Fall 7-31-2022

APPLICATION OF DEEP LEARNING FOR MEDICAL SCIENCES AND EPIDEMIOLOGY DATA ANALYSIS AND DIAGNOSTIC MODELING

Somenath Chakraborty
University of Southern Mississippi

APPLICATION OF DEEP LEARNING FOR MEDICAL SCIENCES AND EPIDEMIOLOGY DATA ANALYSIS AND DIAGNOSTIC MODELING

by

Somenath Chakraborty

A Dissertation
Submitted to the Graduate School,
the College of Arts and Sciences
and the School of Computing Sciences and Computer Engineering
at The University of Southern Mississippi
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy

Approved by:

Dr. Beddhu Murali, Committee Chair
Dr. Andrew Sung
Dr. Sarah B. Lee
Dr. Amer Dawoud
Dr. Sungwook Lee

August 2022

Published by the Graduate School

ABSTRACT

Machine Learning and Artificial Intelligence have made significant progress concurrent with new advancements in hardware and software technologies. Deep learning methods heavily utilize parallel computing and Graphical Processing Units(GPU). It is already used in many applications ranging from image classification, object detection, segmentation, cyber security problems and others. Deep Learning is emerging as a viable choice in dealing with today's real-time medical problems. We need new methods and technologies in the field of Medical Science and Epidemiology for detecting and diagnosing emerging threats from new viruses such as COVID-19. The use of Artificial Intelligence in these domains is becoming more accepted, not as a replacement for current medical practices but rather as an enhancing and augmenting tool to current practices. This dissertation is about the Application of Deep Learning for Medical Sciences and Epidemiology Data analysis and Diagnosis Modeling. The work started with an analysis of existing data, then focused on developing diagnosis models using both supervised and unsupervised approaches. This resulted in a renal transplantation recommender system, breast cancer diagnosis model and COVID-19 diagnosis models. The recent completed work for COVID-19 diagnosis modeling leverage the potential of more advanced deep neural network model and vision transformer. The results are very promising and outperformed all existing models. The final diagnosis model has three kinds of dataset. For Chest X-Ray(CXR) image dataset training accuracy is 95.9376 %, validation accuracy is 96.1667 % and testing accuracy is 95.1250 %. For Chest CT image dataset training accuracy is 96.5474 %, validation accuracy is 95.7302 % and testing accuracy is 97.0588 %. For the Combined CXR and Chest CT image dataset training

accuracy is 95.6665 %, validation accuracy is 96.6961% and testing accuracy is 97.8859 %.

ACKNOWLEDGMENTS

It is my immense pleasure to thank all the people who inspired me constantly, motivated me and provided help and support whenever needed.

First and foremost, I like to thank my supervisor and my Ph.D. dissertation committee chair Dr. Beddhu Murali, who is instrumental in my Ph.D. journey and provide constant help and support whenever needed. I have had the privilege to have him as my mentor. He guided me throughout my Ph.D. journey and actively helped me in the methodology development phase, thoroughly revised every paper I wrote and provided insightful feedback. He is an excellent mentor from whom I learned a lot.

I also like to thank committee member, Dr. Andrew Sung, who motivated me throughout my Ph.D. journey and always provided the necessary support.

I also received immense help from our school director and my Ph.D. committee member, Dr. Sarah B. Lee. Dr. Amer Dawoud and Dr. Sungwook Lee also helped me to improve my work by providing valuable feedback. I also express my sincere gratitude to Dr. Karen S. Coats, Associate Provost and Dean of the Graduate School for her immense support and to Dr. Janet Donaldson, Associate Dean for Research and Graduate Education, for strengthening me though leadership skills and other professional activities.

I would like to thank many friends and professors at my previous institutions who always stayed on my side so that I can stay strong and do some meaningful work.

