Predictive Data Analytics in Healthcare

Thesis Submitted
By

Priyanka Saha

Doctor of Philosophy
(Engineering)

Department of Computer Science and Engineering
Faculty Council of Engineering and Technology
JADAVPUR UNIVERSITY
KOLKATA, INDIA

2024

Index no: 001/18/E

1. **Title of thesis:** Predictive Data Analytics in Healthcare

2. Name, Designation, and Institution of the Supervisor:

Dr. Sarmistha Neogy
Professor
Department of Computer Science and Engineering
Jadavpur University, Kolkata
West Bengal, 700032, India

3. List of Publications

i. Conference Publications

a) **P. Saha,** S. Patikar and S. Neogy, "A Correlation - Sequential Forward Selection Based Feature Selection Method for Healthcare Data Analysis," 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2020, pp. 69-72, doi: 10.1109/GUCON48875.2020.9231205.

b) **Saha, P.,** Neogy, S. (2022). An Overview of Pulmonary Tuberculosis Detection and Classification Using Machine Learning and Deep Learning Algorithms. In: Dua, M., Jain, A.K., Yadav, A., Kumar, N., Siarry, P. (eds) Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-5747-4_72

c) **P. Saha,** "An Ensemble CNN-Dempster Shafer based tuberculosis detection from chest x-ray images," 2022 IEEE Calcutta Conference (CALCON), Kolkata, India, 2022, pp. 228-232, doi: 10.1109/CALCON56258.2022.10060463

d) S. Patikar, **P. Saha,** S. Neogy and C. Chowdhury, "An Approach towards prediction of Diabetes using Modified Fuzzy K Nearest Neighbor," 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2020, pp. 73-76,doi: 10.1109/GUCON48875.2020.9231066.

e) Patikar, S., **Saha, P.,** Neogy, S., Chowdhury, C. (2022). Review of Body Vitals Monitoring Systems for Disease Prediction. In: Biswas, S., Chowdhury, C., Acharya, B., Liu, CM. (eds) Internet of Things Based Smart Healthcare. Smart Computing and Intelligence. Springer, Singapore. https://doi.org/10.1007/978-981-19-1408-9_9

ii. Journal Publications

Published

a) **Saha, P.,** Neogy, S. Concat_CNN: A Model to Detect COVID-19 from Chest X-ray Images with Deep Learning. *SN COMPUT. SCI.* **3**, 305 (2022). https://doi.org/10.1007/s42979-022-01182-1

Communicated

a) **Saha, P.,** Neogy, S. A correlation Whale Optimization Based Feature Selection approach for Parkinson Disease Classification. Submitted to Springer Journal of Medical Systems

b) **Saha, P.,** Neogy, S. A Gabor filter based Nested U-Net approach for Breast MRI segmentation. Submitted to Journal of Imaging Informatics in Medicine.

4. List of Presentation in International Conference

a) **P. Saha,** S. Patikar and S. Neogy, "A Correlation - Sequential Forward Selection Based Feature Selection Method for Healthcare Data Analysis," 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON), Greater Noida, India, 2020, pp. 69-72, doi: 10.1109/GUCON48875.2020.9231205.

b) **Saha, P.,** Neogy, S. (2022). An Overview of Pulmonary Tuberculosis Detection and Classification Using Machine Learning and Deep Learning Algorithms. In: Dua, M., Jain, A.K., Yadav, A., Kumar, N., Siarry, P. (eds) Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-5747-4_72

c) **P. Saha,** "An Ensemble CNN-Dempster Shafer based tuberculosis detection from chest x-ray images," 2022 IEEE Calcutta Conference (CALCON), Kolkata, India, 2022, pp. 228-232, doi: 10.1109/CALCON56258.2022.10060463

5. List of patents: None

Statement of Originality

I, Priyanka Saha registered on 4th June 2018 do hereby declare that this thesis entitled **"Predictive Data Analytics in Healthcare"** contains literature survey and original research work done by the undersigned candidate as part of Doctoral studies.

All information in this thesis have been obtained and presented in accordance with existing academic rules and ethical conduct. I declare that, as required by these rules and conduct, I have cited and referred all materials and results that are not original to this work.

I also declare that I have checked this thesis as per the "Policy on Anti Plagiarism, Jadavpur University, 2019", and the level of similarity as checked by iThenticate software is 6 %.

Signature of Candidate: Priyanka Saha

Date: 28.2.24

Certified by Supervisor:
(Signature with date, seal)

Sarmistha Neogy 28.02.2024

Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata-700032

Dr. Sarmistha Neogy

Professor

Department of Computer Science and Engineering,

Jadavpur University, Kolkata, 700032, India

Certificate from the Supervisor

This is to certify that the thesis entitled **"Predictive Data Analytics in Healthcare"** submitted by **Smt. Priyanka Saha,** who got her name registered on **4th June,2018,** for the award of Ph.D. (Engineering) degree of Jadavpur University, is absolutely based upon her own work under the supervision of **Dr. Sarmistha Neogy** and that neither her thesis nor any part of the thesis has been submitted for any degree/diploma or any other academic award anywhere before.

1. Sarmistha Neogy. 28.02.2024

Professor
Computer Sc. & Engg. Department
Jadavpur University
Kolkata-700032

Dr. Sarmistha Neogy
Professor,
Department of Computer Science and Engineering,
Jadavpur University,
Kolkata,700032, India
(Supervisor)

Acknowledgement

First and foremost, I extend my extensive gratitude and heartfelt thank you to my esteemed supervisor, Dr. Sarmistha Neogy, Professor, Department of Computer Science and Engineering, Jadavpur University, Kolkata. Her unwavering guidance, dedicated assistance, invaluable insights, and continuous encouragement throughout my research period have been instrumental in steering and completing this research, contributing significantly to the fulfillment of my Doctor of Philosophy. Without her encouragement, enthusiasm, profound perspectives on research, and unwavering support this journey towards my Ph. D would not have been possible.

I want to convey my earnest thank you to Dr. Nandini Mukhopadhyay Madam, Professor, HOD, Department of Computer Science and Engineering Department, Jadavpur University, Kolkata and Dr. Samiran Chattopadhyay Sir, Vice Chancellor, Techno India University, Kolkata, for their constant support and encouragement in this work.

I am sincerely grateful to the other members of my thesis committee for their invaluable and constructive feedback, which has substantially enhanced the quality of the thesis.

I am also thankful to Jadavpur University for providing me with such wonderful work environment to illuminate my knowledge quest and help me groom my academic and professional life. I am thankful to Computer Science and Engineering Department for providing me with the opportunity and resources to carry out the research work. All faculties, staff, research scholars, and students in this department have helped me a lot in this journey. Juniors like Bidisha, Arundhati, Srabani, Manjarini, Asif, Sounak are amazing. They helped me a lot in carrying out my work. I am also grateful for the fellowship I received from State Fellowship, government of West Bengal, India during the span of my work.

I am deeply thankful for my husband, Joydeep, son Riddhideep daughter Adyatrayee for their unwavering support and to my parents and my in-laws to encourage me and support me in my hard times. Without my family support, it is not possible for me to complete my thesis. Finally, I want to thank all my family members, friends, near and dear one for their constant support throughout my journey.

Dedicated to my Parents, Husband, Son, Daughter...

ABSTRACT

Over the last decade, AI has become an integral part of many industries specially in Healthcare Industry. The intention of AI inventions was to replicate human intelligence and aims to reduce human workload. A highly intelligent Computer Aided System will help healthcare professionals in faster decision making, optimize solution, and cost-effective system. At the same time, AI focuses on eliminating unnecessary data and make prediction based on important features. Medical data can be analyzed in numerous ways in different labels. Data Analytics is a procedure to analyze raw data and made decision based on hidden information in the data. Predictive analytics make decision about future from previous knowledge. Predictive Data analytics has promising capability of unearthing meaningful insight from data and use the information for predicting future events. Therefore, utilizing the strength of Machine Learning (ML) and Deep Learning (DL), the aim of this work is to classify disease (Parkinson, Tuberculosis, COVID19, Breast Cancer, Bone Fracture) and predict the possible outcome for fast disease detection. The objective of the work is to sustain the quality and quantity of important healthcare data for patient's health status. The work focuses on determining the important features of patient data that will reflect the prediction accuracy. Here, an optimized and evidential based predictive approach is proposed for classifying patient's data correctly. The work also addressed data imbalance problem and how to mitigate it. The work here is carried out in three steps. First, the study investigates the strength of machine learning in Parkinson disease classification different by applying two feature selection methods and the effect of feature selection on enhancing the predictive nature of machine learning models. Secondly, the important aspects of deep learning models are explored, and the efficiency of deep learning approaches on predicting infectious diseases from patient's image data is carried out. Next, section discusses an image segmentation-based approach for proper identification and enhance the disease prediction. The focus of this work is to check these approaches for real time data flow where decision is required to made fast. This work explores different aspects of ML and DL application in healthcare disease data. Designing of more robust models for disease prediction will lead to better utilization of resources and will help in proper analysis of data. This understanding will provide a flow the current research in a better proportion. So, this work can be expanded to predict any disease where feature selection is important for fast detection.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Over the last decade, massive research has been done on Artificial Intelligence (AI) in the healthcare industry. The aim is to increase its ability to address key challenges like treatment Quality, Treatment Cost, Patient Wellbeing, among others. The purpose of AI invention was to replicate human intelligence and reduce human workload with Artificial Operators.AI was introduced as a technology to make better decisions in uncertain and complex environment [1]. It is felt that AI has the potential to transform the Healthcare Infrastructure. AI shows a way to elaborate and augment human thinking power, decision making ability, and human creativity power [2]. So, developing a highly intelligent Computer Aided System will helps healthcare professionals optimize their decision-making process. It will provide them by proving them a faster and easier information interpretation system that is necessary to analyze critical information. At the same time, it will eliminate unnecessary data or scenarios. Over the years AI has been through series of transformations from early rule-based system to current Machine Learning and Deep Learning Based systems. AI has evolved since its inception in 1951 by Christopher Strachey.

Two main pillars of incorporating Information Technology with Healthcare Industry are Electronic Health Records [3] and Decision-Making Systems [4]. Information mining in healthcare can significantly increase patient quality of life, boost the efficiency of AI systems, thereby helping in saving numerous patient lives. Healthcare data can be generated and stored in different forms like, Electronic Health Record (EHR), pharmacy details, ubiquitous digital sensors across whole population, pathologic lab tests, diagnostic information, and so on, as shown in figure 1.1.

Medical data can be analyzed in numerous ways at various levels. While processing Electro Cardiography (ECG) data, a conventional alert system can be introduced to alert about health condition of the patient if parameters are out of range. Data from different resources can be used for merging and then fed onto a system to generate diagnostic and inference based on preset rules of the system.

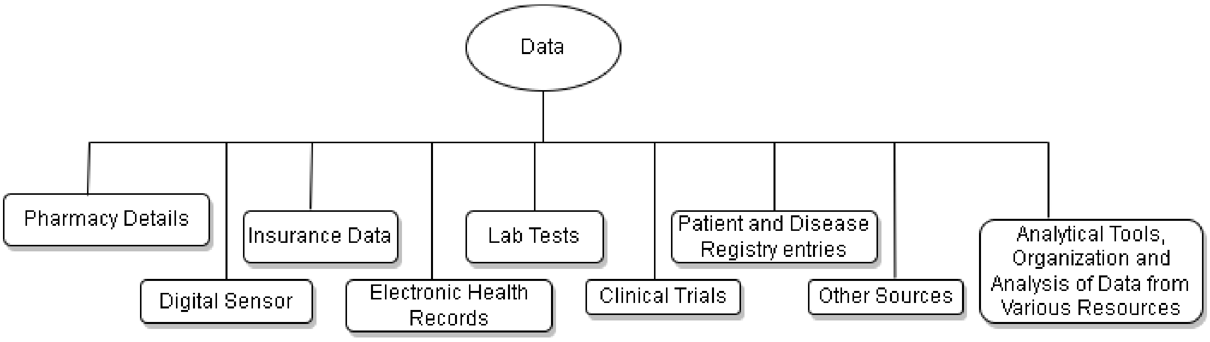


Figure 1.1: Different Data Resources in Healthcare Industry

These systems are built in tree like hierarchy and helpful for finding a realistic explanation for the symptoms. These rule-based systems are called Expert systems. Expert systems are seen as like human decision-making process where knowledge is gained from previous learning. Different research is carried out to make this system more reliable and effective. The aim of these expert systems is to generate early diagnosis and fast treatment. Though there are some ambiguities about the exact role and influence of AI in healthcare, an amalgamation of trust, privacy, and security of data will have a positive impact on human health [5].

1.1Data Analytics – Contribution in Decision Making

One major issue of implementing AI in healthcare is comprehensive understanding and unfolding of medical traits using ML and DL [17]. Data mining is a process to analyze big set of information to find relevant patterns and use these patterns for future prediction or finding the probability of future events. Data Analytics is a process to analyze raw data and making conclusion based on facts while analyzing the data. Data Analytics is a vast area that incorporates different types of data analysis. It is a way to unearth trends and metrics that would otherwise not be possible to analyze from the vast sea of information. Data analytics helps to optimize performance and thus has huge impact on healthcare industry. It provides way to reduce cost and store large amount of information. The process of data analytics is divided into different steps:

Step 1: First step is to understand the data, how data is grouped based on some criteria such as disease type, patient type, demographic area, etc.

Step 2: The second important step is to collect data from different sources like personal data, online resource, hospitals, or healthcare center.

Step 3: The data organization is a vital part for any analysis to start. Unorganized data will lead to confusion and delay the process further. So, data processing and storing in proper format is utmost important for effective result.

Step 4: Ambiguity in data will affect the result. Data is needed to be cleaned, thereby eliminating duplicate, missing values, or error.

The type of result and information from data depends on the type of analysis that has been done on the data. In Data Science, there are four types of data analysis, such as Descriptive, Prescriptive, Diagnostic and Predictive.

1.1.1 Descriptive Analytics

Descriptive Analytics can be defined as a statistical interpretation. It is applied to analyze historical data to recognize hidden patterns and relationship among these patterns. It helps to describe past events, phenomenon, or past outcome [15]. By analyzing past events or outcome, it provides industries or business a perfect scenario based on past trends. This analysis is best suited in financial statement analysis, social media engagement, generating financial reports.

1.1.2 Prescriptive Analytics

It is a variation of analytics which suggests decision based on future opportunities and helps to mitigate future risks [13]. It prescribes the best possible solution to take while considering the whole future possibilities [6]. Application areas of Prescriptive Analytics are fraud detection, product development and improvement, algorithmic content creation, investment decision.

1.1.3 Diagnostic Analytics

Diagnostic Analytics, as its name suggests, is used to find fundamental causes of problems. It tries to find out why the problem has occurred [14]. It often infers information from descriptive analytics for what occurred before analyzing and why it occurred. Retail and manufacturing industry, human resource and health sector are some major implementation areas of diagnostic analytics.

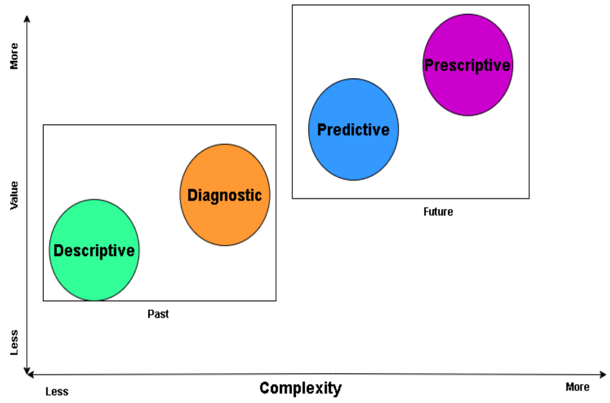


Figure 1.2: Types of Data Analytics based on Value and Complexity

1.1.4 Predictive Analytics

Predictive analytics is a new form of advanced analytics which makes prediction about future, based on analyzing old or previous knowledge. It uses statistical interpretations, data mining techniques, and machine learning techniques for accurate insightful predictions. Using Predictive Data Analytics, complex business problems can be addressed by extracting important and potential information. Application includes Finance, Retail, Healthcare, Manufacturing.

The success of any analytics and its efficacy depends on the techniques and methods used to analyze information and the ability to provide reliable and meaningful insight about the data [7]. Application of data analytics would be beneficial for Healthcare Industry in different aspects like better diagnosis of disease, reduction of health cost, improvement in patient care, prediction of disease to prevent its spread, less re-admission at hospital, optimizing resource utilization and improved drug supply chain [8]. So, a data driven information analysis system is necessary to quickly learn from everyday clinical care data for a better personalized prediction system.

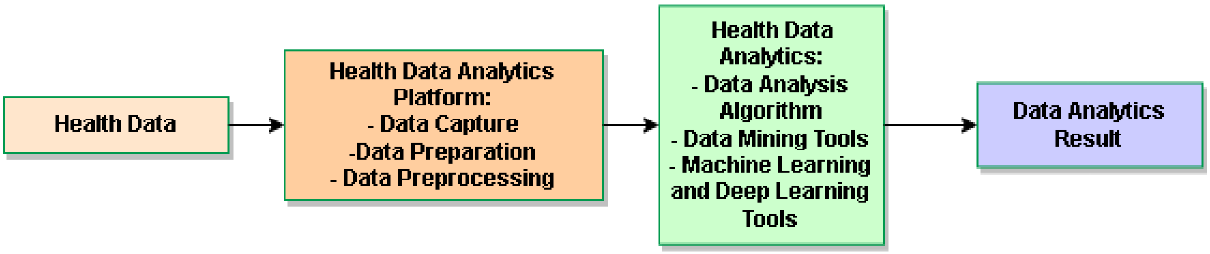


Figure 1.3: Flowof data analytics

1.2Advantages of Data Analytics in Healthcare

Data Analytics has the promising capability of providing meaningful insight from different types of data. It is aimed to help in predicting future events. It emphasizes to reform fraud detection ability, risk reduction, saving life though improved prediction accuracy [9]. Different stake holders like payers, providers, and patients can be benefitted through data analytics and how they can be benefitted are described below.

1.2.1 Healthcare Provider

Healthcare Providers are the main recipients and get maximum help from Data Analytics in Healthcare. Use of Electronic Health Records has made data accessibility easy from any area making it possible to analyze and have maximum advantage of profitability, cost reduction, usefulness without compromising the quality of service. It enables secure data sharing. Data analytics helps to diagnose severe disease well ahead and healthcare providers can plan their services accordingly. Health centers can get the details of every activity, health condition of every patient by analyzing different perspective.

1.2.2 Payer

Using data analytics, payers formulate plan for Health Preventive Programs thus improving patient’s health insurance and quality of life. It allows to structure cost effective medical processing of any disease and its risk factors. It enables to implement preventive measures while informing patients about potential risk factors of the disease.

1.2.3 Patients

Patients are the main pillars of Healthcare Industry. So, providing best solution to them in terms of best healthcare providers and doctors, analyzing the treatment procedure and its effectiveness, medicine’s effectiveness, and quality check, and helping to pick the best service is the main aim of Data Analytics.

1.3Opportunities of Data Analytics in Healthcare

Healthcare Industry is one of the main beneficiaries of Data Analytics. Since last decade, applying AI tools to analyze patient data and extracting valuable information from them have led tremendous opportunity for doctors and patients [16]. The opportunities healthcare can avail and benefit from Data Analytics are summarized below:

1.3.1 Healthcare Service Improvement

AI driven Data Analytics in healthcare helps doctors with a decision support system while dealing with huge number of patients. From large amount of data, analytics helps to discover certain trends based on any demographic area so that if any pandemic is spread over area, detection and early treatment would be very easy to implement. Disease detection become easier, effective, and cost-effective leading to improvement in health and lifestyle. Data analytics is also very much popular to analyze human genome which can mitigate numerous genetic diseases well ahead.

1.3.2 Additional Support to Healthcare Professional

In many rural and remote areas of India, the number of doctors and treatment equipment are far less than actual required. So, a decision support system would be helpful to detect disease at early stage and taken care of. Doctors can have a clear picture of medical history of each patient for better diagnosis and treatment. Detection of any epidemiological risks can be done early, and an improved control system will control reaction rates and pathogenic spots. Patients at risk of any life-threatening disease can be treated on emergency basis. Data analytics will also help healthcare professional to track the occurrence of any specific disease of a patient, can predict disease progress, estimate the risk factors, can investigate drug doses, and overall side effects.

1.3.3 Support Scientific Research Activities

Whenever a new drug or vaccine is introduced, it becomes very tedious job to keep track of every drug or vaccine administrated patient for their health status. Data Analytics would help in analyzing these data and predict the efficacy of the drug or vaccine. It helps to detect the side effects of certain drug or vaccine on a group of patients. It helps to identify patients with special biological features that will require some clinical trail in future. Predictive analytics would also help to go through intense research to design better drug and medicines that may be helpful for wide range of patients.

1.3.4 Support on Business and Management

Data analytics helps to locate fast unauthorized financial operations to prevent it from further abusing of resources. It predicts a solution to high treatment cost for any patient. It helps to stop unnecessary medical activities to eliminate errors and prevent abuse.

1.4Challenges of Data Analytics in Healthcare

Medical data by its nature is very complex which requires proper approaches for data collection, data store, data analysis. It also requires presenting a proper interpretation of data to healthcare professionals, business partners and patients. The road to proper health data analytics is not smooth and filled with challenges, including visualization, security, data integrity and other different problems. A few challenges faced by healthcare industry while implementing data analytic program have been identified below. Approaches to mitigate these challenges to achieve a data driven prediction and decision-making system, are also identified.

1.4.1 Data Collection

Data collection requires data to be collected in proper format using proper devices. But unfortunately, in many healthcare service providers, data is collected through some improper data governance habit. Capturing data in clean, accurate, and correct format for use in decision-making systems is a battle to overcome. In a study [10] of ophthalmology clinic, electronic health record (EHR) data and patient disease data match was just 23.5%, where the number of symptoms reported by patient varied from HER data. Poor way of handling EHR data, heavy workflow, and insufficient understanding of why data quality is important leads to low quality and plague data collection. This situation can be improved by prioritizing important data capture routine which can be demonstrated using entitled projects, prioritizing data governance and expert health data management professionals, and developing a clinical documentation program that will assist clinicians how to collect data for proper analysis.

1.4.2 Data Cleaning

Healthcare providers are least bothered about the effect of quality of data on data analysis process. Improper data can lead to misguide a predictive data analytic project. Data cleaning also known as data scrubbing ensures data is correct, accurate, relevant, consistent, and not corrupted through noise. While data cleaning is still done manually, some projects are conducted to scrub data through some automated scrub tools that use certain rules to contrast, compare data. These tools are inspired by sophisticated Machine Learning tools that contributes to rapid advancement, reducing data cleaning time, and provide high level of accuracy and integrity in crucial health data.

1.4.3 Data Storage

As healthcare data size is increasing exponentially every day, some data storage providers are not being able to keep this huge amount of data due to its impact on increasing storage cost. While many organizations have their own data storage capacity which provide controlled access, security, and on time data retrieval process, an on-site storage network may incur expensive cost to maintain and provide data across healthcare departments. Cloud storage is a popular option as it is cost effective and reliable. A large number of healthcare industry keep their data on cloud-based storage Infrastructure [11]. Modern cloud servers are equipped with disaster recovery system, easier fetching option through organizations and careful in terms of data security that understands the importance of different security and privacy acts across the world [12]. Many organizations now a days rely on a hybrid approach for data storage that are more flexible according to their varying needs.

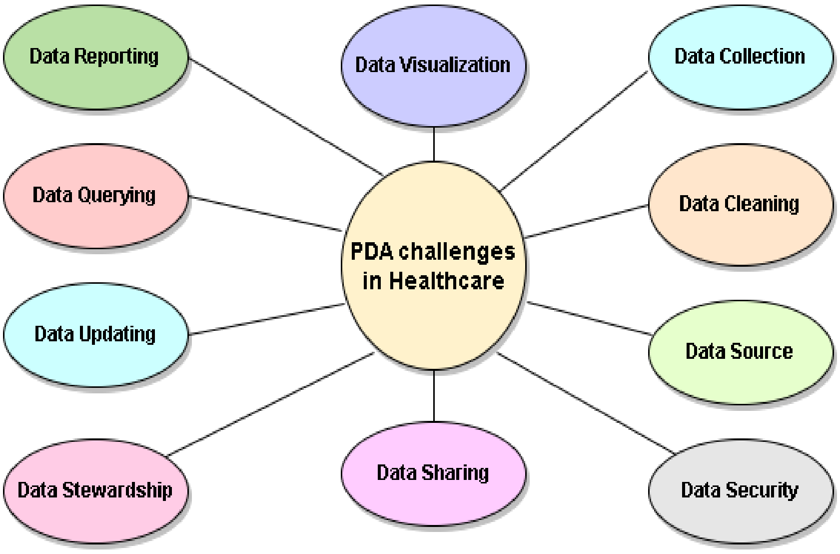


Figure 1.4: Challenges of predictive data analytics in healthcare

1.4.4 Data Security

Data security is the most important aspect for any healthcare organization especially in the era where data breaches, data hack are common. Different attacks include phishing, malware, accidental exposure of data in public space, to name a few. Data is exposed to infinite number of threats and vulnerabilities. The HIPAA Security Rule [12] discusses a list of safeguards for health organization data that include some including transmission security, authentication protocols, data integrity. These safeguards provide security aspects using firewalls, encrypting sensitive data, using multi-level authentication process, and anti-virus software.

1.4.5 Data Querying

Proper data query is utmost important for expecting the right answer to a question. Data querying is a fundamental step for data analytics and data reporting, Healthcare industry must overcome some data query challenges to proceed a meaningful data analysis process. Healthcare query system is required to fetch the intended information from the system rather than fetching the entire repository data. This process will not occupy large amount of data and will provide a complete picture of individual patient's health. Many healthcare organizations use Structured Query Language (SQL) to query into large relational datasets. This process is effective for accuracy and standardization of the data.

1.4.6 Data Stewardship

Healthcare data has a long life because for prolonged diseases, previous data is required for treatment purpose. Data understanding like why, when, how, by whom is important for data analysis. So, a metadata-based data governance is required. Healthcare industry should maintain a data steward to manage the metadata for development A data steward maintain all data elements in standard definitions, formats, size, and documents appropriately for data query purpose.

1.4.7 Data Reporting

After successful data query, an effective result of query generation procedure is required in clear concise for concerned query. Data report quality depends on the accuracy and integrity of data. Poor data will lead to misleading report, which can have detrimental effect on patient's health. Reporting is a prerequisite for data analysis. Healthcare organizations should carefully plan to review their report to ensure that database administrators can generate important information that is required for certain patient's health scenario.

1.4.8 Data Visualization

A proper data visualization can make it easier for doctors to understand the information and use it accordingly. Color coding is a popular data visualization tool that provides immediate response. So, healthcare organizations should practice a well-crafted data presentation practice, so that it can illustrate the important aspects of data. Low quality graphics, overlapped texts, convoluted flowcharts, can mislead the recipient leading to misinterpretation of data. Some data visualization tools are Pie Chart, Histogram, Bar Charts, Heat Maps.

1.4.9 Data Updating

Many health parameters like blood pressure, blood oxygen saturation level, hemoglobin level, change frequently according to patient's health status. But information like address, phone number, age, gender may not require change or frequent change. Understanding the behavior of data, it is a challenge for any organization to monitor these changes at regular intervals and record these changes in datasets. So, an automated updating process to manage the integrity of data without damaging the quality of data will be helpful. Duplicate records should be carefully handled as it may affect the decision-making process.

1.4.10 Data Sharing

Data sharing with external users is essential as it is required for overall population health management system. Data interoperability is a major concern for organizations of all shape and size. Healthcare industry need to cross the barrier of data sharing using data sharing tools and strategies.

1.5 Motivation

In this thesis, the work concentrates on prediction of disease classification based on machine learning and deep learning approaches on healthcare disease data. In healthcare, disease is considered as a particular abnormal condition that affects structure of human body parts. Disease is also referred to as a health condition that are associated with certain signs and symptoms. Disease can be categorized as infectious disease, hereditary disease, deficiency disease, and physiological disease. In the context of predictive data analytics in healthcare, this work essentially focuses on identifying disease from a patient's medical data and predicting its class based on the classification result. Different machine learning and deep learning models are to be investigated on patients' medical records. The overall healthcare disease classification models need to handle huge data, preprocess the data, generate proper classification result based on patients' health data and many more aspects. On the other hand, predictive analysis model should be able to generate faster result to help healthcare professionals for timely and correct treatment. Predictive nature of Machine learning and deep learning applications plays a crucial role for monitoring public health. It encompasses three major key aspects of population health such as-

- **Identifying Chronic Disease**

  Through predictive analytics, people can be identified and treated before any chronic condition can arise. It is therefore some analytical measure that categorizes patients based on characteristics like demographic, heredity, age, past concerns, and many others.

- **Identifying Pandemic Outbreak**

In recent time, predictive analytics was proved as boon and has shown its capability to identify pandemic outbreak COVID19. Using the power of predictive analytics a Canadian Company Bluedot identified the unusual pneumonia cases in Wuhan, China, on 30^{th}December 2019 before formal announcement from WHO [19].

- **Monitoring Critical patient's Health Status**

Monitoring their vitals in critically ill patients is an important task. Predictive analytics provides a great help in constant monitoring of these patients. If there are any changes in vital, it helps by indicating signal and making more informed decision.

- **Improved Patient Engagement**

With predictive analytics, a patient's noncompliance is detected well in advance, and proper measures can be described to keep patient healthy.

1.6 Objective of the Present Work

It is very much evident from the discussion that predictive analysis is very important part of healthcare. Getting some useful insight from patient data through some predictive analytical process is very beneficial. Therefore, keeping the usefulness of both machine learning and deep learning in mind, predictive models are important aspect nowadays. For that, a rigorous analysis of ML and DL applications are to be carried out. The predictive models can be proposed based on research study.

The work presented in the thesis has introspected different aspects of predictive analysis based on machine learning and deep learning and discusses the following requirements. The objectives are:

1. Maintaining the quality and quantity of healthcare data related to patient's health status.
2. Determining important features of the patient data that will reflect prediction quality.
3. Designing optimization and evidence based predictive models so that patient's health data can be correctly classified.
4. Balancing the data quantity for every classification problem to cope up with data imbalancing problem.

The objective of this research has been carried out in three steps. Initially, the study investigates the strength of machine learning in disease classification analyzing different feature selection methods and the effect of feature selection on enhancing the predictive nature of machine learning models. Secondly, the important aspects of deep learning models are explored, and the efficiency of efficiently deep learning approaches on predicting infectious diseases from patient's image data is carried out. Next, an image segmentation-based approach is furnished to properly identify and enhance the disease prediction.

1.7 Thesis Organization

The thesis is organized as follows. **Chapter 2** presents a literature survey of Machine Learning and Deep Learning based disease classification. The chapter is divided into two parts: First part discusses about the previous studies and research on how machine learning based implementations classify patient data based on their class levels. In second part, different research works based on deep learning applications are discussed. It focuses on deep learning supervised applications proposed in the literature. **Chapter 3** focuses on machine learning based supervised disease classification approaches. Feature selection is an important part for any machine learning based implementations. As relevant features have direct impact on machine learning model performance for any classification problem, different feature selection-based approaches are discussed for disease classification. The proposed models are compared with state-of the-art models and existing approaches. For experimental purpose, Parkinson disease is considered. **Chapter 4** considers deep learning based supervised disease classification approaches from image data. Image data should be noise free and its pixel intensity largely affects the model's performance. In this chapter, data imbalacing problem is addressed and images are preprocessed for contrast enhancement. Image classification approaches based on concatenation of different feature extractors of deep learning methods is discussed. The chapter includes classification of COVID19 images, and a comparative study of performance evaluation is presented with an experimental setup. A deep learning based evidential classifier is presented for tuberculosis disease classification. Performance of a transfer learning models on collected bone fracture image dataset is also evaluated and analyzed. The proposed approaches are compared with existing studies on these image datasets. **Chapter 5** discusses image segmentation-based disease classification approach. Image segmentation simplifies an image representation which becomes more meaningful and easier to analyze. Breast Cancer image segmentation dataset is considered for this study. An edge detection-based image segmentation model is proposed with experimental setup and results. The chapter then discusses the merits and demits of the proposed framework. **Chapter 6** concludes the research works, discusses the issues faced during implementation, and highlights some future scopes and open research issues.

Chapter 2

Literature Survey

Healthcare is a wide area that deals with all sorts of medical services to meet the medical and healthcare emergency demands of people. In healthcare, maintaining and restoring every individual's health by providing proper detection, diagnosis, and prevention of disease by trained and licensed professionals in fields such as medicine, dentistry, public health, clinical psychology. Over the past decade, healthcare industry has emerged as one of the fastest-growing industries with lots of challenges as well as opportunities. Machine Learning (ML) and Deep Learning (DL) technologies, a prominent research field has helped to build system trying to mimic a human intelligence. Machine Learning (ML) and Deep Learning (DL) are inter disciplinary fields that amalgamate fields like Data Processing, Mathematics, Statistics, and Knowledge Analysis. ML extracts data out of data, analyzes the data and discovers patterns. These patterns help to detect and improve the understanding of the disease by medical professionals. The application of ML in healthcare is vast, and it is proving beneficial for both patient and healthcare professionals. Healthcare data analysis includes various types of diseases like diabetes, cancer, cardiovascular diseases, and recent pandemic COVID diagnosis, among others. Decision making, automated hospital management system, providing medical care are some important ML applications in healthcare [20]. Medical care and decision making involve the processes of Medical Image Diagnosis, Behavioral Modification, Data Gathering from Crowd, Improved Radiotherapy, Clinical Trails, Personalized Disease Diagnosis and Treatment [21,22]. Medical Data is vast and include different forms such as clinical data, EHR data, omics data [23], etc. Since last decade, research is being carried out and many ML systems have been proposed for better understanding of data and better decision making. A forecast as mentioned in [24] for the period of 2020 to 2027, the global AI market for healthcare is expected to grow at a CAGR of 4.41 % from 2020 to 2027 to reach $ 51.3 billion by 2027.

In healthcare the amount of data is enormous. For critically ill patients, data generated per day is huge as their vitals are changing over time and huge number data is being generated. So, it is quite impossible for any individual to process the data in a stipulated time and provide proper diagnosis and treatment in time. In recent times Machine Learning, Deep Learning, Big Data, Artificial Intelligence, Pattern Recognition, Cloud Computing can be and are all used together for better healthcare. Using wearable vital measuring sensor devices, data can be transferred to hospital data storage as well as cloud storage. Data stored in cloud can be analyzed further for disease detection [25]. ML helps in data extraction and pattern recognition from massive amount of data. It uses different algorithms to predict future treatment of patients [26]. Numerous research is carried out to show how multi-layer inference algorithms of Machine Learning and Deep Learning improve both efficiency and accuracy in data analysis [27].

In this chapter, the primary goal is to study and explore how ML and DL improve healthcare data metrics and how ML and DL predicts the occurrence of disease from health data. The aim is to evaluate different supervised ML and DL applications, their opportunities, and challenges with respect to healthcare data.

2.1 Objective of Machine Learning and Deep Learning in Healthcare

The origin of Machine Learning was primarily for Computer Gaming and Artificial Intelligence. It was pioneered by an IBM employee, Arthur Samuel in 1959 [28]. It was first proposed as a self-teaching computer. The main concern of Machine Learning algorithm was to develop a program that can be trained on data and use the data for learning purpose. It helps to make necessary prediction and helps to act without explicitly being programmed. ML helps to replace rule-based system by data driven system. This was necessary mainly because the amount of data was growing exponentially. In case of disease analysis, it is necessary to analyze crucial information for critically ill patients. Application of DL is more beneficial for recognizing any problem, and its data driven nature performs better in any type of problem. Thus, ML and DL are more convenient for in-time decision making. Some important applications of Machine Learning and Deep Learning algorithms in healthcare are:

- **Predictive Analysis:** Machine Learning algorithm analyzes data from different sources and predicts possible outcomes such as any chronic disease, presence of tumor, presence of fracture, etc. This helps to analyze and identify high risk patients and helps to take necessary steps to avoid any adverse situation.
- **Population Health Management:** Machine Learning helps to analyze data collected from large demographic area to identify possible disease trends and patterns which can help in public health development initiatives.
- **Diagnosis and Treatment:** Machine Learning algorithms help to analyze medical images such as X-rays, MRI, CT-scan which further helps to diagnose disease and identify appropriate treatment accordingly.

- **Medicine Prediction:** Machine Learning can be used to predict the course of treatment that is most likely to be effective for a patient based on their medical history and genetics.

2.2 Healthcare Data

Healthcare data consists of collection of patient's health information. These datasets contain wide range of information starting from demographic information of the patient, different diagnostic results, patient's medical history, treatment undergone, and medication history. Healthcare data is complex and challenging due to its size and types of information. However, researchers used different machine learning and deep learning algorithms to analyze data and extract important pattern.

The main backbone of Machine Learning and Deep Learning is data. The more relevant and accurate the data, the more accurate predictions can be made. As data plays an important role in ML systems efficiency, different types of comprehensive collections of data play important role in the overall output of the technique. Different types of healthcare data are discussed below.

- **Clinical Data**

Clinical data is collected during the ongoing process of treatment. It consists of laboratory tests, Radiological Imaging test, different allergic test and so on. Different research works have been carried out on clinical data for prediction and detection of critical diseases.

- **Sensor Data**

Sensor data is the outcome of a device that senses, detect, and responds to some type of input from the physical environment. Different sensors are used in Intensive Care Units, Hospital wards, lab equipment, dental tabs, and in home care products. Sensor data can be numerical value, categorical value or can be more complex.

- **Omics Data**

It is a collection of huge complex and high dimensional data. It consists of genomic, transcriptomic, and proteomic data. These types of data are best handled by Machine Learning and Deep Learning algorithms.

- **Genomic Data**

Genomic Data is a collection of data related to structure and function of an organism's genome. It consists of gene expression, DNA data, Copy number Variation, and Sequence number.

- **Transcriptomic Data**

Transcriptomic is a collection of RNA transcript, including coding and noncoding in an individual or a population of cells [42]. These samples are used to analyze and extract, and different datasets are generated.

- **Proteomic Data**

Proteomics is a vast scale study of proteomes. A proteome is a set of proteins found in an organism, system, or biological content. Proteomic data is represented as actual function molecules in the cell. Different works had been carried out by researchers with this data type.

2.2.1 Literature survey on Different types of healthcare data

2.2.1.1 Work on Clinical Data

An ML algorithm was proposed by Tahmassebi et al. for prediction of pathological complete response (pcr) to neoadjuvant chemotherapy using Multiparametric Magnetic Resonance Imaging (mpMRI). It predicts the survival outcome of breast cancer patients [30]. They took 38 women patients with breast cancer and underwent mpMRI test. The proposed model uses 8 classifiers such as support vector machine, random forest, logistic regression, linear discriminant analysis, decision tree, stochastic gradient descent, adaptive boosting, and extreme gradient boosting (XGBoost). Area under curve was evaluated for each feature of pcr. From the results, it was seen that XGBoost produces the best result with higher accuracy for Residual Cancer Burden (RCB) and Disease Specific Survival (DSS) and Logistic Regression for Recurrence Free Survival (RFS). A new improved classification approach for prediction of Hepato Cellular Carcinoma (HCC) patients have been proposed by Kayal et al. [31]. An evaluation was carried out on 165 patients. Out of 49 risk factors 15 risk factors are responsible for HCC. Their experiment on the 165 patients proved that Deep Neural Network has significant higher accuracy than Support Vector Machine (SVM) and Unsupervised KNN.

Zheng et al. proposed a model for detection of type 2 Diabetes Mellitus from Electronic Health Data (EHR). Total 300 patients' data are evaluated, and 114 features are extracted on which different Machine Learning algorithms such as Decision Tree, Random Forest, Logistic Regression, SVM, KNN were applied. From the results it is seen that SVM produces better accuracy of (96%) than other algorithms [32]. Evangelia et al. proposed a classification model based on computer assistant by combining profusion MRI data and convolutional MRI data for diagnosing different types of brain tumors and grading of gliomas. Three classifiers such as Support Vector Machine with Recursive Feature Elimination (SVM-RFE), Linear Discriminant Analysis and K Nearest Neighbor were applied on 102 brain tumor patient data. From their experiment, it is observed that SVM RFE exhibits an accuracy of 85% for tumor classification and 88% for grading of gliomas [33]. Andrew et al. [34] implemented five machine algorithms such as Random Forest, Artificial Neural Network with a single hidden layer, Linear Support Machine, Logistic Regression, and Gradient Boosting Machine on EHR data based on Delirium Risk Prediction. A total of 18223 patients' samples were taken for the evaluation. From the results it is shown that Gradient Boosting Algorithm produced better result than others with AUC value of 85.5%. Maryam et al. [35] proposed a heart failure prediction model based on EHR data. A dataset of 5044 patients' data was taken, and features were extracted to calculate the survival score with the help of Cox Proportional Regression Model. Patients who survived for five years and died after five years are excluded from calculation of survival score and different machine learning models like Random Forest, Support Vector Machine, Logistic Regression, ADABoost, and Decision Tree were applied. From the experiment they came to conclusion that Logistic Regression performed well on extracted features with 11% improved AUC curve value.

2.2.1.2 Work on Sensor Data

Different research works have been carried out considering sensor data by the following authors. Luca et al. [36] proposed a ML model for detection of Parkinson's Disease from data streams collected from wearable sensors. Data was collected from 20 individual patients each wearing 6 wearable sensors. Total 13 tasks were performed by every individual every day for 2 weeks. A total of 41,802 data were collected. Convolutional Neural Network (CNN) and Random Forest classifiers were applied for detection of bradykinesia and tremor. Random forest classifier performed well with RO-AUC value of 73 % for bradykinesia detection and 79% tremor detection. David et al. [37] proposed a ML model for Developmental Risk Array (AD) and Typical Development (TD) detection in infants. Infants are divided into two sets such as 0-6 months and 6-12 months. Inertial movements of infants were detected using Opal Sensors fixed on infant's ankle. Support Vector Machine, Logistic Regression and AdaBoost are used for prediction. SVM performed well for0-6-month infants with an accuracy at 90% and ADABoost performed for 6–12-month-old infants with accuracy of 83%. Diana et al. [38] implemented four machine learning algorithms Decision Tree, Logistic Regression, DeepNets, and Ensemble for detection of fall in aged people. 3D-axis accelerometer is fitted on 6lowPAN wearable device which collected data and features are extracted using sliding window technique. Ensemble Algorithm performs better with accuracy of 94%. Shamsul et al. [39] proposed a model for recognizing human activity using data from body sensors. Sensor data was collected, and features are extracted using Kernel Component Analysis (KPCA) and Linear Discriminant Analysis (LDA). Extracted features were trained on Deep Belief Network with 40 hidden layers. Deep Belief Network performed well for activity detection with an accuracy of 97.5%.

2.2.1.3 Work on Genomic Data

Patrick et al. [40] used ML algorithms for hazard characterization in microbial risk assessment. DNA isolation and sequencing data were collected, and feature extraction was done. Random forest, Support Vector Machine, Logic Boost were applied on extracted features. Logic Boost performed well, and 75% accuracy is obtained. Marcus et al. [41] used XGBoost Machine Learning model for prediction of antimicrobial concentration level among patients. A non-typhoidal Salmonella genomes dataset with 5278 samples were collected. XGBoost with gradient ensemble boosting method is applied and an accuracy of 95% is obtained.

2.2.1.4 Work on Transcriptomic Data

Carly et al. [43] proposed a model which considered multiple gene expression dataset to identify gene signature for tuberculosis detection. Four datasets are integrated with 1164 patient samples. After feature extraction, machine learning algorithms like Random Forest, Support Vector Machine with polynomial kernel and Partial Least Square Discriminant analysis are applied. Random Forest outperformed other methods here with an accuracy of 95%. A hybrid approach for profiling of single-cell RNA sequencing data is proposed by Suhas et al. [44]. The model implements a deep unsupervised single cell clustering algorithm which uses features generated by deep learning model. The proposed model exhibits accuracy of 96%.

2.2.1.5 Work on Proteomic Data

Christine et al. [45] proposed a model to identify a set of critical proteins that have an association with FLT3-ITD mutation in newly diagnosed acute myeloid leukemia patients. A deep learning model with auto encoders in a form of hierarchical model is implemented here. Dimensionality reduction reduced the number of proteins from 291 to 20. The proposed model performed well with an accuracy of 97%. Bishnupriya et al. proposed a hybrid model for the prediction of protein structure class. A hybrid model including SkipGram based word2hovac and Atchleys space II, III, IV for electron ion interaction were implemented for amino acid sequence representation [46]. Stockwell transformation is used for feature extraction for time and frequency domain. Deep Recurrent Neural Network was implemented for classification. Six datasets including three small size datasets comprising 498, 277, 204 and large size samples comprising PDB25, 640, and FC699 are considered. Their proposed model performed well with accuracies of 95.9%, 94.9%, 85.36%, 84.2%, 94.3% and 93.1% for both sized datasets.

2.3 Review of Machine Learning based Disease Classification

Machine Learning Algorithms can be categorized into three parts: Supervised Learning, Unsupervised Learning, and Reinforcement Learning as shown in figure 2.1.

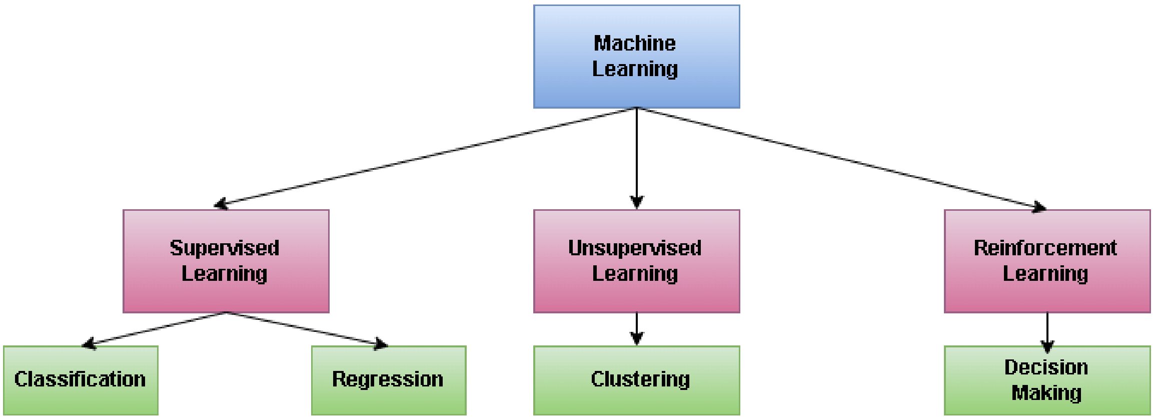


Figure 2.1: Different types of Machine Learning Algorithm

Supervised Learning algorithm allows collecting data and produces an output from previous experience. It works with known input data i.e., labeled data. Labeled data is data that are tagged with correct classification class or answer. Supervised Learning, as its name suggests, has a presence of a supervisor. Here ML model is trained with labelled data. After training, a new set of data is provided to be tested, and supervised ML algorithms produce result from labelled data. Whereas Unsupervised ML Algorithm allows the model to work with non-labeled data. It acts on the data without guidance. It produces results by finding similarities, patterns, and exiting differences among data without any prior knowledge. Unlike supervised learning algorithms, no supervisor is provided while training a model. Therefore, this learning process is supposed to find the hidden information from unlabeled data by itself. In Reinforcement learning, desired behavior is rewarded and undesired behavior is punished. In Supervised Learning technique, if an algorithm predicts the correct label of given input data it is called Classification model. Whereas in Regression model, the correlation between dependent and independent variables are established. It helps in predicting continuous variables. The aim of Regression algorithm is to find a mapping function that maps the input variable to the corresponding continuous output variable. Some widely used supervised machine learning algorithm application areas are Diabetic Prediction, Heart Disease Prediction, Speech Recognition [29], etc. Unsupervised Machine Learning methods are utilized to improve prevention strategies, identifying risk factors, and to deliver personalized therapies and targeted patient care.

Now a days, due to the growth of high throughput data, dimensionality and size of data are major concerns for both data mining and machine learning model implementation. Extraction of useful patterns and knowledge from this data has become very challenging. Features are considered as an independent variable that acts as an input to the ML algorithm. So, feature selection is an important step for any ML application. Different feature selection processes and their use in different research works are described below.

2.3.1 Feature Selection

Feature Selection is a process of obtaining a subset of relevant features from original dataset that is aimed to enhance an ML model's performance. The aim is to reduce the dimension of data while considering important features. Feature selection helps in model's interpretability. It also reduces the overfitting thus reducing overfitting problem. Methods of feature selection can be categorized into two types –

2.3.1.1 Filter Methods

Filter method categorizes attributes by evaluating them on some statistical metrics and prioritizing the most significant features based on metric results. Statistical attributes such as mutual information, correlation, or chi-square test are more significant. Nagarajan et al. used filter-based feature selection method for heart disease prediction [47].

2.3.1.2 Wrapper Methods

Wrapper method represents feature selection approach utilized by machine learning and data analysis. It evaluates the learning quality based on different evaluation metrics with different subsets of features from the original set of features. Based on the model performance, optimal set of features are selected which exhibits better performance than any other set of features or subset. Wrapper methods are computationally expensive as it takes time to evaluate possible subsets of features [48].

2.3.2 Feature Extraction

Machine Learning and Deep Learning model best works when relevant features are extracted from the dataset. Feature extraction is a crucial part of any machine learning algorithm. Feature extraction is a process that transforms raw data into features that provides strong potential for pattern recognition and classification. The primary objective of feature extraction is to point out vital attributes or traits or information from raw data that defines the data most. These attributes will serve as inputs to the machine learning model to execute the required task. Numerous feature extraction methods such as autoencoders, Principal Component Analysis (PDA), Linear Discriminant Analysis (LDA), t-distributed stochastic neighbor embedding (t-SNE) are available which can efficiently extract important datapoints or features.

2.3.2.1 Principle Component Analysis (PCA)

Principle Component Analysis is a dimensionality reduction technique which is used to reduce the dimensionality of large dataset to a smaller one retaining valuable information of the large dataset. It is defined as an orthogonal linear transformation that transforms the data to a new coordinate such that the greatest variance by some scalar projection of data lies to first coordinate which is also called as first principal component, the second greatest variant on second coordinate, the third greatest variant on third coordinate and so on [49,50].

2.3.2.2 Linear Discriminant Analysis (LDA)

Linear Discriminant Analysis (LDA) is a supervised dimensionality reduction technique used in pattern recognition, machine learning, and different statistical evaluation. It converts high dimensional data to a low dimensional data while optimizing the distinction level of different classes. It works well for classification, feature extraction and multi class data visualization [51,52].

2.3.2.3 Autoencoder

Autoencoder is an unsupervised artificial neural network used for feature extraction, dimensionality reduction. It consists of an encoder and decoder that together forms a neural network which first compress the image through encoder and reconstructs the image through decoder with minimal loss of information. It is widely used for anomaly detection, image denoising, pretraining of unsupervised neural network [53].

2.3.2.4 t-SNE

t-SNE is a nonlinear dimensionality reduction method used for visualization of high dimensional data. It conserves the basic structure of data which helps in preserving the distance between adjacent data points during dimensionality reduction. It is highly effective for clustering, finding patterns, and different structures in intricate datasets [54].

2.4 Supervised Machine Learning Algorithms for Disease Classification

2.4.1 Support Vector Machine (SVM)

Support Vector Machine is a widely used supervised machine learning algorithm for regression, outlier detection and classification tasks. It can be used for image classification, spam detection, text classification, face detection, gene expression analysis, and anomaly detection. SVM is very effective in managing high dimensional data and different nonlinear relationships. SVM exhibits high performance while dealing with large datasets. Its data training process is comparatively easy. The disadvantage of SVM is that it requires large training time which furthermore increases if the training samples number increases [55]. Many researchers have experimented with SVM in healthcare data. Most popular SVM application areas in healthcare are illness detection and prevention, facial recognition, speech recognition, image recognition.

Xingyu et al. in their study proposed an analytical technique for prediction of patients with severe acute myocardial infraction using SVM based on electronic medical records in the Medical Information Marketplace for Intensive Care (MIMIC)-III database [56]. An automated prognostic risk prediction model was proposed using SVM and the model was evaluated for five cross validations. Their proposed model exhibits 92.2% accuracy and 98% AUC. Its confidence interval was less with higher accuracy and better performance when compared with other algorithms such as Decision tree, KNN, Kernel Naïve Bayes, and RUSBoost. An improved stacked SVM model was proposed for early heart failure prediction by Liaqat et al. [57]. Their model has an accuracy range from 57.85% to 91.83%. A fuzzy support vector machine was used to detect coronary heart disease by Mehrbakhsh et al. [58]. Their experiment result reveals significant reduction in computational time of disease diagnosis in compared to non-incremental learning algorithm.

