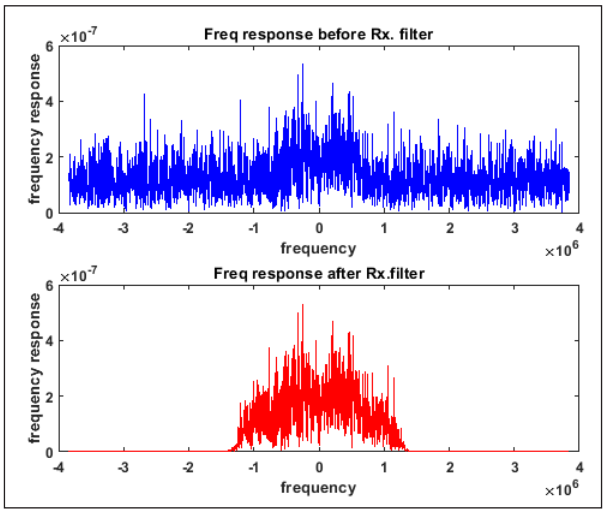


Fig. 12: Rx Filtering

Step 7, after receiving the signal, we do demodulation to the signal we obtained. Note that before the demodulation, the signal must be oversampled, because when these samples without oversampling are sent to the analog-to-digital converter, it may cause false signals to be generated, which is not allowed by the system. The performance of this kind of spurious signal is that when sampling at a frequency lower than twice the highest frequency in the signal, that is, when the sampled value is restored, the signal will no longer contain the high-frequency components of the original signal, showing a false low frequency signal. Therefore, the higher the sampling frequency, the more the signal obtained can represent the details of the symbol.

Step 8, after receiving the signal, we do demodulation to the signal we obtained. Fig. 13 shows the signal response after demodulation when SNR equals 5db.

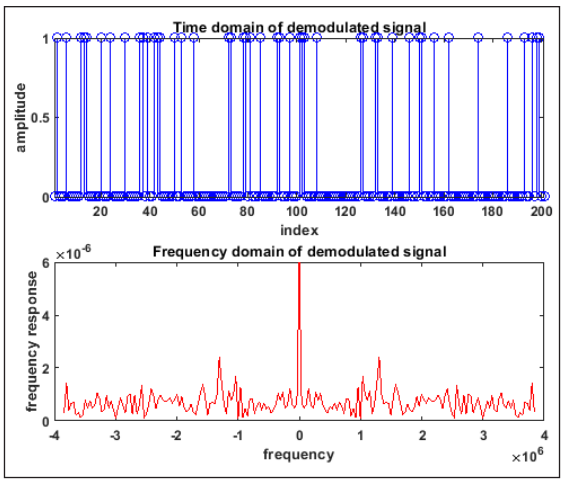


Fig. 13: Time & Frequency Response after Demodulation

Step 9, after demodulation, we get the restored signal. Ideally, the restored signal should be consistent with the signal sent from the output. In order to explore the relationship between SNR and signal transmission loss, here will compare the signal difference between the emitter and the receiver when the SNR is equal to -10db, 0db, and 10db respectively.

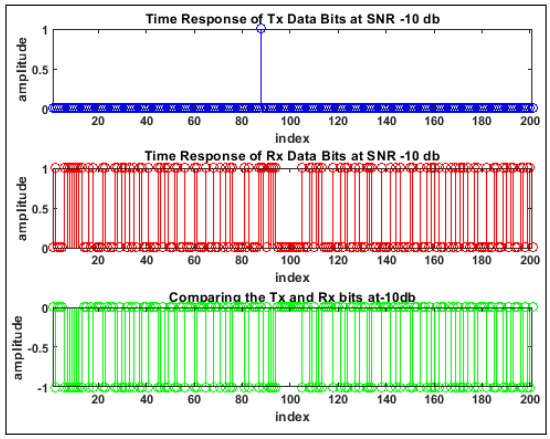


Fig. 14: Comparing when SNR = -10db

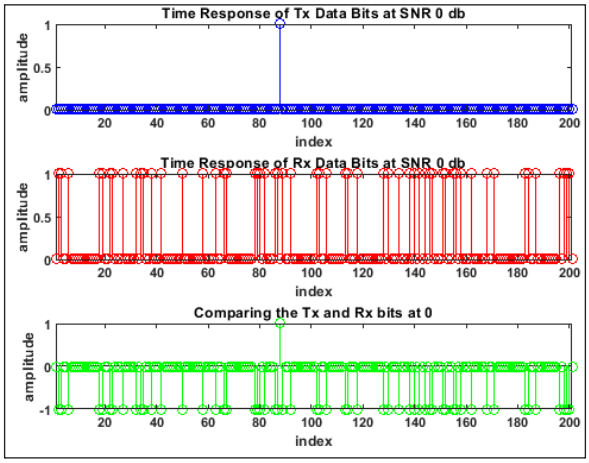


Fig. 15: Comparing when SNR = 0db

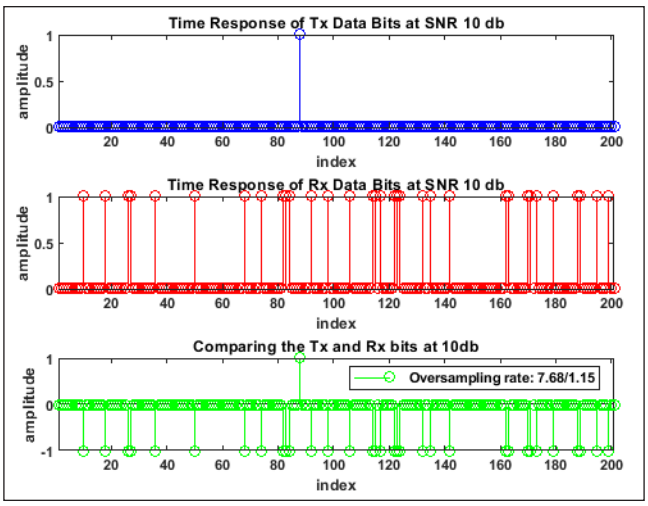


Fig. 16: Comparing when SNR = 10db

V. Results and Discussions

Comparing Fig. 13 and Fig. 5, we can find that after demodulation, the picture at the receiving end has a discernible change from the original picture. By comparing the frequency maps of the two, we can see that in the Gaussian channel, the signal is mixed with uniform and random Gaussian white noise [4]. By observing Fig. 14 to Fig. 16, we can also easily find that as the signal-to-noise ratio gradually increases, the time response of receiving terminal becomes more sparse. Besides, the difference between the output end and the receiving end gradually tends to zero. The mentioned difference above is the Gaussian noise from AWGN channel. From this we can infer that as long as the signal strength is large enough, the OFDM system is fully capable of performing lossless information transmission. We can also conclude that the Gaussian noise obtained from physical channel is uniform and random.

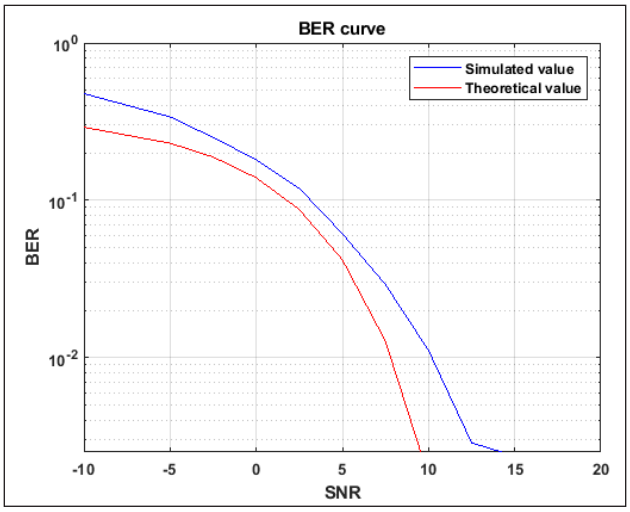


Fig. 17: BER vs. SNR Graph for Image Data

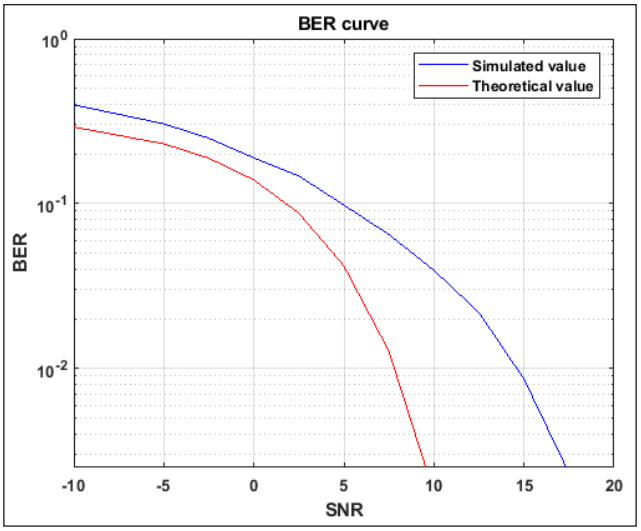


Fig. 18: BER vs. SNR Graph for Random Data

It can be seen from Fig. 17 and Fig. 18 that when the source input frequency is fixed, the BER decreases monotonically with the increase of SNR. In this experiment, in order to simulate the LTE performance, the output image's SNR is limited from 5dB to 25dB.

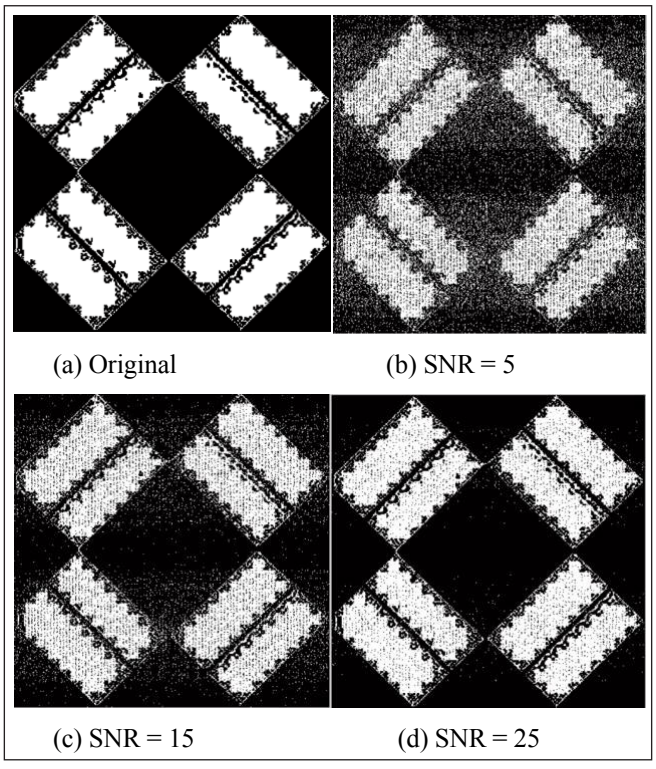


Fig. 19: Comparison at the Receiving Terminal under different SNR

From the above Fig. 19, we can find that as the SNR increases, the Gaussian noise contained in the picture of the receiving terminal decreases. For the test picture used in the current

experiment, when the value of SNR is between 0-5, the system simulation has the highest degree of excellence, which is closest to the theoretical curve.

In addition, by comparing the output of Fig. 17 and Fig. 18, we found that when the system input is random multi-frequency noise, the BER result is roughly the same as the test image. However, as the SNR increases, the goodness of fit of the BER of multi-frequency signals is lower than that of a single frequency.

VI. Conclusion

This article first briefly introduced the basic ideas of the OFDM system. Then the working principle of the OFDM system is described in detail according to the module sequence in Fig. 17 and Fig. 18. Then use experiments to analyze the information transmission performance of the OFDM system under the AWGN channel. The results obtained from the experiment also verify the advantages and disadvantages of OFDM mentioned above. In addition, the experiment can further explore the transmission performance of OFDM under different conditions by adding other channel tests.

The most valuable development direction of OFDM in the future is multi-antenna technology. Because multi-antenna technology can ideally increase system capacity and highlight system characteristics, and can significantly improve network stability and reliability, and greatly increase signal coverage, it is especially suitable for use in Internet and multimedia services. The MIMO-OFDM system combines MIMO technology and OFDM technology, which greatly improves the performance of the system. In the future, only if the shortcomings are overcome, OFDM can play a greater role in the post-5G era.

Acknowledgment

The authors would like to cordially thank all the teachers of the Department of Computer Science and Engineering, Islamic University, Bangladesh and most importantly our beloved parents for their continuous support.

References

[1] W. Stallings, *Wireless Communications & Networks,* 2nd ed. USA: Prentice-Hall, Inc., 2004.

[2] H. Viswanathan, and M. Weldon, "The past, present, and future of mobile communications," *Bell Labs Technical Journal*, vol. 19, pp. 8-21, 2014, doi: 10.15325/BLTJ.2014.2335491.

[3] Q. K. Ud Din Arshad, A. U. Kashif, and I. M. Quershi, "A review on the evolution of cellular technologies," *2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, 2019, pp. 989-993, doi: 10.1109/IBCAST.2019.8667173.

[4] Y. Li, and G. Cheng, "Fourth generation wireless communication network," *2013 3rd International Conference on Consumer Electronics, Communications and Networks*, 2013, pp. 312-315, doi: 10.1109/CECNet.2013.6703334.

[5] R. Prasad, and S. Hara, "An overview of multi-carrier CDMA," *Proceedings of ISSSTA'95 International Symposium on Spread Spectrum Techniques and Applications*, 1996, vol. 1, pp. 107-114, doi: 10.1109/ISSSTA.1996.563752.

[6] S. B. Weinstein, "The history of orthogonal frequency-division multiplexing [History of Communications]," *IEEE Communications Magazine*, vol. 47, no. 11, pp. 26-35, Nov. 2009, doi: 10.1109/MCOM.2009.5307460.

[7] H. Lian, R. Zhao, B. Hu, and H. Pang, "Simulation and analysis of OFDM communication system," *2010 The 2nd International Conference on Industrial Mechatronics and Automation*, 2010, pp. 677-680, doi: 10.1109/ICINDMA.2010.5538218.

[8] S. Coleri, M. Ergen, A. Puri, and A. Bahai, "Channel estimation techniques based on pilot arrangement in OFDM systems," in *IEEE Transactions on Broadcasting*, vol. 48, no. 3, pp. 223-229, Sept. 2002, doi: 10.1109/TBC.2002.804034.

[9] S. Weinstein, and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," *IEEE Transactions on Communication Technology*, vol. 19, no. 5, pp. 628-634, Oct. 1971, doi: 10.1109/TCOM.1971.1090705.

[10] T. Hwang, C. Yang, G. Wu, S. Li, and G. Ye Li, "OFDM and its wireless applications: A survey," *IEEE Transactions on Vehicular Technology*, vol. 58, no. 4, pp. 1673-1694, May 2009, doi: 10.1109/TVT.2008.2004555.

[11] Y. Wu, and W. Y. Zou, "Orthogonal frequency division multiplexing: A multi-carrier modulation scheme," *IEEE Transactions on Consumer Electronics*, vol. 41, no. 3, pp. 392-399, Aug. 1995, doi: 10.1109/30.468055.

[12] A. R. James, R. S. Benjamin, S. John, T. M. Joseph, V. Mathai, and S. S. Pillai, "Channel estimation for OFDM systems," *2011 International Conference on Signal Processing, Communication, Computing and Networking Technologies*, 2011, pp. 587-591, doi: 10.1109/ICSCCN.2011.6024619.

[13] L. Gopal, and M. L. Sim, "Performance analysis of signal-to-noise ratio estimators in AWGN and fading channels," *2008 6th National Conference on Telecommunication Technologies and 2008 2nd*

Malaysia Conference on Photonics, 2008, pp. 300-304, doi: 10.1109/NCTT.2008.4814291.

[14] A. Mráz, and L. Pap, "General performance analysis of M-PSK and M-QAM wireless communications applied to OFDMA interference," in *2010 Wireless Telecommunications Symposium (WTS)*, 2010, pp. 1-7, doi: 10.1109/WTS.2010.5479673.

[15] M. Batariere, K. Baum, and T. P. Krauss, "Cyclic prefix length analysis for 4G OFDM systems," *IEEE 60th Vehicular Technology Conference, 2004. VTC2004-Fall. 2004*, 2004, vol. 1, pp. 543-547, doi: 10.1109/VETECF.2004.1400066.

[16] R. Prasad, *OFDM for Wireless Communications Systems*, Artech, 2004.

[17] L. M. Rajeswari, and S. K. Manocha, "Design of data adaptive IFFT/FFT block for OFDM system," *2011 Annual IEEE India Conference*, 2011, pp. 1-5, doi: 10.1109/INDCON.2011.6139435.

[18] H. Al-Ghaib, and R. Adhami, "On the digital image additive white Gaussian noise estimation," *2014 International Conference on Industrial Automation, Information and Communications Technology*, 2014, pp. 90-96, doi: 10.1109/IAICT.2014.6922089.

[19] L. Galleani, "Time–frequency analysis of the impulse response," *IEEE Transactions on Signal Processing*, vol. 67, no. 5, pp. 1280-1295, Mar. 2019, doi: 10.1109/TSP.2018.2890365.

International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 10-16
http://www.publishingindia.com/ijkbcs/

Find_S Algorithm: To Detect Node Behaviour in Ad Hoc Network

Samara Mubeen
Assistant Professor, Department of IS&E, Jawaharlal Nehru National College of Engineering, Shivamogga, Karnataka, India. Email: samaramubeen@jnnce.ac.in

Abstract: As humans have different behaviour good and bad. Some of the behaviour the human has kind, selfish, cooperative, uncooperative, gentle, soft hearted, back biting, jealous etc. Similarly, nodes in the Ad Hoc network have different behaviour like selfish nodes, regular nodes, malicious nodes etc. Transferring the information from the source node to the destination node via intermediate nodes is done using different routing algorithms. Routing algorithms find out the optimal path to reach the destination. They fail to identify the behaviour of the nodes when transferring the information. Identification of the nodes whether they are good or bad is a difficult task. In this paper different dataset having different behaviour are collected. Find_S algorithm is used to analyze the dataset and give the correct hypothesis as output. The hypothesis consists of different attributes like nature of the agent, energy level, type of agent, and finally, is the decision of the node.

Keywords: Dataset, Find_S algorithm, Node behaviour, Regular node, Selfish node.

I. Introduction

Ad Hoc network is infrastructure less network in which the nodes communicate with one another directly. The device themselves act as the router for forwarding of the packet from one node to another node until it reach the destination. As human being has different behaviour at different time, even the nodes in the Ad Hoc network have different behaviour. Some of the behaviour of the nodes is regular node, selfish node and malicious node. The regular node is the one which forwards the packet to neighbouring node until all the energy within it is exhausted. The selfish node is the one which behaves as regular node but after sometime the node stop forwarding the packet even the energy is there for forwarding. The selfish nodes want to preserve their resource. The malicious node is the one which will cause many problems in the network. Some of the behaviour exhibited by the malicious node are dropping of the packet, by performing unnecessary operation draining of the battery, simply consumes or drops the packet and does not forward it. To analyze this behaviour, the different dataset are considered.

Ever since computers were invented, the question arises whether they might be made to learn. To understand how to program them to learn, to improve automatically with experience, imagine computers learning from medical records which treatments are most effective for new diseases, houses learning from experience to optimize energy costs based on the particular usage patterns of occupants, or personal assistants learning the evolving interests of their users in order to highlight especially relevant stories from the online morning newspaper. A successful understanding of how to make computers learn would open up many new uses of computers and new levels of competence and customization. As the understanding of computers continues to mature, it seems inevitable that machine learning will play an increasingly central role in computer science and computer technology.

Find_S algorithm is used for finding the best hypothesis. The Find_S algorithm illustrates a way in which the more_general_than partial ordering can be used to organize the search for an acceptable hypothesis. The search moves from hypothesis to hypothesis, searching from the most specific to progressively more general hypotheses along one chain of the partial ordering. The algorithm of Find_S is given below.

Algorithm 1: Find_S Algorithm

1. Initialize h to the most specific hypothesis in H.
2. For each positive training instance x
   For each attribute constraint a_i in h
   If the constraint ai is satisfied by x
   Then do nothing
   Else replace a_i in h by the next more general constraint that is satisfied by x
3. Output hypothesis h

Paper is organized as follows: Introduction in Section I, Literature Review in Section II, System Design of Regular and Selfish Nodes in Ad Hoc Network considered in Section III,

in Section IV Implementation and Result Analysis followed by Conclusion in Section V.

II. Literature Review

Mohammad Ahmed Ahmed Al-JAoufi *et al.*, "Study of selfish node incentive mechanism with a forward Game node in wireless Sensor Networks", International Journal of Antennas and Propagation. Selfish node do not cooperate in forwarding of the packets and preserve their own limited resource [1]. Evolutionary game theory concept implemented into the nodes of the wireless sensor network to improve the reliability and stability of the network. The results of the simulation show that they use incentive mechanism for forwarding packets. By this stability and reliability of wireless sensor networks are improved.

Enrique Hernandrez Oralla *et al.*, "Improving selfish node detection in MANETs using a collaborative watchdog", IEEE Communications Letters, vol. 16, No. 5 May 2012 [2]. MANET do not use pre existent infrastructure. To save its resource selfish nodes will not forward their packets. Watchdogs are used to detect selfish nodes to improve accuracy and reduction of detection time. Collaborative watchdog can reduce the overall detection time with a reduced overhead.

Sonja Buchegger and Jean-Yves Le Boudee, "Self policing Mobile Ad Hoc Networks by Reputation System", The performance of the Mobile Ad Hoc network decrease due to the presence of selfish or malicious nodes. Some problems are eliminated by using incentives or secure routing by crypto-graphy [3]. By using a reputation system in all nodes makes them detect misbehaviour and use of second hand information.

Neenavath Veeraiah and B. T. Kaishna, "Selfish node detection IDSM based approach using Individual master cluster node", Manet is a group of infrastructure less network. Every node in the network is responsible for forwarding packets to its neighbouring nodes. The selfish node behaviour presence leads to partition of the network and makes a negative impact in the operation of the network. There are lots of techniques to identify the selfish node where more computational resources and time consuming process to identify selfishness of the node [4]. The selfish node is detected for single node and the clustering is used to increase the efficiency and also reduces the network energy consumption which leads in the reliable quality services throughout the network.

Shailender Gupta, C. K. Nagpal and Charu Singla, in MANET node should perform the community service truthfully [5]. In community service each node relays data packets of other nodes by spending their resources. A selfish node present in the MANET uses the network resources for its own profit but avoid helping others for preserving its resources. If number of selfish nodes increases in MANET, will eventually disrupt the network. Here the impact of selfish nodes concentration on quality of service in MANETs.

Naveen Kumar Gupta, Ashish Kumar Sharma and Abhishek Gupta [6], Wireless interfaces will be common in Ad Hoc networks, here relaying of packets among nodes are done without base stations. Selfish nodes are present in the Ad Hoc network that will degrade the whole network. A comparison study of different methods which will detect the increase in selfish node and decrease in the false detection rate. To detect them a simulation model is used which will detect the increase of selfish node and decrease of false detection rate.

Ashraf Al Sharah, Mohammad Alhaj and Mohammad Hassan [7], In mobile Ad Hoc network nodes organizing themselves in a non-centralized form for forwarding of the information. Selfish node exists in the network which will utilize the resource from for their own benefits. In this paper a selfish dynamic punishment scheme is used for using the cooperative repeated game in Mobile Ad Hoc network, by using this approach the selfish nodes are motivated to cooperate. The cooperative punishment is pertain to all network nodes, so that punished node is simulated for cooperating with other nodes in the MANET.

MANET becomes an encouraging area for innovative work of wireless communication system [8]. The Intrusion detection system observes for doubtful actions inside a system and takes action against them. In their research work a specially designed MANET's known as Enhanced Adaptive Acknowledgement system is discussed. This scheme solves all the problems of existing intrusion detection systems and the presence of false misbehaviour report.

New routing techniques which incorporate security attributes as parameter into Ad Hoc route discovery called Security-Aware Ad Hoc routing [9]. SAR enables the use of security as a metric to improve the relevance of the routes discovered by Ad Hoc routing protocols. Framework enables applications to adapt their behaviour according to the level protection available on communicating nodes in an Ad Hoc network. A wireless Ad Hoc network is a temporary network in which nodes are moving arbitrary in the places that have no network infrastructure [10]. Here a new routing algorithm Ad Hoc On-Demand Distance Vector with Black-hole Avoidance used to avoid black hole attack. Non-malicious nodes gradually isolate the black-hole nodes based on the values collected in their legitimacy table and avoid them while making path between source and destination.

A security threats against Ad Hoc routing protocols, examining AODV and DSR. A solution is proposed where no network infrastructure is pre-deployed [11]; a small amount of security coordination is expected. The protocol ARAN is based on certificates and successfully defeats for all identified attacks. In MANET's any node under attack exhibits an anomalous behaviour called malicious behaviour [12]. Here the entire operation of a network gets disturbed. In this work malicious node behaviour is defend and security solution are presented which are used in furnishing a secure and reliable communication in Ad Hoc.

III. System Design of Regular and Selfish Nodes in Ad Hoc Network

The Fig. 1 shows the behaviour of the regular node in Ad Hoc network, when the nodes are regular. The source node broadcast the hello packet to know the nodes welling to forward the packets, in response the nodes will send the acknowledgement to the source node willing to forward the packet. The source node will send the packet to the neighbouring node. The neighbouring node will send the packet to its neighbour until the final destination node is reached.

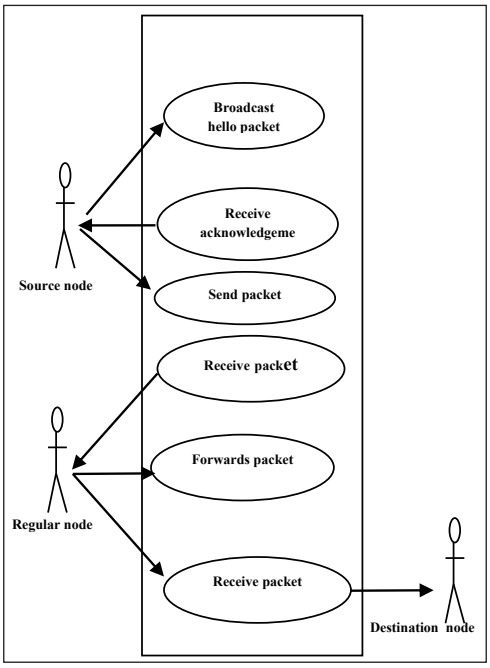


Fig. 1: Behaviour of Regular Node

Fig. 2 shows how the node will behave when selfish node is present in the network. The source node will broadcast the hello packet; the selfish node will send the acknowledgement will to forward the packet. Selfish node will receive the packet forwarded from the source node, due to its selfish behaviour it will not forward the packet to the neighbouring node as result the packet will not be received by the destination node. If the source node identifies the presence of selfish node the packet can be forwarded to the regular present in the network and the packet will reach the destination node.

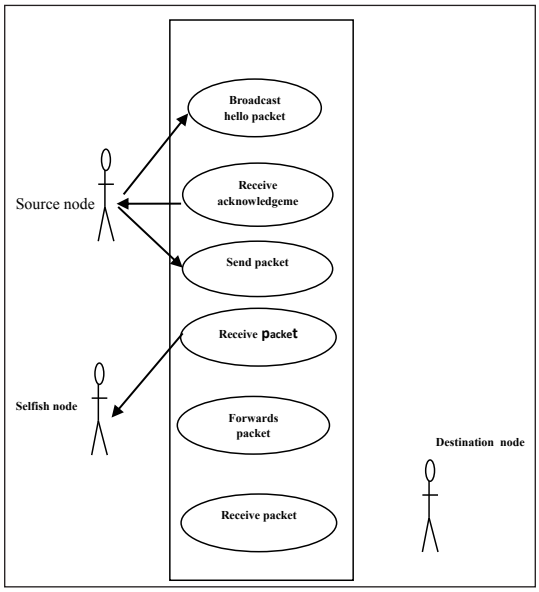


Fig. 2: Behaviour of Selfish Nodes

IV. Implementation and Result Analysis

The Fig. 3 shows the flowchart which represents the implementation done in order analysis the behaviour of the nodes in Ad Hoc.

The flowchart begins with dataset set given as input to the Find_S algorithm. A hypothesis is generated as output; this is checked with the regular behaviour of the node. If the node is regular then the packet is forwarded to the node and if not the packet is not forwarded to the node in the Ad Hoc network.

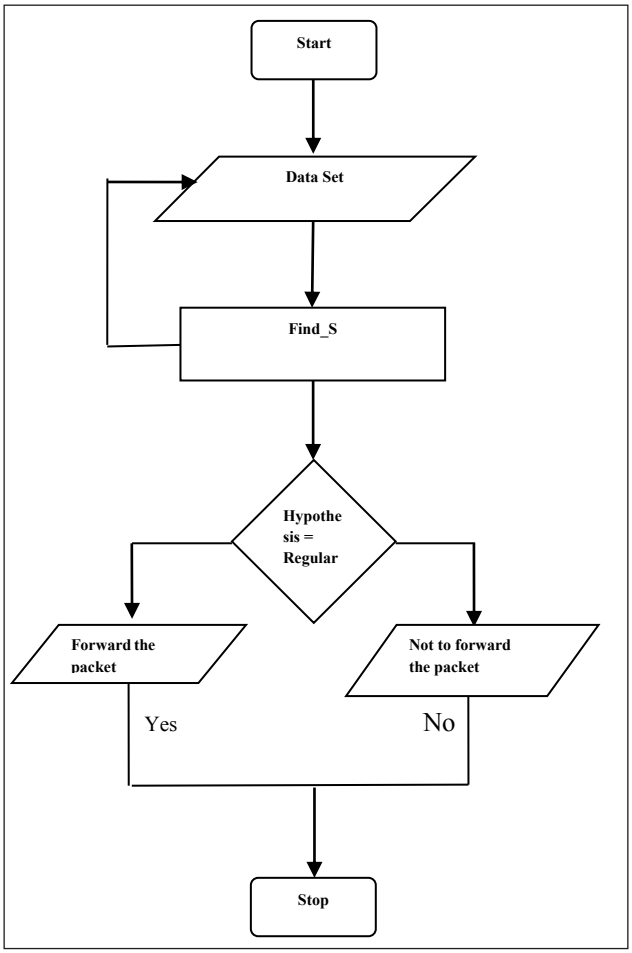


Fig. 3: Flowchart for Making Decision for Forwarding of the Packet

Case Studies on Different Dataset for Finding the Behaviour of the Node using Find_S Algorithm

In implementation four different cases are discussed having same attributes and different values. The attributes considered for dataset for finding the behaviour of the nodes are nature of the agent of the node whether they are good or bad, good means these nodes always helps in forwarding of the information to other nodes, bad means these nodes are not involved in forwarding of the information. The second attribute considered is the energy level in which the energy level can be high or medium represented as HM, energy level low is represented as L. The type of agent is classified as regular node represented as R or selfish node represented as S. Decision of node is last attributes in the table which justify the node behaviour whether to forward the information to next node is represented by F or not forwarding of the information is represented by NF which not forwarding.

Case 1: When Nodes are Regular and Packets can be Forwarded

The Table I shows dataset of the behaviour of regular and selfish nodes. The dataset consist of good and bad nodes by nature having energy level high or medium or low level, agent type is regular or selfish and the behaviour of the node is given by F or NF. The dataset consist of good natured nodes, bad natured nodes are countable. This dataset is given as input to the Find_S algorithm.

Table I: Behaviour of Regular Nodes Chosen to the Forward the Packet

Nature of the Agent	Energy Level	Type of Agent	Decision of Node
Good	HM	R	F
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Good	L	R	NF
Bad	HM	R	NF
Good	L	S	NF
Bad	HM	R	NF
Good	HM	R	F
Good	HM	R	F
Good	HM	R	F
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Good	L	S	NF
Good	HM	R	F
Good	L	S	NF
Bad	HM	R	NF
Bad	L	S	NF
Good	HM	R	F
Good	H	R	F
Good	L	S	NF
Bad	L	R	NF
Good	H	S	NF
Good	H	S	NF
Bad	H	S	NF
Good	H	R	F

The output of the Find_S algorithm is given below:

OUTPUT: [Good,?, R], Decision to forward = F

The hypothesis obtained as output shows that if the dataset has good nodes by nature along with energy level can be low, high or medium and the agent is regular. The final decision is to forward the information to next node in the network.

Case 2: When Nodes are not Regular and Forwarding of Packets can be Avoided

Table II is the dataset which consist of same attributes with different values. The dataset consists of bad and good natured nodes.

Table II: Behaviour of Regular Nodes Chosen to the Not to Forward the Packet

Nature of the Agent	Energy Level	Type of Agent	Decision of Node
Good	L	S	NF
Good	HM	R	F
Bad	HM	SR	NF
Good	HM	R	F
Good	L	R	F
Bad	HM	SR	NF
Bad	L	S	NF
Bad	HM	R	NF
Good	HM	R	F
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Good	L	S	NF
Good	HM	S	NF
Good	L	S	NF
Bad	L	R	NF
Good	H	S	NF
Good	H	S	NF
Bad	H	R	NF

This is fed to the Find_S algorithm the output obtained from the decision is given below:

OUTPUT: [Bad, ? ,?] Decision to forward = NF

The nature of the agent is bad, having energy level low, medium or high and the type of agent can either be selfish or regular, the decision is not to forward the packet to this node.

Case 3: When Nature of Nodes is Good and Energy Level is Low Forwarding of Packets can be Avoided

The Table III shows the dataset in which there are both good and bad natured nodes by nature, their energy level is low and type of node is selfish.

Table III: Behaviour of Regular Nodes Not Regular Chosen to the Not to Forward the Packet

Nature of the Agent	Energy Level	Type of Agent	Decision of Node
Good	HM	R	F
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Good	L	R	NF
Bad	HM	R	NF
Good	L	S	NF
Bad	HM	R	NF
Good	HM	R	F
Good	HM	R	F
Good	HM	R	F
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Good	L	S	NF
Good	HM	R	F
Good	L	S	NF
Bad	HM	R	NF
Bad	L	S	NF
Good	HM	R	F
Good	H	R	F
Good	L	S	NF
Bad	L	R	NF
Good	H	S	NF
Good	H	S	NF
Bad	H	S	NF
Good	H	R	F

When this dataset is given to the Find_S algorithm the hypothesis obtained is shown below:

Output: [?, ?, ?] Decision to forward = NF

The nature of agent is good, energy level is low, high or medium

and the type of agent is selfish or regular. These types of nodes having the behaviour pattern are avoided from forwarding of the information.

Case 4: When Nodes by Nature are Good, Energy Level is Low, Medium or High

Table IV shows the good or bad nature of nodes having energy levels low and high or medium represented by HM. This dataset is fed as input to the Find_S algorithm.

Table IV: Behaviour of Regular Nodes Not Regular Chosen to the Not to Forward the Packet

Nature of the Agent	Energy Level	Type of Agent	Decision of Node
Good	L	S	NF
Good	HM	S	NF
Bad	HM	SR	NF
Good	HM	S	NF
Good	L	S	NF
Bad	HM	SR	NF
Bad	L	S	NF
Bad	HM	R	NF
Good	HM	S	NF
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Bad	HM	R	NF
Good	HM	R	F
Good	L	S	NF
Good	HM	S	NF
Good	H	S	NF
Good	H	S	NF
Bad	H	S	NF
Good	H	R	F

The hypothesis obtained is given below:

Output hypothesis [?, ?, ?] Decision to Not forward = NF

The node having these values is not allowed to forward the information.

Discussion on the Result Obtained

The above cases show the different behaviour of the nodes. By knowing the behaviour of the nodes in a particular environment, the presence of the selfish nodes can be completely eliminated. By taking precaution earlier from the test cases available. This will help to overcome problem of identification of the selfish nodes in the network of Ad Hoc network provided that the dataset of the about the behaviour of the nodes are available earlier.

V. Conclusion

As human have different behaviour similarly nodes in Ad Hoc network are having different behaviour, in this paper different dataset is used, having attributes like Nature of the Agent, Energy Level, Type of Agent and Decision of node are considered. Using Find_S algorithm the behaviour of the node is recorded for each run. By this we can classify the node as selfish or regular node in Ad Hoc network. The elimination of the regular node. If the nodes which are selfish can be completely eliminated from the network before forwarding of the packets or information is done, by doing in this way speed of the packet reaching the destination is speeded and along with that misbehaving nodes are completely eliminated from the Ad Hoc network.

References

[1] Mohd. A. A. Al-Jaoufi, Y. Liu, Z.-J. Zhang, and L. Uden, "Study on selfish node incentive mechanism with a forward game node in wireless sensor networks," *International Journal of Antennas and Propagation*, vol. 2017, 2017.

[2] E. Hern´andez-Orallo, M. D. Serrat, J.-C. Cano, C. T. Calafate, and P. Manzoni, "Improving selfish node detection in MANETs using a collaborative watchdog," *IEEE Communications Letters*, vol. 16, no. 5, pp. 642-645, May 2012, doi: 10.1109/LCOMM.2012.030912.112482.

[3] S. Buchegger, and J.-Y. Le Boudec, "Self-policing mobile ad-hoc networks by reputation systems," in National Competence Center in Research on Mobile Information and Communication Competence Center in Research on Mobile Information and Communication Systems (NCCR-MICS).

[4] Neenavath, and B. T. Krishna, "Selfish node detection IDSM based approach using individual master cluster node," in *2nd International Conference on Inventive Systems and Control*, 2018.

[5] S. Gupta, C. K. Nagpal, and C. Singla, "Impact of selfish node concentration," *International Journal of Wireless & Mobile Networks*, vol. 3, no. 2, Apr. 2011.

[6] N. K. Gupta, A. K. Sharma, and A. Gupta, "Selfish behaviour prevention and detection in mobile ad-hoc network using intrusion prevention system (IPS)," *International Journal of Research Review in Engineering Science and Technology*, vol. 1, no. 2, pp. 31-34, Sep. 2012.

[7] A. Al Sharah, Mohd. Alhaj, and Mohd. Hassan, "Selfish dynamic punishment scheme: Misbehavior detection

in MANETs using cooperative repeated game," *International Journal of Computer Science and Network Security*, vol. 20, no. 3, pp. 158-173, Mar. 2020.

[8] V. U. Raut, and M. S. Mahindrakar, "A comprehensive survey on intrusion detection in MANET," *International Journal of Information Technology and Knowledge Management*, vol. 2, no. 2, pp. 305-310, Jul.-Dec. 2010.

[9] S. Yi, P. Naldurg, and R. Kravets, "Security-aware ad hoc routing for wireless networks," in *Proceedings of ACM MOBIHOC*, Oct. 2001, pp. 299-302.

[10] Y.-C. Hu, D. B. Johnson, and A. Perrig, "SEAD: Secure efficient distance vector routing for mobile wireless ad hoc networks," in *Proceedings of the 4th IEEE Workshop on Mobile Computing Systems and Applications*, Callicoon, NY, Jun. 2002, pp. 3-13.

[11] K. Sanzgiti, B. Dahill, B. N. Levine, L. Y. Clay, S. Elizabeth, and M. Belding-Royer, "A secure routing protocol for ad hoc networks," in *Proceedings of the 10th IEEE International Conference on Network Protocols (ICNP'02),* Nov. 2002, pp. 78-87.

[12] Z. Alfawaer, and S. Al Zoubi, "A proposed security subsystem for ad hoc networks," *International Forum on Computer Science Technology and Applications*, IEEE, 2009.

International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 17-21
http://www.publishingindia.com/ijkbcs/

To Study the Hypervisor Scanner Model with ANN for Cloud Systems

Anshu Mali Bhushan[1*], Harsh Kumar[2] and Devendra Singh Chauhan[3]
[1]Research Scholar, Computer Application Department, Himalayan Garhwal University, Pauri Garhwal, Uttarakhand, India. Email: hianshumali@gmail.com
[2]Faculty of Computer Application, Himalayan Garhwal University, Pauri Garhwal, Uttarakhand, India. Email: research@hgu.ac.in
[3]Assistant Professor, Department of Physics, Dr. PDBH Government PG College, Kotdwar, Uttarakhand, India. Email: dr.devchauhan78@gmail.com
*Corresponding Author

Abstract: The detection part is called the Hypervisor Scanner, which is programmed to detect malicious insiders. Using the feed forward neural network, the hypervisor scanner is generated and trained with a supervised learning algorithm referred to as the Levenberg-Marquardt algorithm. The three criteria are considered for service level agreement, such as bandwidth requirement, memory consumption and storage space. The Hypervisor Scanner can detect malicious insiders in cloud systems that breach SLA and suffer from an insecure cloud administrative domain that lacks control over the cloud service provider (CSP). Here, the Hypervisor Scanner is constructed using the biologically inspired classification approach referred to as artificial neural network modelling. ANN teaching uses the Levenberg-Marquardt learning algorithm. The LM algorithm works by minimising the average square error to boost the detection system. For detecting malicious insiders, the Hypervisor Scanner uses threshold values for SLA parameters. It is known from performance review and comparison that the Hypervisor Scanner is the appropriate one with high detection accuracy and low false alarm rate for the detection of malicious insiders. This can therefore be effective, robust and realistic in the detection of malicious insiders in and around the cloud world.