DEDICATION

I would like to dedicate this work to my parents without whom nothing would be possible and also express my sincere gratitude to my supervisor Dr. Beddhu Murali, who always stayed beside me whenever I needed it most. His enormous, excellent guidance played an instrumental role in the completion of my Ph.D., alongside his insightful feedback which helped me grasp new knowledge and sharpen my programming skills.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ML	Machine Learning
DL	Deep Learning
CNN	Convolutional Neural Network
DNN	Deep Neural Network
Chest X-Ray	Chest Radiography or Lung Radiography
CXR	Chest X-Ray
CT	Computed Tomography scan
Chest CT	Computed Tomography of Chest
OT	Organ Transplantation
GAOGB	Genetic Algorithm-based Online Gradient Boosting
PAM	Partitioning Around Medoids
MRI	Magnetic Resonance Imaging
MBC	Metastatic Breast Cancer
CDC	Centers for Disease Control and Prevention
RF	Random Forest
SVM	Support Vector Machine
MLP	Multi-Layer Perceptron
NMPS	Novel Medical Prognosis System
GA	Generic Algorithm
SARS CoV-2	Severe Acute Respiratory Syndrome Coronavirus-2

WHO	World Health Organization
PHNNM	Parallel Phase Neural Network Model
LT	Liver Transplantation
GPU	Graphical Processing Unit
PD	Peritoneal Dialysis
SRTR	Scientific Registry of Transplant Recipients
UNOS	United Network for Organ Sharing
ULAM	UNOS Liver Allocation Model
PVA	Polyvinyl Alcohol
MDP	Markov Decision Process
RL	Reinforcement Learning
QoL	Quality of Life
MSE	Mean Squared Error
MAE	Mean Absolute Error
CUDA	Compute Unified Device Architecture
RFR	Random Forest Regressors
HPC	High Performance Computing
NSF	National Science Foundation
RT-PCR	Real-Time Reverse Transcription Polymerase Chain Reaction
EML	Ensemble Machine Learning
AI/ML	Artificial Intelligence/Machine Learning
HR.Net	High Resolution Network
RNN	Recurrent Neural Network

ANN	Artificial Neural Network
ResNet	Residual Neural Network
API	Application Programming Interface
RBF	Radial Basis Function
ROC Curve	Receiver Operating Characteristic Curve
AUROC	Area Under the Receiver Operating Characteristics
VGG	Visual Geometry Group
BEiT	Bidirectional Encoder representation from Image Transformers

CHAPTER I - INTRODUCTION

This chapter starts with the motivation behind the present work and then presents a brief overview of the fields of Artificial Intelligence and Machine Learning and Deep Learning, in particular. After that it presents an overview of the datasets available and used in the present work. We have many publications and many are under review. One conference paper was awarded the best paper award and a journal article is featured at World Health Organization(WHO)'s website for global research on COVID-19 pandemic.

1.1 Motivation

Initially, the work started with the exploration of possible domain of artificial intelligence and machine learning research fields. Extensive literature studies on supervised learning methodologies, semi-supervised learning methodologies, unsupervised learning methodologies helped to formulate the path by which we can design our model and solve different kinds of problems which are of current interest. The initial work started with the renal transplantation model design, then the design of predictive model for breast cancer was developed both using unsupervised and supervised approaches. When the COVID-19 pandemic broke out our focus shifted to designing machine learning models to help in the understanding of this pandemic and come up with practical solutions. The initial problem was the lack of information and data related to the COVID-19 pandemic. But, gradually the research community, medical fraternity and other resources made data available online so we can start analyzing those and build up our problem solving models. We were able to develop many models to design diagnosis model for COVID-19 using deep learning methodologies. We initially

developed a prognosis model for COVID-19 detection using Chest X-ray(CXR) image data, then we designed an ensemble model, then we designed a prognosis model using Chest Computer Tomography(CT) images. Recently, we also successfully developed integration of transformative and convolutional neural network using a combination of CXR, CT images.

1.2 Artificial Intelligence

"AI is the study of complex information processing problems that often have their roots in some aspect of biological information processing. The goal of the subject is to identify solvable and interesting information processing problems and solve them."

− David Marr, Computational neuroscientist and physiologist.

"The intelligent connection of perception to action"

− Rodney Brooks, Artificial Intelligence Scientist

"Actions that are indistinguishable from a human's"

− Alan Turing, Father of theoretical computer science and artificial intelligence

Artificial Intelligence(AI) is revolutionizing every aspect of our daily life and is already used widely in a number of disciplines. The term AI is used in a more generic sense in today's computational work. It integrates and give birth to many revolutionary technologies like neural networks, deep learning, representation learning, reinforcement learning to name a few. One could say that it all started form the 'Turing Test' experiment by Alan Turing in 1950, where in order to pass the Turing test a machine's action should become distinguishable from a human's actions as observed by another human. Once this goal was set, a paradigm shift happened in the above fields with

revolutionary research by thousands of scientists all around the world. The following Fig. 1.1 illustrates the recent scenario of research area in an abstract way.