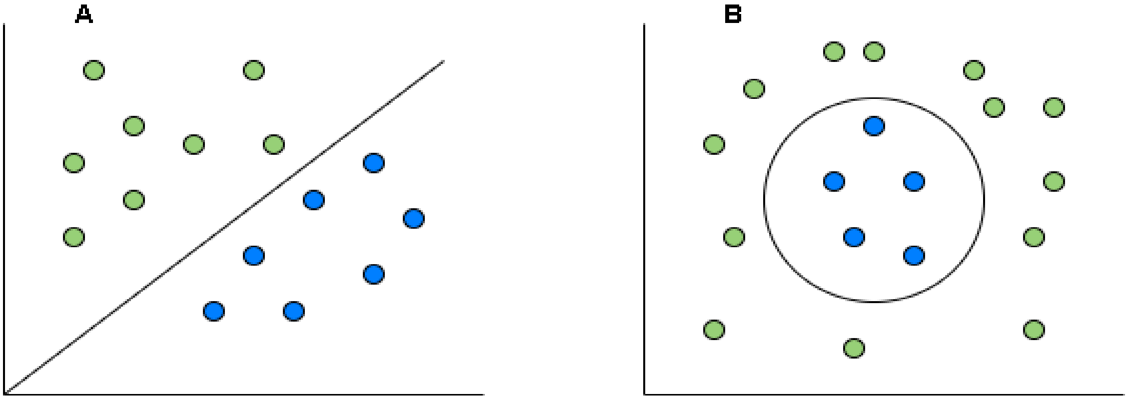


Figure 2.2: Support Vector Machine

2.4.2 Decision Tree

A decision tree classifier is a graphical tree representation of possible alternative outcomes and their possible end values. It consists of calculation of probabilities of few courses of action. It is a non-parametric algorithm utilized for classification and regression tasks. Training of decision tree starts with training data points with their related category labels. The training set is divided recursively into subsets based on feature labels. Data in each subset is much cleaner than its parent set. The internal nodes of the decision tree act like a test feature and each branch node represents test results and each leaf node represents class labels. To identify a test sample’s category label, transversal is done starting from root to leaf node and identifying the sample’s category label. Decision tree is easiest to implement as no prior domain knowledge and parameter settings is required. Hence its execution is fast and simple. It supports incremental learning that is immutable due to the alternative functions based on each internal node [59]. Decision Tree execution time is higher if the number of features are high. As it evaluates every potential branch at each level of the tree, it is computationally expensive [60]. Ching-Chin et al. proposed a systematic approach based on decision tree which considers the most important attributes in a dataset to identify the patients eligible for telehealth service. They considered NHIRD dataset which has huge outpatient and inpatient records. Decision tree-based model HDTTCA exhibits better performance in terms of accuracy, precision, and specificity than logistic regression algorithm [61]. Asim et al. [62] used a fuzzy rule-based approach with decision tree predictive model for diagnosing heart disease. Their model accuracy was calculated at 88% which is significantly high. Mohd Najib et al. [63] used decision tree to reduce data size by converting data to condensed form where sensitive information can be preserved, and accuracy will be higher for mobile health technology.

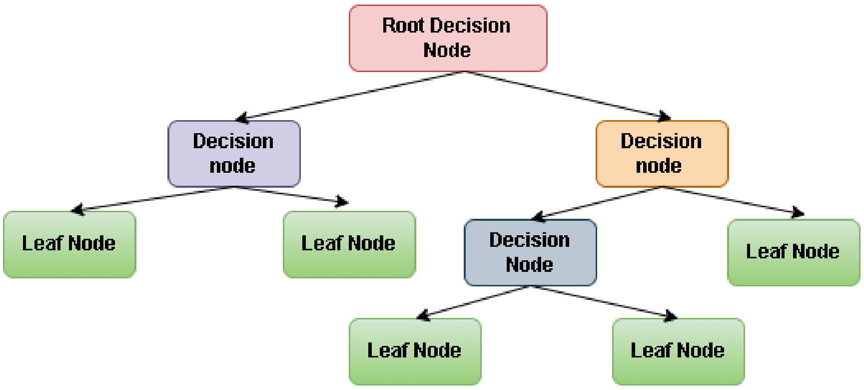


Figure 2.3: Decision Tree with Root, Decision and Leaf Nodes

2.4.3 Naïve Bayes

Naïve Bayes is a supervised Machine Learning algorithm used for classification task. It is a part of generative learning algorithms where the model is supposed to distribute the inputs to its corresponding class or label. It has one parent node and some independent child nodes which represent simple Bayesian Network [64]. Naïve Bayes works with probability classification method which multiplies the individual probability of each attribute value pair. It has shorter computational time and removes unnecessary attributes. Naïve Bayes works well on multi class classification problem and works better for incremental learning [65]. One demerit of Naïve Bayes is its lower rate of accuracy compared to other supervised machine learning algorithms. It requires many training records for better performance. Different studies had used Naïve Bayes as classifier for disease prediction. Hossam et al. [66] proposed a skin disease detection system based on Naïve Bayes Classifier which outperformed other Machine Learning methods with an accuracy ranging from 91.2 % to 94.3%. Gupta el al. [67] used heart disease dataset through some feature selection methods and obtained an accuracy of 88.16% on test datasets. Repaka et al. used a Naïve Bayes classifier in association with secure data transfer by employing Advanced Encryption Standard (AES) [68]. The authors used different techniques such as Sequential Minimal Optimization (SMO), Bayes Net (BN), Multi-Layer Perceptron along with Naïve Bayes to compare the performance accuracy. From the experiment it was shown that Naïve Bayesian Classifier performs better with an accuracy of 89.77% with fastest executional time.

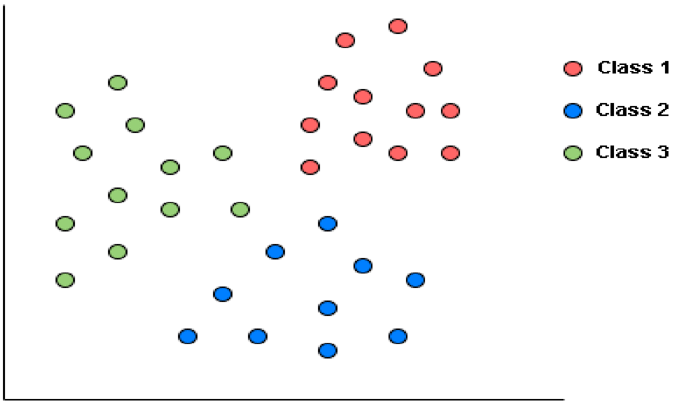


Figure 2.4: Naïve Bayes Classifier

2.4.4 K Nearest Neighbors (KNN)

K Nearest Neighbor is a classification as well as regression algorithm which works with an idea of finding similar datapoints with similar level or values. It identifies an unknown pattern by assigning a k value and finding k nearest neighbors of the input data point based on the distance of each neighbor from the input data point. KNN works well on multi class classification problem and data training phase is faster than any other algorithm [69]. One drawback of KNN is it takes more computational time for classification and its lack of ability to assign k value [70]. Further if too many unwanted features are there, it can lead to poor performance. Different researchers used KNN for disease prediction and diagnosis. Jabbar et al. [71] proposed an algorithm based on KNN with genetic algorithm for effective classification. Their algorithm had enhanced accuracy for heart disease prediction. Shouman et al. incorporated different algorithms into KNN for diagnosis of cardiac disease. According to their experiment, KNN produced higher accuracy than Neural Network for heath disease prediction [72].

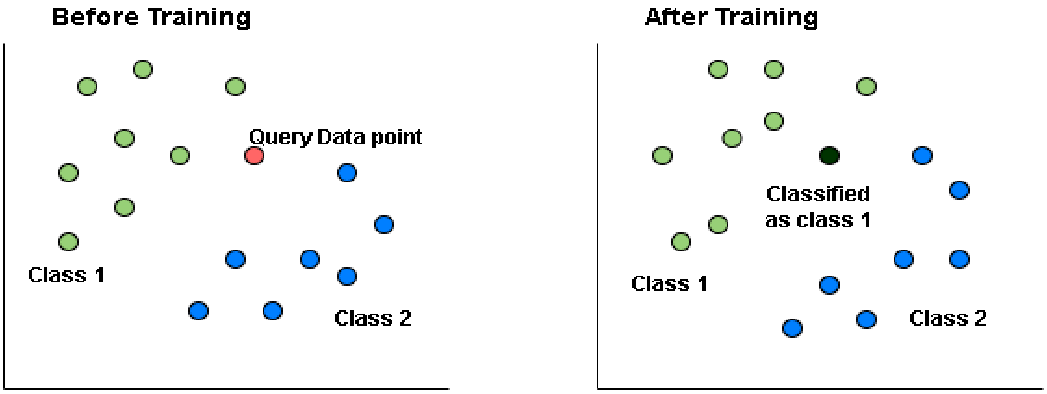


Figure 2.5: K-Nearest Neighbor (KNN)

2.5 Supervised Machine Learning Algorithms for Regression

Supervised regression technique is a process that can provide a continuous response as prediction. The aim of supervised regression algorithm is that it will predict a specific value of an outcome or data point rather than classifying the data point. Input dataset is divided into train and test set, and a continuous output or a targeted value is predicted by regression algorithms. Some popular regression machine learning algorithms are linear regression, logistic regression, support vector regression, ensemble method.

2.5.1 Linear Regression

Linear regression technique is a process to predict the value of a variable based on the values of other variables. The attribute whose value is required to be predicted is called dependent variable. The variables that are used to predict the value of dependent attribute are called independent variables. Linear regression works best with small dataset with high accuracy where it can be easily understandable and interpretable. The drawback of linear regression is that it has too many one-to-one relationships among attributes. So, it will not work with non-linear relational data. Linear regression will also have performance issue in highly non-linear problems where the relation between input and output cannot be approximated linearly. Though it has limitation, researchers try to implement Linear Regression for comparison purpose and for proper problem understanding. Petrou et al. used Linear regression for Healthcare Resource Utilization and achieved an accuracy of 95% [73].

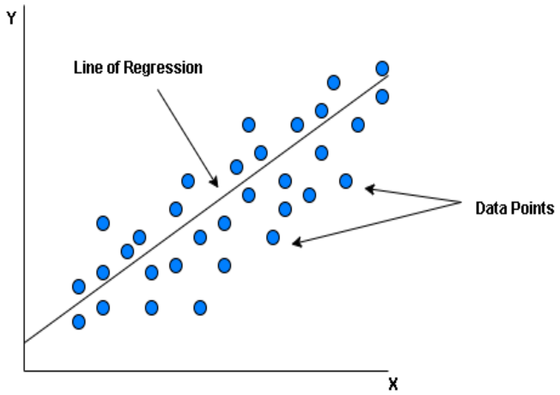


Figure 2.6: Simple Linear Regression

2.5.2 Logistic Regression

Logistic regression is used to predict discrete class labels. It evaluates the probability of an event occurring, based on a dataset’s independent variable. It uses a logistic function for classifying a label’s binary outcome either 0 or 1. So the output variable indicates the category of the input data point. Lemon et al. used logistic regression for predicting health related behavior [74].

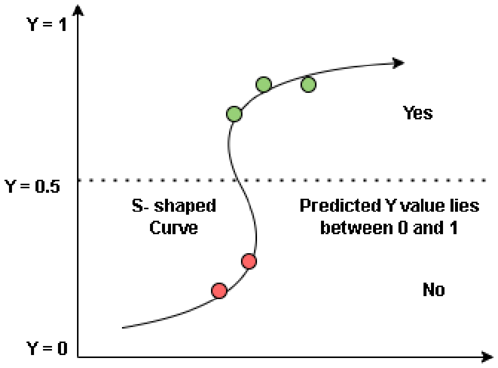


Figure 2.7: Logistic Regression

2.5.3 Support Vector Regression

Support Vector Regression is a machine learning algorithm for regression analysis [78]. The aim of SVR is to find a function that maps the input variable and the continuous target variable which will tend to minimize the prediction error. SVR tries to search a hyperplane where data points will be best suited in a continuous space. SVR can handle non-linear relationship between input data points and the target value by implementing a kernel function that maps the data points to a high dimensional space [79]. Advantage of using SVR is its computational complexity that does not get affected by the dimensionality of data. It has high prediction accuracy rate with good generalization ability. Sharin et al. [80] used SVR for predicting and analyzing COVID-19 data and obtained accuracy near about 94% from their experiment. Claris et al. proposed a single forecasting model using Gradient Boosting Methods, SVR and Generalized Addictive Model [81]. In figure 8, solid line indicates the hyperplane, and the distance between two dotted line denotes the boundary for two different classes.

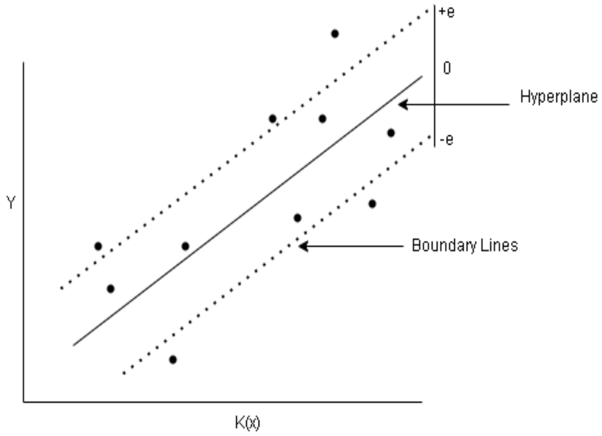


Figure 2.8: Support Vector Regression Graph.

2.6 Ensemble Methods

Ensemble means a collection of methods. In Ensemble machine learning algorithm, a collection of ML algorithms is combined to form a new model. Here each ML model is called base model which works as a feature extractor from the input data. Ensemble methods have high accuracy than any one of its base models as different models' strengths are used together for higher predictive power. Because of combining different models, the ensemble training process is much more complicated. Two popular ensemble methods are Bagging and Boosting. Example of Bagging Ensemble model includes Random Forest model and Boosting Ensemble model has GDBT, XGBOOST, ADABoost. Ensemble model can avoid overfitting problem. A single ML model considers different hypotheses that will forecast the training data less accurately for predicting unknown data for small dataset. Ensemble method eliminates the risk of considering unnecessary hypotheses which improve overall performances. Ensemble method will reduce the problem of reaching a local minimum by clubbing algorithms as one which can perform a local search that will be within the optimal space. It is also suitable for large dataset execution. Different ensemble methods are proposed by researchers for health data analysis. Ganie et al. used bagging, boosting, and voting ensemble methods for early-stage diabetes prediction [75]. Nguyen et al. proposed a deep stacked generalization ensemble algorithm on x-ray dataset [76]. An ensemble method was used for predicting patients' average weekly average expenditure on pain medication [77].

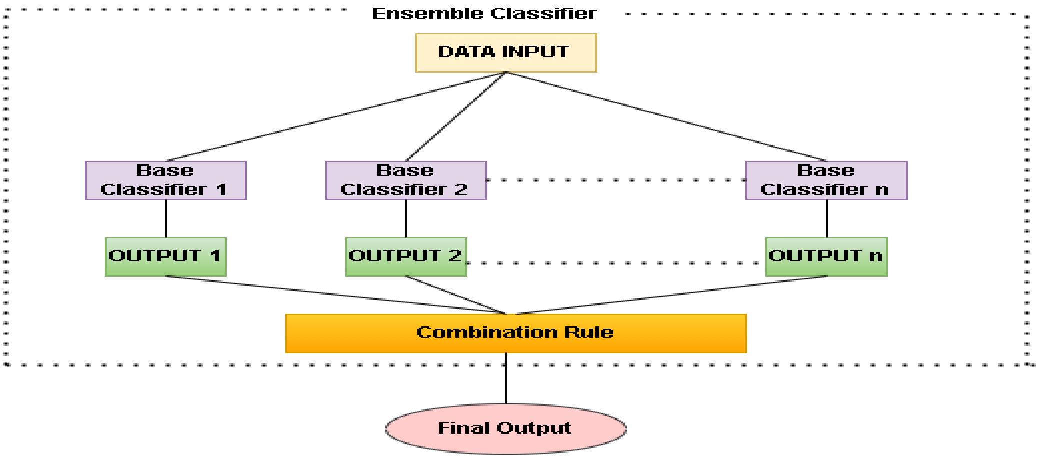


Figure 2.9: Architecture of Ensemble Method with Base Classifiers

2.7 Deep Learning Approaches in Healthcare Data

Since past decade, a vast range of Artificial Intelligence and Machine Learning approaches have been used for medical data analysis. Among different Machine Learning models, Deep Learning (DL) models or better known as Deep Neural Networks are more popular while dealing with huge datasets [63]. Deep Learning is an approach where multiple layers are used for feature learning and data gets passed through a cascaded series of layers for classification. Deep learning models are more accurate when large datasets are provided and can outperform many classical machine learning algorithms. Deep Learning models also have great performance while analyzing image, test, video dataset and natural language processing. As deep learning has a scope for continuous improvement in its methodology, it is more popular in analyzing healthcare data. For biomedical and healthcare data, different initiatives have been conducted such as Google's Deep Mind [83], IBM's Watson [84], etc. have been developed. These are computer-based support systems for analyzing health data.

Deep Learning applications can be categorized into two parts. Some applications are based solely on deep learning models, while some applications are in hybrid mode i.e., deep learning in association with machine learning models. Weakness and strength of Deep learning models alone and in association with classical machine learning models are studied extensively by researchers. Four main architectures of deep learning such as Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Networks (DBN) and Autoencoders are widely used for disease prediction and detection. In the following section, contribution of research works with deep learning models and their performance and drawbacks are discussed.

2.7.1 Deep Learning Algorithms and Architecture

A deep learning model is a simple Artificial Neural Network (ANN) architecture with input layer, one or more hidden layer, and output layer [85]. In a traditional Deep Neural Network (DNN), a weighted and bias-corrected input value is sent via a non-linear activation function. DNN tries to optimize the weight of each layer so that the overall weight of network is optimized so that the loss function also can be minimized. DL is vastly used for speech recognition, health monitoring, drug discovery, medical image analysis, to name a few.

2.7.1.1 Convolutional Neural Network (CNN)

Convolutional Neural Network is a supervised deep learning model. It is widely used for health image analysis. Three main layers of CNN are Convolutional layer, Pooling Layer, and fully connected layer. The input image is fed to convolutional layer filters which extract information and create a feature map for subsequent layers. Pooling layer tries to reduce the size of feature map while retaining important features and tries to minimize the weight. This process is called down sampling or sub sampling. Various types of pooling methods such as Max Pooling, Average Pooling, Global Pooling are there. Lastly fully connected layer transforms two-dimensional feature map into one dimensional vector for classification. However due to lack of computing power and data volume, CNN was not developed to a greater extent. But from 2012, the overwhelimng response of some objective factors and performance of CNN in various predictive applications, CNN gained popularity in computer vision field. Different transfer learning based CNN models like VGGNet[86], ResNet[87], GoogleNet[88], DenseNet[89], Mobile Net[90] gained popularity over past decade and used extensively for several healthcare based applications. CNN can be used for different types of data, such as processing of images with some spatial local correlation data, for example, MRI, chest X-ray, CT-scan, and other image data. Whether a single image has been used for prediction or a series of images used for auxiliary prediction [100], CNN can process images by extracting features and combining fusion technologies. Its performance is better than traditional machine learning models [101].

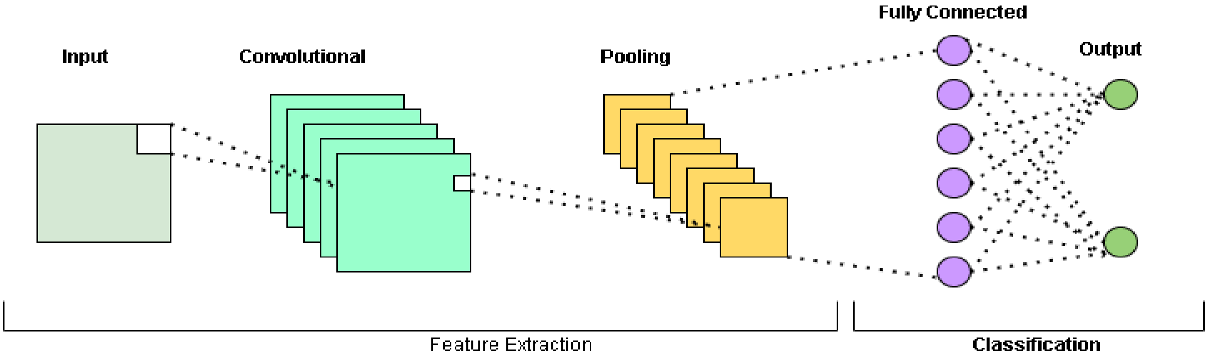


Figure 2.10: Convolution Neural Network (CNN) Architecture

2.7.1.2 Recurrent Neural Network (RNN)

Recurrent Neural Network is used for pattern recognition for sequential data like speech, handwriting, text data. RNN contains a cyclic structure where recurrent computations are done in hidden units with the help of their cyclic connection. Previous layer data is stored in a state vector and these state vectors are used to compute current layer output. Cyclic connection of each hidden layer uses recurrent computation for sequentially processing the input data. One drawback of RNN is the vanishing gradient problem while training data. Long short-term Memory (LSTM) network which is a type of RNN can store data sequence for long time using Gated Recurrent Unit (GRU). Efficiency of LSTM with GRU is dependent on the type of input data and how complex the problem is.

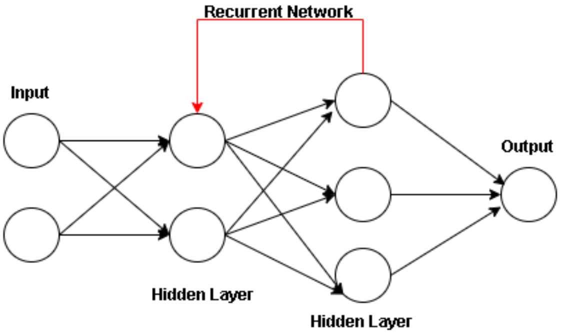


Figure 2.11: Recurrent Neural Network (RNN) architecture

2.7.1.3 Deep Belief Network (DBN)

Deep Belief Network is a form of generative graphical model, which is used to analysis high dimensional data. It is a hybrid multi-layer network that has both directed and undirected conditions. Top two layers have undirected connection and all other layers have directed connections. DBN is a stack of Restricted Boltzman Machines (RBMs) that are trained in greedy approach. RBM layers communicates with previous and subsequent layers. Yang et al. proposed a DBN model for medical big data analysis [126].

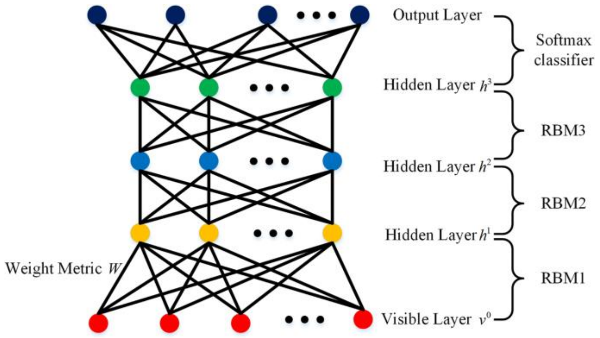


Figure 2.12: Deep Belief Network (DBN) architecture [125]

2.7.1.4 Generative Adversarial Networks (GAN)

GAN is a synthetic image generation tool based on game theory. The ideology of game theory lies in the way a network is trained. GAN can be implemented using two networks – Generative Network and Discriminative Network The idea is based on competition and back propagation algorithm is used to optimize the generation and discriminant network. The generating network is used for learning the actual distribution of samples, and the discriminating network is mainly used for distinguishing the real samples and samples generated by the generating network. The architecture of Generative Network based on deconvolutional neural network, and the discrimination network is based on Convolutional GAN was first proposed in 2014. Since then, the performance of GAN is constantly increasing regardless of image size or image fidelity. The application area of GAN includes image processing [102], improving image resolution, image-to-image translation [103], image restoration [104], medical image fusion. Application of GAN in healthcare data is gradually increasing and achieving remarkable results [105]. Though traditional GAN has some instability problems such as gradient disappearance [106,107], model collapse during training, modern improved methods that mitigates these problems have exhibited great performance [108]. Application of GAN for healthcare is still at its early stage. It has shown great performance in some medical applications [109]. With further solutions of its deficiencies, GAN will prove to a great help in healthcare applications.

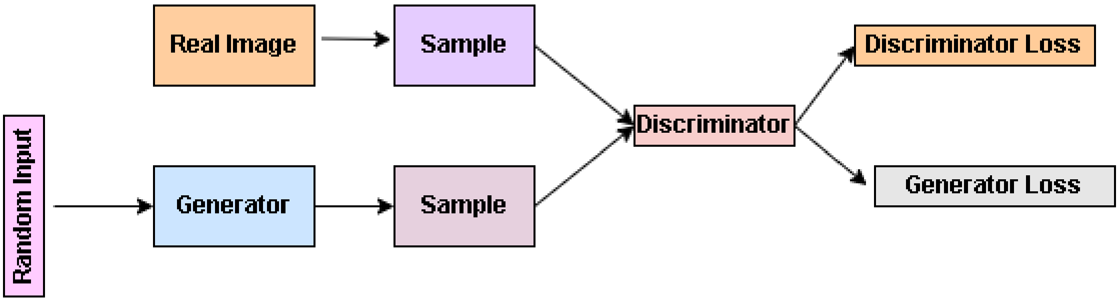


Figure 2.13: Generative Adversarial Network (GAN)

2.7.1.5 AutoEncoders (AE)

In real world, large amount data remain unlabeled due to lack of expertise. It promotes the need of Unsupervised Learning Algorithms. Sometimes it is difficult to label exactly the data collected, and some automated model is required to classify the data to discover potential value of the data [110]. Furthermore, medical data can be lost or distorted, or incomplete. Autoencoder is an unsupervised model that performs efficient feature extraction, dimensionality reduction, noise removal and helps in restoration of original data [111]. An autoencoder model can be combined with a variety of models due to its high degree of freedom [112]. In real world applications, many AE architectures like, Sparse Autoencoder [113], Variational Autoencoder [114] are constructed based on type of data. Those have proved efficient in processing low recognized medical data.

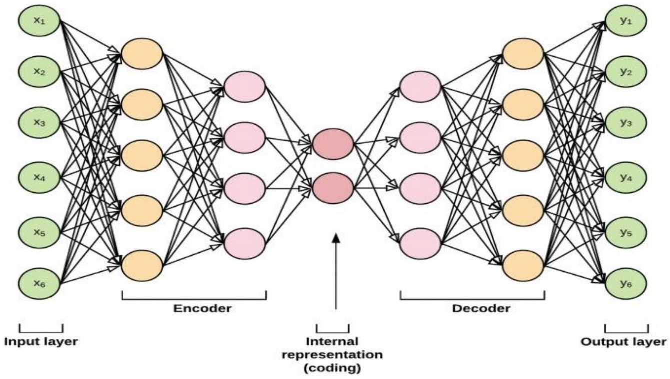


Figure 2.14: Autoencoder Architecture [124]

In real world applications, different research methods used combination of multiple models for better performance. Models are selected based on problem definitions. Ensemble methods of CNN and RNN are proposed in [115] where CNN is used to process horizontal data, RNN for vertical data, and further combined with Bagging classifier. The model performed significantly well. Transfer learning models have become more popular. The drawback of deep learning is that its data dependence is high, and it requires large amount of data. More transfer learning models can be combined with different DL models [116]. In [117], deep transfer learning model is proposed to analyze magnetic resonance images, which will effectively help healthcare professionals in diagnostic work. Transfer learning models applied on cancer image lesion recognition application showed great performance [118]. With the evolving nature of medical imaging technology, transfer learning will help the medical research field. Below table 2.1 describes some recent applications of Deep Learning models on various disease prediction.

Author Name	**Disease Name**	**Model Details**	**Year**
Arooj et al. [91]	Heart Disease	DCNN	2022
Mehmood et al. [92]	Heart Disease	CNN	2021
Sharma et al. [93]	Heart Disease	CNN	2022
Sudha et al. [94]	Heart Disease	Hybrid CNN + LSTM	2023
Shankar et al. [95]	Cardiovascular Disease	CNN	2020
Ordas et al. [96]	Cardiovascular Disease	CNN+SparseAutoEncoder	2023
Sharma et al. [97]	Breast Cancer	Ensemble CNN	2022
Gupta et al. [98]	Breast Cancer	CNN + LR + SVM	2020
Alissa et al. [80]	Parkinson Disease	CNN	2021
Choi et al. [99]	Parkinson Disease	CNN	2017
Ayon et al. [119]	Diabetes	Deep Neural Network	2019
Manokaran et al. [120]	COVID 19 detection	DenseNet201	2021
Navaneeth et al. [121]	Kidney Disease	Improved CNN	2020
Bhaskar et al. [122]	Kidney Disease	CNN-SVM	2019
Ma et al. [123]	Kidney Disease	Heterogeneous ANN model	2020

Table 2.1: Different Deep Learning model implementation for different disease detection and prediction

2.8 Conclusion

The development of healthcare industry has several modern aspects with the incorporation of big data and Internet-of-Things. In the era of huge multi-dimensional and multi form of unstable medical data, machine learning and deep learning have proved as a boon for future development. In the literature review, a series of popular machine learning and deep learning predictive models used for different disease in studies done by different researchers are discussed. Though some limitations and problems still exist while implementing ML and DL models. However, individual disease prediction on generic level in combination of precision medicine and combining Internet of Thing for developing intelligent medical prediction model will boost automated healthcare service. In near future, more diversified disease detection and prediction models will evolve and a transition from today’s auxiliary disease detection system to data driven decision making system will take over the world. As complexity of disease will increase, more complex and diverse ML and DL model will emerge. These models will be able to learn from complex data and will be able to produce significant result for complex medical system. Application of ML and DL will enhance the development of healthcare industry and clinical applications will be easy while promoting development of overall mankind healthcare.

Chapter 3

Machine learning Based Disease Classification

Medical Disease Detection and Prediction with machine learning is one of the biggest application areas of AI. AI's primary job is to develop algorithms and models to analyze if the systems behavior is appropriate in disease diagnosis. Medical Professionals identify disease or any critical condition in a patient based on a patient's health status, symptoms, etc. Detection of any health ailment from patient's medical history and symptoms require proper trained health professionals. Otherwise, it will be difficult to diagnose, as many symptoms may appear ambiguous. Therefore, regions where there is lack of health experts, proper health service and timely intervention are difficult to provide. As manual diagnosis of every ailment of patient is time consuming and prone to error, there is a chance to over diagnose or misdiagnosis a patient more often. If misdiagnosis is a problem, proper course of treatment will lack, and patient's condition may deteriorate. According to a study published in year 2022 [102], around 5.2 million injuries occur in India due to medical errors which result in preventable death around 3 million. Factors that lead to misdiagnosis are lack of evident symptoms which are often unnoticeable, disease may be omitted and mistaken and not considered for analysis and neglecting conditions of rare disease as well. So, the use of machine learning in healthcare starting from disease diagnosis to assuring safety of patient, is becoming both advantageous and popular. Many researchers have studied the benefits of machine learning based disease diagnosis and prediction, which prove to be inexpensive and time-efficient [103]. While traditional diagnosis of disease is time consuming, and costly, it also requires constant human intervention, ML based disease detection and prediction systems have no such limitations, and do not get exhausted after some time as human do. So, large number of patient's data can be analyzed for disease prediction. For efficient implementation of ML algorithms, healthcare data should be in tabular format where each patient's vital details are maintained or image data such as x-rays, MRI, CT-Scan could also be considered [164].

ML, being a subset of AI, the application of different mathematical functions produce result for classification and regression task, that are difficult for human to analyze in short span of time. The influence of machine learning algorithms in disease detection and prediction embellishes the use of technology in healthcare. The difficulties in applying ML in healthcare are addressing data imbalance, ML model's interpretation based on disease data, and utilization of proper ML ethics in healthcare domain. These challenges are required to be handled for proper integration of ML in healthcare [165].

Therefore, the purpose of analyzing the strength of ML in healthcare disease detection is to enable a machine learning based disease prediction system that will choose the most appropriate ML methods for a particular disease prediction and thereby increase the speed of disease classification and diagnosis process. In the present work, Parkinson disease is considered.

Parkinson's Disease (PD) is the second most common neurodegenerative disease which comes after Alzheimer's disease or dementia. Parkinson disease causes damage to dopamine-secreting neuron cells in substantia nigra. Population above 60 years of age are more prone to get affected by this disease [127]. Dopamine which acts on the striatum is responsible for the regulation of body motor functioning. Dopamine deficiency can lead to some major motor dysfunctions such as, muscle spasms, tremors, gait disturbance, slow gait, speech change, writing change, impaired posture, rigid muscle. Some non-motor actions such as depression, cognitive power degeneration, nervous system disorders are common side effect of dopamine deficiency [128]. Diagnosis of Parkinson disease is done by some cardinal motor symptoms such as rigidity, bradykinesia, and resting tremors [129]. If 50 % of dopaminergic cells in the substantia nigra are degenerated, such motor malfunctions occur. So early detection and prediction of PD is necessary. Some non-motor manifestations such as sleep disorder, cognitive decline, rise of depression, autonomic symptoms increase the onset of motor features [130]. Parkinson disease with prodromal or premotor lasts on an average of 20 years. So, there is an increasing demand to detect PD based on these premotor manifestations. Patients with early-stage PD can significantly improve by taking initial dose of drug as prescribed by doctors. However, if PD is detected in advanced stage, drug dose and frequency of drug dose administration increases and some new symptoms such as gait freezing, dyskinesia, and functional disability of patient also worsens. All these lead to degeneration of non-motor symptoms such as nervous system disorder, sensory changes [131]. Though PD is not life threating, complications such as sudden fall, swallowing disorder, gait movement abnormality significantly increase the functional mobility limitation [132].

Since last few decades, machine learning techniques are hugely applied for Parkinson disease detection which leads to significant increase in detection and prediction accuracy by means of different scenarios such as voice and speech signal analysis [133], different handwritten pattern analysis [134], and different neuroimaging technique [135,136].

In a recent review of work which focuses on applying machine learning algorithms for detection and prediction of Parkinson’s disease, it is shown that out of 209 studies, 168 focuses on detecting PD patients, and 2 research studies is focused on detection of PD based on non-motor symptoms[137]. Mabrouk et al, implemented five machine learning models on non-motor features and combined the results using DatSCAN with motor features. Prashant et al, gained high classification accuracy using Support Vector Machine while considering non-motor symptoms, cerebrospinal fluid markers, and dopamine transporter imaging using Single Photo Emission Computed tomography (SPECT) [139]. Zhang et al. [140] proposed a combination of non-motor symptoms in association with other clinical image and bio fluid bio makers for PD sub typing. ChunGu et al. [141] proposed a prediction study based on depression arising on newly diagnosed PD patients. Armañanzas et al. used classification algorithms for non-motor symptoms [142].

As evident from the above not enough studies have been carried out for detection of PD based on non-motor symptoms. If PD is detected in preclinical stage by recognition of non-motor actions, it may help in slowing down the disease progression. So, higher detection accuracy using non-motor symptoms with the help of machine learning algorithms will quicker the decision-making process of general practitioners. Additionally, building impactful healthcare system with the help of interpretable models will boost overall healthy mankind.

The current work focuses on two proposed generic frameworks for Parkinson Disease detection, prediction, and optimization. Two case studies are presented for Parkinson disease classification problem. First case study proposes an approach based on a Correlation-Sequential Forward selection-based feature selection method for Parkinson Disease detection and data analysis. The second case study proposes an approach based on a Filter – WOA wrapper-based feature selection approach for Parkinson disease detection and classification. For implementation purpose, data sources and data processing processes are elaborated first, and the correlation matrix is defined for filter-based feature selection methods, and finally two case studies are implemented and analyzed.

3.1 Data Source Description

The current implementation is based on three public datasets. First dataset is UCI’s Oxford Parkinson Disease Detection Dataset by Little et al. [143] and New HandPD dataset by Pereira et al. [144]with two variants – HandPD Spiral dataset and HandPD Meander Dataset.

3.1.1 UCI Parkinson Disease Dataset

This dataset is comprised of a range of bio medical voice measurements collected from 31 people, out of which 23 are diagnosed with Parkinson disease [143]. Each column in the table represents a particular voice measure while each row indicates one of 195 voices recording from those 31 individuals. In the dataset, healthy people are set to 0 as status while status 1 is reserved for PD. The dataset is in ASCII CSV format. Below the features of the dataset are discussed:

1. **Name -** ASCII subject name and recording number
2. **MDVP:Fo(Hz) -** Average vocal fundamental frequency
3. **MDVP:Fhi(Hz) -** Maximum vocal fundamental frequency
4. **MDVP:Flo(Hz) -** Minimum vocal fundamental frequency

Below five features describes the variation in Frequency.

5. **MDVP:Jitter(%) -** Percentage of cycle-to-cycle variability of the period duration
6. **MDVP:Jitter(Abs) -** Absolute value of cycle-to-cycle variability of the period duration
7. **MDVP: RAP -** Relative measure of the pitch disturbance
8. **MDVP: PPQ -** Pitch perturbation quotient
9. Jitter:DDP - Average absolute difference of differences between jitter cycles

Below five features indicates the variation in amplitude.

10. **MDVP:Shimmer -** Variations in the voice amplitude
11. **MDVP:Shimmer(dB) -** Variations in the voice amplitude in dB
12. **Shimmer: APQ3 -** Three-point amplitude perturbation quotient measured against the average of the three amplitudes
13. **Shimmer: APQ5 -** Five-point amplitude perturbation quotient measured against the average of the three amplitudes
14. **MDVP: APQ** - Amplitude perturbation quotient from MDVP
15. **Shimmer:DDA -** Average absolute difference between the amplitudes of consecutive periods

The dataset contains two features that measures the ratio of noise to tonal components of voice.

16. **NHR** - Noise-to-harmonics Ratio
17. **HNR** - Harmonics-to-noise Ratio

Features for three nonlinear dynamical complexity measures are:

18. **RPDE** - Recurrence period density entropy
19. **D2** - correlation dimension
20. **DFA** - Signal fractal scaling exponent

Features for three nonlinear measures of fundamental frequency variation:

21. **spread1** - discrete probability distribution of occurrence of relative semitone variations
22. **spread2** - Three nonlinear measures of fundamental frequency variation
23. **PPE** - Entropy of the discrete probability distribution of occurrence of relative semitone variations

Feature that indicates the condition of a patient whether it is healthy or diagnosed with PD:

24. **Status -** Health status of the subject (one) - Parkinson's, (zero) – healthy

3.1.2 New Hand-PD Dataset

This dataset is comprised of hand-written exam details from 66 individuals sub-grouped into two parts: healthy and PD patient. It consists of 594 images divided into 6 classes such as Healthy Meander, Healthy Spiral, Healthy Circle, Patient Circle, Patient Meander, and Patient Spiral [166]. From this dataset, two variety of data is considered for binary classification - Hand PD Meander (with two class- Healthy and PD patient), and Hand PD Spiral (with healthy and PD patient class).Tabular form of data in csv format is considered. Features of the corresponding CSV file are discussed below:

1. **IMAGE_NAME**: name of the image
2. **ID_EXAM:** identifier of the exam (handwritten form.
3. **ID_PATIENT**: identifier of the patient
4. **GENDER:** M = male and F = female
5. **RIGH/LEFT-HANDED**: R = right-handed and L = left-handed
6. **AGE:** age (years)
7. **RMS:** root mean square (Equation 3 of the paper)
8. **MIN_BETWEEN_ET_HT:** the maximum difference between ET and HT radius
9. **MAX_BETWEEN_ET_HT:** the maximum difference between ET (exam template) and HT (handwritten trace) radius
10. **STD_DEVIATION_ET_HT:** standard deviation of the difference between ET and HT radius
11. **MRT:** mean relative tremor.
12. **MAX_HT:** maximum HT radius
13. **MIN_HT:** minimum HT radius
14. **STD_HT:** standard deviation of HT radius
15. **CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT:** the number of times the difference between ET and HT radius changes from negative to positive, or vice-versa.
16. **GENDER:** M = male and F = female
17. **CLASS_TYPE:** 1 = Healthy and 2 = Parkinson affected patient

3.1 Data Preprocessing Stage

It is observed that data may not always be in the appropriate form or format or quantity that is required for analysis and subsequent prediction. So, there may be data imbalance. Data imbalance appears to be a predominant and inherent problem. Imbalance data refers to the number of instances for each class being different in a classification problem. Balanced dataset acts like fuel for machine learning models and exhibit better performance. In imbalance dataset, it is considered that data is generally biased towards one class [145]. So, if an algorithm is trained on that data, classification result will also be biased towards that class. The model will learn more from biased instances as opposed to the instances of the minority class. The distribution of an imbalanced dataset is categorized by the difference of instances or data points of each class. In machine learning application, imbalance data poses challenges for binary classification in terms of class separation, feature correlation, and evaluation of the model. Whereas, a balanced dataset refers to a dataset where distribution of instances in every class is equal. Balanced dataset makes training of a model much easier, and it helps a model to become non-biased towards one class. ML is a data centric approach of AI, where data acts like a powering tool of the model. Better performance from the model is achieved when more data is used for training. However, obtaining right data and ensuring data quality is essential for improving performance of the model. Balancing data can have two approaches - either collect more real-world data for minority class or generate synthetic data. There may be some constraints while collecting more real-world data such as, lack of resources, demographic constraints, etc. So, synthetic data generation is a reliable way to balance a dataset.

Resampling is a technique to down sample or up sample the majority or minority class. If instances of minority class are increased, it is called Oversampling. If numbers of instances are deleted randomly from majority class, it is called under sampling.

3.2.1 SMOTE: Synthetic Minority Oversampling Technique

One popular technique of oversampling is Synthetic Minority Oversampling Technique (SMOTE) [146,147]. SMOTE helps in mitigating bias and captures important features from minority class for generating samples of that class. By this way SMOTE helps in better model performance leading to better accuracy in prediction. SMOTE generates new virtual training instances by linear interpolation for minority class. These synthetic instances are generated by randomly selecting one or more of the k-nearest neighbors for every instance of the minority class [148]. After this oversampling process, data is reconstructed, and ready for model implementation. Below the steps involved for data generation are discussed:

Step1: Denote the minority class as A, where each instance x, $\boldsymbol{x \epsilon}$ **A.**

Step 2: Calculate Euclidean Distance between x and every other sample in A for obtaining K-nearest Neighbors of x.

Step 3: Set sample rate N according to the imbalance proportion.

Step 4: For each $x \epsilon$ A, N samples (x_1, x_2,…….,x_N) are selected randomly based on its K-neighbor and construct subset $A_{1.}$

Step 5: For each $x_k \epsilon A_{1,}$ where k = 1 2,….. N, new samples are generated using the formula:

$$x' = x + \text{rand}(0,1) * | x - x_k |, \qquad (3.1)$$

Where rand (0, 1) represents random number 0,

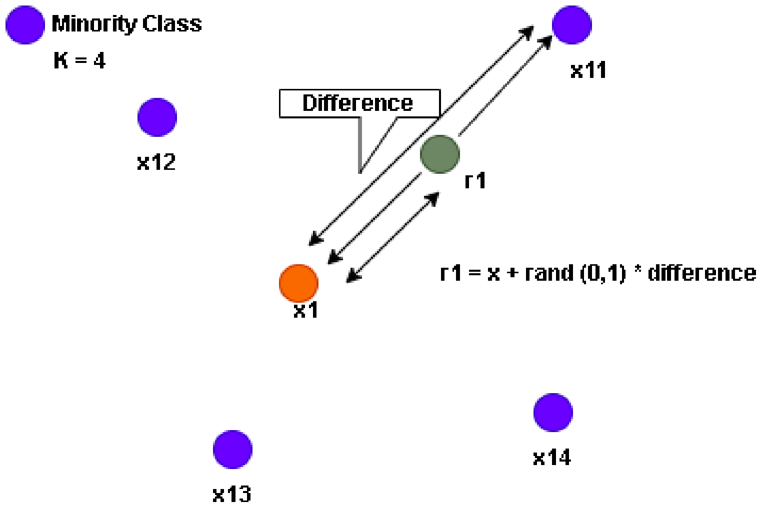


Figure3.1: SMOTE data balancing for Synthetic data generation

3.2.2 ADASYN Data balancing Algorithm

Though SMOTE is good for data balancing, it suffers from sample overlapping, noise interference, blindness while selecting neighbor. To overcome this problem, Adaptive Synthetic Sampling Approach or ADASYN is used [149]. It takes density distribution into consideration which decides the number of samples to be generated for each sample which is difficult to learn. It helps in changing decision boundary adaptively based on the samples which are difficult to learn. Below Steps of ADASYN are discussed:

Step 1: For any dataset, total number of majority N^- and minority N^+ are recorded. Calculate the ratio of minority to majority as *d*,

$$d = \frac{N+}{N-} \qquad (3.2)$$

Step 2: If d is lower than the threshold value, initialize the oversampling algorithm.

Step 3: Total number of synthetic data is generated as:

$S = (N^- - N^+) * d$, where S is the total number of minority data generated, d is the ratio of minority to majority. If $d = 1$ dataset is perfectly balanced.

Step 4: For every minority sample x_i, Euclidean distance is calculated for K nearest Neighbor and ratio r_i is calculated as

$$r_i = \frac{\#majority}{k}, \qquad (3.3)$$

where each minority instance should be associated with different neighborhood.

r_i value here denotes the level of dominance of the majority class in each neighborhood. Higher r_i indicates more majority class samples and more difficult to learn.

Step 5: Normalize r_i values such that summation all r_i values is equal to 1.

$$\hat{r}_i = \frac{r_i}{\sum r_i} \text{ and } \sum \hat{r}_i = 1 \tag{3.4}$$

Step 6: Calculate the number of synthetic instances to generate per neighborhood.

$$S_i = S\hat{r}_i \tag{3.5}$$

If r_i value is higher for neighborhoods dominated by majority class instances, more synthetic minority class instances will be synthesized for those neighborhoods.

Step 5: Generate S_i data for each neighborhood. First, consider the minority instance for the neighborhood, x_i. Select randomly another minority instance within the neighborhood, xz_i. The new synthetic example can be calculated using:

$$s_i = x_i + (x_{zi} - x_i) * \text{rand}\ (0,1), \tag{3.6}$$

where rand produces a random number between 0-1, s_i represents new synthetic data instance, x_i and x_{zi} represents two minority data instances within same neighborhood. So, it can be interpreted that new synthetic data instances can be created based on linear combination between x_i and x_{zi}.

.

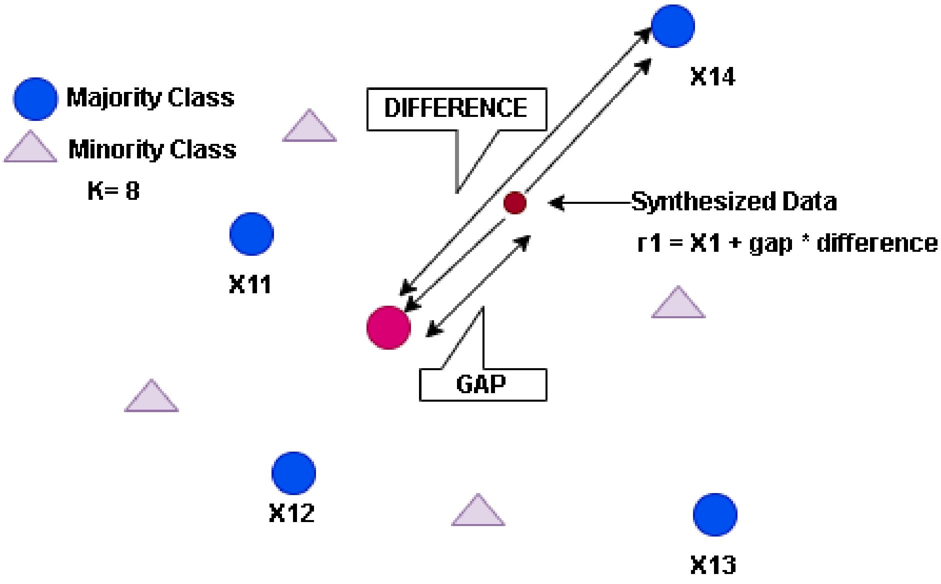


Figure 3.2: ADASYN Data Balancing technique for Synthetic data generation

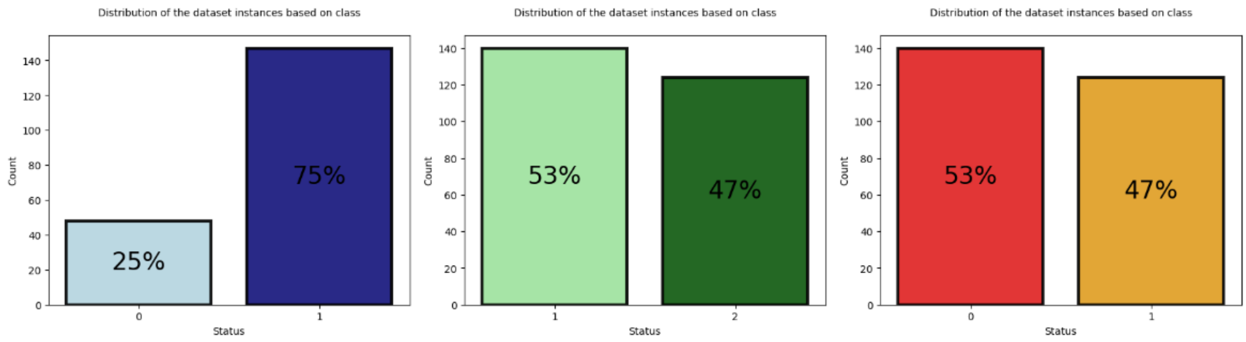


Figure 3.3(a) **Figure 3.3(b)** **Figure 3.3(c)**

Figure 3.3: Bar plot of 3.3(a) Data Imbalancing for Parkinson UCI dataset, 3.3(b) HandPD- Spiral Dataset, and 3.3(c) HandPD Meander Dataset

From the above figure 3.3, Bar plots of each class is shown for each dataset. Here class 0 indicates healthy people, and class 1 indicates Parkinson Patient. In figure 3.3(a), it is evident that data imbalance exists as instances for class 0 is only 25% and instances for class 1 is 75% for UCI dataset. Whereas for HandPD spiral dataset, instances for class 1 is 53% and for class 2 is 47% and for HandPD Meander dataset, instances for class 0 is 53% and instances for class 1 is 47%. So, for Parkinson UCI dataset, percentage of data imbalance is higher than other two datasets – HandPD Spiral and HandPD Meander. ADASYN SMOTE Data balancing algorithm is applied only for UCI Parkinson dataset because for other two datasets, less data imbalance is observed. After applying Smote algorithm class balancing obtained for UCI Parkinson dataset is shown in figure 3.4.

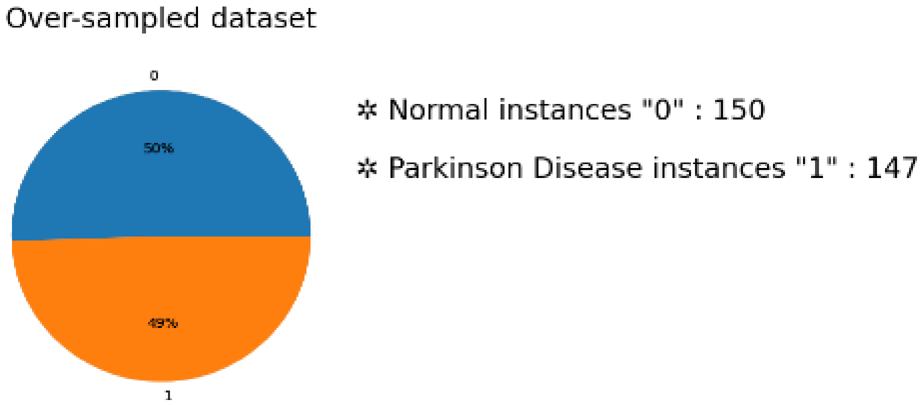


Figure 3.4: Oversampled Dataset after ADASYN Balancing Application on UCI Parkinson Dataset

3.3 Feature Selection

Feature selection is a process to find the best subset of features that helps to build an optimized model that perform better with feature subset than model trained with whole dataset. Feature is considered as an attribute which can have high impact on the problem. Each machine learning model depends on feature engineering which can be classified as feature selection and feature extraction. In feature selection a subset of features from the existing features are considered for model evaluation while in feature extraction new features are incurred from the exiting features. Feature selection is a way to reduce the input dataset for the model by selecting important features and thus reducing overfitting of the model.

As data is collected from different sources, it may be noisy, sometimes irrelevant data and some unimportant data too. As huge number of data will slow down the learning process of the model, it may not perform well if noisy and irrelevant data is trained. So, selecting the best features from the data for the model is an important step.

Feature selection helps in eliminating the curse of dimensionality, helps in simpler model interpretation, reduces overfitting while enhancing generalization process, and most importantly reduces training time.

There are two types of feature selection techniques:

- Supervised Feature Selection

It considers the target variable and uses for labelled data.