Keywords: Architecture, Cloud computing, Commercial and technological, Hypervisor Scanner.

I. Introduction

Due to the evolutionary growth of numerous existing methods and computing facilities, such as distributed networks, services, applications and infrastructures, including network collection, computing technologies and storage resources, cloud computing allows the development of a commercial and technical service model. While cloud computing is growing in academia, enterprise, and industry, it is still an evolving paradigm. Via optimization, cloud computing has the ability to minimise costs and increase the benefits of operations and economic features (Catteddu & Hogben, 2009) [1]. In addition, cloud computing greatly strengthens and allows Internet infrastructures to create a universal computing model. This potentially evolved computing model will turn out to be a great failure model without adequate protection and privacy solutions built for cloud systems. It is imperative to ensure information security controls when using the internet, cloud computing and wireless applications. Therefore, the protection and privacy issues of this unique model are major obstacles to its adoption (Bruening & Treacy, 2009) [2]. The architecture of the cloud is shown in Fig. 1. The figure illustrates the characteristics of cloud computing systems and how the functionality of the cloud computing model can be easily accessed by users.

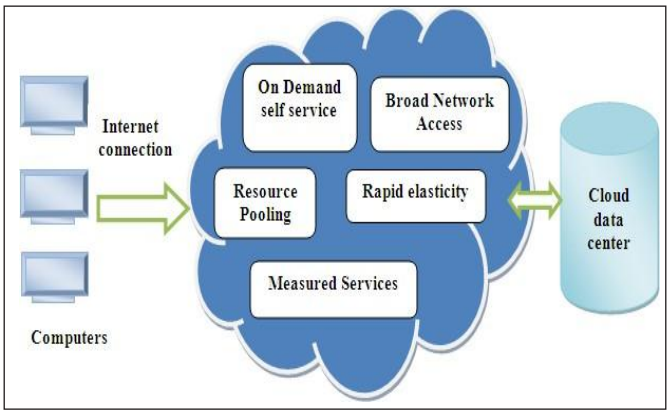


Fig. 1: Cloud Computing Architecture

The main objective of this research is to recognise and study the security problems that can impact the performance of cloud systems. The aim of the research work is to define a security mechanism that can be used to obtain prominent security solutions. As a result, this dissertation also aims to provide a malware detection component based on Hypervisor that is

appropriate for the detection of malicious activities against cloud technology in a dynamic cloud environment.

A. Hypervisors and Virtualization

Virtualization is an important feature that makes it easier for customers to access conceptual infrastructure and services as isolated VMs (Takabi *et al.,* 2010) [3]. A virtual machine monitor or hypervisor is a component of software for platform virtualization that enables multiple operating systems to run simultaneously on a host device. As this facilitates the generation of virtualized sharing services, the attack surface is thus strengthened. Therefore, a mechanism is required to guarantee strong VM isolation, arbitrated sharing, and stable VM communications. This can be done by using a versatile framework for access control and monitoring that handles VM control and sharing capabilities within a cloud host (Chen *et al.,* 2010) [4].

B. Layered Architecture of Cloud System in Virtual Environment

Cloud draws more individuals to use web services as well as misuse cloud services and its tools with faster growth of web-based services. A survey was conducted on business and technological problems that directly impact cloud computing efficiency at design and implementation levels (Nirmala & Sridaran, 2012) [5]. Cloud apps run outside the firewall and transfer to a public domain that has a good security awareness. Furthermore, the dynamic existence of virtual machines makes the process of monitoring more complex. In order to use the detection system, the hypervisor-based technology that encourages and regulates the execution of VMs and middleware can also be used. Hence, in this dissertation, the Hypervisor-based malware detection component is proposed to detect unauthorised access to malware in the cloud environment. In Fig. 2, the Hypervisor virtual cloud architecture is shown.

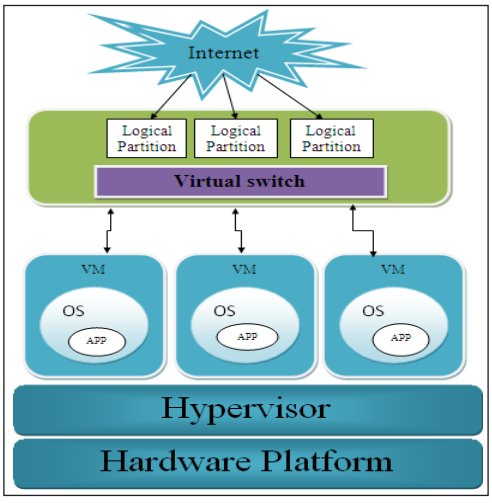


Fig. 2: Layered Cloud Architecture with Hypervisor

Virtualization offers web-based software and IT companies with improved and streamlined performance for applications in a cost-effective manner. This makes distribution of cloud service much simpler and scalable. Virtualization can be applied to anything in a cloud environment that relies on storage, operating system, network, device, and software/hardware virtualization forms of virtualization. One of the cloud providers, Infrastructure as a Service, provides data virtualization (IaaS). This allows many, on a pay-use basis, to have access to vast quantities of cloud resources. Virtualization suits are prepared for cloud computing with the three most critical virtualization properties: Partitioning, Isolation and Encapsulation. Virtualization (Bento & Aggarwal, 2012) [6] is described as the virtual description of something such as hardware, operating system, software, etc. Virtualization (Jin *et al.,* 2013) [7] is a method for providing parallel and interactive access to a large pool of information centre that supports many instances of control over multiple OS that successively generates hardware virtualization Hypervisor requires many OS instances to share the hardware facilities it is hosted on.

It addresses the similarities and differences between cloud computing and multiprogramming. Virtualization allows multiple operating systems to share CPU resources where they are handled by multiple OSes who have no knowledge of them. Normally the OSes are supposed to be managed by the underlying hardware. In cloud systems, different operating systems are managed and executed by the hypervisor layer.

The virtual machine runs a mechanism for the user on a single physical device. Virtual machines can transfer their states dynamically. The VM will move unconditionally from one platform to another. It is possible to clog, poise or infertile VM. Virtual machines run multiple programmes on a single hardware platform that are stored at different host device locations. Because of the complex existence of cloud technology (Vinothina *et al.,* 2012) [8], it is possible to migrate virtual computing environments and scale their properties over a multi-domain infrastructure. The operating system that is installed and executed on a virtual machine is the Guest OS. The Hypervisor controls the guest operating system that operates on virtual machines (Sabahi, 2012) [9]. Hypervisor enables numerous OS instances to share the hardware facilities on which it is hosted.

C. Hypervisor Layer is Used to Deploy Malware Detection System for Cloud

Web-based services and cloud services are being built more rapidly, attracting many users to use web services and hack cloud services and resources. They may have a serious awareness of security as cloud apps run outside the firewall and move on to the public domain. In addition, the complex design of virtual machines makes the control process more difficult. Therefore, it is possible to use a virtual machine monitor or hypervisor-based technology. It can be used to use the detection

system as it progresses and monitors the execution of VMs and middleware. The Hypervisor-based malware detection system is therefore proposed in this research in order to recognize unauthorised access to cloud systems. The malware detection system based on Hypervisor is an intrusion detection system provided to the hypervisor layer. The Hypervisor-based malware detection component is designed to observe both host and network-based activities in the cloud infrastructure, based on the rapid flow of huge amounts of data. This conducts several tasks in the virtualization layer, such as tracking, managing and evaluating, instead of being directly deployed in virtual machine monitoring. In order to take appropriate action against malicious activities, the analysis and control of the Hypervisor-based malware detection component is provided to the cloud administrative module in the cloud environment. The malware detection component based on Hypervisor consists of two features that can distinguish it from conventional intrusion detection systems, including

- Isolation from attackers and
- Transparency to attackers.

The isolation of the malware detection component based on Hypervisor from attackers secures it from direct attacks. The portion of hypervisor-based malware detection is clear to attackers, as it is both host and network-based IDS without virtual machines being directly accessible. The malware detection component based on the Hypervisor should guarantee virtual machine security and ensure virtualization layer safety. In the detection method, most of the attacks in the cloud environment often seek to compromise. But this malware detection component based on the hypervisor is built in such a way that it is not neutralised or compromised by any action. In virtual networks, Hypervisor monitors VMs and middleware by allowing VMs to be inspected by the host system. The Hypervisor is positioned between the virtualization layer and the kernel. Hypervisor Detector can discover any form of attack by actively and passively monitoring cloud environments, thereby ensuring that cloud components have not been compromised. The Detector Hypervisor minimises VM transparency.

II. Research Methodology

The Proposed Hypervisor Scanner Developed with ANN

Even if the existence of the malicious insider is not open, this chapter defines the malicious insider as an authorised user who breaches the Service Level Agreements (SLA) and infuriates the hacking of cloud infrastructure by affecting cloud services and resources. In order to use cloud services, consumers must acquire a legal conformity with CSP called Service Level Agreements. According to SLA, the cloud service provider (CSP) (Maurer *et al.,* 2012) [10] is responsible for the efficient provision of services and resources. CSP must also be guaranteed to mitigate breaches of SLA and optimise consumption of energy. Huge numbers of users have to share cloud services. This limitation demands that the computing activity of a single user does not accumulate cloud resources. In certain cases, by sending a broad resource request to access the cloud resources, malicious insiders attempt to break SLAs. The three attributes of SLA (Maurer *et al.,* 2012) are addressed in this section, such as 1) Bandwidth requirement, 2) Utilisation of memory and 3) Space space. Threshold values are allocated to each factor, such as memory consumption and storage space, as shown in Table I, and for bandwidth requirements. The dishonest cloud administrator, who is deficient in CSP management, is unable to analyse user behaviours. Often called the hypervisor layer, the virtual machine monitor (VMM) provides abstraction between real and virtual machines that can monitor and supervise virtual machine activities. By taking this into account, at the Hypervisor layer, the detection component called Hypervisor Scanner is built to track the virtual clients requested for services and resources.

Table I: SLA Attributes and Their Threshold Values

Sr. No.	Attribute for SLA	Threshold Limit
1	Bandwidth	8Mbits/s
2	Memory utilization	500MB
3	Storage capacity	1200GB

III. Results and Discussion

In order to have a better classification (Abdel-Azim *et al.,* 2009) [11], the Hypervisor Scanner is modelled using the Artificial Neural Network modelling technique on whether the user is malicious or normal with greater accuracy. In this model, for training ANN, the Levenberg-Marquardt (LM) back propagation training algorithm (Linda *et al.,* 2009) [12] is used. Kenneth Levenberg and Donald Marquardt created the Levenberg-Marquradt algorithm (Yu & Wilamowski, 2012) [13] to provide a better solution to the issue of minimising a nonlinear algorithm. The writers (Pradeep *et al.,* 2011) [14] addressed the relevance of ANN's LM training algorithm. Based on the above-mentioned SLA features, the Hypervisor Scanner part is skilled and tested to determine the malicious insider that violates the SLA. How this method works is explained in Algorithm.

IV. Hypervisor Scanner Modelled with ANN

The hypervisor scanner is constructed using the Levenberg-Marquardt back propagation algorithm for ANN modelling. For larger datasets, the conventional classification techniques involve significant computing power, memory and CPU resources. The common learning method (Yang & Fung, 2008) [15] for machine learning to improve the efficiency of the intelligent system is ANN. The artificial neural network (Abdel-Azim *et al.,* 2009) [11] is a very important modelling

tool for accurately classifying (Donghai & Weiwei, 2013) [16], especially where conventional analysis is difficult. ANN is an efficient method for mathematically simulating the capacities of the human brain. The neural network consists of the essential processing unit, a group of neurons. Since multiple neurons function at the same time, the brain can stimulate and generate better results quickly. The Artificial Neural Network works with multiple neurons simultaneously to boost their efficiency, according to the same methodology. Either through supervised or unsupervised learning methods, the ANN must obtain and foresee. The related patterns between input data and corresponding target values can easily be learned by ANN. Input and output specifications are necessary for the supervised back propagation learning technique, where the training takes place iteratively by using a collection of training samples.

The neuron (Pradeep *et al.,* 2011) [14] is the central operating entity with an input layer, a collection of hidden layers and an output layer for the neural network. The weight is the number for an artificial neuron and it represents the neuron's weight value. The input values are totally applied and altered by weights. The activation function finally regulates the output amplitude. The neuron model for neural networks is shown in Fig. 3.

This model of ANN has a structure of three layers of network 1. The input layer is equivalent to the number of parameters 2 with three input neurons. Hidden layer with 2 computational neurons that are hidden in order to determine the number of neurons in the secret layer, there is no clear formula. There are certain laws of thumb, anyway, for the neurons to determine. The formula used in this work is

$$N = \sqrt{(m+2)n} + 2\sqrt{n/(m+2)}$$

In which the number of input neurons is n and the number of output neurons is m. 3 Layer of output of one output neuron. Artificial neural network simulation for the Hypervisor Scanner as shown in Fig. 4.

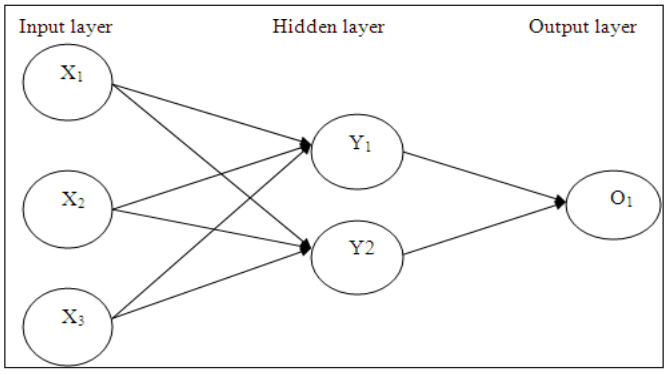


Fig. 3: Model of Neuron in ANN

The input signals flowing from the input layer in the feed forward network take the information from original data that advances through any hidden neurons in hidden layers, eventually reaching the output units. The input variable values are placed in the input units during the network execution and the secret and output layer nodes are gradually executed. To generate the output of the node, the activation value is passed through the activation function. By summing the weighted output values of the nodes in the preceding layer, each node in the layer calculates their activation value. The outputs of the output layer serve as the output of that network after execution of the entire network.

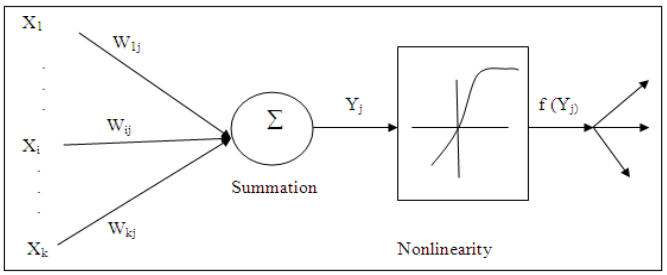


Fig. 4: ANN Modelling for Hypervisor Scanner

By using a simulation, the Hypervisor Scanner architecture is assessed. MatLab R2010 with Intel core2 processor operating at 2GHz, 2GB of RAM is used for Hypervisor Scanner simulation. Levenberg Marquardt's trainlm function is used to train the network, and is also the quickest algorithm for back propagation. The synthetic dataset is generated with characteristics such as bandwidth requirement, memory consumption and storage capacity with 500 data for experiment and study. The machine is trained to perform 100 iteration numbers. The dataset is split into three separate sets, called datasets for preparation, testing and validation. For networks, the training dataset is used to learn the patterns that occur in the dataset. To assess the skill of the qualified network, the testing dataset is used. By using the validation dataset, the performance of the network is assessed. Training uses 70 percent of the data in this model, testing uses 15 percent of the data, and validation uses 15 percent of the details.

V. Summary

This proposed method uses a better governing system called Hypervisor Scanner to detect malicious insiders in a virtual cloud setting where the cloud is controlled by a single cloud management domain coupled with a lack of cloud service provider (CSP) power. Here, the Hypervisor Scanner is designed using the best classification technique called Artificial Neural Network Modelling. To train ANN, the Levenberg-Marquardt (LM) back propagation algorithm is used. For experiments, the three attributes of SLA are considered, such as bandwidth requirement, memory consumption and storage space. In terms of bandwidth requirements, memory usage and storage space, users request more resources and are branded as malicious insiders. The experimental results show that the proposed Hypervisor Scanner generates mean values of squared error and regression that are approximately equal to zero and one

respectively. It can be found from a performance comparison that the Hypervisor Scanner detects malware behaviour with a minimum error. The ANN-designed Hypervisor Scanner has high detection precision and a low false negative rate. This can therefore be reliable, robust and realistic in order to detect malicious insiders in and around the cloud world.

References

[1] D. Catteddu, and G. Hogben, "Cloud computing: Benefits, risks and recommendations for information security," European Network and Information Security Agency, Technical Report, 2009.

[2] Bruening and Treacy, "Cloud computing: Privacy, security challenges," Bureau of National Affairs, 2009.

[3] H. Takabi, J. B. D. Joshi, and G.-J. Ahn, "Security and privacy challenges in cloud computing environments," *IEEE Security & Privacy*, vol. 8, no. 6, pp. 24-31, Nov.-Dec. 2010, doi: 10.1109/MSP.2010.186.

[4] Y. Chen, V. Paxson, and R. H. Katz, "What's new about cloud computing security," University of California, Berkeley, Report No. UCB/EECS-2010-5, Jan. 20, 2010.

[5] A. P. Nirmala, and R. Sridaran, "Cloud computing issues at design and implementation levels - A survey," *International Journal of Advanced Networking and Applications*, vol. 3, no. 6, p. 1444, 2012.

[6] A. M. Bento, and A. K. Aggarwal, *Cloud Computing Service and Deployment Models: Layers and Management*. USA: IGI Global, 2012, pp. 1-365.

[7] H. Jin, F. Xu, F. Liu, L. Liu, B. Li, and B. Li, "iAware: Making live migration of virtual machines interference - Aware in the cloud," *IEEE Transactions on Computers*, vol. 63, no. 12, pp. 3012-3025, 2014.

[8] V. Vinothina, R. Sridaran, and P. Ganapathi, "A survey on resource allocation strategies in cloud computing," *International Journal of Advanced Computer Science and Applications*, vol. 3, no. 6, pp. 97-104, 2012.

[9] F. Sabahi, "Secure virtualization for cloud environment using hypervisor-based technology," *International Journal of Machine Learning and Computing*, vol. 2, no. 1, pp. 39-45, 2012.

[10] M. Maurer, D. Borgetto, G. Da-Costa, J.-M. Pierson, and I. Brandic, "Energy-efficient and SLA-aware management of IaaS clouds," in *2012 Third International Conference on Future Systems: Where Energy, Computing and Communication Meet (E-Energy),* IEEE, May 2012, pp. 1-10.

[11] M. Abdel-Azim, A. I. Abdel-Fatah, and M. Awad, "Performance analysis of artificial neural network intrusion detection systems," in 2009 *International Conference on Electrical and Electronics Engineering, ELECO 2009*, IEEE, 2009, pp. II-385.

[12] O. Linda, T. Vollmer, and M. Manic, "Neural network based intrusion detection system for critical infrastructures," in *2009 International Joint Conference on Neural Networks*, IEEE, 2009, pp. 1827-1834.

[13] H. Yu, and B. M. Wilamowski, "Efficient and reliable training of neural networks," in *2009 2nd Conference on Human System Interactions*, IEEE, May 2009, pp. 109-115.

[14] J. Pradeep, E. Srinivasan, and S. Himavathi, "Diagonal based feature extraction for handwritten character recognition system using neural network," in *2011 3rd International Conference on Electronics Computer Technology,* IEEE, 2011, vol. 4, pp. 364-368.

[15] P. J. Yang, and C. C. Fung, "Artificial intelligence in malware - Cop or culprit?," 2008.

[16] Weiwei, and Donghai, "Cascade disturbance rejection control of the uncertain nonlinear systems with nonlinear parameterization," in *2013 American Control Conference,* IEEE, 2013, pp. 265-271.

International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 22-26
http://www.publishingindia.com/ijkbcs/

Various Credit Card Fraud Detection Techniques based on Machine Learning

Chinnu Maria Varghese[1*], Deepika M. P.[2] and Abraham Varghese[3]
[1]PG Student, Computer Science and Engineering, Adi Shankara Institute of Engineering & Technology, Kerala, India. Email: chinnumariav@gmail.com
[2]Associate Professor, Computer Science and Engineering, Adi Shankara Institute of Engineering & Technology, Kerala, India. Email: deepika.it@adishankara.ac.in
[3]Faculty, University of Technology & Applied Science, Muscut, Oman. Email: abraham.v@hct.edu.com
*Corresponding Author

Abstract: Use of mobile devices facilitates more people to do online shopping using credit cards. As a result, Internet shopping has become a popular method of making daily purchases. Online shopping offers benefits like efficiency, convenience and greater selection as well as better pricing. With the number of transactions by credit cards are increasing rapidly transaction fraud are also increasing. A fraud activity results in financial loss to the individuals. Therefore financial institutions provide more value and demand for fraud detection applications. We need to also make our systems learn from the past submitted frauds and make them fit for adapting to future new methods of frauds. This paper shares the concept of frauds related to credit cards and their various types. We have considered various techniques available for a fraud detection system such as, Hidden Markov Model (HMM), Bayesian Network, Hybrid Support Vector Machine (HSVM), K-Nearest Neighbor (KNN), Naïve Bayes, Logic regression, Decision Tree and a feedback mechanism. This paper covers existing and proposed models for credit card fraud detection and focus to find the best method by comparing different techniques on the basis of quantitative estimations like accuracy, detection rate and false alarm rate.

Keywords: Bayesian network, Hybrid support vector machine, K-Nearest neighbor, Logic regression, Naïve Bayes.

I. Introduction

In the course of the most recent couple of years, e-commerce has become an essential part of the global retail framework. Like many other industries, the retail landscape has gone through a considerable change following the advent of the internet, and thanks to the on-going digitalization of modern life, consumers from virtually every country now profit from the perks of online transactions. As internet access and adoption are quickly expanding around the world, the number of online buyers keeps climbing every year. According to Statista study, around two billion people purchased goods through online, and e-retail sales surpassed 4.2 trillion U.S dollars worldwide during the year 2020 [1].

Today use of credit card in developing countries has become a common scenario. Credit cards can be used for shopping, paying bills, loan payments and for online transactions. However, as the number of credit card users grows, so does the number of credit card fraud instances. Credit card related cheats cause worldwide a loss of billions of dollars. Fraud can be defined as any activity with the intent of deception to obtain financial gain by any manner without the knowledge of the cardholder and the issuer bank. With the advancement of mobile phones, online shopping becomes a popular mode of everyday buys. However, the Internet environment is open, online shopping systems have problems, and criminals can use some bad techniques suchas Trojan and pseudo base-station. All these ends up in a serious increase of credit card fraud events. When a criminal steals or cheats the information of the cardholder, the criminal can utilize the credit card to consume. Credit Card fraud can be done in numerous ways such as by phishing, by producing fake cards, by stolen cards, by cloning the original site, by changing the magnetic strip present at the card which contains the user's information, by skimming or by stealing data from a merchant's side. With continued advancement in fraudulent strategies it is important to develop effective models to combat these frauds in their initial stage only, before they can take to completion. However, the main issue in constructing such a model is that the number of fraudulent transactions is quite tiny in comparison to the total number of transactions, making the task of discovering a fraudulent transaction in aneffective and efficient manner rather difficult.

Detecting a fraud is a complicated computational task. The number of parameters to select, cluster and classify are

tremendous and classification of parameters will determine the success of any fraud detection technique. Moreover a transaction can't be exactly classified as a fraud or a genuine one by the existing systems; they just find the chances of a transaction being fraud based upon the elaborated study of customer's behavior, their spending habits and also analyzing the previously committed frauds and identifying their patterns. Limited time period to determine the transaction is genuine or fraud and processing of large number of parameters during training are the two main challenges in making a decision.

A good fraud detection system should have the following properties:

- The number of wrong classifications should be minimum that is identifying the frauds accurately.
- It should be able to detect the fraud while it is in transit.
- It ought not term any genuine transaction as fake.

In our paper we have tried to study some techniques that can be used in a fraud detection system and have tried to do a comparative study to show which technique performs better under what scenarios.

II. Related Work

Credit-card-based purchases can be categorized into physical card and virtual card. In a physical-card based purchase, the cardholder submits his card physically to a dealer for making a payment. In this kind of purchase, an attacker has to steal the credit card for fraudulent transactions. If the cardholder doesn't understand the loss of card, it can lead to a considerable financial loss to the credit card organization. In the virtual card based purchase, only some important information about a card is required to make the payment. Such buys are normally done on the internet or via telephone. A scammer only has to know the card details to conduct fraud in these types of transactions. Most of the time, the certified cardholder is not aware that another person has seen or stolen his card information. The best way to find this kind of fraud is to analyze the spending patterns on each card and to sort out any irregularity with respect to the "usual" spending patterns. Fraud detection based on the analysis of existing purchase data of cardholder is a promising method to lessen the rate of successful credit card frauds. Since humans tend to exhibit specific behavioristic profiles, each cardholder can be addressed by a set of patterns containing information about the typical purchase category, the time since the last purchase, the amount of money spent, etc. Deviation from these patterns poses a risk to the system.

In 2008, Abhinav Srivastava and Amlan Kundu [2] in their paper proposed a Hidden Markov Model (HMM). Intheir model they model sequence of operations in credit card transaction processing using a Hidden Markov Model (HMM) and show how it can be used for the detection of frauds. An HMM is first trained with the normal pattern of acardholder. If an incoming credit card transaction is notaccepted by the trained HMM with sufficiently high probability, it is viewed as fake. At the same time, try to ensure that genuine transactions are not rejected. The different steps in credit card transaction processing are addressed as the underlying stochastic process of an HMM. For the observation symbols ranges of transaction amount is used and types of item is considered as HMM states .The paper proposed a method for discovering the spending profile of cardholders, as well as utilization of this knowledge in deciding the value of observation symbols and initial estimate of the model parameters. The paper also explain how the HMM can recognize whether an incoming transaction is fake or not. Test results show the performance and effectiveness of system and demonstrate the usefulness of learning the spending profile of the cardholders. Accuracy of the system is nearly 80 percentages over a wide variation in the input data. The system is also applicable for large volumes of transaction handlings.

V. Filippov and L. Mukhanov [3] have proposed a Fraud Detection System in their paper. The use of credit cards is very common in modern day society. However, the number of credit card fraud cases is continually increasing despite the chip cards worldwide integration and existing protection systems. This is why the issue of fraud detection is vital at this point. A general description of the developed fraud detection system and examinations between models based on using of artificial intelligence are given. In the fraud detection system two modules FDS ONLINEP and FDS OFFLINEP are used for fraud detection. The FDS ONLINEP module is used for online fraud detection. FDS OFFLINE module permits the system to detect fraud among transactions that have been already authorized and grouped in the FDS ONLINEP module. For the storage of incoming transactions, statistical data for corresponding models, results of classification and generic parameters a FDS Data Warehouse is used. Module FDS ALERT is used for alerting credit card holders in case of fraud detection by the FDS ONLINEP module using SMS or email messages. For building statistical data the FDS BUILDSTATP module is used. A clustering model is developed based on the use of the parameters' data clusterization regions. In this system 24 real parameters of transactions are used for classification. All of them are discrete. It is impossible to use the Naive Bayesian Classifier based on the discrete distribution, the normal distribution and the kernel density estimation for this type of fraud detection. To solve this problem they developed clusterization algorithm which gives an input attribute value 1 when a real value of this parameter hits into some region of data clusterization, or else it has value 0. Finally the evaluative testing of transactions have been generated for training process and the testing process

John O. Awoyemi, Adebayo O. Adetunmbi and Samuel A. Oluwadare [4] in their paper in 2017 did a detailed comparative study of Credit card fraud detection using Machine Learning Techniques. Data mining technique is one prominent techniques used in solving credit fraud detection problem. Credit card fraud detection is the way toward recognizing those transactions that

are fraudulent into two classes of legitimate and fraudulent transactions. The paper examines the performance of Naïve Bayes, K-Nearest Neighbour and Logistic regression on highly skewed credit card fraud data. A hybrid technique of under- sampling and oversampling is carried out on the skewed data. The performance of fraud detection in credit card exchanges is greatly influenced by the sampling approach on dataset, selection of variables and detection techniques used. A hybrid technique of under-sampling and oversampling is completed on the skewed data. The three methods are applied on the raw and pre-processed data. The work is carried out in Python. The performance of the techniques is assessed based on accuracy, sensitivity, specificity, precision, Matthews correlation coefficient and adjusted classification rate. The outcomes shows of ideal accuracy for Naïve Bayes, K-Nearest Neighbour and Logistic regression classifiers are 97.92%, 97.69% and 54.86% respectively. The similar results show that K-Nearest Neighbour performs better compared to Naïve Bayes and Logistic regression techniques. They concluded through experiments that Random Forest technique shows most accuracy followed by Logistic Regression and Support Vector.

In 2017 J. Vimala Devi and K. S. Kavitha [5] did a detailed study of different Fraud Detection methods in Credit Card Transactions by using Classification Algorithms. This paper mainly focused on three algorithms Support Vector Machine, Random Forest and Decision Tree and has compared the results of these algorithms. These three classification techniques are chosen since the dataset is huge, an imbalanced one. Each of these methods has its own merits and demerits depending on the application. Therefore based on the implementation, the decision tree classification algorithm is displaying the best accuracy for the given dataset compared to other classification algorithms. The actual dataset is considered since the data is highly imbalanced one, classification may be done by oversampling the data.

Specific crime within the banking industry is credit card fraud. Credit card use has been expanded because of the fast development of E-business strategies. Credit card fraud also increased simultaneously. Avoidance is better than detection. So the current system prevented the credit card fraud by recognizing fraud in the application of the Credit card. To overcome the limitations like scalability issues, extreme imbalanced class and time constraints of the existing system, V. Mareeswari and G. Gunasekaran [6] proposed new algorithm along with the existing algorithm. Those limitations are overwhelmed by hybrid support vector machine (HSVM) along with communal and spike detection for credit card application fraud detection. Credit card fraud detection is done at the initial stage of credit card application and general fraud and crime activities are predicted by this system. In order to perform the identification of frauds, system uses the hybrid support vector machine (HSVM) for computing the weight of the each attribute for communal and spike detection for credit card application fraud detection. The system is working with existing two layers Communal Detection (CD) and Spike Detection (SD) along with HSVM for computing weight of each attribute of applicant's application. It uses to identify for legal behavior and data errors of the applicant. The CD algorithm performs with communal data of the applicant andworks in real time by exact or similar matches between categorical data, giving scores. Blacklist or attribute- oriented approach on the variable - size set of attributes is done by SD. After CD evaluation, SD takes care of the further process.

Support vector machine is the most used method for the pattern recognition and classification. In their approach it performs prediction and classification on the credit card dataset and classifies into two classes; fraud and genuine transaction. Credit card application is an online application process under which users enter their personal info for registration. According to proposed model when the user feeds the data to the online application form, this data is matched with the Whitelist (WL) database by CD algorithm, if it is a valid then pass it to SD for verifying civil score according to the Blacklist (BL) Database. Prior to this data matching process the attributes are learned by proposed algorithm and then the dual optimization process is performed during this the attribute are divided based on their priorities. CD and SD algorithm are computing weight of the attribute by using hybrid support vector machine prediction method. The application for a credit card is granted or else it is rejected based on the verifying of the both algorithm's outcomes. The main target that focused on this system is to preserve the credit fraud in the initial stage of the credit life cycle. The implementation of this algorithm in order to perform the identification of frauds, this system uses the Hybrid Support Vector Machine (HSVM) for computing the weight of the each attribute for communal and spike detection for credit card application fraud detection.

Credit card fraud detection is a significant method to prevent fraud events, which is usually categorized into anomaly detection and classifier-based detection techniques. Anomaly detection means calculating the distance between the data points in space. By calculating the distance between the incoming transaction and the cardholder's profile, an anomaly detection method can filter any incoming transaction which is inconsistent with the cardholder's profile. The classifier-based detection technique utilizes some supervised learning methods to train a classifier on thebasis of the given normal transactions and fraud ones. The supervised learning extracts fraud features from fraud transactions. However, both of these techniques have limitations. For the anomaly detection, it has no capacity to portray fraud features although it can portray cardholders' transaction behaviors. For the classifier-based detection, it fails to recognize different normal behaviors from different cardholders although it can catch fraudsters' behaviors. Transaction habits of an individual vary regularly since they are easily influenced by their incomes, resources, ages and characters. Thus their distribution evolves over time because of seasonality and new attack methodologies. This is known

as the problem of concept drift that is difficult to be resolved by the above detection methods. On the other hand, both of them don't know about the adaptive capacity of the model. For instance, a person may involve some new transaction behaviours in a particular period which has never occurred in his/her history. The vast majority of the proposed methods just keep the recent instances for model training, but don't consider the adaptiveness of the model.

Facing the existing challenges J. Changjun Jiang, Jiahui Song and Guanjun Liu [7] proposed a feedback mechanism which can adapt to the transaction of cardholder's behaviours seasonally. This paper proposes a novel fraud detection method that consists of four stages. To enrich a cardholder's behavioural patterns, first use clustering method, all cardholders are isolated into three groups depending up on the transaction amount, that is high, medium and low. Then introduces a sliding-window-based technique to aggregate the transactions in each group, i.e., derive a set of extra features from windows to characterize a cardholder's behavioural patterns. After pre-processing features, train a set of classifiers for each group using the data comprising of each specific behavioural pattern and extracted fraud features. Finally, use a classifier set prepared for a group is assigned to each cardholder in the group as his/her own behavioural patterns, and the classifier with the most rating score is viewed as his/her recent behavioural pattern. Based on the three classifier sets, propose a fraud detection method in which a feedback mechanism is taken to take care of the concept drift problem.

III. Comparison

Table I: Comparison Table

Paper Name	Comparison		
	Method	Advantages	Disadvantages
Credit Card Fraud Detection using Hidden Markov Model [2]	Detection of frauds using a Hidden Markov Model (HMM)	Scalable for handling large volumes of transactions	Suffer performance degradation
Credit Card Fraud Detection System [3]	Developed a fraud detection system and clustering model	Clustering model- allows fast monitoring of incoming transactions	The model is less accurate
Credit Card Fraud Detection using Machine Learning Techniques [4]	Three classifier models based on Naive Bayes, K-Nearest Neighbour and Logistic regression are used	Hybrid sampling greatly improves the performance	Logistic regression did not show better improvement
Fraud Detection in Credit Card Transactions by using Classification Algorithms [5]	Decision tree, Random forest and Support Vector Machines (SVM) are used to classify the data	Decision tree classification algorithm is displaying the best accuracy	Sometimes require long raining times and be deficient in available memory, if dealt with large databases
Prevention of Credit Card Fraud Detection based on HSVM [6]	Uses hybrid support vector machine (HSVM) for computing the weight of the each attribute	Improved the efficiency as well as performance less training time	Not suitable for large datasets
Credit Card Fraud Detection: A Novel Approach using Aggregation Strategy and Feedback Mechanism [7]	A feedback mechanism is introduced	Solved the concept drift problem. Increase the average recall and accuracy	Low performance

IV. Conclusion

Although there are several fraud detection techniques available today but none is able to detect all frauds completely when they are actually happening, they usually detect it after the fraud has been committed. This happens because a very minuscule number of transactions from the total transactions are actually fraudulent in nature. So we need a technology that can detect the fraudulent transaction when it is taking place so that it can be stopped then and there and that too in a minimum cost. So the major task of today is to build an accurate, precise and fast detecting fraud detection system for credit card frauds that can detect not only frauds happening over the internet like phishing and site cloning but also tampering with the credit card itself i.e. it signals an alarm when the tampered credit card is being used. The biggest disadvantage is that the existing procedures

will not produce the same outcomes in various contexts. They produce superior outcomes with one type of dataset while producing poor or unacceptable results with others.

The HMM and Logic regression suffer some sort of performance degradation. But HMM is scalable for handling large volumes of transactions. Some techniques like clustering model and Naive Bayesian Network though have high detection rates and gives high accuracy they are very expensive to train. Some like KNN and HSVM gives excellent results with small datasets but are not scalable to large datasets. The problem of concept drift can be solved by using a Feedback mechanism. Some techniques like decision tree and support vector displaying the best accuracy and gives better results on sampled and pre-processed data. A solution to these gaps by creating a hybrid of various techniques that are already used in fraud detection to cancel out their limitations and get enhanced performance.

References

[1] D. Coppola, "E-commerce worldwide - statistics & facts," 2021. [Online]. Available: https://www.statista.com/topics/871/online-shopping/

[2] A. Srivastava, A. Kundu, S. Sural, and A. Majumdar, "Credit card fraud detection using Hidden Markov Model," *IEEE Transactions on Dependable & Secure Computing*, vol. 5, no. 1, pp. 37-48, 2008.

[3] V. Filippov, L. Mukhanov, and B. Shchukin, "Credit card fraud detection system," *IEEE International Conference on Cybernetic Intelligent Systems,* 9-10 Sep. 2008, pp. 1-6.

[4] J. O. Awoyemi, O. A. Adetunmbi, and A. Samuel, "Credit card fraud detection using machine learning techniques," *International Conference Computer Networking and Informatics (ICCNI)*, 2017, pp. 1-9.

[5] J. V. Devi, and K. S. Kavitha, "Fraud detection in credit card transactions by using classification algorithms," *International Conference on Current Trends in Computer, Electrical, Electronics and Communication (ICCTCEEC)*, 2017.

[6] V. Mareeswari, and G. Gunasekaran, "Prevention of credit card fraud detection based on HSVM," *International Conference on Information Communication and Embedded Systems (ICICES)*, 2016.

[7] C. Jiang, J. Song, and G. Liu, "Credit card fraud detection: A novel approach using aggregation strategy and feedback mechanism," *IEEE Internet of Things Journal*, 2018.

International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 27-32
http://www.publishingindia.com/ijkbcs/

Classification of Spices using Machine Learning Techniques

Yukti Gupta[1*], Haneet Kour[2], Jatinder Manhas[3] and Vinod Sharma[4]
[1]M.Tech. Student, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: yuktigupta19@gmail.com
[2]PhD Research Scholar, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: haneetkour9@gmail.com
[3]Sr. Assistant Professor, Department of Computer Science and IT, Bhaderwah Campus, University of Jammu, Jammu and Kashmir, India. Email: manhas.jatinder@gmail.com
[4]Professor, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: vnodshrma@gmail.com
*Corresponding Author

Abstract: Machine learning (ML) has played a significant role in pattern recognition including fruits and vegetables classification. In this paper, comparative analysis of various ML techniques have been carried out for the identification of Spices. For the current work, ML techniques namely Naïve Bayes (NB), Decision Tree (DT), K-Nearest Neighbor (KNN), Random Forest (RF) and Support Vector Machine (SVM) have been undertaken. The main aim of the current study is to find out the most appropriate ML approach for Spices recognition. The experimental study has been performed on primary dataset of Spices. This dataset consists of 1000 images of five different Spices including clove, green cardamom, cinnamon, black pepper and curry leaf. The performance of the ML techniques have been analyzed on the basis of four parameters i.e. accuracy, precision, recall and f1-score. Out of five implemented ML models, best performance has been predicted by SVM approach with accuracy of 94.5%, precision of 95%, and recall of 94% with f1-score of 0.95.

Keywords: Decision tree, K-Nearest neighbor, Machine learning, Spices recognition, Support vector machine.

I. Introduction

India is referred to the Land of Spices and it is the world's largest manufacturer, consumer and exporter of Spices. It commands a formidable position in the world Spice trade. The total value of Spices produced in India is about 7000 crores annually. The country produces about 75 out of 109 types of Spices listed by the International Organization for Standardization (ISO) [1].

Every Spice has its own flavor and essence, and its addition or omission can genuinely make or destroy a dish. Spices include remarkable nutritional values and have numerous health benefits including antioxidant and anti-inflammatory properties, glucose-lowering effects, appetite control, regulates metabolism, aids in weight loss, improves memory and brain function. Spices are being used by several medical industries such as cosmetic, pharmaceutical and aromatic as perfumery [2].