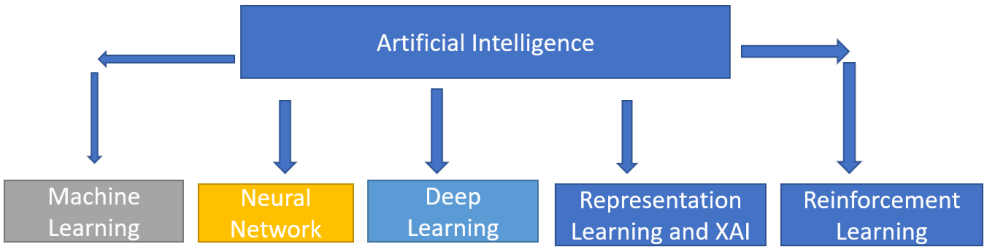

Figure 1.1 The abstract representation of Artificial Intelligence and related branches

The artificial intelligences comprises of many technologies and theoretical fields, like Machine learning, Neural Network, Deep Learning, representation learning and explainable Artificial Intelligence(XAI). Also, reinforcement learning is getting popularity in the field of simulations, agent based learning, multi-agent learning, game design, etc. In the following sections will gives brief introduction about those fields which we actually use in this work and have relevant publication related to this thesis.

1.3 Machine Learning

Machine learning has evolved a lot in the last two decades and we now also have more sophisticated hardware with large memory resources available both in local machines and in the virtual cloud machine. Data centers, cloud services have contributed a lot for the growth of machine learning as the core of machine learning involves working with data. Huge amount of data processing is an integral part of machine learning domain. Data could be in any form like raw data, satellite data, image data, video data, audio data or any combination of multimedia data or signals. If we wish to design a

machine learning model we need to have the data, essentially a large amount of data can help a lot in designing the model in a robust way. The design and performance of the model also depends heavily on the preprocessing of the data as well. This varies a lot depending upon which type of model we are focusing on. For supervised machine learning methodologies the data annotations and labeling is an important aspect of design the model. Preprocessing deals with how we can refine the data so that it should fit appropriately with the model, especially when we have many missing values or null values. Practically machine learning problems deals with realistic problem solving and designing model with the help of data. So, meaning data extraction, effective feature extraction and automatic processing is an integral part of the system, but those realistic problems have many missing values or null values in their parameters. The parameters are the attributes by which we understand the system and try to model using different model design techniques such as whether it is regression problem, classification problem, segmentation problem or object detection problem. Most datasets for all kinds of realistic problem domains have those missing values and null values. So, we can write codes which can delete those values or replaced it with some constant values. In many times we also perform many operations to convert non-numerical categorical data, like yes or no, benign or malignant to have a numerical value so that computer system can understand it in proper way as data is considered as a vector or tensor in the machine learning model. Another kinds of data preprocessing involves with the operation of standardizing the dataset. It require deep understanding of the problem domain and how we are going to standardize those values. Understanding the data and preprocessing also help to recognize which model is best suited for that particular dataset, whether it require nonlinear model

or linear model or simple mathematical model and statistical model can solve the problem. The analysis of the dataset and its complexity is the fundamental part of any machine learning model design. The Machine learning model require a board range of mathematical and statistical understanding as well. Like for regression problem analysis, we often use linear algebra and vector calculus. Likewise for classification problem, optimization theories and matrix decomposition methods are widely used.

1.4 Deep Learning

Deep Learning(DL) nowadays is getting more popularity as it has successfully provided leading results in many regression, classification and other kinds of problem settings.