- Unsupervised feature Selection

It ignores the target variable and can be considered for unlabeled data.

Among the supervised feature selection methods, filter method is popular. Filter method is based on statistical measures. It does not rely on learning algorithm and acts like a preprocessing step for features. The advantage of filter method is it requires low computational time and mitigates overfitting of data. However, filter method tends to select high number of features or even it can select whole feature set as subset. So, a threshold value is needed to be to select a subset.

3.3.1 Filter based Feature Selection Method

Filter Feature selection is a process of applying statistical tests to inputs with a specified output. The goal of feature selection process is to determine which features are more predictive in nature and has direct influence on decision making. Therefore, feature selection is a process, that eliminates irrelevant and redundant features from the dataset thus reducing the complexity and dimensionality of data. In a nutshell, filter method for feature selection is a process to select a subset of features based on features relationship with target variable. On other hand, filter-based feature selection is not dependent on any machine learning algorithms.

Features in a Parkinson disease dataset are not always correlated to each other. Using temporal and spatial features will always lead to overfitting [151]. Overfitting means the model fits well for training data rather than testing data. For Parkinson disease prediction some features are learnt well and can fit training data. However, these relevant features are too tough to be recognized in test data well. To avoid the problem of overfitting, Pearson Correlation Coefficient (PCC) [151] is used to reduce the number of irrelevant features. Pearson Correlation is considered as best method for measuring the association among features in a dataset. It relies on the method of covariance. It provides an insight about the correlation or magnitude of the feature's association and direction of feature relationship. As Pearson correlation measures linear relationship, it is relevant for features measured from an interval scale. So, for Parkinson disease classification problem, Pearson Correlation is considered for filter method.

Pearson Correlation is used to evaluate the linear correlation between two variables say X, Y. Pearson Correlation is computed as:

$$\rho_{xy} = \frac{COV(X,Y)}{\sigma X \sigma Y} = \frac{E\left[(X-\mu_x)(Y-\mu_y)\right]}{\sigma X \sigma Y} \quad (3.7)$$

where COV (X, Y) is defined as covariance of X, Y. σ_X and σ_Y are defined as standard deviations of X and Y, respectively. μ_X and μ_Y are the means for X and Y, E represents expectations. The value of Y varies from -1 to +1. If Y value is 1, it indicates X is positively linear correlated to Y. If Y value is 0, it indicates X is not at all correlated to Y. If Y is -1, it implied X is negatively linear correlated to Y. If value of ρ_{xy} is greater than equal to 0.8, it indicates X and Y have extremely strong correlation to each other.

Therefore, value of ρ_{xy} is set to .75 for selecting highly correlated features from the datasets.

Here in this case the correlation between each feature and class variable is considered for subset generation. If the correlation value between a feature and the class variable is greater than 75%, the feature is added to the feature subset.

Thus, after application of ADASYN data resampling technique on UCI Parkinson Disease dataset, next step in the implementation is to generate Correlation Matrix using Pearson Correlation Coefficient. After applying Pearson Correlation in Parkinson UCI Dataset, HandPD dataset for Spiral and HandPD dataset for Meander, respective Correlation Matrices provided in figure 3.5,3.6,and 3.7 are obtained.

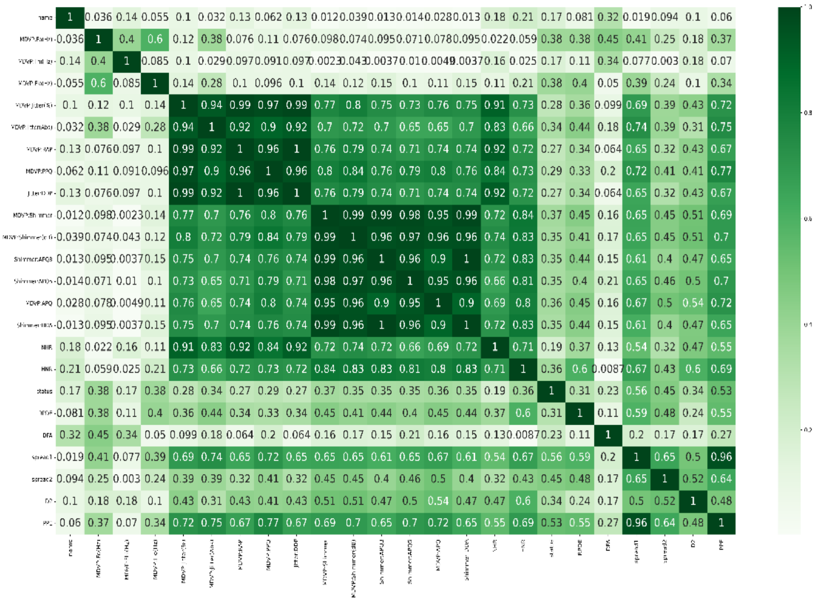


Figure 3.5: Pearson Correlation Coefficient Matrix for UCI dataset

Figure 3.6: Pearson Correlation Coefficient Matrix for HandPD – Spiral dataset

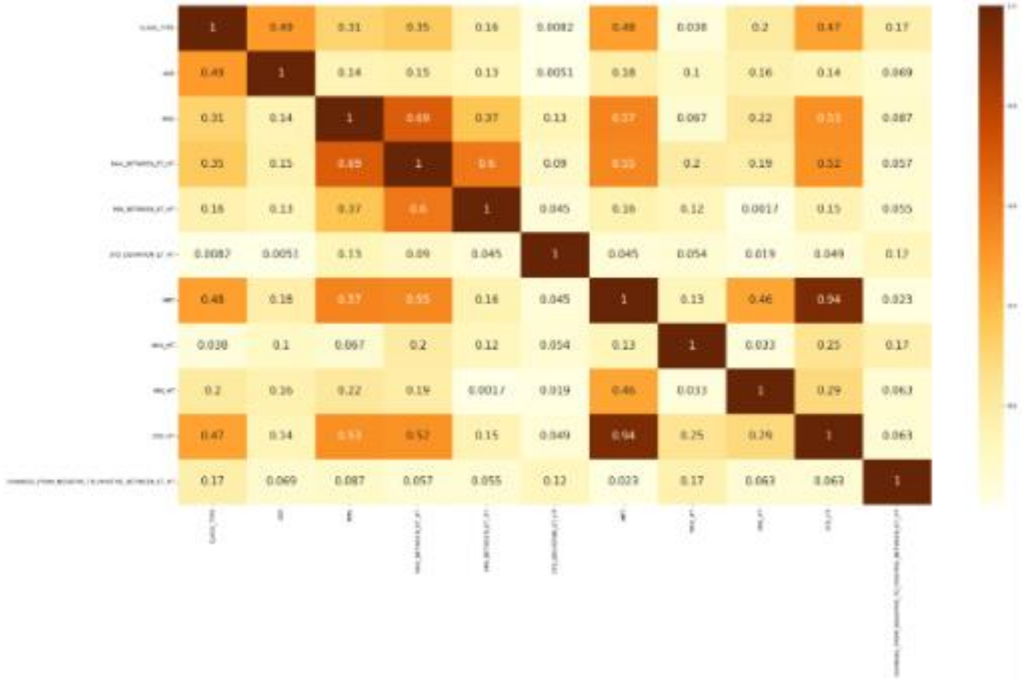


Figure 3.7: Pearson Correlation Confusion Matrix for HandPD - Meander Dataset

3.4 Case Study 1 – A Correlation - Sequential Forward Selection Based Feature Selection Approach

As informative feature has great impact on accuracy of a classification model, quite a few research works related to efficient feature selection have been carried out. Below are discussed some feature selection methods which are proposed in the literature. The authors in [4] described an iterative hybrid filter – wrapper feature selection approach for document clustering. They implemented a greedy algorithm for feature selection in each step and thus selected a competent set of features with expectation maximization-based algorithm. In the work in [5], the authors have implemented a hybrid feature selection method for supervised classification based on Laplacian Score ranking combined with a wrapper strategy. In their proposed method KNN and C4.5 performs better reaching on average precision around 87% and required less running time than any other feature selection algorithm. The proposed model in [6] for filter-wrapper hybrid method is divided into two phases. In first phase features are ordered according to a reasonable criterion and then best features are selected based on final criterion. The method tries to eliminate errors in filter method and increase the efficiency of wrapper method. As it integrates ranking results of multi filter methods, it avoids the uncertainty of using single criterion method. In the filter wrapper-based hybrid feature selection model described in [7], mutual information is used to rank each feature and each class and then k highest relevant features to the classes are chosen. In second phase while applying wrapper method, Shapley value is chosen to evaluate the contribution of features to the classification task in a feature subset. In [8], the aim of the authors was to achieve a fast, affordable, and objective diagnosis of the genetic variant of oligodendroglioma with a novel data mining approach combining a feature selection and ensemble-based classification. So, the authors applied a global optimization-based hybrid filter wrapper feature selection method which increases the chance to effectively diagnose brain tumor.

3.4.1 Filter Wrapper based Method

Filter and wrapper are two independent methods to choose a feature subset. They can work well on dataset for feature selection. Here, a Filter Wrapper based feature selection approach is proposed It is based on correlation-based feature selection and sequential forward feature selection concepts. These two concepts are clubbed to enhance the accuracy and link them closely to provide better feature subset for machine learning model application.

3.4.1.1 Correlation Based Feature Selection

Correlation based feature selection selects a subset of features that are highly correlated with classification, and uncorrelated to each other [150]. As filter method selects feature independent of any machine learning algorithm, they are selected on basis of score in various statistical methods for their correlation with outcome variable. According to variety of data i.e., continuous, or categorical data, different correlation coefficients can be applied to them, the table 3.1 provides all correlation coefficients that are appropriate for continuous and categorical data.

Features	**Continuous**	**Categorical**
Continuous	Pearson’s Correlation	LDA (Linear Discriminant Analysis)
Categorical	Anova	Chi-Squared

Table 3.1: Filter Methods

As discussed in section 3.3, Pearson’s Correlation is implemented to obtain correlation matrixes for three datasets. It helps to obtain linear dependency between two continuous variables and its value ranges from +1 to -1[138]. Value tending to +1 means variables are highly correlated i.e., if value of variable X increases, value of variable Y also increases. Value tending to -1 indicates variables are negatively correlated and value equals to 0 means no linear dependency. As, the cutoff for Pearson correlation is set to .75, features that are highly correlated to the class variable are selected for feature subset. This means features that have correlation with class variable greater than 75% is selected and features that have lower correlation value is discarded. This process eventually decreases the number of features and makes the learning process much faster. Since healthcare data requires timely prediction of occurrence of any unusual event, so a proper feature subset will effectively speed up the learning process and provide answers in timely manner.

3.4.1.2 Sequential Forward Feature Selection

The basic difference between a filter method and wrapper method is that filter method does not need any machine learning model to select a subset of features; whereas in wrapper method, a model is first evaluated based on a subset of features and according to the evaluation more features are added or removed from the subset [139]. Wrapper method is computationally expensive and requires more time to ultimately decide the final feature subset. Sequential forward selection starts with a model with some no. of features. It is an iterative process, which keeps adding feature in every iteration until the accuracy of the model does not improve with further addition of a feature [140].

3.4.2 Proposed a Correlation – Forward Sequential based Feature Selection Approach

In the proposed correlation – forward sequential based feature selection model, Pearson's Correlation is first implemented to evaluate correlation between each feature. As discussed in previous section 3.3, after application of Pearson's correlation, a subset of features that is further considered for sequential forward feature selection is obtained. Here the final feature subset value is set to 5 i.e., after sequential forward feature selection the feature subset size is 5. Initially it was the workflow of the algorithm is illustrated in figure 3.8:

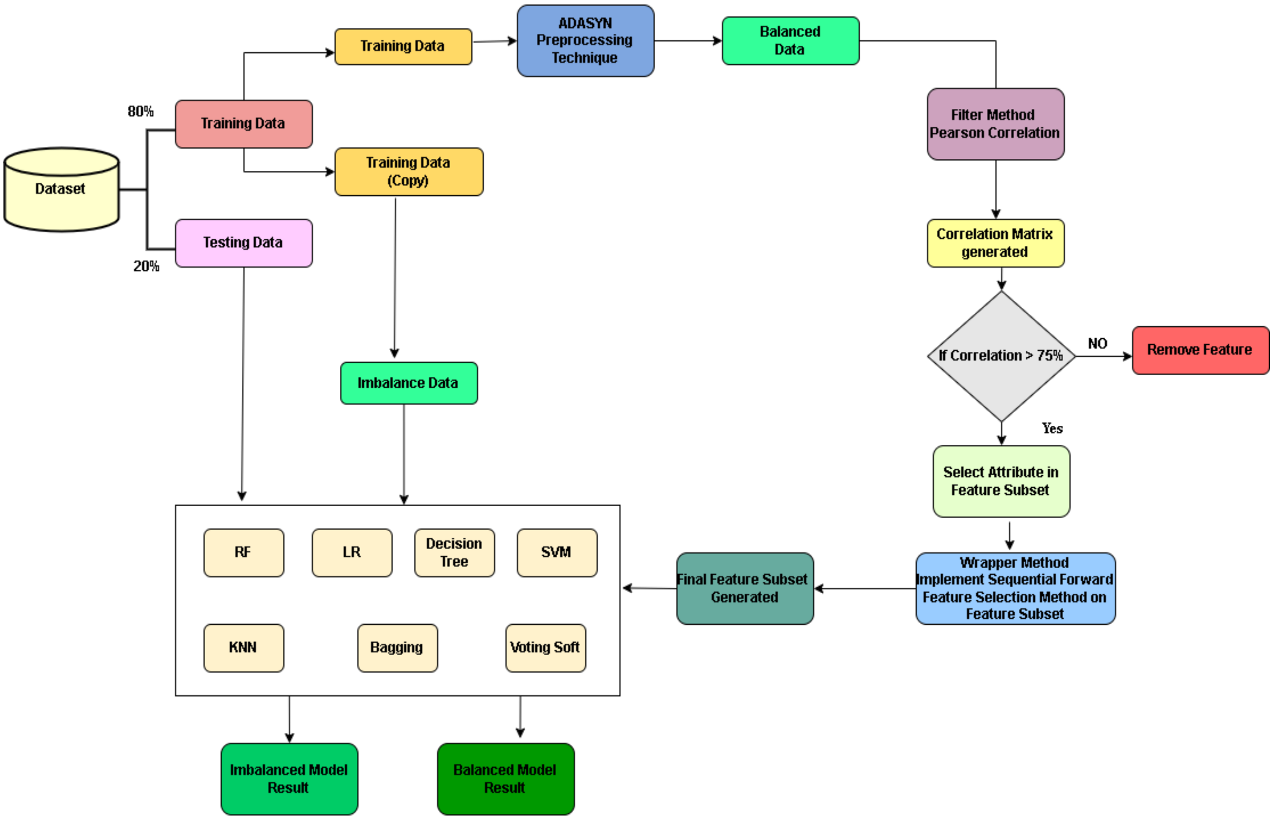


Figure 3.8: Proposed Correlation – Sequential Forward Selection based Feature Selection Approach

3.4.3 Implementation Results on Parkinson Disease Dataset and Analysis

The proposed approach is implemented on Parkinson disease dataset from UCI machine learning repository dataset, HandPD Spiral dataset, and HandPD Meander Dataset. Machine learning models such as -KNN, Decision Tree, Random Forest, LR, SVM, Bagging, Voting Soft (KNN, SVM and DT as base classifier) are implemented on these datasets after proposed feature selection process. The aim of the experiment is to find how accurately a machine learning model can predict the occurrence of Parkinson disease disorder with the selected feature subset. In the correlation-based filter method, Pearson's Correlation is applied to the original datasets and correlation matrices are generated as shown in figure 3.5, 3.6 and 3.7. The features obtained after filter-based correlation feature selection process is listed below:

- **Feature subset for UCI Dataset**

MDVP:Fo(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz), MDVP:Jitter(%), RPDE,DFA, spread1, spread2 , D2, STATUS

- **Feature subset for Hand-PD Spiral dataset**

CLASS_TYPE, AGE, RMS, MAX_BETWEEN_ET_HT, MIN_BETWEEN_ET_HT, MIN_HT, STD_DEVIATION_ET_HT, CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT, MRT, MAX_HT.

- **Feature subset for Hand-PD Meander dataset**

CLASS_TYPE, AGE, RMS, MAX_BETWEEN_ET_HT, MIN_BETWEEN_ET_HT, MIN_HT, STD_DEVIATION_ET_HT,CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT, MRT, MAX_HT.

After having a feature subset from correlation-based feature selection, it is used for sequential forward based feature selection process. As sequential forward based feature selection is a wrapper-based method, KNN (k=5) is used as classifier for wrapper-based feature selection. The final subset size is mentioned for all datasets is 5 excluding the classification feature. The feature subsets obtained after sequential forward feature selection is mentioned below:

- **Feature subset for UCI Dataset after Sequential forward Feature Selection**

MDVP:Fo(Hz), MDVP:Fhi(Hz), MDVP:Flo(Hz), MDVP:Jitter(%), D2

- **Feature subset for Hand-PD Spiral dataset after Sequential forward Feature Selection**

MAX_BETWEEN_ET_HT,MIN_BETWEEN_ET_HT,STD_DEVIATION_ET_HT, CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT, MRT.

- **Feature subset for Hand-PD Meander dataset after Sequential forward Feature Selection**

MAX_BETWEEN_ET_HT,MIN_BETWEEN_ET_HT,STD_DEVIATION_ET_HT, CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT, MRT.

The final feature subsets are processed for model training and testing. The datasets are split into 80:20 for training and testing purposes. Equations of all evaluation metrics are mentioned below:

$$\textbf{Accuracy} = \frac{TN+TP}{TN+TP+FN+FP}\ \% \tag{3.8}$$

$$\textbf{Sensitivity (True Positive Rate)} = \frac{TP}{TP+FN}\ \% \tag{3.9}$$

$$\textbf{Specificity (True Negative Rate)} = \frac{TN}{TN+FP}\ \% \tag{3.10}$$

$$\textbf{Precision (Positive predicted Rate)} = \frac{TP}{TP+FP}\ \% \tag{3.11}$$

$$\textbf{F1 Score} = \frac{TP}{TP+\frac{1}{2}(FP+FN)} \tag{3.12}$$

Table 3.2 shows the performance of the proposed Correlation-Forward Feature selection-based method on machine learning models on three datasets.

Model Name	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**CrossVal Score**
LR	80.0	84.62	73.33	78.57	80.25
SVM(Linear)	83.33	91.67	73.33	81.48	80.25
KNN(K=5)	86.67	84.38	90.0	87.1	91.88
Random Forest	91.67	96.3	86.67	91.23	94.33
Decision Tree	86.67	84.38	90.0	87.1	93.24
Bagging	91.67	96.3	86.67	91.23	91.94
Voting Soft	86.67	92.31	80.0	85.71	90.33

Table 3.2: Performance Evaluation on UCI Parkinson Dataset

From table 3.2, it is evident that Bagging and Random Forest Classifier performs best on UCI Parkinson Dataset with an accuracy of 91.67%. The cross-validation score for Bagging and Random Forest are 91.94% and 94.33%. Along with RF and Bagging classifier, Decision Tree, KNN, Voting Soft classifier exhibit good accuracy of 86.67%. Cross validation scores for DT, KNN, and Voting Soft classifiers are 93.24%,91.88% and 90.33% respectively. The confusion matrix that is illustrated in figure 3.9 shows that Random Forest Classifier correctly classified 29 True Negative instances, and 26 True Positive instances. Only 5 instances are classified as wrongly.

Model Name	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**CrossVal Score**
LR	75.47	71.43	89.29	79.37	83.38
SVM(Linear)	90.57	89.66	92.86	91.23	96.67
KNN(K=5)	90.57	89.66	92.86	91.23	96.67
Random Forest	94.34	90.32	100	94.92	99.05
Decision Tree	96.23	93.33	100	96.55	99.07
Bagging	94.34	90.32	100	94.92	99.05
Voting Soft	88.33	92.59	83.33	87.72	85.22

Table 3.3: Performance Evaluation on HandPD Spiral Dataset

In case of Hand-PD spiral dataset, Decision Tree obtained highest accuracy of 96.23%with a cross validation score of 99.07%. Random forest and Bagging Classifier performed well and achieved accuracy of 94.34% and cross validation score of 99.05%. From figure 3.10, the confusion matrix shows the number of True Positive and True Negative cases are 23 and 28 respectively. So, the proposed feature selection approach is fair enough for Parkinson disease detection for Hand-PD spiral dataset.

Model Name	**Accuracy**	**Precision**	**Recall**	**F1 Score**	**CrossVal Score**
LR	83.02	90.0	72.0	80.0	81.99
SVM(Linear)	84.91	90.48	76.0	82.61	81.49
KNN(K=5)	86.79	87.5	84.0	85.71	76.73
Random Forest	86.79	95	76.0	84.44	80.09
Decision Tree	71.7	67.86	76.0	71.7	79.59
Bagging	75.47	87.5	56.0	68.29	80.54
Voting Soft	71.7	77.78	56.0	65.12	77.75

Table 3.4: Performance Evaluation on HandPD Meander Parkinson Dataset

For Hand-PD meander dataset, KNN and Random Forest have achieved better accuracy of 86.79% than other classifier used. The confusion matrix for this dataset is shown in figure no. 3.11. Here, KNN has manages to predict 25 True Negative cases and 21 True positive cases whereas RF classified 27 cases as True Negative and 19 as True Positive.

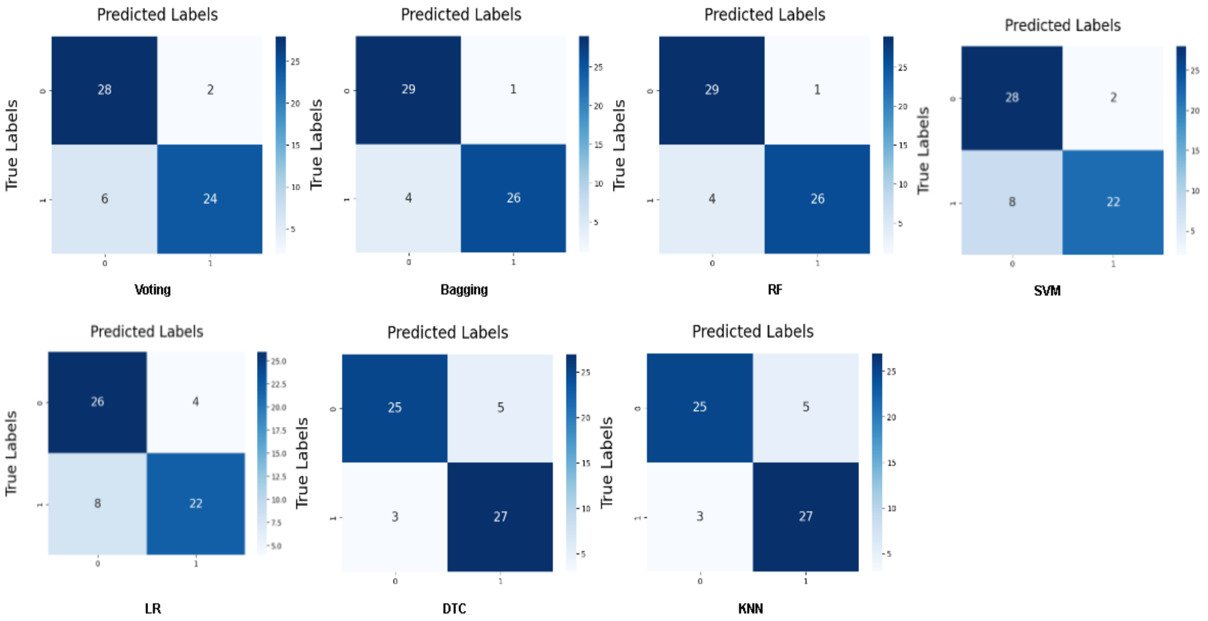


Figure 3.9: Confusion Matrix for KNN, SVM, LR, DTC, RF, Bagging, Voting Soft classifiers for UCI Parkinson Dataset

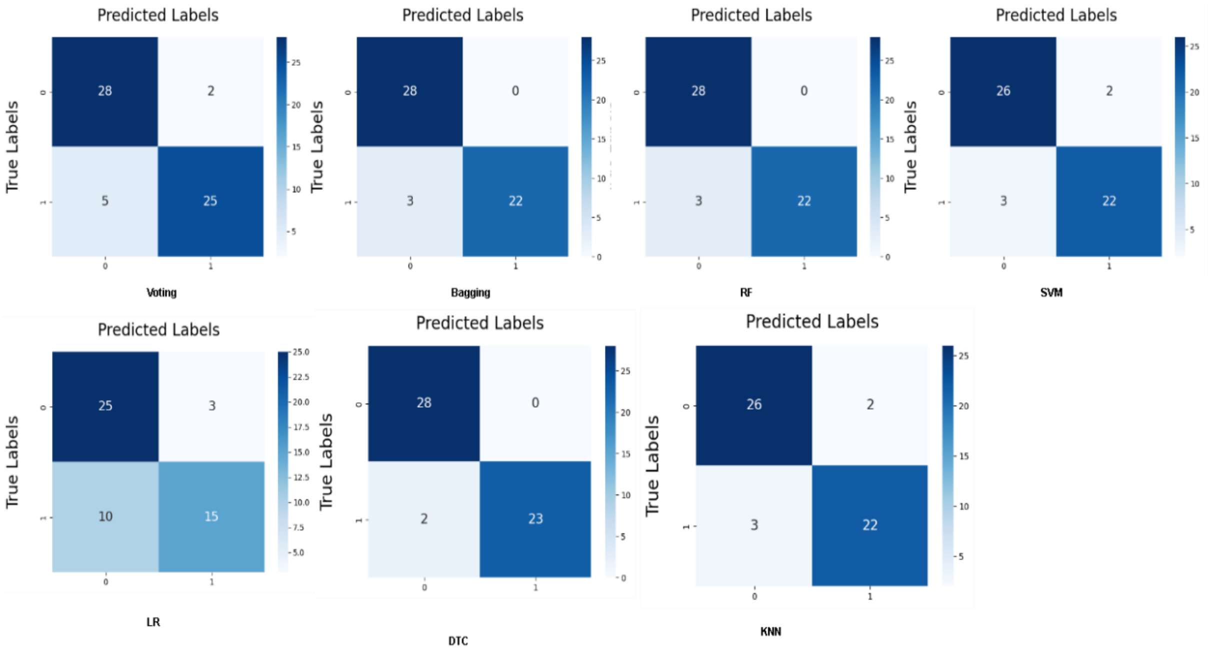


Figure 3.10: Confusion Matrix for KNN, SVM, LR, DTC, RF, Bagging, Voting Soft classifiers for HandPD Spiral Parkinson Dataset

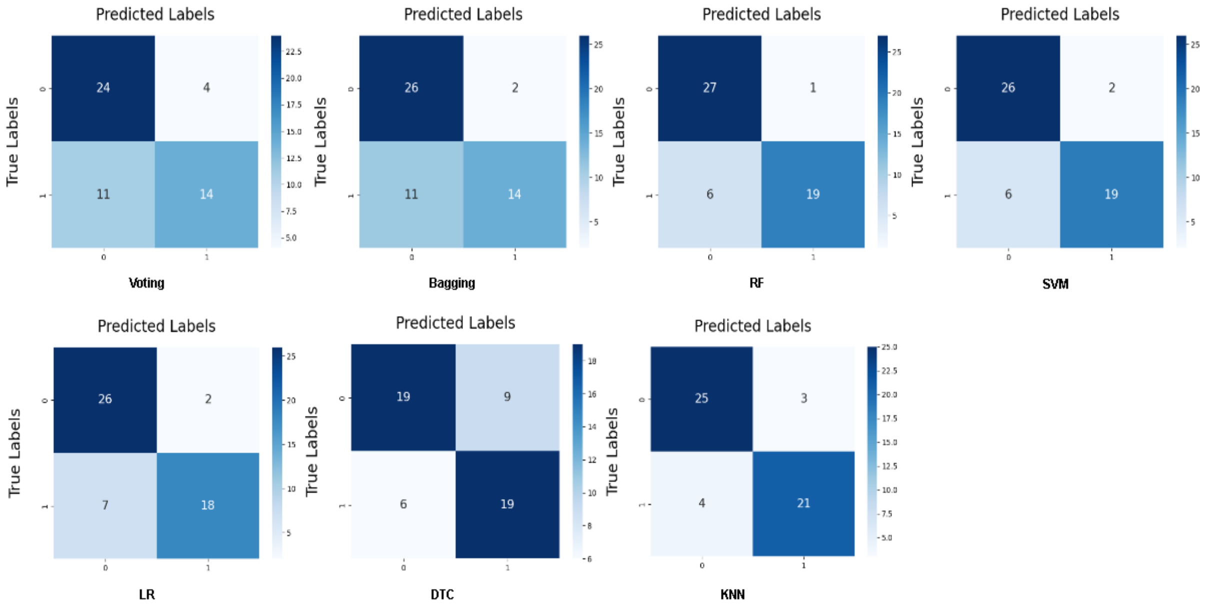


Figure 3.11: Confusion Matrix for KNN, SVM, LR, DTC, RF, Bagging, Voting Soft classifiers for HandPD Meander Parkinson Dataset

The proposed Correlation – Sequential forward Based Feature Selection approach performed well for HandPD spiral and UCI dataset. This provides an insight how feature selection can help in performance enhancement for ML applications. Further feature selection helps to quicken the execution process, thus helping in faster health data evaluation. The drawback of the proposed approach is that in Forward Selection based Wrapper method, the number of features is required to be mentioned beforehand. In this case, it is mentioned as 5. So, the final feature subset will have five features. In this process there might be a scenario where some important features might get missed due to feature number limitation. In next section, another Feature Selection based ML application for Parkinson Data Analysis is proposed. This will mitigate the deficiencies of this model.

3.5 Case study 2: A Correlation-Whale Optimization based Feature Selection Approach for Parkinson disease prediction

3.5.1 Whale Optimization based Wrapper feature Selection Method

Obtaining an optimal subset after feature selection method is a crucial step for machine learning. In exhaustive approach, all possible subsets can be generated to obtain the best subset. But if the dataset size is big with N number of features, then 2N solutions will be generated which will be time consuming for evaluating result with high computational cost. Another way to find the best subset is the heuristic approach which uses heuristic information for optimal search. But heuristic search does not guarantee optimal solution. It helps to find acceptable solution in a stipulated time. Different Nature inspired algorithms such as swarm based, physics based, and evolutionary based are mostly meta heuristic in nature. Moreover, many researchers have tried to use stochastic methods for solving feature selection problems.

Whale Optimization Algorithm is based on stochastic population-based algorithm proposed by Mirjalili and Lewis in the year 2016 [168]. It used the concept of bubble net feeding for the foraging behavior of humpback whale. Humpback whale hunts for its food close to the surface while trapping the prey in a bubble-like net. They swim in a 6-shape path for creating the net.

The algorithm is implemented in three phases: in first phase is also called prey encircling, second phase is called spiral bubble net attacking phase where a spiral net is formed for encircling the prey, and the third phase or exploration phase, searching of the prey is randomly done. The detail of each phase is discussed below [168].

Phase 1: Prey Encircling Phase

In this phase, the algorithm is initialized with an initial best search agent. It is assumed that the current solution is best, and it assumes the location of the prey is almost close to the solution. Remaining agents consequently update their position in search of the location towards the best search agent. This idea can be expressed as:

$$\vec{D} = |\vec{C}.\vec{X}^*(t) - \vec{X}(t)| \quad (3.13)$$

$$\vec{X}(t+1) = \vec{X}^*(t) - \vec{A}.\vec{D} \quad (3.14)$$

Where t shows current iteration, X* represents current best solution obtained so far, X represents current situation, || is absolute value, represents element by element multiplication. A and C are coefficient vectors that are calculated as:

$$\vec{A} = 2\vec{a}.\vec{r} - \vec{a} \quad (3.15)$$

$$\vec{C} = 2.\vec{r} \quad (3.16)$$

Where value of $\vec{a}$ decreases linearly from 2 to 0, over the number of iterations, and r is a random vector whose value lies between [0,1].

Here the modeling helps the agent in location updating in region for current best solution while imitating encircling the prey. It can further extend its search space in n dimensions, and the hyper cubes movement will be done by agent while searching for the best solution.

Phase 2: Bubble Net Attacking Phase

In Bubble Net Attacking Phase or exploitation phase, two steps are involved.

Step 1: Shrink of Encircle Technique

Here the value of $\vec{a}$ is decreased which leads consequently to decreasing fluctuation range of $\vec{A}$ by $\vec{a}$. This act suggests that a is randomly placed in [-$\vec{a}$, $\vec{a}$] where a value decreases from 2 to 0 over the span of optimization time. As the randomness of $\vec{A}$ lies in between [-1,1], the search agent's location can be analyzed based on agent's past location and agent's current best location. Following figure shows how the possible location is shifting from (X, Y) to (X^*, Y^*) based on 0, A and 1 in a 2-dimentional space.

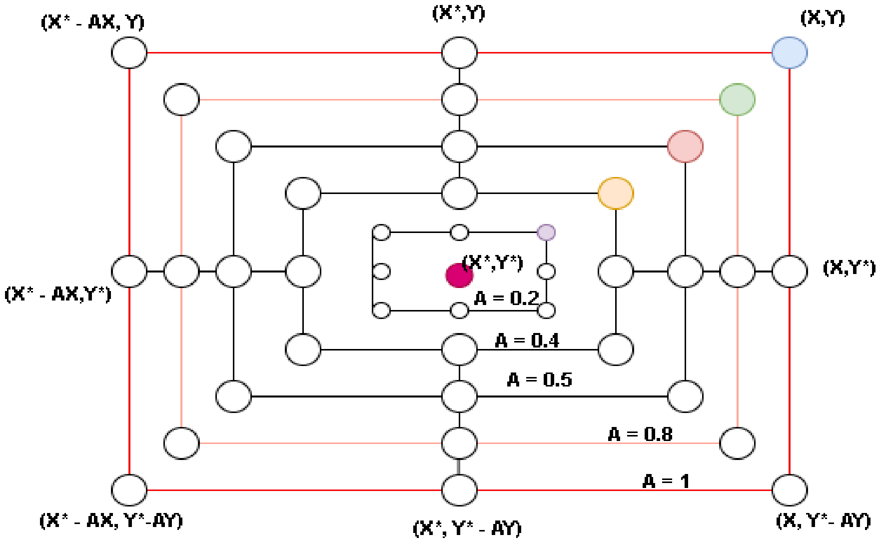


Figure 3.12: WOA Shrinking of Encircle Technique [167]

Step 2: Updating of Spiral Position

Here, the distance between the position of the whale and its prey is calculated. The equation of the spiral is formulated from the calculated distance and the movement of the whale is simulated in helix shape. The movement can be equated as:

$$\vec{X}(t+1) = \vec{D}^{t}.e^{bl}.\cos(2\Pi l) + \vec{X}^{*}(t) \qquad (3.17)$$

$$\vec{D}' = |\vec{X}^{*}(t) - \vec{X}(t)| \qquad (3.18)$$

Where b represents constant for obtaining a logarithmic spiral shape, l represents a random number in the region range between [-1, 1]. Equation 18 represents the distance between i^{th} whale and its prey. The whale moves for its prey concurrently by applying circle shrinking in a spiral path. To implement two mechanisms – shrinking of encircle and upward spiral path calculation, there is 50% chances to choose between two conditions during optimization process as discussed below:

$$\begin{array}{ll} \vec{X}(t+1) = \vec{X}^{*}(t) - \vec{A}.\vec{D}, & \text{if } p < 0.5 \\ \vec{X}(t+1) = \vec{D}^{t}.e^{bl}.\cos(2\Pi l) + \vec{X}^{*}(t), & \text{if } p \geq 0.5 \end{array} \qquad (3.19)$$

Where p denotes a random number range from [-1,1].

Phase 3: Exploration Phase:

In this phase, WOA gains global optimization. Here whales search prey according to their location from each other randomly. The value of A is assigned randomly from [-1,1] to help the search agent to move away from the reference whale. It means the value of A should be greater than +1 or lesser than -1. Position of the search agent is randomly assigned which allows WOA to do global search. The process of exploration is expressed in equation as below:

$$\vec{D} = |\vec{C}.X_{rand} - \vec{X}|, \tag{3.20}$$

$$\vec{X}(t+1) = \vec{X}_{rand} - \vec{A}.\vec{D}, \tag{3.21}$$

Where $\vec{X}_{rand}$ is a random position for a random whale which has been selected from current location. Figure 3.13 shows the mechanism of WOA exploration.

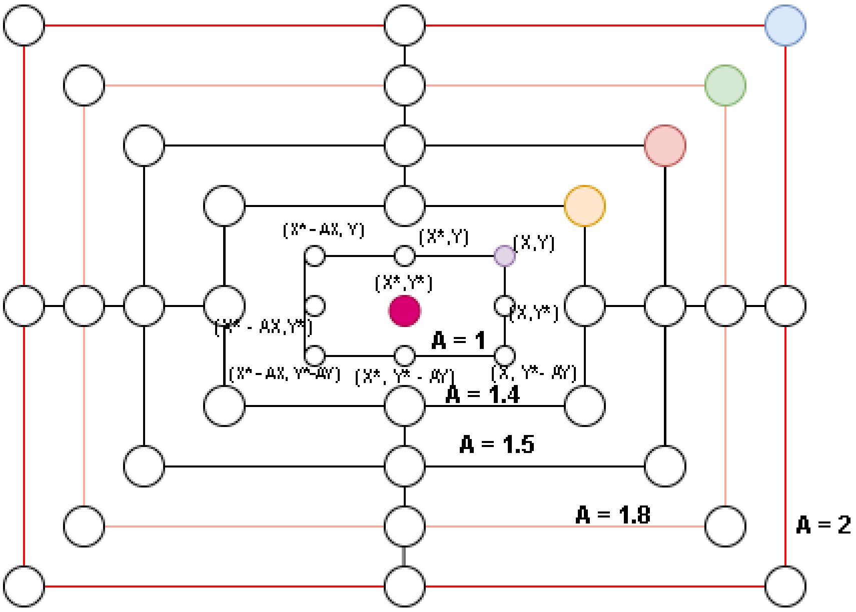


Figure 3.13: WOA Exploration Algorithm [167]

3.5.2 Proposed Correlation – WOA based Feature Selection Approach for Parkinson Disease Classification

Whale optimization-based feature selection method is implemented here as a wrapper-based feature selection method. In wrapper method, a classification algorithm is used to guide the feature selection mechanism based on some optimization algorithm. Therefore, here for implementation purpose, KNN(K=5) classifier is used. In this work, a hybrid approach of Filter - WOA based feature selection method is taken. For filter method, where feature selection is independent of any classification model, Pearson Correlation Coefficient is selected as the filter. If the correlation between a feature and class level is higher than a threshold value, the feature is selected for feature subset. Highly correlated feature will affect classification accuracy. Presently, if features are correlated above 75% with the class level, that feature is selected.

After applying filter-based feature selection as mentioned above, Whale Optimization Algorithm is implemented as wrapper-based method to select optimal feature subset. From the proposed WOA feature selection approach, feature subset is continuously updating until best optimized feature subset is obtained. It mimics the concept of the whales that are continuously trying to update their position in search of best agent based on equation no. (3.13) and (3.14). The solution is a continuous vector with the same dimension as the dataset. The solution vector has continuous value and bounded by value [0,1]. In the present implementation KNN, SVM, LR, Random Forest, Decision Tree, Bagging, Voting Soft are used as classifiers and these are used to save selected feature set. The fitness function used here is based on the classification performance and feature subset size. Fitness evaluation values are represented in binary format. Hence, the Fitness function combining the classification performance and feature subset size is represented as [167]:

$$\mathbf{f_\theta = \alpha. E + (1- \alpha)\frac{\Sigma i\, \theta i}{N}} \qquad \mathbf{(3.22)}$$

where N is the number of features obtained after Filter based Pearson Correlation Analysis, f_θ is the fitness function, where θ is a feature vector of size N with 0/1 elements that describes selected or unselected feature subset. E is a classification error and α is a constant that manages the tradeoff between classification error and number of features in the subset. To enhance the classification accuracy, we have set α is set to 1. Figure 3.14 describes whole workflow of proposed hybrid filter-WOA based feature selection approach and figure 3.15 represents a flowchart of WOA feature selection process.

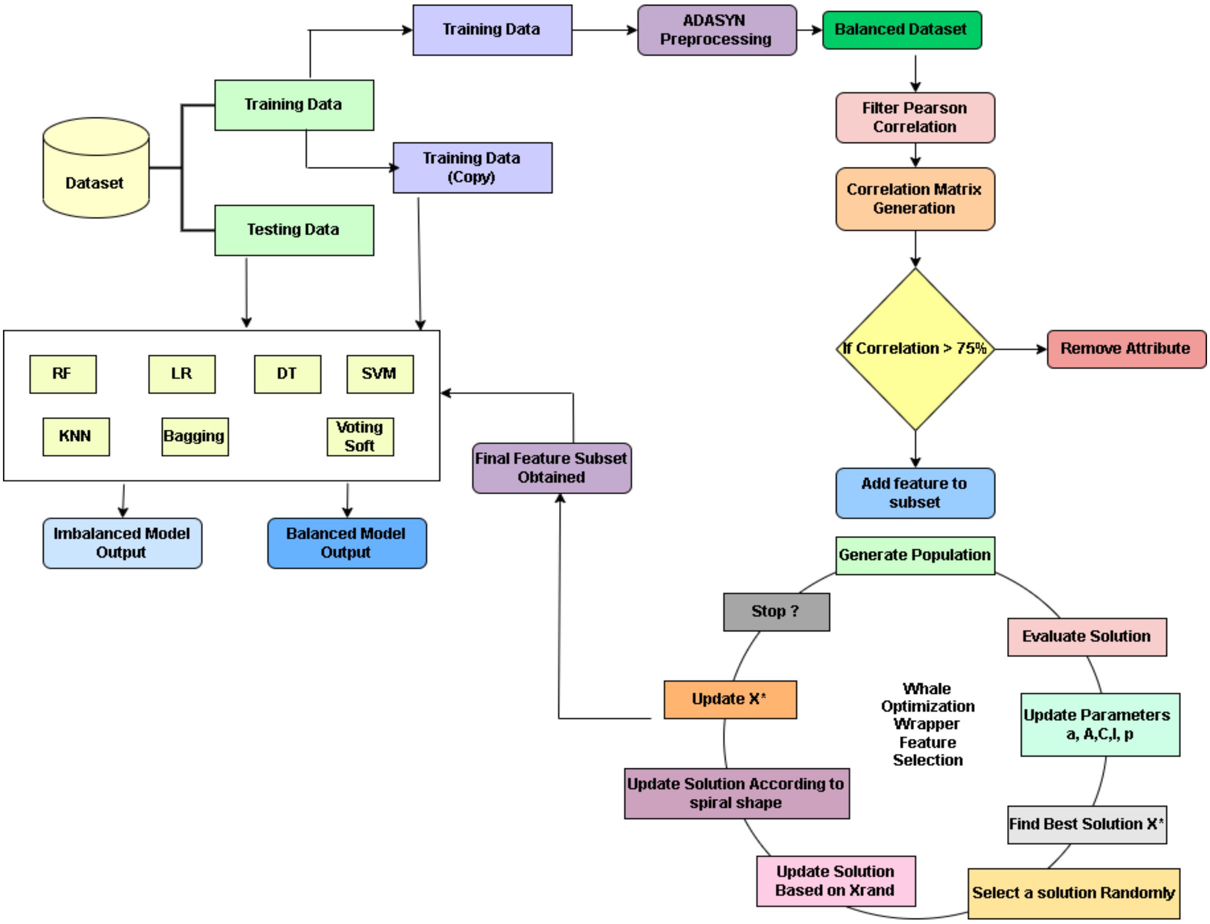


Figure 3.14: Proposed Correlation -WOA Based Feature Selection Approach

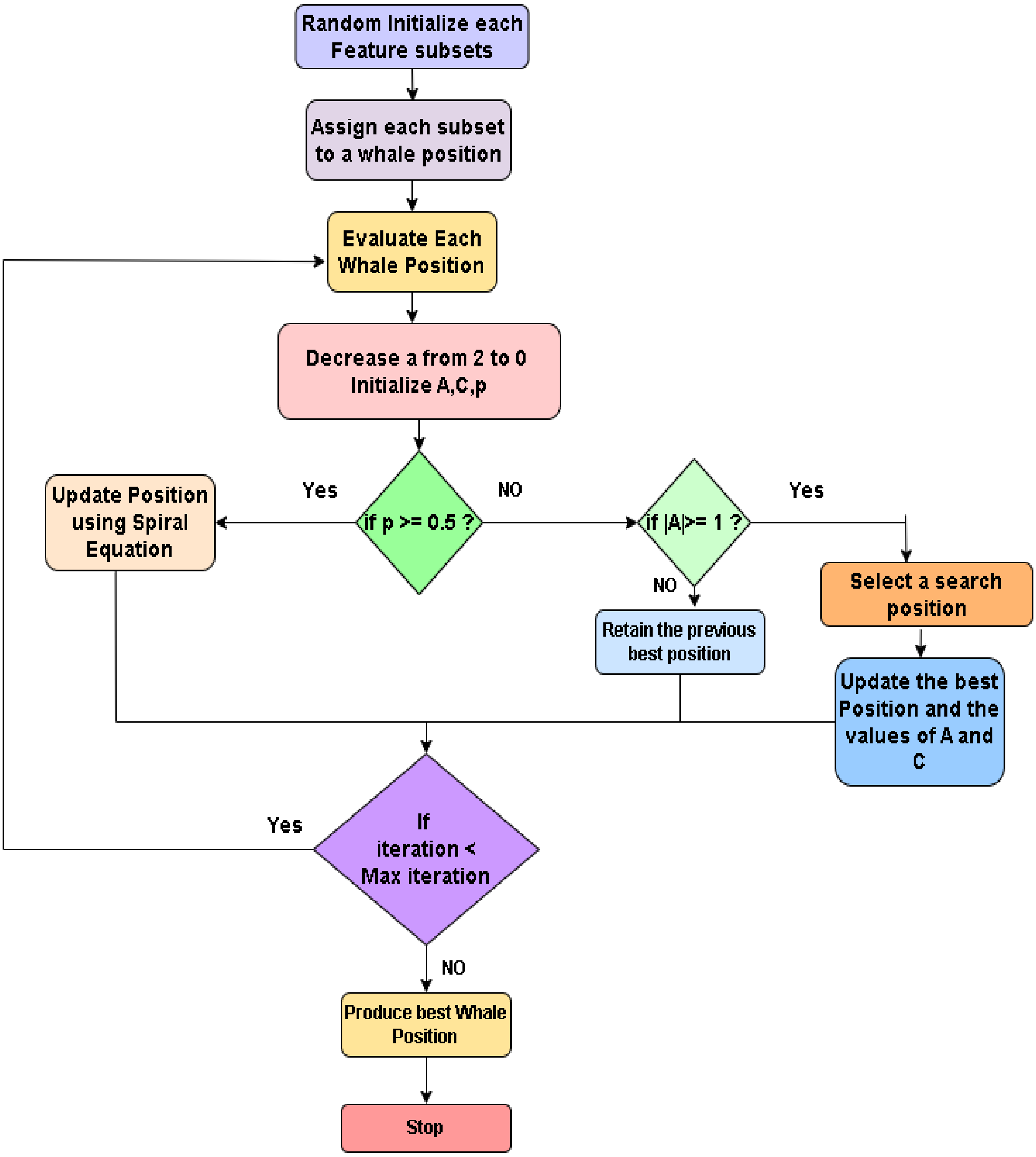


Figure 3.15 : flowchart of WOA feature selection process.

3.5.3 Implementation Result and Analysis

The experiments are stimulated using python tools Pandas, Scikit_learn, Matplot, Numpy, Seaborn libraries in Google Colab platform. The model is tested on three datasets – UCI Parkinson Disease Dataset, and HandPD dataset. For comparison purpose, the experiment is first conducted on full dataset where no data preprocessing or feature selection is done. Each machine learning model considered in this work, Linear Regression, Decision Tree, Support Vector Machine, Random Forest, K-Nearest Neighbor, Bagging, Voting-Soft are implemented on full dataset and evaluated based on evaluation metrics – confusion matrix, accuracy, Precision, F1 score and ROC-Area Under Curve (AUC). Receiver Operating Characteristics or ROC [32] curve is a probabilistic curve and AUC is used to measures the area under this curve. It is a used to measure the ability of a model to distinguish between classes.

Model	Accuracy	Precision	Recall	F1 score	Cross Validation	ROC-AUC
LR	69.8	66.67	72.0	69.23	75.28	.737
SVM	66.04	62.07	72.0	66.67	74.83	.727
DT	58.49	56.0	56.0	56.0	76.28	.583
RF	77.36	78.26	72.0	75.0	75.76	.809
KNN	75.47	71.43	80.0	75.47	71.99	.732
Bagging	79.25	79.17	76.0	77.55	75.3	.901
Voting Soft	67.92	65.38	68.0	66.67	75.78	.732

Table 3.5: Machine Learning models performance on UCI Parkinson dataset without oversampling

Model	Accuracy	Precision	Recall	F1 score	Cross Validation	ROC-AUC
LR	75.47	77.78	75.0	76.36	76.75	.76
SVM	75.47	80.0	71.43	75.47	76.3	.76
DT	71.7	72.41	75.0	73.68	72.99	.715
RF	75.47	75.86	78.57	77.9	76.32	.85
KNN	79.25	84.0	75.0	79.25	68.72	.81
Bagging	73.58	71.88	82.14	76.77	75.8	.84
Voting Soft	69.81	71.43	71.43	71.43	72.03	.80

Table 3.6: Machine Learning model performance on HandPD – Spiral dataset without oversampling

Model	**Accuracy**	**Precision**	**Recall**	**F1 score**	**Cross Validation**	**ROC-AUC**
LR	66.04	60.61	80	68.97	74.81	.748
SVM	62.26	57.14	80	66.67	74.33	.748
DT	77.36	74.07	80	76.92	79.59	.799
RF	79.25	76.92	80	78.43	79.09	.865
KNN	73.58	67.74	84	75.0	79.57	.821
Bagging	83.02	80.77	84	82.35	81.45	.857
Voting Soft	81.67	83.1	80.0	81.53	75.14	.843

Table 3.7: Machine Learning model performance on HandPD- Meander dataset without oversampling

Tables 3.5,3.6,3.7 shows the results of the models directly applied on the imbalanced dataset without any feature selection method. Here the models are trained on the entire dataset with all features and without any feature selection method. Now the proposed method has been implemented on three datasets. From figures 3.5,3.6,3.7 the correlation matrices are obtained for all three datasets. Features having correlation greater than 75% with output label are selected and obtained a first feature subset. The size and details of selected features are described in section 3.4.3. Whale Optimization Algorithm is implemented to further create a subset of features that are optimized and will generate maximum accurate performance for each classifier. KNN, SVM, LR, DT, RT, Bagging, Voting Soft Classifiers are implemented on the feature subsets generated for three datasets considered – UCI Parkinson Dataset, HandPD Spiral, and HandPD Meander.

Final feature subset size for all three datasets as computed by Whale optimization-based wrapper feature selection method is mentioned below.

- **Final Feature subset for UCI Dataset**

MDVP:Flo(Hz), MDVP:Jitter(%), RPDE,DFA, D2

- **Final Feature subset for Hand-PD Spiral Dataset**

MAX_BETWEEN_ET_HT, MIN_BETWEEN_ET_HT, STD_DEVIATION_ET_HT, MRT.

- **Final Feature subset for Hand-PD Meander Dataset**

MAX_BETWEEN_ET_HT,MIN_BETWEEN_ET_HT,CHANGES_FROM_NEGATIVE_TO_POSITIVE_BETWEEN_ET_HT

The table 3.8,3.9 and 3.10 show details of all performance measure for three datasets.