It is quite difficult to distinguish between Spices as many Spices look similar. The recognition of Spices used in our daily basis food is gaining more importance in our daily life. Research on Spices recognition and classification is vital for several economic sectors, both for the wholesale and retail markets, as well as for the processing industries.

Consuming healthy and good quality fruits and vegetables are the utmost necessity of the purchaser. Hence, automation in food industries are developing nowadays because it is improbable for humans to manually audit the fruits and vegetables as it requires a large number of workers along with a lot of time and effort. In recent years, numerous machine learning (ML) algorithms have been applied with various different feature description methods for fruits and vegetables classification in several real-world applications [3].

In this study, comparative analysis of several ML techniques including Naive Bayes (NB), Decision Tree (DT), K-Nearest Neighbor (KNN), Random Forest (RF) and SVM (Support Vector Machine) have been carried out for the recognition of Spices. The main motive of the current research work is to implement ML based system for the detection of Spices for food image researchers and dietician to make an appropriate analysis of nutrition and other type of health hazards. Classifying a particular kind of Spice will enable us to distinguish it from another kind. For the current study, five different classes of

Spices including clove, green cardamom, cinnamon, black pepper and curry leaf have been considered.

This paper is divided into five sections: Section II illustrates the literature related to the current work, Section III demonstrates materials and methodology adopted in the current work, Section IV shows experimental results and analysis, and Section V explains the conclusion and future scope.

II. Literature Review

In the recent past, ML techniques have been applied in agricultural domain. These approaches have been found to be effective for the recognition of fruits and vegetables.

In the year 2012, Zhang and Wu [4] used multi-class kernel support vector machine (kSVM) for fruit classification. The SVM approach with max-wins-voting strategy and gaussian radial basis kernel achieved the classification accuracy of 88.2%.

Next year, Sanjayan *et al.* (2013) [5] employed ML techniques for the evaluation of the nutritional value of ripened fruits of a commonly known medicinal plant Morinda tinctoria. The work targeted on the estimation of Ash content, protein, carbohydrate, vitamins and mineral content of Morinda tinctoria fruit. The difference in nutritional content between fresh and dried fruits had been also evaluated. The ash content for fresh and dry fruits are 4% and 1.6% respectively.

After two years, Surya and Priya (2015) [6] proposed a food image recognition system for calculating the calorie and nutrition values using SVM classifier. In this approach, the area, volume and size of the undertaken food item has been used to measure the calorie and nutrition precisely.

Femling *et al.* (2018) [7] depicted MobileNet based approach for recognizing fruit and vegetables within the retail market using photographs captured with a video camera connected to the system. The proposed system achieved maximum accuracy of 97%.

Dheir *et al.* (2019) [8] introduced a system for classifying nuts types. This study applied Convolutional Neural Network (CNN) algorithm to the dataset of five classes of Nuts that contains 2868 images. This method achieves excellent results with accuracy of 98%.

Bhattacharya and Mukherjee (2020) [9] implemented Neural Network based spices image classification system for Indian food images. The system is based on segmentation and recognition of spices images that identifies spice types. The dataset contains 4 types of spices containing a total of 80 spices images. The experimental study achieved the accuracy of 96.7%.

In the same year, Rojas-Aranda *et al.* (2020) [10] developed an image classification method based on lightweight CNN. A new dataset of images is designed belonging to three different classes of fruits, within or without plastic bags. Various input features have been added to the CNN architecture to improve classification accuracy. The experimental results predicted the 95% classification accuracy for fruits without plastic bag, and 93% for fruits inside plastic bag.

Next year, Shaikh *et al.* (2021) [11] applied faster Region-based Convolutional Neural Network (RCNN) model to detect and classify fruits as affected or unaffected based on their surface. The experimental study achieved accuracy of 60-75% for healthy apple, 60-70% for bad apple, 85-99% for healthy pear, 80-98% for bad pear, 80-97% for healthy banana and 70-80% for bad banana.

In the same year, Patil *et al.* (2021) [12] created grading and classifying techniques for dragon fruits with ML and Deep Learning (DL) algorithms i.e. SVM, CNN and Artificial Neural Network (ANN) based on a thorough review of algorithms available for fruit quality detection and classification using numerous characteristics of fruits and vegetables.

Zakeri *et al.* (2021) [13] presented a computer vision-based approach for grading jujube fruits using ML techniques. This method takes into account most of the important pricing factors and can be used to increase farmers' profits. They used various classifiers and training methods to get the best results, and with the help of DT, they were able to achieve a classification accuracy of 98.8%.

III. Materials and Methods

The goal of the current research is to identify Spices using ML techniques. This paper presents a comparative study of different ML techniques for the classification of Spices in order to find out the most suitable technique for Spices classification. The experimental study has been done on Python platform using *Spyder Interface* [14]. Fig. 1 represents the methodology for the current study.

A. Data Collection

The proposed study has been done on primary dataset as the domain of work is new and no dataset is available on the internet. For the current study, five different classes of Spices including *clove*, *green cardamom*, *cinnamon*, *black pepper* and *curry leaf* have been collected.

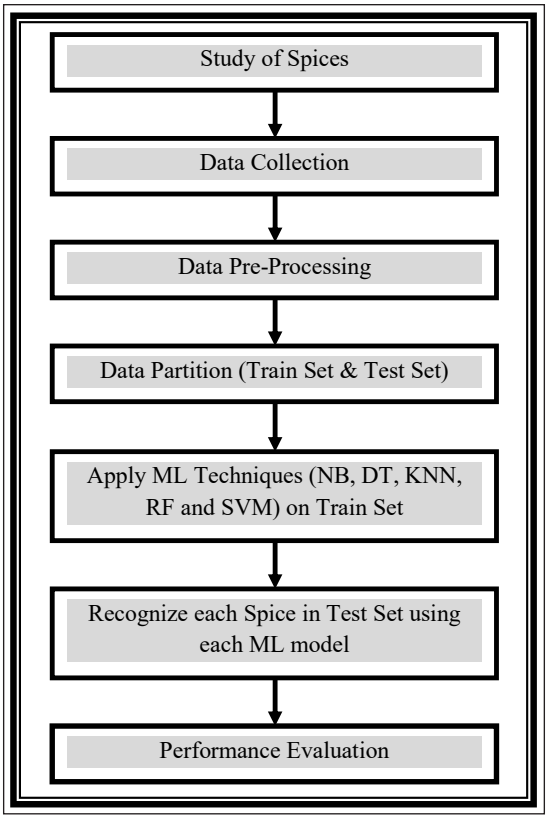


Fig. 1: Work Flow of Current Research

Each Spice image has been captured by using high quality camera and these images have been taken from different angles in order to introduce sufficient variations. White sheet of paper has been used as background of each image. The collected dataset consists of 1000 images, belonging to five different classes of undertaken Spices. There are 200 images of each Spice taken in this study. Fig. 2 depicts Sample of Spices in the collected dataset.

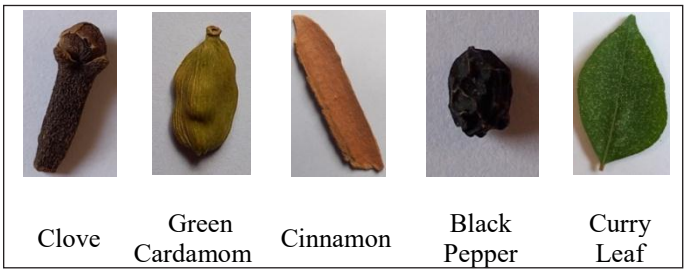


Fig. 2: Sample of Spices Dataset

B. Data Pre-Processing and Data Partition

The data pre-processing steps include cropping, renaming and resizing of the images to the standard dimensions, in order to reduce computational load. The collected images of Spices are in the RGB format and these images have been converted to gray-scale format.

After data pre-processing, data has been partitioned in the ratio 80:20 into two sets: *Train Set* and *Test Set*. Train Set consists of 800 images whereas Test Set contains 200 images.

C. Implementation using Machine Learning Techniques

After data partition, ML techniques have been implemented on Train Set to build the model for Spices recognition. For the current work, five ML techniques namely NB [15], DT [16], KNN [17], RF [18] and SVM [19] have been undertaken. For the implementation of each ML model, feature extraction approach i.e. Discrete Wavelet Transform (DWT) [20] has been applied. DWT has been implemented at Level-1 using *Haar* approach; which computes four coefficients i.e. cA (approximation coefficient), cH (horizontal coefficient), cV (vertical coefficient) and cD (diagonal coefficient). For each coefficient, five features i.e. contrast, energy, dissimilarity, homogeneity, and correlation have been computed. Thus, for each Spice image, total 20 features have been extracted and these features have been saved in *.csv* file. The final size of *.csv* file becomes 1000*21 since the dataset has 1000 images and 20 are the input extracted features and 01 is the output class label.

Each undertaken ML technique has been applied to build the Spices recognition model using Train Set. For each ML technique, the extracted features have been fed as input feature vector and the class label has been used as predictors.

In this study, NB algorithm has been predicted on the basis of the probability of Spices by applying Bayes theorem. It assumes that the occurrence of certain feature of Spice image is independent of the occurrence of other features.

In DT model, decision nodes are used to make any decision and leaf nodes are the output of these decisions. Decisions have been taken on the basis of features of the Spices dataset using criterion parameter with entropy value as a metric for impurity.

KNN model has been implemented using ‘feature similarity’ for predicting the values of new data points and the value of k = 7 with *Euclidean* distance metric has been taken.

For the current work, RF technique has been constructed using 20 small decision trees known as estimators on data samples with max depth 5 and gets the prediction from each of them. By means of voting best solution has been selected.

SVM approach has created best decision boundary (hyperplane) that segregated the 5 different classes of Spices in multidimensional space using ECOC (error correcting code output) with kernel function.

IV. Result and Analysis

The main objective of this study is to develop a Spices Recognition framework based on ML techniques in order to recognize different Spices. To test the performance of the

proposed system, experiments were performed on a dataset of 1000 images of five Spices that includes: clove, green cardamom, cinnamon, black pepper and curry leaf. The training dataset contains 800 images whereas the testing dataset contains 200 images. DWT feature extractor has been employed to compute the texture features: correlation, energy, contrast, homogeneity, and dissimilarity.

All the undertaken ML approaches have been evaluated based on four performance measures i.e. accuracy, recall, precision and f1-score. Accuracy depicts about how accurately these ML techniques can recognize the given spice. In addition to it, it is required to avoid the misclassification of the given spice by undertaken ML approach. To deal with such issues, recall and precision parameters have been considered. Recall tells about *false negative* whereas precision tells about *false positive*. It is needed that the undertaken ML technique should predict lowest false positive and false negative. If one technique depicts *high false positive and low false negative*; and the other one predicts *low false positive and high false negative*, then f1-score is employed to evaluate the model performance as it computes the harmonic mean of precision and recall.

From the experimental results, it has been found out that all the ML techniques can be applied for the Spices classification. Fig. 3 presents experiment results for NB. The experiment results for DT are shown in Fig. 4. Fig. 5 represents the experiment results for KNN. Fig. 6 depicts experiment results for RF. Fig. 7 shows the experiment results for SVM.

```
Confusion Matrix:
[[25  1  0 12  0]
 [ 4 33  1  1  0]
 [ 1  0 31  0  8]
 [ 0  2  0 38  0]
 [ 6  0  3  0 34]]
Classification Report:
              precision    recall  f1-score   support

         0.0       0.69      0.66      0.68        38
         1.0       0.92      0.85      0.88        39
         2.0       0.89      0.78      0.83        40
         3.0       0.75      0.95      0.84        40
         4.0       0.81      0.79      0.80        43

    accuracy                           0.81       200
   macro avg       0.81      0.80      0.80       200
weighted avg       0.81      0.81      0.80       200

Accuracy: 0.805
```

Fig. 3: Experimental Results for NB

```
Confusion Matrix:
[[30  1  0  1  1]
 [ 1 43  2  2  0]
 [ 1  1 30  1  4]
 [ 4  1  0 33  0]
 [ 0  0  5  0 39]]
Classification Report:
              precision    recall  f1-score   support

         0.0       0.83      0.91      0.87        33
         1.0       0.93      0.90      0.91        48
         2.0       0.81      0.81      0.81        37
         3.0       0.89      0.87      0.88        38
         4.0       0.89      0.89      0.89        44

    accuracy                           0.88       200
   macro avg       0.87      0.87      0.87       200
weighted avg       0.88      0.88      0.88       200

Accuracy: 0.875
```

Fig. 4: Experimental Results for DT

```
Confusion Matrix:
[[33  0  0  6  0]
 [ 3 34  0  0  0]
 [ 0  1 32  0  7]
 [ 1  0  0 40  0]
 [ 0  0  0  0 43]]
Classification Report:
              precision    recall  f1-score   support

         0.0       0.89      0.85      0.87        39
         1.0       0.97      0.92      0.94        37
         2.0       1.00      0.80      0.89        40
         3.0       0.87      0.98      0.92        41
         4.0       0.86      1.00      0.92        43

    accuracy                           0.91       200
   macro avg       0.92      0.91      0.91       200
weighted avg       0.92      0.91      0.91       200

Accuracy: 0.91
```

Fig. 5: Experimental Results for KNN

```
Confusion Matrix:
[[22  1  0  3  0]
 [ 1 47  0  1  0]
 [ 0  2 38  0  1]
 [ 4  0  0 40  0]
 [ 0  0  1  0 39]]
Classification Report:
              precision    recall  f1-score   support

         0.0       0.81      0.85      0.83        26
         1.0       0.94      0.96      0.95        49
         2.0       0.97      0.93      0.95        41
         3.0       0.91      0.91      0.91        44
         4.0       0.97      0.97      0.97        40

    accuracy                           0.93       200
   macro avg       0.92      0.92      0.92       200
weighted avg       0.93      0.93      0.93       200

Accuracy: 0.93
```

Fig. 6: Experimental Results for RF

```
Confusion Matrix:
[[23  0  0  3  0]
 [ 1 47  0  1  0]
 [ 1  1 39  0  0]
 [ 4  0  0 40  0]
 [ 0  0  0  0 40]]
Classification Report:
              precision    recall  f1-score   support

         0.0       0.79      0.88      0.84        26
         1.0       0.98      0.96      0.97        49
         2.0       1.00      0.95      0.97        41
         3.0       0.91      0.91      0.91        44
         4.0       1.00      1.00      1.00        40

    accuracy                           0.94       200
   macro avg       0.94      0.94      0.94       200
weighted avg       0.95      0.94      0.95       200

Accuracy: 0.945
```

Fig. 7: Experimental Results for SVM

Table I presents experimental results of five ML techniques for Spices classification. Fig. 8 shows the comparative analysis of all the undertaken ML techniques. NB predicted lowest value for accuracy of 80.5%, precision of 81%, recall and f1-score of 81% and 0.80. DT model achieved an accuracy of 87.5%, precision of 88%, recall of 88% and f1-score of 0.88. KNN model attained an accuracy of 91%, precision of 92%, recall of 91% and f1-score of 0.91. RF model obtained an accuracy of 93%, precision of 93% and recall of 93% with f1-score of 0.93. SVM model showed best results as it achieved higher value for accuracy, precision, recall and f1-score being 94.5%, 95%, 94% and 0.95 respectively.

TABLE I: EXPERIMENTAL RESULTS OF FIVE ML TECHNIQUES FOR SPICES RECOGNITION

Traditional ML Technique	Accuracy	Precision	Recall	F1-Score
NB	80.5%	81%	81%	0.80
DT	87.5%	88%	88%	0.88
KNN	91%	92%	91%	0.91
RF	93%	93%	93%	0.93
SVM	94.5%	95%	94%	0.95

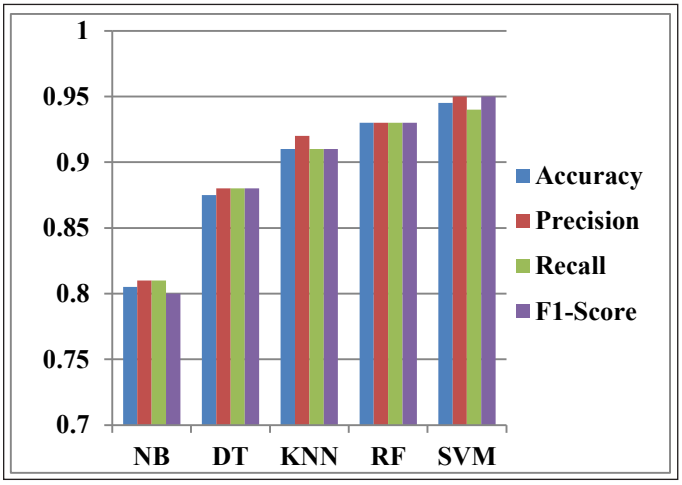


Fig. 8: Comparative Analysis of All Undertaken ML Techniques

V. CONCLUSION AND FUTURE SCOPE

From experimental results, it has been concluded that ML approaches can be used for the detection of Spices. The main purpose of the current work is to implement various ML techniques for the recognition of different Spices in order to reveal out the better approach for Spices classification. Out of all the ML techniques taken in this study, SVM has been found to be more effective technique for the Spices recognition with an accuracy of 94.5%, precision of 95% and recall of 94% with f1-score of 0.95. This study is limited to only five Spices. Future work will involve recognition of other different varieties of Spices using combined features more efficiently.

References

[1] K. Ponmozhi, and R. Jose, "A review on spices classification and recognition," *International Journal of Mathematics and Computer*, vol. 12, no. 1, pp. 25-32, 2022.

[2] A. Jiang, "Health benefits of culinary herbs and spices," *Journal of AOAC International*, vol. 102, no. 2, pp. 395-411, 2019, doi: 10.5740/jaoacint.18-0418.

[3] Machine Learning. [Online]. Available: https://en.m.wikipedia.org/wiki/Machine_learning#:~:text=Machine%20learning%20(ML)%20is%20the,a%20part%20of%20artificial%20intelligence

[4] Y.-D. Zhang, and L. Wu, "Classification of fruits using computer vision and a multiclass support vector machine," *Sensors*, vol. 12, no. 9, pp. 12489-12505, 2012, doi: 10.3390/s120912489.

[5] V. Anuradha, A. Praveena, and K. Sanjayan, "Nutritive analysis of fresh and dry fruits of Morinda tinctoria," *International Journal of Current Microbiology and Applied Sciences*, vol. 2, no. 3, pp. 65-74, 2013.

[6] R. Surya, and S. S. Priya, "Food image recognition using SVM classifier for measuring calorie and nutrition values," *International Journal of Scientific & Engineering Research*, vol. 6, no. 4, pp. 324-328, 2015.

[7] F. Femling, A. Olsson, and F. Fernandez, "Fruit and vegetable identification using machine learning for retail applications," *International Conference on Signal-Image Technology & Internet-Based Systems (SITIS)*, 2018, doi: 10.1109/SITIS.2018.00013.

[8] I. M. Dheir, A. S. A. Mettleq, and A. A. Elsharif, "Nuts types classification using deep learning," *International Journal of Academic Information Systems Research (IJAISR)*, vol. 3, no. 12, pp. 12-17, 2019.

[9] S. Bhattacharya, and J. Mukherjee, "Recognisation of Indian spices based on the combination of features and comparison using neural network," *International Research Journal of Engineering and Technology (IRJET)*, vol. 7, no. 6, pp. 2418-2423, 2020.

[10] J. Aranda, J. Varela, J. Tello, and G. Ramirez, "Fruit classification for retail stores using deep learning," *in K.* F. Mora, J. A. Marín, J. Cerda, J. Carrasco-Ochoa, J. Martínez-Trinidad, and J. Olvera-López (Eds.) Pattern Recognition, MCPR 2020, *Lecture Notes in Computer Science,* vol. 12088, Springer Nature Switzerland AG, , 2020, doi: 10.1007/978-3-030-49076-8_1.

[11] H. Shaikh, Y. Wagh, S. Shinde, and S. Patil, "Classification of affected fruits using machine learning," *International Journal of Engineering Research & Technology (IJERT)*, vol. 9, no. 3, 2021.

[12] P. Patil, S. Lande, V. Nagalkar, S. Nikam, and G. Wakchaure, "Grading and sorting technique of dragon fruits using machine learning algorithms," *Journal of Agriculture and Food Research,* vol. 4, no. 3, 2021. [Online]. Available: https://doi.org/10.1016/j.jafr.2021.100118

[13] A. Zakeri, R. Hedayati, M. Khedmati, and M. Taghipour-Gorjikolaie, "Classification of jujube fruit based on several pricing factors using machine learning methods," 2021. [Online]. Available: https://doi.org/10.48550/arXiv.2111.00112

[14] Spyder Interface. [Online]. Available: https://anaconda.org/anaconda/spyder

[15] C. Prasad, "An efficient heart disease detection system utilizing Naïve Bayes classification," *Journal of Applied Information Science*, vol. 9, no. 1, pp. 22-25, 2021.

[16] B. Poornima, "An improved decision tree classification for breast cancer detection with optimal parameters," *Journal of Applied Information Science*, vol. 9, no. 1, pp. 19-21, 2021.

[17] G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, "KNN model-based approach in classification," *OTM Confederated International Conferences*, 2003, doi: 10.1007/978-3-540-39964-3_62.

[18] Gokul S., and N. Banu P. K., "Analysis of phishing detection using logistic regression and random forest," *Journal of Applied Information Science*, vol. 8, no. 1&2, pp. 7-13, 2020.

[19] A. MS, and Manjesh R., "Fruit recognition using SVM technique," *International Journal of Science and Research (IJSR)*, vol. 9, no. 2, pp. 1210-1214, 2020, doi: 10.21275/SR20212155833.

[20] S. Thakral, and P. Manhas, "Image processing by using different types of discrete wavelet transform," *Advanced Informatics for Computing Research*, 2018.

International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 33-39
http://www.publishingindia.com/ijkbcs/

A Comprehensive Survey on Support Vector Machines for Intrusion Detection System

Akram Salim Khanfar[1], Firdous Ahmad Lone[2] and MD Moizuddin[3*]
[1]Faculty, King Saud University, Saudi Arabia. Email: akram.khanfar2011@gmail.com
[2]Faculty, King Saud University, Saudi Arabia. Email: lonefirdous686@gmail.com
[3]Faculty, King Saud University, Saudi Arabia. Email: moizqa12@gmail.com
*Corresponding Author

Abstract: Machine learning is a widely interdisciplinary field centered on theories from cognitive science, computer science, statistics, optimization and many other theoretical and mathematical disciplines. Classification is a supervised learning technique used in machine learning to evaluate a given dataset and to create a model that divides data into a desired and distinct number of groups. The strength of SVMs lies in their use of nonlinear kernel features that map input into high-dimensional spaces of features implicitly. We'll address the value of SVMs in this survey article. Discussing their SVM tuning parameters as well. The main purpose of this paper is to include detailed studies on SVM implementations by contrasting the current ML models with the SVM versions, also poses the problems of the intrusion detection method of the support vector machines, and also this paper provides researchers with a summary of the SVM that assists in their future analysis.

Keywords: Data Mining (DM), Intrusion Detection System (IDS), Machine Learning (ML), Optimization, Support Vector Machines (SVMs).

I. Introduction

The Machine learning is a branch of Artificial Intelligence, and the development of techniques, methods, and algorithms is an evolving field. These algorithms allow machines to understand the processes, assignments, and decisions that are made. In Machine learning, classification is supervised approach and using the classification techniques helps to classify the data into classes so that it can easily be identified. Classification is one of the techniques for machine learning that allows to group data to extract characteristics and forecast future effects. There are plenty of machine learning various algorithms in order to distinguish the data into classes. Support Vector Machine is among the most popular techniques for constructing models for machine learning. With less power to store, it has tremendous accuracy.

Machine learning (ML) algorithms for computational solutions can be used to perform efficiently. In recent years, there have been plenty of work carried out in the field of Support vector machines (SVM). SVM have shown good results in classification, generalization performance on many problems.

This paper is organized as follows: Section II presents the SVM's associated work as well as the SVM's findings and Section III presents Support Vector Machine and also covers the importance of SVM, SVM's tuning parameters. In terms of the Intrusion Detection Method, Section IV outlines the study of different SVM implementations. The SVM challenges are outlined in Section V.

Related Work

SVM is one of the powerful supervised methods for solving the classification and regression problems and also to provide the optimal solution [1] [2]. Through each point, it has the capacity to solve even the shortest classification problem. SVM offers solutions to problems relating to the fitting of training problems within a personal computer or workstation's storage capacity, since SVM does not need any matrix equations and is less likely to have issues with numerical formats [2]. SVM is an effective tool for target object selection and detection in a medical imaging device and microcalcifications [3]. SVM provides the best results on application to face detection and reuters collection and also give the new technique for implementing the SVM algorithm efficiently. SVMs have attractive qualities, such as classification accuracy, computational models, simple geometric interpretation, and stronger intrusion detection efficiency [4]. In classification methods, the performance of the SVM depends on understanding the necessary parameters and soft-margin coefficient of the kernel function [5]. The previous studies showed that, relative to linear SVM algorithms that use a single CPU, parallel SVMs can reach large speeds [6]. SVMs have been extended to several machine learning tasks. From an adequate set of kernel functions, it creates learning frameworks and architectures [7]. By introducing the information geometric of Riemannian geometry structure induced by the kernel, it is possible to improve the SVMs classifier performance [7]. With Dynamic Time Warping distance measurement, an

improved SVM scheme was used as a feature for the SVM Classifier [8] [9]. SVM classifiers are used in many fields for text classification [10], facial components detection and tracks facial and emotional expression recognition [11] [12]. SVM identify sets of genes with a common function using expression data analysis [13] and also provides best performance compared to Parzen windows and Fisher's linear discriminant [14]. SVMs were used to conduct malfunction categorization [15]. With a rule-based decision tree (RBDT), a multi-class support vector machine classification model is used to identify the faults of water quality sensors due to its reliability and generalization [16] [17].

II. Support Vector Machine

SVM is a supervised algorithm. The key goals of SVM are classification and regression. It's based on the concept of statistics and Vapnik-Chervonenkis dimensions [18] [19] [20]. The main purpose is to find the optimal hyperplane by dividing the data points into two components and maximizing the margin, in this way it solves the classification and regression problems. In 2D, hyperplane is line and in 3D it is a plane (also called n-dimensional line) [18]. This process contains data from training and testing data from research. The algorithm generates an optimal hyperplane in training data (supervised learning), which categorizes new instances and then the evaluation process is carried out from the constructed model. SVM takes minimal training area and less processing time and also avoids overfitting problems [21] [18].

Suppose we have two groups of labels and they are plotted on a graph as seen in Fig. 1 below.

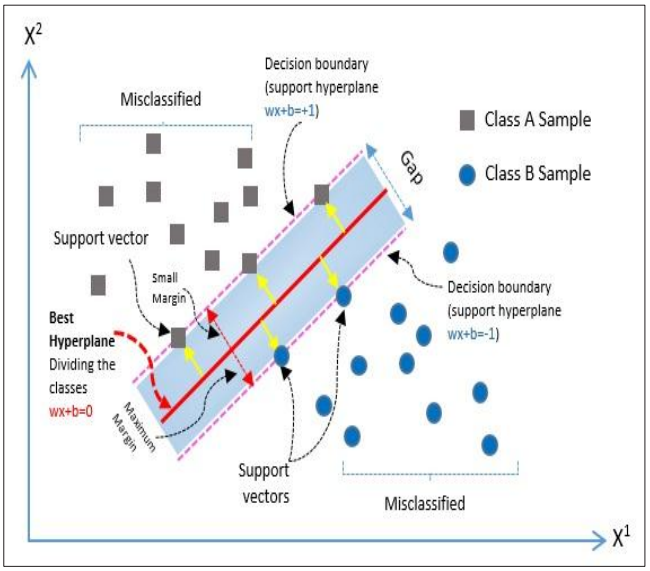


Fig. 1: The Concept of SVM-Hyperplane

Mathematically, for a hyperplane which divides the given data into two classes is represented by an equation as follows:

$$wx + b = 0 \quad (22)$$

A. Importance of Support Vector Machines

SVMs have been shown to work well, like analysis of remote protein sequence homology and recognition of gene transcription sites, it is used to examine expression data. It helps to overcome the challenges caused by high-dimensional data on gene expression data. One explanation for their effectiveness is that SVMs are targeted at minimizing error rates directly.

B. SVM Tuning Parameters

The categorization of groups in real-world situations is time-consuming. In order to overcome this, we use the parameters such as kernel, regularization, gamma, margin which are called tuning parameters or varying parameters.

C. Kernel

It is the backbone of SVM algorithm. It can resolve any complicated task. The SVM uses kernel-defined mathematical functions. These functions, such as linear, non-linear, polynomial, radial and sigmoid, take the data input training set and transform a non-linear decision into a high-dimensional space linear equation. In general, 2D data is converted into 3D data, as shown in Fig. 2.

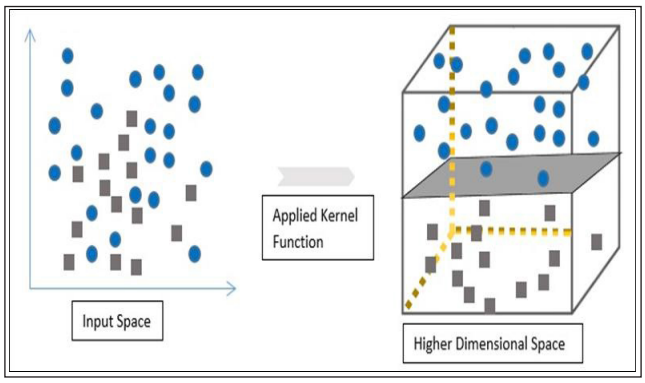


Fig. 2: Applied Kernel Function to Input Space

Mathematically, Kernel Function Equation is represented as:

$$K(x)= 1, \text{ if } ||x||<=1 \quad (22)$$

$$K(x)=0, \text{ Otherwise} \quad (22)$$

D. Regularization

Regularization is a computational approach that seeks to make a model structure as easy as possible to create. The optimized model will minimize the implications of overfitting at a small cost of performance. It is a way to reduce the model's complexity by penalizing the overfitting loss function. In mathematical

terms it is adding the sum of the weights to the cost function and also represented as:

$$J = \frac{1}{2}\sum_{i=1}^{M}(d_i - y_i)^2 + \lambda\frac{1}{2}\|w\|^2 \quad (17)$$

E. Gamma

Gamma is a kernel parameter of the RBF. It is known as the decision boundary. The greater Gamma value means more curvature, the decision region is high, and less curvature is indicated by the lower value, and the decision region is smaller. The gamma equation is mathematically expressed as:

$$\gamma = -\frac{1}{2\sigma^2} \quad (23)$$

F. Margin

It is the distance between the decision boundary to the closest data point in the given set of class. Models with higher margins yields better classification and performance.

G. Types of SVM Classifier

Support vectors are actually the coordinates that plots each data item as points, which are closest to the hyperplane. SVM choses the extreme points and these extreme points are called support vectors. There are two type of SVM Classifiers: Linear SVM, Non-linear SVM as shown in Fig. 3.

Linear SVM: In linear SVM, the data points are classified into two classes by a straight line, which divides the two classes. In this the data is linearly arranged and can be easily separated by a straight line. In linear, we use x and y as dimensions.

Non-Linear SVM: We cannot separate the non-linear data by a straight line, so to divide non-linear data a third dimension z is required.

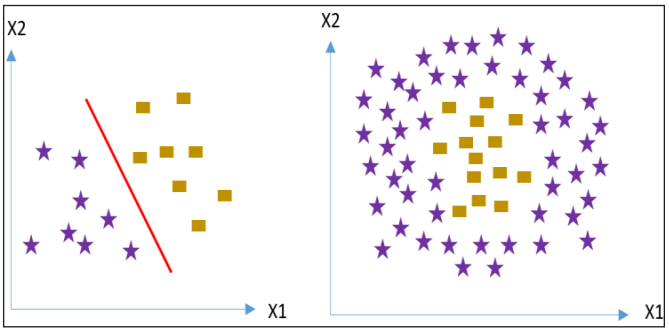


Fig. 3: Linear & Non-Linear SVMs (Ghosh et al., 2019)

III. Some of the Related Work of SVM in Terms of Intrusion Detection System

There are several SVM implementations using datasets. A list of the SVM implementations most used as seen in Table I. In the Table I, highlights the work carried out during the year 2016 to 2020.

Table I: Implementations of SVM in Terms of Intrusion Detection System

Challenges	Addressed by
Shortcomings related to the accuracy, number of selected features, and execution time	(Safaldin *et al*., 2020)
Managing changing data	(Jackson, 2002)
Low accuracy in weighted majority voting (WMO) approach	(Aburomman & Ibne Reaz, 2016)
Need complex features such as multi-classifier and feature selection in IDS	(Yang *et al*., 2016)
Lead to a long detection delay in the practical application scenario	(Gao *et al*., 2019)
Does not give the detailed information on the structure and characteristics of the malware	(Vinayakumar *et al*., 2019)
False positive rate	(da Costa *et al*., 2019)
Time cost in the data optimization stage and support for online processing	(Ren *et al*., 2019)
Large scale network will require additional infrastructure	(Taher *et al*., 2019)
High dimensionality of problems	(Tavara, 2019)
Address severe class imbalance, network traffic variability	(Gu & Lu, 2020)
Only a few numbers of works have been designed for detecting anomalies in the hosts	(Hosseinzadeh *et al*., 2020)

Challenges	Addressed by
Curse of dimensionality, Irrelevant features	(Hosseinzadeh *et al*., 2020)
Bolt loosening detection	(F. Wang *et al*., 2020)
Algorithmic complexity	(Cervantes *et al*., 2020)
Development of optimal classifiers for multi-class problems	(Cervantes *et al*., 2020)
The selection of kernel for a problem, Choosing good quality kernel parameters	(Nayak *et al*., 2015)
Speed and size in training and testing, training for very large datasets	(Tavara, 2019)
There are no theories concerning how to choose good kernel functions in a data-dependent way	(Byun & Lee, 2002)
Parallel algorithmic approaches for implementation of SVMs	(Tavara, 2019)
Exceedingly high time complexity in DTW computation	(Thapanan Janyalikit, 2016)

Conducted a search of the keyword 'Support Vector Machines' on numerous search engines like IEEE, Google Scholar, Elsevier, and Springer and the result shown in the below Table II.

Table II: Searched Keyword Support Vector Machines

Support Vector Machines	Results
Google Scholar	2,200,000
IEEE	63,369
Elsevier	8,896
Springer	24,737

IV. Challenges of SVM

Many researchers have noted a number of challenges in data mining science. Some of these are shown in Table III and need further focus from study.

Table III: Implementations of SVM in Terms of Intrusion Detection System

Authors & Year	Proposed Model/Method	Dataset Used	AD	FS	E. Criteria
(Yang *et al*., 2016)	ICPSO-SVM	KDD Cup 1999	Yes	Yes	FPR, DR
(Rebai, 2016)	ML-MKL	UCI dataset COIL-20	Yes	Yes	Accuracy, TP, TN
(Aburom man & Reaz, 2017)	LDA-PCA	KDD99	Yes	Yes	Accuracy, FP
(Liang *et al*., 2019)	Clustering-SVM Ensemble Method	NSL-KDD	Yes	Yes	Accuracy, Time, DR, FAR
(Safaldin *et al*., 2020)	GWOSVM-IDS	NSL-KDD	Yes	Yes	Accuracy, No. of features, Time, FR, DR
(H. Wang *et al*., 2017)	LMDRT-SVM	Gure-KDD dataset	Yes	Yes	Accuracy
(Al-Qatf *et al*., 2018)	STL-IDS	NSL-KDD	Yes	Yes	Accuracy
(Gu *et al*., 2019)	DT-EnSVM	NSL-KDD	Yes	Yes	Accuracy, DR, FAR
(Saleh *et al*., 2019)	Hybrid IDS	KDD Cup 1999, NSL-KDD, Kyoto 2006+ dataset	Yes	Yes	DR, Sensitivity, Specificity, Precision
(Tao et al., 2018)	FWP-SVM-GA	KDD Cup 1999	Yes	Yes	FPR, FNR, DR, Accuracy
(Kabir et al., 2018)	OA-LS-SVM	KDD Cup 1999	Yes	Yes	Accuracy, FAR
(Kavitha & Elango, 2020)	GRRF-FWSVM	KDD Cup 1999	Yes	Yes	Precision Recall F-Score
(Al Shorman et al., 2020) the number of Internet of Things (IoT	GWO-OCSVM	N-BaIoT	Yes	Yes	Average detection time, TPR, FPR

Authors & Year	Proposed Model/Method	Dataset Used	AD	FS	E. Criteria
(Roopa Devi & Suganthe, 2020)	HGWCSO with ETSVM	NSL-KDD	Yes	Yes	Precision, Recall, Sensitivity, Specificity, Accuracy
(Roopa Devi & Suganthe, 2020)	ROC and Confusion Matrix	NSL-KDD	Yes	Yes	Accuracy, Error, time
(Mighan & Kahani, 2020)	hybrid SAE–SVM scheme	NSK KDD, Kyoto, CDMC 2012, KDD Cup 1999 and UNB ISCX 2012	Yes	Yes	Accuracy, Recall, Time, Precision, F-measure
(Kumar & Ramasamy, 2020)	CSO-SVM algorithm	NSL-KDD	Yes	Yes	Accuracy, Recall Sensitivity Specificity Precision
(Ye *et al*., 2019)	GOA-SVM	KDD Cup, different datasets	Yes	Yes	Time, Accuracy

V. Conclusion and Future work

SVM are based on the concept of statistical learning theory. In SVM the inputs are placed in 3D space, where the different class groups are mapped using the mathematical functions. For each parameter SVM represents an optimal solution. This SVM algorithm is different from the other algorithms in terms working and in mapping the inputs on space. Kernel is important parameter on which this SVM works. In this paper, we have discussed the related works of SVM and highlighted the challenges of SVM, which will be helpful for the researchers in their future work. Lots of research work is ongoing to extend the scope and increase the performance and accuracy in detecting the Intrusions in a network. In future, we will focus on SVM applications and comparison of SVM methods with other machine learning techniques.

References

[1] J. Cervantes, F. Garcia-Lamont, L. Rodríguez-Mazahua, and A. Lopez, "A comprehensive survey on support vector machine classification: Applications, challenges and trends," *Neurocomputing*, 2020, doi: 10.1016/j.neucom.2019.10.118.

[2] S. Ghosh, A. Dasgupta, and A. Swetapadma, "A study on support vector machine based linear and non-linear pattern classification," *Proc. Int. Conf. Intell. Sustain. Syst. ICISS 2019*, no. Iciss, 2019, pp. 24-28, doi: 10.1109/ISS1.2019.8908018.

[3] Q. Yang, H. Fu, and T. Zhu, "An optimization method for parameters of SVM in network intrusion detection system," *Proc. - 12th Annu. Int. Conf. Distrib. Comput. Sens. Syst. DCOSS 2016*, 2016, pp. 136-142, doi: 10.1109/DCOSS.2016.48.

[4] I. Rebai, Y. BenAyed, and W. Mahdi, "Deep kernel-SVM network," *2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC, Canada, 2016, pp. 1955-1960.

[5] A. A. Aburomman, and M. B. I. Reaz, "Ensemble of binary SVM classifiers based on PCA and LDA feature extraction for intrusion detection," *Proc. 2016 IEEE Adv. Inf. Manag. Commun. Electron. Autom. Control Conf. IMCEC 2016*, 2016, pp. 636-640, doi: 10.1109/IMCEC.2016.7867287.

[6] D. Liang, Q. Liu, B. Zhao, Z. Zhu, and D. Liu, "A clustering-SVM ensemble method for intrusion detection system," *2019 8th Int. Symp. Next Gener. Electron. ISNE 2019*, vol. 2, no. 2, pp. 1-3, 2019, doi: 10.1109/ISNE.2019.8896514.

[7] M. Safaldin, M. Otair, and L. Abualigah, "Improved binary gray wolf optimizer and SVM for intrusion detection system in wireless sensor networks," *J. Ambient Intell. Humaniz. Comput.*, 2020, Art. no. 0123456789, doi: 10.1007/s12652-020-02228-z.

[8] H. Wang, J. Gu, and S. Wang, "An effective intrusion detection framework based on SVM with feature augmentation," *Knowledge-Based Syst.*, vol. 136, pp. 130-139, 2017, doi: 10.1016/j.knosys.2017.09.014.

[9] M. Al-Qatf, Y. Lasheng, M. Al-Habib, and K. Al-Sabahi, "Deep learning approach combining sparse autoencoder with SVM for network intrusion detection," *IEEE Access*, vol. 6, no. c, pp. 52843-52856, 2018, doi: 10.1109/ACCESS.2018.2869577.