Deep learning is the way by which we introduce more hidden layers in a neural network and design algorithms which can process the data in a faster way, say, using multiple GPUs, than the traditional machine learning methods. Deep learning generally uses parallel processing concepts and due to the advancement of multi-core central processing unit(CPU) and Graphical Processing Unit(GPU) those theoretical concepts are now implemented exploiting these advanced hardware support. Deep multiple layers in the neural network architecture of DL systems are called as Deep Neural Network(DNN). It is more useful for complex data representation understanding and extract feature which are more robust, and accurate. It is currently providing breakthrough results in intensive data analytics problems including Big-data problems, image processing and image classification, image segmentation and image object detection, video image analysis and video processing problems, audio analysis and speech recognition, Natural language Processing(NLP), cyber security to name a few. These applications cover different

industry sectors like computer vision, public health, education, transportation, robotics, airline industries, drug discovery, bio-informatics, cyber security and others. Deep learning also gaining remarkable success in the field of agent learning, multi-agent learning, Simulation design, reinforcement learning and other kinds of representation learning. Deep learning is very much synonymous with advanced statistical machine learning and advanced artificial intelligence techniques.

1.5 Medical Data Analysis and Prognosis Model Design

The use of Machine Learning(ML) and Artificial Intelligence(AI) in the field of medical and epidemiological data analytics and diagnostic modeling, open new window of hopes and revolutionize the field. The perspective of solving problems in this domain changed remarkable by the use of automated systems and diagnostic tools. Artificial Intelligence (AI) has been brought into medical practice because it is helping to save thousands of human lives. AI has become the essential choice for solving emerging medical problems. As novel viruses challenge conventional solutions, a trending introspection in the field of Medical Science and Epidemiology constantly draws on AI-powered resources such as innovative real-time tools to make up for identifiable shortcomings. Thus, AI provides an indispensable, complementary support to current medical practice, empowering the global health system to respond effectively to present and future emergencies. AI provides innovative real-time tools with essential support to boost the responsiveness of medical practice to life-threatening challenges. The interconnection of unattended network devices incorporating sensors (aka things) form the Internet of Things (IoT) whose functions can be bolstered using AI, especially in medical fields. The things and the AI that powers them make up the Artificial

Intelligence of Things (AIoT). In medical practice, image data capturing devices such as smartphones and other wearable devices are useful for measuring human health parameters. The measuring process involves obtaining and monitoring real time health data such as parameter values obtained from patients or any healthy human being. Heart, diabetes, and blood temperature parameters are some examples of the parameter values. Human health related data can be obtained by using our smartphone and some wearable devices the human health parameters can be easily measured. Those parameter values received from the patient or any healthy human being could be as follows,

1) Heart Parameters: Like heart rate, oxygen saturation, heart rate variability even Electrocardiography(ECG).
2) Blood Pressure and Temperature sensor etc.
3) Diabetics parameters: Like glucose measurement, heart rate, etc.
4) Image capturing devices

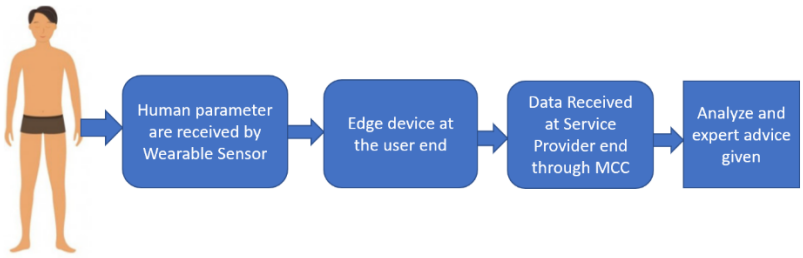

Figure 1.2 *The Remote Human health monitoring system*

Figure 1.2 shows how with the integration of internet of thing(IoT), big-data, mobile cloud computing we can easily automate the process of monitoring human health by AI techniques. It would process the data in real time, then after analyzing and identifying the issue it would provide recommended guidelines along with a report of different health parameters. The availability of wide range of medical open access

datasets online and the availability of approval-based datasets online, ranging from radiological image dataset, different disease related analytical dataset has provided much help to build different deep learning models rapidly with high efficiency. Many image-capturing systems are able to produce radiological image data and feeding such images to a Deep Learning model could lead to the detection of many kinds of diseases in near real-time. At the least they can provide valuable information to radiologists regarding that disease. One of our recent publications is in this regard where an automated system can detect COVID-19 using Chest X-Ray images. This research is already featured on the official website of the World Health Organization(WHO) on COVID-19 Global literature on Coronavirus Disease.