Model	Accuracy	Precision	Recall	F1 score	Cross Validation	ROC-AUC
LR	81.13	90.91	71.43	80.0	75.78	.884
SVM	81.13	90.91	71.45	80.4	76.26	.871
DT	83.02	80.65	89.29	84.75	77.21	.826
RF	86.79	92.0	82.14	86.79	80.54	.897
KNN	88.68	95.83	82.14	88.46	74.85	.905
Bagging	84.91	88.46	82.14	85.19	77.21	.895
Voting Soft	86.79	95.65	78.57	86.27	78.66	.933

Table 3.8: Performance evaluation of proposed Correlation – WOA based feature Selection method on Hand-PD Spiral Parkinson Dataset

Model	Accuracy	Precision	Recall	F1 score	Cross Validation	ROC-AUC
LR	83.33	81.25	86.67	83.87	80.92	.894
SVM	80.0	82.14	76.67	79.31	78.82	.897
DT	90.0	92.86	86.67	89.66	80.92	.900
RF	95.0	96.56	93.33	94.92	88.97	.992
KNN	91.67	90.32	93.3	91.8	83.06	.946
Bagging	96.67	100	93.33	96.55	86.67	.996
Voting Soft	91.67	93.1	90.0	91.53	85.14	.943

Table 3.9: Performance evaluation of proposed Correlation – WOA based feature Selection method on UCI Parkinson Dataset

Model	Accuracy	Precision	Recall	F1 score	Cross Validation	ROC-AUC
LR	81.67	85.19	76.67	80.7	80.16	.866
SVM	81.67	85.19	76.67	80.7	76.06	.860
DT	83.33	88.46	76.67	82.14	81.87	.833
RF	88.33	92.59	88.33	87.72	85.22	.963
KNN	93.33	96.43	90.0	93.1	80.14	.956
Bagging	86.67	92.31	80.0	85.71	82.26	.938
Voting Soft	86.67	92.31	80	85.71	86.67	.911

Table 3.10: Performance evaluation of proposed Correlation – WOA based feature Selection method on hand-PD Meander Parkinson Dataset

From the table 3.8, it is seen that KNN classifier works better than all classifiers with accuracy of 88.68% for Hand-PD spiral dataset and AUC_ROC value is .905. For UCI dataset Bagging worked well with accuracy of 96.67% and cross validation score of 86.67%, and AUC_ROC value is .996. In case of Hand-PD Meander dataset, KNN has obtained highest accuracy of 93.33% with AUC_ROC value of .956. Therefore, for three datasets, the proposed approach achieved better accuracy than previous filter -sequential based feature selection approach.

Figures 3.16, 3.17 and 3.18 shows all the confusion matrices for all classifiers for three datasets. The proposed model can perform well for all three datasets in terms of TP and TN cases. For comparison purpose, few state of art models represented in different works have been considered as shown in table 3.11 for the UCI and Hand-PD datasets. The proposed approach performs better than the considered models with better accuracy and ROC values. So, the proposed approach can be helpful for emergency healthcare data evaluation.

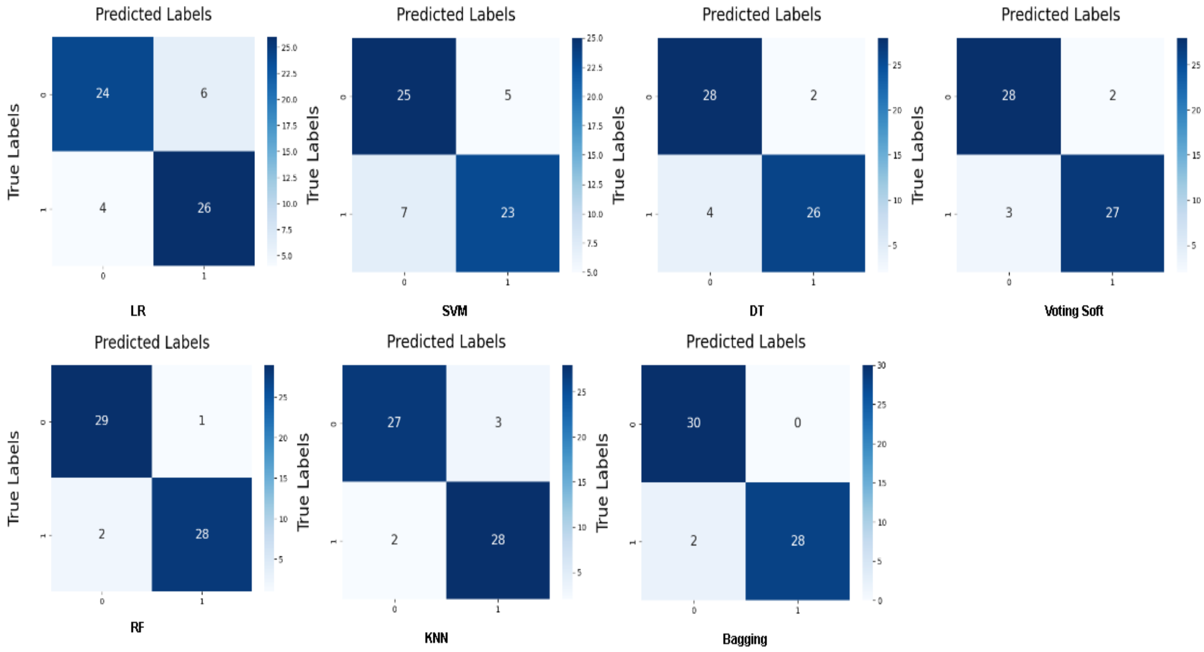


Figure 3.16: Confusion matrix for UCI dataset for proposed correlation - WOA feature selection approach.

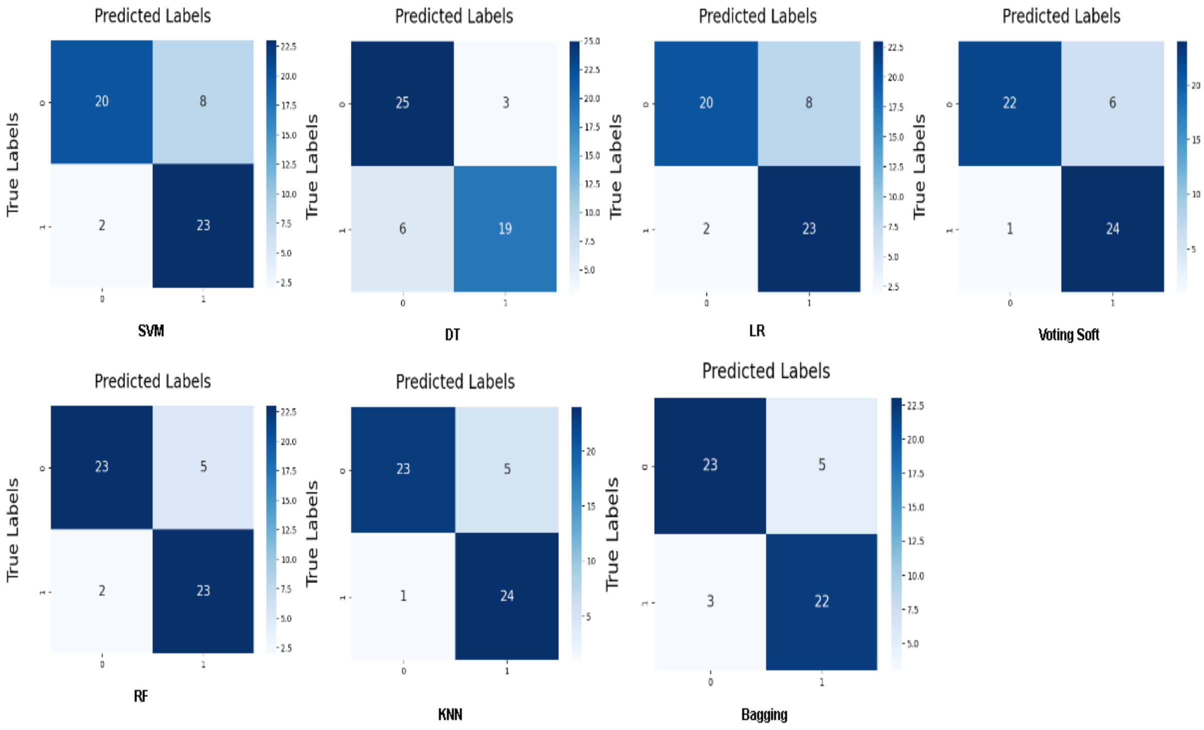


Figure 3.17: Confusion Matrix for Hand-PD Spiral Dataset for proposed correlation-WOA feature selection approach.

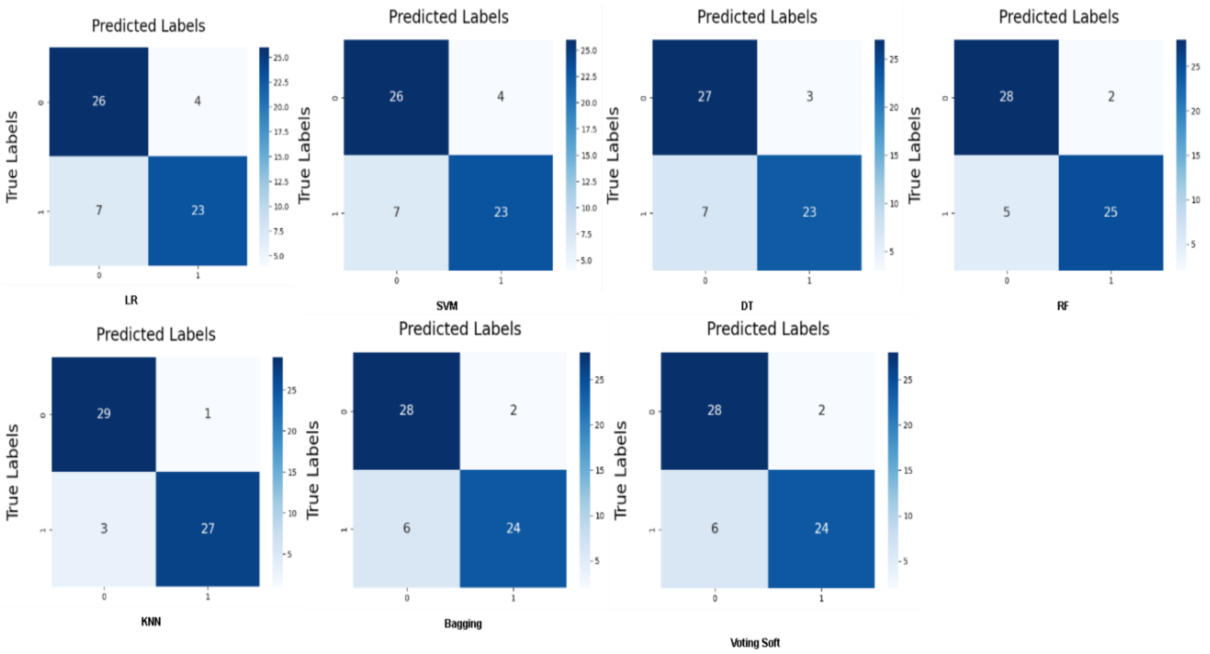


Figure 3.18: Confusion Matrix for Hand-PD Meander Dataset for proposed correlation -WOA feature Selection approach.

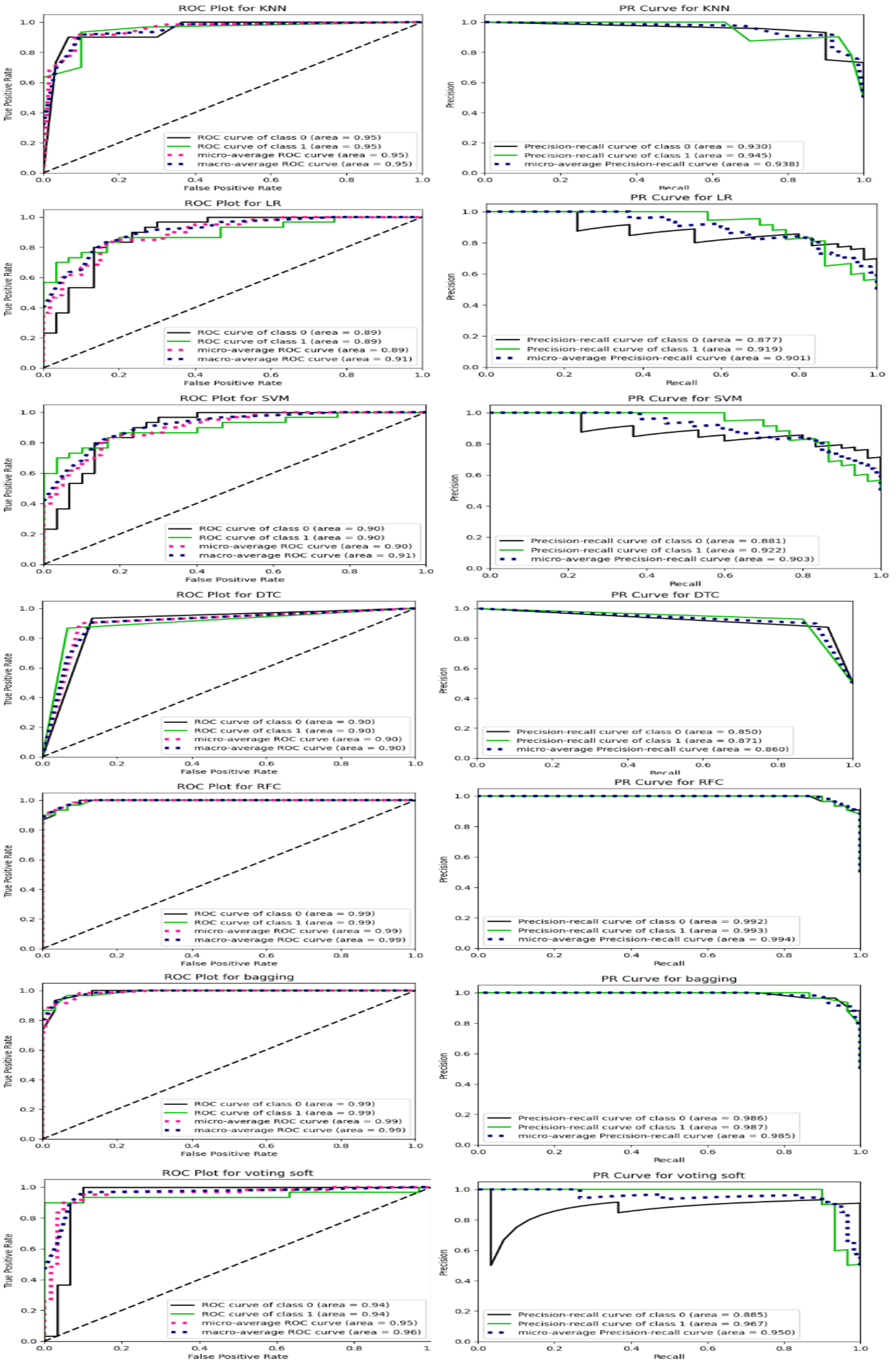
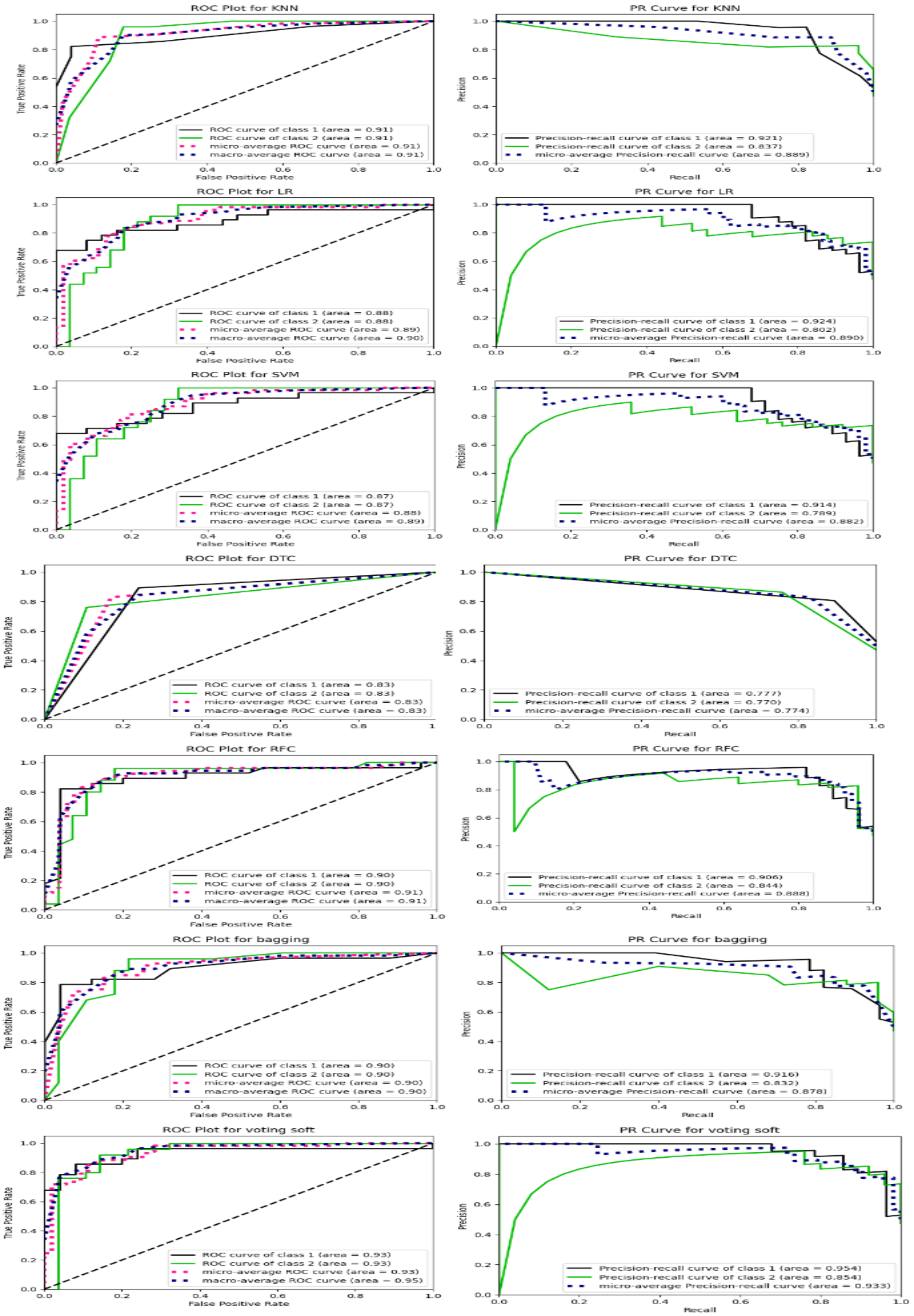


Figure 3.19: ROC- PR graph of UCI dataset


ROC Plot for KNN

PR Curve for KNN

ROC Plot for LR

PR Curve for LR

ROC Plot for SVM

PR Curve for SVM

ROC Plot for DTC

PR Curve for DTC

ROC Plot for RFC

PR Curve for RFC

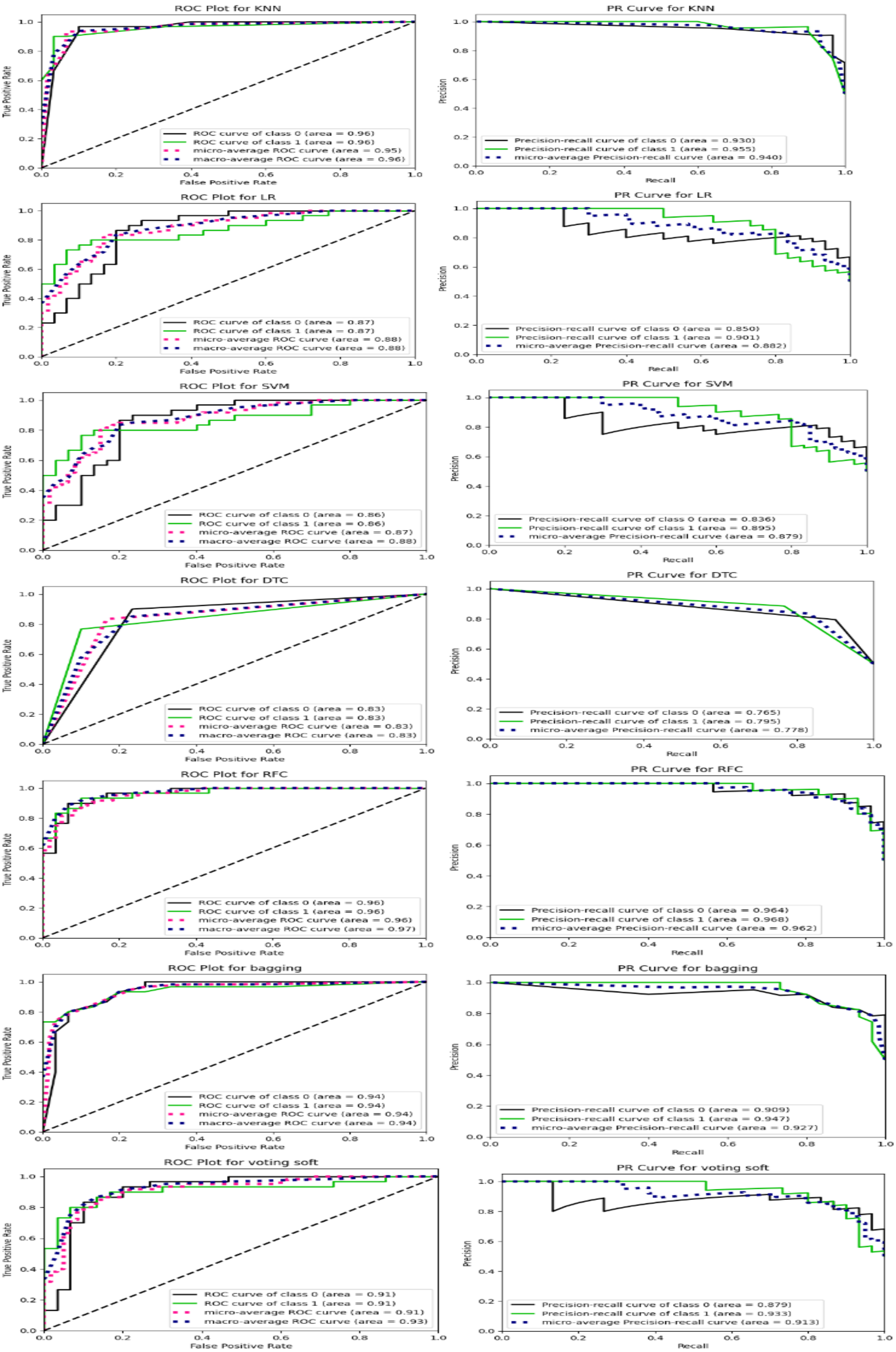
ROC Plot for bagging

PR Curve for bagging

ROC Plot for voting soft

PR Curve for voting soft

Figure 3.20: ROC-PR graph of Hand-PD spiral dataset


ROC Plot for KNN

PR Curve for KNN

ROC Plot for LR

PR Curve for LR

ROC Plot for SVM

PR Curve for SVM

ROC Plot for DTC

PR Curve for DTC

ROC Plot for RFC

PR Curve for RFC

ROC Plot for bagging

PR Curve for bagging

ROC Plot for voting soft

PR Curve for voting soft

Figure 3.21: ROC-PR graph for Hand-PD Meander dataset

State of the art Model	Dataset Name	Year	Model Details	Performance Evaluation
Ali et al. [152]	HandPD	2019	Chi2 with AdaBoost	Accuracy – 76.44%, Sensitivity – 70.94%
De Souza et al. [153]	HandPD	2018	Naïve Bayes, SVM with cross-validation	Accuracy – 85.54%
Sheibani et al. [154]	UCI Dataset	2019	Ensemble learning with 10 cross validations	Accuracy – 90.6%, Sensitivity – 95.8%
Yaman et al. [155]	UCI Dataset	2020	KNN, SVM	SVM: Accuracy = 91.25%, Precision = 0.9125 Recall = 0.9125
Senturk et al. [156]	UCI Dataset	2020	CART, SVM, ANN	SVM accuracy – 93.84%
Gunduz et al. [157]	UCI Dataset	2019	CNN	cross validation Accuracy = 0.869
Sayaydeha et al.[158]	UCI Dataset	2019	EFMM-OneR	cross validation Accuracy = 94.21%
Ribeiro et al[159]	HandPD	2019	Bidirectional Gated Recurrent Units	The Spiral dataset: Accuracy = 89.48% The Meander dataset: Accuracy = 92.24%
Anand et al [160]	UCI Dataset	2018	Logistic regression, KNN, naïve Bayes, SVM, DT, RF, DNN with 10-fold cross validation	KNN accuracy = 95.513%
Marar et al[161]	UCI Dataset	2018	Logistic regression, KNN, SVM, naïve Bayes, DT, RF, ANN	Accuracy = 94.87% Specificity = 96.55%
Proposed Filter - WOA based Feature selection	UCI dataset, Hand-PD Spiral, Hand-PD Meander dataset	2023	KNN, SVM, DT, LR, RF,Bagging, Voting Soft	UCI: Bagging – Accuracy -96.67, Spiral – KNN- 88.68%, Meander – KNN – 93.33%

Table 3.11: Performance Comparison with State of art models with proposed approach

3.6 Discussion

The problem in current healthcare industry is that the patient doctor ratio is not optimum in most of the regions in India. Unavailability of medical person in emergency causes delay in disease detection and thus further delay in treatment. Using an automatic disease detection system will accelerate treatment process. From the proposed approaches it can be summarized that if machine learning is incorporated during disease detection and prediction tasks, healthcare system will be more helpful and robust. It will improve patient outcome and reduce unnecessary tests and treatment thus reducing overall healthcare cost. Healthcare professionals will be able to get a clear insight on risk and emergency situations of the patient. Therefore, overall patient outcome will be improved. Early diagnosis is a boon for patients. Identifying the risk factors for any specific disease will help in improving patient’s condition. Thus, the proposed approaches are reliable with fair performance in all the considered dataset and will be helpful in improving patient’s outcome. In near future, machine learning implementation in healthcare will provide tremendous help for disease prediction.

Although it is now well known that disease diagnosis can be made faster and more accurate with the help of Machine Learning models. However, it does not seem sufficient for many other disease detections. In such cases, particularly, as large number of data (including images) is available, it seems appropriate to utilize deep learning approaches towards disease detection (via classification) for faster diagnosis.

The next chapter details out approaches using Deep Learning taken in the present work for classification.

Chapter 4

Deep Learning based Disease Classification

Nowadays healthcare organizations of every type and size are inclining towards Artificial Intelligence implementation to provide better patient care while reducing treatment cost and improving treatment efficiency. Healthcare is revolving into a new era where abundant data is playing important role to deliver right treatment to right patient in right time. Hence consideration of considering several aspects of patient data including patient's lifestyle data, Electronic Health Record (EHR), molecular traits [169,170] is required. Healthcare industry is incorporating AI based tools to have a firm grasp of health data for disease detection and prediction. Deep learning is a branch of Artificial Intelligence that has very easily became a transformative and reliable source to analyze data with precision. It provides a clear insight of data and relation among its attributes. Deep learning which is also known as hierarchical learning uses a layered algorithmic architecture for analyzing data. Deep Learning architecture is based on biological neurons. As biological neurons relate to one another, deep learning layers are also connected with each other for passing information and each subsequent layer's node is activated when it receives information from its neighboring nodes. In basic deep learning model, each layer is assigned with a specific transformation task and data traverses through the layers to optimize the output. Huge availability of biomedical data imposes tremendous challenges and brings new opportunities in healthcare industry. Analyzing the association among all different data is a basic problem for developing reliable data driven tools. Although existing machine learning models exhibit great performance in medical data, there are still many challenges in full understanding of medical data owing to its high dimensionality, sparsity and irregularity, heterogeneity, temporal dependency. So deep learning in healthcare would be helpful for its enhanced performance, end-to-end learning models with integrated feature learning module. It has the capability to handle large complex multi modal data. Deep learning research should address several challenges related to healthcare data like time-dependent data, sparse data, noisy data, heterogeneous data. So, to accelerate clinical decision support system, improved deep learning methods are necessary to handle healthcare data workflow.

Data representation in deep learning and the way data is learnt, are different from traditional deep learning model. In deep learning, different patterns are learnt from data using multiple processing layer of neurons with multiple levels of abstraction [171]. The main difference between DL and a traditional Artificial Neural Network (ANN) is the number of hidden layers, connection between each layer and the capability to extract meaningful abstractions from the input. Every layer of deep learning results in a representation of pattern extracted based on the data it receives as input from its previous layer.[172]. The key aspect of deep learning model is that the learning of the features is in each layer is done using general purpose learning procedure. In figure 4.1 the difference between deep learning and ANN is shown. While ANN is comprised of three layers and middle layer is responsible for transformation towards the final output, deep learning architectures are comprised of several layers of neural network. Layer wise training allows DL to fine tune efficiency and helps in extracting deep features from inputs to obtain high level features that enhance accuracy of the model.

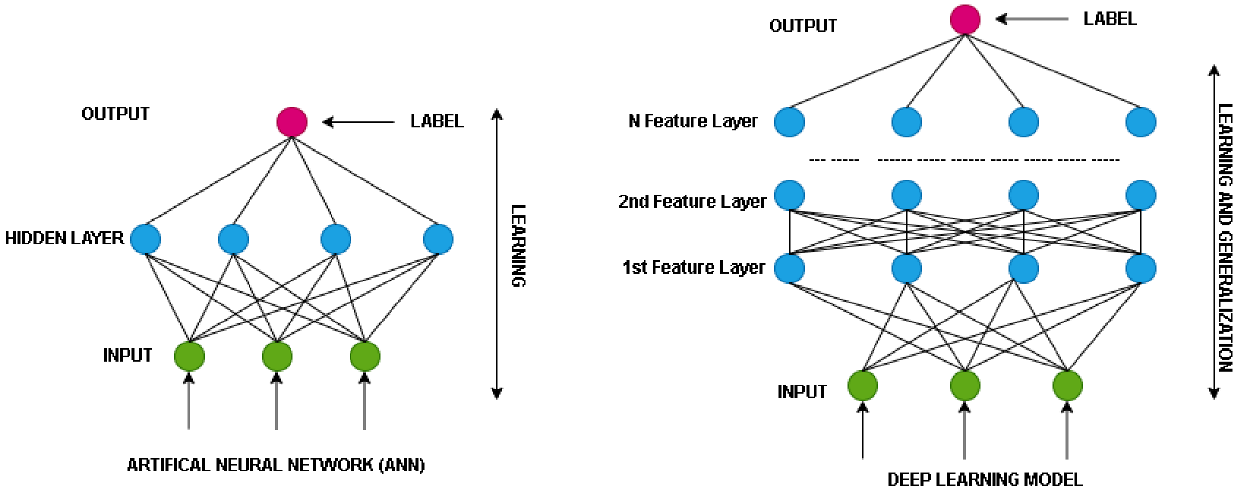


Figure 4.1: Architectural difference between Artificial Neural Network and Deep Learning Model

Therefore, applications of deep learning in medical data highlighting the important aspects will impact healthcare service. Several applications of Deep Learning in healthcare are drug discovery, medical imaging and diagnosis, personalized treatment, to name a few. In this work, Pulmonary Tuberculosis, Covid detection and Bone Fracture detection from chest x-ray images are considered for proposed deep learning model implementation.

4.1 Case Study 1: A CNN -Dempster Shafer based classification approach for Pulmonary Tuberculosis Detection from Chest X-ray images

Pulmonary Tuberculosis (TB) is a potential noxious disease that is spread by a bacterium and mainly affects the lungs. It is considered as a leading respiratory disease worldwide. Pulmonary Tuberculosis can be life threatening if not detected in right time followed by proper treatment. According to worldwide survey, China, India, and Indonesia have the highest number of cases every year [173]. According to a report published by Ministry of Health and Family welfare, Government of India, in 2022, over 24.22 Lakh Tuberculosis cases were registered suppressing the pre-COVID level report [174]. Hence, early detection and prevention are the only keys to help in decreasing the number of cases. In the modern era of Artificial Intelligence, Computer Aided Diagnosis System (CAD) is rapidly revolving the pattern of disease diagnosis and treatment. With the help of Deep learning algorithms, strong modelling of radiographic images is helpful for analysis of the data. It will thus help in proper classification and subsequent prediction.

A bacterium named Mayobacterium Tuberculosis which spreads though air is the main reason of Pulmonary Tuberculosis. Detection is done by examining microscopic images of sputum for acid fast bacilli i.e., observing pus samples or sputum for confirming the presence of mycobacterium TB. A designated highly equipped laboratory is required to examine the presence of bacteria in this process. It is time-consuming and requires more resource and expertise to analyze. But most of the countries lack the infrastructure to do these acid-fast bacilli test. Sputum Smear Microscopy is another feasible solution for pulmonary tuberculosis detection [175]. It is fast and a straightforward approach and popular technique. As the number of Tb cases are increasing, chest X-Ray is an effective, hassle free and affordable solution these days. Other methods that are used for tuberculosis detection are Tuberculin Skin test, culture test, Interferon Gamma Release Assay (IGRA), GeneXpert etc. [176]. Generally, a sputum smear image dataset contains three types of images- image with single bacilli, images with touching or overlapping bacilli and debris. A touching bacilli image consists of several single bacilli objects that are overlapped with each other. To detect the severity of TB, detection of the bacilli is important. But most of the research works are carried out single bacilli object [177,178]. So, an automated system is required to analyze faster, and make accurate decision making and control the number of cases.

4.1.1 Literature Review

As manual detection is a tedious job, different intelligent diagnosis models are proposed for pulmonary tuberculosis detection. Models based on learning algorithms have shown proper accuracies and many researchers have proposed models for automatic pulmonary tuberculosis detection using chest X-Ray (CXR) images. Different research works have been carried out by researchers for pulmonary tuberculosis detection and prediction. A review has been done for comparative study of Machine learning and Deep Learning based disease classification approaches in several research works [254]. Table 4.1 describes different deep learning model implementations and compares their performance on TB dataset in different studies.

Authors	**Data Description**	**Classification Approach**	**Performance Evaluation**
Lillian Muyama (2021) [179]	Ziehl-Neelsen sputumsmear microscopic images. Total Records – 148	VGGNet, GoogLeNet Inception v3	Accuracy - 79.6%, 76.8%
Xukun Li (2020) [180]	Affiliated Hospital of Zhejiang University Total Records – 501	DenseVoxNet, 3D U-Net, and V-Net networks are used for image segmentation.	Precision - 93.7% , Recall - 98.7%
Ray-I Chang (2020) [181]	Tao-Yuan General Hospital, Taiwan Total Records – 16503 images form 1727 samples.	An Encoder – Classifier model approach has been adopted A fully connected. network (FCN) with its parameter set θf *is* used as a classifier to classify v into one of the n classes	Accuracy and precision – 99%
Mustapha Oloko-Oba (2020) [182]	Shenzhen tuberculosis dataset Total Records – 662	A ConvNet consisting of 6 convolutional. layers	Accuracy - 87.8%
Muhammad Ayaz (2020) [183]	Montgomery and Shenzhen Dataset available at U.S. National Library	An ensemble method that combines hand-crafted features with deep features (convolutional neural network-based).	Accuracy - 93.47%, 90.6%in LR, 97.59% in CNN as level 0 classifier
Hery Yugaswara (2020)[184]	public health center, Jakarta Total records – 81	LR, KNN, NB, SVM, RF, NN LDA were implemented on the preprocessed dataset with parameter fine tuning.	Accuracy - Lies in 97 % to 99%
Muhammad Ali (2019) [185]	Gulab Davi Hospital ,	Deep Residual Network with with ReLU, OxfordNet	Sensitivity - 82.08%, 84.91 % , Specificity - 93.80 %, 93.02%
Manoj Raju, (2019) [186]	Montgomery County, Maryland, Shenzhen Hospital in Shenzhen, A medical college in India	Using WEKA toolkit, SVM, Logistic Regression, Naïve Bayes and C5.0 were implemented.	Accuracy - 81 %-84%
Moumen El-Melegy (2019) [187]	ZNSM-iDB Dataset, Total Records – 500,	faster R-CNN was proposed which combines a CNN	Precision - 82.6%, Recall -98.3% -
Alfonso Hern´andez (2019) [188]	Montgomery County's Tuberculosis screening program Total Records – 800,	Pre-trained models ResNet50, InceptionV3, VGG16, VGG19	Accuracy- 79% to 86%
Sonaal Kant (2018) [189]	ZiehlNeelsen Sputum smear Microscopy image DataBase	five-layer full convoluted Neural Network (CNN)	Precision - 67.55%
Rahib H. Abiyev (2018) [190]	National Institutes of Health—Clinical Center, Total Records – 112,120	BPNN, CpNN, CNN was used as a classifier.	Accuracy - 80.04% to 92.4%

Table 4.1: Literature Review of Tuberculosis Detection Using Deep Learning Models

Author	Dataset Description	Classification Approach	Performance Evaluation
Tawansongsang Karnkawinpong (2018) [191]	National Library of Medicine (NLM) and the Ministry of Public Health, Thailand Total Records - 3310	AlexNet, VGG-16, CapsNet	Accuracy - 86.86% to 90.63%
Seelwan Sathitratanacheewin (2018) [192]	Shenzhen No.3 Hospital in Shenzhen, Guangdong providence, China, NIH Clinical Center, Bethesda, Maryland, USA	InceptionV3 and pretrained DCNN are used to classify TB CXR images.	Accuracy - 98.45 %, 85.02%
Yadini Pérez López (2017) [193]	Not Mentioned Total Records – 492, 9770 patches are extracted from focus smear microscopic images	Three different CNN architectures	Accuracy – 97%
K.C. Santosh (2017) [194]	Montgomery County Hospital collection; Shenzhen Hospital, China collection; and Indian (IN) collection, New Delhi.	Ensemble Voting based classifier with three base classifiers such as Bayesian Network, MLP and RF were used.	Accuracy – 91%
Paras Lakhani (2017) [196]	Montgomery County, Shenzhen, China Belarus TB Public Health Program, Thomas Jefferson Hospital	AlexNet and GoogleNet	Accuracy- First set -78.3 %, Second set – 90 %
Stefan Jaeger (2013) [195]	Department of Health and Human Services of Montgomery County (MC), Maryland, Shenzhen Hospital, China Japanese Society of RadiologicaTechnology(JSRT) Dataset	For the segmented lung field, texture and shape features are computed and feed as input to a binary classifier. Using a decision rule and thresholds	ROC- 99%
Ebenezer Priya (2012) [197]	Not Mentioned Total Records – 100	active contour method and Differential Evolution based Extreme Learning Machines (DE-ELM). Further, significant features are derived using PCA and KPCA (Kernel PCA).	Accuracy - 97.5 %, 95 %
Bram van Ginneken (2002) [198]	political asylum in The Netherlands, University of Chicago Hospitals Maryland, USA, Dataset1 – 616, Dataset2- 200,	KNN is applied in for each training set of each region.	Sensitivity - 86%, 97%

Table 4.1(Contd): Literature Review of Tuberculosis Detection Using Deep Learning Models

Various pulmonary tuberculosis detection and classification approaches have been discussed by different researchers which starts from collecting data, data preprocessing, feature engineering, classifier implementation and finally measuring accuracy of their proposed work. Common problem faced by researchers during their research was unavailability of labeled data. Labeling of collected data requires time and expert supervision. So, this is a major concern while building a model for pulmonary tuberculosis detection. The next section discusses data collection process for this case study.

4.1.2 Data Collection and Data Preprocessing

Use of chest x-ray images is very popular for tuberculosis detection. X-ray tests are cost effective and widely available in any demographic area. It is easily available for analysis. Pulmonary Tuberculosis is considered as an infection of tuberculosis in lungs. So, for this study chest x-ray images are considered. For implementation purposes, tuberculosis chest x-ray datasets available online at Kaggle is considered. The dataset consists of a total of 3500 normal chest x-rays and 700 tuberculosis images [199]. The dataset is curated by a team of researchers from Qatar University, Doha, Qatar, and University of Dhaka, Bangladesh, in association with Hannad Medical Corporation's doctor. For implementation purpose, the image dataset is divided into 80:20 ratio for training and testing purposes, respectively.

For better performance of any deep learning model, input images should be noise free and should be unambiguous. While capturing images from an X-ray machine and collecting images in dataset format, images may blur due to the technique of data collection. So, contrast enhancement is a way to improve image quality for further processing. A chest x-ray image contains rib structure as foreground and back muscles as background of the image. But in some cases, the outline of the rib bones is not visible enough to distinguish background and foreground image. Different contrast enhancement algorithms have evolved since last decade. Main goal of any image enhancement technique is to first improve appearance of images for better visual interpretation, and secondly increasing performance of the designated task. As most contrast enhancement technique relies on histogram-based modification, it can be applied locally or globally.

Contrast Limited Adaptive Histogram Equalization Algorithm (CLAHE) is an enhancement technique that obtains global optimal contrast by enhancing local contrast [200,201]. CLAHE first divides the images into small equal size areas called tiles. Bilinear interpolation is used to remove false boundaries between surrounding tiles. Algorithm for CLAHE [202] is described below:

CLAHE Algorithm [271]

Step 1: Input chest x-ray Image.

Step 2: Segment input into same shape tiles.

Step 3: Compute histogram for each tile.

Step 4: Apply Triangular Fuzzy Membership (TFM) to compute Clip Limit.

Step 5: Limit the contrast based on computed clip limit.

Step 6: Output enhanced image.

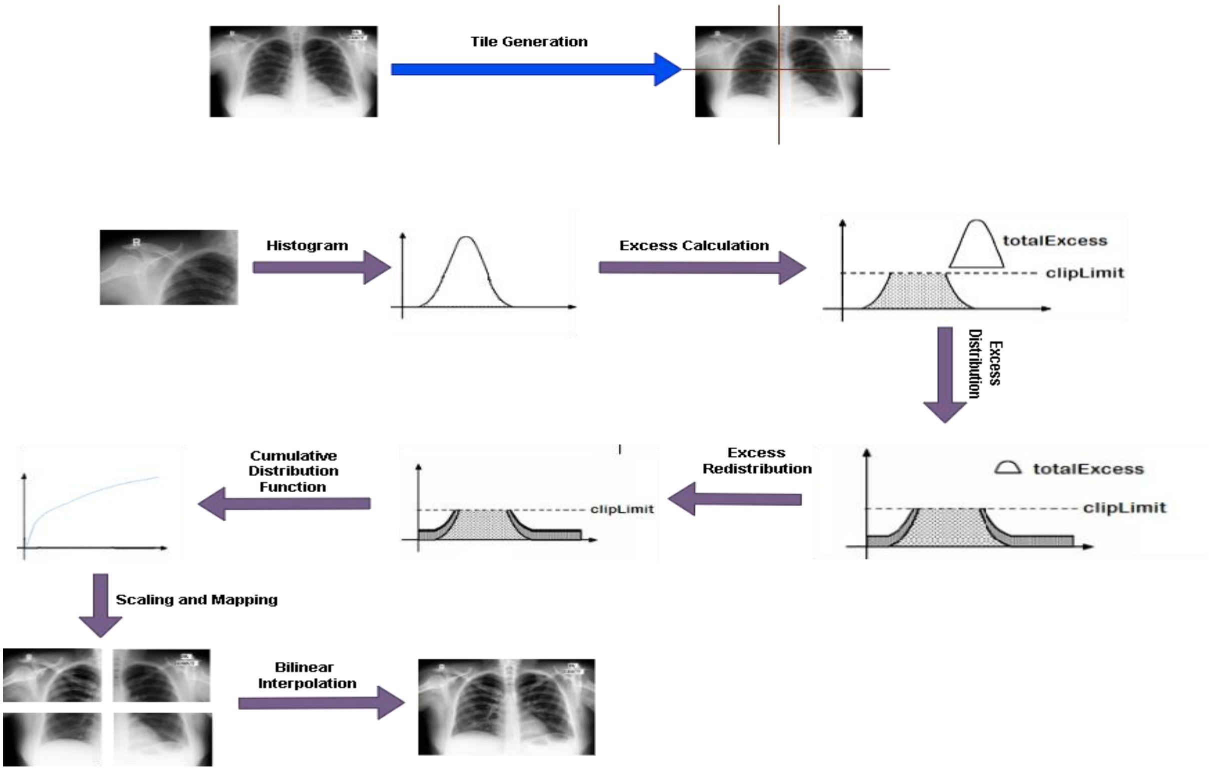


Figure 4.2: Process of CLAHE image enhancement technique [270]

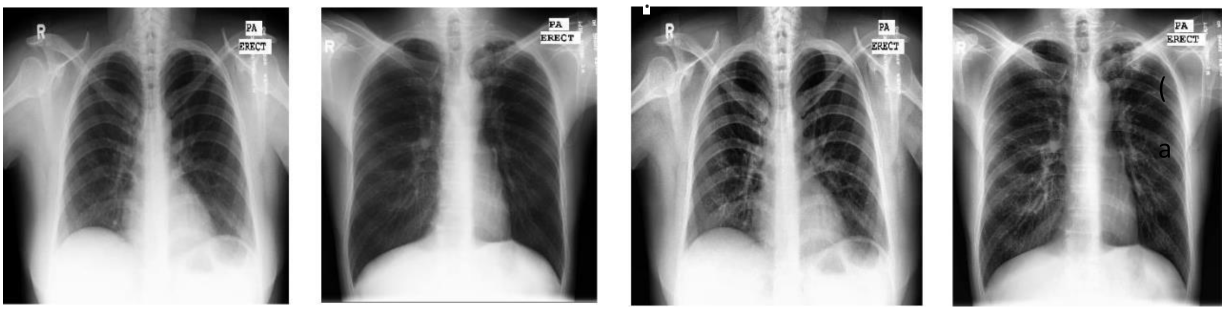


Figure 4.3(a) **Figure 4.3(b)** **Figure 4.3(c)** **Figure 4.3(d)**

Figure 4.3: Application of CLAHE algorithm on Normal and Tuberculosis infected chest x-ray images

From the above figures, 4.3(a) and 4.3(b) shows normal and TB infected lung x-rays. Figure 4.3(c) and 4.3(d) shows images after CLAHE enhancement technique.

4.1.3 Proposed CNN-Dempster Shafer based classification approach

4.1.3.1 Proposed Concatenated - CNN approach

For image data analysis, Convolutional Neural Network (CNN) is till now considered to be best as it effectively extracts important information for patten recognition, object detection, image classification and so on [203]. The proposed approach for tuberculosis detection takes input images of size 224 x 224 pixels in RGB format. Three CNN models with varying convolutional layers are implemented a feature extractor. For each CNN feature extractor, input image is first fed as input to convolution layer. Main function of a convolutional layer is to detect edges in an image using filters. Filters help to discard unwanted information. To understand feature representation of deep CNNs' used as feature extractor, the operation of CNN is discussed below.

Consider an input image $z = (z_1, \ldots, z_D)$ consisting of D input maps or channels z_i $(i = 1, \ldots, D)$ with size H ×W. A convolutional layer consists of several convolution filters or kernels that are used to extract feature maps from z. As an image consists of pixels, the influence of nearby pixels is calculated using some filter. A convolution filter or kernel is represented as a small matrix used to operate a convolution operation to each input map. The filter or kernel map is sided over the map and an element-wise multiplication is performed with a part of the input map while keeping the kernel on. Multiplication results are summed up into a single value, and bias is added to the summed value [204]. So, the processes of a convolutional layer, with e convolution filters or kernels with size a × b, are represented as

$$\mathbf{c^{j} = f(b^{j} + \Sigma_{i}\, w^{i,j} * z^{i}),} \tag{4.1}$$

where b^j is the bias of the kernel of w^{ij}, $w^{i,j}$ represents the convolution filter between the i-th input map and the j-th output map, z^i is the ith input map of size H x W, $*$ denotes the convolution operation; $i = 1, \ldots, D$, c^j is the j-th output feature map of size $\frac{H-a+1}{r} X \frac{W-b+1}{r}$, $j = 1, \ldots, e$, r represents the stride value with which the filter or kernel strides over the input map z^i. f represents here an activation function. ReLU activation function is used here It is represented as:

$$\mathrm{ReLU}(x) = \max(0, x) \tag{4.2}$$

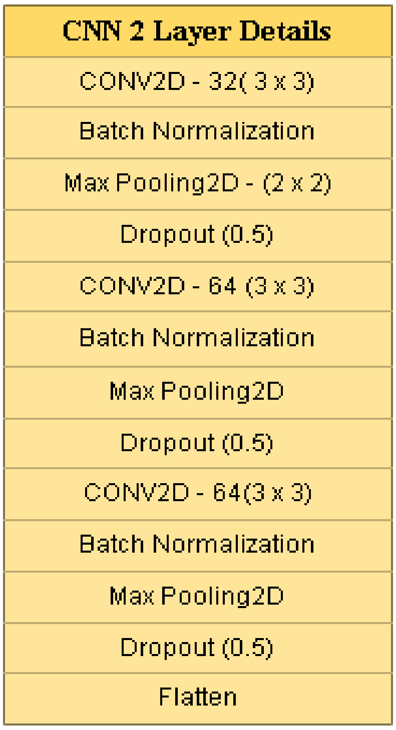
Normalization is a data processing tool used for scaling numeric data without distorting its shape. Batch normalization is a process to make neural network faster and more stable by performing standardizing and normalizing operation on input from previous stage. Normalized data should have mean zero and standard deviation 1. After batch normalization operation, feature map is generated which is further used for pooling operation. The pooling operation is done by sliding a two-dimensional filter over each channel of feature map. Further summarization of the features lying within the region of filter is done [205]. A feature map having dimensions of n_h x n_w x n_c, the output after pooling layer operation is described below:

$$\mathbf{P} = \frac{n_h - f + 1}{s} \times \frac{n_w - f + 1}{s} \times n_c \tag{4.3}$$

Where $\mathbf{n_h}$ - height of feature map, $\mathbf{n_w}$ - width of feature map, $\mathbf{n_c}$ - number of channels in the feature map, **f -** size of filter, **s -** stride length. After pooling layer, the output is further processed for overfitting. Dropout is a regularization technique where certain nodes are disregarded in a layer at random while training. It eliminates overfitting by ensuring no units are independent of one another. After pooling layer operation, the output vector is flattened using flatten layer. It converts multi-dimensional feature vector to a one-dimensional array. Layer details of CNN as feature extractor is defined below in table 4.2(a), 4.2(b), and 4.2(c).

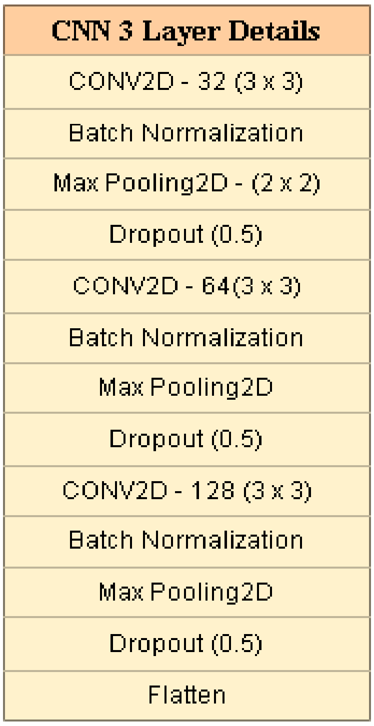
CNN 1 Layer Details
CONV2D - 16(3 x 3)
Batch Normalization
Max Pooling2D - (2 x 2)
Dropout (0.5)
CONV2D - 32 (3 x 3)
Batch Normalization
Max Pooling2D
Dropout (0.5)
CONV2D - 64(3 x 3)
Batch Normalization
Max Pooling2D
Dropout (0.5)
Flatten

Table 4.2(a)



CNN 2 Layer Details
CONV2D - 32(3 x 3)
Batch Normalization
Max Pooling2D - (2 x 2)
Dropout (0.5)
CONV2D - 64 (3 x 3)
Batch Normalization
Max Pooling2D
Dropout (0.5)
CONV2D - 64(3 x 3)
Batch Normalization
Max Pooling2D
Dropout (0.5)
Flatten

Table 4.2(b)



CNN 3 Layer Details
CONV2D - 32 (3 x 3)
Batch Normalization
Max Pooling2D - (2 x 2)
Dropout (0.5)
CONV2D - 64(3 x 3)
Batch Normalization
Max Pooling2D
Dropout (0.5)
CONV2D - 128 (3 x 3)
Batch Normalization
Max Pooling2D
Dropout (0.5)
Flatten

Table 4.2(c)

Table 4.2: Details of three CNNs' as feature extractor

4.1.3.2Dempster Shafer Theory based Classifier

Three CNN models are implemented as feature extractors to obtain important features from input image of size 224 x 224 and three feature vectors are obtained. After CNN operation, the feature vectors of each CNN are concatenated along the channel dimension. The output from the concatenated layer is fed as input to the Dempster Shafer based Evidence layer. The D-S theory is a versatile evidence theory that combines different types of evidence obtained from different sources. It is considered as an alternative way to probabilistic models, where uncertainty is represented in mathematical format. The notable step of Dempster Shafer based classifier is that it allocates a probability mass to intervals or sets. The features vector from the concatenated layer is fed to D-S layer for the construction of mass functions. The steps in Dempster Shafer classification approach are explained below:

- **Step 1: Set Frame of Discernment**

The frame of discernment for Tuberculosis dataset ɵ is defined as {Normal, Tuberculosis}. The power set of ɵ, E can be set as: {Normal} = class 0, {Tuberculosis} = class 1, {Normal, Tuberculosis}.

- **Calculate class membership value**

The simplest way to calculate class membership value is to determine each feature's maximum and minimum value for each class. So, for each class, each feature's maximum and minimum values are calculated and a range of [min, max] is set.

- **Allocate mass function**

A mass function can be defined as a belief representation. It can be used to derive all other representations. A mass function m can be defined as m: P(E) →[0,1], which assigns a mass value to each hypothesis A belonging to E of frame of discernment. Thus, m(A), is defined as the amount of belief strictly assigned to hypothesis A.