[10] J. Gu, L. Wang, H. Wang, and S. Wang, "A novel approach to intrusion detection using SVM ensemble with feature augmentation," *Comput. Secur.*, vol. 86, pp. 53-62, 2019, doi: 10.1016/j.cose.2019.05.022.

[11] A. I. Saleh, F. M. Talaat, and L. M. Labib, "A hybrid intrusion detection system (HIDS) based on prioritized k-nearest neighbors and optimized SVM classifiers," *Artif. Intell. Rev.*, vol. 51, no. 3, pp. 403-443, 2019, doi: 10.1007/s10462-017-9567-1.

[12] P. Tao, Z. Sun, and Z. Sun, "An improved intrusion detection algorithm based on GA and SVM," *IEEE Access*, vol. 6, pp. 13624-13631, 2018, doi: 10.1109/ACCESS.2018.2810198.

[13] E. Kabir, J. Hu, H. Wang, and G. Zhuo, "A novel statistical technique for intrusion detection systems," *Futur. Gener. Comput. Syst.*, vol. 79, pp. 303-318, 2018, doi: 10.1016/j.future.2017.01.029.

[14] G. Kavitha, and N. M. Elango, "An approach to feature selection in intrusion detection systems using machine learning algorithms," *Int. J. e-Collaboration*, vol. 16, no. 4, pp. 48-58, 2020, doi: 10.4018/IJeC.2020100104.

[15] A. Al Shorman, H. Faris, and I. Aljarah, "Unsupervised intelligent system based on one class support vector machine and grey wolf optimization for IoT botnet detection," *J. Ambient Intell. Humaniz. Comput.*, vol. 11, no. 7, pp. 2809-2825, 2020, doi: 10.1007/s12652-019-01387-y.

[16] E. M. Roopa Devi, and R. C. Suganthe, "Enhanced transductive support vector machine classification with grey wolf optimizer cuckoo search optimization for intrusion detection system," *Concurr. Comput.*, vol. 32, no. 4, pp. 1-11, 2020, doi: 10.1002/cpe.4999.

[17] S. N. Mighan, and M. Kahani, "A novel scalable intrusion detection system based on deep learning," *Int. J. Inf. Secur.*, 2020, Art. no. 0123456789, doi: 10.1007/s10207-020-00508-5.

[18] D. V. Kumar, and V. Ramasamy, "Improved intrusion detection classifier using cuckoo search optimization with support vector machine," *ICTACT J. Soft Comput.*, vol. 10, no. 2, pp. 2029-2034, 2020, doi: 10.21917/ijsc.2020.0287.

[19] Z. Ye, Y. Sun, S. Sun, S. Zhan, H. Yu, and Q. Yao, "Research on network intrusion detection based on support vector machine optimized with grasshopper optimization algorithm," *Proc. 2019 10th IEEE Int. Conf. Intell. Data Acquis. Adv. Comput. Syst. Technol. Appl. IDAACS 2019*, vol. 1, pp. 378-383, 2019, doi: 10.1109/IDAACS.2019.8924234.

[20] J. Jackson, "Data mining; A conceptual overview," *Commun. Assoc. Inf. Syst.*, vol. 8, 2002, doi: 10.17705/1cais.00819.

[21] A. A. Aburomman, and M. B. I. Reaz, "A novel SVM-kNN-PSO ensemble method for intrusion detection system," *Appl. Soft Comput. J.*, vol. 38, pp. 360-372, 2016, doi: 10.1016/j.asoc.2015.10.011.

[22] X. Gao, C. Shan, C. Hu, Z. Niu, and Z. Liu, "An adaptive ensemble machine learning model for intrusion detection," *IEEE Access*, vol. 7, pp. 82512-82521, 2019, doi: 10.1109/ACCESS.2019.2923640.

[23] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," *IEEE Access*, vol. 7, pp. 41525-41550, 2019, doi: 10.1109/ACCESS.2019.2895334.

[24] K. A. P. da Costa, J. P. Papa, C. O. Lisboa, R. Munoz, and V. H. C. de Albuquerque, "Internet of things: A survey on machine learning-based intrusion detection approaches," *Comput. Networks*, vol. 151, pp. 147-157, 2019, doi: 10.1016/j.comnet.2019.01.023.

[25] J. Ren, J. Guo, W. Qian, H. Yuan, X. Hao, and H. Jingjing, "Building an effective intrusion detection system by using hybrid data optimization based on machine learning algorithms," *Secur. Commun. Networks*, vol. 2019, 2019, doi: 10.1155/2019/7130868.

[26] K. A. Taher, B. M. Y. Jisan, and M. M. Rahman, "Network intrusion detection using supervised machine learning technique with feature selection," *1st Int. Conf. Robot. Electr. Signal Process. Tech. ICREST 2019*, 2019, pp. 643-646, doi: 10.1109/ICREST.2019.8644161.

[27] S. Tavara, "Parallel computing of support vector machines: A survey," *ACM Comput. Surv.*, vol. 51, no. 6, 2019, doi: 10.1145/3280989.

[28] J. Gu, and S. Lu, "An effective intrusion detection approach using SVM with Naive Bayes feature embedding," *Comput. Secur.*, vol. 103, 2021, Art. no. 102158, doi: 10.1016/j.cose.2020.102158.

[29] M. Hosseinzadeh, A. M. Rahmani, B. Vo, M. Bidaki, M. Masdari, and M. Zangakani, "Improving security using SVM-based anomaly detection: Issues and challenges," *Soft Comput.*, vol. 25, pp. 3195-3223, 2021, doi: 10.1007/s00500-020-05373-x.

[30] F. Wang, Z. Chen, and G. Song, "Monitoring of multi-bolt connection looseness using entropy-based active sensing and genetic algorithm-based least square support vector machine," *Mech. Syst. Signal Process.*, vol. 136, 2020, Art. no. 106507, doi: 10.1016/j.ymssp.2019.106507.

[31] J. Nayak, B. Naik, and H. S. Behera, "A comprehensive survey on support vector machine in data mining tasks: Applications & challenges," *Int. J. Database Theory Appl.*, vol. 8, no. 1, pp. 169-186, 2015, doi: 10.14257/ijdta.2015.8.1.18.

[32] H. Byun, and S. W. Lee, "Applications of support vector machines for pattern recognition: A survey," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 2388, pp. 213-236, 2002, doi: 10.1007/3-540-45665-1_17.

[33] P. S. Thapanan Janyalikit, "Intelligent information and database systems: 8th Asian conference, ACIIDS 2016 da Nang, Vietnam, March 14-16, 2016 proceedings, Part I," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 9621, pp. 616-625, 2016, doi: 10.1007/978-3-662-49381-6.

International Journal of Knowledge Based Computer Systems
10 (1) June 2022, 40-44
http://www.publishingindia.com/ijkbcs/

Mango Leaf Diseases Detection using Deep Learning

Amisha Sharma[1*], Rajneet Kaur Bijral[2], Jatinder Manhas[3] and Vinod Sharma[4]
[1]M.Tech. Student, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: amishasharma26111997@gmail.com
[2]PhD Research Scholar, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: rajbijral@gmail.com
[3]Sr. Assistant Professor, Department of Computer Science and IT, Bhaderwah Campus, University of Jammu, Jammu and Kashmir, India. Email: manhas.jatinder@gmail.com
[4]Professor, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India.
Email: vnodshrma@gmail.com
*Corresponding Author

Abstract: Diseases and pests cause great economic loss to the mango industry every year. The detection of various mango diseases is challenging for the farmers as the symptoms produced by different diseases may be very similar, and may be present simultaneously. This research paper is an attempt to provide the timely and accurate detection and identification of mango leaf diseases. Convolutional Neural Networks are end-to-end learning algorithms which perform automatic feature extraction and learn complex features directly from raw images, making them suitable for a wide variety of tasks like image classification, object detection, segmentation etc. In the proposed study, we develop a Convolutional Neural Networks based model for detection and classification of mango leaf diseases at the initial stages. Data augmentation is performed on a collected dataset. We applied data augmentation techniques like rotation, translation, reflection and scaling. Convolutional Neural Networks model has been trained on the augmented data for detection and classification of mango leaf diseases. The proposed CNN based model attains 90.36% of accuracy. The results validate that the proposed method is effective in detecting various types of mango leaf diseases and can be used as a practical tool by farmers and agriculture scientists.

Keywords: Convolution Neural Network (CNN), Crop, Deep learning, Image classification, Mango.

I. Introduction

About 40% of mangoes production is produced by India and stands first in the various mango growing countries of the world. Amongst the fruit crops in India vital place gets occupied by mangoes and plays an essential role in the economy of the country. About 30% to 40% of the crop yield got infected by various diseases and due to unaided eye perception the mango leaf disease went undetected**.** The different diseases affecting mango leaves cannot get acknowledged by the farmers which cause less production of mango fruit. Different diseases [1] cause different effects on mango crops. Some cause white patches and some cause black and all these patches seem over the surface of the leaf or early grown fruits as well while some other diseases cause white fungal powder on leaves and some affect the young leaves and shoots also. All these different types of illnesses need to be discerned in the initial stages and should be managed before it grows more and causes a severe loss to the plant life. To get this detection done in the initial stage, farmers and agricultural scientists need to keep an eye continuously on the plant parts which is a sluggish process. For the advance detection of disease in the plants some technique is needed as the prior acclaim of disease is the first step in the detection and expansion of mango diseases. Conventional ways to identify diseases is time consuming and expensive as it needs the expertise, knowledge and continuous monitoring. Still it lacks correct recognition of disease because of the complex structure and pattern of the leaf. With the advent of computational methods in the field of image recognition [2] and classification [3] this problem can be solved with greater accuracy. By using technology one can detect diseases on a large scale. In the case of mango leaves; there are various types of diseases [4] present like powdery mildew, anthracnose, red rust etc. In the present work, a deep learning (DL) based model has been proposed for the classification of various mango leaf diseases (powdery mildew, anthracnose, red rust) at the initial stages. Accuracy, Recall, Precision and F-Score have been used to evaluate the model.

II. Literature Review

Early Disease Detection and discussion are important for better yield and quality of crops. Disease Plants can lead to the huge Economic Losses to the Individual farmers. From the extensive literature review it has been found that various Machine Learning (ML) and Deep Learning (DL) based techniques have been used for recognition and classification of mango leaf disease to prevent the loss of harvest [5] a novel segmentation approach is proposed in this study to segment the diseased part by considering the vein pattern of the leaf. Afterward, features were extracted and fused using canonical correlation analysis (CCA)-based fusion. As a final identification step, a cubic support vector machine (SVM) is implemented to validate the results. The highest accuracy achieved by this proposed model is 95.5%. In 2013 a deep Convolutional Neural Network [6] was presented to identify 14 crop species and 26 diseases (or absence thereof) by using a public dataset of 54,306 images of diseased and healthy plant leaves collected under controlled conditions. The trained model achieves an accuracy of 99.35%. [7] Proposed the DL based approach for image recognition and examined the three main architecture of the Neural Network: Faster Region-based Convolutional Neural Network (Faster R-CNN), Region-based Fully CNN (R-CNN) and Single shot Multibook Detector (SSD). System Proposed can detect the different types of disease efficiently and have the ability to deal with complex scenarios. The accuracy of 94.6% was attained. [8] Built a model for plant diseases and pests detection, and put forward a comparison with traditional plant diseases and pests detection methods. [9] Proposed a work that includes finding a solution to the problem of 38 different classes of plant diseases detection using the simplest approach while making use of minimal computing resources to achieve better results compared to the traditional models. VGG16 training model is deployed for detection and classification of plant diseases. Neural network models employ automatic feature extraction to aid in the classification of the input image into respective disease classes. This proposed system has achieved an average accuracy of 94.8% indicating the feasibility of the neural network approach even under unfavorable conditions. [10] Used a pre-trained CNN architecture called AlexNet that is modeled for automatic feature extraction and classification. The system is developed with MATLAB and achieves an accuracy rate of the detection of 99% and 89% for Grape leaves and Mango leaves respectively.

III. Methodology

A deep learning based model is designed for the proposed framework and for this study; three commonly found diseases are considered viz. Anthracnose, Red rust and Powdery mildew. Data is collected from the Sher-e-Kashmir University of Agriculture Sciences and Technology, Jammu (SKUAST-J) where they grow different varieties of fruits for educational and research purposes. There are around 980 images of 4 classes i.e. Normal, Anthracnose, Red rust and Powdery mildew. Fig. 1 shows some images of the collected dataset. The brief flowchart of the proposed model is shown in Fig. 2.

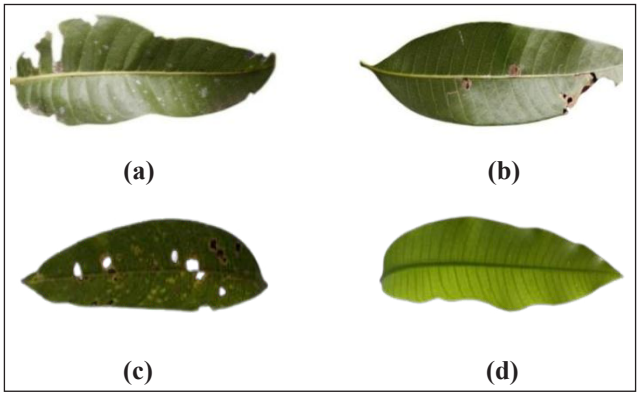


Fig. 1: Showing the Samples of Images of Mango Leaves (a) Anthracnose (b) Powdery Mildew (c) Red Rust and (d) Normal in Collected Dataset

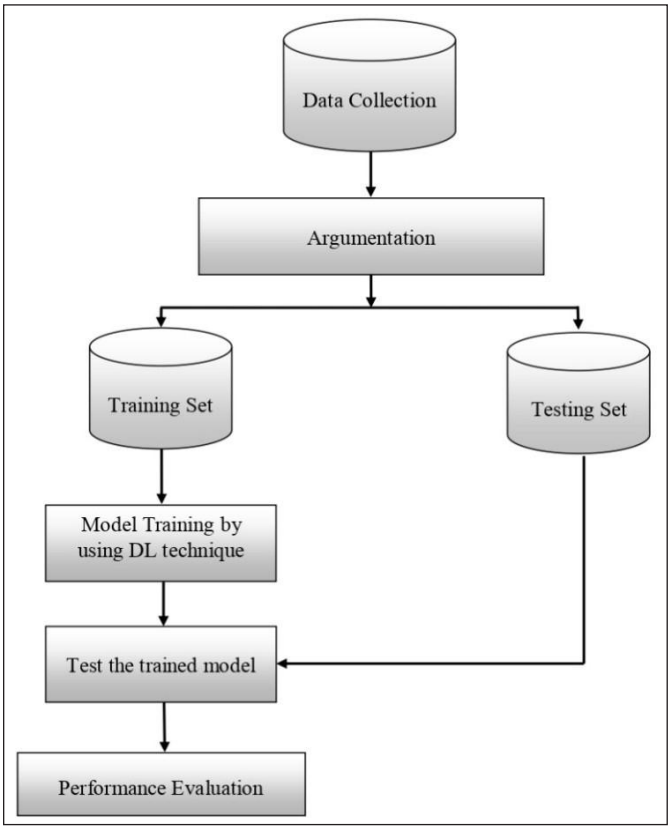


Fig. 2: Research Methodology for the Mango Leaf Diseases Detection using CNNs

Deep Learning

DL is the subset of machine learning (ML) [22] that emulates the functioning of the human brain in coursing data and

producing patterns for use in decision making. It is a broader and more advanced part of ML methods. It is also called deep neural learning or deep neural networks. DL technique uses various layers consisting of nonlinear units. Every layer utilizes the output of the previous layer and considers it as its input [11].

Convolutional Neural Networks

Convolutional Neural Network (CNNs) [23] is a specific type of artificial neural networks (ANNs) that uses perceptrons, a ML unit algorithm, for supervised learning, to analyze data. CNNs are capable of performing complex jobs with images, sounds, texts, videos, etc, and most commonly used for analyzing visual imagery. CNNs use a wide variety of multilayer perceptrons designed that require minimal preprocessing. They are also known as shift invariant, based on their shared-weights architecture and translation invariant characteristics. CNNs are best for predictions and most widely used than other algorithms. It consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of CNNs typically consist of a series of Convolutional layers that convolve with a dot product. The activation function is commonly a RELU layer (rectified linear unit), and is subsequently followed by additional convolutions such as pooling layers, fully connected layers and normalization layers, referred to as hidden layers because their inputs and outputs are masked by the activation function and final convolution layer [12]. The basic architecture of the CNNs is shown below in Fig. 3.

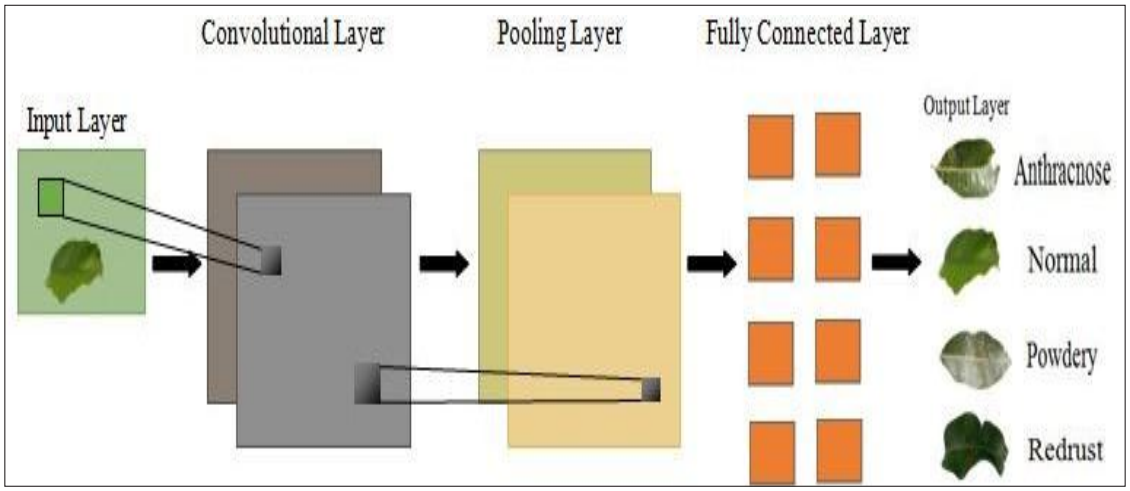


Fig. 3: Basic Architecture of Convolutional Neural Networks (CNNs)

Convolution Layer: Convolutional layer performs action on the input layer and convolves it and submits its result to the following layer. Here the series of images i.e. different mango leaves images acts as Input. This process is similar to the response of a neuron present in a human body. Sometimes a "kernel" is passed over the image, viewing a few pixels at a time (for example, 3X3 or 5X5). Convolution operation is there as well which is a dot product of the original pixel values with weights defined in the filter. The results obtained are then summarized into one number that represents all the pixels the filter observed [13].

Activation Layer: Matrix generated by the convolution layer is much smaller in size than the original image. This matrix passes through an activation layer that introduces non-linearity to allow the network to train itself via back propagation. The activation function generally used is ReLU [14].

Pooling Layer: Pooling layer is another building block of CNN. In this layer there is further down sampling and reduction of the size of the matrix. Over the results of the previous layer, a filter is passed and selects one number out of each group of values (typically the maximum value, this is called max pooling). This allows the network to train much faster and efficiently by focusing on the major information in each feature of the image [15].

Fully Connected Layer: A one-dimensional vector which is representing the output of the previous layers acts as input for this layer. Its output is a list of probabilities for different possible labels attached to the image (e.g. Anthracnose, Red rust, Powdery mildew and Probability [16]. The label that receives the highest probability is the classification or detection decision.

Output Layer: Output layer in CNN is that layer in which the input from the other layers is flattened and sent so as to transform the output into the number of classes as desired by the problem undertaken [17].

There may be multiple activation and pooling layers, depending on the CNN architecture.

Performance Evaluation

The performance of the model will be analyzed on the basis various parameters derived from its confusion matrix represented as under in Table I.

TABLE I

Actual Result			
Predicted Result		LOW	HIGH
	LOW	True Positive	False Positive
	HIGH	False Negative	True Negative

True Positive (TP): These are the correctly predicted positive values which mean that the value of the actual class is yes and the value of predicted class is also yes.

True Negative (TN): These are the correctly predicted negative values which means that the value of actual class is no and the value of predicted class is also no.

False Positive (FP): When the actual class is no but the predicted class is yes.

False Negative (FN): When the actual class is yes but the predicted class is no.

Accuracy - Accuracy is the most intuitive performance measure and it is simply a ratio of correctly predicted observation to the total observations. [18]

Precision - Precision is the ratio of correctly predicted positive observations to the total predicted positive observations [19].

Recall - Recall is the ratio of correctly predicted positive observations to the all observations in actual class – yes [20].

F1 Score - F1 Score is the weighted average of Precision and Recall. Therefore, this score takes both false positives and false negatives in observations [21].

These performance measures can be calculated as shown in the Table II below.

TABLE II

Parameter	Formula
Accuracy	(TP + TN) / (Total Cases)
Recall	TP / (TP + FN)
Precision	TP / (TP + FP)
F1 Score	(2 * Recall * Precision) /(Recall + Precision)

IV. Results and Discussion

In the proposed study, a dataset of 980 images is collected from SKUAST-J. Dataset consists of 4 classes i.e. Normal, Anthracnose, Red rust and Powdery mildew. Argumentation is performed on the collected dataset to increase the number of images for model training and testing. Then the splitting of the dataset has been performed to train and test the dataset. 80% of the data has been used to train the model and the rest 20% of images have been used to test the model. A CNN based model has been developed to classify the images. The model has been trained with epochs equal to 40 and the batch size of 35. The test accuracy of 90.36% is attained. The training and Validation accuracy graph attained by the proposed model is shown below in Fig. 4. The Training and Validation loss graph is shown in Fig. 5. A generalized model is built in the proposed study as the validation loss is very less and the Validation accuracy is more than the training accuracy as shown in Fig. 5. The confusion matrix thus generated by the model shown in Table III. The various parameters are used to measure the performance of the model like precision, recall and F1 score and the values of these for the proposed model are 0.9, 0.90 and 0.9 respectively.

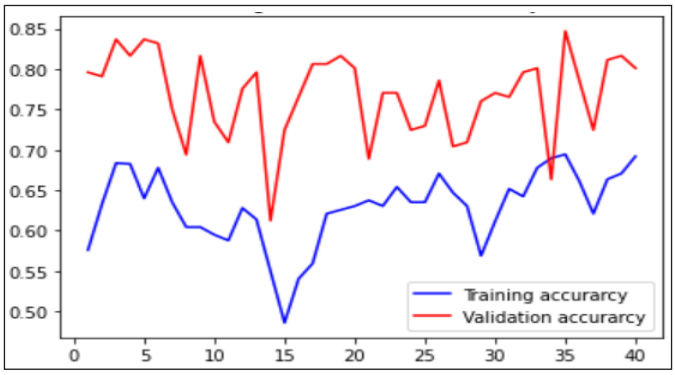


Fig. 4: Shows the Training and Validation Accuracy of the Proposed Model

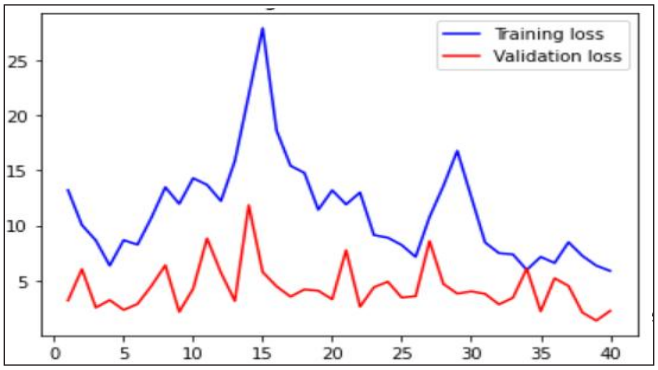


Fig 5. Shows the Training and Validation Loss for the Proposed Model

TABLE III: SHOWS THE CONFUSION MATRIX OF THE PROPOSED CNNS BASED FRAMEWORK

Predicted class / True class	Anthracnose	Normal	Powdery	Red rust
Red rust	3	3	2	50
Powdery	7	0	40	0
Normal	0	47	0	0
Anthracnose	40	0	4	0

V. Conclusion

Mango is one of the most cultivated fruit crops in India. So it is important to protect it and detect the various diseases in the initial stages. For that, a model based on deep learning approach called CNN is proposed for the identification of 3 different plant leaf diseases, detection and recognition systems. This approach utilized a minimum set of layers to identify the diseases of four classes. The CNN is trained with SKUAST-J dataset. The Model works with accuracy of 90.36%. In order to increase accuracy and to make the proposed model more efficient in future, the number of the dataset can be increased, the concept of transfer learning can also be used. Moreover, the model can be made more efficient to detect images in the natural background as well.

References

[1] S. Gulvani, and R. Patil, "Deep learning for image based mango leaf disease detection," *International Journal of Recent Technology and Engineering (IJRTE*), vol. 8. no. 3S3, pp. 54-56. 2019.

[2] M. M. Krishna, M. Neelima, M. Harshali, and M. V. G. Rao, "Image classification using deep learning," *International Journal of Engineering & Technology*, vol. 7, no. 2.7, pp. 614-617, 2018.

[3] G. Geetha, S. Samundeswari, G. Saranya, K. Meenakshi, and M. Nithya, "Plant leaf disease classification and detection system using machine learning," *Journal of Physics: Conference Series,* vol. 1712, 2020.

[4] O. Parkash, and A. K. Mishra, "Important diseases of mango and their effect on production," *Biological Memoirs*, vol. 18, no. 1&2, pp. 39-55, 1992.

[5] R. Saleem, J. H. Shah, M. Sharif, M. Yasmin, H.-S. Yong, and J. Cha, "Mango leaf disease recognition and classification using novel segmentation and vein pattern technique," *Appl. Sci.*, vol. 11, no. 24, 2021, Art. no. 11901.

[6] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Frontiers in Plant Science*, 2016.

[7] S. Kumar, V. Chaudhary, and S. Chandra, "Plant disease detection using CNN," *Turkish Journal of Computer and Mathematics Education,* vol. 12, no. 12, 2021.

[8] J. Liu, and X. Wang, "Plant diseases and pests detection based on deep learning," *Plant Methods*, vol. 17, 2021, Art. no. 22.

[9] R. Rinu, and S. H. Manjula, "Plant disease detection and classification using CNN," *International Journal of Recent Technology and Engineering (IJRTE),* vol. 10, no. 3, pp. 152-156, 2021.

[10] U. S. Rao, R. Swathi, V. Sanjanaa, L. Arpitha, K. Chandrasekhar, Chinmayi, and P. K. Naikb, "Deep learning precision farming: Grapes and mango leaf disease detection by transfer learning," *Global Transitions Proceedings*, vol. 2, no. 2, pp. 535-544, 2021.

[11] H. Muresan, and M. Oltean, "Fruit recognition from images using deep learning," *Acta Univ. Sapientiae Inform.*, vol. 10, pp. 26-42, 2018.

[12] S. Indolia, A. Goswami, S. P. Mishra, and P. Asopa, "Conceptual understanding of convolutional neural network," *Procedia Computer Science*, vol. 132, pp. 679-688, 2018.

[13] A. Khan, A. Sohail, U. Zahoora, and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural network," 2017.

[14] R. Yamashita, M. Nishio, R. K. G. Do, and K. Togashi, "Convolutional neural network: An overview and application in radiology," *Insights into Imaging*, vol. 9, pp. 611-629, 2018.

[15] R. Nirthika, S. Manivannan, A. Ramanan, and R. Whang, "Pooling in convolutional neural network for medical image analysis: A survey and an empirical study," *Neural Comput Appl.*, vol. 34, no. 7, pp. 5321-5347, 2022.

[16] L. Alzubaidi, J. Zhang, A. J. Humaidi, A. Al-Dujaili, Y. Duan, ... and L. Farhan, "Review of deep learning: Concepts, CNN, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, 2021.

[17] E. Irmark, "Implementation of convolutional network approach for COVID-19 disease detection," *Physiological Genomics*, vol. 52, no. 12, pp. 590-601, 2020.

[18] S. Sorooshain, N. F. Aziz, A. Ahmad, S. N. Jubidin, and N. M. Mustapha, "Review on performance measurement systems," *Mediterranean Journal of Social Sciences*, vol. 7, no. 1, 2016.

[19] D. M. W. Powers, "Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation," *International Journal of Machine Learning Technology*, vol. 2, no. 1, pp. 37-63, 2011.

[20] C. Goutte, and E. Gaussier, "A probabilistic interpretation of precision, recall and F-score, with implication for evaluation," 2014.

[21] M. Sokolova, N. Japkowicz, and S. Szpakowicz, "Beyond accuracy, F-score and ROC: A family of discriminant measures for performance evaluation," 2014.

[22] S. Rajesh, P. Amudhavalli, and Dhanapriya A. P., "Prediction of air pollution using supervised machine learning," *Journal of Applied Information Science*, vol. 10, no. 1, pp. 10-16, 2022.

[23] Chinmaya H. S., J. M. Balaji, G. N. Sharma, and N. Divakar, "Dragonfly-net: Dragonfly classification using convolution neural network," *Journal of Applied Information Science*, vol. 10, no. 1, pp. 60-66, 2022.

International Journal of Knowledge Based Computer Systems
13 (1) 2025, 01-09
www.publishingindia.com//ijkbcs/

Potato Leaf Disease Detection using Machine Learning

Palak Jasrotia

M.Tech. Student, Department of Computer Science and IT, University of Jammu, Jammu and Kashmir, India. Email: palak9799@gmail.com

Abstract: The Paper "Potato Leaf Disease Detection using Machine Learning" focuses on the accurate identification of diseases in potato leaves to support early intervention and enhance crop yield. The process begins with Data Collection, where a dataset of 2,170 potato leaf images is sourced from Kaggle. This dataset is categorized into three classes: Early Blight, Late Blight, and Healthy. Next, Data Pre-Processing is undertaken to ensure the images are cleaned, resized, and normalized, preparing them for effective analysis. During Feature Extraction, relevant features from the images are identified to represent the data meaningfully. The dataset is then split into two subsets: the Train Image Set with 1,736 images, and the Test Image Set with 434 images, to facilitate model training and evaluation respectively. Various Machine Learning Techniques are applied, including k-Nearest Neighbors (KNN), Support Vector Machine (SVM), Naive Bayes, Decision Tree, and Random Forest. Each trained model undergoes rigorous testing to evaluate its performance. To assess the effectiveness of each model, Performance Metrics such as accuracy, precision, recall, and F1-score are computed. The model exhibiting the highest accuracy is subjected to Feature Selection using Neighborhood Component Analysis (NCA) to enhance its performance further. Ultimately, a Hybrid Model is developed by combining the strengths of the individual models, aiming to improve overall accuracy and robustness in disease detection. This comprehensive approach integrates multiple machine learning techniques and feature selection methods, offering a robust solution for potato leaf disease detection. The project demonstrates the potential of machine learning in agricultural applications, contributing to more efficient and precise disease management.

Keywords: Decision tree, Feature extraction, Feature selection, K-nearest neighbor, Machine learning, Naive bayes, Potato leaf disease, Random forest, Support vector machine.

I. Introduction

Agriculture productivity plays a very significant role in the Indian economy. Agriculture is the major occupation in India. The economy of our country highly depends on agriculture and its associated products. India stood first in the world with the highest net cropped area followed by the US and China. Various pests and diseases affect the plant growth, quantity and quality of the product. So, it is very necessary to detect the disease at an early stage of the growth of the plant. India is a cultivation nation, more than 60% of the population of India has agriculture and its productivity as their main occupation. In India, agriculture is dependent on the monsoon. When the monsoon is good, agriculture productivity is good; when the monsoon is less, productivity is less or not in good condition [1].

Potato is one of the most cultivated crops. Worldwide potatoes have their cultivation priority as a staple food. For successful potato production, a strong food security system can be developed as it is a great source of vitamins and minerals. However, several diseases affect potato production and degrade agricultural development. Therefore, disease detection in an early stage can provide a better solution for successful crop cultivation [2].

II. Literature Review

Bhuyar (2014) [3] proposed an approach where different classification algorithms such as J48, Naïve Bayes, and Random forest algorithm were applied to soil dataset to predict its fertility. J48 algorithm gave better result with an accuracy of 98.17% than other algorithms.

Singh *et al.* (2015) [4] performed a review of the literature on various image processing techniques for detecting leaf disease. The writers wanted to speed up identification and detection of plant diseases while lowering the subjectivity that comes with naked-eye observation. They presented an algorithm that uses picture segmentation to automatically identify and categorize plant leaf diseases. The impact of HSI, CIELAB, and YCbCr color spaces on disease spot detection was examined by the authors. To identify the disease spot, the Otsu technique was applied to the color component after an image was smoothed using a median filter. The suggested method was not put to the test on any datasets.

He *et al.* (2015) [5] deep residual networks were discussed in relation to image recognition challenges. As shallow representations for image retrieval and classification, VLAD and Fisher Vector are cited in the paper's literature review as related concepts. The writers also covered the advantages of encoding residual vectors over original vectors for vector quantization. The Multigrid method was also mentioned as a method for solving partial differential equations (PDEs) by reformulating systems as sub-problems at multiple scales, where each subproblem is accountable for the residual solution between a coarser and a finer scale. This method is used in low-level vision and computer graphics.

Rajeshwari and Arunesh (2016) [6] performed a comparative analysis of ML algorithms i.e., Naive Bayes, JRIP and J48 for prediction of soil types. The experiments were performed on soil data consisting of 110 samples using data analytics tool R. The experimental results predicted that JRIP algorithm performed better as it gave highest accuracy of 98.18% with kappa statistic of approximate 1.0.

Sujata (2016) [7] proposed a model to estimate the crop yield in order to improve the value and gain of farming area using data mining techniques.

Islam *et al.* (2017) [8] suggested a method to identify diseases from images of potato plant leaves by combining image processing and machine learning. The authors classified diseases using the 'Plant Village' database of openly accessible plant images. Using the suggested method, the research classified 300 images of diseases with a 95% accuracy. The paper also highlights how crucial contemporary phenotyping and plant disease detection are to assuring food security and sustainable agriculture.

Tiwari *et al.* (2020) [9] proposed a methodology in which VGG19 was used for feature extraction, and multiple classifiers, including logistic regression, SVM, KNN, and neural networks, were applied for classification. Logistic regression, in combination with VGG19, achieved a classification accuracy of 97.8%. The proposed approach outperformed previous models, achieving a higher classification accuracy (97.8%). Various approaches, including image segmentation and traditional image processing techniques, have been used for disease detection.

Tarik *et al.* (2021) [10] proposed a methodology that involves the use of Convolutional Neural Networks (CNNs) for detecting various types of potato diseases. The authors collected a dataset consisting of 2034 images of healthy and diseased potato leaves. The images were preprocessed using OpenCV, and a CNN model architecture was employed for disease classification. The CNN model achieved a high accuracy rate of 84% for disease detection. The training, testing, and validation processes were discussed, indicating successful results. The authors mention the use of transfer learning to enhance the model's performance, adjusting pre-trained weights to match the desired output dimensions.

Afzaal *et al.* (2021) [11] explored the use of artificial intelligence, specifically convolutional neural networks (CNNs) and deep learning, for the early detection of early blight disease in potato plants. They collected a dataset of images of healthy and diseased potato plants from four fields, using different lighting

conditions. The CNNs used in the study were Google Net, VGG Net, and Efficient Net, trained using the PyTorch framework. The images were classified into three classes (2-class, 4-class, and 6-class) to accurately identify the disease at different growth stages. Results showed that Efficient Net and VGG Net outperformed Google Net in terms of disease identification accuracy.

Rashid *et al.* (2021) [12] proposed PDDCNN method, applied to the Potato Leaf Disease (PLD) dataset, achieved an impressive accuracy of 99.75%, demonstrating high precision, recall, F1-score, and an excellent ROC curve. With minimal parameters, it surpasses existing methods in computational efficiency. Future extensions include multi-disease detection on a single leaf, disease localization, severity estimation, dataset enhancement, an IoT-based real-time monitoring system, and a website/ mobile application for broader accessibility. Cross-dataset evaluations on Plant Village and PLD datasets, with and without augmentation, further validated the superiority of the PDDCNN method.

Singh and Kaur (2021) [13] proposed a methodology that includes a description of the necessity of image segmentation in analyzing colored images. Implementation of the K-means algorithm for segmentation. Steps involved in the K-means algorithm. Mentioning of distance metrics like Minkowskis Manhattan, and Euclidean distances. Explanation of the importance of feature extraction in reducing dimensionality. Adaptation of the binary classification for multiclass classification in the context of plant disease detection. Use of the proposed framework on the 'Plant Village' dataset. Achievement of an overall accuracy of 95.99%. dataset for three popular crops (tomato, corn and potato) and the accuracy was 90.01%.

III. Research Methodology

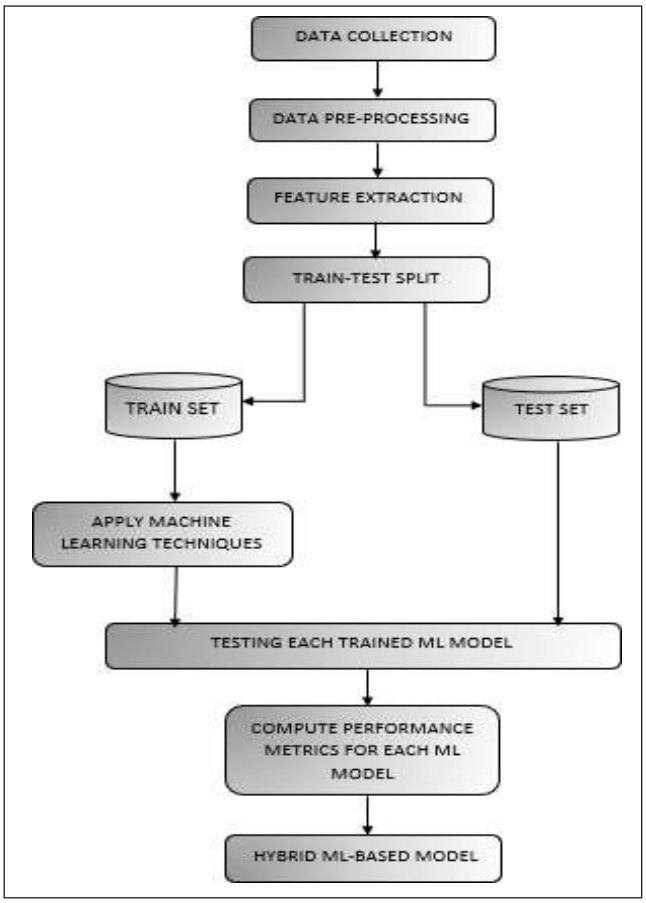


Fig. 1: Methodology for the Current Study

A. Data Collection

The dataset contains images of potato leaves in .jpg format and has been sourced from Kaggle (https://www.kaggle.com/datasets/emmarex/plantdisease) [14].

TABLE I: NUMBER OF IMAGES IN DATASETS

Class	No. of Images
Early Blight	750
Late Blight	693
Healthy	727
Total	2170

It comprises a total of 2,170 images, which are categorized into three distinct classes. The "Early Blight" class contains 750 images, representing the highest number of images among the classes. The "Late Blight" class has 693 images, making it the class with the fewest images. The "Healthy" class includes 727 images. Together, these classes form a comprehensive dataset for analyzing and distinguishing between early blight, late blight, and healthy conditions in plants. Samples of Potato Leaf dataset are shown in Fig. 2.

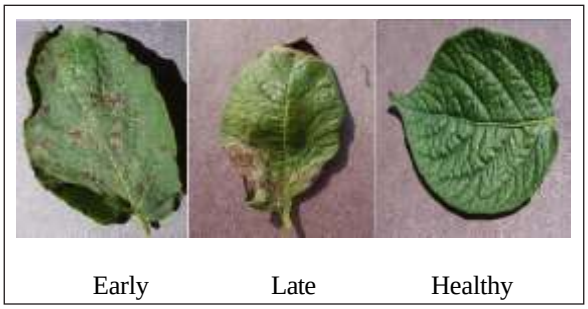


Fig. 2: Sample of Collected Potato Leaf Images

B. Data Pre-Processing

After data collection, preprocessing operations and filtering were performed on the collected images of the dataset. The images were filtered using median filtering to smooth the edges and remove noise. Median filters are particularly useful in reducing random noise, especially when the noise amplitude probability density has large tails and periodic patterns [15]. The dataset was then divided into an 80:20 ratio for training and testing purposes. The Train Image Set comprised 1,736 images, while the Test Image Set included 434 images.

RGB image is converted into grayscale image using grayscale () function.