CHAPTER II - LITERATURE REVIEW

This chapter would provide the literature related to our research work and showing the research potential, limitations and shortcoming. Our journey started with the reading and conceptualization the famous work of Alan Turing[1], where he proposed to consider the question, “Can machines think?” with his elaboration of imitation game. This article was published in 1950 at the British Journal named Mind. Then we make comprehensive understanding from the artificial intelligence book[2], which mainly focuses on modern approach and practices in this field. Then we study different application domains of AI, namely education[3], Communication[4], Transportation[5], Agriculture[6], Entertainment, Finance[7], Public health sector[8-10]. The book chapter[11] on category identification is also review many works related to feature generation. The next section will shows the literature review related to our work on application of deep learning for medical sciences and epidemiology data analytics and diagnostic modeling.

2.1 Survival Prediction Model for Renal Transplantation

Successful Organ Transplantation (OT) saves thousands of life all over the world and it is now a common practice everywhere. Hence, there are many research works done for different kinds of OT. Renal Transplantation, Liver Transplantation, Heart Transplantation, Islet transplantation for Type-1 Diabetic patients etc. are now very popular and successful treatment facilities available for medical practitioners around the world. F. Locatelli et. al. [12] describe about the basic areas of renal treatment, where they are emphasizing on peritoneal dialysis (PD) and renal transplantation as a successful

medical treatment. On average, 95 transplants take place each day in the USA alone. Scientific Registry of Transplant Recipients (SRTR), USA, records all ongoing evaluation of the scientific and clinical status of solid organ transplantation. In the USA in the year 1995, the United Network for Organ Sharing (UNOS) invented the first computerized simulation of the liver allocation process, known as the UNOS Liver Allocation Model (ULAM).Various medical prognosis (MP) models mainly based on statistical data analytics modeling. Z. Qi et al. [13] describe in their paper how Polyvinyl Alcohol (PVA) can be applied in Islet transplantation for Type-1 Diabetic patients though there are not much research work done on the long-time successful rate of Survival of the patient with very accurate MP models. A. K. Iyer et al. [14] presents a biologically based discrete-event simulation model for liver transplantation. Their approach is a variant of Alagoz, O., Bryce, C. L., Shechter et al. [15] model which is based on Empiric Natural History Model. Alagoz, Oguzhan, et al. [16]. They address the problem of optimally timing a living-donor liver transplant to maximize the patient's total reward, such as quality-adjusted life expectancy. This approach is based on Markov Decision Process(MDP). The drawback of this kind of reward-oriented modeling is very complex to implement and as MP is hugely depends on patient’s historical data, so reinforcement learning (RL) approach does not yields fruitful results. J. H. Fielder [17] describes a lot of issues related to organ transplants. D. Medved, P. Nugues et al. [18] address the issue of waiting time related to heart transplant. Through they discuss about 10 most significant parameters predicting the patient status for the three different time points, but not explore the survival of the patient in a long run. A. Ravikumar et al. [19] proposed a MP model of renal transplantation using Support Vector Machine (SVM). Though they determine

survival critic factor using Cox regression model, but their approach lacks long term survival analysis with more detail data and parameter. L. Al-Ebbini et al. [20] proposed a hybrid technique to understand different parameters and predict the survival which they termed as quality of life (QoL) for lung transplantation. Their approach mainly focused on parameter ranking rather than survival analysis. C. G. Raji et al. [21] presents their approach for survival of liver transplantation using multi-layer perceptron. Their approach lacks speed in implementation as they not proposed and implemented their approach in parallel mode.

2.2 Medical Prognosis System for Breast Cancer

There were two different types of work we had completed in this regard. The following section will describe the literature review for each of the work which was completed regarding Medical Prognosis System design for Breast Cancer.

2.2.1 Breast Cancer Dataset Analysis using Unsupervised Learning

The research work that was done by S. Muthukumaran et al.[22], Nepomuceno et al.[23], Doruk Bozdag et al.[24] use Bi-clustering approach. V Sangeetha [25] et al. investigate their research using genetic algorithm and Spiking Neural Networks to generate cluster in their breast cancer dataset which are very prosing in this domain of research. Their integration of genetic algorithm and Spiking Neural Network helps to develop the prognosis system. T. Padhi [26] et al., D. Verma et al. [27] uses weka Software tool to get the cluster for breast cancer dataset. L. M. Naeni et al. [28], Taosheng Xu et al.[29] uses gene expression and biomarker to investigate gene expression profiling to cluster the cancer dataset. A. L. Fijri et al. [30], P. H. S. Coutinho et al. [31] describe in their paper some kinds of Fuzzy analysis to cluster the dataset. D.