In D–S theory, some important functions are basic probability-assignment function, Belief function (BEL),and the likelihood function (PLS).These three functions are described below:

Basic probability Assignment function: $\sum A\in\Theta\ m(A)=1, m(\phi)=0$ (4.4)

Belief Function: $2^{\Theta}\rightarrow[0,1], \mathrm{Bel}(A)=\sum_{B\in A} m(B)$ (4.5)

In Equation (4.4), A is considered as a hypothesis in the classification framework where $A\subseteq\Theta$ and

m(A) is defined as basic probability-assignment function. In Equation (4.5), Bel(A) is considered

as the sum of all basic distribution probability functions of all subsets of A

The plausibility pl(A) is defined as the amount of probability not strictly committed to A. Therefore, Belief is considered as a lower bound of probability and plausibility is seen as an upper bound of probability.

- **Apply Dempster Combination Rule**

For all $A\subseteq\Theta$ and considering n no. of masses $m_1, m_2, \cdots, m_n$, Dempster's rule is calculated by the following equations:

$$m_1\oplus m_2\oplus\cdots\oplus m_n(A)=1/K\sum_{A1\cap A2\cap\cdots\cap An} m_1(A_1)\cdot m_2(A_2)\cdot\ldots\cdot m_n(A_n) \quad (4.6)$$

$$K=1-\sum_{A1\cap A2\cap\cdots\cap An}=\phi m_1(A_1)\cdot m_2(A_2)\cdot\ldots\cdot m_n(A_n) \quad (4.7)$$

Where K is defined as the amount of conflict between two masses. Thus, using Dempster Rule of Combination (DRC), the mass values for all features are combined using DRC strategy, which results in overall mass for all considered hypotheses. The hypothesis that contains the highest belief value is used to classify the image. If the hypothesis does not contain a single class, further steps are implemented.

- **Classify Data**

For each attribute calculate the standard deviation for each class and calculate the feature selection value as:

$$\text{Feature selection Value (FSV)} = \frac{sd(class1)x\ sd(class2)x\ sd(class3)\ldots\ldots\ldots sd(classn)}{sd(class1\cup class2\ \cup class3\cup \ldots\ldots classn)} \tag{4.8}$$

Where n is number of classes. The feature f that has the smallest FSV value is selected, and the absolute difference d between the input image’s feature f value and mean value of feature f is calculated for each class:

$$d_i = |\ f_i - mean(f_i)|,\ i = 1,\ldots,\ n \tag{4.9}$$

The test image is classified to the class which has the smallest d value.

Figure 4.4 and 4.5 represent the Dempster Shafer classification flowchart and proposed concatenation CNN -Dempster Shafer based classification approach respectively.

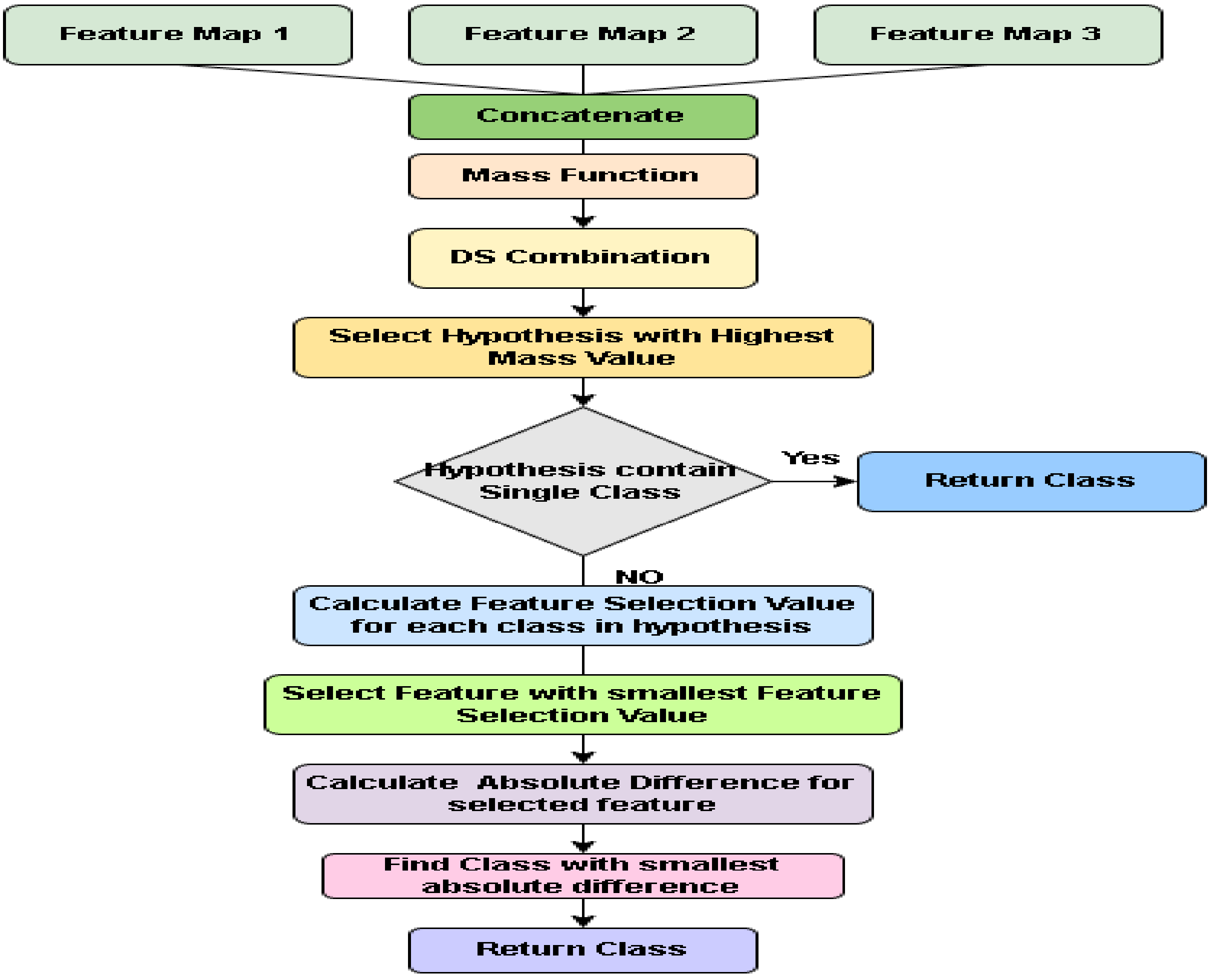


Figure 4.4: Dempster Shafer Classification flowchart

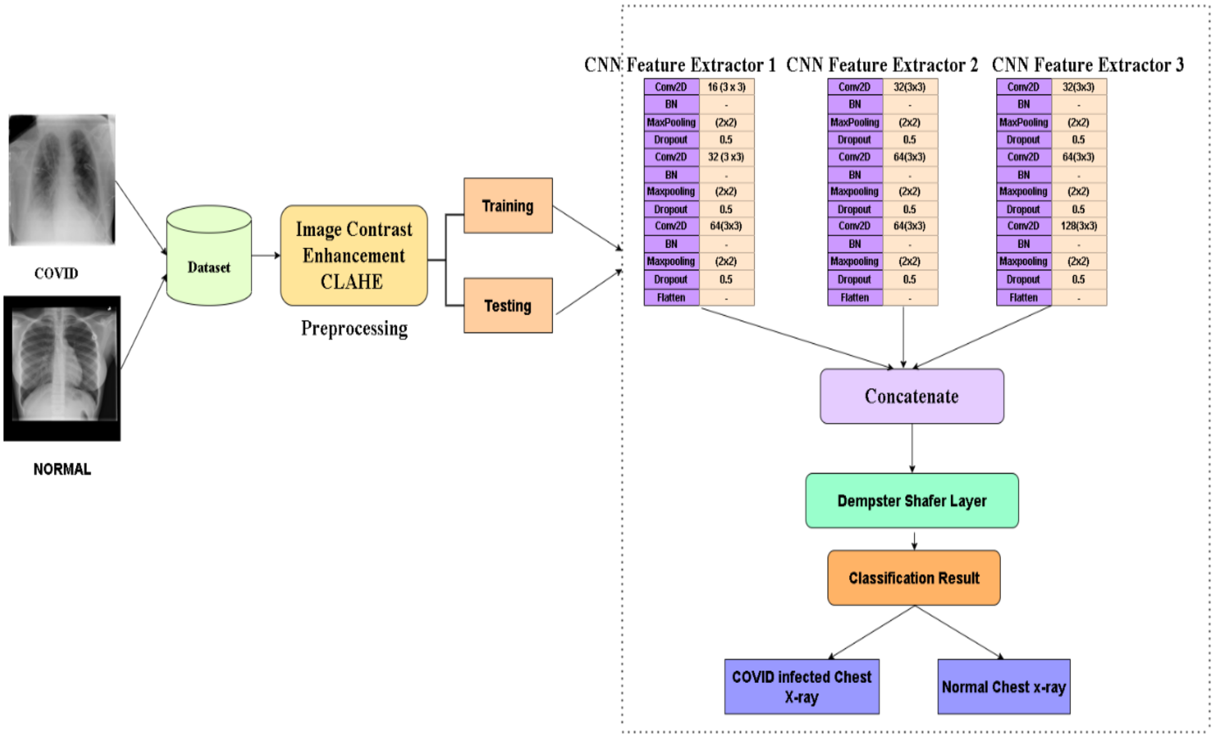


Figure 4.5: Proposed concat-CNN Dempster Shafer based Classification approach.

4.1.4 Implementation details and result analysis

In DL, optimizers are algorithms that are used for model's parameter adjustment during training to minimize the loss function. Optimizers help model to learn from data by iteratively updating bias and weights. For training the proposed approach, Adaptive Moment Estimation (Adam) optimizer with learning rate at 0.001 is used. Adam optimizer is an extension of stochastic gradient decent algorithm. This algorithm has low memory requirement, faster running time, and requires less tuning. Formula for Adam optimizer is described below [39]:

$$mt = \beta_1 m_{t-1} + (1 - \beta_1)\left[\frac{\delta L}{\delta wt}\right] v_t = \beta 2 v_{t-1} + (1-\beta_2)\left[\frac{\delta L}{\delta wt}\right]^2 \qquad (4.10)$$

To evaluate the effectiveness of the proposed approach, the proposed approach is trained for five-fold cross validation with splitting of images in different sets of images. In each fold, a separate set of images is used for training and testing purpose while keeping the validation set as fixed. Proposed model is analyzed based on performance metrics like Accuracy, Precision, Recall, AUC. Classification accuracy defines the percentage of correctly predicted TB images divided by total number of predicted images, whereas Recall also known as True Positive Rate signifies the ratio between the number of samples correctly classified as positive and to total number of positive images in that class. Precision refers to the ratio of total number of true positive images to total number of positive predictions.

$$\text{Classification Accuracy} = \frac{TP+TN}{TN+TP+FP+FN} \qquad (4.11)$$

$$\text{Recall} = \frac{TP}{TP+FN} \tag{4.12}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{4.13}$$

The proposed model was implemented using Google Colab GPU (Tesla K80), Python 3.7, TensorFlow 2.7.0 and Keras 2.7.0 libraries.on Intel® Xeon® CPU E5-2697 v4 @ 2,30GHzand 64 GB RAM, with a 16 GB NVIDIA GeForce GTX 1080 GPU.

Below table 4.3 shows accuracy, precision, recall, and AUC obtained for each fold. The proposed approach is trained for 30 epochs for each fold. An epoch describes the training of a model with all the training data for one cycle.

Fold	**Accuracy**	**Precision**	**Recall**	**AUC**
1	94.64 %	94.64%	94.64%	97.66%
2	91.78%	91.78%	91.78%	98.31%
3	96.78%	96.78%	96.78%	99.1%
4	94.64%	94.64%	94.64%	99.29%
5	93.21%	93.21%	93.21%	97.48%
Average	94.21%	94.21%	94.21%	98.29%

Table 4.3: Performance evaluation for 5 folds of Proposed CNN-Dempster Shafer approach

Average Accuracy obtained in the proposed approach is 94.21%, average Precision is 94.21%, average Recall is 94.21% and Average AUC is 98.29%. This shows the efficacy and robustness of the proposed approach. Table 4.3 shows the Accuracy, Precision, Recall, and AUC obtained for all 5 folds. The proposed approach's efficacy has been compared with state-of-the-art transfer learning models and other models proposed in different studies. Different deep CNN Transfer Learning models such as InceptionV3, ResNet50, DenseNet2011models have been compared with their pretrained weight and bias and trained in the considered dataset configuration. These models are also validated using 5 cross validation and trained for 30 epochs for each fold. A comparative study was done in table 4.4 based on Accuracy, Precision, Recall, and AUC.

Fold Number	State of the Art Model	Accuracy	Precision	Recall	AUC
Fold 1	InceptionV3	63.57	63.57	63.57	73.73
	ResNet50	72.14	72.14	72.14	76.73
	DenseNet201	71.07	71.07	71.07	84.01
	Proposed	**94.64**	**94.64**	**94.64**	**97.66**
Fold 2	InceptionV3	89.28	89.16	89.29	97.08
	ResNet50	76.42	76.42	76.42	82.92
	DenseNet201	89.99	89.99	89.99	95.46
	Proposed	**91.78**	**91.78**	**91.78**	**98.31**
Fold 3	InceptionV3	99.64	99.64	99.64	99.99
	ResNet50	93.92	93.92	93.92	98.45
	DenseNet201	95.71	95.71	95.71	98.60
	Proposed	**96.78**	**96.78**	**96.78**	**99.1**
Fold 4	InceptionV3	98.92	98.92	98.92	99.98
	ResNet50	91.42	91.42	91.42	94.28
	DenseNet201	92.50	92.50	92.50	98.18
	Proposed	**94.64**	**94.64**	**94.64**	**99.29**
Fold 5	InceptionV3	97.85	97.85	97.85	99.56
	ResNet50	99.28	99.28	99.28	99.97
	DenseNet201	82.14	82.14	82.14	83.50
	Proposed	**93.21**	**93.21**	**93.21**	**97.48**
Average	InceptionV3	89.85	89.85	89.85	94.07
	ResNet50	86.64	86.64	86.64	90.47
	DenseNet201	86.28	86.28	86.28	91.97
	Proposed	**94.21**	**94.21**	**94.21**	**98.29**

Table 4.4: Comparative Study of Performance of Different Transfer Learning Models for TB Classification

As mentioned earlier different experiments were conducted for classification of Tuberculosis infected chest x-ray images and Normal Chest X-ray images. The comparative performance of different transfer learning models for binary classification is shown in Table 4.4. It is evident from table no 4.4 that all the state of art transfer learning models performed well while classifying Tuberculosis and Normal images. It is evident that for every fold the proposed concatenated CNN-Dempster Shafer classification approach exhibits better performance in classifying the chest X-ray images. The training per epoch for the proposed model is 5s whereas training time per epoch for InceptionV3, ResNet50, and DenseNet201 are 12s, 14s, and 9s. Figure 4.6 shows graphs of the training and validation

Accuracy, Precision, Recall and AUC versus epochs for the proposed concatenated CNN-Dempster Shafer classifier. It is also evident that the network stabilizes with the lowest loss while dealing with contrast enhanced images.

In summary, concatenated CNN-Dempster Shafer classification approach produced highest classification accuracies of respectively. It is clearly seen that the model performs well for any computer aided classifier which can significantly help in fast analysis of Tuberculosis disease by healthcare professional after obtaining the chest x-ray images.

This state-of-the-art performance of the proposed approach was evaluated with the recently published research works in the Tuberculosis detection from Chest X-ray. Table 4.5 summarizes the implementation details along with performance of studies carried out in different research works. In study [209], the detection accuracy was reported to be 98.6%. However, a large dataset is considered in their application. The proposed approach has shown consistency about accuracy, precision, recall result through 5-fold cross validation. The accuracy of the proposed approach has been enhanced through a histogram algorithm CLAHE that provide proper visualization of images while classification. The proposed model obtained accuracy around 97.8% which proves its versatility and robustness.

As it is well known, chest x-ray images are useful and meaningful input for diagnosing various types of diseases that particularly affect chest, lung. During the Covid period, chest x-ray images serve the purpose of providing information about spread of the virus. The images helped in assessing the severity of the spread. Therefore, in the next section the handling of chest x-ray images for COVID detection is detailed out.

Author Name	Published Year	Dataset Details	Classification Method used	Number of Images	Performance Matrix
NIkarika et al. [249]	2019	MC and CHN	SVM	805	Specificity 100% and AUC 96%
Pasa et al. [250]	2019	Mc, CHN, Belarus Dataset	Optimized CNN1	1111	AUC – 81% for MC, 80% for CHN and 92.5% for combined
Ojasvi et al. [251]	2018	ResNet	MC and CHN	805	Accuracy – 94.89%
Sayeda et al. [252]		Transfer Learning	MC and CHN	805	AUC – 85%
Quang et al. [272]	2019	DenseNet model with parameter tuning	MC and CHN	805	AUC – 94% for MC and 82% for CHN
Alfonso et al.[188]	2019	Transfer Learning Models	MC and CHN	805	Accuracy- 86%
Rahman et al. [209]	2020	Pretrained CNN models	NLM,Belarus, NIAID TB portaland RSNA	3500	Accuracy 98.6%
Muhammad Ayaz (2020) [183]	2021	CNN-Ensemble	Montogomary and Shenzhen		93.47%, 90.6%in LR, 97.59% in CNN
Proposed Method [253]	2022	Ensemble CNN-Dempster Shafer Classifier	NLM,Belarus, NIAID TB portaland RSNA	3500	Accuracy – 94.21%

Table 4.5: Comparative study with state of art approaches from different research works.

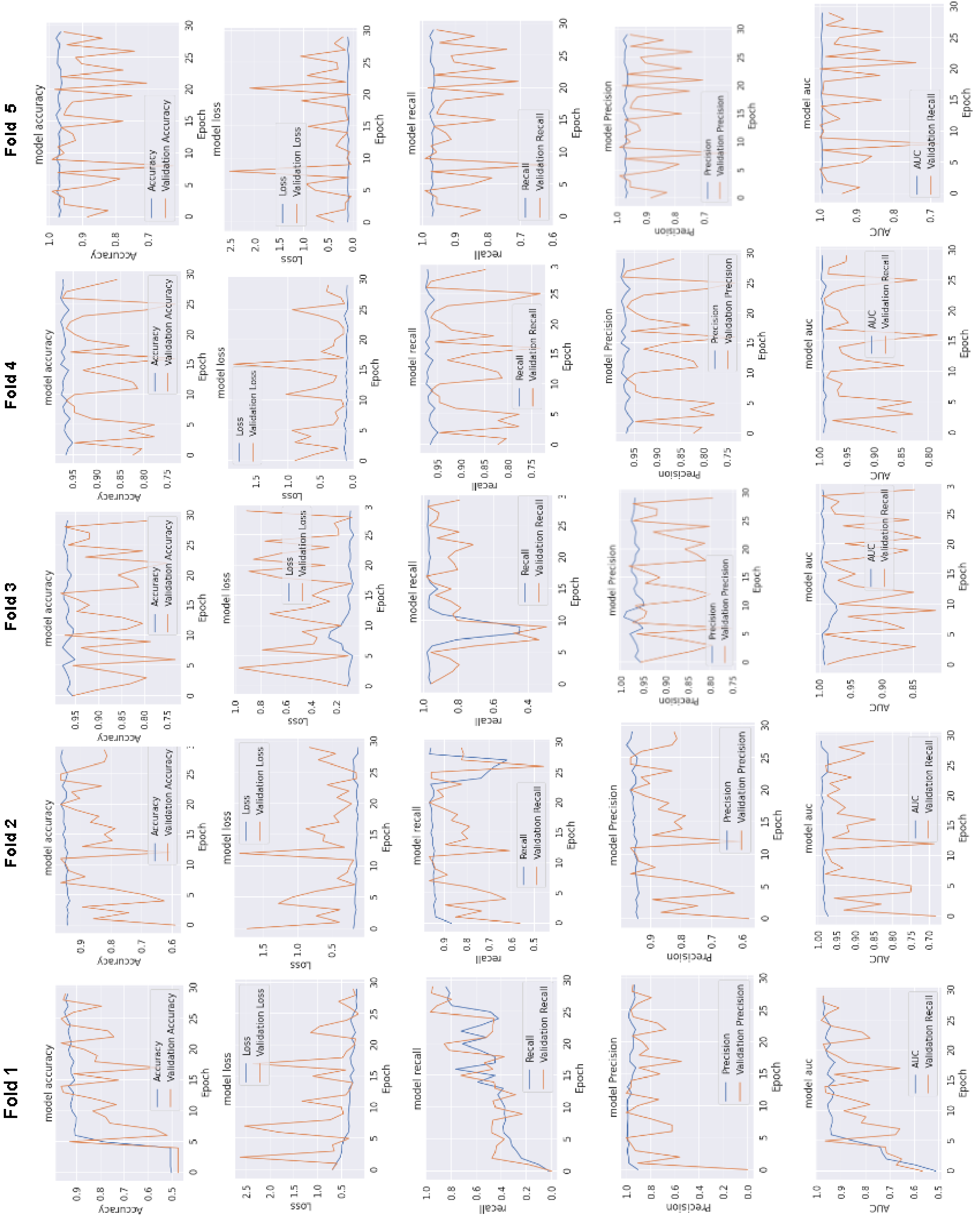


Figure 4.6: Accuracy, Loss, Precision, Recall, and AUC graph for 5-fold cross validation.

4.2 Case study 2: A concatenated- CNN approach for COVID19 detection from Chest X-ray images

Coronaviruses are a large group virus which cause respiratory distress. COVID19 or Coronavirus 2 is a new strain of this virus group that was previously not found in human body. This new strain of virus is named as ‘COVID-19’virus as the first case of this virus was found in 2019. According to WHO’s confirmed report, first case of novel coronavirus (n-COV) was detected in Wuhan City, the largest metropolitan area of Hubei province in China in December 2019 [210]. Covid19 virus is a single strand RNA that causes moderate to severe respiratory infection. In case of mild to moderate infection, infected person will recover within 14 days. It is a variant of SARS-COV2 which are contagious and can spread rapidly among proximity of infected person. around 45 million people have been infected and around 0.5 million people died due to COVID-19 as of December,2023 in India [211]. Since the inception of the disease, be it developed countries like USA, China, Italy, England etc. or developing countries or underdeveloped countries, all suffered due to huge case load and shortage of healthcare equipment, facilities, and lack of effective and efficient approach for proper diagnosis. SARS-COV2 spreads rapidly through droplet from mouth, nose from an infected person to an uninfected person. Four variants of this virus – alpha, beta, gamma, delta are marked as variants of concern, whereas Eta, Iota, Kappa, Lambda falls under variant of interest [212].

Clinical Detection of Covid19 largely depends on a genetic test known as Reverse Transcription Polymerase Chain Reaction (RT-PCR). Nasopharyngeal swabs are taken from infected patient for diagnosis purpose and RNA of the virus is extracted [213]. It takes around 2 hours for detection with long turnaround time. So, it is costly, time consuming and complicated. If the swabs are not collected properly its accuracy rate becomes exceptionally low and might show negative result though the person is positive [214]. Rapid antigen testing is an easy-to-use process with less infrastructure requirement and equipment. It checks the presence of one or more antibodies in the infected person’s body as a response against the virus. Though it is cost friendly, easy to perform, quick response, and able to detect active SARS-COV2 virus, its sensitivity and specificity is much lower than regular RT-PCR test [215]. The best solution is use of radiographic images like chest X-rays and chest CT-Scans. The advantage of using radiographic images is its wide availability all over the world. It is cost effective, and diagnosis is done at the time of testing. The main concern while using chest X-ray images is availability of trained doctors or healthcare professional. Research on infected chest x-ray images has found that some unknown features like vague darkening spots or ground glass opacity helps in COVID19 detection [216]. However, analyzing every chest x-ray individually at the time of pandemic will delay the detection process thereby delay the treatment procedure. Moreover, in rural places, availability of efficient radiologist is a problem for timely intervention and detection. So, an automatic COVID detection system or rather a Computer Aided COVID detection system is an important requirement at present scenario.

4.2.1 Literature review

While the pandemic was still creating havoc on mankind, different researchers come with different ideas and proposed different systems for early detection and prevention of the disease. Among them applications of Deep Learning algorithms have shown promising results. In healthcare sector, deep learning [217] has shown tremendous development in automatic disease prediction such as tuberculosis detection [218,219], cancer detection [220,221], tumor detection [222,223], bone fracture detection [224,225], genome sequence analysis [226,227,228] etc.

In their study, Abbas et al. [229], proposed a model named DeTraC that classifies chest x-ray images. The model is based on deep CNN models. Class decomposition layer was used to simplify the dataset structure. ResNet pretrained model was used to pretrain the model and then class-composition layer was used for parameter fine tuning. The dataset size was 185 images where 105 images are virus infected and 80 are normal chest x-ray images. The authors have achieved 95.12% accuracy, 97.91% sensitivity and 91.87% specificity. Das et al [230], had proposed an ensemble of multiple state of the art CNN models like InceptionV3, ResNet50V2, and DenseNet201. All these models are trained individually and then training results are combined using weighted average ensemble technique. The dataset containing 538 COVID+ images and 468 COVID- images is divided into train, test and validation set. The accuracy of the proposed model is 91.62%. Wang et al. [231] proposed a custom deep learning architecture named COVIDNet that classifies images into three categories: Covid infected patients, healthy patients, and pneumonia patients. Their model is designed using a lightweight projection-expansion-projection-extension (PEPX) pattern which obtained a classification accuracy of 94%. Apostolopoulos et al. [232] proposed COVID19 identification system using five pretrained deep learning models. Transfer learning models VGG19, Inception, Xception, MobileNet, and Inception_ResNet_V2 are used for classification. They achieved accuracy around 96.78%. Their binary classification results are better than multiclass classification. A hybrid model combining Xception and ResNet50_V2 model was proposed by Rahimzadeh et al. [233]. They used a dataset of 6054 x-ray images and achieved an accuracy of 91.4 %. Fu et al. [234] explored two characteristics of SARS-COV2-dynamic viral shedding and serological characteristics. They identified the risk factors associated with poor recovery and lung lesion resolutions. They studied the relationship between pro-inflammatory response, viral shedding, and lung lesion evolution. According to their study, 58 % of the infected patients will have poor recovery and prolonged interval of viral shedding. In another study, Cheng et al [235], compared the performances of different ensemble CNN models like majority voting, unweighted average, and Bayes Optimal Classifier. Afsar et al [236], have proposed a model named COVID-CAAPS which is based on capsule network to detect covid infection. They combined different convolution and capsule layers to address the problem of class-imbalance. With smaller number of trainable parameters, they obtained an accuracy of 95.7%, sensitivity of 90% and specificity of 95.8%.

4.2.2 Data Description and Preprocessing

In this section details of the data source, dataset description, image preprocessing techniques are discussed.

4.2.2.1 Data Collection

For performing the proposed research work, chest x-ray images from different online open sources were collected. These dataset sources contain chest x-ray images having different infection manifestations including COVID 19 viral infection. These images are collected from all over the world and contain different chest x-ray images like viral pneumonia, bacterial pneumonia, normal chest x-ray images. Chest x-ray images of viral infections like COVID19, SARS, MERS-CoV, varicella, Influenza, Herpes and bacterial infections like streptococcus, staphylococcus, Klebsiella, Escherichia coli, Mycoplasma are included in these datasets with detailed medical information. Different views of these images, like: Anterior-Posterior view (AP), Lateral View(L), and Posterior Anterior (PA) view are available. As the current focus is to view the condition of the lungs for any infection, Posterior-Anterior (PA) views suit best for this purpose. Hence images having PA views only are considered and others are discarded. Even though there are many viral and bacterial categories of images, for the experiment purpose, three classes– COVID19 positive, Normal chest x-ray, and viral pneumonia are considered. Normal chest X ray (CXR) images are considered as class 0, COVID19 positive as class 1, and viral pneumonia as class 2. To maintain class balancing, the dataset is created with equal number of images in each class, i.e., 500 covid Chest x-ray images, 500 normal chest x-ray images, and 500 chest x-ray images of viral pneumonia. The images are obtained from the following open sources:

1. 150 images from GitHub which has images from all three classes i.e., 50 images from each class (https://github.com/ieee8 023/covid -chest xray-dataset)
2. 55 covid images from GitHub(https://github.com/Figure-1-COVID-chestxray-dataset)
3. 250 covid images, 300 normal images and 270 Pneumonia images(https://www.kaggle.com/tawsifurrahman/covid19-radiography-database)
4. 95 covid images and 75 normal images(https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images)
5. 75 normal images and 230 pneumonia images (https://www.kaggle.com/paulimothymooney/chestxray-pneumonia)

Table 4.6 sums up the per class image inatances.

Types of Images	No. of images
Healthy or Normal (Class 0)	**500**
COVID-19 (Class 1)	**500**
Non COVID19 viral pneumonia (Class 2)	**500**

Table 4.6: Dataset Description of COVID

4.2.2.2 Data Preprocessing

As the images are collected from different sources, images are of different shapes and their brightness and contrast are different. So, the images are first resized to 224x224 shape. To improve the contrast of all images, the image processing technique Contrast Limited Adaptive Histogram Equalization technique (CLAHE) [248] is implemented at first**.** Here the entire image is divided into fixed number of tiles, contrast of each tile is enhanced, and then bilinear interpolation is used to remove the artificial boundaries within the images. CLAHE is used to enhance the lungs region of each image for better understanding. The images are next normalized by a factor of 1/255. The images are then shuffled for better distribution and the whole dataset is divided into train, validation, and test set in a ratio of 70:10:20 respectively. Below 4.7 projects the training, testing and validation set structure.

Dataset	Healthy (Class 0)	COVID +ve (Class 1)	Non-COVID viral pneumonia (Class 2)	Total
Training	360	360	360	1080
Testing	40	40	40	120
Validating	100	100	100	300

Table 4.7: Image Distribution for training, Validation, and testing

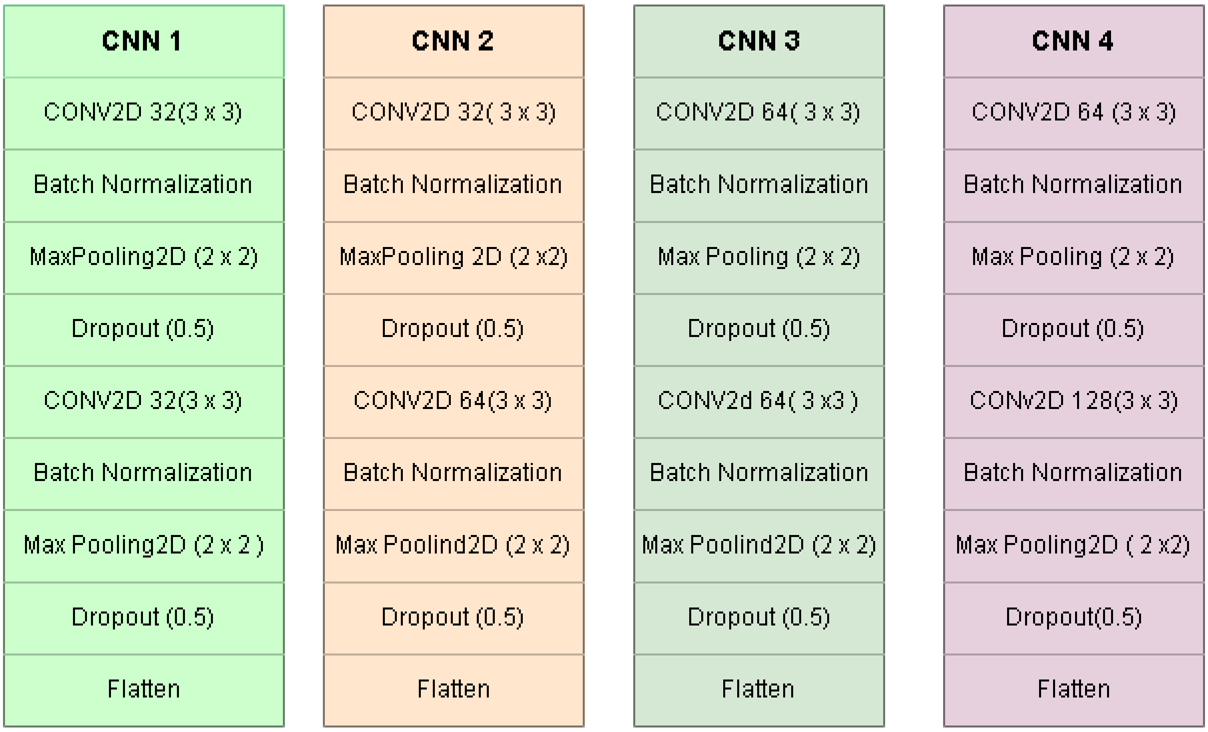
4.2.3 Proposed Concatenated CNN approach

The present work considers a concatenation approach of Convolutional Neural Network for COVID19 detection using chest x-rays. Concatenation of CNNs' helps to combine the features extracted by CNNs' to enhance the classification process more accurately. Four different CNNs with different convolutional layer architecture are implemented as feature extractors from the training dataset. This section discusses the proposed concatenated CNN architecture, necessary optimization process, and parameter fine tuning stages.

4.2.3.1CNNs as Feature Extractor

In this work four CNN models are used as feature extractors. These CNNs have been used to learn good representations of the images like edges, shapes, rotation, spatial representation. These models extract relevant features from the chest x-ray images to obtain a proper knowledge about normal, COVID infected and pneumonia characteristics. The preprocessed training images of size 224 x 224 pixel are provided as input to all four CNNs. Feature maps are generated from the last feature extraction layer of all extractors. As all CNNs extract same size of feature maps, a concatenation function is used to generate an enhanced quality of semantic features. A concatenated convolutional neural network (Concat_CNN) is proposed by using all extracted feature maps. Concatenation and connecting this concatenation layer to a fully connected layer with 32 neurons in association with ReLU Activation function which is further connected to a classifier layer completes the proposed approach [255]. The proposed model is implemented as a multi-class classification problem as data is categorized in three classes as COVID19, Normal, and Pneumonia. Each feature extractor CNN is composed of two convolutional layers with varying filter size of 3 x 3. Each convolution layer is followed by a batch normalization layer, maxpooling layer with pool size (2 x 2) followed by a dropout rate at 0.5. ReLU activation function as showed in eq. 4.2 is used in each convolution layer. Rectified activation function is one of the most popular activation functions which directly maps the inputs to the outputs if it is positive, else it maps zero. It performs better than vanishing gradient and works faster.

Layer details of each CNN is represented in the Table 4.8.



CNN 1	CNN 2	CNN 3	CNN 4
CONV2D 32(3 x 3)	CONV2D 32(3 x 3)	CONV2D 64(3 x 3)	CONV2D 64 (3 x 3)
Batch Normalization	Batch Normalization	Batch Normalization	Batch Normalization
MaxPooling2D (2 x 2)	MaxPooling 2D (2 x2)	Max Pooling (2 x 2)	Max Pooling (2 x 2)
Dropout (0.5)	Dropout (0.5)	Dropout (0.5)	Dropout (0.5)
CONV2D 32(3 x 3)	CONV2D 64(3 x 3)	CONV2d 64(3 x3)	CONv2D 128(3 x 3)
Batch Normalization	Batch Normalization	Batch Normalization	Batch Normalization
Max Pooling2D (2 x 2)	Max Poolind2D (2 x 2)	Max Poolind2D (2 x 2)	Max Pooling2D (2 x2)
Dropout (0.5)	Dropout (0.5)	Dropout (0.5)	Dropout(0.5)
Flatten	Flatten	Flatten	Flatten

Table 4.8: Four CNN layer Details

A concatenation function concatenate is used to merge all feature extractors output. An interpretation layer or a fully connected layer with 32 neurons is further added followed by the final output layer.

Activation function *Softmax* is used in the output layer as it acts best in multiclass classification problem. The equation of *Softmax* function can be defined as:

$$\text{Softmax}(z_i) = \frac{\exp(z_i)}{\Sigma j \exp(z_j)} \tag{4.14}$$

In each convolutional layer, a kernel initializer 'he_uniform' is used.

Kernel initializers are used to stabilize the weight initialization and distribution for neural network in each stage of iteration.

Figure 4.7 illustrates the proposed concatenated CNN approach for COVID19 detection.

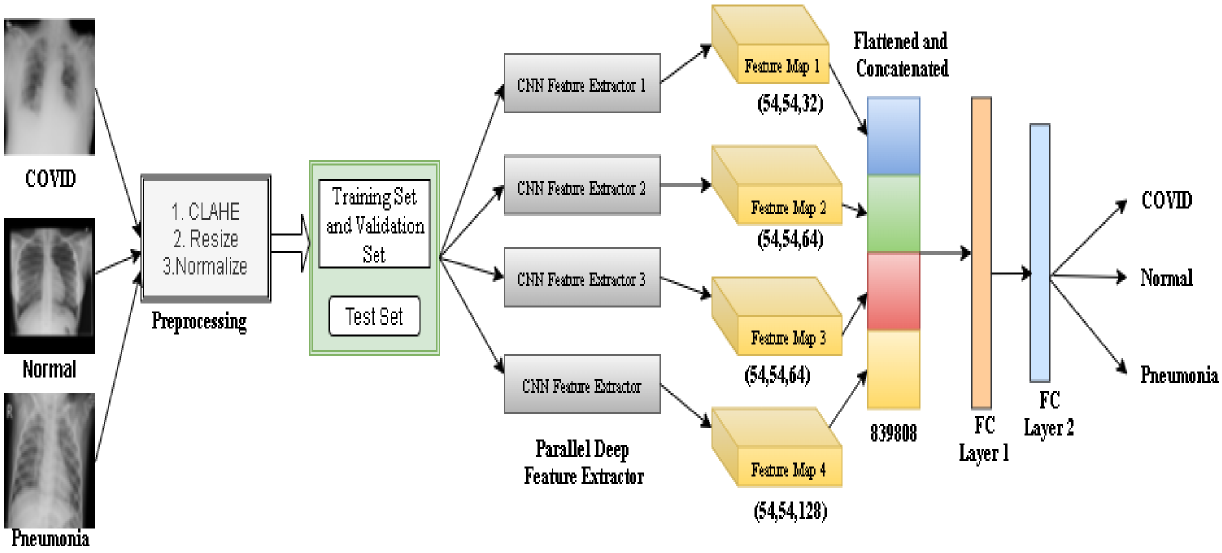


Figure 4.7: Proposed Concat_CNN approach

4.2.4 Proposed model training

The following experiment is conducted on Google Colab GPU (Tesla K80), Python 3.7, TensorFlow 2.7.0 and Keras 2.7.0 libraries are used. The proposed model is compiled using Adam optimizer provided by Keras library with the parameters as learning rate $\alpha = 0.0001$. Categorical Crossentropy Loss function is used as the model is a multiclass classification problem. The model has been trained for 50 epochs with Early stopping callbacks with patience for 10 epochs. To train the model effectively, 5-fold cross validation is considered in addition to splitting of images into different sets. In each fold, a different set of images is used as training and testing set while keeping the validation image set fixed.

4.2.5 Implementation Result and Analysis

4.2.5.1 Performance Metrics

The proposed approach has been evaluated by considering the following four performance evaluation metrics – Classification Accuracy, Precision, Recall AUC [76] as discussed in equations 4.11, 4.12, and 4.13. From the implementation scenario, TP stands for True Positive, TN stands for True Negative, FP stands for False Positive, FN stands False Negative. COVID+ cases that are correctly classified are called True positive cases in the confusion matrix. If COVID+ cases are falsely classified as COVID- then these are called False Negative. Same wise, if COVID- cases are classified correctly then these are called True Negative and COVID- cases falsely classified as COVID+ are known as False Positive.

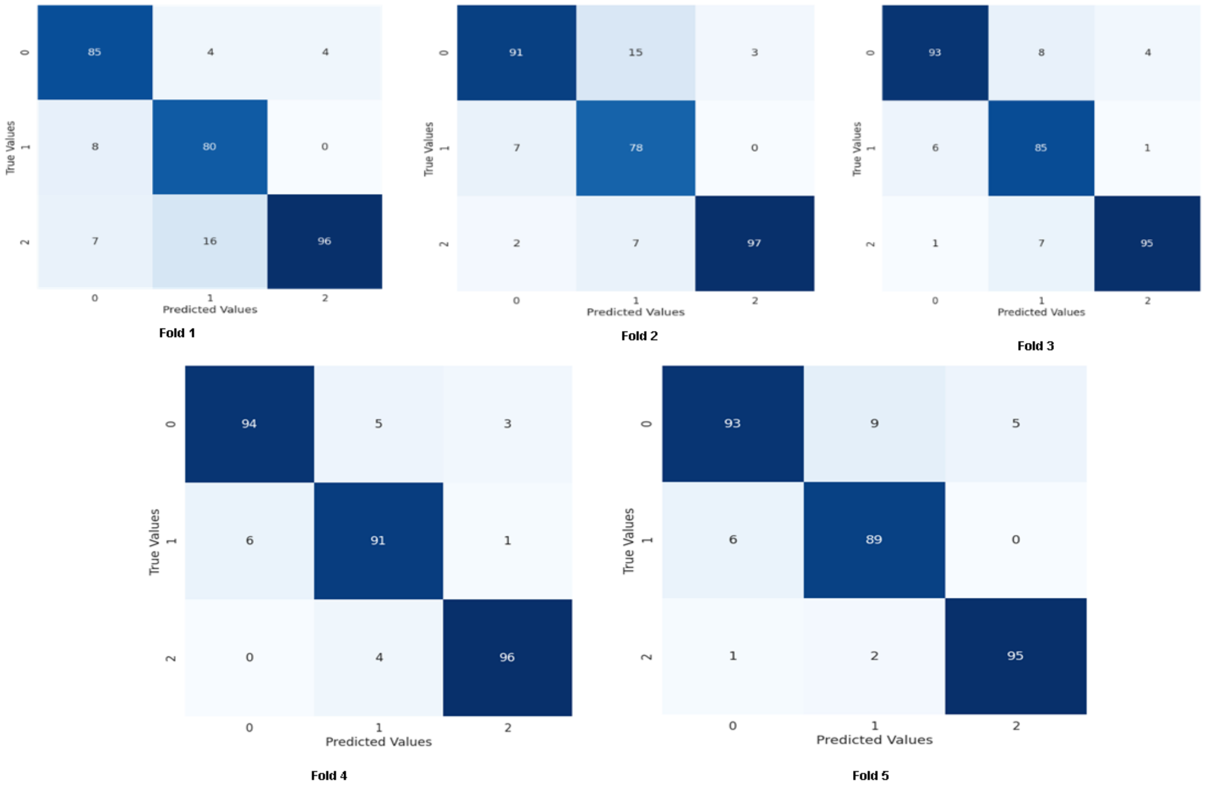
4.2.5.2 Performance Evaluation and Benchmark comparison

In this work, a concatenation-based CNN (Concat_CNN) model for COVID19 disease prediction is proposed. For comparison purpose the proposed model is compared results with other contemporary works done for COVID19 disease prediction using chest x-rays. The present work is also compared with individual transfer learning algorithms performance. This provides a better understanding of the proposed model's performance with respect to other contemporary works and benchmark transfer learning algorithms. As the model is trained with 5-fold cross validation and for 50 epochs with learning rate at 0.0001 and Adam optimizer, same training approach is adopted for all transfer learning algorithm implementation. VGG16[240], InceptionV3[243], ResNet50[241], and DenseNet201[242] transfer learning algorithms are considered. All model evaluation results are saved as .h5 files.

Table 4.9 demonstrates an overview of Accuracy, Precision, Recall and AUC of the benchmark transfer learning algorithms and the proposed model. It is clearly understood from the table that the proposed model exhibits better results in most of the folds in terms of accuracy, precision, recall and AUC. Average Accuracy obtained by our model is 96.31%, average Precision is 95.8%, average Recall is 92.99% and average AUC is 98.02%. It helps to prove the efficiency and robustness of the proposed approach. Figure 4.8 shows the confusion matrix of the proposed model obtained in all five folds. Each confusion matrix provides an overview of consistent performance of the model in all folds.

Fold No	**Model**	**Accuracy**	**Precision**	**Recall**	**AUC**
1	VGG16	87.22%	89.55%	90.74%	91.38%
	InceptionV3	78.55%	89.73%	90.12%	92.34%
	Resnet50	80.11%	82.97%	87.99%	88.80%
	DenseNet121	79.89%	90.56%	92.66%	93.45%
	Proposed Approach	**93.44%**	**91.69%**	**88.33%**	**96.56%**
2	VGG16	89.99%	91.22%	91.86%	93.82%
	InceptionV3	93.44%	93.59%	94.66%	96.44%
	Resnet50	90.55%	91.82%	89.66%	94.77%
	DenseNet121	96.88%	94.88%	92.08%	95.67%
	Proposed Approach	**95.22%**	**93.55%**	**92.00%**	**96.48%**
3	VGG16	94.17%	95.58%	92.00%	94.10%
	InceptionV3	94.11%	95.89%	94.89%	95.54%
	Resnet50	95.33%	96.86%	92.45%	95.09%
	DenseNet121	96.33%	97.25%	94.33%	99.36%
	Proposed Approach	**97.22%**	**97.57%**	**93.99%**	**95.8%**
4	VGG16	96.77%	96.71%	96.00%	95.49%
	InceptionV3	95.34%	96.20%	95.00%	97.28%
	Resnet50	96.44%	96.98%	96.57%	96.77%
	DenseNet121	96.54%	95.45%	95.78%	98.68%
	Proposed Approach	**97.66%**	**97.28%**	**95.66%**	**99.32%**
5	VGG16	96.77%	97.67%	96.85%	99%
	InceptionV3	96.88%	96.00%	94.07%	98.56%
	Resnet50	97.24%	97.48%	95.00%	97.11%
	DenseNet121	97.55%	98.07%	98.33%	98.56%
	Proposed Approach	**98.00%**	**98.95%**	**94.99%**	**99.15%**
Average	VGG16	92.98%	94.14%	93.43%	94.75%
	InceptionV3	91.66%	94.28%	93.78%	96.03%
	Resnet50	91.93%	93.22%	92.33%	94.50%
	DenseNet121	93.43%	95.24%	94.63%	97.14%
	Proposed Approach	**96.31 %**	**95.8%**	**92.99%**	**98.02%**

Table 4.9: Summary of all Performance Comparison with Transfer Learning Algorithms



Fold 1 Fold 2 Fold 3 Fold 4 Fold 5

Figure 4.8: Confusion Matrix of proposed Concat_CNN approach for five folds

While analyzing the confusion matrix of all folds, it is observed that the concatenated CNN model detects most of the cases correctly. From this confusion matrix, it is evident that the model clearly identified true positive and true negative values efficiently. In figure 4.9, Accuracy, Loss, Precision, Recall and AUC graphs are shown for all five folds. In each fold all the metrics exhibit better results than the previous folds. The proposed model also gives consistent high precision which again proves the sturdiness of the model.

Also, to demonstrate the proposed concatenated CNN model's efficacy towards detection of COVID cases based on chest x-rays, a comparison is made in Table 4.10 with some recent studies and their implementation details are also mentioned. Among the different research works, some have shown promising results whereas some have faced difficulty with a small dataset [246, 247]. Some have even proposed new methods which show promising results. From the Table 4.10 it can be observed that the proposed Concat_CNN model has the capability to yield better results than state-of-the-art deep learning models [233,238,239,245] and implementations.

Author's Name	Dataset Description	Model Description	Training Details	Accuracy	Precision	Recall
Tahmina Zebin [238]	802 CXR images	VGG16(Augmentation) VGG16(GAN Augmentation) ResNet50(GAN Augmentation) EfficientNetB0(GAN Augmented)	80% training, 20% testing	88%, 90%, 94.3% 96.8%	93%	90%
Abdullahi Umar Ibrahim [239]	COVID19 – 371 Viral pneumonia – 4237, Bacterial Pneumonia – 4237, Healthy - 2882	Pretrained AlexNet Model for two-way, three way and four-way classification	70% for training, and 30% for testing	Average 95.5% around all classes	Average 93.11% around all classes	Average 96.56% around all classes
Mohammad Rahimzadeh [233]	Normal – 8851, Pneumonia – 6054, Covid - 180	A concatenation network of ResNet50V2 and Xception (5-fold cross validation)	-	Overall accuracy 91.4% in five folds.	-	-
Amit Kumar Das [245]	training images - 771,Testing images – 118validation – 117	An ensemble network of DenseNet201, ResNet50V2, InceptionV3 (5-fold cross validation)	-	Accuracy – 91.62%		Average 95.09 for COVID +ve, average 88.33 for COVID -ve
Liinda Wang [246]	Total – 13975, Covid- 5538, 8066	COVIDNet	Train: Test – 50:1	93.03%		COVID+ - 91%, COVID- - 94%Z
Asif Iqbal Khan [247]	Total – 594, COVID – 284, Normal - 310	Proposed a new model CoroNet.	-	95.3%	98.60%	97.5%

Table 4.10: Comparative study of COVID detection using chest x-ray images.

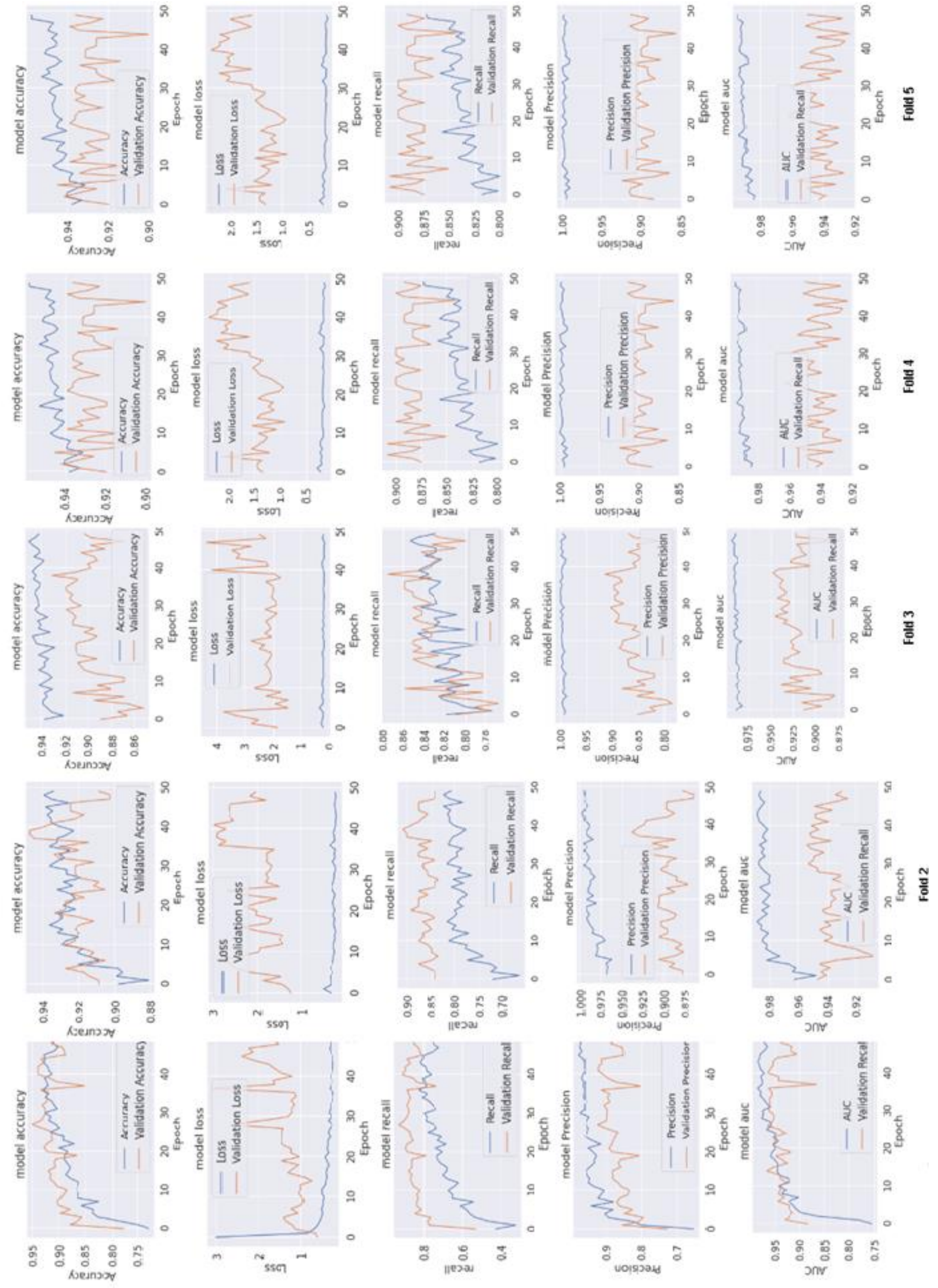


Figure 4.9: Accuracy Precision, Recall, Loss, AUC graph of Proposed Concat_CNN approach

4.3 Case study 3: Bone fracture detection using Transfer Learning Models

Bone fracture can be defined as complete or partial break of human bone. Fractures are caused when more force is applied to a bone than the bone can lay hold of. Causes of bone fracture include fall from height, trauma, or sudden force on body. Bone fractures can also be caused by diseases such as osteoporosis, bone cancer that causes bone weakness. Clinical epidemics can be explained as musculoskeletal suffering or bone fracture. Radiological tests that are effective in fracture detections are X-ray, Magnetic Resonance Imaging (MRI), or Computed Tomography Scan (CT-Scan). In maximum cases, X-ray is the first and foremost important step for interpreting fractures in bone. Despite the limitations of X-rays due to its constrained projection view and limited clinical observations, it acts as primary evidence for any abnormality in bone. Bone fractures can be classified into different categories such as greenstick, transverse, spiral, oblique, compression, segmental, and comminuted. The radiological interpretation of bone fracture is required to be accurate to provide proper clinical support and treatment. The increasing amount of research works on this bone fracture domain suggests importance of developing AI based systems for radio diagnosis. Deep Learning has proven to be beneficial for capturing important spatial information from image, reconstruction of the information for image analysis, and cost-effective applications.