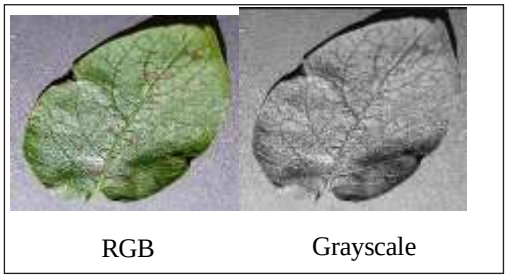


Fig. 3: RGB to Grayscale

TABLE II: CLASS LABELS OF POTATO LEAVES

Class Labels	Potato Leaves
Class 0	Early Blight
Class 1	Late Blight
Class 2	Healthy

C. Implementation of the Framework

The framework begins with the conversion of RGB images into grayscale images using the grayscale () function. This step simplifies the image data and prepares it for feature extraction. In the feature extraction phase, features of the Potato Leaves dataset are extracted using the Gray-Level Co-Occurrence Matrix (GLCM). The GLCM is a statistical method used to examine texture by considering the spatial relationship of pixels within an image. Specifically, it calculates how often pairs of pixels with specific values occur in a specified spatial relationship, thereby creating a GLCM [16].

To characterize the texture of the images, the greycomatrix () function is used to generate the GLCMs, and the greycoprops () function extracts specific features from these matrices. The features calculated from the GLCM include Angular Second Moment (ASM), Contrast, Dissimilarity, Homogeneity, Energy, and Correlation. These features are extracted at four different angles: 0°, 45°, 90°, and 150° [17].

In addition to the GLCM features, the framework computes additional statistical features such as mean, standard deviation, variance, skewness, and kurtosis. In total, 29 features are extracted by applying GLCM feature extraction at the specified angles. This comprehensive set of features provides a detailed representation of the texture and statistical properties of the Potato Leaves dataset, which is crucial for subsequent analysis and modeling steps.

D. Technique Used for Feature Extraction

In this section, Features of the Potato Leaf datasets have been extracted using the Gray-Level Co-Occurrence Matrix (GLCM).

E. Implementation of Gray-Level Co-Occurrence Matrix (GLCM)

The framework for feature extraction begins with converting RGB images to grayscale using the grayscale () function. This step simplifies the data and prepares it for the next phase. The feature extraction process involves the use of the Gray-Level Co-Occurrence Matrix (GLCM), a statistical method that examines texture by considering the spatial relationship of pixels. The GLCM, also known as the gray-level spatial dependence matrix, characterizes the texture of an image by calculating how often pairs of pixels with specific values occur in a specified spatial relationship within the image. The GLCMs are created using the greycomatrix () function. Subsequently, the greycoprops () function extracts specific features from the GLCM. These features include Angular Second Moment (ASM), Contrast, Dissimilarity, Homogeneity, Energy, and Correlation. To capture comprehensive texture information, the features are extracted at four different angles: 0°, 45°, 90°, and 150°. Additionally, statistical features such as mean, standard deviation, variance, skewness, and kurtosis are computed. In total, 29 features are extracted by applying GLCM feature extraction at the specified angles, providing a detailed representation of the texture and statistical properties of the Potato Leaves dataset.

F. Texture Features

At first level decomposition of GLCM, texture features likeASM, contrast, homogeneity, correlation, dissimilarity and energy have been calculated for each coefficient. A total of 29 features Classification of Potato Leaves using Machine Learning Techniques for each image were calculated.

Angular Second Moment (ASM), also known as Energy, is a feature derived from the Gray-Level Co-Occurrence Matrix (GLCM) used to characterize the texture of an image. It measures the uniformity or smoothness of the texture.

$$ASM = \sum i \sum j P(i, j)2$$

Contrast is a measure of intensity contrast between a pixel and its neighbor over the whole image.

$$Contrast = | iji - j| \; 2 \; p(i, j)$$

Homogeneity measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$Homogeneity = 1 \; 1+|i-j| \; ij \; 2 \; p(i, j)$$

Correlation is a measure of how to correlate a pixel to its neighbor over the whole image.

$$Correlation = i-\mu ij-\mu jpi,j\sigma i\sigma jij$$

Dissimilarity is a measure of distance between pairs of pixels in the region of interest.

$$Dissimilarity = i - jjpi,ji$$

Energy provides the sum of squared elements in the GLCM.

$$Energy = Pi,ji,j \; 2 \; [18]$$

After the implementation of GLCM, all the computed features have been saved in .csv file. The resulting size of .csv file becomes 2170*30 since the dataset contains 2170 images and 29 are the input extracted texture features and 01 is the output class label. Later, the resulting dataset is divided in the ratio 80:20, into two sets, train set and test set. Thus, train set contains 1736 images and test set contains 434 images.

G. Technique Used for Feature Selection

To further enhance the performance of the Random Forest model, Neighborhood Component Analysis (NCA) feature selection was applied, aiming to achieve better results by identifying and utilizing the most relevant features for classification.

H. Implementation of Neighborhood Component Analysis (NCA)

Neighborhood Component Analysis (NCA) was applied to the dataset for feature selection. NCA is a supervised dimensionality reduction method that enhances the predictive power of the model by selecting the most relevant features. Neighborhood Component Analysis for Feature Selection (NCAFSR) is implemented to select the most relevant features from the dataset before training a classification model. The process begins by importing necessary libraries such as pandas, numpy, and scikit-learn modules. Three datasets, representing different stages or conditions, are loaded from CSV files using pd.read_csv. Any missing values in these datasets are filled with the mean of the respective columns. The data is then normalized using the MinMaxScaler from scikit-learn's preprocessing module to scale the features between 0 and 1. After normalization, the data is combined into a single array and shuffled to ensure random distribution.

The combined data is split into features (X) and labels (Y), and then into training and testing sets using train_test_split. Feature selection is performed using Neighborhood Component Analysis for Feature Selection and Reduction (NCAFSR), which fits the model to the training data to learn the weights of the features. These weights help in transforming the training and testing sets by emphasizing more relevant features. A Random Forest Classifier is then trained on the transformed training set, and predictions are made on the transformed testing set. The performance of the classifier is evaluated using a confusion matrix, classification report, and accuracy score, providing insights into the model's effectiveness.

The confusion matrix illustrates the performance of a classification model across three classes. Class 0 has 138 true positives, 7 false negatives, and 11 false positives. Class 1 shows 124 true positives, 11 false negatives, and 15 false positives. Class 2 exhibits 134 true positives, 2 false negatives, and 3 false positives. The classification report provides detailed metrics for precision, recall, and F1-score for each class. Class 0 achieves a precision of 0.93, recall of 0.94, and F1-score of 0.93, based on 147 instances. Class 1 has a precision of 0.85, recall of 0.90, and F1-score of 0.87, encompassing 138 instances. Class 2 shows perfect precision (1.0), recall of 0.96, and F1-score of 0.98, covering 149 instances. The overall accuracy of the model is 0.91, with macro-averaged precision, recall, and F1-score also at 0.91, calculated across all 434 instances. The weighted averages for precision, recall, and F1-score are consistently 0.91. The "Accuracy" score provided is 0.9124423963133641, indicating the overall effectiveness of the model in correctly predicting classes across the dataset. It provides a complete workflow for detecting potato leaf diseases using machine learning. It demonstrates the importance of data preprocessing, feature selection, and the application of an effective classifier to achieve accurate and reliable results. By applying NCA for feature selection and using a Random Forest classifier, the model is optimized for better performance in detecting various potato leaf diseases.

I. Performance Metrics Used

In order to evaluate the performance of all the ML trained models for crop yield prediction, four important measures namely *accuracy*, *recall*, *precision*, and *F-score* has been computed. These measures can be calculated from confusion matrix described as under:

As multiple (three) class labels (Low, Mid, High) are there in both the wheat and mustard crop yield prediction problem, thus 3*3 confusion matrix is used. It is a technique for summarizing the performance of classification algorithms.

TABLE III: CONFUSION MATRIX

	Predicted Class			
		Low (L)	Mid (M)	High (H)
Actual Class	Low (L)	TP_L	E_{LM}	E_{LH}
	Mid (M)	E_{ML}	TP_M	E_{MH}
	High (H)	E_{HL}	E_{HM}	TP_H

- *True Positive (TP):* These are the correctly predicted positive values which mean value of both actual and predicted class is yes. True positive will be diagonal values for every class TP_L, TP_M, TP_H.
- *True Negative (TN):* These are the correctly predicted negative values which mean value of both actual and predicted class is no. True negative for a certain class will be the sum of all columns and rows excluding that class's column and row.

  TN for class *Low*, $TN_L = TP_M+E_{MH}+E_{HM}+TP_H$
  TN for class *Mid*, $TN_M = TP_L+E_{LH}+E_{HL}+TP_H$

  TN for class *High*, $TN_H = E_{LM}+E_{LH}+TP_M+E_{MH}$
- *False Positive (FP):* When value of actual class is no but the predicted class is yes. False positive (FP's) for a class is the sum of values in the corresponding column (excluding the TP).

  FP for class *Low* = $E_{ML}+E_{HL}$

  FP for class *Mid* = $E_{LM}+E_{HM}$

  FP for class *High* = $E_{LH}+E_{MH}$
- *False Negative (FN):* When value of actual class is yes but the predicted class is no; False Negative (FP's) for a class is the sum of values in the corresponding row (excluding the TP).

  FN for class *Low* = $E_{LM}+E_{LH}$
  FN for class *Mid* = $E_{ML}+E_{MH}$

  FN for class *High* = $E_{HL}+E_{HM}$
- *Accuracy:* Accuracy is the ratio of correctly predicted observations to the total observations. Accuracy is great measure but only when one has symmetric datasets where value of false positive and negative are almost same.

  Accuracy = TP+TN/TP+FP+FN+TN
- *Precision:* Precision is the ratio of correctly predicted observations to the total predicted positive observations. High precision leads to low false positive rate.

  Precision = TP/(TP+FP)
- *Recall:* Recall is the ratio of correctly predicted positive observations to the all observations in the actual class-yes.

  Recall = TP/(TP+FN)
- *F1-Score:* F1 Score is the weighted average of precision and recall. Mostly used when there is an uneven class distribution.

  F1-Score = 2*Recall*Precision / (Recall + Precision) [18]

IV. RESULTS AND DISCUSSION

The performance of various machine learning techniques was evaluated using several metrics. The K-Nearest Neighbors (KNN) algorithm achieved an accuracy of 77.1%, with a precision of 80%, recall of 77%, and an F-Score of 0.777. The Support Vector Machine (SVM) method demonstrated an accuracy of 79.4%, precision of 85%, recall of 79%, and an F-Score of 0.80. Decision Trees (DT) exhibited an accuracy of 80.8%, with both precision and recall at 81%, and an F-Score of 0.81. The Random Forest (RF) classifier showed a significant improvement, attaining an accuracy of 84.1%, precision of 85%, recall of 84%, and an F-Score of 0.84. Lastly, Naive Bayes (NB) presented an accuracy of 79%, precision of 83%, recall of 79%, and an F-Score of 0.78.

TABLE IV: EXPERIMENTAL RESULTS FOR ML TECHNIQUES UNDER STUDY

ML Techniques	Accuracy	Precision	Recall	F-Score
KNN	77.1%	80%	77%	0.77
SVM	79.4%	85%	79%	0.80
DT	80.8%	81%	81%	0.81
RF	84.1%	85%	84%	0.84
NB	79%	83%	79%	0.78

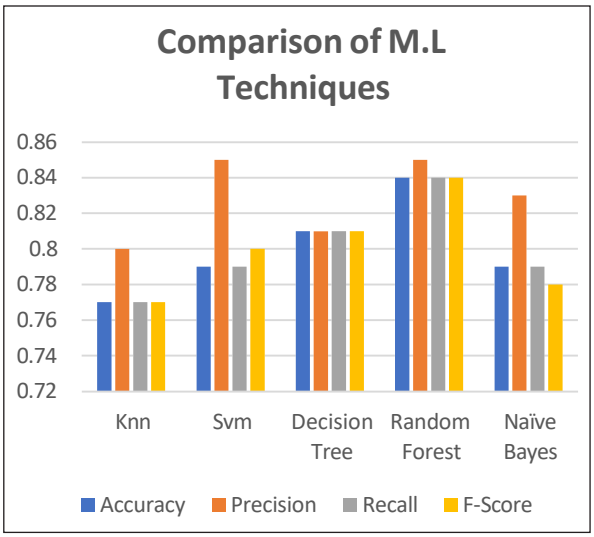


Fig. 4: Comparative Analysis of All ML Techniques

The Hybrid ML-Based Model, specifically the Modified Random Forest (Modified RF), demonstrated superior performance compared to the individual machine learning techniques. The Modified RF model achieved an impressive accuracy of 91.2%, with precision, recall, and F-Score all at 91%. This indicates that the hybrid approach significantly enhances the model's predictive capabilities.

TABLE V: EXPERIMENTAL RESULTS FOR HYBRID ML-BASED MODEL UNDER STUDY

ML Techniques	Accuracy	Precision	Recall	F-Score
Hybrid ML-Based Model	91.2%	91%	91%	0.91

V. CONCLUSION

In this research, we explored the efficacy of machine learning techniques for detecting potato leaf diseases based on texture features extracted from grayscale images using the Gray-Level Co-Occurrence Matrix (GLCM). The study encompassed three main conditions: Early Blight, Late Blight, and Healthy leaves, each characterized by distinct textural patterns captured through GLCM analysis. Our findings demonstrated that the Random Forest classifier, augmented by Neighborhood Component Analysis for Feature Selection (NCAFS), achieved the highest accuracy of 84.1% among the tested models. This approach not only effectively classified potato leaf diseases but also highlighted the importance of robust feature extraction and selection methods in enhancing classification accuracy. The results underscore the potential of machine learning in precision agriculture, offering a promising tool for early disease detection and proactive crop management strategies. Future research could explore larger datasets, integrate more sophisticated feature extraction techniques, and validate the model in field conditions to further enhance its applicability and impact in agricultural settings.

REFERENCES

[1] V. P. Gaikwad, and V. P. Musande, "Potato plant leaf disease detection using CNN model," *European Chemical Bulletin*, vol. 12, no. 1, pp. 516-527, 2023.

[2] M. A. R. Nishad, M. A. Mitu, and N. Jahan, "Predicting and classifying potato leaf disease using K-means segmentation techniques and deep learning networks," *Procedia of Computer Science*, vol. 212, pp. 220-229, 2022, doi: https://doi.org/10.1016/j.procs.2022.11.006.

[3] V. Bhuyar, "Comparative analysis of classification techniques on soil data to predict fertility rate for Aurangabad district," *International Journal of Emerging Trends & Technology in Computer Science*, vol. 3, no. 2, pp. 200-203, 2014.

[4] V. Singh, Varsha, and A. K. Misra, "Detection of unhealthy region of plant leaves using image processing and genetic algorithm," *International Conference on Advances in Computer Engineering and Applications*, 2015, doi: https://doi.org/10.1109/icacea.2015.7164858.

[5] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, doi: https://doi. org/10.1109/cvpr.2016.90.

[6] V. Rajeshwari, and K. Arunesh, "Analyzing soil data using data mining classification techniques," *Indian Journal of Science and Technology*, vol. 9, no. 19, pp. 1-4, 2016.

[7] R. Sujatha, "A study on crop yield forecasting using classification techniques," *IEEE*, 2016.

[8] M. Islam, K. Wahid, and P. Bhowmik, "Detection of potato diseases using image segmentation and multiclass support vector machine," *IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2017, doi: https://doi.org/10.1109/ CCECE.2017.7946594.

[9] D. Tiwari, M. Ashish, N. Gangwar, A. Sharma, S. Patel, and S. Bhardwaj, "Potato leaf diseases detection using deep learning," *Proceedings of the International Conference on Intelligent Computing and Control Systems (ICICCS 2020) IEEE Xplore*, 2020, doi: https://doi. org/10.1109/iciccs48265.2020.9121067.

[10] M. I. Tarik, S. Akter, A. A. Mamun, and A. Sattar, "Potato disease detection using machine learning," *Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV)*, 2021, doi: https://doi.org/10.1109/ icicv50876.2021.9388606.

[11] H. Afzaal, A. A. Farooque, A. W. Schumann,N. Hussain, A. McKenzie-Gopsill, T. Esau, F. Abbas, and B. Acharya, "Detection of a potato disease (early blight) using artificial intelligence," *Remote Sensing*, vol. 13, no. 3, 2021, doi: https://doi.org/10.3390/rs13030411.

[12] J. Rashid, I. Khan, G. Ali, S. H. Almotiri,M. Alghamdi, and K.Masood, "Multi- level deep learning model for potato leaf disease recognition," *Electronics*, vol. 10, no. 17, 2021, doi: https://doi.org/10.3390/ electronics10172064.

[13] A. Singh, and H. Kaur, "Potato plant leaves disease detection and classification using machine learning methodologies," *IOP Conference Series: Materials Science and Engineering*, 2021, doi: https://doi. org/10.1088/1757-899x/1022/1/012121.

[14] Data Collection. [Online]. Available: https:// www.kaggle.com/datasets/emmarex/ plantdisease

[15] K. Padmavathi, and K. Thangadurai, "Implementation of RGB and grayscale images in plant leaves disease detection - Comparative study," *Indian Journal of Science and Technology*, vol. 9, 2016, doi: https://doi. org/10.17485/ijst/2016/v9i6/77739.

[16] Implementation of the Proposed Framework. [Online]. Available: https://www.mathworks. com/help/images/texture-analysis-using-the- gray-level-co-occurrence-matrix-glcm.html

[17] Gray-Level Co-Occurrence Matrix (GLCM). [Online]. Available: https:// medium.com/@girishajmera/feature- extraction-of-images-using-glcm-gray-level cooccurrence-matrix-e4bda8729498

[18] Y. Gupta, H. Kour, J. Manhas, and V. Sharma, "Classification of spices using machine learning techniques," *International Journal of Knowledge Based Computer Systems*, vol. 10, no. 1, pp. 27-32, 2022.

International Journal of Knowledge Based Computer Systems
13 (1) 2025, 10-30
www.publishingindia.com//ijkbcs/

Contemporary Trends in Plant Health Care and Agriculture Extension Services

Jyoti Lakhani

Department of Computer Science, Maharaja Ganga Singh University, Bikaner, Rajasthan, India.
Email: jyotilakhani@mgsubikaner.ac.in

Abstract: The present communication is a survey of digital tools related to plant health care which is beneficial for farmers. The computational methods are used by researchers these days for developing agricultural extension services. This research used more than 200 research papers in which nine diverse research issues are observed in this survey. We have identified 9 research areas in plant health care in which computational methods can be used. Image processing is the most common computational method used by researchers in solving plant health care related issues. The computational methods are also identified based on each research area separately. We have identified most commonly used computational tool for each of the nine plant health care issue. One of the observations is that the researchers use Decision Support Systems tools for sharing knowledge about plant health care. The results of this study will be helpful for computer scientists to select research gaps in plant health care related issues.

Keywords: Agriculture extension services, Computational methods, Digital tools, Plant health care.

I. Introduction

Modern technological evolution plays a vital role in the field of plant health management and plant care. Farmers can get instant help with information for crop management and disease management by using current digital technologies. Farmers can get acquainted with support services, facilities, schemes, and the schemes of the government on time with modern digital technologies that help them to recognize plant diseases early, cure ailments of plants timely, and improve crop production. Farmers can also get advisory services from scientists with the help of digital tools to adapt to modern agricultural practices. Fig. 1 represents applications of information technology for sustainable agriculture practices. These services include climate prediction, an online marketplace, field-level decision support systems, pest management, disease management, damage analysis, and plant health monitoring. Modern computational tools and technologies used in the published literature are studied. This information will also be helpful for researchers interested in plant health care services.

Fig. 1: Applications of IT for Sustainable Agriculture Practices

II. Digital Tools and Techniques

A. Knowledge Based Tools and Techniques

a. Knowledge Sharing

Knowing modern agricultural tools is necessary for farmers. The world faces local atmospheric

Uncertainties and variable weather patterns due to constant climatic change. Early and on-time information about local weather conditions is essential for a farmer to know. It is possible only when farmers can use modern tools and technology. The Internet of Things (IoT) can fill the gap between technology and the traditional agriculture system. Technological shifts toward agriculture and farming show a positive impact on agriculture services. Governments and NGOs are trying to fill this gap between farmers and technology by developing knowledge-sharing portals and websites. Knowledge is shared through meetings and discussions by the farmers traditionally. Farmers can also share their observations with their peers while working in farms and marketplaces. Modern technological innovations can provide farmers with an efficient and fast agriculture information- sharing platform. WeFarm [1] is a farmer-to-farmer digital network in which farmers can communicate with each other via SMS and online chats. Nearly

1.8 million small-scale farmers are using this platform. Agtech company Verdant connects African farmers to provide agricultural insight and extension services. It uses mobile technology to increase market accessibility [2]. International Fund for Agricultural Development (IFAD) and International Development Research Centre (IDRC) jointly develop a knowledge networking for rural development in the Asia Pacific. It is a knowledge-sharing tool [3]. Online and offline knowledge sharing can be effective if websites, blogs, and digital radios are used instead of specialized applications. Busoga Rural Open Source Development Initiative (Kampala, Uganda) (BROSDI) broadcasts a live monthly program that facilitates farmers about effective farming practices [4]. The knowledge disseminated by audio blogs can reach a wider audience [5]. WorldSpace radio uses two satellites named- AfriStar and AsiaStar, to broadcasts more than 100 digital-quality audio channels for farmers around the world. Agricultural blogging allows quick dissemination of information to the farmers. The Voices of Africa blog [6] and Kisan blog [7] show the contribution of blogging in sharing the experiences of rural farmers. The Collecting and Exchange of Local Agricultural Content (CELAC) blog is another blog where farmers and agriculture practitioners can post articles [8]. Some public and private organizations use mobile technology to provide market information to farmers. Tradenet. biz is an enterprise that covers 15 countries and 500 markets to offer information related to supply-chain, from price updates, harvests, transport, trading offers, disease outbreaks, and weather. The poor farmers in Tanzania use mobile phones to access market information in real-time under the First Mile Project [9-14]. It is a joint venture of the Government of Switzerland and the Government of Tanzania.

b) Climate Predictions

Due to climate change, fluctuations in the temperature and rainfall observed impose a diverse effect on agricultural production. The North-Eastern United States developed free, online decision-making tools for farmers in a Cornell Climate Smart Farming program. The developed tools combine local weather stations and agricultural data and inform farmers that are useful in decision-making [15]. The United States Environmental Protection Agency also developed a toolkit for farmers. The developed climatic tools include the growing degree day calculator [16], water deficit calculator [17], Nitrogen management tool [18], crop planting scheduler [19], drought monitor [20-21], EPA's climate resilience evaluation and awareness tool [22] for annual total precipitation, annual average temperature. The South-East Climate Consortium developed a web tool that deals with volatile weather patterns in Florida [23]. The Climate Predictability Tool is a software package designed by Columbia University for the seasonal Climate forecast Model [24]. Machine learning can also play role in Climate prediction [25-27]. Example of some of the contemporary tools are - outbreak risk prediction [28], seasonal precipitation [29-31], weather prediction [32], climate change prediction [33, 34], solar radiation prediction [35-36], to predict plant growth and yield [37-38]. Some researchers use GIS-based machine learning for climate downscaling prediction [39]. Fig. 2 is showing some eTools helpful for sustainable agricultural practices.

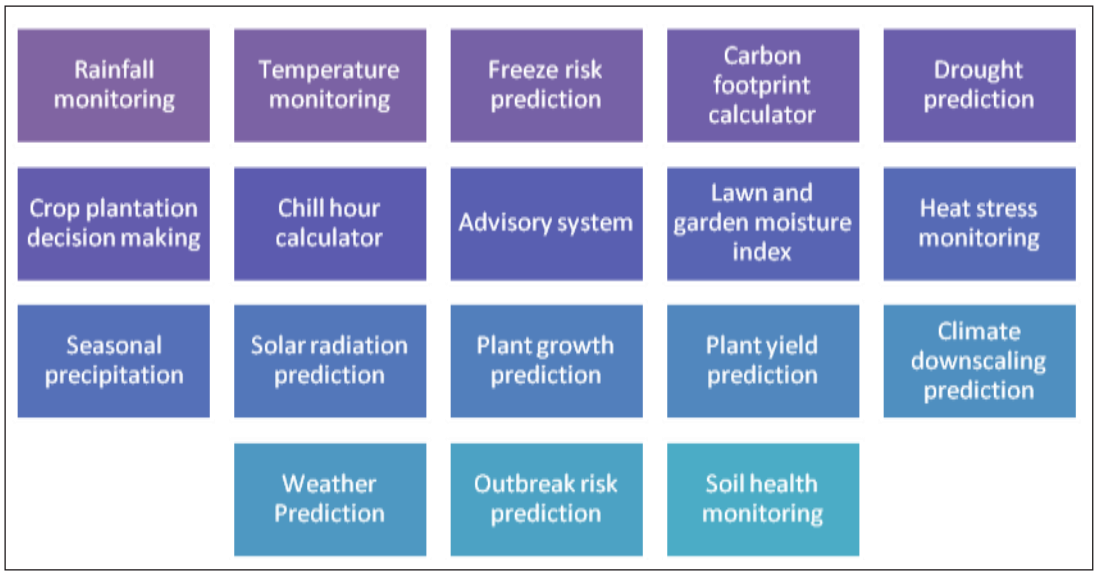


Fig. 2: eTools for Sustainable Agriculture Practices

c) Linkages with Agriculture Research Scientists and Local Advisors

Active Linkages between researchers and farmers are necessary. Agriculture-based research could be successful if it is helpful for farmers. This linkage of research and farmers such interactions lead to the design and delivery of appropriate technology practically suitable for the stakeholders [40]. Many gaps between agricultural research and extensions cause the complexity of agriculture research products [41-42]. Communication and digital knowledge-share are effective for this [43].

d) Field-Level Decision Support System

Applications of decision support systems in agriculture have increased rapidly in the last decade [44-46]. A decision support system (DSS) is an interactive software-based system used to help decision-makers to compile information from a combination of raw data to identify and solve problems. It also makes an optimized decision [44, 45]. An effective DSS could be paper-based, software, or a mobile app. A study suggests that software is 28% more useful for a farmer than paper-based (22%) DSS, followed by mobile app based (10%) [47]. Nowadays, the evolution of agriculture steps into Agriculture 4.0, thanks to the employment of current technologies like the Internet of Things [48], Big Data [49], Artificial Intelligence, Cloud Computing [50], and Remote Sensing [51-53]. The applications of these technologies can improve the efficiency of agricultural activities significantly. A machine learning-based DSS could be more effective [54-55]. A decision support system used for farm management [56], assessment and management of soil functions [57], enhancing crop productivity [45, 110, 112], farm machinery management [58-59], agro-technology transfer [46], improving the efficiency of water use [60]. The Watson Decision platform for agriculture combines IBM's advanced capabilities in artificial intelligence, the Internet of things, and cloud computing [61].

e) Tools for Plant Protection

Early control of pests is a must in agriculture. Pesticides can harm the crop and human health, hazardous to the environment also. It is required to inform farmers about sustainable plant protection techniques [62].

f) Sustainable Plant Protection

The global impact of climate change is influencing plant production and constraining farmers' change crop monitoring and management strategies. Farmer finds difficulty in the present climatic scenario to get ecological and economic ROI (return on investment).

One factor is to protect plants from plant disease outbreaks. For this, a farmer must concentrate on plant type, soil-weather conditions, predictions and forecasting of disease outbreaks, and optimal timing and quantity of using pesticides. Smart field monitoring viz IoT-based soil monitoring systems improve the soil as per the crop [63] and proper irrigation [64-65].

g) Online Marketplace for Sustainable Plant Protection

To have sustainable plant products, farmers must have sustainable plant protection products such as biofertilizers and organic pesticides at affordable prices. All India Network Project on biofertilizers (aicrp BNF) is an example of such an initiative of The Indian Council of Agricultural Research (ICAR) [66]. Indian Farmers Fertilizer Cooperative Limited (IFFCO) [67] and e-Urvarak is a dashboard developed by National Informatics Center, India, ensures adequate and on-time delivery of sustainable fertilizers to the farmers and facilitates easy monitoring by various stakeholders [68].

B. Tools for Plant Health Care

a) Plant Health Care is an Important Issue

The production is directly dependent on the healthy crop. It is desirable to keep plants healthy by early detection of diseases and attack of pests. Continuous monitoring of plants is the main factor for disease monitoring. AI-enabled plant disease detection [69-70] and monitoring systems [71-72] are helpful in disease detection [73-74], monitoring [71-72], classification [75], and pest detection [76]. The computational techniques used in plant disease detection and monitoring digital tools are neural networks [73, 77], multimedia sensor networks [74], remote sensing [78], biosensors [79], image processing [69], and deep learning [75-76].

b) Detecting, Monitoring and Controlling Pest Threats

Climate change affects the behavior, distribution, life cycle, and outbreak potential of pests. An effective pest monitoring system is essential as an embedded feature of climatic predictions to eradicate potential threats. During the outbreaks, a real-time monitoring system must have pest risk prediction data and pest distribution observations. An effective digital pest control and monitoring system can have several modules as pest detection [76, 78, 80-85], pest management [86-88], pest monitoring [89], pest controlling [89-90], pest classification [85, 91], pest surveillance [80, 92-93] and, damage detection [94-95]. Researchers use deep learning [76, 96-97], remote sensing [78], SVM classification [84, 98], computer vision [81, 98, 99-101], Supervised learning [102], clustering [103], intelligent mobile applications [104], radio [105], multi-object detection method [104], image processing [83], bio-inspired methods [106], artificial intelligence [88], wireless sensors [90], AIoT [82] and electronic nose [107-108], artificial intelligence [109], sensing technologies [110] for pest detection and monitoring. PestinaNet [111] is an example of a real-time crop pest detection system.

i) Pest Recognition

Pest recognition is finding pest attacks in plants using various computational techniques [112-113]. Researchers attempt to successfully recognize pests in plant leaves [114-115], roots [89] [102], tuber [89], fruits [116], and trunk [115]. Diverse computational techniques used by researchers in pest recognition. Some of this are-image processing [114, 117], attention-embedded lightweight network [112], artificial intelligence [119], SVM [120], K-means clustering [91, 121], machine learning [91], deep learning [116], embedded system [122], machine vision [99].

ii) Damage Analysis

Another issue related to pest control and management is detecting the amount of damage by pests [94, 95, 123]. Farmers use information about the severity of the infection and the damage to the plant. It may help farmers decide to apply for medicine or discard the plant. The computation techniques used by researchers for damage analysis are pattern recognition [94], RBF-SVM algorithm [124], artificial nose [125-126], remote sensing [127], computer vision [128], and image processing [123].

iii) Pest Management

An initiative is required to develop response plans to tackle crop health threats [129]. Digital pest management [130] helps farmers to elevate pest control by automating the monitoring of pests. These Intelligent and automatic systems improve pest monitoring and support pest managers smartly. An example of this is rodent traps using IoT Technology [131]. The classification and detection of insects in crops using intelligent pest management [132] is an efficient way of pest control. There is evidence of using climate-smart pest management systems [86], Automatic Moth detection [133], and Environment monitoring [134] in the literature. Information technology tools like Drones [135], Satellite images [136], and agricultural aircraft [137-138] for pest control and management system. The computing technologies used for this purpose include image processing [132, 133, 138-139], decision support systems [148], deep learning [132,140], GIS [141], and web technology [142].

c) Detection, Monitoring, Management and Controlling Plant Diseases

i) Disease Detection

Disease detection is the process of detecting diseases from infected plants. Many disease detection tools are discussed in the literature [143-144, 70, 79]. The digital techniques used for disease detection in plants are image processing [69, 145-152], sensors [153], robotics [154], deep learning [76, 149, 151-152, 155-161], hyperspectral imaging and image sensing [162-164], Internet of Things (IoT) [165-166], neural networks [167-168], machine learning [150,169], orthogonal learning [168], evolutionary computation [168], swarm intelligence [168], support vector machine (SVM) [170-171], soft computing [146, 149]. The AuToDiDAC [172] and pathology [154] are plant disease detection tools designed for plant disease detection.

ii) Disease Monitoring and Management

Combining modern plant disease monitoring and management tools with disease detection tools is required. It is beneficial for the development of integrated high-tech crop protection systems in the future [72, 173]. The computational methods used by the researchers for this purpose are IoT [174-178], neural networks [77, 179], robotics [154, 180], unmanned aerial vehicles [178, 181], soft computing [182], image processing [182- 183], remote sensing imaginary [184-186], deep learning [187], geographical Information system (GIS) [188], machine learning [176], data mining [189], wireless sensor network [178, 190], wearable sensors [191], online decision support system (DSS) [192] and, hyperspectral technology [193]. A playhouse is a plant monitoring system that uses a convolutional neural network (CNN) for disease monitoring [77]. Chlorophyll fluorescence imaging is another tool to monitor the progress of a root pathogen in a perennial plant [183].

A decision support system was developed by researchers for decision-making [194-195]. An expert system is used for this purpose [196]. Quant is a software to quantify disease severity [197-198]. IoT tools have been used in literature by researchers for detecting and controlling plant diseases [174].

III. Discussion

Pest detection, pest monitoring, disease detection, and disease monitoring issues are common research areas in the literature. Monitoring of field, soil and irrigation, damage detection, damage prediction, and predictions of real-time early pest approaches are some issues ignored by IT professionals. The development of intelligent computational methods for plant health care is required. Table I indicates the observed research areas related to plant health care and the computational methods used for that issue.

TABLE I: COMPUTATIONAL METHODS USED IN DIFFERENT RESEARCH AREAS RELATED TO PLANT HEALTH CARE

Sr. No.	Major Research Area	Sub Areas	Computational Method Used
1	Knowledge Based Tools & Techniques	Knowledge Analysis Knowledge Sharing	Decision Support System [44, 45-46, 50, 56-59] GIS [39] Digital Knowledge Sharing Tools [43] Mobile Apps [47] Internet of Things [48, 61] Big Data [49] Artificial Intelligence [50, 61] Cloud Computing [50, 61] Remote Sensing [51-53] Machine Learning [54-55]
2	Field Monitoring	Field Monitoring	Smart Applications [202-204]
		Soil Monitoring	IoT [63]
		Irrigation Monitoring	IoT [64-65] Decision Tool [205]
		Crop Productivity Monitoring	Decision Support System [206]
3	Plant Health Monitoring	Disease Detection Disease Classification	Neural Networks [73, 87] Multimedia Sensor Networks [74] Remote Sensing [78] Biosensors [79] Image Processing [69] Deep Learning [75-76]
4	Pest Detection	Pest Detection	Deep Learning [76, 96-97] Remote Sensing [78] SVM Classification [98, 114] Computer Vision [81, 98-101] Supervised Learning [102] Clustering [103] Smart Phones Applications [104] Radio [101] Multi Object Detection Method [104] Image Processing [83] Bio-Inspired Methods [106] Artificial Intelligence [19, 109] Wireless Sensors [90] AIoT [82] Electronic Nose [107-108] Sensing Technologies [110]
5	Pest Recognition	Pest Recognition	Image Processing [114, 117] Attention-Embedded Lightweight Network [118] Artificial Intelligence [119] SVM [120] K-means Clustering [91, 121] Machine Learning [91] Deep Learning [116] Embedded System [122] Machine Vision [99]

Sr. No.	Major Research Area	Sub Areas	Computational Method Used
6	Pest Management (Monitoring, Controlling, and Surveillance)	Climate-Smart Pest Management Systems Automatic Pest Detection Environment Monitoring	Drones [137] Satellite Images [136] Agricultural Aircraft [137-138] Image Processing [132-133, 138-139] Decision Support System [148] Deep Learning [132, 140] GIS [141] Web Technology [142]
7	Damage Analysis	Damage Detection Damage Severity Analysis	Pattern Recognition [94] RBF-SVM Algorithm [124] Artificial Nose [125-126] Remote [127] Computer Vision [128] Image Processing [123]
8	Plant Disease Detection	Plant Disease Detection	Image Processing [69, 112, 145-147, 149-152] Sensors [153] Robotics [154] Deep Learning [76, 149, 151, 155-157, 158-161, 199, 201] Hyperspectral Imaging and Image Sensing [162-164] Internet of Things (IoT) [165-166] Neural Networks [167-168] Machine Learning [150,169] Orthogonal Learning [168] Evolutionary Computation [168] Swarm Intelligence [168] Support Vector Machine (SVM) [170-171] Soft Computing [149]
9	Disease Monitoring and Management	Disease Monitoring Disease Management Disease Control	IoT [172, 176-178] Neural Network [77, 179] Robotics [154, 180] Unmanned Aerial Vehicles [178,181] Soft Computing [182] Image Processing [182-183] Remote Sensing Imaginary [184-186] Deep Learning [187] Geographical Information System (GIS) [188] Machine Learning [176] Data Mining [189] Wireless Sensor Network [178, 190] Wearable Sensors [191] Online Decision Support System (DSS) [192] Hyper Spectral Technology [193] Expert Systems [196]

It is required to identify computational methods before solving an issue related to plant health. A review of more than 200 already published research papers has been performed to answer this query. Fig. 3 shows commonly used computational methods in already published research papers for plant health care. Fig. 3 is a frequency distribution chart of computational methods by computer scientists used for plant health care issues. It is observed that image processing and image analysis is the most frequently used computational method for solving plant health care issues, followed by the internet of things (IoT) and decision support system (DSS).

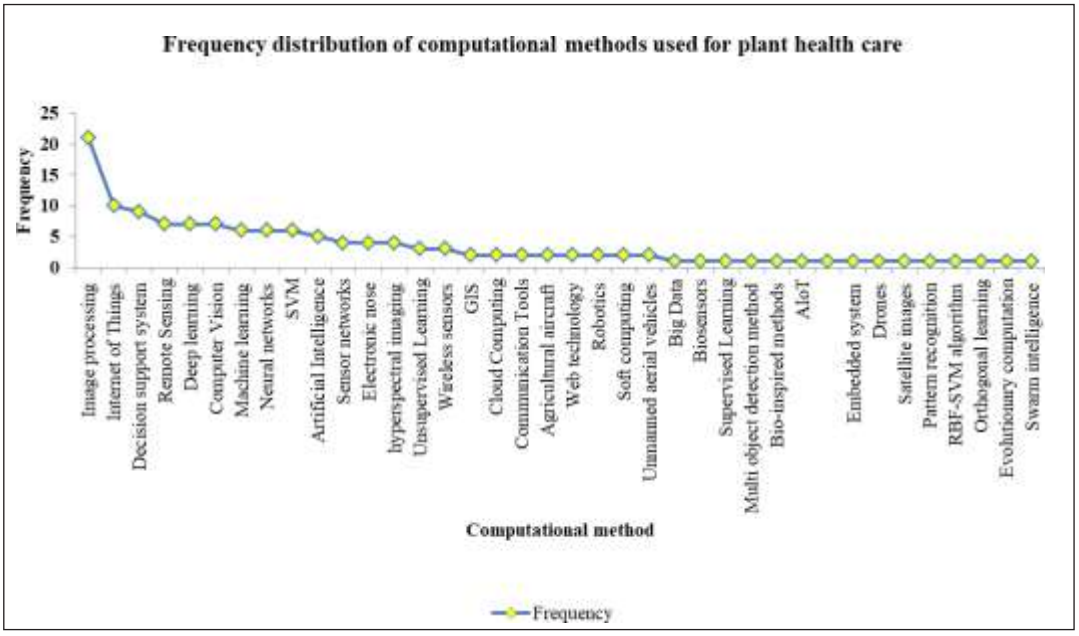


Fig. 3: Frequency Distribution Graph for Computational Methods Opted by Researchers for Development of Plant Health Care Related Tools and Applications

Fig. 4 is showing the share of various computational methods used in different research areas observed in the survey. For knowledge Analysis and sharing of plant related information, decision support systems (35%) and remote sensing (13%) is most commonly used, followed by cloud computing (9%) and artificial intelligence (9%). For Field Monitoring related issues smart application tools and internet of things shares 50-50% part.

In plant health management applications, Biosensors (36%), neural networks (18%) and deep learning are most commonly used by the researchers. For pest detection, computer vision (19%), deep learning (12%) and support vector machines (8%) are given importance. For pest monitoring and management systems agriculture aircrafts (15%) and drones (8%) and, GIS (8%) are used with support of image processing (31%) and deep learning (15%) etc. For pest recognition image processing (18%) and unsupervised learning (18%) is most commonly used by the researchers. Electronic nose is a modern technology used in 29% research papers reviewed for damage analysis. Some other methods used for this damage analysis are- pattern recognition (14%), support vector machines (14%), remote sensing (14%), image processing (14%) and computer vision (14%). For plant disease detection deep learning (29%) is most commonly used followed by image processing (24%).