Wu, L. Sheng et al. [32] and Chen D et al. [33] describe in their paper an ensemble algorithm for clustering cancer data. This algorithm is based on the k-medoids or partitioning around medoids (PAM) algorithm.

All this existing research mainly focus on predicting the class in known perspective rather than investigation of the fundamental unsupervised mechanism to segregate the datapoint in the dataset.

Investigation of the dataset in unsupervised way is very import in terms of designing more accurate model for data labeling which further can be used by classification and prediction problem.

2.2.2 Breast Cancer Prognosis Model Design

While had acquired the knowledge of breast cancer data analysis and implementing our unsupervised clustering analysis which were presented in the previous section, now we were thinking more holistically and try to develop the breast cancer prognosis model using machine learning technique. We had observed that Medical specialists need to depend on a host of pathology tests after initial biopsy. It takes a lot of time like weeks and sometimes months before they decide whether they should go for lumpectomy operation or mastectomy operation, or chemotherapy would be the better choice, or radiotherapy, otherwise endocrine therapy needed or not which was fully depends on their diagnosis and treatment and lot of clinical, pathological and recommender system guidelines [34 -36]. Due to this there is potential research demands to automatically identify and detect whether a patient's tumor is malignant or benign with the help of statistical machine learning model or ensemble learning model. Jafarpisheh, Nafisi and Teshnehlab [37] describe a combination of machine learning approaches for breast

cancer. Naveen, Sharma and Ramachandran Nair [38] presents an ensemble machine learning model. Amitha and Selvamani [39] provides a survey on automatic breast cancer grading of histopathological images. They claim that computer-aided image analysis systems can be used to automatically find their breast cancer grade. Imani, Chen, Tucker and Yang [40] describes survival analysis of cancer recurrences based on random forest modeling. Elnahas, Hussein and Keshk [41] proposed an ensemble technique using artificial neural network to im-prove classification accuracy. Kumari and Krishna [42] provides a survey of different machine learning model which are in used to design the prognosis of the diseases. Zhu W, Xie L, Han J and Guo X [43] presents a comprehensive review of cancer prognosis system by use of deep learning methods. B. Fu, P. Liu, J. Lin, L. Deng, K. Hu and H. Zheng [44] describe an early-stage breast cancer prognosis model, they named it as MP4Ei Framework. Huang S, Yang J, Fong S, Zhao Q [16] analyses in their review work about research on the application of Artificial Intelligence (AI) to cancer diagnosis and prognosis. They also give a brief discussion about the ad-vantages of using AI in this field. H. Lu, H. Wang, S.W. Yoon [45] proposed a novel genetic algorithm-based online gradient boosting (GAOGB) model for incremental breast cancer (BC) prognosis where it has been adopted with Gradient Descent based loss minimization and Generic Algorithm(GA) based parameter optimization schemes.

2.3 COVID-19 Prognosis Model Design

When the COVID-19 pandemic broke out we started gathering all sort of information related to COVID-19 from all the sources available. We then started to collect the dataset whichever available that time and later on increasing the size of the dataset as the pandemic continues and cases increases day by day. The dataset we had collected

consists of Chest X-Ray(CXR) image data and Chest Computed Tomography(CT) scan image data.

The deep neural network model we work have divided into four research sections as in each research direction we follow different kinds of methodology to build the COVID-19 prognosis model. So, the following section would describe the literature review which are related to this four different kinds of research work related to COVID-19.