The implementation of DL techniques in comparison to traditional radiology techniques, offer speedup in diagnosis, efficiency in diagnostic procedure, and reducing the workload from doctors to system. However, DL suffers from some vulnerabilities and inconsistency in observation process. In academic research platform, DL models perform well while suppressing human observation for fracture detection and analysis. Since last decade, different research works are carried out for bone fracture detection and proposed models in different studies have high capacity than simple neural networks. Some current research works that used the power of DL for fracture detection is discussed in the next section.

4.3.1 Literature Review

The process of bone fracture diagnosis is not only limited to application of deep learning algorithms in different tasks, but it also includes the path for DL algorithms evolvement, rich representation of specific task, and generation of heat maps for panoramic fracture visualization. In [256], authors used faster R-CNN(Inceptioin-ResNet) for wrist fracture detection from 7356 wrist x-ray images. For classifying pelvic fracture, an encoder -decoder structured network had been represented by researchers in [257] using 789 anterior-posterior pelvic x-ray images. WrisNet model was implemented for detection of wrist fracture [258]. Their dataset consists of 4346 wrist x-rays. The model used Generative Adversative Network (GAN) along with a fracture extraction module and a detection module. A combination of DenseNet121 and Gradcam analysis were implemented on 25505 radiographic images for hip fracture localization [259]. A multiclass classification model is implemented for fracture detection in hand, wrist, elbow, knee, shoulder, foot, ankle and pelvic [260]. The authors named their model FAMO model (ResNet-101 and FPN) for fracture detection. Different researchers used deep learning algorithms for fracture detection. ResNet-152 was implemented for classifying proximal humerus fracture from 1891 shoulder AP x-ray images [261]. Fracture in distal radius region was detected using Faster-RCNN in association with InceptionV4 model from 2340 AP wrist radiographic images [262]. Deep learning models are great help for localization task. Fracture localization is important if there is a need of surgery for fracture. 8329 images are used for scaphoid localization using DCNN(EfficientNetB3) and the implementation is analyzed using Gradcam analysis process [263]. Hip and pelvic area fractures are localized using a MIL-FCN (DensetNet121) network from 4410 x-ray images. The authors mined the ROI region for further analysis [264]. An extensive version of U-Net was implemented by researchers for localizing 11 types of fractures from135845 radiographic images. The areas considered for localization process are knee, spine, foot, hip, elbow, shoulder, femur, ankle, humerus, pelvis, and tibia [265]. A detection mechanism was analyzed for sternum fracture detection from 1221 labeled images. A combination of R-CNN, attention mechanism, and convolutional layer was implemented for further analysis [266].

As theoretical Deep Learning application for fracture detection is always ahead for its applications and shows significant growth in bone fracture detection, there are many aspects of DL models that are required to be explore for fast detection and classification process. In next section, different transfer learning models are analyzed on the curated bone fracture dataset and results are discussed.

4.3.2 Data Collection and Dataset Generation

In this work, bone fracture detection problem is considered as a binary classification problem where an input x-ray image is required to be classified as into fractured or not fractured. The dataset contains total 3500 images of both fractured and non-fractured bone. The dataset contains X-rays of elbow, wrist, humerus, shoulder, hand, forearm, and finger. Some x-ray images are collected through local hospital and diagnostic centers and others are collected from website Radiopedia [269]. The entire dataset structure is described in figure 4.10. The amount of image collected and online resourced is projected in Table 4.11. After curation of bone dataset, the images are contrast enhanced for better understanding of bone and background tissue. The dataset is further divided into 80:20 ratio for training and testing purposes. A csv file has been created to keep track of all images in training and testing folder with bone category and bone label. If a bone label is positive, it indicates bone fracture and Negative indicate no fracture.

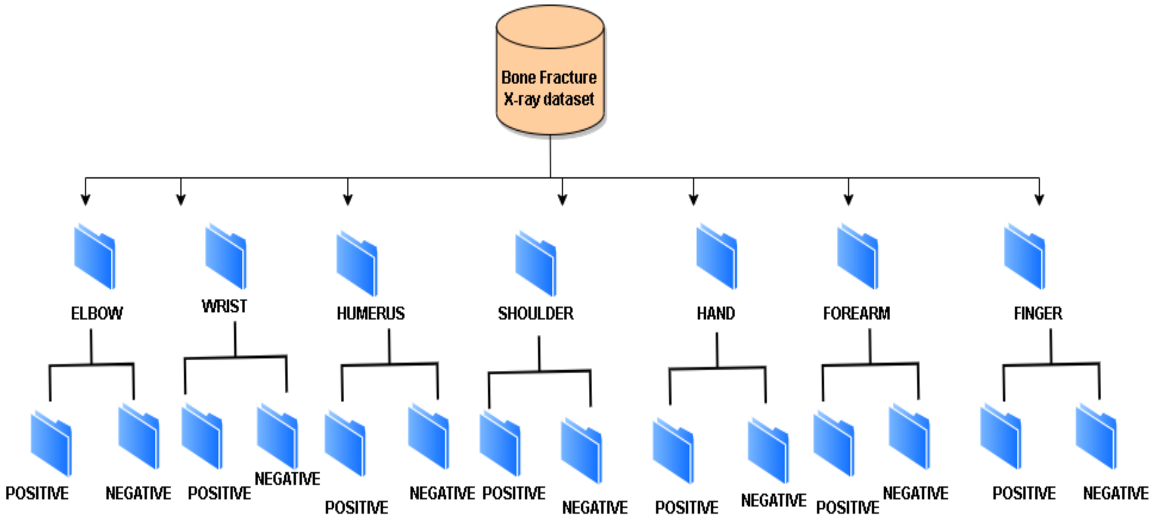


Figure 4.10: Bone Fracture dataset Configuration Details

Area	**Collected**	**Online**
Elbow	**278**	**222**
Wrist	**348**	**152**
Humerus	**222**	**278**
Shoulder	**381**	**119**
Hand	**266**	**234**
Forearm	**182**	**318**
Finger	**394**	**106**

Table 4.11: Collected and Online resourced images for all classes in Bone Fracture dataset.

4.3.3 Transfer Learning Model application

In deep learning territory, transfer learning has proved to be more widespread. The basic concept of building a transfer learning model is to consider a previously trained model on a large dataset and use its training knowledge. Training a deep CNN model from scratch with millions of parameters is very time consuming and requires high computational environment. Weights and parameters of the trained model are used for the new model. This perspective is very important and helpful when the dataset size is comparatively small. Transfer learning models are highly popular in computer vision and Natural Language processing field. The benefit of using transfer learning models is it delivers result fast while incurring lower computational cost. In this study, three pre trained transfer learning models are implemented InceptionV3, ResNet101, DenseNet121. All these models were pre-trained on the ImageNet dataset. Any transfer learning model consists of two parts: a feature extraction part and a classification part with fully connected layers. All three models are discussed below:

- **InceptionV3 Model**

InceptionV3 is a popular pretrained CNN model with 48 layers. Input size of this model is 299 x 299. The layer details of the model are shown in figure 4.11. The model extracts important features from input image and classifies the features in second part. The model consists of over 20 million parameters. The model is contended with symmetrical and asymmetrical building blocks, where each block has various convolutional layers, batch normalization, max and average pooling, dropouts, concatenation, and fully connected layers. The model's fully connected layer use softmax activation function to classify the images.

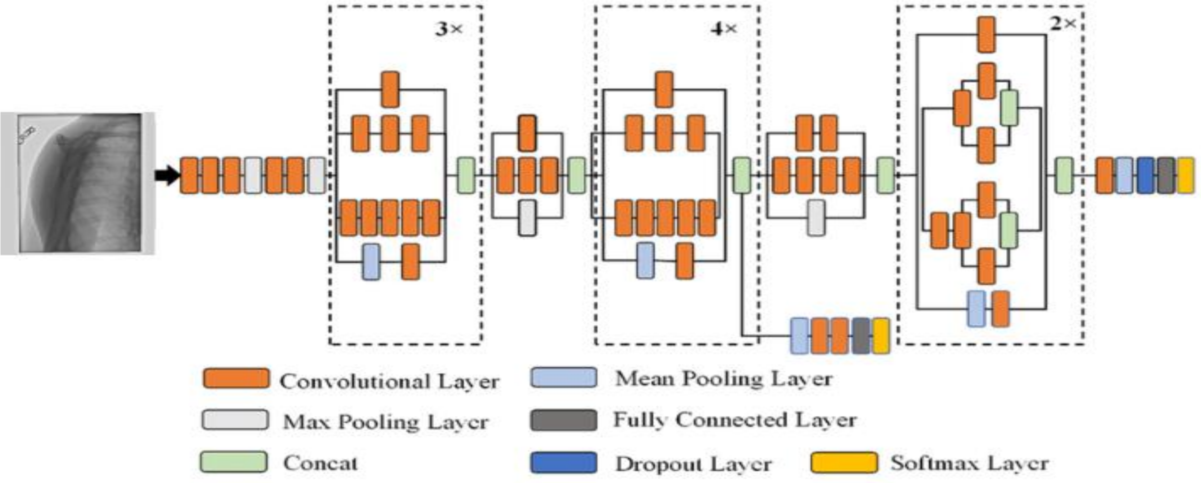


Figure 4.11: InceptionV3 Architecture [267]

- **ResNet50 Model**

ResNet50 is an advanced CNN that has 50 deep layers. The model's input image size is required to be 224 x 224. Each deep learning model suffers from vanishing or exploding gradient problem. When the gradient becomes 0 or too large, this problem arises. With increasing number of layers, training and test error rate also increase. Hence to solve the problem of vanishing gradient problem, residual network was proposed. The skip connection of residual networks skips few layers for training and connects to output directly. The advantage of this network is it allows the network for residual mapping than its underlying mapping process. The architecture of ResNet50 is explained in figure 4.12 below.

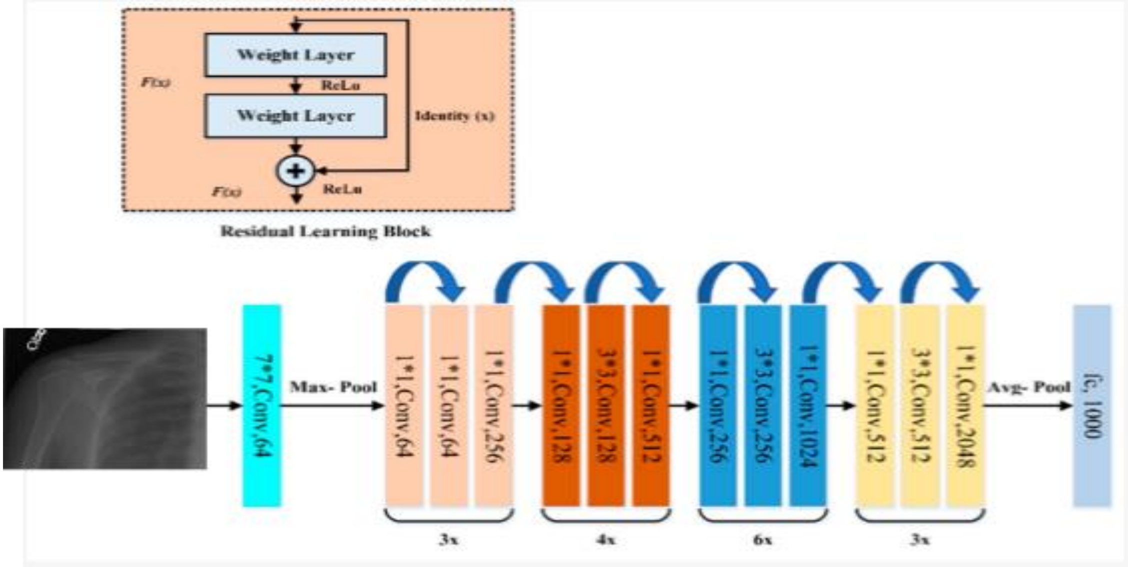


Figure 4.12: ResNet50 Architecture [267]

- **DenseNet121 Model**

DenseNet was proposed as an advanced CNN architecture. DebseNet121 is among one of the DenseNet group of models for performing image classification. It has 120 convolutional layer and 4 Average pooling layers and one fully connected layer. In normal CNN, if the number of layers increases, the model is prone to vanishing gradient problem. DenseNet solved this problem by simplifying the connectivity pattern between layers. In this model each layer is connected directly to every other layer. If the number of layers is L, L(L+1)/2 direct connections are present in the model. Four major components of DenseNet model are connectivity, dense blocks, growth rate, and bottleneck layers. The architecture of the model is shown in figure 4.13.

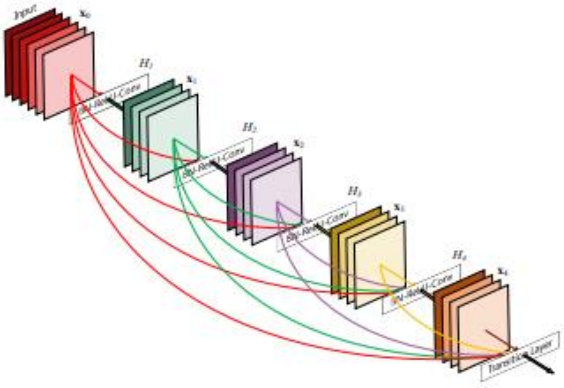


Figure 4.13: DenseNet Block architecture [268]

As discussed earlier, above three transfer learning models such as InceptionV3, Resnet50, and DenseNet121 are implemented on the dataset and results are observed. The flow of classification of bone fracture using transfer learning model is shown in below figure 4.14.

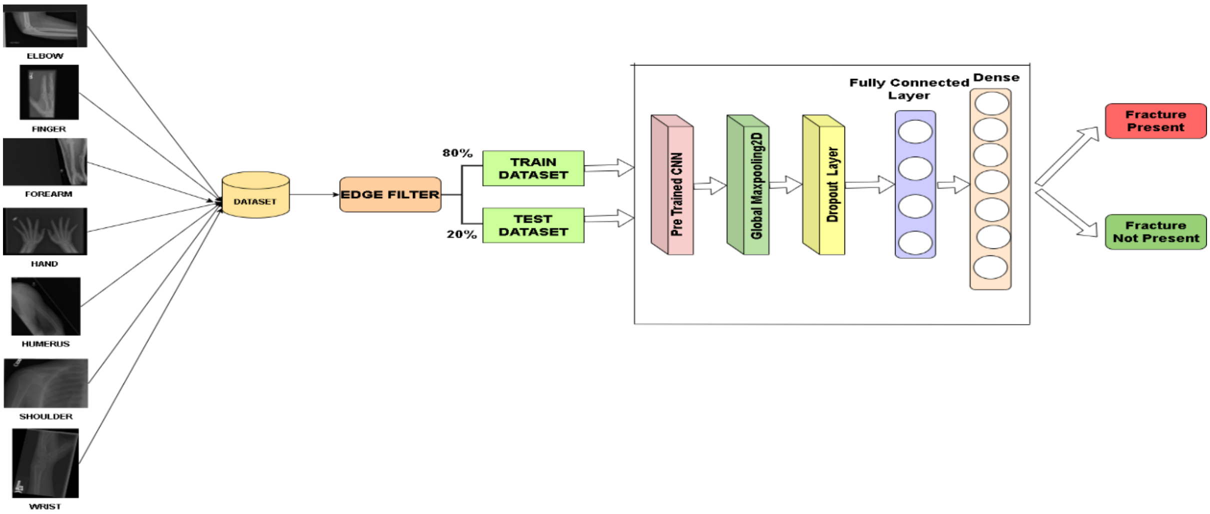


Figure 4.14: Proposed Transfer Learning based approach for bone fracture detection

4.3.4 Implementation Details and Result Analysis

All transfer learning models are implemented on Google Colab GPU (Tesla K80) using Python 3.7, TensorFlow 2.7.0 and Keras 2.7.0 libraries. Adam optimizer was used with a learning rate of 0.0001. Binary Cross entropy loss function was used. All the models are evaluated based on Accuracy, Precision, Recall, AUC. Here, another metric Cohen Kappa is used to find the similarity between the predicted value and the actual ground truth. The equation of Cohen kappa can be defined as:

$$\textbf{Cohen Kappa Score} = \frac{p_o - p_e}{1 - p_e} \qquad \textbf{(4.16)}$$

Where p_o represents observed probability and p_e represents as expected probability. The performance of each table is discussed in below table.

Model Name	**Accuracy**	**Precision**	**Recall**	**AUC**	**Cohen-Kappa**
InceptionV3	**75.89**	**75.8**	**79.99**	**82.4**	**.67**
ResNet50	**74.5**	**74.5**	**75.6**	**78.1**	**.63**
DenseNet121	**78.97**	**79.8**	**71.55**	**85.3**	**.58**

Table 4.12 Performance evaluation of Transfer learning models on Bone fracture dataset.

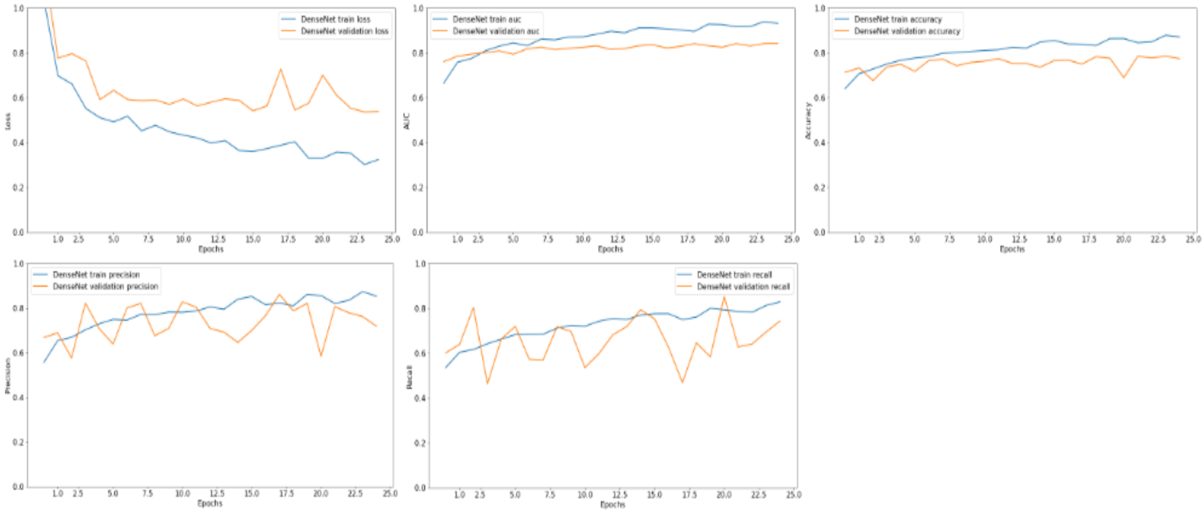


Figure 4.15: Graph representation of all metrics for DenseNet201

From table 4.12, it is seen that Densenet121 model has achieved an accuracy of 78.97%, precision 79.8%, recall of 71.55%, AUC of 85.3% and Kappa score of .58. It is clearly seen that the models are still now not able to provide more accurate results than the current results. A kappa score above .75 is a perfect for assuming the observed probability and expected probability similar. One main drawback of this implementation is the dataset size. The dataset contains low number of images for each bone type. Hence, this is one of the problems in implementing this classification problems. Hence, the dataset size has impact on the model’s performance.

Therefore, in future, the size of the dataset is required to be increased for properly utilizing the strength of deep learning for bone fracture detection.

4.4 Discussion

In this chapter, the proposed models for Pulmonary Tuberculosis Detection and COVID19 detection with deep Convolutional Neural Network are discussed firstly. At first, an automated tuberculosis detection system using chest x-ray with the help of CNN -Dempster Shafer approach is detailed out. As tuberculosis remains an epidemic and number of cases are high, an automated TB detection system will help in early detection and more accurate detection. The proposed model has achieved 94.21% accuracy. So, an automated TB detection system is required that would reduce human error, release excess load on manual diagnosis thereby reducing time for early detection and prevention.

Next, a concatenation-based CNN (Concat_CNN) model is implemented to classify chest x-ray images into three classes: normal, COVID, and pneumonia. Four CNN networks are used as feature extractors to obtain feature vector and are concatenated to get more accurate classification. The dataset has been curated from different open sources. The class balancing is maintained by choosing equal no. of images for all classes in the training phase. The proposed approach has an average classification accuracy of 96.31 % between five folds. The proposed approach seems promising for medical diagnosis as number of COVID cases are still significantly high. So, a computer based COVID detection system will accelerate the detection and classification and will further help to ease the case load problem of medical professionals.

So, these state of-the-art performances will be very helpful as fast diagnostic tools, which will save numerous people from suffering from Tuberculosis, COVID19 or bone fracture. They will prove as great help for mankind where people die every year due to improper and delayed diagnosis.

In the last case study, a transfer learning-based bone fracture approach is implemented. Transfer learning models like DenseNet121, InceptionV3, ResNet50 are implemented on the curated dataset. It is seen that DenseNet121 performed well than other transfer learning classifiers. The main drawback of this implementation is dataset size. So, in future the dataset size is needed to grow for further implementations.

As discussed, this chapter has considered X-ray images for various purposes, like detection of different pulmonary diseases, as well as bone fracture. However, the work in this chapter has not considered one of the frequent illness, like breast cancer that is prevalent over the world. The next chapter focuses on detection of breast cancer by segmenting and analyzing the MRI images using other deep learning approaches.

Chapter 5

Medical Image Segmentation based Classification

Medical Imaging is a process in the field of healthcare where medical professionals recreates various parts of the body for treatment and diagnostic purposes. Medical Imaging includes non-invasive techniques which helps to get an overview of internal structure of human body [273]. These images help in diagnosing different medical condition and diseases. Medical images such as x-rays, CT-Scan, MRI, Ultrasonography, Positron Emission Tomography (PET) requires proper knowledge and training for interpretation. Segmentation of image is an important task that would perform layer by layer observation of organs and lesion for disease detection. Prior to AI era, manual segmentation was used which was time consuming and imposed heavy workload on healthcare professionals. Manual segmentation also involves bias as it is subjective to opinions of different doctors which may vary widely. For any critical image segmentation, a board of doctors were involved which was a time-consuming process where fast diagnosis is essential. Since AI influence is increasing in medical field, medical images can be computed and stored on computers for building an automated segmentation and analysis system. Earlier, traditional methods like mathematical formulas, and edge detection filters were main tools for composing an automated medical image analysis system [274]. Different Machine Learning models such as Random Forest [275], SVM [276], Regression Algorithms [277] are now used for feature selection. Due to advancement of AI technologies, deep learning models showed significant improvement and has tremendous capability in image processing and Computer Vision tasks. Deep learning application in medical image can be classified into four parts: Image Classification, Object Detection, Object Localization, and Image Segmentation. Image Classification has huge contribution in medical image analysis. Image classification will indicate presence or absence of a disease in images. Whereas in localization, presence of an organ or anatomical object is investigated from an image. In detection task, identification of a lesion or object is an important task. Image segmentation is the task of converting an image into different segments based on region of pixel values. It analyzes 2D and 3D image to obtain segmentation, extraction, and reconstruction of image in three-dimensional space. It identifies the required organ or lesion from the pixel values of the whole image and discriminates the organ from the background. Image segmentation helps in quantitative analysis and Computer Aided Detection system. In mathematical terms, image segmentation can be described with the help of set theory as:

A medical image is I and C_i is the set of similarity constraints where i = (1, 2…N), the segmentation of image I can be described as:

$$\bigcup_{X=1}^{N} \boldsymbol{Rx} = \mathbf{I},\ \mathbf{R_x} \cap \mathbf{R_y} = \emptyset,\ \forall \mathbf{x} \neq \mathbf{y},\ \mathbf{x}, \mathbf{y} \in [\mathbf{1}, \mathbf{N}] \tag{5.1}$$

where Rx defines a set of all pixel values in communication similarity Constraint C_i (I – 1, 2,…N) i.e., the image area. Ry defines a set of pixel values. x and y are used to differentiate different regions of the image. N is a positive number not less than 2 which indicates the number of regions after image segmentation. The process of medical image segmentation can be described as follows:

Step 1: Preprocess the image which includes standardization of input image and, perform image augmentation if the size of the dataset is small.

Step 2: Divide the dataset into training, validation and testing set. Training set is used to train the model, validation set is used for adjusting hyperparameters of the model, test set is used for varying and testing the effectiveness of the model.

Step 3: Implement image segmentation model on training set obtained in step 2 to segment the images. The required organ or bone or lesion is segmented from the background image.

Step 4: Evaluate the performance of the model and analyze the effectiveness of the segmentation model on the medical image dataset.

Image segmentation is an integral part of computer vision research and has become an important part of medical image understanding. Segmentation refers to the division of an image into several disjoint areas based on features like grayscale, geometric shape, color, and spatial texture. Different branches of image segmentation are semantic segmentation, panoramic segmentation, instance-based segmentation. With advancement of neural network architecture day by day, image segmentation method is improving with significant accurate results. With the advancement of deep learning models in segmentation methods, traditional segmentation methods like threshold-based segmentation method [278], region-based segmentation method [279], edge-detection based segmentation method [280] are still relevant based on the problem. They can be used as preprocessing step for further deep learning-based segmentation. Traditional image segmentation methods use the concept of digital image processing and mathematical equations to segment images. These are simple to implement, and fast to segment, but the efficacy of the process is not dependable. In recent years, Deep Learning based Image Segmentation models have achieved tremendous growth and have surpassed in accuracy with respect to all traditional models. UNet [281], Region based Convolutional Neural Network (RCNN) [282], Mask-RCNN [283], DeconvNet [284], RefineNet [285] are some popular deep segmentation models which possess strong capability of segmentation and edge detection in 2D- 3D space.

In computer vision application, deep learning is widely used for number recognition, pattern recognition, dimensionality reduction, image segmentation, object tracking. Convolutional Neural Network (CNN) has feature extraction capacity and feature expression capability which are helpful for performing segmentation task. Use of convolutional operation in segmentation task has achieved tremendous success. Deep Learning based image segmentation methods can be categorized as FCN [286], GAN [287], U-Net [288]. All these models use the concept of convolutional operation for image segmentation.

U-Net was first proposed by Ronneberger et al. [288] based on the concept of Fully Convolutional Neural Network which is the most widely used model for medical image segmentation. The concept of U-Net was first presented in year 2015 at MICCAI conference and has been cited in thousands of papers Different variants of U-Net are proposed over the years since it was proposed. U-Net network architecture is shaped as letter U with skip connections. It has an encoder-decoder structure. The encoder has four modules, each having two convolutional layers. After each module, maxpooling layer is used to downsampling. So, encoder part comprises the down part of letter U. The decoder also has four modules. The resolution of the image is increased by upsampling and finally the prediction is made based on each pixel values. The input image size used was 572 x 572, and the segmented image has size of 388 x 388. The output image is smaller than the input image because of the segmentation. The model has no fully connected layer, but only convolutional layer for feature extraction, Downsampling is used for minimizing image resolution, skip connection, whereas Upsampling is used for re-resolution of the image. The architecture of U-Net as proposed in [288] is shown in figure 5.1.

.

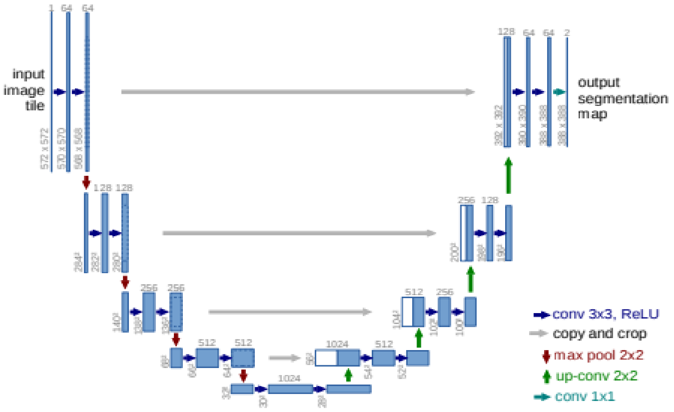


Figure 5.1: UNet Architecture [288]

In this chapter, an Edge Detection based Gabor Nested U-Net approach for medical image segmentation and classification has been proposed for breast cancer segmentation. As early detection of breast cancer will surely increase decrease the mortality rate among woman, so detection of unusual masses in breast is important for proper diagnosis. This chapter is divided into following sections: Section 5.2 describes a literature survey about Breast Cancer Image Segmentation as found in different works. Section 5.3 contains the Data Source, Data Description and Preprocessing techniques. Section 5.4 discusses the proposed Gabor Filter Based Nested U-net architecture, and Section 5.5 contains implementation details along with the comparison of performance of the proposed model with state-of-the-art models. Section 5.6 concludes the chapter.

5.1Literature Review

Breast Cancer segmentation with Magnetic Resonance Imaging (MRI) is an important step in early detection of any anomaly and treatment. According to a breast cancer statistics survey in India, breast cancer contributes to 14% among all cancers in Indian women [300]. 50% of Indian women are diagnosed with cancer at stage 3 or stage 4 and the survival rate becomes less with higher stages. If breast anomalies like tumor or cancer lesions can be diagnosed at early stage, survival rate becomes high. MRI techniques like T1/T2 weighing Image, Diffusion Weight Image (DWI), Dynamic Contrast Enhanced MRI (DCE-MRI) are valuable source of information for breast tissue study [301]. MRI image data analysis for tumor or lesion detection is a difficult and challenging task due to its large volume. Moreover, different imaging artifacts like incorrect positioning of patient during test, background noise, chemical shift, Magnetic Susceptibility, different movements by patient, aliasing defect, can affect imaging quality. So, a Computer Aided AI driven diagnosis system is important to accurately analyze Breast MRI. Therefore, different studies and works had been carried out by different researchers.

Most of the automatic segmentation process are based on U-Net architecture-based CNN model. Ronneberger et al. used U-Net for biomedical image segmentation [288]. U-Net extends the architecture of a CNN model in association with up-sampling and down sampling layers. Long et al. adopted fully convolutional networks such as Alexnet, VGGnet, GoogLeNet and fine-tuned their transfer learning representation into segmentation process [302]. In [303], Cicek et al. applied a 3D U-Net architecture to train 3D volume data with sparse annotations. 3D models help to extract both intra slice and inter slice tumor specifications. A two-stage U-Net architecture was implemented where first stage used a refined U-Net to describe breast ROI, and second stage had an enhanced U-Net model with modified activation function, group normalization, a dense residual module, and the convolutional block in decoder side being replaced with a recurrent attention block. In [304], Piantadosi et al. modified the original U-Net model by mapping the output feature map into a single channel with zero padding and applying batch normalization after each convolutional layer. The model is used to analyze 2D MRI in colonal, transversal, and sagittal planes using voting ensemble approach. An ensemble of three U-Net model is used for effective primary lesion segmentation [305]. Yue et al. [293] proposed a standalone architecture called Res-UNet by adding residual blocks in the encoder part and combining with U-Net skip connections.

In summary, it can be concluded that an encoder-decoder architecture is very important in image segmentation process. Different U-Net models or ensemble of U- Net models were proposed in different studies. Also, there are different implementation of semantic segmentation process for medical image segmentation process.

5.2 Case Study 1: Proposed Gabor Edge Detection Based Nested U-Net approach for Image Segmentation

In digital image, edges are described as local significant change of intensity. An edge can be seen as a set of connected pixels that is present as a boundary between two distinct regions. There are different types of edges such as horizontal edge, vertical edge, and diagonal edge. Edge detection is a process of segmenting a digital image into different regions of pixel values. It is widely used in different AI implementations like pattern recognition, feature extraction, and image morphology. In healthcare, edge detection is an integral part for human organ segmentation. It allows to focus on important regions of interest for further analysis. Digital image may contain noise due to error in imaging process. Therefore, it is necessary to eliminate those errors and enhance every region of image for better outcome.

Different edge detection filters based on derivations in a digital image such as Gradient based edge detection, Gaussian based edge detection are there. Gradient based edge detection operators use first order derivatives of image. Sobel, Prewitt, and Robert are Gradient based edge detection operator. On the other hand, Gaussian based operators use second order derivative of image for edge detection. It includes Canny and Laplacian edge detectors. Though these operators are simple to implement, have time efficient computational process, and detect smooth edges, but they have their own limitations. Sobel filter is sensitive to noise, detects thick, rough edges that may affect accuracy of model prediction, and fails to preserve diagonal points. Prewitt edge operator performs well for detecting horizontal and vertical edges but lacks in diagonal edge detection. Its magnitude of coefficient is fixed. Robert Edge operator detects edges in diagonal position but suffers from noise sensitivity. Canny edge detection is popular among optimal edge detection but lacks in false zero crossing. It is also time consuming and involves high computational complexity. Another gaussian based filter Laplacian is good to detect edges in all directions. But it succumbs to localization error for curved edges. It sometime produces false edges due to noise sensitivity.

Due to constraints of all other filters, Gabor filter is preferred and used for further image segmentation task. Gabor filter is considered as a linear filter popular for texture analysis task. It analyzes whether there is a specific frequency towards a specific direction in the region of interest. Correct edge detection in medical image will have huge impact in disease diagnosis and prediction. Therefore, Gabor filter is suitable for pixel wise edge detection, and it is flexible to detect edges in all directions. As it is a linear filter used for texture analysis, it can analyze any specific frequency in an image with a specific direction in a local region. Gabor filter is highly effective in capturing both spatial and frequency characteristics from an image. It responds at points of texture change. The parameters in Gabor filter helps in producing large magnitude and results edges with bright intensity. Below the functionality of gabor filter is described.

5.2.1 Gabor Filter Bank

Gabor filter is a hand-crafted linear filter which is used for many computers vision application like texture analysis and edge detection. Gabor filter has the capability of a bandpass filter that enhances the characteristics of a specific pattern or edge by elevating the edge or pattern and suppressing the background of the image. Gabor filter is basically considered as a filter bank designed to respond to various frequencies and orientation of objects to detect texture and edge within the images. Edge detected by Gabor filter exhibits similarities with human visuality. A gabor filter is an amalgamation of two components – a Gaussian function and a sinusoidal wave function [307]. So, a gabor filter can be interpreted as a specific orientation and frequency based sinusoidal signal modulated by gaussian wave. A Gabor function can be formulated as [308]:

$$\mathbf{g\ (x, y) = w_r\ (x, y)\ s(x, y)} \tag{5.2}$$

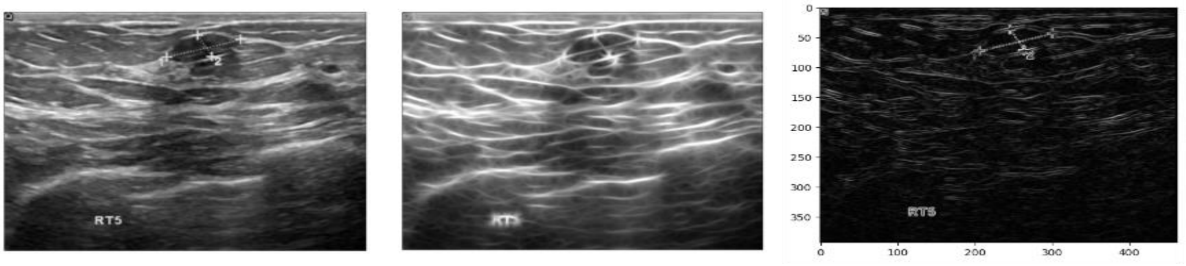
where $w_r(x, y)$ is a Gaussian function and $s(x, y)$ is a sinusoidal function. Equation 2 can be reformulated in a 2D convolutional filter space using the following formula [308]:

$$\mathbf{g\ (x, y, \sigma, \theta, \lambda, \gamma, \psi)} = e^{\left(-\frac{x^2+\gamma^2 y^2}{2\sigma^2}\right)} e^{\left(i\left(2\pi\frac{x}{\lambda}+\psi\right)\right)} \tag{5.3}$$

where θ is the filter orientation, σ represents standard deviation of Gaussian function, λ denotes wavelength of the sinusoidal factor, γ is called spatial aspect factor, and ψ defines as the phase offset.

For the proposed Gabor filter based Nested U-Net approach, gabor filter is used for the first convolutional layer of the encoder part in the model. As the gabor filter is represented as a gabor filter bank, the number of gabor filters for first convolutional layer will be same as the number of output channels of the first convolution layer. As the proposed Nested U-Net encoder's first layer has 32 channel or output, 32 Gabor filters are used.

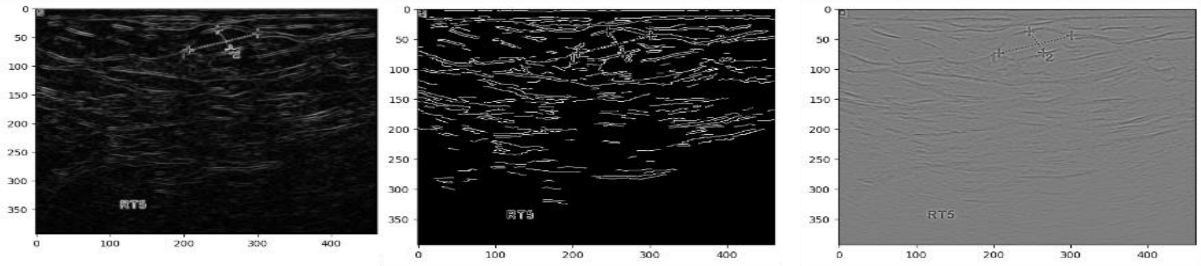
Figure 5.2 shows the edge detection difference after using different edge detection filters and how Gabor filter outperforms all other filters. Filters such as Gabor, Canny, Sobel, Laplacian, Prewitt are implemented for edge detection on a breast cancer malignant MRI image. From the figure 5.2, it is evident that Gabor filter rightfully detects all edges without being affected by noise and the contrast of image is also enhanced due to correct edge detection. This is helpful for further analysis.



Actual Breast Cancer Malignant MRI

Gabor Filter

Sobel



Prewitt

Canny

Laplacian

Figure 5.2: Comparison of different Edge detection filters performance on Breast cancer malignant MRI image

5.2.2 Proposed Nested U-Net Model

Medical Image segmentation, as the name signifies, is a process of partitioning an image into regions based on pixel values. Segmentation of medical image provides numerous clinical benefits. It helps healthcare professionals with fast data processing time, proper measurements, and job specific visualization. It is proven to be a great help in obtaining an insight into abnormal region of human organ or tissue. Image segmentation process can be classified into two categories – Semantic segmentation and instance segmentation. Semantic segmentation focuses on pixel wise segmentation whereas in instance segmentation different objects are identified within the same category. Semantic segmentation was first modeled using deep neural network [306]. Later Fully Convolutional Network was proposed with the concept of encoder and decoder path. [286]. Using the concept of FCN, U—Net architecture was proposed for biomedical image analysis. U-Net model can work on image data such as MRI, CT-Scan, PET- Scan, Optical Coherence Tomography (OCT). The U-Net model is comprised of two parts – encoder part for extraction of contextual and semantic features, and decoder part for spatially increasing the resolution of feature map while reducing the dimension of feature for pixel wise classification. The contraction path generates a low-level fine feature map using the convolution power whereas the expansion path generates a high-level semantic feature map [288]. Skip connection in U-Net is an important part as it copies the output of each stage at contraction path or encoder side and maps it into corresponding expansion path or decoder side. However there remains a semantic gap between the features used for encoder and decoder side. In U-Net the decoder may find it difficult to retain fine grain details of the image and produce proper segmented image.

To eliminate this semantic gap between encoder and decoder, a nested U-Net model was proposed for more efficient medical image segmentation task in [290]. This model is based on the concept of dense skip connection. The objective of this model is to capture fine grain details effectively in encoder path while generating high resolution feature maps and fusion these feature maps correctly with semantically generated feature maps from decoder path. Deep supervision is a process that involves extra supervision to the intermediate layers of deep networks. It can significantly reduce the problem faced while training and make easy the optimization process by avoiding the vanishing gradient problem.

The main difference of the Nested U-Net model from existing U-Net model is the implementation of skip pathways that connects two subnetworks and implementation of deep supervision. In classical U-Net Model, the feature map generated by the encoder is directly fed into decoder. But in case of the proposed Nested U-Net architecture, the feature map undergoes a block of convolutional layers of the pyramid label. In the skip connection pathway between two nodes of encoder and decoder respectively, a block of convolutional layers is introduced where a concatenation layer is present after each convolutional layer that adds the output of previous convolutional layer of same block with corresponding up-sampled block output. In the proposed Gabor edge detection based nested-Net approach, each convolutional block consists of a convolutional layer followed by Batch Normalization and Relu activation Layer. This block structure for proposed approach is different from the block structure described in [290] where each convolutional block consists of multiple convolution layers and a dropout layer thus increasing the number of model parameters. The convolutional block in each skip pathway is used to introduce encoder semantic feature map to the corresponding decoder block. In this way, the optimization process is simplified by making encoder feature map and corresponding decoder feature map similar. The mathematical formulation of Skip Connection pathway is discussed below.

Let output of node $X_{i,j}$ be denoted as $x_{i,j}$ where i indicates the index of down sample encoder layer and j denotes the convolutional layer of the block along with skip connection pathway. So the feature map represented by $x_{i,j}$ can be represented as[290]:

$$\mathbf{x_{i,j} = H(x_{i-1,j}),\ if\ j = 0,\ or\ x_{i,j} = H\left(\left[\left[x^{i,k}\right]_{k=0}^{j-1}, u\left(x^{i+1,j-1}\right)\right]\right)\ of\ j > 0} \qquad \textbf{(5.4)}$$

where function H (·) denotes a convolution function with the association of an activation function, U (·) indicates an up-sampling layer, and [ ] signifies the concatenation layer. If j=0, the node receives only one input from encoder previous layer, and if j=1, node receives two inputs from two consecutive level of encoder sub network. If the value of $j > 1$, the corresponding node gets $j + 1$ input, where j inputs come previous j nodes located in same skip pathway, and remaining one output is obtained from up sampled output of lower skip connection pathway. As convolutional block for each skip connection pathway is used, all previous feature maps are gathered at current node. Figure 5.3 shows how feature map traverses through the skip connection pathway. Introduction of deep supervision helps the model to accurately average the output obtained from segmentation branches and the selection of final segmentation map helps in speed gain and model pruning. Different full feature maps are generated at different semantic levels which are represented as $x_{0, j}$ where $j \epsilon \{1,2,3,4\}$ that are manageable at deep supervision.

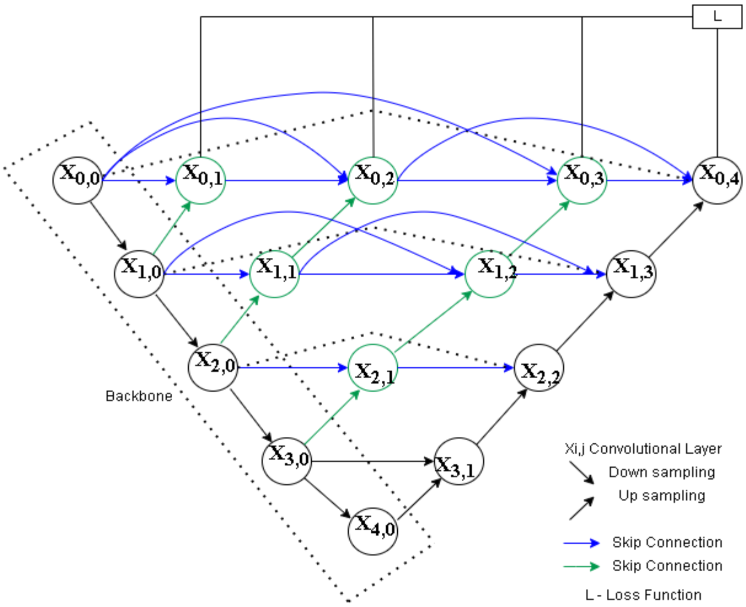


Figure 5.3: Nested U-Net [290]

While training, the encoder path is used to extract features from the input images using down sampling image resolution mechanism. This process doubles the filter space. A convolution layer with kernel size 3 × 3 is used to extract important features from the input images, and then extracted images are passed through Batch Normalization layer and ReLU activation layer. In this first layer of Nested U-Net Gabor filter bank is used as kernel initializer to extract the features using Gabor filter. Average max pooling layer with filter size 2 x 2 is introduced after each convolution block in encoder. This process is repeated throughout the first half of the Nested U-Net approach until maximum depth is obtained. From the second encoding layer, a transposed convolution layer with filter size 2 x 2 is introduced. This helps in up sampling of output for all subsequent convolution blocks. This process helps to prepare the feature maps for skip-connection operation to other layers subsequently.

Once the maximum depth of the network is obtained, the feature maps generated from the decoder's path will be transferred to the decoder side for classifying each pixel value of the feature map into one of the classes or background. Here the mechanism is opposite from the first half of the Nested U-Net architecture because the filter space will be decreasing at each level doubling the image resolution. The skip connections used in the first half of the architecture will be linked up to its corresponding level on decoder side. Each convolutional layer helps in merging encoder into corresponding decoder's feature map with the help of a subsequent Convolutional Transpose layer. This process is used to up sample the feature map data. A secondary segmentation layer is used next to last layer of skip connection pathway. The segmentation output of this layer is added applying elementwise addition operation with the final segmentation output to accurate the final model's output. In the classical Nested U-Net architecture as proposed in [290], the authors have introduced a loss function comprising of binary cross entropy and dice coefficient loss. Binary Crossentropy Loss function is defined as:

$$\text{Binary Crossentropy (L)} = -\frac{1}{N}\sum_{i}^{N} \sum_{j}^{M} yi,j \log(pij) \qquad \textbf{(5.5)}$$

The main difference between U-Net and Nested U-Net architecture can be defined like this: Nested U-Net has convolutional layers with skip connection pathways, and it will be trained under deep supervision which will help in training and model pruning, dense skip connections in skip connection pathway. This will help in improving gradient flow at each layer.

Figure 5.4 elaborates the proposed Gabor nested U-Net approach for image segmentation. The dataset is first augmented to increase the dataset size. It is further divided into training and testing set into 80:20 ratio. The proposed approach is trained with the training data, and the performance is evaluated on test dataset. The proposed approach finally predicts the mask of the required segmented region.

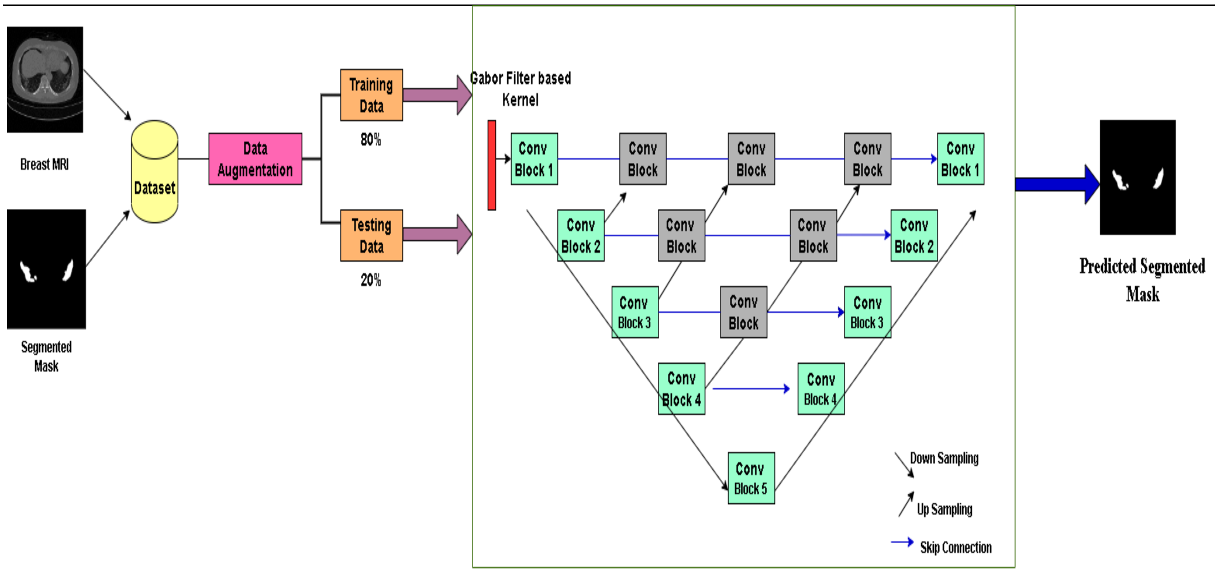


Figure 5.4: Proposed Gabor filter based Nested U-Net image segmentation approach.

5.2.3 Data Description and Data Preprocessing

For implementation purpose, Breast Cancer Ultrasound Dataset is considered [289]. Breast cancer is one of the most common diseases among women worldwide. So early detection and prediction of this disease is most important for overall mankind health. Ultrasound scan is one of the easiest ways to locate breast cancer. The dataset [289] is categorized into three parts classes Benign, Malignant, and Normal. Breast Ultrasound images are great sources for classification, segmentation, and detection. Ultrasound images for this dataset was collected among women aged between 25 to 75. 600 female patients were considered. The dataset has total 830 images of 500 x 500-pixel size. The number of images for class benign is 487, the number of images for class malignant is 210, and number of images of normal class is 133. For segmentation, the mask of all three categories like normal, benign, malignant is required. Mask can be defined as the region of interest. The considered dataset has masks for all three categories. As seen in figure 5.5(a), the number of instances (images) for each class is less for implementing any deep learning segmentation model, hence data augmentation method is applied. This generates synthetic images for each class for increasing per class instances. The parameters for data augmentation are horizontal flip is true, rotation range is 40, shear range is set at 0.2, zoom range is 0.2, brightness ranges between 0.5 to 1.5. So, after data augmentation, the number of instances as shown in figure no. 5.5(b) for benign, malignant, and normal class are 1311,630, and 399 respectively. For the implementation purpose, the images are resized into 128 x128 x 3 where 3 denotes the images in RGB format and corresponding mask of the images of size 128 x 128 x 1 where 1 denotes the images in Grayscale format.

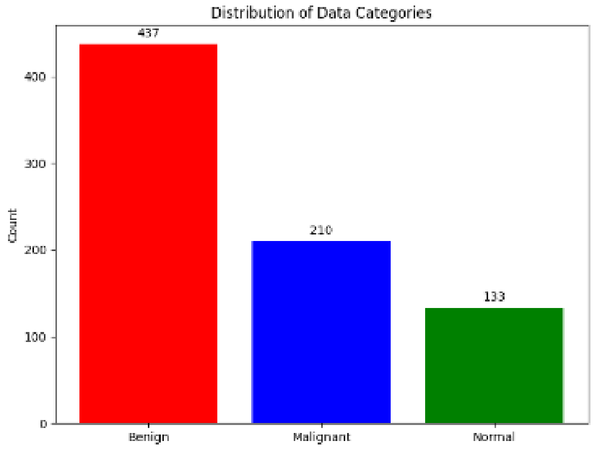


Figure 5.5(a): Data distribution for Benign, Malignant, and Normal Class

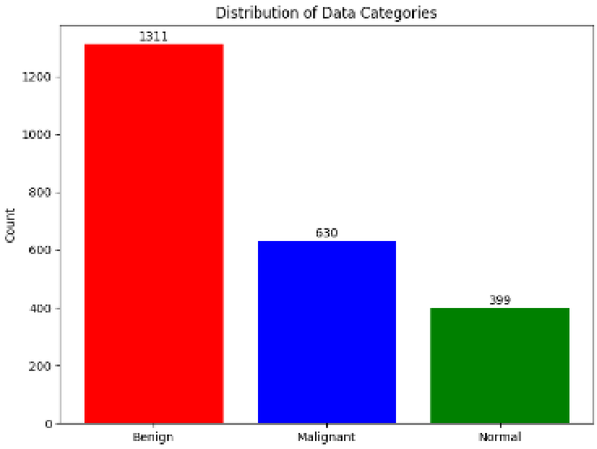


Figure 5.5(b): Data Distribution for Benign, Malignant and Normal Class after Data Augmentation

5.2.4 Implementation Details and Performance Analysis

The kernel parameters for each gabor filter are set uniformly to predefined range. The range for sigma, gamma, theta, lambda, and psi are defined in table 5.1. In proposed Nested-U-Net model, the range of each parameter is divided into 32 parts uniformly and for each channel or output of first convolution layer, a gabor filter is generated using these parameters.