IV. Conclusion

Plant health is an important issue that has direct effect on crop production. In the present survey of more than 200 papers it was observed that image processing, followed by Decision Support System (DSS), Internet of Things (IoT), computer vision and deep learning are contemporary computational

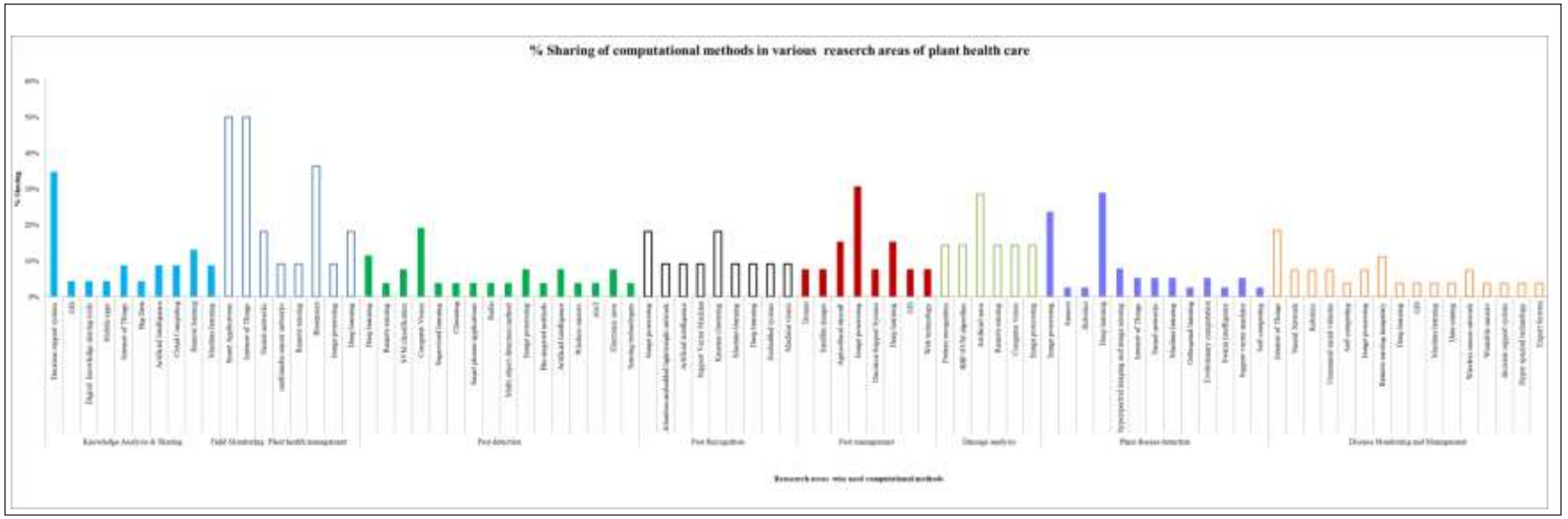


Fig. 4: A Graph Showing Research Area Wise % Sharing of Used Computational Method in Development of Plant Health Care Related Tools and Applications

tools used by researchers for various research areas related to plant health care. Among nine identified research areas, plant disease detection and disease monitoring is most frequently researched. For each observed nine research areas in the literature, the most used computational method identified. The area most used tool duo for each of the observed research area are: (i) Knowledge Analysis & Sharing – DSS (35%), (ii) Field Monitoring – IoT (50%), (iii) Plant Health Management – Biosensors (36%), (iv) Pest Detection – Computer Vision (19%), (v) Pest Recognition – Image Processing (18%), (vi) Pest Management – Image Processing (31%), (vii) Damage Analysis – Artificial Nose (29%), (viii) Plant Disease Detection – Deep Learning (29%) and (ix) Disease Monitoring and Management – IoT (19%). The present research can be used to identify research gaps in terms of plant health issues and used computational methods.

REFERENCES

[1] Wefarm - The World's Largest Platform for Small-Scale Farmers. [Online]. Available: https://wefarm.com/about

[2] https://verdant.ng/farmers/

[3] Facilitator's Guide. [Online]. Available at: https://www.fsnnetwork.org/sites/default/files/ifad_km.pdf

[4] Audio Blog for Knowledge Sharing. [Online]. Available: http://audioblog.podbean.com/

[5] WorldSpace Community Radio. [Online]. Available: www.worldspace.com

[6] Voices of Africa Blog. [Online]. Available: http://voicesofafrica.africanews.com

[7] Kisan Blog. [Online]. Available: http://kisan.wordpress.com

[8] Collecting and Exchange of Local Agricultural Content Blog. [Online]. Available: http://celac.wordpress.com

[9] C. Lightfoot, H. Gillman, U. Scheuermeier, and V. Nyimbo, "The first mile project in Tanzania," *Mt. Res. Dev.*, vol. 28, no. 1, pp. 13-17, 2008.

[10] P. Njenga, S. Willilo, and J. Hine, "First mile transport challenges for smallholder tomato farmers along Ihimbo–Itimbo Road, Kilolo District Tanzania," *AFCAP Report, AFCAP/TAN2015C*, 2015.

[11] J. Hine, P. Njenga and S. Wilolo, "Overcoming the first mile - Lessons from farmers in Kenya & Tanzania".

[12] A. Nazir, D. Vervoort, and C.L. Reddy, "From the first mile to the last: Challenges of the global surgical supply chain," *The American Journal of Surgery*, vol. 222, no. 4, pp. 709-711, 2021.

[13] P. Njenga, G. Wahome, and J. Hine, "Pilot study on first mile transport challenges in the onion smallholder sector," *AFCAP Report, AFCAP/GEN/147*, 2014.

[14] M. Skjöldevald, "Small-scale farmers and the shift in the food trading paradigm: A comparison of two rice supply chains in Babati district, Tanzania," 2008.

[15] U.S. Climate Resilience Toolkit. [Online]. Available: https://toolkit.climate.gov/tools

[16] Climate Smart Farming CSF Growing Degree Day Calculator. [Online].

[17] Climate Smart Farming CSF Water Deficit Calculator. [Online].

[18] Climate Smart Farming Adapt-N Nitrogen Management Tool. [Online].

[19] Climate Smart Farming CSF Winter Cover Crop Planting Scheduler. [Online].

[20] Climate Smart Farming CSF New York State / Northeast Drought Atlas. [Online].

[21] Climate Smart Farming CSF Caribbean Drought Atlas. [Online].

[22] Document Display | NEPIS | US EPA. [Online].

[23] AgroClimate – Tools for Managing Climate Risk in Agriculture. [Online].

[24] J. M. Simon, K. T. Michael, S. Lulin, and G. M. Ángel, "Climate predictability tool version 16.1.2," *Columbia University Academic Commons*, 2019.

[25] D. Rolnick, P. L. Donti, L. H. Kaack, K. Kochanski, A. Lacoste, K. Sankaran, A. S. Ross, N. Milojevic-Dupont, N. Jaques, A. Waldman-Brown, and A. S. Luccioni, "Tackling climate change with machine learning," *ACM Comput. Surv. (CSUR)*, vol. 55, no. 2, pp. 1-96, 2022.

[26] K. S. Krishna, and N. Padhy, "Climate prediction by using a machine learning approach: A case study," in *The Role of IoT and Blockchain: Techniques and Applications*. CRC Press, 2022.

[27] Y. Liu, D. Li, S. Wan, F. Wang, W. Dou, X. Xu, S. Li, R. Ma, and L. Qi, "A long short-term memory-based model for greenhouse climate prediction," *Int. J. Intell. Syst.*, vol. 37, no. 1, pp. 135-151, 2022.

[28] Z. Tianyu, R. Fethi, B. Ali, C. Xin, P. Hye-Young, Y. Lina, and R.M. Chandini, "Use of automated machine learning for an outbreak risk prediction tool," *Inform. Med. Unlocked*, vol. 34, no. 101121, 2022.

[29] J. Yang, Y. Xiang, J. Sun, and X. Xu, "Multi-model ensemble prediction of summer precipitation in China based on machine learning algorithms," *Atmosphere*, vol. 13, no. 9, p. 1424, 2022.

[30] V. S. Monego, J. A. Anochi, and H. F. de Campos Velho, "South America seasonal precipitation prediction by gradient-boosting machine-learning approach," *Atmosphere*, vol. 13, no. 2, p. 243, 2022.

[31] A. Ghazikhani, I. Babaeian, M. Gheibi, M. Hajiaghaei-Keshteli, and A. M. Fathollahi-Fard, "A smart post-processing system for forecasting the climate precipitation based on machine learning computations," *Sustainability*, vol. 14, no. 11, p. 6624, 2022.

[32] B. Bochenek, and Z. Ustrnul, "Machine learning in weather prediction and climate analyses - Applications and perspectives," *Atmosphere*, vol. 13, no. 2, p. 180, 2022.

[33] M. Kadkhodazadeh, M. ValikhanAnaraki, A. Morshed-Bozorgdel, and S. Farzin, "A new methodology for reference evapotranspiration prediction and uncertainty analysis under climate change conditions based on machine learning, multi criteria decision making and Monte Carlo methods," *Sustainability*, vol. 14, no. 5, p. 2601, 2022.

[34] H. Li, Y. Yang, H. Wang, P. Wang, X. Yue, and H. Liao, "Projected aerosol changes driven by emissions and climate change using a machine learning method," *Environ. Sci. Technol.*, vol. 56, no. 7, pp. 3884-3893, 2022.

[35] S. Ghimire, R.C. Deo, D. Casillas-Pérez, and S. Salcedo-Sanz, "Boosting solar radiation predictions with global climate models, observational predictors and hybrid deep- machine learning algorithms," *App. Energy*., vol. 316, p. 119063, 2022.

[36] J. Mustafa, S. Husain, U. A. Khan, and M. M. Akhtar, "Prediction of diffuse solar radiation using machine learning models based on sunshine period and sky-clearness index for the humid-subtropical climate of India," *Environ. Prog. & Sustainable Energy*, vol. 42, no. 29, p. e13973, 2022.

[37] G. Nagarajan, and K. Mahesh, "Machine learning based predict plant growth and yield in greenhouse environments," in *Technology Innovation in Mechanical Engineering*, Springer, Singapore, 2022.

[38] W. Zhou, Y. Liu, S. T. Ata-Ul-Karim, Q. Ge, X. Li, and J. Xiao, "Integrating climate and satellite remote sensing data for predicting county-level wheat yield in China using machine learning methods," *Int. J. App. Earth Obs. Geoinf.*, vol. 111, p. 102861, 2022.

[39] S. Mohapatra, M. Kundu, and S. Mohanty, "Climate downscaling and prediction using GIS-based machine learning," in *Proc. Second International Conference on Computer Science, Engineering and Applications (ICCSEA)*, IEEE, 2022, pp. 1-6.

[40] D. Byerlee, and R. Tripp, "Strengthening linkages in agricultural research through a farming systems perspective: The role of social scientists," *Exp. Agric.*, vol. 24, no. 2, pp. 137-151, 1988.

[41] S. Kumar, R. Kumar, and U. Sah, "Linkages among research, extension and farmers - A review," *Agric. Rev. - Agric. Res. Commun. Cen. Ind.*, vol. 22, no. 3/4, pp. 215-221, 2001.

[42] R. Bourgeois, "Structural linkages for integrating agricultural research and extension," 1990.

[43] R. Rivera-Huerta, G. Dutrénit, J. M. Ekboir, J. L. Sampedro, and A. O. Vera-Cruz, "Do linkages between farmers and academic researchers influence researcher productivity? The Mexican case," *Res. Pol.*, vol. 40, no. 7, pp. 932-942, 2011.

[44] M. Rinaldi, and Z. He, "Decision support systems to manage irrigation in agriculture," *Adv. Agron.*, vol. 123, pp. 229-279, 2014.

[45] Z. Zhai, J. F. Martínez, V. Beltran, and N. L. Martínez, "Decision support systems for agriculture 4.0: Survey and challenges," *Comput. Electron. Agric.*, vol. 170, no. 105256, 2020.

[46] Y. W. Jame, and H. W. Cutforth, "Crop growth models for decision support systems," *Can. J. Plant Sci*., vol. 76, no. 1, pp. 9-19, 1996.

[47] C. D. Rose, W. J. Sutherland, J. William, C. Parker, M. Lobley, M. Winter, C. Morris, S. Twining, C. Ffoulkes, and T. Amano, "Decision support tools for agriculture: Towards effective design and delivery," *Systems*, vol. 149, pp. 165-174, 2016.

[48] F. J. Ferrández-Pastor, J. M. García-Chamizo, M. Nieto-Hidalgo, J. Mora-Pascual, and J. Mora-Martínez, "Developing ubiquitous sensor network platform using internet of things: Application in precision agriculture," *Sensors*, vol. 16, no. 7, p. 1141, 2016.

[49] S. Wolfert, L. Ge, C. Verdouw, and M. J. Bogaardt, "Big data in smart farming - A review," *Agric. Syst.*, vol. 153, pp. 69-80, 2017.

[50] J. A. López-Riquelme, N. Pavón-Pulido, H. Navarro-Hellín, F. Soto-Valles, and R. Torres-Sánchez, "A software architecture based on FIWARE cloud for precision agriculture," *Agric. Water Manage*, vol. 183, pp. 123-135, 2017.

[51] A. Sharifi, "Remote sensing and decision support systems," in *Spatial Statistics for Remote Sensing*, A. Stein, F. Van der Meer, and B. Gorte, Eds. Remote Sensing and Digital Image Processing, Dordrecht: Springer, vol. 1, 1999.

[52] W. Ping, I. R. Farah, and A. Hussain, "A novel decision support system for the interpretation of remote sensing big data," *Earth Sci. Inform.*, vol. 11, pp. 31-45, 2018.

[53] W. Ping, J. Fu, W. Qiao, M. Yasir, S. Hui, M. S. Hossain, and S. Nazir, "Decision support system for hyperspectral remote-sensing data of yellow river estuary, China," *Sci. Program.*, vol. 2021, no. 1376167, 2021.

[54] B. J. Sowmya, A. Singh, Z. Khan, B. Sathvik, S. Seema, and K. G. Srinivasa, "A utility system for farmers to build decision support system on agrometeorological data using machine learning algorithms," in *Proc. 3rd International Conference on Computing Informatics and Networks. Lecture Notes in Networks and Systems*, A. Abraham, O. Castillo, D. Virmani, Eds. Singapore: Springer, 2021.

[55] M. W. P. Maduranga, and R. Abeysekera, "Machine learning applications in IoT based agriculture and smart farming: A review," *Int. J. Eng. Appl. Sci. Technol.*, vol. 4, no. 12, pp. 24-27, 2020.

[56] M. Singh, P. Singh, and S. B. Singh, "Decision support system for farm management," *World Academy of Science, Engineering and Technology*, vol. 39, pp. 346-349, 2008.

[57] M. Debeljak, A. Trajanov, V. Kuzmanovski, J. Schröder, T. Sandén, H. Spiegel *et al.*, "A field-scale decision support system for assessment and management of soil functions," *Front. Environ. Sci.*, vol. 7, no. 115, 2019.

[58] P. Ajewole, and F. Fayose, "Decision support system approach to farm machinery management in Nigeria," in *Proc. Pasae-Niae 2021 Virtual International Conference*, 2021.

[59] S. Fountas, C. G. Sorensen, Z. Tsiropoulos, C. Cavalaris, V. Liakos, and T. Gemtos, "Farm machinery management information system,"

Comput. Electron. Agric., vol. 110, pp. 131-138, 2015.

[60] R. G. Evans, and E. J. Sadler, "Methods and technologies to improve efficiency of water use," *Water Resour. Res.*, vol. 44, no. 7, 2008.

[61] Watson Decision Platform for Agriculture. [Online]. Available at: https://www.ibm.com/downloads/cas/ONVXEB2A

[62] L. Burketova, L. Trda, P. G. Ott, and O. Valentova, "Bio-based resistance inducers for sustainable plant protection against pathogens," *Biotechnol. Adv.*, vol. 33, no. 6, pp. 994-1004, 2015.

[63] N. Ananthi, J. Divya, M. Divya, and V. Janani, "IoT based smart soil monitoring system for agricultural production," in *IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, 2017, pp. 209-214.

[64] R. N. Rao, and B. Sridhar, "IoT based smart crop-field monitoring and automation irrigation system," in *2nd International Conference on Inventive Systems and Control (ICISC)*, 2018, pp. 478-483.

[65] G. Fernandez, K. Vijayakumar, V. Kalyanasundaram, S. Vidyasagar, K. S. D. Selvabharathi, and R. Palanisamy, "Smart soil monitoring and water conservation using irrigation on technology," *Indo. J. Elec. Eng. Comput. Sci.*, vol. 19, no. 1, pp. 99-107, 2020.

[66] All India Network Project on Biofertilizers (aicrp BNF). [Online]. Available: biofertilizers – aicrp-biofertilizers (icar.gov.in)

[67] https://www.iffco.in/en/organic-and-bio-fertilisers

[68] e-Urvarak | National Informatics Centre. [Online]. Available: www.nic.in

[69] S. D. Khirade, and A. B. Patil, "Plant disease detection using image processing," in *International Conference on Computing Communication Control and Automation*, 2015, pp. 768-771.

[70] F. Martinelli, R. Scalenghe, S. Davino *et al.*, "Advanced methods of plant disease detection - A review," *Agron. Sustain. Dev.*, vol. 35, pp. 1-25, 2015.

[71] I. Buja, E. Sabella, A. G. Monteduro, M. S. Chiriacò, L. De Bellis, A. Luvisi, and G. Maruccio, "Advances in plant disease detection and monitoring: From traditional assays to in-field diagnostics," *Sensors (Basel)*, vol. 21, no. 6, p. 2129, 2021.

[72] A. Mohammad-Razdari, D. Rousseau, A. Bakhshipour, S. Taylor, J. Poveda, and H. Kiani, "Recent advances in e-monitoring of plant diseases," *Biosens. Bioelectron.*, vol. 201, p. 113953, 2022.

[73] A. Musa, M. Hamada, F. M. Aliyu, and M. Hassan, "An intelligent plant disease detection system for smart hydroponic using convolutional neural network," in *Proc. IEEE 14th International Symposium on Embedded Multicore/Many-core Systems-on-Chip (MCSoC)*, IEEE, 2021, pp. 345-351.

[74] K. Indumathi, R. Hemalatha, S. A. Nandhini, and S. Radha, "Intelligent plant disease detection system using wireless multimedia sensor networks," in *Proc. International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, IEEE, 2017, pp. 1607-1611.

[75] A. Haridasan, J. Thomas, and E. D. Raj, "Deep learning system for paddy plant disease detection and classification," *Environ. Monit. Assess.*, vol. 195, no. 120, 2023.

[76] J. Liu, and X. Wang, "Plant diseases and pests detection based on deep learning: A review," *Plant Methods*, vol. 17, no. 22, 2021.

[77] N. Radha, and R. Swathika, "A polyhouse: Plant monitoring and diseases detection using CNN," in *Proc. International Conference on Artificial Intelligence and Smart Systems (ICAIS)*, 2021, pp. 966-971.

[78] J. Zhang, Y. Huang, R. Pu, P. Gonzalez-Moreno, L. Yuan, K. Wu, and W. Huang, "Monitoring plant diseases and pests through remote sensing technology: A review," *Comput. Electron. Agric.*, vol. 165, p. 104943, 2019.

[79] Y. Fang, and R. P. Ramasamy, "Current and prospective methods for plant disease

detection," *Biosensors (Basel)*, vol. 5, no. 3, pp. 537-561, 2015.

[80] S. Augustin, N. Boonham, W. J. D. Kogel, P. Donner, M. Faccoli, D. C. Lees, and A. Battisti, "A review of pest surveillance techniques for detecting quarantine pests in Europe," *EPPO Bull.*, vol. 42, no. 3, pp. 515-551, 2012.

[81] T. Liu, W. Chen, W. Wu, C. Sun, W. Guo, and X. Zhu. "Detection of aphids in wheat fields using a computer vision technique," *Biosyst. Eng.*, vol. 141, pp. 82-93, 2016.

[82] C. J. Chen, Y. Y. Huang, Y. S. Li, C. Y. Chang, and Y. M. Huang, "An AIoT based smart agricultural system for pests detection," *IEEE Access*, vol. 8, pp. 180750-180761, 2020.

[83] H. Nagar, and R. S. Sharma, "A comprehensive survey on pest detection techniques using image processing," in *4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, IEEE, 2020, pp. 43-48.

[84] M. A. Ebrahimi, M. H. Khoshtaghaza, S. Minaei, and B. Jamshidi, "Vison-based pest detection based on SVM classification method," *Comput. Electron. Agric.*, vol. 137, pp. 52-58, 2017.

[85] G. Pattnaik, and K. Parvathi, "A review on advanced techniques on plant pest detection and classification," *Smart Intell. Comput. Appl.*, pp. 665-673, 2020.

[86] L. Heeb, E. Jenner, and M. J. Cock, "Climate-smart pest management: Building resilience of farms and landscapes to changing pest threats," *J. Pest Sci.*, vol. 92, no. 3, pp. 951-969, 2019.

[87] C. Frezal, and G. Garsous, "New digital technologies to tackle trade in illegal pesticides," *OECD Trade and Environment Working Papers*, no. 2020/02, Paris: OECD Publishing, 2020.

[88] J. Kaur, K. P. Sahu, and S. Singh, "Optimization of pest management using artificial intelligence: Fundamentals and applications," *Souvenir & Abstracts*, vol. 11, n.d.

[89] J. Kreuze, J. Adewopo, M. Selvaraj, L. Mwanzia, P. Kumar, W. J. Cuellar, and G. Blomme, "Innovative digital technologies to monitor and control pest and disease threats in Root, Tuber, and Banana (RT&B) cropping systems: Progress and prospects," *Root, Tuber and Banana Food System Innovations*, pp. 261-288, 2022.

[90] S. Azfar, A. Nadeem, and A. Basit, "Pest detection and control techniques using wireless sensor network: A review," *Journal of Entomology and Zoology Studies*, vol. 3, no. 2, pp. 92-99, 2015.

[91] J. G. A. Barbedo, "Detecting and classifying pests in crops using proximal images and machine learning: A review," *AI*, vol. 1, no. 2, pp. 312-328, 2020.

[92] F. Awuor, S. Otanga, V. Kimeli, D. Rambim, and T. Abuya, "e-pest surveillance: Large scale crop pest surveillance and control," in *Proc. IST-Africa Week Conference (IST-Africa)*, IEEE, 2019, pp. 1-8.

[93] J. M. Jez, C.N. Topp, G. Silva, J. Tomlinson, N. Onkokesung, S. Sommer, and N. Boonham, "Plant pest surveillance: From satellites to molecules," *Emerging Top. Life Sci.*, vol. 5, no. 2, pp. 275-287, 2021.

[94] Z. Basati, B. Jamshidi, M. Rasekh, and Y. Abbaspour-Gilandeh, "Detection of sunn pest-damaged wheat samples using visible/ near-infrared spectroscopy based on pattern recognition," *Spectrochim. Acta, Part A*, vol. 203, pp. 308-314, 2018.

[95] Y. Zhao, Y. He, and X. Xu, "A novel algorithm for damage recognition on pest-infested oilseed rape leaves," *Comput. Electron. Agric.*, vol. 89, pp. 41-50, 2012.

[96] D. Mondal, K. Roy, D. Pal, and D. K. Kole, "Deep learning-based approach to detect and classify signs of crop leaf diseases and pest damage," *SN Comput. Sci.*, vol. 3, no. 6, pp. 1-12, 2022.

[97] C. Nandhini, and M. Brindha, "Deep learning solutions for pest detection," in *Object*

Detection with Deep Learning Models, Chapman and Hall/CRC, pp. 179-198, 2022.

[98] M. A. Ebrahimi, M. H. Khoshtaghaza, S. Minaei, and B. Jamshidi, "Vision-based pest detection based on SVM classification method," *Comput. Electron. Agric.*, vol. 137, pp. 52-58, 2017.

[99] K. A. Vakilian, and J. Massah, "Performance evaluation of a machine vision system for insect pests identification of field crops using artificial neural networks," *Arch. Phytopathol. Plant Prot.*, vol. 46, no. 11, pp. 1262-1269, 2013.

[100] Z. Qiang, and F. Shi, "Pest disease detection of Brassica chinensis in wide scenes via machine vision: Method and deployment," *J. Plant Dis. Prot.*, vol. 129, no. 3, pp. 533-544, 2022.

[101] E. Önler, "Real time pest detection using YOLOv5," *Int. J. Agric. Nat. Sci.*, vol. 14, no. 3, pp. 232-246, 2021.

[102] D. C. Corrales, J. C. Corrales, and A. Figueroa-Casas, "Towards detecting crop diseases and pest by supervised learning," *Ingenieríay Universidad*, vol. 19, no. 1, pp. 207-228, 2015.

[103] R. Pratheba, A. Sivasangari, and D. Saraswady, "Performance analysis of pest detection for agricultural field using clustering techniques," in *Proc. International Conference on Circuits, Power and Computing Technologies(ICCPCT-2014)*, IEEE, 2014, pp. 1426-1431.

[104] J. W. Chen, W. J. Lin, H. J. Cheng, C. L. Hung, C. Y. Lin, and S. P. Chen, "A smartphone-based application for scale pest detection using multiple-object detection methods," *Electronics*, vol. 10, no. 4, p. 372, 2021.

[105] H. Rware, M. K. Kansiime, I. Mugambi, D. Onyango, J. A. Tambo, C. M. Banda, and J. Godwin, "Is radio an effective method for delivering actionable information for responding to emerging pest threats? A case study of fall armyworm campaign in Zambia," *CABI Agric. Biosci.*, vol. 2, no. 1, pp. 1-11, 2021.

[106] L. Nanni, G. Maguolo, and F. Pancino, "Insect pest image detection and recognition based on bio-inspired methods," *Ecol. Inf.*, vol. 57, no. 101089, 2020.

[107] Z. Zheng, and C. Zhang, "Electronic noses based on metal oxide semiconductor sensors for detecting crop diseases and insect pests," *Comput. Electron. Agri.*, vol. 197, p. 106988, 2022.

[108] W. G. Henderson, A. Khalilian, Y. J. Han, J. K. Greene, and D. C. Degenhardt, "Detecting stink bugs/damage in cotton utilizing a portable electronic nose," *Comput. Electron. Agri.*, vol. 70, no. 1, pp. 157-162, 2010.

[109] S. Miresmailli, D. Badulescu, M. Mahdaviani, R. H. Zamar, and M. B. Isman, "Integrating plant chemical ecology, sensors and artificial intelligence for accurate pest monitoring," *Tomatoes: Agriculture Procedures Pathogen*, 2009.

[110] B. D. Lampson, "Detection and identification of agricultural pests and pest damage using novel sensing technologies," Doctoral dissertation, Clemson University, 2012.

[111] H. Abid, N. Nida, and A. Irtaza, "PestinaNet - A real-time crop pest detection system," in *Proc. 2nd International Conference on Computing and Machine Intelligence (ICMI)*, IEEE, 2022, pp. 1-4.

[112] C. H. Bock, G. H. Poole, P. E. Parker, and T. R. Gottwald, "Plant disease severity estimated visually, by digital photography and image analysis, and by hyperspectral imaging," *Crit. Rev. Plant Sci.*, vol. 29, no. 2, pp. 59-107, 2010.

[113] M. C. Ferreira Lima, M. E. Damascena de Almeida Leandro, C. Valero, L. C. Pereira Coronel, and C. O. Gonçalves Bazzo, "Automatic detection and monitoring of insect pests - A review," *Agriculture*, vol. 10, no. 5, p. 161, 2020.

[114] L. C. Ngugi, M. Abelwahab, and M. Abo-Zahhad, "Recent advances in image processing techniques for automated leaf pest and disease

recognition - A review," *Info. Process. Agri.*, vol, 8, no. 1, pp. 27-51, 2021.

[115] K. Kusrini, A. Setyanto, I. M. A. Agastya, S. Pariyasto, W. W. Widiyanto, and H. Priantoro. "Automatic mango leaf and trunk detection as supporting tool of mango pest identifier (MPI)," in *Proc. 3rd International Conference on Cybernetics and Intelligent System (ICORIS)*, IEEE, 2021, pp. 1-7.

[116] W. Tan, C. Zhao, and H. Wu, "Intelligent alerting for fruit-melon lesion image based on momentum deep learning," *Multimedia Tools Appl.*, vol. 75, no. 24, pp. 16741-16761, 2016.

[117] W. Dawei, D. Limiao, N. Jiangong, G. Jiyue, Z. Hongfei, and H. Zhongzhi, "Recognition pest by image-based transfer learning," *J. Sci. Food Agric.*, vol. 99, no. 10, pp. 4524-4531, 2019.

[118] J. Chen, W. Chen, A. Zeb, A. D. Zhang, and Y. A. Nanehkaran, "Crop pest recognition using attention-embedded lightweight network under field conditions," *Appl. Entomol. Zool.*, vol. 56, no. 4, pp. 427-442, 2021.

[119] M. G. Selvaraj, A. Vergara, H. Ruiz, N. Safari, S. Elayabalan, W. Ocimati, and G. Blomme, "AI-powered banana diseases and pest detection," *Plant Methods*, vol. 15, no. 1, pp. 1-11, 2019.

[120] P. K. Sethy, C. Bhoi, N. K. Barpanda, S. Panda, S. K. Behera, and A. K. Rath, "Pest detection and recognition in rice crop using SVM in approach of bag-of-words," in *International Conference on Software and System Processes*, 2017.

[121] F. Faithpraise, P. Birch, R. Young, J. Obu, B. Faithpraise, and C. Chatwin, "Automatic plant pest detection and recognition using k-means clustering algorithm and correspondence filters," *Int. J. Adv. Biotech. Res.*, vol. 4, no. 2, pp. 189-199, 2013.

[122] T. A. S. Perera, and J. Collins, "Novel embedded system based species recognition system for pest control," *Int. J. Comput. Digital Syst.*, vol. 5, no. 5, 2016.

[123] C. J. Chen, J. S. Wu, C. Y. Chang, and Y. M. Huang, "Agricultural pests damage detection using deep learning," in *Proc. International Conference on Network-Based Information Systems*, Springer, Cham, 2019, pp. 545-554.

[124] J. Zhang, R. Ji, X. Yuan, H. Li, and L. Qi, "Recognition of pest damage for cotton leaf based on RBF-SVM algorithm," *Trans. Chin. Soc. Agric. Mach.*, vol. 42, no. 8, pp. 178-183, 2011.

[125] S. Cui, P. Ling, H. Zhu, and H. M. Keener, "Plant pest detection using an artificial nose system: a review," *Sensors*, vol. 18, no. 2, p. 378, 2018.

[126] B. D. Lampson, Y. J. Han, A. Khalilian, J. K. Greene, D. C. Degenhardt, and J. O. Hallstrom, "Development of a portable electronic nose for detection of pests and plant damage," *Comput. Electron. Agric.*, vol. 108, pp. 87-94, 2014.

[127] P. Gimbarzevsky, "Remote sensing in forest damage detection and appraisal: Selected annotated bibliography," *Pacific Forest Research Centre, Canadian Forestry Service*, 1984.

[128] A. L. P. de Ocampo, and E. P. Dadios, "Integrated weed estimation and pest damage detection in Solanum melongena plantation via aerial vision-based proximal sensing," *Philipp. J. Sci.*, vol. 150, no. 4, pp. 677-688, 2021.

[129] C. K. Panda, "Advances in application of ICT in crop pest and disease management," in *Natural Remedies for Pest, Disease and Weed Control*. Academic Press, 2020, pp. 235-242.

[130] G. Norton, and M. Taylor, "What pest is that? Recent developments in digital pest diagnostics," *Outlooks Pest Manage.*, vol. 21, no. 5, pp. 236-238, 2010.

[131] W. M. Getz, and A. P. Gutierrez, "A perspective on systems analysis in crop production and insect pest management," *Annu. Rev. Entomol.*, vol. 27, no. 1, pp. 447-466, 1982.

[132] W. Li, T. Zheng, Z. Yang, M. Li, C. Sun, and X. Yang, "Classification and detection of insects

from field images using deep learning for smart pest management: A systematic review," *Ecol. Inf.*, vol. 66, no. 101460, 2021.

[133] W. Ding, and G. Taylor, "Automatic moth detection from trap images for pest management," *Comput. Electron. Agric.*, vol. 123, pp. 17-28, 2016.

[134] D. L. Haynes, R. K. Brandenburg, and P. D. Fisher, "Environmental monitoring network for pest management systems," *Environ. Entomol.*, vol. 2, no. 5, pp. 889-900, 1973.

[135] F. H. I. Filho, W. B. Heldens, Z. Kong, and E. S. de Lange, "Drones: Innovative technology for use in precision pest management," *J. Econ. Entomol.*, vol. 113, no. 1, pp. 1-25, 2020.

[136] R. M. Strickland, D. R. Ess, and S. D. Parsons, "Precision farming and precision pest management: The power of new crop production technologies," *J. Nematol.*, vol. 30, no. 4, p. 431, 1998.

[137] S. J. Thomson, and D. L. Sudbrink, "Digital imaging from agricultural aircraft: System configurations and constraints for integrated pest management, weed detection, and determination of crop status," in *Proc. 18th Biennial Workshop on Color Photography and Videography in Resource Assessment*, 2004.

[138] Y. Huang, Y. Lan, and W. C. Hoffmann, "Use of airborne multi-spectral imagery in pest management systems," *Agricultural Engineering International: CIGR Journal*, vol. 1, 2008.

[139] Y. Koumpouros, B. D. Mahaman, M. Maliappis, H. C. Passam, A. B. Sideridis, and V. Zorkadis, "Image processing for distance diagnosis in pest management," *Comput. Electron. Agric.*, vol. 44, no. 2, pp. 121-131, 2004.

[140] Y. He, H. Zeng, Y. Fan, S. Ji, and J. Wu, "Application of deep learning in integrated pest management: A real-time system for detection and diagnosis of oilseed rape pests," *Mobile Inf. Syst.*, vol. 1, 2019.

[141] A. A. Beckler, B. Wade French, and L. D. Chandler, "Using GIS in areawide pest management: A case study in South Dakota," *Trans. GIS*, vol. 9, no. 2, pp. 109-127, 2005.

[142] P. Damos, "Modular structure of web-based decision support systems for integrated pest management - A review," *Agron. Sustainable Dev.*, vol. 35, no. 4, pp. 1347-1372, 2015.

[143] S. Sankaran, A. Mishra, R. Ehsani, and C. Davis, "A review of advanced techniques for detecting plant diseases," *Computers and Electronics in Agriculture*, vol. 72, no. 1, pp. 1-13, 2010.

[144] A. K. Mahlein, E. C. Oerke, U. Steiner, and H. W. Dehne, "Recent advances in sensing plant diseases for precision crop protection," *Eur. J. Plant Pathol.*, vol. 133, no. 1, pp. 197-209, 2012.

[145] J. G. Arnal Barbedo, "Digital image processing techniques for detecting, quantifying and classifying plant diseases," *SpringerPlus*, vol. 2, no. 1, pp. 1-12, 2013.

[146] S. Kumar, and R. Kaur, "Plant disease detection using image processing - A review," *Int. J. Comput. Appl.*, vol. 124, no. 16, 2015.

[147] V. Pooja, R. Das, and V. Kanchana, "Identification of plant leaf diseases using image processing techniques," in *Proc. IEEE Technological Innovations in ICT for Agriculture and Rural Development (TIAR)*, IEEE, 2017, pp. 130-133.

[148] P. Trematerra, "Aspects related to decision support tools and integrated pest management in food chains," *Food Control*, vol. 34, no. 2, pp. 733-742, 2013.

[149] G. G. Hungilo, G. Emmanuel, and A. W. Emanuel, "Image processing techniques for detecting and classification of plant disease: A review," in *Proc. International Conference on Intelligent Medicine and Image Processing*, 2019, pp. 48-52.

[150] N. Gobalakrishnan, K. Pradeep, C. J. Raman, L. J. Ali, and M. P. Gopinath, "A systematic review on image processing and machine

learning techniques for detecting plant diseases," in *Proc. International Conference on Communication and Signal Processing (ICCSP)*, IEEE, 2020, pp. 465-468.

[151] S. P. Mohanty, D. P. Hughes, and M. Salathé, "Using deep learning for image-based plant disease detection," *Front. Plant Sci.*, vol. 7, no. 1419, 2016.

[152] S. K. Tichkule, and D. H. Gawali, "Plant diseases detection using image processing techniques," in *Online International Conference on Green Engineering and Technologies (IC-GET)*, IEEE, 2016, pp. 1-6.

[153] C. A. Berdugo, R. Zito, S. Paulus, and A. K. Mahlein, "Fusion of sensor data for the detection and differentiation of plant diseases in cucumber," *Plant Pathol.*, vol. 63, no. 6, pp. 1344-1356, 2014.

[154] Y. Ampatzidis, L. De Bellis, and A. Luvisi, "iPathology: Robotic applications and management of plants and plant diseases," *Sustainability*, vol. 9, no. 6, p. 1010, 2017.

[155] K. P. Ferentinos, "Deep learning models for plant disease detection and diagnosis," *Comput. Electron. Agric.*, vol. 145, pp. 311-318, 2018.

[156] J. G. Barbedo, "Factors influencing the use of deep learning for plant disease recognition," *Biosyst. Eng.*, vol. 172, pp. 84-91, 2018.

[157] J. Chen, D. Zhang, and Y. A. Nanehkaran, "Identifying plant diseases using deep transfer learning and enhanced lightweight network," *Multimedia Tools Appl.*, vol. 79, no. 41, pp. 31497-31515, 2020.

[158] M. Loey, A. ElSawy, and M. Afify, "Deep learning in plant diseases detection for agricultural crops: A survey," *Int. J. Serv. Sci. Manage. Eng. Technol.* (IJSSMET), vol. 11, no. 2, pp. 41-58, 2020.

[159] S. Sladojevic, M. Arsenovic, A. Anderla, D. Culibrk, and D. Stefanovic, "Deep neural networks based recognition of plant diseases by leaf image classification," *Comput. Intell. Neurosci.*, 2016.

[160] S. Hernández, and J. L. López, "Uncertainty quantification for plant disease detection using Bayesian deep learning," *Appl. Soft Comput.*, vol. 96, p. 106597, 2020.

[161] M. Afify, M. Loey, and A. Elsawy, "A robust intelligent system for detecting tomato crop diseases using deep learning," *Int. J. Software Sci. Comput. Intell. (IJSSCI)*, vol. 14, no. 1, pp. 1-21, 2022.

[162] S. Thomas, M. T. Kuska, D. Bohnenkamp, A. Brugger, E. Alisaac, M. Wahabzada, and A. K. Mahlein, "Benefits of hyperspectral imaging for plant disease detection and plant protection: A technical perspective," *J. Plant Dis. Prot.*, vol. 125, no. 1, pp. 5-20, 2018.

[163] A. Ahmad, D. Saraswat, and A. El Gamal, "A survey on using deep learning techniques for plant disease diagnosis and recommendations for development of appropriate tools," *Smart Agric. Technol.*, p. 100083, 2022.

[164] L. Wan, H. Li, C. Li, A. Wang, Y. Yang, and P. Wang, "Hyperspectral sensing of plant diseases: Principle and methods," *Agronomy*, vol. 12, no. 6, p. 1451, 2022.

[165] T. Singh, D. Singh, and S. S. Bedi, "Monitoring and detecting plant diseases using cloud-based internet of things," in *Integration and Implementation of the Internet of Things Through Cloud Computing*. IGI Global, 2021, pp. 217-235.

[166] A. Khattab, S. E. Habib, H. Ismail, S. Zayan, Y. Fahmy, and M. M. Khairy, "An IoT-based cognitive monitoring system for early plant disease forecast," *Comput. Electron. Agric.*, vol. 166, no. 105028, 2019.

[167] J. Boulent, S. Foucher, J. Théau, and P. L. St-Charles, "Convolutional neural networks for the automatic identification of plant diseases," *Front. Plant Sci.*, vol. 10, p. 941, 2019.

[168] A. Darwish, D. Ezzat, and A. E. Hassanien, "An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis," *Swarm Evol. Comput.*, vol. 52, p. 100616, 2020.

[169] X. Yang, and T. Guo, "Machine learning in plant disease research," *European Journal of BioMedical Research*, vol. 3, no. 1, p. 6, 2017.

[170] T. Rumpf, A. K. Mahlein, U. Steiner, E. C. Oerke, H. W. Dehne, and L. Plümer, "Early detection and classification of plant diseases with support vector machines based on hyperspectral reflectance," *Comput. Electron. Agric.*, vol. 74, no. 1, pp. 91-99, 2010.

[171] R. Kaur, and S. S. Kang, "An enhancement in classifier support vector machine to improve plant disease detection," in *Proc. IEEE 3rd International Conference on MOOCs, Innovation and Technology in Education (MITE)*, IEEE, Oct. 2015, pp. 135-140.

[172] D. S. Tan, R. N. Leong, A. F. Laguna, C. A. Ngo, A. Lao, D. M. Amalin, and D. G. lvindia, "AuToDiDAC: Automated tool for disease detection and assessment for cacao black pod rot," *Crop Prot.*, vol. 103, pp. 98-102, 2018.

[173] J. Quinn, "Computational techniques for crop disease monitoring in the developing world," in *Proc. International Symposium on Intelligent Data Analysis*, Springer, Berlin, Heidelberg, 2013, pp. 13-18.

[174] A. Rehman, T. Saba, M. Kashif, S. M. Fati, S. A. Bahaj, and H. Chaudhry, "A revisit of internet of things technologies for monitoring and control strategies in smart agriculture," *Agronomy*, vol. 12, no. 1, p. 127, 2022.

[175] Y. Shi, Z. Wang, X. Wang, and S. Zhang, "Internet of things application to monitoring plant disease and insect pests," in *Proc. International Conference on Applied Science and Engineering Innovation*, Atlantis Press, 2017, pp. 31-34.

[176] G. Nagasubramanian, R. K. Sakthivel, R. Patan, M. Sankayya, M. Daneshm, and A. H. Gandomi, "Ensemble classification and IoT-based pattern recognition for crop disease monitoring system," *IEEE Internet Things J.*, vol. 8, no. 16, pp. 12847-12854, 2021.

[177] Shi, "Monitoring and discrimination of plant disease and insect pests based on agricultural IoT," in *Proc. 4th International Conference on Information Technology and Management Innovation*, Atlantis Press, 2015, pp. 112-115.

[178] D. Gao, Q. Sun, B, Hu, and S. Zhang, "A framework for agricultural pest and disease monitoring based on internet-of-things and unmanned aerial vehicles," *Sensors*, vol. 20, no. 5, p. 1487, 2020.

[179] E. Ivliev, V. Demchenko, and P. Obukhov, "Automatic monitoring of smart greenhouse parameters and detection of plant diseases by neural networks," in *Robotics, Machinery and Engineering Technology for Precision Agriculture*, Springer, Singapore, 2022, pp. 29-36.

[180] H. Rizk, and M. K. Habib, "Robotized early plant health monitoring system," in *Proc. IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society*, IEEE, 2018, pp. 3795-3800.

[181] K. Neupane, and F. Baysal-Gurel, "Automatic identification and monitoring of plant diseases using unmanned aerial vehicles: A review," *Remote Sens.*, vol. 13, no. 19, p. 3841, 2021.

[182] S. Zhang, W. Huang, and H. Wang, "Crop disease monitoring and recognizing system by soft computing and image processing models," *Multimedia Tools and Appl.*, vol. 79, no. 41, pp. 30905-30916, 2020.

[183] D. A. Ivanov, and M. A. Bernards, "Chlorophyll fluorescence imaging as a tool to monitor the progress of a root pathogen in a perennial plant," *Planta*, vol. 243, no. 1, pp. 263-279, 2016.

[184] L. Yuan, Z. Bao, H. Zhang, Y. Zhang, and X. Liang, "Habitat monitoring to evaluate crop disease and pest distributions based on multi-source satellite remote sensing imagery," *Optik*, vol. 145, pp. 66-73, 2017.

[185] M. C. Mothapo, T. Dube, E. Abdel-Rahman, and M. Sibanda, "Progress in the use of geospatial and remote sensing technologies in the assessment and monitoring of tomato crop diseases," *Geocarto Int.*, vol. 37, no. 16, pp. 4784-4804, 2022.

[186] J. Zhang, L. Yuan, J. Wang, J. Luo, S. Du, and W. Huang, "Research progress of crop diseases and pests monitoring based on remote sensing," *Trans. Chin. Soc. Agric. Eng.*, vol. 28, no. 20, pp. 1-11, 2012.

[187] M. Brahimi, M. Arsenovic, S. Laraba, S. Sladojevic, K. Boukhalfa, and A. Moussaoui, "Deep learning for plant diseases: Detection and saliency map visualisation," in *Human and Machine Learning*, Springer, Cham, 2018, pp. 93-117.

[188] X. Wu, C. Wang, Z. Xu, and Y. Hu, "Application of GIS technology in monitoring and warning system for crop diseases and insect pests," in *Proc. Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Monitoring and Assessment of Natural Resources and Environments*, SPIE, 2018, vol. 7145, pp. 391-399.

[189] A. Gamal, G. I. Sayed, A. Darwish, and A. E. Hassanien, "A new proposed model for plant diseases monitoring based on data mining techniques," in *Plant Bioinformatics*, Springer, Cham, 2017, pp. 179-195.

[190] F. Nabi, S. Jamwal, and K. Padmanbh, "Wireless sensor network in precision farming for forecasting and monitoring of apple disease: A survey," *Int. J. Inf. Technol.*, pp. 1-12, 2020.

[191] Z. Li, Y. Liu, O. Hossain, R. Paul, S. Yao, S. Wu, and Q. Wei, "Real-time monitoring of plant stresses via chemiresistive profiling of leaf volatiles by a wearable sensor," *Matter*, vol. 4, no. 7, pp. 2553-2570, 2021.

[192] C. Von Kröcher, and M. Röhrig, "Monitoring of plant pests and diseases as a base of the Germany-wide online decision support system ISIP," *J. Für Verbraucherschutz und Lebensmittelsicherheit*, vol. 2, no. 1, pp. 50-51, 2007.

[193] H. Zhang, H. Hu, and H. Zhang, "Monitoring methods of crop diseases and pests based on hyperspectral technology," *Int. J. Simul. Syst. Sci. Technol.*, vol. 17, no. 11, p. 2, 2016.

[194] A. Y. Dong, Z. Wang, J. J. Huang, B. A. Song, and G. F. Hao, "Bioinformatic tools support decision-making in plant disease management," *Trends Plant Sci.*, vol. 26, no. 9, pp. 953-967, 2021.

[195] K. Alemu, "The role and application of bioinformatics in plant disease management," *Adv. Life Sci. Technol.*, vol. 28, pp. 28-33, 2015.

[196] J. W. Travis, and R. X. Latin, "Development, implementation, and adoption of expert systems in plant pathology," *Annu. Rev. Phytopathol.*, vol. 29, no. 1, pp. 343-360, 1991.

[197] F. X. R. Do Vale, E. I. F. Filho, J. R. Liberto, and L. Zambolim, "Quant - A software to quantify plant disease severity," in *International Workshop on Plant Disease Epidemiology*, vol. 1, p. 161, 2001.

[198] K. P. Singh, T. Aravind, A. K. Srivastava, and C. S. Karibasappa, "Decision-making tools for integrated disease management," in *Emerging Trends in Plant Pathology*, Springer, Singapore, 2021, pp. 703-724.

[199] J. Lakhani, and D. Harwani, "Deep convolution neural network based prediction and classification of blight disease of potato," *J. Phytol. Res.*, vol. 34, no. 2, 2021.

[200] B. S. Dhillon, and T. S. Thind, "Information and communication technology for effective integrated pest management," *Curr. Sci*., vol. 110, no. 3, p. 287, 2016.

[201] R. Alguliyev, Y. Imamverdiyev, L. Sukhostat, and R. Bayramov, "Plant disease detection based on a deep model," *Soft Comput.*, vol. 25, no. 21, pp. 13229-13242, 2021.

[202] M. K. Gayatri, J. Jayasakthi, and G. S. Mala, "Providing smart agricultural solutions to farmers for better yielding using IoT," in *IEEE Technological Innovation in ICT for Agriculture and Rural Development (TIAR)*, IEEE, 2015, pp. 40-43.

[203] S. Sadowski, and P. Spachos, "Wireless technologies for smart agricultural monitoring

using internet of things devices with energy harvesting capabilities," *Comput. Electron. Agric.*, vol. 172, p. 105338, 2020.

[204] M. Ashifuddin Mondal, and Z. Rehena, "IoT based intelligent agriculture field monitoring system," in *Proc. 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, IEEE, 2018, pp. 625-629

[205] M. Sachidananda, S. Rahimifard, and D. P. Webb, "A decision tool for improving manufacturing water usage efficiency," *Int.J. Comput. Integr. Manuf.*, vol. 33, no. 9, pp.823-839, 2020.

[206] A. J. Churi, M. R. Mlozi, H. Mahoo, S. D. Tumbo, and R. Casmir, "A decision support system for enhancing crop productivity of smallholder farmers in semi-arid agriculture," *Int. J. Inf.*, vol. 3, no. 8, 2013.

International Journal of Knowledge Based Computer Systems
13 (1) 2025, 31-37
www.publishingindia.com//ijkbcs/

Methodologies for Optimized Event and Incident Processing in Cybersecurity

Cheryl Ann Alexander[1*] and Lidong Wang[2]

[1]Institute for IT Innovation and Smart Health, Mississippi, USA.
Email: cannalexander68@gmail.com
[2]Institute for Systems Engineering Research, Mississippi State University, Mississippi, USA.
Email: lidong@iser.msstate.edu
[*]Corresponding Author

Abstract: Cybersecurity is an essential function of all enterprises. Threat mitigation is necessary to prevent cyber risks from costing the enterprise much more money and resources. The enterprise should establish a robust and vigorous cybersecurity program. Every incident is an event but not every incident should demand an urgent response. Much of the value of incident processing and incident reporting lies in the free-text section of the incident report. In healthcare, incident processing is a key function of the IT department. By determining whether an attack's detection is imminent, IT can utilize threat mitigation to reduce costs and resources associated with cybersecurity threats. Artificial intelligence, big data, and machine learning have been utilized in threat mitigation. In healthcare, this is an essential function of the IT department to reduce the likelihood of Ransomware attacks. This paper discusses event processing, incident processing, and the tools used to mitigate any attacks.

Keywords: Artificial intelligence (AI), Cybersecurity, Healthcare, Incident processing, Information, System, Threat mitigation.

I. Introduction

Every incident constitutes an event and selected events should not demand an urgent response. Information concerning patient safety incidents has been broadly identified by incident reporting systems. The free-text section of incident reports constitutes much of the value of incident reports. To address the gravity of incidents within an enterprise, researchers developed a decision-making scoring system that determined the severity of incidents using the semantic features of the text in incident reports comparing the results with the opinions of experts. Next, researchers calculated a severity term score, a severity report score, and a severity group score. However, it is necessary to research and investigate a suitable number of incident reports necessary to evaluate groups utilizing the severity report score, therefore, further research is necessary. The data volume necessary to decide the trends corresponding to the tendencies of the central limit theorem is the key to this approach [1].

A collection of automated tools such as vulnerability scanners, monitoring and logging tools, and antivirus software has been investigated. Decision trees for each category of tools were executed to assist in finding the way through huge amounts of information [2]. Table I [3] shows some threats (with a high probability of incident occurrences) during the COVID-19 pandemic.

TABLE I: SOME THREATS DURING COVID-19

Aspects	Threats
Actions on objectives	Financial fraud, data theft, personal information theft, password stealer, ransom, disturbance
Installation	Backdoor & persistence, Lokikbot Trojan, Trickbot Trojan, AZORult Info stealer, Ransomware Samas, GradCrab

Article can be accessed online at
www.publishingindia.com

Aspects	Threats
Exploitation	RDP (Remote Desktop Protocol) brute force, Drive-by compromise
Delivery	Email phishing, SMS (short message service) phishing, fake testing apps, attacks against the teleworking infrastructure, attacks against health organizations, fraudulent domains: Corona-virus-map.com; Apps: COVID-19 tracker

In healthcare, incident reporting is a key process for the information technology (IT) department and depends on many factors. Clinical staffs carry mobile devices in most cases, have access to facility computers including email, and have access to programs that could threaten the very livelihood of the medical center community. Having a robust IT reporting system for incidents is a crucial function of the IT department. Clinical staff, ancillary staff, and management staff should have access to most incident reporting systems. In case of a threat such as phishing emails, Ransomware, or other malware from suspicious or malicious actors, the staff should be taught to report so IT staff can minimize the damage and halt any threat to patient care disruption.

II. Event/Incident Processing, Attacks Detection, and Threat Mitigation

Event prediction in big data forces the discovery and assimilation of interconnected techniques that focus on challenges, comprised of the following: 1) heterogeneous multi-output estimates [4], 2) complex needs amid prediction outputs [5], 3) a real-time run of prediction jobs) [6], and 4) trials in the big data event. The taxonomy of event prediction and related practices are shown in Table II [7].

Table II: Event Prediction and Techniques

Categorization (Goals)	Further Categorization (Output Forms of Goals)	Techniques
Location prediction	Point-based	● Unsupervised: spatial scan, network scan ● Supervised: spatial autoregressive, spatial multi-task learning, ● geo-featured classification
	Raster-based	● Trajectory destination prediction ● Spatial clustering, embedding, and convolution
Time prediction	Continuous time	● Survival analysis ● Point process ● Regression
	Discrete-time	● Indirect manner ● Direct manner
	Occurrence	● Anomaly detection ● (Auto-) regression ● Classification
Semantic prediction	Semantic sequence	Step 1: Event representation Step 2: Event causality inference Step 3: Future event inference
	Associate-based	● Frequent set mining ● Decision list
	Causality-based	● Whole-sequence classification: model-based, feature-based, prototype-based ● Next-event generation: descriptive-based, attributed-based

Joint prediction	Location and time	● Point-based: spatiotemporal point process, spatial temporal Gaussian process ● Raster-based: RNN, 3D CNN, CNN+RNN, spatial temporal CRF
	Semantic and time	● Time expression extraction ● Temporal association rule ● Multi-variate time-series forecasting
	Location, time, and semantics	● Tensor-decomposition-based methods ● Crowd wisdom-based methods: crowdsource systems, model ensembles ● Future event expression identification

It is important to the facility for IT staff to monitor specific functions of a potential attack. Programs in the cloud and the network are highly susceptible to malicious actors and events likely to become threats. In-depth-detection of an incident or event may be accomplished by employing one or more of the subsequent detection methods [8]:

- Network or cloud intrusion detection.
- Behavior monitoring—using machine learning, this function is an advanced correlation engine, and/or behavioral biometrics. It can be accomplished using the following techniques:
  - Service and infrastructure monitoring.
  - Network protocol analysis.
  - Network flow analysis.
- Privilege escalation detection—used to detect privilege escalation and these techniques embrace, but are not constrained to the following:
  - Host intrusion detection.
  - File integrity monitoring.
  - Attempted unauthorized user access detection.
  - Monitoring SaaS Services such as Office 365 or G Suite.
- Event correlation—gathering data from application logs or host logs and analyzing it to identify relationships and can be achieved using the following techniques:
  - Security incident and event management (SIEM).
  - Malicious host communication detection.
  - Using a centralized dashboard that selects threats based on user inputs.
- Physical security.

Threat mitigation is crucial to all enterprises, especially in hospitals. Greater risks and damage will occur very quickly without mitigation. Mitigation is a complicated issue and is related to current resources, available techniques or tools, and individuals with experience and skills. Basic cyber threats should be evaluated as to their severity. The simple methods of cybersecurity threat mitigation contain, but are not limited by the subsequent techniques [8]:

- Risk assessment
- Access control system
- Intrusion detection system
- Video surveillance system
- Policies and procedures
- Automated device data wiping—this technique uses tools that activate routine data wiping of departing employees' devices.

When Security Information and Events Management (SIEM) ends, Security Orchestration, Automation, and Response (SOAR) systems pick up the incident response procedure delivering an automated and arranged response during the Identification Phase, as well as the Containment, Eradication, and Recovery Phases. SOAR is a critical component of cybersecurity threat mitigation when disparate tools are integrated within a common platform. SOAR is a product line category of the Security Operations and Automation Platform Architecture (SOAPA). An

SOAPA platform puts together technologies through data collection, processing, analytics, and security operations. Table III describes an SOAPA system [9].

TABLE III: A MULTI-FUNCTIONAL SOAPA SYSTEM

Functions and Components	Description and Details
Data services functions	● Historical data ● Database management system (DBMS) ● Security logs ● Data loss prevention ● On premises & cloud-based services/ data, network traffic, etc. ● Other (network analyzers, alerts/ sensor systems, vulnerability scanners, honeypots/probes)
Security operations functions	● SIEM ● SOAR ● Intrusion detection system (IDS) ● Intrusion prevention system (IPS) ● Extended detection and response (XDR) ● Threat intelligence platform (TIP) ● Unified threat management (UTM)
Analytic functions	● Analytic AI/ML models
Integration functions	● Communication services ● Message/data formatting ● Processing ● Transformations
User interface & experience (UIX)/ Management Station	● User interface (UI) ● Security Operations Center (SOC) analysts ● Proactive SOAPA dashboard (notifications, alerts/events, attack predictions, and reporting)

Because AI/ML does not rely on static signatures utilized in conventional anti-virus systems, it is perfect for malware detection and anti-virus [10] [11]. Useful for finding and classifying phishing and malware emails, artificial neural network (ANN)-based models can [12] attain full AI/ ML empowerment of SOAR, and reinforcement learning (RL) models can be forced to control security orchestration, automation, and response. Autonomous software agents making observations and taking sequential actions optimally, with or without incomplete previous knowledge of the operational environment, can also be RL and consequently be predominantly flexible for deployment in real-time and dynamic cybersecurity environments [13].

III. EVENT AND INCIDENT PROCESSING IN HEALTHCARE

Healthcare generally uses two kinds of event predictions: 1) individual-level for clinical longitudinal events and, 2) population-level disease outbreaks and epidemics. Extensive research on disease outbreaks for many different types of diseases and epidemics such as flu, H1N1, and COVID-19 has been conducted for population-level events [14]. Due to the rapid growth of massive surveillance data from government agencies, social media datasets, and other public datasets, a massive increase in the use of data-driven approaches to directly learn predictive mapping has developed [15]. Research has begun to focus on the individual level, including the longitudinal predictive analysis of individual health-related incidents such as adverse drug events [16] and sudden illnesses (e.g., strokes).

Incident reports can be submitted by physicians, providers, nurses, pharmacy technicians, biomedical equipment technicians, nursing assistants, administrative assistants, security guards, radiologic technologists, etc., depending on the specific situations of incidents, physicians, nurses, nursing assistants, radiologic technologists, biomedical equipment technicians, pharmacists, administrative assistants, security guards, etc. [1]. Yokohama City University Medical Center, Japan developed an incident reporting system. Anonymous, the incident reporting system included information about the incidents, like the professions, incident level, location, free-text content regarding the actual occurrence, and a measure to prevent future incidents. Currently, seven nationally defined incident levels exist in the system including Level 0—Incident with no direct impact on patients, Level 1—Incident with no substantial harm to patients, Level 2—Incident with only minor impairment to patient care requiring no further treatment, Level 3a—Substantial damage

to patients making it necessary to render further treatment to prevent a prolonged hospital stay, Level 3b—Significant damage to patient care models requiring further treatment including a prolonged hospital stay, Level 4—Permanent disability related to the accident, Level 5—Death related to accident. A Japanese hospital must report to the Japan Council for Quality Healthcare and publicly announce an incident that is Level 3b or above [17].

IV. Event and Incident Processing in a Large Medical Center

Charleston Regional Medical Center is a large medical center/hospital that serves patients in Mississippi, USA. In the Medical Center, events can be patient data theft, financial fraud, insurance fraud, password theft, personal health information theft, email phishing, etc. Ransomware occurrence, patient data theft, financial fraud, insurance fraud, personal health information theft, mobile device tampering, etc. are often treated as incidents. All incidents and events should be reported to the IT department, management, upper management, and risk management at first occurrence.

There are generally the following steps in the event/incident processing:

- *Event/Incident Detection:* At the first notice of a breach during patient care or management review of patient data or billing information.
- *Logging:* Need to be logged via a report in the IT department.
- *Categorization:* Categorized according to the level of incident severity and whether it will shut down patient care such as with Ransomware; for example, with Ransomware all patient care or billing is shut down by malicious actors with the intent to scam monetary reward for returning function. With a breach of patient data, this is categorized as a high priority because patient data is protected. Billing data is also protected.
- *Prioritization:* Low (e.g., email phishing); Medium (e.g., theft of passwords or medical equipment); and High (Ransomware, patient data theft, insurance fraud, billing fraud, etc.)
- *Response:* The police and IT departments need a detailed ongoing analysis and risk management program.
- *Escalation (Level 1, 2, or 3 Support):* Level 1 (such as email phishing and spam mail)—probably will not harm patient care; Level 2 (e.g., theft of medical equipment and mobile devices)—may shut down patient care; Level 3 (Ransomware, patient data theft, insurance fraud or billing fraud, etc.)—will shut down patient care.
- *Resolution/Recovery:* Will need to either pay the ransom or find a roundabout way to return patient care ability. Block email from personal accounts.

The Event and Incident Processing System is being implemented in the Medical Center to meet operational needs and improve cybersecurity. Annual training regarding event/incident processing and cybersecurity in healthcare is provided for all employees and contractors. The Event and Incident Processing System consists of the following components:

- *Data Services:* Network traffic, security logs, sensor systems, user behaviors, etc.
- *Big Data Analytics and AI/ML:* Data analytics, predictions, anomaly detection.
- *Dashboard/Display:* Events/incidents alerts or notifications.
- *Incidents/Security Reporting:* Internal/external reporting, agencies, third parties, and partners/collaborators.

V. Conclusion

Every incident is an event, although some events may not warrant an urgent reaction. As a result, incident reporting systems should be widely adopted for information collection regarding patient safety incidents. Most of the value for incident reports lies in the free-text section. In healthcare, incident reporting can be a critical function of the IT department and may depend on many factors. Clinical staff can carry mobile devices, have access to medical center

computers including email, and maintain access to programs that could threaten the very livelihood of the medical center community. A robust IT reporting system for incidents is a key function of the IT department. Clinical staff, ancillary staff, and management staff should have access to most incident reporting systems. In case of a threat such as phishing emails, Ransomware, or other malware from suspicious or malicious actors, the staff should be taught to report so IT staff can minimize the damage and halt any threat to patient care disruption. More hospitals are using big data in patient care. However, big data has many challenges and should be carefully considered by experts. Many hospitals are using AI in daily medical center functions. Using AI can alert the IT department of most threats before they become serious threats to halt all patient care functions. The IT department is using all manner of functions to identify threats and halt all threats to patient care before the threat becomes irreparable.

Acknowledgement

The authors would like to express thanks to Technology and Healthcare Solutions, USA for its help and support.

Conflict of Interest

The authors would like to announce that there is no conflict of interest.

Ethics

In this article, ethical principles related to scientific research articles are observed. The corresponding author confirms that both authors have read, revised, and approved the paper.

References

[1] H. Uematsu, M. Uemura, M. Kurihara, T. Umemura, M. Hiramatsu, F. Kitano, T. Fukami, and Y. Nagao, "Development of a novel scoring system to quantify the severity of incident reports: An exploratory research study," *Journal of Medical Systems*, vol. 46, no. 12, p. 106, 2022, doi: https://doi.org/10.1007/s10916-022-01893-1.

[2] G. Iakovakis, C.-G. Xarhoulacos, K. Giovas, and D. Gritzalis, "Analysis and classification of mitigation tools against cyberattacks in COVID-19 era," *Security and Communication Networks*, no. 1, p. 3187205, 2021, doi: https://doi.org/10.1155/2021/3187205.

[3] ENISA, ENISA Threat Landscape Report - 2020, 2021. [Online]. Available: https://www.enisa.europa.eu/topics/threat-risk-management/threats-and-trends

[4] N. Ramakrishnan, P. Butler, S. Muthiah, N. Self, R. Khandpur, P. Saraf, W. Wang *et al.*, "Beating the news' with EMBERS: Forecasting civil unrest using open source indicators," in *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Minin*g, 2014, pp. 1799-1808.

[5] Y. Matsubara, Y. Sakurai, C. Faloutsos, T. Iwata, and M. Yoshikawa, "Fast mining and forecasting of complex time-stamped events," in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2012, pp. 271-279.

[6] F. Salfner, M. Lenk, and M. Malek, "A survey of online failure prediction methods," *ACM Computing Surveys (CSUR)*, vol. 42, no. 3, pp. 1-42, 2010, doi: https://doi.org/10.1145/1670679.167068.

[7] L. Zhao, "Event prediction in the big data era: A systematic survey," *ACM Computing Surveys (CSUR)*, vol. 54, no. 5, pp. 1-37, 2021, doi: https://doi.org/10.1145/3450287.

[8] B. K. Schwab, "Insider threat management: Operating environments, detection methods and mitigation strategies," *Journal of Physical Security*, vol. 14, no. 1, pp. 13-34, 2021.

[9] J. Kinyua, and L. Awuah, "AI/ML in security orchestration, automation and response: Future research directions," *Intelligent Automation & Soft Computing*, vol. 28, no. 2, pp. 527-545, 2021, doi: https://doi.org/10.32604/iasc.2021.016240.

[10] M. Beaninger, and A. Sogabe, "Recognizing the analyst: Building machine learning models for the SOC," 2018. [Online]. Available: https://www.fireeye.com/blog/threatresearch/2018/06/buildmachine- learning-models-for-the-soc.html

[11] D. Krisiloff, "Churning out machine learning models: Handling changes in model predictions," 2019. [Online]. Available: https:// www. f i reeye. com/ blog/ threat- research/2019/04/churning-out-machine- learning-models-handling-changes-in-model- predictions.html

[12] R. Trifonov, O. Nakov, and V. Mladenov, "Artificial intelligence in cyber threats intelligence," in *2018 International Conference on Intelligent and Innovative Computing Applications (ICONIC)*, IEEE, 2018, pp. 1-4.

[13] M. H. Ling, K.-L. A. Yau, J. Qadir, G. S. Poh, and Q. Ni, "Application of reinforcement learning for security enhancement in cognitive radio networks," *Applied Soft Computing*,vol. 37, pp. 809-829, 2015. doi: https://doi.org/10.1016/j.asoc.2015.09.017.

[14] F. Petropoulos, and S. Makridakis, "Forecasting the novel coronavirus COVID-19," *PloS One*, vol. 15, no. 3, p. e0231236, 2020, doi: https:// doi.org/10.1371/journal.pone.0231236.

[15] B. Adhikari, X. Xu, N. Ramakrishnan, and B. A. Prakash, "Epideep: Exploiting embeddings for epidemic forecasting," in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 2019, pp. 577-586.

[16] S. Santiso, A. Casillas, A. Pérez, M. Oronoz, and K. Gojenola, "Adverse drug event prediction combining shallow analysis and machine learning," in *Proceedings of the 5th International Workshop on Health Text Mining and Information Analysis (Louhi)*, 2014, pp. 85-89.

[17] T. Abe, H. Sato, and K. Nakamura, "Extracting Safety-II factors from an incident reporting system by text analysis," *Cureus*, vol. 14, no. 1, p. e21528, 2022, doi: https://doi.org/10.7759/ cureus.21528.

International Journal of Knowledge Based Computer Systems
13 (1) 2025, 38-46
www.publishingindia.com//ijkbcs/

Bridging the Gap: Developing an Alumni Association Platform to Enhance Student Guidance and Institutional Connectivity

Ajay Jaiswal[1*], Priya Solanki[2], Rohit Rathore[2], Shreejith Pillai[2], Sneha Jain[2], Tisha Berchha[2] and Vaibhav Toke[2]

[1]Assistant Professor, Department of Computer Science and Engineering, Prestige Institute of Engineering Management and Research, Indore, Madhya Pradesh, India.
Email: ajay.jaiswal55555@gmail.com
[2]Student, Department of Computer Science and Engineering, Prestige Institute of Engineering Management and Research, Indore, Madhya Pradesh, India.
[*]Corresponding Author

Abstract: During their collegiate journey, students face uncertainties and challenges that could be alleviated through guidance from alumni; however, many senior graduates do not maintain active connections with junior students, leading to a significant communication gap. During recruitment periods, students require insights into career paths and potential employers, yet the lack of a systematic record of alumni achievements and a centralized communication platform often leaves them unsupported. Alumni, on the other hand, may lose contact with their alma mater, missing out on opportunities to engage with current students and support ongoing initiatives. To address these issues, we propose a comprehensive alumni association platform. This platform aims to bridge the communication gap, enabling alumni-student mentorship, showcasing alumni success stories, providing a job portal, and promoting alumni events. By fostering meaningful connections and career advancement, the platform cultivates a lasting relationship between alumni and their institution.

Keywords: Alumni association platform, Data management, Donation portal, Event coordination, Job opportunities, Mobile and web application, User-friendly interface.

I. Introduction

Alumni associations serve a crucial function in fostering enduring connections between graduates and their alma mater, often functioning as an intermediary to provide networking, mentorship, and professional development opportunities. Robust alumni networks contribute to both individual career advancement and the institution's reputation; however, many institutions encounter challenges in maintaining active alumni engagement. Over time, graduates may become disconnected from their alma mater, unaware of events, student support opportunities, or avenues for financial contribution. For current students, the absence of a structured alumni network limits access to valuable guidance, particularly during recruitment periods when insights from experienced alumni regarding industry trends and potential employers are invaluable.

In response to these challenges, this study proposes the development of a dynamic Alumni Association platform specifically designed for Government Engineering College. This platform will facilitate connections between alumni and students through an accessible web and mobile interface, enabling mentorship, professional networking, and community building. Key features include a secure donation portal, an alumni directory, job postings, and the

capability to share success stories. Furthermore, the platform will host event announcements and registration, enabling alumni to maintain involvement in institutional activities and fostering a vibrant, supportive alumni-student ecosystem. This solution not only enhances alumni engagement but also empowers students, thereby strengthening the overall legacy of the institution.

II. Literature Survey

Extensive research has been conducted on the function of alumni associations and their digital management systems, underscoring their pivotal role in preserving institutional-graduate relationships. These platforms serve to augment networking opportunities, foster professional growth, and promote philanthropic initiatives, while incorporating features that bolster both alumni involvement and the career advancement of current students.

- *Enhancing Alumni Growth and Engagement:* A study underscores the significance of alumni associations in facilitating institutional growth through the cultivation of alumni as ambassadors and mentors. The research proposes the incorporation of alumni into governance and decision-making processes to enhance their commitment and influence, thereby augmenting the university's reputation and societal impact [1].
- *Data Collection and Profile Management:* Effective alumni data management is essential for accurate engagement. An alumni tracking system developed at Tribhuvan University addresses challenges in managing diverse data sources, incorporating self-updating profiles for enhanced tracking accuracy. This system supports a dynamic, up-to-date database of alumni, enabling users to view the career trajectories of their cohort members, thereby fostering active participation in the alumni network [2].
- *Integrated Media and Cloud-Based Solutions:* The Integrated Media Model presents a cloud-based solution for alumni associations, utilizing a "1+M+N" model for multi-level engagement. This model incorporates functionalities for establishing distinct alumni identities and provides scalable cloud data management to facilitate effective alumni relations across diverse geographical regions and institutional departments [3].
- *Modular Platform Design for Alumni Engagement:* The Alumni Connect Hub, developed for Sipna College, demonstrates the efficacy of modular platform design. Key features encompass profile management, event registration, donation options, and job boards, providing alumni and students with an interactive and user-centric experience for career support and institutional contributions [4].
- *Career Support and Feedback Systems:* A modular alumni portal developed for JD College emphasizes job postings, event updates, and feedback systems, aiming to enhance alumni-student connections and facilitate students' career advancement through consistent interactions with alumni [5].
- *Data Security and Scalable Solutions:* An alumni portal designed for Nigerian universities prioritizes secure and scalable data management. Through the implementation of secure login protocols and planned donation integrations, this platform addresses data security concerns while facilitating alumni engagement [6].
- *Mentorship-Focused Smart Alumni System:* The Smart Alumni System (SAS) integrates social networking features with data mining to facilitate student-alumni mentor matching based on shared interests and career backgrounds. This system fosters individualized mentorship and collaborative opportunities for students and alumni [7].
- *Data-Driven Insights with Regression Models:* The alumni network platform developed at the University of Antique utilizes regression models to enhance user satisfaction by analyzing predictors such as navigation ease and mobile accessibility. The platform also

proposes artificial intelligence and virtual reality integration to augment engagement through personalized experiences and advanced security measures [8].

TABLE I: SUMMARY OF KEY FEATURES ACROSS ALUMNI MANAGEMENT SYSTEMS

Feature	Study 1	Study 2	Study 3	Study 4	Study 5	Study 6	Study 7	Study 8
Profile Management	✓	✓	✓	✓	✓	✓	✓	✓
Networking & Mentorship	✓	Limited	✓	✓	✓	✓	Advanced	✓
Career Support	Limited	✓	✓	✓	✓	✓	✓	✓
Event Management	✓	No	✓	✓	✓	✓	✓	✓
Donation Management	✓	No	✓	✓	No	Planned	No	✓
Data Security	Emphasized	Emphasized	Emphasized	Emphasized	Emphasized	Emphasized	Yes	Advanced
Data Analytics	No	Limited	No	✓	No	No	✓	Advanced
AI/ML Integration	No	No	No	No	No	No	Planned	✓
User Engagement Features	Basic	Limited	Moderate	High	High	Moderate	High	High

III. EVALUATION CRITERIA

This section delineates the metrics and criteria employed to evaluate the efficacy and technical performance of alumni management systems, with a focus on user engagement and technological functionality.

A. User Engagement Metrics

The effectiveness of alumni management systems is measured by their capacity to engage users meaningfully and consistently. Key metrics include:

- *Engagement Rate:* The proportion of active users relative to the total registered alumni. For instance, systems such as the Smart Alumni System achieve higher engagement rates by offering personalized mentorship features that promote consistent utilization.
- *User Satisfaction:* Satisfaction levels are influenced by ease of navigation, relevance of resources, and accessibility. Regression analysis in the University of Antique Platform identified navigation clarity and mobile accessibility as primary determinants of user satisfaction.

B. Technological Effectiveness

- *Security and Privacy:* Ensuring the protection of alumni data is paramount. Platforms such as the Nigerian Universities System emphasize secure login protocols and encrypted data storage, fostering trust and ensuring data integrity. Software risk assessment methodologies can mitigate uncertainties in platform performance and security [9].
- *Scalability:* Scalability ensures that systems can accommodate growing alumni networks without compromising performance. Cloud-based solutions, such as the Integrated Media Model, excel in managing large-scale data operations while remaining accessible to diverse alumni communities. Tools for software estimation, like the Project Estimator Tool, emphasize the necessity for accurate budgeting in alumni platform development [10].

IV. COMPARATIVE ANALYSIS OF KEY STUDIES

This section examines the commonalities and distinctive aspects of the reviewed studies, emphasizing the prioritized features and the application of technologies.

A. Common Features Across Alumni Systems

- *Profile Management:* A fundamental feature in all reviewed systems, enabling alumni to update their information. The Alumni Portal and Tracking System demonstrates the significance of self-updating profiles for maintaining accurate alumni databases [11].
- *Networking and Mentorship:* Facilitating professional connections is a priority. Systems such as the Smart Alumni System excel by utilizing data mining to match alumni mentors with students based on shared interests and career trajectories.
- *Career Support:* Platforms such as the Alumni Connect Hub integrate job posting and career support features, offering alumni opportunities to contribute professionally while assisting students in transitioning into the workforce.

B. Impact of Emerging Technologies

Technological advancements enhance alumni platforms' capabilities:

- *AI and Machine Learning:* The University of Antique Platform utilizes AI to provide personalized user recommendations, enhancing engagement through tailored content, content-based and collaborative filtering techniques, as explored in prior research, are pivotal in enhancing platform recommendations [12]. The integration of AI techniques, as discussed in computing studies, can enhance mentorship and data analysis features [13].
- *Cloud Computing:* The Integrated Media Model employs cloud-based infrastructure to manage alumni data efficiently, ensuring scalability and accessibility.
- *3D Game Engine:* Immersive technologies, as explored in 3D game engine studies, could offer innovative solutions for interactive and engaging alumni platforms, enhancing the user experience [14].

C. Opportunities for Future Development

Several gaps in current systems provide opportunities for innovation:

- *Block Chain Technology:* For secure donation management and transparent transactions.
- *Advanced AI Features:* For real-time mentorship recommendations and automated event management.

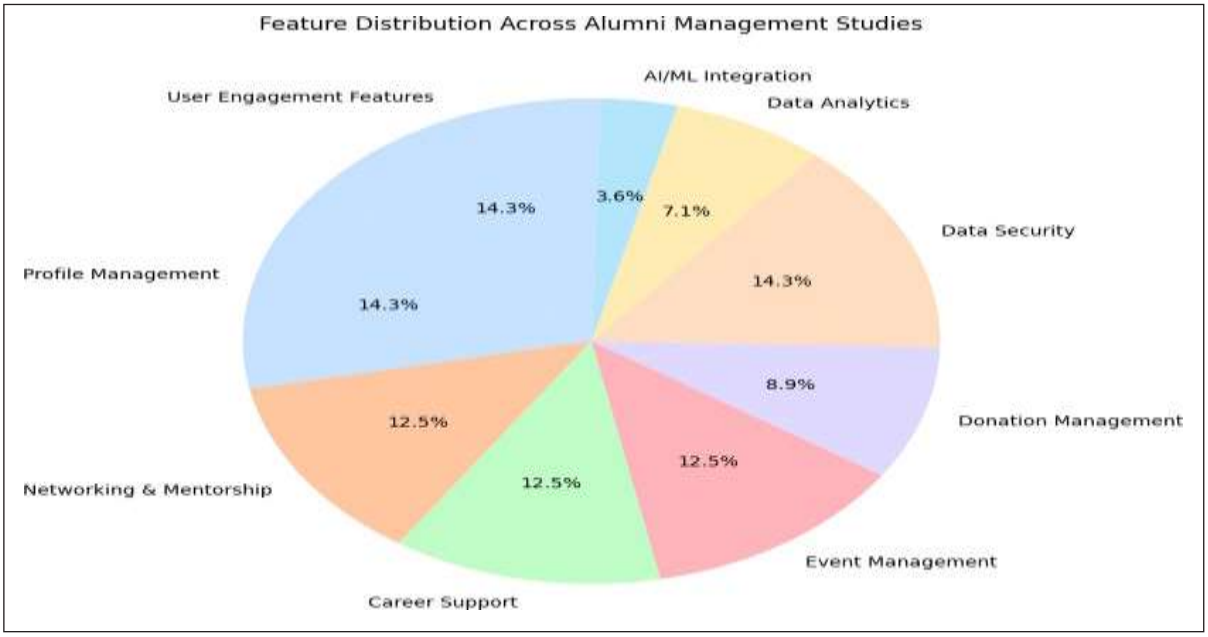


Fig. 1: Pie Chart of Most Common Features in Alumni Systems

This chart should illustrate the frequency of each feature (e.g., Networking & Mentorship, Career Support) across the studies, showing readers which functionalities are prioritized.

V. Summary of Observations

This section synthesizes the insights derived from the comparative analysis, focusing on commonalities, challenges, and areas for potential improvement.

A. Commonly Implemented Features

- *Universal Features:* Profile Management, Networking, and Career Support are implemented in nearly all systems, underscoring their significance in fostering meaningful alumni engagement.
- *Technological Essentials:* Data security and scalability are emphasized in every system, reflecting the necessity for secure and robust platforms.

B. Challenges and Limitations

- *Data Privacy:* Ensuring data security remains a significant challenge, particularly for systems with extensive alumni networks.
- *User Engagement:* Platforms frequently encounter difficulties in maintaining long-term engagement, especially among geographically dispersed alumni.

Table II: Evaluation of Features and Limitations in Alumni Systems

Feature	Advantages	Limitations
Profile Management	Keeps alumni data up-to-date and accurate.	Relies on alumni participation to update data.
Networking	Builds professional connections and mentorship.	Requires active user engagement.
Career Support	Provides job opportunities and professional guidance.	Limited to specific industries or alumni roles.
Event Management	Simplifies event planning and registration.	Engagement often limited to local alumni.
Donation Management	Encourages financial contributions.	Security concerns with online payment systems.
Data Security	Protects sensitive alumni information.	Costly to implement and maintain securely.

VI. Proposed Model, Testing and Results

A "literature review" is a systematic method for identifying, evaluating, and synthesizing all sources relevant to a specific research question, activity, or phenomenon of interest. Observational or experimental studies that directly relate to the subject under investigation often form the basis of secondary research. In this case, the literature review provided insights into key features, challenges, and innovative solutions within alumni management systems.

Based on the findings, a conceptual framework for an alumni management platform was proposed, focusing on bridging the communication gap between alumni and students, enhancing career support, and streamlining event and donation management. Mobile-based systems like highlight the importance of flexibility and accessibility in alumni platforms [15]. The following sections describe the proposed model in detail and outline the outcomes achieved.