Parameter Name	**Parameter Symbol**	**Range**
Sigma	σ	(2,21)
Theta	θ	(0,360)
Lambda	Λ	(8,100)
Gamma	Γ	(0,300)
Psi	Ψ	(0,360)

Table 5.1: Gabor filter parameter details with value range for proposed approach

Google Colab GPU (Tesla K80), Python 3.7 have been used for this implementation purpose. TensorFlow 2.7.0 and Keras 2.7.0 libraries are utilized for the proposed Gabor filter based Nested U-Net approach implementation on Google COLAB platform. As discussed in earlier section, the dataset on Breast Cancer MRI dataset is used for implementation. The original images and its corresponding mask are resized into 128 x 128-pixel values. The augmented dataset is divided into 80:20 ratio for training and testing purpose. The performance metrices, Accuracy, Loss, Recall, Precision, AUC, Jaccard, Dice Coefficient are discussed below with an early stop mechanism on testing set. Adam Optimizer is used for compilation with a learning rate at 3e-4(0.0003). The model has been trained for 100 epochs.

$$\textbf{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{5.6}$$

$$\textbf{Precision} = \frac{TP}{TP+FP} \tag{5.7}$$

$$\textbf{Recall} = \frac{TP}{TP+FN} \tag{5.8}$$

$$\textbf{Jacard} = \frac{|A \cup B|}{|A \cap B|} \tag{5.9}$$

$$\textbf{Dice\_Coefficient} = \frac{2TP}{2TP+FP+FN} \tag{5.10}$$

$$\textbf{IOU\_Coefficient} = \frac{TP}{TP+FP+FN} \tag{5.11}$$

Where TP stands for True Positive, FP stands for False Positive, FN stands for False Negative, and TN stands for True Negative.

Model Name	Accuracy	Precision	Recall	AUC	Jacard	Dice Coefficient	IOU Score
UNet	96.81%	84.78%	70.10%	95.56%	53.36	68.68%	43.14%
Residual Unet	96.18%	81.90%	63.20%	93.93%	47.60%	64.39%	38.01%
Attention Unet	95.59%	81.00%	53.90%	91.36%	38.40%	55.40%	27.66%
Proposed Approach	99.44%	96.79%	95.99%	.99	91.22%	96.39%	77.45%

Table 5. 2: Performance Evaluation of Proposed approach with State-of-the-art Models.

For comparison purpose, other segmentation models like classical U-Net, Attention U-Net, Residual U-Net are also implemented using the same dataset. All the models are reimplemented according to each model’s best interpretation available. All the models are compared based on accuracy as overall segmentation goodness which can be defined as a ratio of the total number of correctly classified pixel values and total number of pixels in the image. Precision defines the relevance of pixel values from all retrieved pixel values, Recall is the percentage of correctly identified pixel values that belongs to the class of interest Dice Similar Coefficient is to find similarity between original mask and segmented mask. Jaccard Similarity Index is used to measure the similarity and diversity between original and segmented mask. Intersection Over Union is used to measure the overlap between the predicted bounding box and ground truth bounding box. To compare the obtained breast mask and original breast mask quantitatively, the measures are considered based on True Positive, True Negative, False Positive and False Negative and the overlap region measure. Table 5.2 shows the performance metrices of all state-of-art models and the proposed approach. From the table below it is evident that proposed Gabor Filter Based Nested U-Net architecture excels in almost all performance parameters than other state of art models. It obtained an accuracy of 99.44%, Precision of 96.79 %, Recall of 95.99%, Jaccard Coefficient of 91.22%, Dice coefficient of 96.39%, and IOU of 77.45%. Figure 6 shows the Performance of all metrics for 100 epoch training.

The outcome of the predicted segmentation mask can be seen graphically. The original MRI image, Original Mask, and predicted Segmented Mask are shown in figure 5.7. Although the proposed model can segment the Mask, some undesirable areas are present in the segmented image. The images reveal that the proposed model is capable of accurately segmenting the required region. Also, if the overlapping metrics such as Dice Coefficient and Jaccard Index are considered, the results are quite satisfactory with values of 96.36% and 91.22% respectively. As Dice similarity coefficient measures the reproducibility of manual segmentation and automated probabilistic segmentation, 96.36% value of dice coefficient for our proposed method proves that proposed approach is segmenting the required region correctly. The implementation of the model shows that the output of Gabor filter based Nested U-Net is accurate for breast cancer segmentation.

Also, to demonstrate the proposed approach’s efficiency towards Breast Cancer detection from MRI images, a comparison is shown in table 5.3 with some recent proposed segmentation models and their implementation details in different studies.

Author Name	Year	Dataset Details	Model Details	Performance Evaluation
Piantadosi et al. [291]	2019	132 DCE- MRIs of 66 non-cancerous patients	2 Consecutive U-Net, 3 class UNet	Dice Coeff – 90.49%
Ma et al. [292]	2020	Bilateral breast Pre contrast MRI of 100 patients.	Unet and Patch DCNN	Dice – 87.00%, Jaccard – 77.6%
Yue et al. [293]	2022	100 Breast Cancer Patient	Res-UNet	Dice – 0.89,
Adoui et al. [294]	2019	DCE-MRI 43 cancer patients	Seg- Unet and U-Net	Seg U-Net IOU – 68.88%, U-Net IOU – 76.14%
Li et al. [295]	2019	DCE-MR 33 cancer patient	3D modified FuseNet Based on UNet	Dice – 77.6%, Sensitivity – 84.4%
Zhang et al. [296]	2019	MRI of 272 patients	A MHL Framework using U-Net architecture	Dice-71.76%, Sensitivity – 75.04%
Peng et al. [297]	2022	Breast Cancer MRI's of 590 patients	DenseVOXNet	Dice for T1C – 90.49, T1 – 85.07
Khaled et al. [298]	2022	46 cases of TCGA- BRCA dataset.	Multi Modal Ensemble learning using modified U-Net	Dice coeff - .79%
Hirsh et al. [299]	2022	Breast Cancer MRI-2555, Normal MRI – 60108	3D U-Net	Dice Coeff – 77%
Proposed Approach	2024	Total 830 images Benign – 487, Malignant – 210, and Normal- 133	Proposed gabor-Nested U-Net	Accuracy – 99.44%, Dice_coeff – 96.39%

Table 5.3: Comparison of Proposed approach with state of the-art research works.

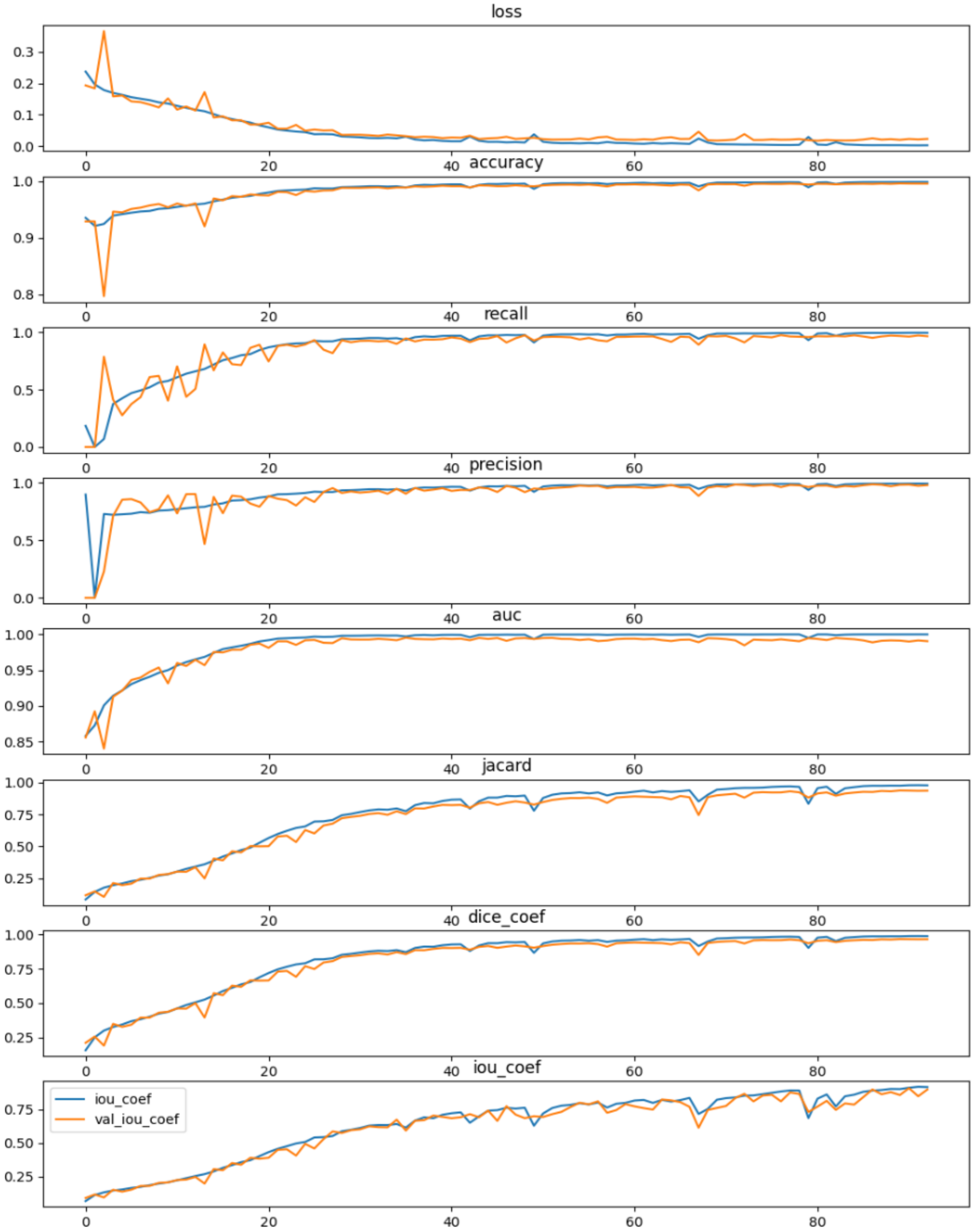


Figure 5.6: Accuracy, Loss, Precision, Recall, AUC, Jacard, Dice Coefficient, IOU Coefficient graph for Proposed Gabor Filter based Nested U-Net Approach

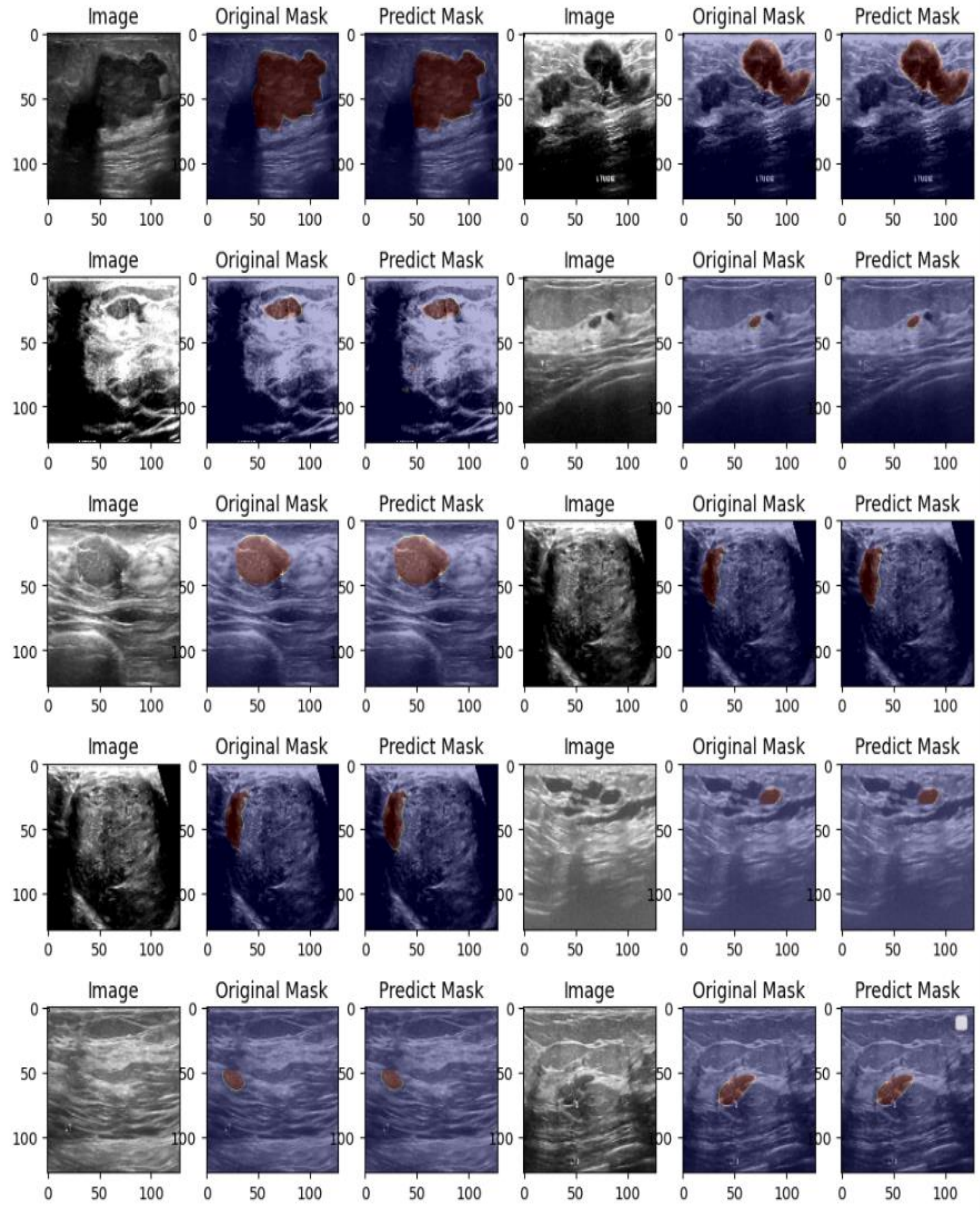


Figure 5.7: Original Image, Original Mask and Predicted Mask

5.3 Discussion

Early detection of breast cancer can substantially increase the rate of survivors. In the age of AI, medical image analysis has become more reachable, cost effective, almost real time, and more accurate while focusing on extraction of important features from large dataset. The proposed Gabor Filter Based Nested U-Net architecture aim to analyze breast cancer MRI images and compares its efficiency with respect to other classic U-Net models and state of-art models described in different research papers. This present work includes processing and generation of augmented images to eliminate class imbalance, detection, and segmentation of Breast Cancer MRI images from its corresponding mask. Despite achieving promising result by proposed approach, several challenges must be mitigated before incorporating DL image segmentation methods into clinical practice. Proper knowledge for interpretation of DL Image Segmentation Model is required for unambiguous detection and segmentation results. These systems need proper interpretation and the ability for addressing ethical matters. So, the proposed approach has potential for adopting AI based systems for Breast Cancer diagnosis based on Imaging Dataset. This hence indicates the urge of an AI integrated fully automated Architecture that can be reliable for Breast Cancer detection with minimum human interaction.

Chapter 6

Conclusion

This final chapter aims to summarize the issues of predictive analytics in healthcare with respect to machine learning and deep learning. The chapter also provides a summary of the research work presented in this thesis while mentioning the major contributions and the significant results. The last section of this chapter focuses on the future scopes of the present work.

In recent times, AI based application in healthcare have become common in practice. Because of the growing nature of healthcare data and growing difficulty in analyzing these data in time, computer aided automated (CAD) prediction systems are fast developing to meet the need. Application of Machine learning and Deep Learning have potential advantages including fast and proper treatment, accurate disease diagnosis, improved healthcare facilities in remote areas. But one major concern in utilization of predictive nature of ML and DL, is the quality of data. As the amount of data is huge, proper processing of data while maintaining the integrity of data is a huge concern in current scenario. Manual processing of huge data is not possible, and it poses a threat if not processed wisely. So, how ML and DL algorithms predict and detect disease from different type of healthcare data is needed to be emphasized. This thesis focuses to explore and address how ML and DL centric disease prediction systems will help e-health applications in future.

6.1 Problems Identification and Analysis

Initially the focus of the thesis is to analyze different ML and DL based application on healthcare disease data. It is well known healthcare disease data analysis is important for patient’s well being and course of treatment. Sometimes healthcare data may be ambiguous that may affect or delay the disease detection process. If data is not correctly validated, it may create confusion in the learning process of a model. Now a days, numerous facets of clinical practice believe in different computational tools for anomaly identification, regulating patient care, evaluating disease aspects, and for many more tasks. In case of supervised learning process, a model trained on a dataset may learn some specific features for a particular data and label pair, which may not be true for other type of data with same modality. A generalized model may be fragile for many reasons such as different noise sources for different types of data, selection bias present in the model, calibration effect and many more. Building a generalized model is a prime concern for clinical research. At the same time, a model working for one healthcare organization may fail for other healthcare organization due to implementational and environmental differences. Therefore, the key challenge is to understand the issues while training, validating, and testing a model. Below some problems that are identified while furnishing the research work are mentioned.

- **Complexity in Data**

Clinical data for ML and DL application comes in different formats such as clinical notes, prescriptions, radiological images, electrocardiograms, and others. Before performing any analysis of a model, data should be collected from reliable source, cleaned, and validated. This data processing task itself is a huge data science task rather than model implementation and optimization. So, assuring the data quality is an important pillar for any ML or DL application. Erroneous value, missing data, absence of validation data will mislead the prediction of disease. So, validation of data should be done with multiple experts. ML or DL models are robust for few incorrectly labeled data, but it is best to main the data quality. Another major concern of data is the complexity of medical data. Data from ECG, EEG should be correctly interpreted and labeled by experts for proper model learning.

- **Using Standard Model**

After data processing, defining a model and its implementation is an important step. Maintaining a baseline standard model is essential to analyze performance tradeoff between simple model and complex models. Complex models tend to produce more accurate prediction results while simple model is easier to interpret. Many ML and DL studies do not include any baseline comparison. Lack of comparison makes it difficult to analyze if complex model is more robust than simple model. In healthcare application, improvement in performance is most valuable for accurate disease prediction.

- **Performance Metrics**

There are many ways for analyzing the performance of a model. In case of a binary classifier, a confusion matrix identifies all the outcomes in a 2 x 2 table. These four values True Positive (TP), True Negative (TN), False Positive (FP), False Negative (FN) can be combined in different ways to generate different metrics such as accuracy, precision, recall etc. In classification problem, a single performance metric is more convenient than any performance curve while comparing different models. A perfect model can cover entire sensitivity-specificity area, so area under Receiver Operating Characteristics (ROC) can provide a single number that can be used for model ranking.

- **Imbalance dataset**

Performance metrics of a model captures different aspects of a model. But these get complicated if there is data imbalance in class distribution. Due to imbalance, model may become prone to be biased towards the class having higher number of instances. Another problem with imbalance dataset is that it shifts the AUC-ROC metric which makes the model look like more responsive.

- **Model Training Metrics**

ML and DL models have different parameters such as optimizer, learning rate, loss function other than the performance metrics. However, many complex models have a set of parameters that fits the training data well. Choices of a model training are not directly involved with model parameters but have effect on model performance. These parameters are model optimizer, model architecture, data processing steps etc. These parameters have direct influence on model's performance from a single digit to a distribution over different hyperparameters. Moreover, training data and its associate labels have huge contribution in a model's performance under uncertainty.

- **Trails and its Associated Uncertainty**

In any scientific research, repeated trails play important role to estimate uncertainty. This process is critical for any ML or DL application. Sampling process of training and testing data with unknown distribution reflects in model's uncertainty. Repeated trials lead to performance deviation. Performance distribution is crucial in clinical health application as it measures potential deployment risk as the model could be inappropriate for standard healthcare morns.

- **Data Partitioning**

There are lot of discussions about the partitioning process of a dataset into training, testing, and validation dataset. The goal of partitioning is to minimize training and testing errors. Carefully crafting a testing dataset is necessary as inconsistent data splitting tends to put a model in non-generalized state.

- **Training and Cross validation**

Most ML or DL techniques are complicated and takes many decisions while training. Different sets of hyperparameters produce a list of models. So, validation set is required to statistically test these models and rank them. So, the size of validation set is important to differentiate between two regularization measures.

- **Mitigating Overfitting and Bias**

ML or DL models tend to suffer due to small datasets. Overfitting is defined as fitting training data very well and predicting test data poorly. As overfitting reflects the performance differences of a model in different datasets, it requires to produce performance metrics across training, validation, and testing set. An accurately fitted model captures required features. Establishing a well fitted model is important in healthcare applications as the model may sometimes tend to ignore random features and focuses on attributes that helps in robust health outcome.

Besides mitigating overfitting or underfitting, a model is required to be unbiased towards a particular datatype. Bias will not let to sample a small dataset. Moreover, due to small data size, transfer learning models will produce ambiguous results leading to decision inequality or even ambiguity.

- **Model Prediction Explanation**

One important aspect of a predictive modeling is to understand a ML or DL models prediction process. Assessing the feature dominance in accuracy, and how input and outputs are related to each other. While there are many methods to interpret a model, many ML techniques do not interpret the model’s prediction power. Instead, it relies on performance metrics for defining the model’s effectiveness.

- **Comparing performances**

When it comes to compare different model’s performances, it is required to train each model on same training set, all preprocessing steps should be same, and the computational environments should be same. Eliminating the experimental difference will ensure that the performance difference is purely based on different model’s approach.

6.2 Contribution

Application of ML or DL in healthcare has gone through numerous challenges that are required to be explored. Due to advancement of technology in health sector, predictive analysis with the help of ML and DL will be able to produce fast, dependable results. It will also have the capability to handle large complex data. Therefore, automated ML or DL based systems will provide high quality of patientcare. ML and DL provide novel approaches and systematic programming environment for analyzing data in real time, allowing to utilize ML’s prediction power to its fullest. Many healthcare data contain attributes such as name, patient id, demographic information etc. that are not always relevant for disease analysis. So, a feature selection process is necessary to predict accurate result based on the important features that defines a certain condition in a patient. Hence, a feature selection method will not only accelerate the prediction process but also increase prediction accuracy. For healthcare data, patients' critical data is very sensitive. So, selection of these sensitive data that affects the output is utmost important. Keeping all the requirements in mind, feature selection approaches are proposed and how these approaches are affecting the performance of different Machine Learning models are discussed in this work.

Image data is a great source of information in healthcare industry. It provides an overview of human organs and any abnormalities among them. So, analyzing these image data is important step for disease detection. In case of any bone fracture, or infectious disease affecting lungs, or any presence of tumor in any organ of human body, Deep Learning is a great help in identification and prediction of chronic disease, any physiological abnormality (like fracture), epidemic outbreak, etc. and smart health monitoring. For application of deep learning, preprocessing of image is an important step. Images collected from different sources may get distorted, or noisy. An enhanced image provides a clear picture of an organ and its abnormalities. Therefore, preprocessing of image is necessary. In the proposed approaches, image contrast is first enhanced using Contrast Limited Histogram Equalization Algorithm. Contrast enhancement helps to distinguish the foreground and background of the image. Due to recent pandemic, COVID19 data is an important area to explore the strength of Deep Learning. A concatenation of CNNs' is used for COVID19 detection from chest x-ray images. Each CNN in this approach is used as a feature extractor to extract important features from input image. Another epidemic disease, Tuberculosis is used for another study in deep learning. An evidence-based CNN approach is discussed to identify the tuberculosis infection presence in chest x-ray image. Bone fracture is a physiological disorder. To detect a fracture bone from x-ray image, images of different types of bone fracture are collected from different sources and used in the experiments. Due to lack of adequate number of images, online bone fracture x-ray images are also used to build a dataset. Different transfer learning algorithms are implemented on this data to detect if a bone is fractured or not.

Image segmentation is a process to segment out important regions of interest for further analysis. It has proved to be helpful in tumor detection, mass detection, border detection. Edge detection is an important process for any segmentation task as it helps to localize the region of interest. Hence, an edge detection-based image segmentation approach is considered for breast tumor detection and prediction for benign or malignant.

The main contribution of the research work is discussed as follows:

- Proposing feature selection approaches and evaluation of machine learning models on obtained feature subset for Parkinson disease data.
- Prediction of Tuberculosis disease from chest x-ray data using an evidence based Convolutional Neural Network.

- Detection of COVID19 disease from chest x-ray dataset applying a concatenated CNN model.
- Use of transfer learning models for bone fracture detection.
- Segmenting Breast tumor tissues from breast MRI data and classification under benign or malignant.

6.3 Result Significance

Different Machine learning algorithms are used in association with different feature selection approaches in the experiments to predict the presence of Parkinson disease in a patient in chapter 3. In this experiment, two feature approaches are proposed. In first approach, a hybrid Filter -Sequential Forward based feature selection process is used where Random Forest and Bagging Classifier performed well with an accuracy of 91.67% for UCI dataset. In case of Hand-PD spiral dataset, Decision Tree classifier achieved an accuracy of 96.23%, precision of 93.33%, recall of 100% and for Hand-PD meander dataset, KNN and random forest achieved highest accuracy of 86.79%, precision of 87.5%, Recall 76%. The confusion matrix for each ML model on three datasets are shown in figure 3.9, 3.10, and 3.11. However, sequential -forward feature selection approach has a limitation. The final feature subset size is to be mentioned earlier. So, a filter -Whale optimization-based feature selection approach is described. Here, bagging classifier performed well on UCI dataset with an accuracy of 96.67%. Confusion matrix for each dataset is described in figure 3.16, 3.17, and 3.18. ROC and PR curve for each class and each classifier is explained in figure 3.19,3.20, and 3.21. These feature selection approaches are compared with no feature selection process where ML models are directly implemented on whole dataset. From experimental results, it is analyzed that feature selection approaches performed well while implementing ML models than no feature selection ML model implementation.

The strength of CNN is analyzed on different types of image data in the next chapter 4. The first approach of this work was to detect Tuberculosis from chest x-ray images. So, an evidence-based CNN was implemented. The proposed model is trained for 5 folds and 30 epochs per fold. An average accuracy of 94.21% is achieved. The proposed approach is compared with transfer learning models, and it is seen that the proposed model performed well than transfer learning models. A comparative study is presented in table 4.5 with state-of-art models implemented on same dataset. In second approach for Deep Learning based application, Covid19 is predicted from chest x-ray images. The proposed concatenation-based approach gained an accuracy of 96.31% for 5-fold validation. It is also compared with existing transfer learning-based models such as VGG16, INceptionV3, ResNet50, DenseNet121 and summary of their performance is discussed in table 4.9. Lastly, in third application of deep learning, the strength of transfer learning models is compared on bone fracture dataset. The results of transfer learning models are not satisfactory due to dataset size.

The experiment with image segmentation process is discussed in chapter 5 and how image segmentation helps in disease detection is discussed. The proposed edge detection-based image segmentation process achieved its accuracy of 99.44%, precision 96.79%, recall 95.99%, Jaccard index value 91.22%, Dice coefficient 96.39%, and IOU score 77.45%. Figure 5.5 shows how accurately the proposed approach predicts the masks for input images that are almost same as original masks.

6.4 Future Scope

In present work, different ML and DL based disease detection application are explored, discussed and few aspects are implemented for prediction task. The immediate concern of this work is to check these approaches for real time data flow where decision is required to be made fast. This work explores different aspects of ML and DL application in healthcare disease data. Designing of more robust models for disease prediction will lead to better utilization of resources and will help in proper analysis of data. This understanding will provide a flow of the current research in a better proportion. So, this work can be expanded to predict any disease where feature selection is important for fast detection.

The CNN based approaches like evidence-based CNN classifier, and concatenation of CNN can be examined for different types of image data the performances.

For life threating disease detection, image segmentation approach can be expanded in more detailed version for detailed and fast prediction.

The future scope of this work is wide as prediction of any health abnormality will always benefit mankind. This may benefit a person or community as a whole.

Bibliography

[1] Jiang F, Jiang Y, Zhi H, Dong Y, Li H, Ma S, Wang Y, Dong Q, Shen H, Wang Y. Artificial intelligence in healthcare: past, present and future. Stroke Vasc Neurol. 2017 Jun 21;2(4):230-243. doi: 10.1136/svn-2017-000101. PMID: 29507784; PMCID: PMC5829945.

[2] Tekkeşin Aİ. Artificial Intelligence in Healthcare: Past, Present and Future. Anatol J Cardiol. 2019 Oct;22(Suppl 2):8-9. doi: 10.14744/AnatolJCardiol.2019.28661. PMID: 31670713.

[3] Häyrinen K, Saranto K, Nykänen P. Definition, structure, content, use and impacts of electronic health records: a review of the research literature. Int J Med Inform. 2008 May;77(5):291-304. doi: 10.1016/j.ijmedinf.2007.09.001. Epub 2007 Oct 22. PMID: 17951106.

[4] Buntin MB, Burke MF, Hoaglin MC, Blumenthal D. The benefits of health information technology: a review of the recent literature shows predominantly positive results. Health Aff (Millwood). 2011 Mar;30(3):464-71. doi: 10.1377/hlthaff.2011.0178. PMID: 21383365.

[5] Hoffman, Robert R., Johnson, Matthew, Bradshaw, Jeffrey M., Underbrink, Al, 2013. Trust in automation. IEEE Intell. Syst. 28 (1), 84–88.

[6] Poornima,S., Pushpalatha,M.,A survey on various applications of prescriptive analytics,International Journal of Intelligent Networks,Volume 1,2020,Pages 76-84,ISSN 2666-6030,https://doi.org/10.1016/j.ijin.2020.07.001.

[7] Bose R. Competitive intelligence process and tools for intelligence analysis. Ind Manag Data Syst.2008;108(4):510–28.

[8] Tsai, CW., Lai, CF., Chao, HC. *et al*. Big data analytics: a survey. *Journal of Big Data* **2**, 21 (2015). https://doi.org/10.1186/s40537-015-0030-3

[9] Davenport TH. Big data at work: dispelling the myths, uncovering the opportunities. Boston: Harvard Business School Publishing; 2014.

[10] https://healthitanalytics.com/news/mismatched-symptoms-call-ehr-data-integrity-into-question

[11] https://www.techtarget.com/searchhealthit/

[12] https://www.techtarget.com/searchhealthit/https://www.techtarget.com/searchhealthit/

[13] Lepenioti,K., Bousdekis, B., Apostolou, D., Mentzas,G.,Prescriptive analytics: Literature review and research challenges, International Journal of Information Management,Volume 50,2020,Pages 57-70,ISSN 0268-4012,https://doi.org/10.1016/j.ijinfomgt.2019.04.003.

[14] Ridzuan,F., Zainon,W.,Diagnostic analysis for outlier detection in big data analytics, Procedia Computer Science, Volume 197, 2022, Pages 685-692, ISSN 1877-0509,https://doi.org/10.1016/j.procs.2021.12.189.

[15] Hussain, S., and Lee, S., "Visualization and descriptive analytics of wellness data through Big Data," 2015 Tenth International Conference on Digital Information Management (ICDIM), Jeju, Korea (South), 2015, pp. 69-71, doi: 10.1109/ICDIM.2015.7381878.

[16] Hariri, R.H., Fredericks, E.M. & Bowers, K.M. Uncertainty in big data analytics: survey, opportunities, and challenges. *J Big Data* **6**, 44 (2019). https://doi.org/10.1186/s40537-019-0206-3

[17] Batko, K., Ślęzak, A. The use of Big Data Analytics in healthcare. *J Big Data* **9**, 3 (2022). https://doi.org/10.1186/s40537-021-00553-4

[18] https://www.statista.com/statistics/263766/total-population-of-india/

[19] Bogoch II, Watts A, Thomas-Bachli A, Huber C, Kraemer MUG, Khan K. Pneumonia of unknown aetiology in Wuhan, China: potential for international spread via commercial air travel. J Travel Med. 2020 Mar 13;27(2):taaa008. doi: 10.1093/jtm/taaa008. PMID: 31943059; PMCID: PMC7107534.

[20] R. Alugubelli, Exploratory Study of Artificial Intelligence in Healthcare, Int. J. Innovat. Eng. Res. Technol. 3 (1) (2016) 1.

[21] Musen, M.A., Middleton, B., Greenes, R.A. (2021). Clinical Decision-Support Systems. In: Shortliffe, E.H., Cimino, J.J. (eds) Biomedical Informatics. Springer, Cham. https://doi.org/10.1007/978-3-030-58721-5_24

[22] Kamdar, J.H., Jeba Praba, J., Georrge, J.J. (2020). Artificial Intelligence in Medical Diagnosis: Methods, Algorithms and Applications. In: Jain, V., Chatterjee, J. (eds) Machine Learning with Health Care Perspective. Learning and Analytics in Intelligent Systems, vol 13. Springer, Cham. https://doi.org/10.1007/978-3-030-40850-3_2.

[23] Bisaso KR, Anguzu GT, Karungi SA, Kiragga A, Castelnuovo B. A survey of machine learning applications in HIV clinical research and care. Comput Biol Med. 2017 Dec 1;91:366-371. doi: 10.1016/j.compbiomed.2017.11.001. Epub 2017 Nov 9. PMID: 29127902.

[24] Healthcare Artificial Intelligence Market- Global Forecast to 2027. https:// www.researchandmarkets.com/reports/4753853/healthcare-artificialintelligence-market-by.

[25] Qureshi, K.N.; Din, S.; Jeon, G.; Piccialli, F. An accurate and dynamic predictive model for a smart M-Health system using machine learning. Inf. Sci. 2020, 538, 486–502

[26] Shailaja, K.M., Seetharamulu, B., & Jabbar, M.A. (2018). Machine Learning in Healthcare: A Review. *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 910-914.

[27] Kang, J.J. Systematic analysis of security implementation for internet of health things in mobile health networks. In Data Science in Cybersecurity and Cyberthreat Intelligence; Springer: Berlin/Heidelberg, Germany, 2020; pp. 87–113.

[28] https://en.wikipedia.org/wiki/Machine_learning

[29] Das, S.; Dey, A.; Pal, A.; Roy, N. Applications of artificial intelligence in machine learning: Review and prospect. Int. J. Comput. Appl. 2015, 115, 31–41.

[30] Tahmassebi A, Wengert GJ, Helbich TH, Bago-Horvath Z, Alaei S, Bartsch R, Dubsky P, Baltzer P, Clauser P, Kapetas P, Morris EA, Meyer-Baese A, Pinker K. Impact of Machine Learning With Multiparametric Magnetic Resonance Imaging of the Breast for Early Prediction of Response to Neoadjuvant Chemotherapy and Survival Outcomes in Breast Cancer Patients. Invest Radiol. 2019 Feb;54(2):110-117. doi: 10.1097/RLI.0000000000000518. PMID: 30358693; PMCID: PMC6310100.

[31] Kayal, C.K., Bagchi, S., Dhar, D., Maitra, T., Chatterjee, S. (2019). Hepatocellular Carcinoma Survival Prediction Using Deep Neural Network. In: Chakraborty, M., Chakrabarti, S., Balas, V., Mandal, J. (eds) Proceedings of International Ethical Hacking Conference 2018. Advances in Intelligent Systems and Computing, vol 811. Springer, Singapore. https://doi.org/10.1007/978-981-13-1544-2_28

[32] Zheng T, Xie W, Xu L, He X, Zhang Y, You M, Yang G, Chen Y. A machine learning-based framework to identify type 2 diabetes through electronic health records. Int J Med Inform. 2017 Jan;97:120-127. doi: 10.1016/j.ijmedinf.2016.09.014. Epub 2016 Oct 1. PMID: 27919371; PMCID: PMC5144921.

[33] Zacharaki EI, Wang S, Chawla S, Soo Yoo D, Wolf R, Melhem ER, Davatzikos C. Classification of brain tumor type and grade using MRI texture and shape in a machine learning scheme. MagnReson Med. 2009 Dec;62(6):1609-18. doi: 10.1002/mrm.22147. PMID: 19859947; PMCID: PMC2863141.

[34] Wong A, Young AT, Liang AS, Gonzales R, Douglas VC, Hadley D. Development and Validation of an Electronic Health Record-Based Machine Learning Model to Estimate Delirium Risk in Newly Hospitalized Patients Without Known Cognitive Impairment. JAMA Netw Open. 2018 Aug 3;1(4):e181018. doi: 10.1001/jamanetworkopen.2018.1018. PMID: 30646095; PMCID: PMC6324291.

[35] Panahiazar M, Taslimitehrani V, Pereira N, Pathak J. Using EHRs and Machine Learning for Heart Failure Survival Analysis. Stud Health Technol Inform. 2015;216:40-4. PMID: 26262006; PMCID: PMC4905764.

[36] Lonini, L., Dai, A., Shawen, N. *et al.* Wearable sensors for Parkinson’s disease: which data are worth collecting for training symptom detection models. *npj Digital Med* **1**, 64 (2018). https://doi.org/10.1038/s41746-018-0071-z

[37] Goodfellow D, Zhi R, Funke R, Pulido JC, Mataric M, Smith BA. Predicting Infant Motor Development Status using Day Long Movement Data from Wearable Sensors. arXiv preprint. 2018; arXiv:1807.02617

[38] Yacchirema D, de Puga JS, Palau C, Esteve M. Fall detection system for elderly people using IoT and Big Data. Procedia computer science. 2018;130:603-610.

[39] Hassan MM, Huda S, Uddin MZ, Almogren A, Alrubaian M. Human Activity Recognition from Body Sensor Data using Deep Learning. J Med Syst. 2018 Apr 16;42(6):99. doi: 10.1007/s10916-018-0948-z. PMID: 29663090.

[40] Njage PMK, Leekitcharoenphon P, Hald T. Improving hazard characterization in microbial risk assessment using next generation sequencing data and machine learning: Predicting clinical outcomes in shigatoxigenic Escherichia coli. Int J Food Microbiol. 2019 Mar 2;292:72-82. doi: 10.1016/j.ijfoodmicro.2018.11.016. Epub 2018 Dec 4. PMID: 30579059.

[41] Alakwaa FM, Chaudhary K, Garmire LX. Deep Learning Accurately Predicts Estrogen Receptor Status in Breast Cancer Metabolomics Data. J Proteome Res. 2018 Jan 5;17(1):337-347. doi: 10.1021/acs.jproteome.7b00595. Epub 2017 Nov 27. PMID: 29110491; PMCID: PMC5759031.

[42] https://en.wikipedia.org/wiki/Transcriptome

[43] Carly A. Bobak, Alexander J. Titus, Jane E. Hill, Comparison of common machine learning models for classification of tuberculosis using transcriptional biomarkers from integrated datasets,Applied Soft Computing,Volume 74,2019,Pages 264-273,ISSN 1568-4946,https://doi.org/10.1016/j.asoc.2018.10.005.

[44] Srinivasan S, Leshchyk A, Johnson NT, Korkin D. A hybrid deep clustering approach for robust cell type profiling using single-cell RNA-seq data. RNA. 2020 Oct;26(10):1303-1319. doi: 10.1261/rna.074427.119. Epub 2020 Jun 12. PMID: 32532794; PMCID: PMC7491323.

[45] Liang CA, Chen L, Wahed A, Nguyen AND. Proteomics Analysis of FLT3-ITD Mutation in Acute Myeloid Leukemia Using Deep Learning Neural Network. Ann Clin Lab Sci. 2019 Jan;49(1):119-126. PMID: 30814087.

[46] Panda, B., Majhi, B. A novel improved prediction of protein structural class using deep recurrent neural network. *Evol. Intel.* **14**, 253–260 (2021). https://doi.org/10.1007/s12065-018-0171-3

[47] Nagarajan, S.M., Muthukumaran, V., Murugesan, R. *et al.* Innovative feature selection and classification model for heart disease prediction. *J Reliable Intell Environ* **8**, 333–343 (2022). https://doi.org/10.1007/s40860-021-00152-3

[48] Li, B.; Zhang, P.l.; Tian, H.; Mi, S.S.; Liu, D.S.; Ren, G.Q. A new feature extraction and selection scheme for hybrid fault diagnosis of gearbox, Expert Systems with Applications, Volume 38, Issue 8, 2011, Pages 10000-10009, ISSN 0957-4174, https://doi.org/10.1016/j.eswa.2011.02.008.

[49] https://en.wikipedia.org/wiki/Principal_component_analysis

[50] Song, F., Guo, Z., & Mei, D. (2010). Feature selection using principal component analysis. In *Proceedings of the 2010 International Conference on System Science, Engineering Design and*

Manufacturing Informatization (pp. 27-30). Yichang, China. https://doi.org/10.1109/ICSEM.2010.14.

[51] Wen, J. (2019). Robust sparse linear discriminant analysis. *IEEE Transactions on Circuits and Systems for Video Technology, 29*(2), 390-403. https://doi.org/10.1109/TCSVT.2018.2799214

[52] Dornaika, F., Khoder, A., Moujahid, A. *et al.* A supervised discriminant data representation: application to pattern classification. *Neural Comput&Applic* **34**, 16879–16895 (2022). https://doi.org/10.1007/s00521-022-07332-z

[53] Luo, X.; Li, X.; Wang, Z.; Liang, J. Discriminant autoencoder for feature extraction in fault diagnosis. Chemometrics and Intelligent Laboratory Systems, Volume 192, 2019, 103814, ISSN 0169-7439, https://doi.org/10.1016/j.chemolab.2019.103814.

[54] Li, M.A.; Luo, X.Y.; Yang, J.F. Extracting the nonlinear features of motor imagery EEG using parametric t-SNE. Neurocomputing, Volume 218, 2016, Pages 371-381, ISSN 0925-2312, https://doi.org/10.1016/j.neucom.2016.08.083.

[55] Bhavsar, H.; Ganatra, A. A comparative study of training algorithms for supervised machine learning. Int. J. Soft Comput. Eng. (IJSCE) 2012, 2, 2231–2307

[56] Zhou X, Li X, Zhang Z, Han Q, Deng H, Jiang Y, Tang C, Yang L. Support vector machine deep mining of electronic medical records to predict the prognosis of severe acute myocardial infarction. Front Physiol. 2022 Sep 29;13:991990. doi: 10.3389/fphys.2022.991990. PMID: 36246101; PMCID: PMC9558165.

[57] Ali, L., Niamat, A., Khan, J.A., Golilarz, N.A., Xingzhong, X., Noor, A., Nour, R., & Bukhari, S.A. (2019). An Optimized Stacked Support Vector Machines Based Expert System for the Effective Prediction of Heart Failure. *IEEE Access*, *7*, 54007-54014.

[58] Nilashi, M., Ahmadi, H., Manaf, A.A. *et al.* Coronary Heart Disease Diagnosis Through Self-Organizing Map and Fuzzy Support Vector Machine with Incremental Updates. *Int. J. Fuzzy Syst.* **22**, 1376–1388 (2020). https://doi.org/10.1007/s40815-020-00828-7

[59] Rokach, L., Maimon, O. (2005). Decision Trees. In: Maimon, O., Rokach, L. (eds) Data Mining and Knowledge Discovery Handbook. Springer, Boston, MA. https://doi.org/10.1007/0-387-25465-X_9

[60] Ray, S. (2019). A Quick Review of Machine Learning Algorithms. *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, 35-39.

[61] Chern, CC., Chen, YJ. & Hsiao, B. Decision tree–based classifier in providing telehealth service. *BMC Med Inform DecisMak* **19**, 104 (2019). https://doi.org/10.1186/s12911-019-0825-9

[62] Pathak, A.K., Arul Valan, J. (2020). A Predictive Model for Heart Disease Diagnosis Using Fuzzy Logic and Decision Tree. In: Elçi, A., Sa, P., Modi, C., Olague, G., Sahoo, M., Bakshi, S. (eds) Smart Computing Paradigms: New Progresses and Challenges. Advances in Intelligent Systems and Computing, vol 767. Springer, Singapore. https://doi.org/10.1007/978-981-13-9680-9_10

[63] Mohamed, W.N.H.W.; Salleh, M.N.M.; Omar, A.H. A comparative study of reduced error pruning method in decision tree algorithms. In Proceedings of the 2012 IEEE International Conference on Control System, Computing and Engineering, Penang, Malaysia, 23–25 November 2012; pp. 392–397.

[64] Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons*, *4*, 51-62.

[65] Dulhare, U.N. Prediction system for heart disease using Naive Bayes and particle swarm optimization. Biomed. Res. 2018, 29, 2646–2649

[66] Balaha, H.M., Hassan, A.ES. Skin cancer diagnosis based on deep transfer learning and sparrow search algorithm. *Neural Comput&Applic* **35**, 815–853 (2023). https://doi.org/10.1007/s00521-022-07762-9

[67] Gupta, A., Kumar, L., Jain, R., Nagrath, P. (2020). Heart Disease Prediction Using Classification (Naive Bayes). In: Singh, P., Pawłowski, W., Tanwar, S., Kumar, N., Rodrigues, J., Obaidat, M. (eds) Proceedings of First International Conference on Computing, Communications, and Cyber-Security (IC4S 2019). Lecture Notes in Networks and Systems, vol 121. Springer, Singapore. https://doi.org/10.1007/978-981-15-3369-3_42

[68] A. N. Repaka, S. D. Ravikanti and R. G. Franklin, "Design And Implementing Heart Disease Prediction Using Naives Bayesian," 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), Tirunelveli, India, 2019, pp. 292-297, doi: 10.1109/ICOEI.2019.8862604.

[69] Duneja, A.; Puyalnithi, T. Enhancing classification accuracy of k-nearest neighbours algorithm using gain ratio. Int. Res. J. Eng. Technol 2017, 4, 1385–1388.

[70] Chen, Z.; Zhou, L.J.; Da Li, X.; Zhang, J.N.; Huo, W.J. The Lao text classification method based on KNN. Procedia Comput. Sci. 2020, 166, 523–528.

[71] M. Akhil jabbar, B.L. Deekshatulu, Priti Chandra, Classification of Heart Disease Using K-Nearest Neighbor and Genetic Algorithm, Procedia Technology, Volume 10, 2013, Pages 85-94, ISSN 2212-0173, https://doi.org/10.1016/j.protcy.2013.12.340.

[72] Deekshatulu, B.; Chandra, P. Classification of heart disease using k-nearest neighbor and genetic algorithm. Procedia Technol. 2013, 10, 85–94

[73] Petrou S, Murray L, Cooper P, Davidson LL. The accuracy of self-reported healthcare resource utilization in health economic studies. Int J Technol Assess Health Care. 2002 Summer;18(3):705-10. doi: 10.1017/s026646230200051x. PMID: 12391960.

[74] Lemon SC, Roy J, Clark MA, Friedmann PD, Rakowski W. Classification and regression tree analysis in public health: methodological review and comparison with logistic regression. Ann Behav Med. 2003 Dec;26(3):172-81. doi: 10.1207/S15324796ABM2603_02. PMID: 14644693.

[75] Ganie, S. M., & Malik, M. B. (2022). An ensemble machine learning approach for predicting type-II diabetes mellitus based on lifestyle indicators. *Healthcare Analytics, 2*, 100092. https://doi.org/10.1016/j.health.2022.100092

[76] Nguyen DK, Lan CH, Chan CL. Deep Ensemble Learning Approaches in Healthcare to Enhance the Prediction and Diagnosing Performance: The Workflows, Deployments, and Surveys on the Statistical, Image-Based, and Sequential Datasets. Int J Environ Res Public Health. 2021 Oct 14;18(20):10811. doi: 10.3390/ijerph182010811. PMID: 34682554; PMCID: PMC8536161.

[77] Kaushik, S., Choudhury, A., Sheron, P. K., Dasgupta, N., Natarajan, S., Pickett, L. A., & Dutt, V. (2020). AI in healthcare: Time-series forecasting using statistical, neural, and ensemble architectures. *Frontiers in Big Data, 3*, Article 4. https://doi.org/10.3389/fdata.2020.00004

[78] Basak, D., Pal, S., & Patranabis, D. (2007). Support vector regression. *Neural Information Processing – Letters and Reviews, 11*, 1-10.

[79] Rivas-Perea, P., Cota-Ruiz, J., Chaparro, D., Venzor, J., Carreón, A., & Rosiles, J. (2013). Support vector machines for regression: A succinct review of large-scale and linear programming formulations. *International Journal of Intelligence Science, 3*(1), 5-14. https://doi.org/10.4236/ijis.2013.31002

[80] Sharin, S. N., Radzali, M. K., & Sani, M. S. A. (2022). A network analysis and support vector regression approaches for visualizing and predicting the COVID-19 outbreak in Malaysia. *Healthcare Analytics, 2*, Article 100080. https://doi.org/10.1016/j.health.2022.100080

[81] Claris, S., & Sigauke, C. (2023). Short-term forecasting of COVID-19 using support vector regression: An application using Zimbabwean data. *American Journal of Infection Control.* https://doi.org/10.1016/j.ajic.2023.03.010

[82] Kalantari, A., Kamsin, A., Shamshirband, S., Gani, A., Alinejad-Rokny, H., & Chronopoulos, A. T. (2018). Computational intelligence approaches for classification of medical data: State-of-the-art, future challenges and research directions. *Neurocomputing, 276,* 2–22. https://doi.org/10.1016/j.neucom.2017.01.126.

[83] DeepMind. DeepMind Technologies Limited, 2020.

[84] Watson, Watson IBM, 2020, p. 1-16.

[85] Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing, 7*(3-4), 197-387.

[86] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[87] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770-778).

[88] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., & others. (2015). Going deeper with convolutions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1-9).

[89] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2261-2269). https://doi.org/10.1109/CVPR.2017.243

[90] Howard, A., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). MobileNets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

[91] Arooj S, Rehman SU, Imran A, Almuhaimeed A, Alzahrani AK, Alzahrani A. A Deep Convolutional Neural Network for the Early Detection of Heart Disease. Biomedicines. 2022 Nov 3;10(11):2796. doi: 10.3390/biomedicines10112796. PMID: 36359317; PMCID: PMC9687844.

[92] Mehmood, A., Iqbal, M., Mehmood, Z. *et al.* Prediction of Heart Disease Using Deep Convolutional Neural Networks. *Arab J Sci Eng* **46**, 3409–3422 (2021). https://doi.org/10.1007/s13369-020-05105-1

[93] Sharma, A, Pal, T, Jaiswal, V, Chapter 12 - Heart disease prediction using convolutional neural network, Editor(s): Ayman S. El-Baz, Jasjit S. Suri, Cardiovascular and Coronary Artery Imaging, Academic Press, 2022, Pages 245-272, ISBN 9780128227060, https://doi.org/10.1016/B978-0-12-822706-0.00012-3.

[94] Sudha, V.K., Kumar, D. Hybrid CNN and LSTM Network For Heart Disease Prediction. *SN COMPUT. SCI.* **4**, 172 (2023). https://doi.org/10.1007/s42979-022-01598-9

[95] Shankar, V., Kumar, V., Devagade, U. *et al.* Heart Disease Prediction Using CNN Algorithm. *SN COMPUT. SCI.* **1**, 170 (2020). https://doi.org/10.1007/s42979-020-0097-6

[96] García-Ordás, M.T., Bayón-Gutiérrez, M., Benavides, C. *et al.* Heart disease risk prediction using deep learning techniques with feature augmentation. *Multimed Tools Appl* **82**, 31759–31773 (2023). https://doi.org/10.1007/s11042-023-14817-z

[97] Sharma, D., Kumar, R., & Jain, A. (2022). Breast cancer prediction based on neural networks and extra tree classifier using feature ensemble learning. *Measurement: Sensors, 24,* Article 100560. https://doi.org/10.1016/j.measen.2022.100560

[98] Gupta, K., & Chawla, N. (2020). Analysis of histopathological images for prediction of breast cancer using traditional classifiers with pre-trained CNN. *Procedia Computer Science, 167,* 878-889. https://doi.org/10.1016/j.procs.2020.03.427

[99] Choi, H., Ha, S., Im, H. J., Paek, S. H., & Lee, D. S. (2017). Refining diagnosis of Parkinson's disease with deep learning-based interpretation of dopamine transporter imaging. *NeuroImage: Clinical, 16,* 586–594. doi: 10.1016/j.nicl.2017.09.010.

[100] Zhang, Y., Liu, Y., Sun, P., Yan, H., Zhao, X., Zhang, L. (2020). IFCNN: A general image fusion framework based on convolutional neural network. Information Fusion, 54, 99–118. DOI 10.1016/j.inffus.2019.07.011.

[101] Sharma, H., Jain, J. S., Bansal, P., Gupta, S. (2020). Feature extraction and classification of chest X-ray images using cnn to detect pneumonia. 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 227–231. Noida, India, IEEE.

[102] Yeh, R. A., Chen, C., Yian Lim, T., Schwing, A. G., Hasegawa-Johnson, M. et al. (2017). Semantic image inpainting with deep generative models. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 5485–5493. Honolulu, HI, USA.

[103] Zhou, Z., Wang, Y., Guo, Y., Qi, Y., Yu, J. (2019). Image quality improvement of hand-held ultrasound devices with a two-stage generative adversarial network. IEEE Transactions on Biomedical Engineering, 67(1), 298–311. DOI 10.1109/TBME.2019.2912986.

[104] Huang, J., Le, Z., Ma, Y., Fan, F., Zhang, H. et al. (2020). MGMDcGAN: Medical image fusion using multigenerator multi-discriminator conditional generative adversarial network. IEEE Access, 8, 55145–55157. DOI 10.1109/ACCESS.2020.2982016.

[105] Bing, X., Zhang, W., Zheng, L., Zhang, Y. (2019). Medical image super resolution using improved generative adversarial networks. IEEE Access, 7, 145030–145038. DOI 10.1109/ACCESS.2019.2944862.

[106] Li, Y., Zhao, J., Lv, Z., Li, J. (2021). Medical image fusion method by deep learning. International Journal of Cognitive Computing in Engineering, 2, 21–29. DOI 10.1016/j.ijcce.2020.12.004.

[107] Yi, X., Walia, E., Babyn, P. (2019). Generative adversarial network in medical imaging: A review. Medical Image Analysis, 58, 101552. DOI 10.1016/j.media.2019.101552.

[108] Hong, H., Li, X., Wang, M. (2019). Gane: A generative adversarial network embedding. IEEE Transactions on Neural Networks and Learning Systems, 31(7), 2325–2335. DOI 10.1109/TNNLS.2019.2921841.

[109] Wang, C., Xu, C., Yao, X., Tao, D. (2019). Evolutionary generative adversarial networks. IEEE Transactions on Evolutionary Computation, 23(6), 921–934. DOI 10.1109/TEVC.2019.2895748.

[110] Kim, J. C., Chung, K. (2020). Multi-modal stacked denoising autoencoder for handling missing data in healthcare big data. IEEE Access, 8, 104933–104943. DOI 10.1109/Access.6287639.

[111] Chen, M., Shi, X., Zhang, Y., Wu, D., Guizani, M. (2017). Deep features learning for medical image analysis with convolutional autoencoder neural network. IEEE Transactions on Big Data, 1, 1. DOI 10.1109/TBDATA.2017.2717439.

[112] Azarang, A., Manoochehri, H. E., Kehtarnavaz, N. (2019). Convolutional autoencoder-based multispectral image fusion. IEEE Access, 7, 35673–35683. DOI 10.1109/ACCESS.2019.2905511.

[113] Zhang, J., Lin, F., Xiong, P., Du, H., Zhang, H. et al. (2019). Automated detection and localization of myocardial infarction with staked sparse autoencoder and treebagger. IEEE Access, 7, 70634–70642. DOI 10.1109/ACCESS.2019.2919068.

[114] Wei, R., Mahmood, A. (2021). Recent advances in variational autoencoders with representation learning for biomedical informatics: A survey. IEEE Access, 9, 4939–4956. DOI 10.1109/ACCESS.2020.3048309.

[115] Dua, M., Makhija, D., Manasa, P. Y. L., Mishra, P. (2020). A CNN–RNN–LSTM based amalgamation for Alzheimer's disease detection. Journal of Medical and Biological Engineering, 40(5), 688–706. DOI 10.1007/s40846-020-00556-1.

[116] Khan, S., Islam, N., Jan, Z., Din, I. U., Rodrigues, J. J. C. (2019). A novel deep learning-based framework for the detection and classification of breast cancer using transfer learning. Pattern Recognition Letters, 125, 1–6. DOI 10.1016/j.patrec.2019.03.022.

[117] Talo, M., Baloglu, U. B., Yıldırım, Ö., Acharya, U. R. (2019). Application of deep transfer learning for automated brain abnormality classification using MR images. Cognitive Systems Research, 54, 176–188. DOI 10.1016/j.cogsys.2018.12.007.

[118] Hu, Z., Tang, J., Wang, Z., Zhang, K., Zhang, L. et al. (2018). Deep learning for image-based cancer detection and diagnosis−A survey. Pattern Recognition, 83, 134–149. DOI 10.1016/j.patcog.2018.05.014.

[119] Ayon, S. I., & Islam, M. M. (2019). Diabetes prediction: A deep learning approach. *International Journal of Information Engineering and Electronic Business, 11*(2), 21-27. https://doi.org/10.5815/ijieeb.2019.02.03

[120] Manokaran J, Zabihollahy F, Hamilton-Wright A, Ukwatta E. Detection of COVID-19 from chest x-ray images using transfer learning. J Med Imaging (Bellingham). 2021 Jan;8(Suppl 1):017503. doi: 10.1117/1.JMI.8.S1.017503. Epub 2021 Aug 23. PMID: 34435075; PMCID: PMC8382139.

[121] Navaneeth, B., Suchetha, M. (2020). A dynamic pooling based convolutional neural network approach to detect chronic kidney disease. Biomedical Signal Processing and Control, 62, 102068. DOI 10.1016/j.bspc.2020.102068.

[122] Bhaskar, N., Manikandan, S. (2019). A deep-learning-based system for automated sensing of chronic kidney disease. IEEE Sensors Letters, 3(10), 1–4. DOI 10.1109/LSENS.2019.2942145.

[123] Ma, F., Sun, T., Liu, L., Jing, H. (2020). Detection and diagnosis of chronic kidney disease using deep learning-based heterogeneous modified artificial neural network. Future Generation Computer Systems, 111, 17–26. DOI 10.1016/j.future.2020.04.036.

[124] Bisong, E. (2019). Autoencoders. In: Building Machine Learning and Deep Learning Models on Google Cloud Platform. Apress, Berkeley, CA. https://doi.org/10.1007/978-1-4842-4470-8_37

[125] Hongmei, W., Pengzhong, L. Image recognition based on improved convolutional deep belief network model. *Multimed Tools Appl* **80**, 2031–2045 (2021). https://doi.org/10.1007/s11042-019-08300-x

[126] Yang, Y. (2020). Medical multimedia big data analysis modeling based on DBN algorithm. *IEEE Access, 8,* 16350-16361. https://doi.org/10.1109/ACCESS.2020.2967075

[127] DeMaagd, G., & Philip, A. (2015). Parkinson's disease and its management: Part 1: Disease entity, risk factors, pathophysiology, clinical presentation, and diagnosis. *P&T, 40*(8), 504-532.

[128] Hilker, R., Schweitzer, K., Coburger, S., Ghaemi, M., Weisenbach, S., Jacobs, A. H., Rudolf, J., Herholz, K., & Heiss, W. D. (2005). Nonlinear progression of Parkinson's disease as determined by serial positron emission tomographic imaging of striatal fluorodopa F-18 activity. *Archives of Neurology, 62*(3), 378–382.

[129] Wüllner, U., Pakzaban, P., Brownell, A.-L., Hantraye, P., Burns, L., Shoup, T., Elmaleh, D., Petto, A. J., Spealman, R. D., & Brownell, G. L. (1994). Dopamine terminal loss and onset of motor symptoms in MPTP-treated monkeys: A positron emission tomography study with 11C-CFT. *Experimental Neurology, 126*(2), 305–309.

[130] Murueta-Goyena, A., Andikoetxea, A., Gómez-Esteban, J. C., & Gabilondo, I. (2019). Contribution of the GABAergic system to non-motor manifestations in premotor and early stages of Parkinson's disease. *Frontiers in Pharmacology, 10,* Article 1294. https://doi.org/10.3389/fphar.2019.01294

[131] Poewe W (2008) Non-motor symptoms in Parkinson's disease. Eur J Neurol 15:14–20

[132] Moscovich, M., Boschetti, G., Moro, A., Teive, H. A. G., Hassan, A., & Munhoz, R. P. (2017). Death certificate data and causes of death in patients with parkinsonism. *Parkinsonism & Related Disorders, 41,* 99–103. doi: 10.1016/j.parkreldis.2017.05.022.

[133] Sakar, B. E., Isenkul, M. E., Sakar, C. O., Sertbas, A., Gurgen, F., Delil, S., Apaydin, H., & Kursun, O. (2013). Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics, 17*(4), 828–834. doi: 10.1109/JBHI.2013.2245674.

[134] Drotiar, P., Mekyska, J., Rektorova, I., Masarova, L., Smekal, Z., & Faundez-Zanuy, M. (2014). Decision support framework for Parkinson's disease based on novel handwriting markers. *IEEE Transactions on Neural Systems and Rehabilitation Engineering, 23*(3), 508–516. doi: 10.1109/TNSRE.2014.2359997.

[135] Nunes, A., Silva, G., Duque, C., Januário, C., Santana, I., Ambrosio, A. F., Castelo-Branco, M., & Bernardes, R. (2019). Retinal texture biomarkers may help to discriminate between Alzheimer's, Parkinson's, and healthy controls. *PLOS ONE, 14*(6), Article e0218826. https://doi.org/10.1371/journal.pone.0218826

[136] Nuvoli, S., Spanu, A., Fravolini, M. L., Bianconi, F., Cascianelli, S., Madeddu, G., & Palumbo, B. (2020). [123I] metaiodobenzylguanidine (MIBG) cardiac scintigraphy and automated classification techniques in Parkinsonian disorders. *Molecular Imaging and Biology, 22*(3), 703–710. doi: 10.1007/s11307-019-01406-6.

[137] Mei, J., Desrosiers, C., & Frasnelli, J. (2021). Machine learning for the diagnosis of Parkinson's disease: A review of literature. *Frontiers in Aging Neuroscience, 13,* Article 184.

[138] Mabrouk, R., Chikhaoui, B., & Bentabet, L. (2018). Machine learning based classification using clinical and DATSCAN spect imaging features: A study on Parkinson's disease and SWEDD. *IEEE Transactions on Radiation and Plasma Medical Sciences, 3*(2), 170–177.

[139] Prashanth, R., Roy, S. D., Mandal, P. K., & Ghosh, S. (2016). High-accuracy detection of early Parkinson's disease through multimodal features and machine learning. *International Journal of Medical Informatics, 90,* 13–21.

[140] Zhang, X., Chou, J., Liang, J. *et al.* Data-Driven Subtyping of Parkinson's Disease Using Longitudinal Clinical Records: A Cohort Study. *Sci Rep* **9**, 797 (2019). https://doi.org/10.1038/s41598-018-37545-z

[141] Gu S-C, Zhou J, Yuan C-X, Ye Q (2020) Personalized prediction of depression in patients with newly diagnosed Parkinson's disease: a prospective cohort study. J Affect Disord 268:118–126. https://doi.org/10.1016/j.jad.2020.02.046

[142] Arman˜anzas R, Bielza C, Chaudhuri KR, Martinez-Martin P, Larran˜aga P (2013) Unveiling relevant non-motor Parkinson's disease severity symptoms using a machine learning approach. ArtifIntell Med 58(3):195–202. https://doi.org/10.1016/j.artmed.2013.04.002

[143] Little,Max. (2008). Parkinsons. UCI Machine Learning Repository. https://doi.org/10.24432/C59C74.

[144] Pereira, C.R., Weber, S.A., Hook, C., de Rosa, G., & Papa, J.P. (2016). Deep Learning-Aided Parkinson's Disease Diagnosis from Handwritten Dynamics. *2016 29th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)*, 340-346.

[145] Kotsiantis, Sotiris &Kanellopoulos, D. &Pintelas, P.. (2005). Handling imbalanced datasets: A review. GESTS International Transactions on Computer Science and Engineering. 30. 25-36.

[146] Chawla, N. V., Bowyer, K. W., Hall, L. O., &Kegelmeyer, W. P. (2002). Smote: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

[147] Elreedy, D., Atiya, A.F. &Kamalov, F. A theoretical distribution analysis of synthetic minority oversampling technique (SMOTE) for imbalanced learning. *Mach Learn* (2023). https://doi.org/10.1007/s10994-022-06296-4

[148] Wang, S., Dai, Y., Shen, J. *et al.* Research on expansion and classification of imbalanced data based on SMOTE algorithm. *Sci Rep* **11**, 24039 (2021). https://doi.org/10.1038/s41598-021-03430-5

[149] Goel, G., Maguire, L., Li, Y., McLoone, S. (2013). Evaluation of Sampling Methods for Learning from Imbalanced Data. In: Huang, DS., Bevilacqua, V., Figueroa, J.C., Premaratne, P. (eds) Intelligent Computing Theories. ICIC 2013. Lecture Notes in Computer Science, vol 7995. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-39479-9_47

[150] Mukaka, M. M. (2012). Statistics corner: A guide to appropriate use of correlation coefficient in medical research. *Malawi Medical Journal, 24*(3), 69–71.

[151] Pearson's Correlation Coefficient. In: Kirch, W. (eds) Encyclopedia of Public Health. Springer, Dordrecht. https://doi.org/10.1007/978-1-4020-5614-7_2569

[152] Ali, L., Zhu, C., Zhang, Z., & Liu, Y. (2019). Automated detection of Parkinson's disease based on multiple types of sustained phonations using linear discriminant analysis and genetically optimized neural network. *IEEE Journal of Translational Engineering in Health and Medicine, 7,* 1–10. https://doi.org/10.1109/JTEHM.2019.2940900

[153] de Souza, J. W. M., Alves, S. S. A., Rebouças, E. d. S., Almeida, J. S., & Rebouças Filho, P. P. (2018). A new approach to diagnose Parkinson's disease using a structural co-occurrence matrix for a similarity analysis. *Computational Intelligence and Neuroscience, 2018,* Article 7613282. https://doi.org/10.1155/2018/7613282

[154] Sheibani, R., Nikookar, E., and Alavi, S. E. (2019). An ensemble method for diagnosis of Parkinson's disease based on voice measurements. J. Med. Signals Sens. 9, 221–226. doi: 10.4103/jmss.JMSS_57_18

[155] Yaman, O., Ertam, F., and Tuncer, T. (2020). Automated Parkinson's disease recognition based on statistical pooling method using acoustic features. Med. Hypoth. 135:109483. doi: 10.1016/j.mehy.2019.109483

[156] KarapinarSenturk, Z. (2020). Early diagnosis of Parkinson's disease using machine learning algorithms. Med. Hypoth. 138:109603.doi: 10.1016/j.mehy.2020.109603

[157] Gunduz, H. (2019). Deep learning-based parkinson's disease classification using vocal feature sets. IEEE Access 7, 115540–115551.doi: 10.1109/ACCESS.2019.2936564

[158] Sayaydeha, O. N. A., and Mohammad, M. F. (2019). "Diagnosis of the Parkinson disease using enhanced fuzzy min-max neural network and OneR attribute evaluation method," in 2019 International Conferenc on Advanced Science and Engineering (ICOASE) (Zakho-Duhok), 64–69. doi: 10.1109/ICOASE.2019.8723870

[159] Ribeiro, L. C. F., Afonso, L. C. S., and Papa, J. P. (2019). Bag of samplings for computer-assisted Parkinson's disease diagnosis based on recurrent neural networks. Comp. Biol. Med. 115:103477. doi: 10.1016/j.compbiomed.2019.103477

[160] Anand, A., Haque, M. A., Alex, J. S. R., and Venkatesan, N. (2018). “Evaluation of machine learning and deep learning algorithms combined with dimentionality reduction techniques for classification of Parkinson’s disease,” in 2018 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT) (Louisville, KY), 342–347. doi: 10.1109/ISSPIT.2018.8642776

[161] Marar, S., Swain, D., Hiwarkar, V., Motwani, N., and Awari, A. (2018).“Predicting the occurrence of Parkinson’s disease using various classification models,” in 2018 International Conference on Advanced Computation and Telecommunication (ICACAT) (Bhopal), 1–5. doi: 10.1109/ICACAT.2018.8933579

[162] Shetty, N,. Patient Safety – Are We Doing Enough?. Archives of Medicine and Health Sciences 10(2):p 157-159, Jul–Dec 2022. | DOI: 10.4103/amhs.amhs_265_22

[163] Ahsan, M.M.; Siddique, Z. Machine Learning-Based Heart Disease Diagnosis: A Systematic Literature Review. arXiv 2021, arXiv:2112.06459

[164] Ahsan, M.M.; E Alam, T.; Trafalis, T.; Huebner, P. Deep MLP-CNN model using mixed-data to distinguish between COVID-19 and Non-COVID-19 patients. Symmetry 2020, 12, 1526.

[165] Ahsan, M.M.; Gupta, K.D.; Islam, M.M.; Sen, S.; Rahman, M.; Shakhawat Hossain, M. COVID-19 symptoms detection based on nasnetmobile with explainable ai using various imaging modalities. Mach. Learn. Knowl. Extr. 2020, 2, 490–504.

[166] https://wwwp.fc.unesp.br/~papa/pub/datasets/Handpd/

[167] M. Sharawi, H. M. Zawbaa, E. Emary, H. M. Zawbaa and E. Emary, "Feature selection approach based on whale optimization algorithm," 2017 Ninth International Conference on Advanced Computational Intelligence (ICACI), Doha, Qatar, 2017, pp. 163-168, doi: 10.1109/ICACI.2017.7974502.

[168] Mafarja, M., & Mirjalili, S. A. (2018). Whale optimization approaches for wrapper feature selection. *Applied Soft Computing, 62,* 441–453. https://doi.org/10.1016/j.asoc.2017.11.006

[169] García Ocaña, M. I., López-Linares Román, K., Lete Urzelai, N., González Ballester, M. Á., & Macía Oliver, I. (2020). Medical image detection using deep learning. In Y. W. Chen & L. Jain (Eds.), *Deep learning in healthcare* (Vol. 171, pp. 1–12). Springer. https://doi.org/10.1007/978-3-030-32606-7_1

[170] López-Linares Román, K., García Ocaña, M.I., LeteUrzelai, N., González Ballester, M.Á., Macía Oliver, I. (2020). Medical Image Segmentation Using Deep Learning. In: Chen, YW., Jain, L. (eds) Deep Learning in Healthcare. Intelligent Systems Reference Library, vol 171. Springer, Cham. https://doi.org/10.1007/978-3-030-32606-7_2

[171] LeCun, Y., Bengio, Y. and Hinton, G. (2015) Deep Learning. Nature, 521, 436-444.http://dx.doi.org/10.1038/nature14539

[172] Bengio Y. Learning deep architectures for AI. Found Trends Mach Learn 2009;2:1–127.

[173] GBD Tuberculosis Collaborators (2018) The global burden of tuberculosis: results from the global burden of disease study 2015. Lancet Infect Dis 18(3):261–284

[174] https://pib.gov.in/PressReleaseIframePage.aspx?PRID=1975724#:~:text=The%20report%20notes%20that%20India's,surpassing%20the%20pre%2DCOVID%20levels.

[175] Khutlang R, Krishnan S, Dendere R, Whitelaw A, Veropoulos K, Learmonth G, et al. Classification of Mycobacterium tuberculosis in images of ZN stained sputum smears. IEEE Trans Inf Technol Biomed 2010;14(4):949–957.

[176] Panicker RO, Soman B, Saini G, Rajan J. A review of automatic methods based on image processing techniques for tuberculosis detection from microscopic sputum smear images. J Med Syst 2016;40(1):1–13.

[177] Khutlang R, Krishnan S, Dendere R, Whitelaw A, Veropoulos K, Learmonth G, et al. Classification of Mycobacterium tuberculosis in images of ZN stained sputum smears. IEEE Trans Inf Technol Biomed 2010;14(4):949–957.

[178] Khutlang R, Krishnan S, Whitelaw A, Douglas TS. Automated detection of tuberculosis in Ziehl-Neelsen stained sputum smears using two one-class classifiers. J Microsc 2010;237(1):96–102.

[179] Muyama L., Nakatumba-Nabende J., Mudali D. (2021) Automated Detection of Tuberculosis from Sputum Smear Microscopic Images Using Transfer Learning Techniques. In: Abraham A., Siarry P., Ma K., Kaklauskas A. (eds) Intelligent Systems Design and Applications. ISDA 2019. Advances in Intelligent Systems and Computing, vol 1181. Springer, Cham. https://doi.org/10.1007/978-3-030-49342-4_6

[180] Li, X., Zhou, Y., Du, P., & et al. (2020). A deep learning system that generates quantitative CT reports for diagnosing pulmonary tuberculosis. *Applied Intelligence*. https://doi.org/10.1007/s10489-020-02051-1

[181] Chang, RI., Chiu, YH. & Lin, JW. Two-stage classification of tuberculosis culture diagnosis using convolutional neural network with transfer learning. J Supercomput 76, 8641–8656 (2020). https://doi.org/10.1007/s11227-020-03152-x

[182] Oloko-Oba M., Viriri S. (2020) Tuberculosis Abnormality Detection in Chest X-Rays: A Deep Learning Approach. In: Chmielewski L.J., Kozera R., Orłowski A. (eds) Computer Vision and Graphics. ICCVG 2020. Lecture Notes in Computer Science, vol 12334. Springer, Cham. https://doi.org/10.1007/978-3-030-59006-2_11

[183] Ayaz, M., Shaukat, F. & Raja, G. Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors. Phys Eng Sci Med (2021). https://doi.org/10.1007/s13246-020-00966-0

[184] Yugaswara H., Fathurahman M., Suhaeri (2020) Experimental Analysis of Tuberculosis Classification Based on Clinical Data Using Machine Learning Techniques. In: Ghazali R., Nawi N., Deris M., Abawajy J. (eds) Recent Advances on Soft Computing and Data Mining. SCDM 2020. Advances in Intelligent Systems and Computing, vol 978. Springer, Cham. https://doi.org/10.1007/978-3-030-36056-6_15

[185] Raju M., Aswath A., Kadam A., Pagidimarri V. (2019) Automatic Detection of Tuberculosis Using Deep Learning Methods. In: Laha A. (eds) Advances in Analytics and Applications. Springer Proceedings in Business and Economics. Springer, Singapore. https://doi.org/10.1007/978-981-13-1208-3_11

[186] Ali M., Arshad W. (2020) Prediction of Tuberculosis Using Supervised Learning Techniques Under Pakistani Patients. In: Khanna A., Gupta D., Bhattacharyya S., Snasel V., Platos J., Hassanien A. (eds) International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, vol 1087. Springer, Singapore. https://doi.org/10.1007/978-981-15-1286-5_4

[187] El-Melegy M., Mohamed D., ElMelegy T. (2019) Automatic Detection of Tuberculosis Bacilli from Microscopic Sputum Smear Images Using Faster R-CNN, Transfer Learning and Augmentation. In: Morales A., Fierrez J., Sánchez J., Ribeiro B. (eds) Pattern Recognition and Image Analysis. IbPRIA 2019. Lecture Notes in Computer Science, vol 11867. Springer, Cham. https://doi.org/10.1007/978-3-030-31332-6_24

[188] Hernández A., Panizo Á., Camacho D. (2019) An Ensemble Algorithm Based on Deep Learning for Tuberculosis Classification. In: Yin H., Camacho D., Tino P., Tallón-Ballesteros A., Menezes R., Allmendinger R. (eds) Intelligent Data Engineering and Automated Learning – IDEAL 2019. IDEAL 2019. Lecture Notes in Computer Science, vol 11871. Springer, Cham. https://doi.org/10.1007/978-3-030-33607-3_17

[189] Kant, S., & Srivastava, M. M. (2018). Towards automated tuberculosis detection using deep learning. In *Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI)* (pp. 1250–1253). IEEE. https://doi.org/10.1109/SSCI.2018.8628800

[190] https://doi.org/10.1155/2018/4168538

[191] Karnkawinpong T, Limpiyakorn Y (2018) Chest X-ray analysis of tuberculosis by convolutional neural networks with affine transforms. ACM international conference proceedings series, pp 90–93. https ://doi.org/10.1145/32971 56.32972 51

[192] Sathitratanacheewin S, Pongpirul K (2018) Deep learning for automated classification of tuberculosis-related chest x-ray: dataset specificity limits diagnostic performance generalizability. arXiv preprint arXiv :1811.07985

[193] López, Y. P., Costa Filho, C. F. F., Aguilera, L. M. R., & Costa, M. G. F. (2017). Automatic classification of light field smear microscopy patches using convolutional neural networks for identifying Mycobacterium tuberculosis. In *Proceedings of the 2017 CHILEAN Conference on Electrical, Electronics Engineering, Information and Communication Technologies (CHILECON)* (pp. 1–5). IEEE. https://doi.org/10.1109/CHILECON.2017.8229512

[194] Santosh, K. C., & Antani, S. (2018). Automated chest X-ray screening: Can lung region symmetry help detect pulmonary abnormalities? *IEEE Transactions on Medical Imaging, 37*(5), 1168–1177. https://doi.org/10.1109/TMI.2017.2775636

[195] Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Zhiyun Xue, Palaniappan K, Singh RK, Antani S, Thoma G, Yi-Xiang Wang, Pu-Xuan Lu, McDonald CJ. Automatic tuberculosis screening using chest radiographs. IEEE Trans Med Imaging.2014 Feb;33(2):233-45. doi: 10.1109/TMI.2013.2284099. Epub 2013 Oct 1. PMID: 24108713.

[196] Lakhani P, Sundaram B (2017) Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology 284:574–582. https ://doi.org/10.1148/radio l.20171 62326

[197] Priya E., Srinivasan S., Ramakrishnan S. (2012) Classification of Tuberculosis Digital Images Using Hybrid Evolutionary Extreme Learning Machines. In: Nguyen NT., Hoang K., Jędrzejowicz P. (eds) Computational Collective Intelligence. Technologies and Applications. ICCCI 2012. Lecture Notes in Computer Science, vol 7653. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-34630-9_28

[198] B. van Ginneken, S. Katsuragawa, B. M. terHaarRomeny, Kunio Doi and M. A. Viergever, "Automatic detection of abnormalities in chest radiographs using local texture analysis," in IEEE Transactions on Medical Imaging, vol. 21, no. 2, pp. 139-149, Feb. 2002, doi: 10.1109/42.993132.

[199] https://www.kaggle.com/tawsifurrahman/tuberculosis-tb-chest-xray-dataset

[200] Campos, G., Mastelini, S., Aguiar, G. *et al.* Machine learning hyperparameter selection for Contrast Limited Adaptive Histogram Equalization. *J Image Video Proc.* **2019**, 59 (2019). https://doi.org/10.1186/s13640-019-0445-4

[201] Reza, A.M. Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology* **38**, 35–44 (2004). https://doi.org/10.1023/B:VLSI.0000028532.53893.82

[202] Venkatesh S., John De Britto C., Subhashini P. & Somasundaram K. (2022) Image Enhancement and Implementation of CLAHE Algorithm and Bilinear Interpolation, Cybernetics and Systems, DOI: 10.1080/01969722.2022.2147128

[203] Huang, S., Lee, F., Miao, R. et al. A deep convolutional neural network architecture for interstitial lung disease pattern classification. Med Biol Eng Comput 58, 725–737 (2020).https://doi.org/10.1007/s11517-019-02111-w

[204] Tong Z,XuP,Denœux P, An evidential classifier based on Dempster-Shafer theory and deep learning, Neurocomputing, Volume 450,2021,Pages 275-293,ISSN 0925-2312,https://doi.org/10.1016/j.neucom.2021.03.066.

[205] Yim, J., Ju, J., Jung, H., Kim, J. (2015). Image Classification Using Convolutional Neural Networks With Multi-stage Feature. In: Kim, JH., Yang, W., Jo, J., Sincak, P., Myung, H. (eds) Robot Intelligence Technology and Applications 3. Advances in Intelligent Systems and Computing, vol 345. Springer, Cham. https://doi.org/10.1007/978-3-319-16841-8_52

[206] G. Shafer, A Mathematical Theory of Evidence, Princeton University, Press, Princeton, 1976.

[207] Hassan, E., Shams, M.Y., Hikal, N.A. *et al.* The effect of choosing optimizer algorithms to improve computer vision tasks: a comparative study. *Multimed Tools Appl* **82**, 16591–16633 (2023). https://doi.org/10.1007/s11042-022-13820-0

[208] Rahib H. Abiyev, Mohammad Khaleel SallamMa'aitah, "Deep Convolutional Neural Networks for Chest Diseases Detection", *Journal of Healthcare Engineering*, vol. 2018, Article ID 4168538, 11 pages, 2018. https://doi.org/10.1155/2018/4168538

[209] Rahman *et al.*, "Reliable Tuberculosis Detection Using Chest X-Ray With Deep Learning, Segmentation and Visualization," in *IEEE Access*, vol. 8, pp. 191586-191601, 2020, doi: 10.1109/ACCESS.2020.3031384.

[210] Rehman MFU, Fariha C, Anwar A, Shahzad N, Ahmad M, Mukhtar S, Farhan UlHaque M. Novel coronavirus disease (COVID-19) pandemic: A recent mini review. Comput Struct Biotechnol J. 2021;19:612-623. doi: 10.1016/j.csbj.2020.12.033. Epub 2020 Dec 31. PMID: 33398233; PMCID: PMC7773542.

[211] https://www.mygov.in/covid-19/

[212] https://www.who.int/en/activities/tracking-SARS-CoV-2-variants/

[213] Bustin SA (2000) Absolute quantification of mRNA using realtime reverse transcription polymerase chain reaction assays. J Mol Endocrinol 25(2):169–193

[214] Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, Wang B, Xiang H, Cheng Z, Xiong Y, Zhao Y, Li Y, Wang X, Peng Z. Clinical characteristics of 138 hospitalized patients with 2019 Novel Coronavirus-infected pneumonia in Wuhan, China. J Am Med Assoc. (2020). https ://doi.org/10.1001/jama.2020.1585.

[215] https://apps.who.int/iris/rest/bitstreams/1323285/retrieve

[216] Xia, W, Shao, J, Guo, Y, Peng, X, Li, Z, Hu, D. Clinical and CT features in pediatric patients with COVID-19 infection: Different points from adults. *Pediatric Pulmonology*. 2020; 55: 1169– 1174. https://doi.org/10.1002/ppul.24718

[217] Rastgarpour, M. and J. Shanbehzadeh. "Application of AI Techniques in Medical Image Segmentation and Novel Categorization of Available Methods and Tools.".

[218] Oloko-Oba M., Viriri S. (2020) Diagnosing Tuberculosis Using Deep Convolutional Neural Network. In: El Moataz A., Mammass D., Mansouri A., Nouboud F. (eds) Image and Signal Processing. ICISP 2020. Lecture Notes in Computer Science, vol 12119. Springer, Cham. https://doi.org/10.1007/978-3-030-51935-3_16

[219] . Dinesh Jackson Samuel, R., Rajesh Kanna, B. Tuberculosis (TB) detection system using deep neural networks. *Neural Comput&Applic* **31,** 1533–1545 (2019). https://doi.org/10.1007/s00521-018-3564-4

[220] Dabral I., Singh M., Kumar K. (2021) Cancer Detection Using Convolutional Neural Network. In: Tripathi M., Upadhyaya S. (eds) Conference Proceedings of ICDLAIR2019. ICDLAIR 2019. Lecture Notes in Networks and Systems, vol 175. Springer, Cham. https://doi.org/10.1007/978-3-030-67187-7_30

[221] Sharma A., Kochar B., Joshi N., Kumar V. (2021) Breast Cancer Detection Using Deep Learning and Machine Learning: A Comparative Analysis. In: Gupta D., Khanna A., Bhattacharyya S., Hassanien A.E., Anand S., Jaiswal A. (eds) International Conference on Innovative Computing and Communications. Advances in Intelligent Systems and Computing, vol 1165. Springer, Singapore. https://doi.org/10.1007/978-981-15-5113-0_39

[222] Irmak, E. Multi-Classification of Brain Tumor MRI Images Using Deep Convolutional Neural Network with Fully Optimized Framework. *Iran J Sci Technol Trans Electr Eng* **45,** 1015–1036 (2021). https://doi.org/10.1007/s40998-021-00426-9

[223] More R.B., Bhisikar S.A. (2021) Brain Tumor Detection Using Deep Neural Network. In: Pawar P.M., Balasubramaniam R., Ronge B.P., Salunkhe S.B., Vibhute A.S., Melinamath B. (eds) Techno-Societal 2020. Springer, Cham. https://doi.org/10.1007/978-3-030-69921-5_9

[224] Kitamura, G., Chung, C. & Moore, B.E. Ankle Fracture Detection Utilizing a Convolutional Neural Network Ensemble Implemented with a Small Sample, De Novo Training, and Multiview Incorporation. *J Digit Imaging* **32,** 672–677 (2019). https://doi.org/10.1007/s10278-018-0167-7

[225] Yadav, D. P., & Rathor, S. (2020). Bone fracture detection and classification using deep learning approach. In *Proceedings of the 2020 International Conference on Power Electronics*

& IoT Applications in Renewable Energy and its Control (PARC) (pp. 282–285). IEEE. https://doi.org/10.1109/PARC49193.2020.236611

[226] Liu, J., Li, J., Wang, H. *et al.* Application of deep learning in genomics. *Sci. China Life Sci.* **63,** 1860–1878 (2020). https://doi.org/10.1007/s11427-020-1804-5

[227] Danilevsky A., Shomron N. (2021) Deep Learning Applied on Next Generation Sequencing Data Analysis. In: Shomron N. (eds) Deep Sequencing Data Analysis. Methods in Molecular Biology, vol 2243. Humana, New York, NY. https://doi.org/10.1007/978-1-0716-1103-6_9

[228] Kotlarz, K., Mielczarek, M., Suchocki, T. *et al.* The application of deep learning for the classification of correct and incorrect SNP genotypes from whole-genome DNA sequencing pipelines. *J Appl Genetics* 61, 607–616 (2020). https://doi.org/10.1007/s13353-020-00586-0

[229] Abbas, A., Abdelsamea, M.M. & Gaber, M.M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Appl Intell* **51,** 854–864 (2021). https://doi.org/10.1007/s10489-020-01829-7

[230] Das, A.K., Ghosh, S., Thunder, S. *et al.* Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. *Pattern Anal Applic* (2021). https://doi.org/10.1007/s10044-021-00970-4

[231] Wang, L., Lin, Z.Q. & Wong, A. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci Rep* **10,** 19549 (2020). https://doi.org/10.1038/s41598-020-76550-z

[232] Apostolopoulos, I.D., Mpesiana, T.A. Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med* **43,** 635–640 (2020). https://doi.org/10.1007/s13246-020-00865-4

[233] Rahimzadeh, M., & Attar, A. (2020). A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Informatics in Medicine Unlocked, 19,* Article 100360. https://doi.org/10.1016/j.imu.2020.100360

[234] Fu S, Fu X, Song Y, Li M, Pan PH, Tang T, Zhang C, Jiang T, Tan D, Fan X, Sha X (2020) Virologic and clinical characteristics for prognosis of severe COVID-19: a retrospective observational study in Wuhan, China. medRxiv.

[235] Ju C, Bibaut A, van der Laan M. The Relative Performance of Ensemble Methods with Deep Convolutional Neural Networks for Image Classification. *J Appl Stat.* 2018;45(15):2800-2818. doi:10.1080/02664763.2018.1441383

[236] Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A (2020) Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images. arXiv preprint arXiv:2004.02696

[237] Cohen JP, Morrison P, Dao L (2020) COVID-19 image data collection. arXiv preprint arXiv:2003.11597

[238] Zebin, T., Rezvy, S. COVID-19 detection and disease progression visualization: Deep learning on chest X-rays for classification and coarse localization. *Appl Intell* **51,** 1010–1021 (2021). https://doi.org/10.1007/s10489-020-01867-1

[239] Ibrahim, A.U., Ozsoz, M., Serte, S. *et al.* Pneumonia Classification Using Deep Learning from Chest X-ray Images During COVID-19. *CognComput* (2021). https://doi.org/10.1007/s12559-020-09787-5

[240] .Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. 2014.

[241] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. 2015.

[242] Huang G, Liu Z, van der Maaten L, Weinberger KQ. Densely connected convolutional networks. 2016.

[243] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al. Going deeper with convolutions. Boston, MA, 2015,: IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2015. p. 19.

[244] Sokolova M, Lapalme G (2009) A systematic analysis of performance measures for classification tasks. Inf Process Manag 45(4):427–437

[245] Das, A.K., Ghosh, S., Thunder, S. *et al.* Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network. *Pattern Anal Applic* **24,** 1111–1124 (2021). https://doi.org/10.1007/s10044-021-00970-4

[246] Wang L, Lin ZQ, Wong A (2020) COVID-Net: a tailored deepconvolutional neural network design for detection of COVID-19 cases from chest X-ray images. Scientific Reports, Vol. 10

[247] Khan AI, Shah J, Bhat M (2020) CoroNet: a deep neural networkfor detection and diagnosis of COVID-19 from chest x-ray images.Comput Methods Programs Biomed 196:105581. Accessed 29 Mar 2020

[248] Reza, A.M. Realization of the Contrast Limited Adaptive Histogram Equalization (CLAHE) for Real-Time Image Enhancement. *The Journal of VLSI Signal Processing-Systems for Signal, Image, and Video Technology* **38,** 35–44 (2004). https://doi.org/10.1023/B:VLSI.0000028532.53893.82

[249] Singh, N., & Hamde, S. (2019). Tuberculosis detection using shape and texture features of chest X-rays. In H. Saini, R. Singh, G. Kumar, G. Rather, & K. Santhi (Eds.), *Innovations in electronics and communication engineering* (Vol. 65, pp. 39–46). Springer. https://doi.org/10.1007/978-981-13-3765-9_5

[250] Pasa, F., Golkov, V., Pfeiffer, F., Cremers, D., & Pfeiffer, D. (2019). Efficient deep network architectures for fast chest X-ray tuberculosis screening and visualization. *Scientific Reports, 9*(1), 1–9. https://doi.org/10.1038/s41598-019-42557-4

[251] Yadav, O., Passi, K., & Jain, C. K. (2018). Using deep learning to classify X-ray images of potential tuberculosis patients. In *Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM)* (pp. 2368–2375). IEEE.

[252] Meraj, S. S., Yaakob, R., Azman, A., Rum, S. N., Shahrel, A., Nazri, A., & Zakaria, N. F. (2019). Detection of pulmonary tuberculosis manifestation in chest X-rays using different convolutional neural network (CNN) models. *International Journal of Engineering and Advanced Technology (IJEAT), 9*(1), 2270–2275.

[253] Saha, P. (2022). An ensemble CNN-Dempster Shafer based tuberculosis detection from chest X-ray images. In *Proceedings of the 2022 IEEE Calcutta Conference (CALCON)* (pp. 228–232). IEEE. https://doi.org/10.1109/CALCON56258.2022.10060463

[254] **Saha, P.**, Neogy, S. (2022). An Overview of Pulmonary Tuberculosis Detection and Classification Using Machine Learning and Deep Learning Algorithms. In: Dua, M., Jain, A.K., Yadav, A., Kumar, N., Siarry, P. (eds) Proceedings of the International Conference on Paradigms of Communication, Computing and Data Sciences. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-16-5747-4_72

[255] **Saha, P.,** Neogy, S. Concat_CNN: A Model to Detect COVID-19 from Chest X-ray Images with Deep Learning. *SN COMPUT. SCI.* **3**, 305 (2022). https://doi.org/10.1007/s42979-022-01182-1

[256] Thian, Y. L., Li, Y., Jagmohan, P., Sia, D., Chan, V. E. Y., & Tan, R. T. (2019). Convolutional neural networks for automated fracture detection and localization on wrist radiographs. *Radiology: Artificial Intelligence, 1*(1), e180001. https://doi.org/10.1148/ryai.2019180001

[257] Lee, C., Jang, J., Lee, S. *et al.* Classification of femur fracture in pelvic X-ray images using meta-learned deep neural network. *Sci Rep* **10**, 13694 (2020). https://doi.org/10.1038/s41598-020-70660-4

[258] Wang, W., Huang, W., Lu, Q. *et al.* Attention mechanism-based deep learning method for hairline fracture detection in hand X-rays. *Neural Comput&Applic* **34**, 18773–18785 (2022). https://doi.org/10.1007/s00521-022-07412-0

[259] Cheng, C. T., Ho, T. Y., Lee, T. Y., Chang, C. C., Chou, C. C., Chen, C. C., Chung, I. F., & Liao, C. H. (2019). Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. *European Radiology, 29*(10), 5469–5477. https://doi.org/10.1007/s00330-019-06167-y

[260] Wu, H. Z., Yan, L. F., Liu, X. Q., Yu, Y. Z., Geng, Z. J., Wu, W. J., Han, C. Q., Guo, Y. Q., & Gao, B. L. (2021). The feature ambiguity mitigate operator model helps improve bone fracture

detection on X-ray radiograph. *Scientific Reports, 11*(1), 1589. https://doi.org/10.1038/s41598-021-81236-1

[261] Chung, S. W., Han, S. S., Lee, J. W., Oh, K. S., Kim, N. R., Yoon, J. P., Kim, J. Y., Moon, S. H., Kwon, J., Lee, H. J., Noh, Y. M., & Kim, Y. (2018). Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. *Acta Orthopaedica, 89*(4), 468–473. https://doi.org/10.1080/17453674.2018.1453714

[262] Gan, K., Xu, D., Lin, Y., Shen, Y., Zhang, T., Hu, K., Zhou, K., Bi, M., Pan, L., Wu, W., & Liu, Y. (2019). Artificial intelligence detection of distal radius fractures: A comparison between the convolutional neural network and professional assessments. *Acta Orthopaedica, 90*(4), 394–400. https://doi.org/10.1080/17453674.2019.1600125

[263] Yoon, A. P., Lee, Y., Kane, R. L., Kuo, C., Lin, C., & Chung, K. C. (2021). Development and validation of a deep learning model using convolutional neural networks to identify scaphoid fractures in radiographs. *JAMA Network Open, 4*(5), e216096. https://doi.org/10.1001/jamanetworkopen.2021.6096

[264] Wang, Y., Lu, L., Cheng, C.-T., Jin, D., Harrison, A. P., Xiao, J., Liao, C. H., & Miao, S. (2019). Weakly supervised universal fracture detection in pelvic X-rays. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2019: Proceedings of the 22nd International Conference* (Vol. 22, pp. 459–467). Springer.

[265] Lindsey, R., Daluiski, A., Chopra, S., Lachapelle, A., Mozer, M., Sicular, S., Hanel, D., Gardner, M., Gupta, A., Hotchkiss, R., & Potter, H. (2018). Deep neural network improves fracture detection by clinicians. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11591-11596. https://doi.org/10.1073/pnas.1806905115

[266] Jia, Y., Wang, H., Chen, W., Wang, Y., & Yang, B. (2022). An attention-based cascade R-CNN model for sternum fracture detection in X-ray images. *CAAI Transactions on Intelligent Technology*, *7*, 658–670.

[267] Ali, L., Alnajjar, F., Jassmi, H. A., Gocho, M., Khan, W., & Serhani, M. A. (2021). Performance evaluation of deep CNN-based crack detection and localization techniques for concrete structures. *Sensors*, *21*(5), 1688. https://doi.org/10.3390/s21051688

[268] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 2261–2269). Honolulu, HI, USA. https://doi.org/10.1109/CVPR.2017.243

[269] https://radiopaedia.org/

[270] https://in.mathworks.com/help/visionhdl/ug/contrast-adaptive-histogram-equalization.html

[271] Sree Vidya, B., & Chandra, E. (2019). Triangular fuzzy membership-contrast limited adaptive histogram equalization (TFM-CLAHE) for enhancement of multimodal biometric images. *Wireless Personal Communications, 106*(2), 651–680. https://doi.org/10.1007/s11277-019-06184-6

[272] Huy, V. T. Q., & Lin, C. -M. (2023). An improved DenseNet deep neural network model for tuberculosis detection using chest X-ray images. *IEEE Access, 11,* 42839–42849. https://doi.org/10.1109/ACCESS.2023.3270774

[273] Kasban, H., El-Bendary, M.A.M. and Salama, D.H., 2015. A comparative study of medical imaging techniques. *International Journal of Information Science and Intelligent System*, *4*(2), pp.37-58.

[274] Hesamian, M.H., Jia, W., He, X. *et al.* Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *J Digit Imaging* **32**, 582–596 (2019). https://doi.org/10.1007/s10278-019-00227-x

[275] Haque, M.R., Islam, M.M., Iqbal, H., Reza, M.S., & Hasan, M.K. (2018). Performance Evaluation of Random Forests and Artificial Neural Networks for the Classification of Liver Disorder. *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*, 1-5.

[276] Islam, M.M., Iqbal, H., Haque, M.R., & Hasan, M.K. (2017). Prediction of breast cancer using support vector machine and K-Nearest neighbors. *2017 IEEE Region 10 Humanitarian Technology Conference (R10-HTC)*, 226-229.
[277] Hasan, Md Kamrul & Islam, Md & Hashem, M.M.A.. (2016). Mathematical Model Development to Detect Breast Cancer Using Multigene Genetic Programming. 10.1109/ICIEV.2016.7760068.
[278] Xu, A., Wang, L., Feng, S., & Qu, Y. (2010). Threshold-Based Level Set Method of Image Segmentation. *2010 Third International Conference on Intelligent Networks and Intelligent Systems*, 703-706.
[279] Cigla, C.; Alatan, A.A. Region-based image segmentation via graph cuts. In Proceedings of the 2008 15th IEEE International Conference on Image Processing, San Diego, CA, USA, 12–15 October 2008; pp. 2272–2275.
[280] Yu-Qian, Z.; Wei-Hua, G.; Zhen-Cheng, C.; Tang, J.-T.; Li, L.-Y. Medical images edge detection based on mathematical morphology. In Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference, Shanghai, China, 17–18 January 2006; pp. 6492–6495.
[281] Yan C, Fan X, Fan J, Wang N. Improved U-Net Remote Sensing Classification Algorithm Based on Multi-Feature Fusion Perception. *Remote Sensing*. 2022; 14(5):1118. https://doi.org/10.3390/rs14051118
[282] Chen, C., Liu, MY., Tuzel, O., Xiao, J. (2017). R-CNN for Small Object Detection. In: Lai, SH., Lepetit, V., Nishino, K., Sato, Y. (eds) Computer Vision – ACCV 2016. ACCV 2016. Lecture Notes in Computer Science(), vol 10115. Springer, Cham. https://doi.org/10.1007/978-3-319-54193-8_14
[283] He K, Gkioxari G, Dollár P and Girshick R, "Mask R-CNN," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 2980-2988, doi: 10.1109/ICCV.2017.322.
[284] Noh, H.; Hong, S.; Han, B. Learning deconvolution network for semantic segmentation. In Proceedings of the IEEE International Conference on Computer Vision, Las Condes, Chile, 11–18 December 2015; pp. 1520–1528.
[285] Lin, G.; Milan, A.; Shen, C.; Reid, I. Refinenet: Multi-path refinement networks for high-resolution semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1925–1934.
[286] Long, J.; Shelhamer, E.; Darrell, T. Fully convolutional networks for semantic segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 3431–3440.
[287] Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; Bengio, Y. Generative adversarial nets. Adv. Neural Inf. Process. Syst. 2014, 27, 2672–2680.
[288] Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.
[289] Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A,Dataset of breast ultrasound images,Data in Brief,Volume 28,2020,104863, ISSN 2352-3409, https://doi.org/10.1016/j.dib.2019.104863.
[290] Zhou, Z., Rahman Siddiquee, M.M., Tajbakhsh, N., Liang, J. (2018). UNet++: A Nested U-Net Architecture for Medical Image Segmentation. In: Stoyanov, D., *et al.* Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support. DLMIA ML-CDS 2018 2018. Lecture Notes in Computer Science(), vol 11045. Springer, Cham. https://doi.org/10.1007/978-3-030-00889-5_1
[291] Piantadosi, G., Marrone, S., Galli, A., Sansone, M., & Sansone, C. (2019). DCE-MRI Breast Lesions Segmentation with a 3TP U-Net Deep Convolutional Neural Network. *2019 IEEE 32nd International Symposium on Computer-Based Medical Systems (CBMS)*, 628-633.
[292] Ma, X., Wang, j., Zheng, X., Liu, Z., Long, w., Zhang, Y., Wei, J., and Lu, Y., “Automated fibroglandular tissue segmentation in breast MRI using generative adversarial networks,” Phys. Med. Biol., vol. 65, no. 10, May 2020, Art. no. 105006.

[293] Yue, W., Zhang, H., Zhou, J., Li, G., Tang, Z., Sun, Z., Cai, J., Tian, N., Gao, S., Dong, J., Liu, Y., Bai, X., & Sheng, F. (2022). Deep learning-based automatic segmentation for size and volumetric measurement of breast cancer on magnetic resonance imaging. *Frontiers in Oncology*, 12, Article 984626. https://doi.org/10.3389/fonc.2022.984626

[294] El Adoui, M.; Mahmoudi, S.A.; Larhmam, M.A.; Benjelloun, M. MRI Breast Tumor Segmentation Using Different Encoder and Decoder CNN Architectures. *Computers* **2019**, *8*, 52. https://doi.org/10.3390/computers8030052

[295] Li,C., Sun, H., Liu, Z., and Wang, M.,"Learning cross-modal deep representations for multi-modal MR image segmentation," in Proc. MICCAI, Shenzhen, China, 2019, pp. 57–65.

[296] Zhang J, Saha A, Zhu Z, Mazurowski MA. Hierarchical Convolutional Neural Networks for Segmentation of Breast Tumors in MRI With Application to Radiogenomics. IEEE Trans Med Imaging. 2019 Feb;38(2):435-447. doi: 10.1109/TMI.2018.2865671. Epub 2018 Aug 16. PMID: 30130181.

[297] Peng C, Zhang Y, Zheng J, Li B, Shen J, Li M, Liu L, Qiu B, Chen DZ. IMIIN: An inter-modality information interaction network for 3D multi-modal breast tumor segmentation. Comput Med Imaging Graph. 2022 Jan;95:102021. doi: 10.1016/j.compmedimag.2021.102021. Epub 2021 Nov 29. PMID: 34861622.

[298] Khaled R, Vidal J, Vilanova JC, Martí R. A U-Net Ensemble for breast lesion segmentation in DCE MRI. Comput Biol Med. 2022 Jan;140:105093. doi: 10.1016/j.compbiomed.2021.105093. Epub 2021 Nov 30. PMID: 34883343.

[299] Hirsch, L., Huang, Y., Luo, S., Rossi Saccarelli, C., Lo Gullo, R., Daimiel Naranjo, I., Bitencourt, A. G. V., Onishi, N., Ko, E. S., Leithner, D., Avendano, D., Eskreis-Winkler, S., Hughes, M., Martinez, D. F., Pinker, K., Juluru, K., El-Rowmeim, A. E., Elnajjar, P., Morris, E. A., Makse, H. A., Parra, L. C., & Sutton, E. J. (2021). Radiologist-level performance by using deep learning for segmentation of breast cancers on MRI scans. *Radiology: Artificial Intelligence*, 4(1), e200231. https://doi.org/10.1148/ryai.200231

[300] https://cytecare.com/blog/breast-cancer/statistics-of-breast-cancer/

[301] Parekh, V. S., Macura, K. J., Harvey, S. C., Kamel, I. R., Ei-Khouli, R., Bluemke, D. A., & Jacobs, M. A. (2020). Multiparametric deep learning tissue signatures for a radiological biomarker of breast cancer: Preliminary results. *Medical Physics*, 47(1), 75-88. https://doi.org/10.1002/mp.13849

[302] Long, J., Shelhamer, E., & Darrell, T. (2015). Fully convolutional networks for semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3431–3440). Athens, Greece.

[303] Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O. (2016). 3D U-Net: Learning Dense Volumetric Segmentation from Sparse Annotation. In: Ourselin, S., Joskowicz, L., Sabuncu, M., Unal, G., Wells, W. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016. MICCAI 2016. Lecture Notes in Computer Science(), vol 9901. Springer, Cham. https://doi.org/10.1007/978-3-319-46723-8_49

[304] Piantadosi, G., Sansone, M., Fusco, R., & Sansone, C. (2020). Multi-planar 3D breast segmentation in MRI via deep convolutional neural networks. *Artificial Intelligence in Medicine*, 103, Article 101781. https://doi.org/10.1016/j.artmed.2019.101781

[305] Khaled, R., Vidal, J., Vilanova, J. C., & Martí, R. (2022). A U-Net ensemble for breast lesion segmentation in DCE MRI. *Computers in Biology and Medicine, 140*, 105093. https://doi.org/10.1016/j.compbiomed.2021.105093

[306] Ciresan, D., Giusti, A., Gambardella, L., & Schmidhuber, J. (2012). Deep neural networks segment neuronal membranes in electron microscopy images. In *Advances in Neural Information Processing Systems* (Vol. 25).

[307] Hu, X.-d., Wang, X.-q., Meng, F.-j., Hua, X., Yan, Y.-j., Li, Y.-y., Huang, J., & Jiang, X.-l. (2020). Gabor-CNN for object detection based on small samples. *Defence Technology*, 16(6), 1116–1129. https://doi.org/10.1016/j.dt.2019.12.002

[308] Özbulak, G., & Ekenel, H. K. (2018). Initialization of convolutional neural networks by Gabor filters. In *2018 26th Signal Processing and Communications Applications Conference (SIU)* (pp. 1–4). IEEE. https://doi.org/10.1109/SIU.2018.8404757