A. Proposed Model

The proposed alumni management system builds upon insights from existing literature and aims to integrate critical functionalities that address current limitations in engagement and data management. Fig. 2 illustrates the flow of the proposed system. Key components of the proposed model are as follows:

- *User Registration and Authentication*
  - Alumni and students register on the platform using unique credentials provided by the institution. The system grants role-specific access to features.
- *Alumni Dashboard*
  - View and respond to student queries, enabling mentorship and guidance.
  - Participate in events through a streamlined registration process.
  - Contribute donations with flexible options (one-time, recurring, or project-specific).
- *Student Dashboard*
  - Access alumni profiles, sorted by field, batch, or geographic proximity.
  - Explore job and internship opportunities posted by alumni.
  - Engage in coding challenges (Code-sphere) and explore research papers and success stories for inspiration.
- *Event Management System*
  - Register for events with payment integration and receive automated receipts.
  - Administrators can manage event participation and track attendee metrics in real time.
- *Data Integration and Dynamic Updates*
  - Alumni and student datasets are sourced from institutional records, including details such as academic history, career achievements, and contact information. As elucidated in previous research, systematic methodologies and precise metrics are imperative for analyzing datasets effectively, mitigating uncertainties, and enhancing predictions [16]. These principles guide the utilization of alumni and student data in platform development.

The system dynamically updates profiles to ensure data accuracy.

B. Testing

The systematic identification, analysis, and integration of functional data play a crucial role in ensuring the operational efficacy of alumni management platforms, as elucidated in extant research on digital systems [17]. The proposed model will undergo rigorous testing to validate its functionality, performance, and user satisfaction. Testing strategies include:

- *Functional Testing*
  - Validate all core features such as profile management, alumni-student chat, job posting, mentorship matching, and donation tracking.
  - Ensure smooth user navigation within both alumni and student dashboards.
- *Performance Testing*
  - Test the system's response time under varying loads and simulate high-traffic scenarios to evaluate scalability on cloud infrastructure.
- *User Acceptance Testing (UAT)*
  - Conduct trials with alumni and students to gather feedback on usability and feature effectiveness.
- *Event Management Testing*
  - Simulate event registrations to ensure accurate invoicing and seamless attendance tracking.
  - Test the automated communication system to ensure all alumni receive event invitations.

C. Results and Expected Outcomes

Although the model is conceptual, based on insights from the literature, the following outcomes are anticipated:

- *Enhanced Alumni Engagement:* Increased participation in events and mentorship programs is expected due to personalized recommendations facilitated by artificial intelligence [13]. Event participation increased significantly as a result of the streamlined registration process and WhatsApp invite integration.
- *Improved Career Support:* The platform enabled direct interactions between students and alumni, resulting in tangible outcomes such

as internships and coding skill development. The job and internship portal assisted students, while alumni shared curated coding challenges (Code-sphere), benefiting students.

- *Increased Donations:* Transparent and secure donation tracking fosters trust, potentially encouraging more alumni to contribute financially.
- *Actionable Insights:* Real-time dashboards provide administrators with data-driven insights to refine engagement strategies.
- *Efficient Event Management:* Event registration and invoicing processes were optimized, reducing administrative time and minimizing errors in participant tracking.

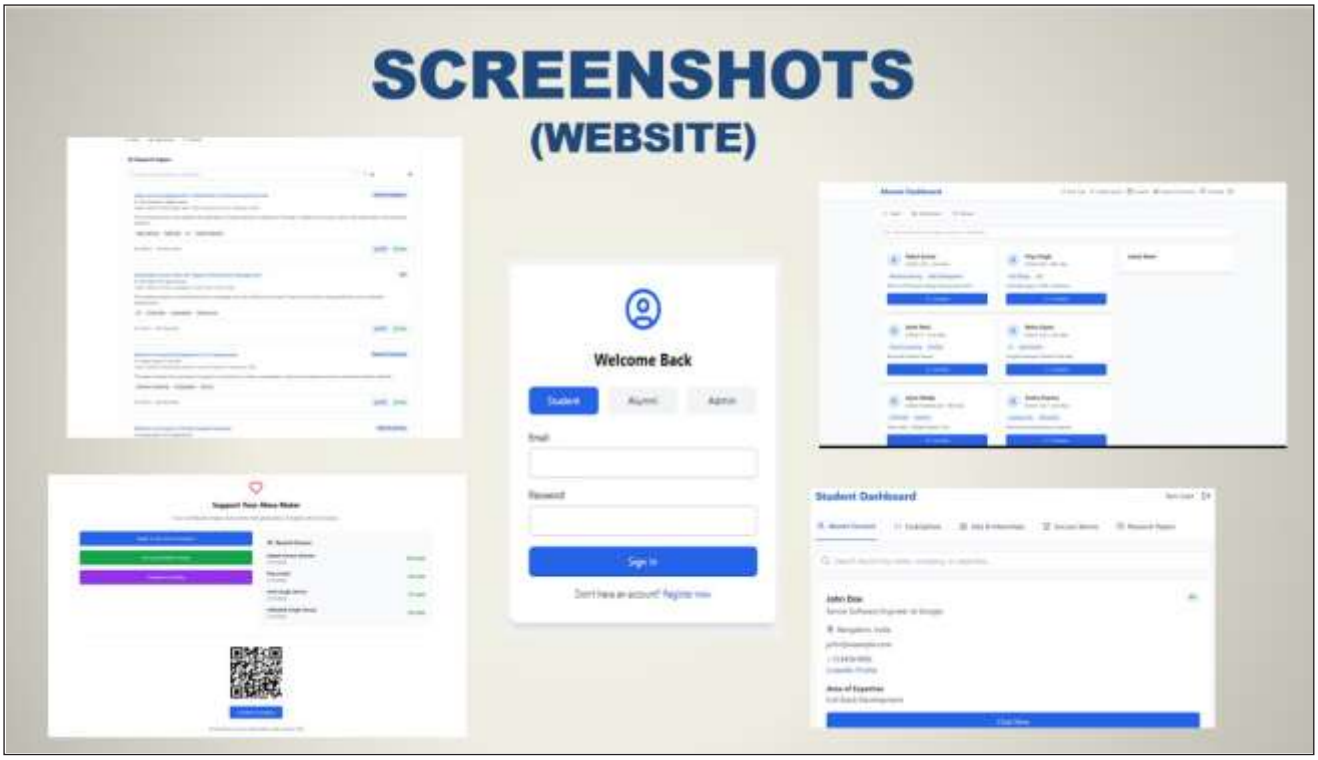


Fig. 2: Snapshot of Proposed Model

VII. Conclusion

Alumni associations serve a crucial function in fostering enduring connections between graduates and their alma mater, often functioning as an intermediary to provide networking, mentorship, and professional development opportunities. Robust alumni networks contribute to both individual career advancement and the institution's reputation; however, many institutions encounter challenges in maintaining active alumni engagement. Over time, graduates may become disengaged from their alma mater, unaware of events, student support opportunities, or avenues for financial contribution. For current students, the absence of a structured alumni network limits access to valuable guidance, particularly during recruitment periods when insights from experienced alumni regarding industry trends and potential employers are essential.

In response to these challenges, this study proposes the development of a dynamic Alumni Association platform specifically designed for Government Engineering College. This platform will facilitate connections between alumni and students through an accessible web and mobile interface, enabling mentorship, professional networking, and community building. Key features include a

secure donation portal, an alumni directory, job postings, and the capability to disseminate success stories. Furthermore, the platform will host event announcements and registration, enabling alumni to maintain involvement in institutional activities and fostering a vibrant, supportive alumni-student ecosystem. This solution not only enhances alumni engagement but also empowers students, thereby strengthening the overall legacy of the institution.

REFERENCES

[1] D. Obeng-Ofori, and H. O. Kwarteng, "Enhancing the role of alumni in the growth of higher education institutions," *International Journal of Multidisciplinary Studies and Innovative Research,* vol. 4, pp. 40-48, 2020.

[2] B. Bista, A. Shakya, B. Joshi, A. Pokhrel, ... and D. S. Baral, "An alumni portal and tracking system," *Journal of the Institute of Engineering*, vol. 16, no. 1, 2021.

[3] J. Wang, and Z. Zhang, "Research on alumni resources development and alumni association work in colleges and universities under "Integrated Media" environment," *Advances in Education, Humanities and Social Science Research,* vol. 6, no. 1, p. 558, 2023.

[4] P. P. Sawai, P. V. Chambhare, A. N. Jaysingpure, A. G. Karhe, D. Rathod, and V. S. Gulhane, "Alumni Connect Hub: A comprehensive alumni management system," *International Journal of Ingenious Research, Invention and Development*, vol. 3, no. 1, pp. 56-64, 2024.

[5] N. Bhandarkar, H. Dudhat, A. Verma, U. Sahare, and M. John, "Alumni web portal for colleges," *International Journal for Research in Applied Science and Engineering Technology*, 2023.

[6] R. Enihe, and V. Omopariola, "Alumni portal system for Nigerian Universities (A case study)," *International Journal of Science and Research*, vol. 12, no. 2, pp. 1047-1055, 2023.

[7] H. Chi, E. L. Jones, and L. P. Grandham, "Enhancing mentoring between alumni and

students via smart alumni system," *Procedia Computer Science*, vol. 9, pp. 1390-1399, 2012.

[8] N. M. Yumen, "Alumni network platform leveraging regression models for data analysis," *African Journal of Biological Sciences*, vol. 6, no. special issue 2, 2024.

[9] A. Jaiswal, and N. Singh, "Software risk assessment process using soft computing methods," in *Proceedings of International Journal of Contemporary Technology and Management (IJCTM)*, ISSN: 2278-8034, Florida-US, vol. 1, no. 1, 2013.

[10] A. Jaiswal, and M. Sharma, "An expert estimator tool to estimate project cost and risk with early stage of function points," *International Journal of Software Engineering & Applications (IJSEA)*, vol. 3, no. 5, Sept. 2012, doi: https://doi.org/10.5121/ ijsea.2012.3512.

[11] A. Jaiswal, K. Bhardwaj, K. Sahu, J. S. Shekhawat, and K. Dhakad, "EmoAssistant – Arya," *Journal of Xidian University*, vol. 18, no. 7, Jul. 2024.

[12] A. Jaiswal, and K. Dutta, "Strategies in recommender systems: A comprehensive study of content-based and collaborative filtering techniques on a custom-curated movie dataset," *International Journal of Computational Engineering Research (IJCER)*, ISSN: 22503005, vol. 14, 2024.

[13] A. Jaiswal, P. Malviya, L. Parihar *et al.*, "Software cost estimation tool: A App based application, estimate the cost of software project," *Computing and Artificial Intelligence*, vol. 2, no. 2, p. 1364, 2024, doi: https://doi. org/10.59400/cai.v2i2.1364.

[14] A. Jaiswal, D. Modi, M. Gupta, M. Mandloi, M. Mehta, and D. Mukheja, "3D game engines as a new reality," *International Journal of Research Studies in Science, Engineering & Technology (IJRSSET)*, ISSN: 2349-4751 (Print) & ISSN: 2349-476X (Online), vol. 1, no. 4, pp. 36-41, Jul. 2014.

[15] A. Jaiswal, A. Goyal, D. Punase, M. Chandwaney, and P. Moharil, "Skill maestro: GRE e-preparation android based system," *International Journal of Advanced Information and Communication Technology (IJAICT)*, ISSN: 2349-6339 (Print), ISSN: 2348-9928 (Online), vol. 2, no. 2, pp. 36-41, Jun. 2015.

[16] A. Jaiswal, J. Raikwal, and C. Chauhan, "Development of software projects: A review of various cost estimation techniques," *Journal of Harbin Engineering University*, ISSN: 1006-7043, vol. 44, no. 7, Jul. 2023.

[17] A. Jaiswal, and M. Sharma, "An expert estimator tool to estimate project cost and risk with early stage of function points," *International Journal of Software Engineering & Applications (IJSEA)*, vol. 3, no. 5, Sept. 2012, doi:https:// doi.org/10.5121/ijsea.2012.3512.

International Journal of Knowledge Based Computer Systems
13 (1) 2025, 47-56
www.publishingindia.com//ijkbcs/

Emerging Trends of Blockchain Technology in the Financial Sector

Anusha Thakur

Independent Researcher, India.
Email: anushathakur89@yahoo.com

Abstract: Transforming the financial world encompasses trailblazing ways to pave the way for tangible business outcomes. Blockchain and ledger technology are increasingly plaguing the industry today regarding efficiency and security. The decentralized ledger technology, underlying cryptocurrencies, offers different ways of organizing the dynamic finance industry. The government and businesses increasingly emphasize implementing state-of-the-art innovativeness, owing to the merits offered. Being one of the leading innovations, blockchain holds the potential to ensure secure and quick trades and transactions at minimal fraud, and risks within the interrelated financial system. Technology has led to a significant push to transform the structure and operational activities of the finance industry, with effective data solutions catering to the contemporary distributed industries. This research topic emphasizes on the factors backing the decisions of financial institutions to adopt blockchain technology and develop the standards and objectives to cater to the changing scenario. It emphasizes on researching the status quo of blockchain technologies in the financial sector. The topic further contemplates a brief analysis of the use cases, probable solutions for interoperability, scalability, and privacy concerns of adopting the technology in the financial domain.

Keywords: Audit, Blockchain, Distributed ledger, Finance, Technology.

I. Introduction

Technology has remodeled the way, businesses operate with bold and new approaches escalating disruptions. Over the past years, industries across the globe have witnessed the most exciting, drastic, and challenging revolution with digitalization and indomitable advancements. However, with the digital currency's value getting varied, the blockchain's potential as a pioneering technology for businesses with its ability to carve through the layers of ineptitude, is gaining significant traction. Blockchain has entered a new age of more practical and wider adoption, with several key players deciphering its long-term potential in businesses. The need to incorporate the latest services, augment client service delivery, and attain efficiency has largely converged on imperatives for exploring ways to redesign operational and product processes.

With blockchain, the financial sector stands on the verge of a major uprise redefining the way businesses interact, trade, and manage financial assets. In today's scenario, trade settlement, credit transformation, payments, and other transactions are largely being accomplished via the latest Internet financial systems owing to the unparalleled benefits offered by this state-of-the-art technology [1]. From security and speed to a decentralized and anonymous nature, blockchain poses as a solution to change the conventional business landscape. Technology has increasingly disrupted the way, traditional financial and banking services operate, with their benefits different from the latest services [2].

The mounting hold of blockchain is reflected in a stout shift among conventional financial firms across the regions. In addition to bolstering cycle-time reductions and substantial efficiencies, this cutting-edge technology also offers complete end-to-end transparency across the financial operations, thereby augmenting predictive operational insights and avenues optimizing the working capital. The technology is also engaged in reshaping a varied range of financial processes such as: financing (invoice factoring, letters of credit, trade finance), warranties, rebates, order-to-cash, procure-to-pay, and intercompany transactions (in case of multiple ERPs) [3].

Blockchain technology is expected to make a difference in any paper-cluttered industry, minimizing extra intermediaries and transaction paperwork. It also offers a single and decentralized data source, letting contributors track financial transactions, trade procedures, and operations chronologically, thereby offering a new level of efficiency, and transparency. Incorporating this technology is strongly expected to push the financial segment towards better data safety, employee productivity, and client experiences. However, in today's highly competitive market landscape, the transformational shift towards blockchain bolstering new revenue streams and business models is anticipated to take some time, owing to the myriad risks such as regulatory compliance, scalability, privacy, interoperability, and lack of standardization issues.

Key Purpose of the Article: The objective of this article is to offer insights into the development of specific technological applications to augment audit trail, transparency, and performance processes. The article discusses the key concepts of centralized and decentralized finance, along with the key use cases of inter-company reconciliations, vendor management, forecasting, and budgeting. In addition, the paper also focuses on addressing the technical challenges, while catering to the benefits of incorporating the technology.

Key Findings: State-of-the-art technologies such as blockchain augment shared operating models, secure inclusive business networks, and efficient processes in the financial sector. The paper briefly emphasizes on the latest trends and advancements in the adoption of blockchain in the financial sector, with different examples of the financial domains incorporating these cutting-edge technologies to revolutionize the same.

II. Literature Review

With businesses becoming decentralized and digitized, the necessity for reliable, transparent, and secure systems is predominantly rising. The digitization of fiscal instruments entailing programmable money, smart contracts, and digital assets, forges unparalleled levels of programmability and connectivity between the holdings, assets, services, and products, reevaluating the processes and creating a new paradigm [4]. The digital financial instruments are tailored to the investor demands, minimizing the costs, and risks, and expanding the investor's market. The financial institutions across are paving the way ahead with the companies working towards eliminating the longstanding friction, offering tangible business outcomes and solutions [5].

With the continual evolution of blockchain technology, the potential to disrupt and revolutionize the financial sector in different industries extends far beyond its use cases in cryptocurrency. The technology owes the capability to endorse financial inclusion by mounting access to financial services for the underbanked and unbanked population segments. In addition, the impact of this technology in finance is expected to continue evolving, with innovations such as DeFi [6]. The decentralized finance (DeFi) platforms, built on blockchain technology, are engaged in offering a range of financial services, such as borrowing, lending, and asset management, without the prerequisite for conventional intermediaries such as banks. By minimizing the barriers to entry and offering cost-effective solutions, the DeFi platforms seamlessly fix the gap between the unbanked and banked populations, facilitating better economic development and financial inclusion [7].

The key financial titans are largely giving importance to DeFi, tokenization, blockchain, and digital investment products, accentuating an ongoing shift and confirming that these concepts are essential tools for future-proofing financial operations [8]. With these institutions continuing to invent and acclimatize, it is evident that blockchain offerings are anticipated to become a vital part of the future of finance, offering unparalleled levels of accessibility, transparency, and productivity.

III. Transformative Potential of Blockchain in Different Segments

From modernizing regulatory compliance to streamlining payment systems, blockchain technology owe the potential to revolutionize conventional business practices and models, steering in a new era of transparency, security, and efficiency in financial services. This pioneering technology endorses financial inclusion, and bolsters economic growth, thereby holistically benefitting individuals and businesses. The potential of blockchain in different areas of the financial sector is illustrated as:

- *Enhancing Trade Finance:* With the changing market scenario, it is becoming essential to reinvent trade finance systems, subsequently opening new opportunities backing the same. Blockchain technology is gaining significant traction as a solution, to solve the trust issues among trading associates in trade finance [9]. This state-of-the-art technology based on the three technological pillars (cryptography, distributed consensus mechanisms, decentralization) augments it to afford the attainment of trust-free economic transactions [10]. Blockchain technology caters to trust-related issues in economic transactions, thereby enabling trading partners to be free from the need to implement mechanisms to convey or signal trust. The transformative impact of this technology is being widely recognized by the financial sectors, to generate new revenue, augment end-user experiences, and deliver process efficiency, thereby minimizing the challenges in the businesses.
- *Revolution of Asset Management:* Blockchain technology, a strategic tool for upholding a competitive edge in the age of digital transformation, is restructuring the asset management sector, with unparalleled levels of security, automation, efficiency, and transparency. By changing the way assets are handled, transacted, and tracked, this technology augments the asset management processes, improves the supply chain traceability, unleashes liquidity for illiquid investments, and opens newer opportunities [11].
- *Transforming Insurance:* Several insurers are largely benefitting from metaverse and blockchain technologies, with casualty, property, and health businesses leading the way. The next-generation products are comprised of policies that cover digital assets that can either be deactivated or activated according to demand. Incorporating blockchain technology streamlines the claims and underwriting processes, thereby minimizing the risk profile and insurers' loss ratios. In addition, the technology backs efficient payments between third parties and insurers, paving the way for transparent claims processes, and streamlined subrogation [12]. Hence, reinsurers, incumbent insurers, and insurtechs increasingly emphasize incorporating cutting-edge innovations to make the processes seamless and efficient.

IV. Blockchain Trends Shaping the Financial Sector

Over the years, blockchain technology has been considered the future of financial markets. It has been labeled transformative and generational owing to the potential applications and its decentralized nature across varied industries, with the capacity to unravel a continuously expanding list of operational issues.

- *Growth of Decentralized Finance:* Decentralized Finance (DeFi) is gaining traction as an emerging model for enabling and organizing cryptocurrency-based financial services in small-sized businesses of developing regions.

- Decentralized finance (DeFi), the usage of blockchain technology in financial services owe the potential to improve access, endorse competition, and augment efficiency in the financial markets to capital for small-size businesses, fueling economic growth. Expansion of the developing DeFi market establishes on whether the matters of the governing authorities, consumers, and investors can be positively catered to, without amending the technology into disuse and inefficiency in the process [13].
- *Adoption of CBDC:* The developed economies are expected to roll out the initial versions of Central Bank Digital Currency (CBDC) by 2025, to augment resilience, efficiency, and transparency in the financial systems. The interoperable CBDCs and their infrastructure across the different regions are gaining significant momentum, with several monetary authorities and banks emphasizing on how to successfully implement CBDCs. According to the reports of the Bank for International Settlements, nearly 9 wholesale and 15 CBDCs are expected to be in circulation by 2030 [14].
- The introduction of CBDCs is anticipated to complement a new dimension of blockchain technology that will revolutionize the future of payments for both individuals as well as businesses. The assimilations in CBDCs embody an archetype shift in the way, governments and financial institutions perceive digital currencies and implement the same.
- *Growing Digital Asset Markets:* With developments in decentralized finance, the volumes of digital assets, and cryptocurrencies along with the market capitalization are expected to witness a tremendous rise over the forecast period.

V. Use Cases of Blockchain in Finance

A. Blockchain Payment Processing

The current payment systems are riddled with the issues of reliance on third-party intermediaries, hefty fees, and slow processing times, pushing the way for efficient and secure systems. A blockchain-based payment system is a decentralized digital infrastructure leveraging blockchain technology to accelerate direct and secure transactions without intermediaries. The technology enables direct transactions, utilizing the consensus algorithm to validate and register the transactions on a distributed ledger. Exploiting cryptographic techniques enables the payment systems to ensure the actuality, privacy, and integrity of the transactions, offering augmented transparency and security in comparison to the conventional payment systems [15]. Integration of blockchain technology is expected to revolutionize overall transparency, security, and the way transactions are processed [16]. This can be illustrated as:

JPMorgan Chase & Co

About the Company: JP Morgan Chase & Co., headquartered in New York City was established in 1799. The company is one of the leading financial institutions offering solutions in private banking, payments, asset management, commercial banking, and investment banking.

Solutions Offered: The company announced the launch of Onyx Unit, the bank's blockchain and digital asset to enable programmable payments using cryptocurrency-style technology, and focusing on enticing business-to-business clients in the domain of payment automation. Onyx Unit, formed in 2020, was the first bank-led blockchain platform for exchanging digital assets, information, and value.

Such features, available for coin blockchain accounts, leverage distributed ledgers and smart contracts to automate payments under explicit conditions [17]. The technology augments programmable payments to offer automation based on specific conditions via distributed ledgers and smart contracts. The Onyx Unit adds more flexibility to process the trading securities and international payments, thereby enabling the clients to use programmable payments to create more precise rules for transactional executions [18]. The different significant products offered by the Onyx Unit can be illustrated as [19]:

- *Liink:* The largest blockchain-based network with growing offerings and memberships.

These information exchange networks augment seamless payment-related data sharing between institutions.

- *JPM Coin:* Blockchain-based account ledger and payment rail. These next-generation payment rails bolstering the right-time value transfers.
- *Onyx Digital Assets:* A blockchain network augmenting value exchange for different types of digital assets.
- *Blockchain Launch:* A center of leadership and excellence in developing and introducing the latest shared networks, applications, and platform technology.

B. Blockchain in Money Transfer

Over the past few years, advancements in blockchain technology are expected to lead to the development of several sophisticated platforms for the seamless process of money transfers. Integrating blockchain technology in the process of money transfers helps augment speed and efficiency, ensure full traceability and transparency, minimize costs, and secure transactional processes [20]. The money transfer applications extend beyond cheaper and faster transactions encompassing financial inclusion, micropayments, and cross-border payments, thereby negating the need for trust in a central authority with minimized chances of fraud and manipulation. However, there are limitations to the adoption of this technology as well, wherein, with the technology reaching a maturity stage, more cost-effective, secure, and efficient ways of revolutionizing financial services are expected to gain a surge [21]. This can be illustrated as:

Chainalysis

About the Company: Chainalysis established in 2014, is headquartered in New York City. The company offers cryptocurrency compliance and investigation solutions to global law enforcement businesses, regulators, and agencies, as they work in amalgamation to fight illegitimate cryptocurrency activities. Chainalysis is the blockchain data platform, backed by Benchmark and other top names in the venture capital.

The company is engaged in offering research, services, software, and data solutions to cyber-security, insurance, financial institutions, exchanges, and government agencies, in approximately 70 countries. The data offered powers market intelligence, compliance, and investigation software to address the increasing consumer accessibility to cryptocurrency and the most high-profile criminal cases [22].

Solutions Offered: Chainalysis emphasizes on building trust between the parties and generating reliable compliance standards at minimized risks. In terms of blockchain, the company has aided in some of the most high-profile criminal cases. The company helps the government and financial institutions investigate blockchain-related activities and carefully observe cryptocurrency exchange. The different tools offered by the company focus on detecting compliance, fraudulent trading, and laundering violations, as well as marking the NFT security risk, to build trust in the blockchain.

The company delivers industry-leading blockchain intelligence by linking the on-chain activities to real-world entities via extensive customer networks, dedicated forensic experts, and sophisticated machine learning. The blockchain intelligence solutions offered by the company can be illustrated as:

Blockchain Intelligence - Mapping the real-world entities to on-chain activity [23]:

- *Breadth:* The ground-reality attributions are gathered with verifiable, directly observable evidence illustrating that an address belongs to a service. Such attributions are the initial focal point to start the real-world entities to the on-chain addresses. The advantages offered include:
  - *i Dynamic Token Support:* Chainalysis backs up all the non-fungible and fungible tokens for the smart-contract-based networks.
  - *ii Clustering Speed:* Efficient tracking of clustering capabilities in the pipeline enables efficient deployment and quick iteration of the clustering heuristics for the blockchains.
  - *iii Onboarding Speed:* The generalized frameworks augment faster integrations with the newer chains.

- *Depth:* Tying the ground-truth attributions to single addresses helps attain the breadth of coverage, necessitating depth. Understanding the addresses and entity controls viagrouping or clustering plays a pivotal role. With the clustering of the addresses, a complete and deeper mapping of the interactions between the entities is created. Different amount of rigor and clustering heuristics involved can be illustrated as:
  - *i Service-Specific Heuristics:* Nearly 100s of service-specific heuristics employed are tailored for a precise entity's architecture.
  - *ii Network-Wide Heuristics:* Network-wide heuristics can be applied to any wallet on a given EVM or UTXO blockchain.
- *Accuracy:* Erroneous data is worse than missing data.
- *Data Quality:* The need for supporting evidence to check the data quality and accuracy.
- *Data Validation:* Validation of the data includes:
  - *i Public Sector Setting the Industry Standard:* A substantial number of chainalysis customers include intelligence agencies, law enforcement agencies, and customers at different levels.
  - *ii Customers Validating the Clusters Daily:* The addresses of several services that are clustered are also the ones using the transaction monitoring services, with no discrepancy found in between the addresses and the data provided by the customers.
- *Speed:* Helps in efficient tracing and discovering of the insights.
  - *i Global Intelligence Obligations:* The company has a team collecting ground-truth attributions daily.
  - *ii Dynamic Token Support:* Supporting the tokens deployed to a chain within seconds.

C. Blockchain Loyalty and Rewards Programs

An important factor bolstering any business's profit is its loyal customers. Reward and loyalty programs act as the keystone of customer retention strategies. From personalized discounts to tiered rewards to point systems, customer loyalty programs help differentiate the brands and create lasting customer relationships across various industries [24]. With technological advancements and evolution in customer expectations, conventional programs are witnessing a shift towards blockchain-based solutions, acting as a value-based service for business operations [25]. This can be illustrated with an example as:

Loyyal

About the Company: Loyyal, headquartered in San Fransico, California, was founded in 2014. The company is the most cutting-edge, Blockchain-as-a-Service platform serving the challenging commercial needs of the loyalty industry for growth and scale at the greatest speed and the lowest probable cost.

Solutions Offered: Loyyal leverages smart contracts and blockchain technology to reinvent the way incentives are created, managed, and rewarded. With blockchain-as-a-service platform, the solutions offered help empower loyalty across different verticals such as retail, banks, airlines, and others. These can be elaborated as:

- *Access Point:* Access point systematizes, and simplifies the partner and merchant management. It enables the loyalty program owners to recruit several offers and rewards compared to their previous ability without any extra internal resources, and further create more enriching rewarding experiences. These self-service operating systems, enable the partners and the merchants to engage with the owners, thereby, seamlessly making the latest connections and unleashing the latest sales channels. The features include:
  - *i Centralized Partnership Management:* The access point centralizes all the reward and merchant partner management, offering a unified view of them, whether managing agreements or conveniently tracking the performance.
  - *ii Plug-and-Play Simplicity:* The plug-and-play solutions are designed for seamless

integration with multiple partners throughout the intuitive operating system, minimizing the complicated setup processes.

iii *Low-Cost Solution:* Affordable solution to manage several commercial relationships. Ensuring that program owners can optimize their partner network without requiring additional resource.

iv *Incremental Revenue:* Unlimited partner connections through a single integration, enables the program owners to focus on the latest revenue-generating partners, overcoming the internal barriers and minimizing friction around administration, onboarding, and approvals.

v *Efficient Resource Allocation:* With streamlined partnership management, effective resource allocation helps minimize the effort and time spent on administrative tasks. It further enables the program managers to emphasize their valuable resources and time on strategic planning and nurturing relationships.

The solutions offered thereby, help drive greater engagement and connectivity across every program via varied Web 3.0-based applications.

- *Xpand Point:* The blockchain-powered engine helps transcend the way loyalty points are exchanged across varied programs. Xpand Point leverages blockchain technology to deliver an efficient, transparent, and secure loyalty points exchange platform. Features of Xpand Point:
  - *i* Seamless Onboarding
  - *ii* Contract Management
  - *iii* Real-Time Processing
  - *iv* Interoperability
- *Reward Point:* The reward point is the reward module, wherein an inclusive range of rewards (such as travel, lifestyle, and leisure rewards) is included.

VI. Challenges in the Implementation of Blockchain in the Financial Services

Along with the benefits of blockchain technology in financial services, its implementation also has challenges. A few of these include:

- *Scalability:* The rising demand for blockchain has led to substantial scalability challenges, transaction latency, and making the system less efficient, hence, impeding their extensive utility and adoption [26]. For businesses emphasizing high throughput, low latency, and quicker transactional processes, scalability poses a major issue [27].
- *Interoperability:* With different blockchain networks in presence, each with its own set of standards and protocols, attaining interoperability poses to be an intimidating task [28]. Lack of interoperability and standardization between different blockchain platforms are anticipated to lead to inefficiencies and hinder data exchange and seamless communication between various systems in the financial services industry.
- *Regulatory Framework:* The rapid rise of blockchain in the financial sector is anticipated to create a lot of unpredictability in regulatory and legal frameworks. The lack of synchronized regulatory frameworks across the jurisdictions leading to an uncertain environment for the investors and businesses operating in the blockchain scope. The absence of a synchronized regulatory framework is expected to lead to regulatory arbitrage, wherein the companies focus on looking for jurisdictions with regulatory milieus for their blockchain-based projects. In addition to this, the need for larger regular coordination and clarity at both international and national levels is gaining significant momentum [29]. Creating clear legal documentation and effective governance plays a vital role in catering to the challenges in blockchain technology by defining rights,

resolving disputes, and delineating decision-making processes.

- *Lack of Awareness:* One of the major challenges in adopting blockchain technology is the lack of awareness, understanding, impact, and areas of technology usage in the respective areas [30].

VII. Conclusion

The integration of blockchain in financial services is not just a trend, but also a revolution set to reshape the way businesses manage digital and tangible assets, with the right regulatory frameworks [31]. With the regulatory bodies becoming more acquainted with its potential, the impact of this pioneering technology is projected to witness a substantial rise.

Blockchain technology has significantly transformed the status quo of businesses across different verticals, guaranteeing transparent, decentralized, immutable, and secure data services. With the adoption of this technology, counterparty risks can be minimized, costs for issuers can be reduced, and the market for investors can be distended [32].

The article delves into the current status, the potential, and the need for blockchain-enabled financial services and applications in businesses. In addition to this, the paper also focuses on the impacts and the use cases of blockchain in the finance industry. The study focuses on the factors underpinning the decisions of financial institutions to adopt blockchain, thereby, enabling the companies to collaborate and develop subsequent strategies and models for their organizations. The Blockchain-based systems empower tailored, cost-effective, and faster issuance of digital securities. In addition to this, the technology helps minimize transactional costs, amplify the transactional scope, augment peer-to-peer transactions, and subsequently generate a new paradigm for decentralized business models. By generating a distributed digital ledger that eternally records transactions in an immutable way, blockchain enables financial institutions to drift away from centralized intermediaries and databases [33]. To stay competitive, it is hence, essential for businesses to make the right investment at the right time, in order to augment their financial services in pace with this ground-breaking technology.

References

[1] J. Allaire, "Blockchain is in from the cold — and stablecoins are set to change the financial system forever," Jan. 15, 2024. [Online]. Available: https://www.weforum.org/stories/2024/01/blockchain-change-world-finance-stablecoins-internet/

[2] R. Valis, "How to increase your revenue using blockchain technology in 2024," Apr. 21, 2024. [Online]. Available: https://dashdevs.com/blog/how-to-increase-your-revenue-using-blockchain-technology-in-2023/

[3] R. de Moll, D. Griswold, D. Hobbs, L. Pawczuk, and P. Podhorecki, "Unleashing blockchain in finance," Mar. 2019. [Online]. Available: https://www2.deloitte.com/us/en/pages/finance/articles/unleashing-blockchain-in-finance.html

[4] Consensys, "Blockchain in financial service: Blockchain use cases," n.d. [Online]. Available: https://consensys.io/blockchain-use-cases/finance

[5] IBM, "Blockchain for financial services," n.d. [Online]. Available: https://www.ibm.com/blockchain/industries/financial-services

[6] S. Plummer, "Blockchain technology disrupts the financial sector: Insights into 2024 trends," Mar. 29, 2024. [Online]. Available: https://www.financemagnates.com/cryptocurrency/blockchain-technology-disrupts-the-financial-sector-insights-into-2024-trends/

[7] I. Adel, "How blockchain is transforming the entire financial services industry," Jun. 7, 2023. [Online]. Available: https://www.forbes.com/sites/forbestechcouncil/2023/06/07/how-blockchain-is-transforming-the-entire-financial-services-industry/

[8] J. Gugl, "Institutional blockchain adoption — The top 30 list of global banks on the forefront," Jun. 27, 2024. [Online]. Available: https://medium.com/riva-markets/institutional-

blockchain-adoption-the-top-30-list-of- global-banks-on-the-forefront-24605fb3a1ed

[9] M. Kowalski, Z. W. Y. Lee, and T. Chan, "Blockchain technology and trust relationships in trade finance," *Technological Forecasting and Social Change*, vol. 166, May 2021.

[10] H. Hyvärinen, M. Risius, and G. Friis, "A blockchain-based approach towards overcoming financial fraud in public sector services," *Business & Information Systems Engineering*, vol. 59, pp. 441-456, 2017.

[11] J. A. D. Mendoza, "Blockchain technology: Powering the transformation of asset management," Sept. 23, 2024. [Online]. Available:https://www.aima.org/article/ blockchain-technology-powering-the- transformation-of-asset-management. html#:~:text=By%20fundamentally%20 changing%20how%20assets%20are%20 handled%2C%20tracked%2C,avenues%20 for%20distribution%2C%20and%20 improves%20supply%20chain%

[12] J. Bramblet, "Ultimate guide to blockchain in insurance," 2024. [Online]. Available: https:// insuranceblog. accenture. com/ ultimate-guide-to-blockchain-in-insurance

[13] R. L. Buenfil, and A. H. Romanowski, "Implications of decentralized finance for competition, antitrust policy and economic growth," Dec. 1, 2022. [Online]. Available: https://www.bakerinstitute.org/research/ implications-decentralized-finance-competition-antitrust-policy-and-economic- growth

[14] A. Clere, "CBDCs: 15 central bank digital currencies by 2030, BIS says," Jul. 10, 2023. [Online]. Available: https://fintechmagazine. com/articles/cbdcs-15-central-bank-digital- currencies-by-2030-bis-says

[15] T. Tamplin, "Blockchain payment system," Jan. 12, 2024. [Online]. Available: https://www. financestrategists.com/wealth-management/ blockchain/blockchain-payment-system/

[16] Cryptomus, "Blockchain in payments: Blockchain payment processing," Jun. 28, 2023. [Online]. Available: https://cryptomus. com/blog/blockchain-payment-gateway-all- you-need-to-know

[17] T. F. T. Report, "JPMorgan Chase goes blockchain: Embraces programmable payments for B2B clients," Nov. 22, 2023. [Online]. Available: https://thefinancialtechnologyreport. com/jpmorgan-chase-goes-blockchain-embraces-programmable-payments-for-b2b-clients/#:~:text=JPMorgan%20Chase%27s%20 Onyx%20unit%20has%20introduced%20pr-ogrammable%20payments,distributed%20 ledgers%20to%20automate%20payment

[18] J. Adams, "How JPMorgan is using blockchain to make B2B payments 'Programmable'," Nov. 11, 2023. [Online]. Available: https:// finance.yahoo.com/news/jpmorgan-using- blockchain-b2b-payments-185753686. html?fr=sycsrp_catchall

[19] J. Chase, "Transforming the way money, information and assets move around the world," 2020. [Online]. Available: https:// www.jpmorgan.com/onyx/index

[20] T. Tamplin, "Blockchain in money transfers | Definition, benefits, challenges," Dec. 18, 2023. [Online]. Available: https://www. financestrategists.com/wealth-management/ blockchain/blockchain-in-money-transfers/

[21] Vegavid, "How blockchain is used in money transfer?," 2024. [Online]. Available: https://vegavid.com/blog/ how-blockchain-is-used-in-money-transfer/#

[22] Chainalysis, 2014. [Online]. Available: https:// www.chainalysis.com/company/

[23] Chainalysis, 2014. [Online]. Available: https:// www.chainalysis.com/blockchain-intelligence/

[24] Antiersolutions, "How blockchain loyalty programs are revolutionizing customer retention in 2024?," 2024. [Online]. Available: https://www.antiersolutions. com/how-blockchain-loyalty-programs-are-revolutionizing-customer-retention-in-2024/

[25] Deloitte, "Making blockchain real for customer loyalty rewards programs," [Online]. Available: https://www2.deloitte.com/

content/dam/Deloitte/us/Documents/financial-services/us-fsi-making-blockchain-real-for- loyalty-rewards-programs.pdf

[26] C. Meijer, "Blockchain and the scalability challenge: Solving the blockchain trilemma," Sept. 26, 2023. [Online]. Available: https://www.finextra.com/blogposting/24941/blockchain-and-the-scalability-challenge-solving-the-blockchain-trilemma

[27] J. Günthner, "The role of blockchain in finance," Apr. 6, 2023. [Online]. Available: https://www.paltron.com/insights-en/the-role-of-blockchain-in-finance

[28] J. Sharma, "Blockchain technology adoption in financial services: Opportunities and challenges," in *Revolutionizing Financial Services and Markets Through FinTech and Blockchain*. USA: IGI Global, 2023, pp. 99-117.

[29] J. Cole, "Regulatory challenges facing blockchain technologies," Apr. 16, 2024. [Online]. Available: https://blockapps.net/blog/regulatory-challenges-facing-blockchain-technologies/

[30] B. Karanjia, S. Lakshman, and S. Goswami, "Blockchain technology in India: Opportunities and challenges," Deloitte, 2017.

[31] sidbi, "Blockchain banking: Revolutionizing finance," n.d. [Online]. Available: https://www.psbloansin59minutes.com/knowledge-hub/blockchain-banking-revolution-india

[32] Mohd. Javaid, A. Haleem, R. P. Singh, R. Suman, and S. Khan, "A review of blockchain technology applications for financial services," *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, vol. 2, no. 3, p. 100073, 2022.

[33] A3Logics, "Blockchain in finance: Benefits, use cases, and challenges," Oct. 11, 2024. [Online]. Available: https://www.a3logics.com/blog/blockchain-in-finance/

[34] R. Choudhury, and B. Weisman, "Digital assets, distributed ledger technology and the future of capital markets," *World Economic Forum: In Collaboration with Boston Consulting Group*, Geneva, May 2021.