2.3.1 COVID-19 Prognosis Model Design using CXR

The Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) causes the disease which was named by the scientist as COVID-19. This COVID-19 disease became the most horrible pandemic in this current century. Initially, we gather the information from World Health Organization(WHO), Center for Disease Control and Prevention, other online news media and sources etc., [46–48]. Many reported source confirmed that COVID-19 initially started in Wuhan City of China in December, 2019[46-48]. Research suggests that Lung involvement is the crucial noticeable manifestation of the disease, ranging from asymptomatic disease or mild pneumonia, to severe disease associated with hypoxia, critical disease associated with shock, respiratory failure and multi-organ failure or death[49]. At the very beginning of the pandemic, due to the lack of knowledge of the virus genome, the detection of the disease was undeveloped. COVID-19 RT-PCR, a real-time reverse transcription polymerase chain reaction (RT-PCR) test, is a very efficient discovery in the qualitative detection of nucleic acid from SARS-CoV-2 in upper and lower respiratory specimens [50]. Still now RT-PCR is considered as the clinical public health standard for the detection of COVID-19. Many emerging issues, still, remain in COVID-19 detection domain. The RT-PCR sensitivity rate is 60–70%,which is not very

much promising and the test is expensive [51]. There is another test also very popular nowadays known as rapid antigen detection test, which takes less amount of time compared to RT-PCR as RT-PCR is very time consuming but with this rapid antigen test we can have the confirmation report within an hour, but the problem is that the sensitivity of this test is very low i.e. 60.5% in average sample cases.[51]. It is also do not able to show the severity of the disease and often confused with other Pneumonia variants. Not capable to distinguish different variant of COVID-19 as well.

In terms of machine learning and artificial intelligence technique is concerns, two very efficient detection methods of COVID-19 is gaining popularity hugely in the prognosis research community. These two approaches uses chest X-ray (CXR) and chest CT-scan [52-60]. Deep Learning methods have revolutionized the field of AI, and due to their potential application in medical diagnosis and prognosis systems, millions of human lives could be saved [61,62]. Radiographic images such as CXR are inexpensive and less time consuming to produce results compared to other clinical or laboratory modalities. Being COVID-19 positive is characterized primarily by patches of ground-glass opacity and consolidations [58]. Therefore, the CXR dataset is useful to test algorithms for detecting COVID-19 and other pulmonary disorders. Successful machine learning approaches could allow for rapid evaluations of chest X-ray images and thus enable radiologists to filter potential candidates in a time-effective manner [59,60]. Studies using deep neural networks have shown the effectiveness of the method in the diagnosis of pneumonia [59]. Another study of deep artificial neural networks using 470,388 fully anonymized institutional adult chest radiographs from three hospitals was able to detect abnormalities from normal adult chest radiographs with a high positive predictive value of 94% [60]. In

the latter study, the average reporting delay was reduced from 11.2 to 2.7 days for critical imaging findings ($p < 0.001$) and from 7.6 to 4.1 days for urgent imaging findings ($p < 0.001$) in the simulation compared with historical data [60].

According to the Centers for Disease Control and Prevention (CDC) [63], the Delta variant of the coronavirus (B.1.617.2) causes more infections and spreads faster than the original SARS-CoV-2 strain. Currently, vaccines remain the best way to reduce the risk of severe illness, hospitalization, and death from COVID-19. However, the CDC cautioned that the effectiveness of the vaccines against new variants that may arise, including the new coronavirus variant Omicron (B.1.1.529), is not clear at this point in time [64].

Detection of COVID-19 using CXR has certain advantages over other methods. Because of the multiple mutations of the virus and the emergence of newer variants such as the Delta variant and the South African Omicron variant (B.1.1.529), there is an urgent need for rapid and automatic detection of COVID-19 using CXR images. If an automated AI/ML model is readily available to health professionals, it will be a substantial improvement in the clinical management of the disease.

2.3.2 Ensemble Model design for COVID-19 Detection

After successfully completed our literature review to design COVID-19 prognosis model design and published our work in a high impact Web of Science journal, then we focused on designing an ensemble model for COVID-19 detection. While researching on deep learning ensemble model design, we investigate many advanced deep neural network model and then study in detail their architectural building blocks, framework and implementation details.

The convolutional neural network(CNN) has more potential in case of classification of complex model design. It is most effectively used for different types of computational problems as well as image classification problems. There are many applications that were implemented in different research papers with the help of deep neural networks. Due to the advancement of different deep neural networks, like ResNet18, ResNet50[65], AlexNet[66], DenseNet[67], VGG16[68], etc., we can make a very sophisticated machine learning model which can detect COVID-19 from CXR, CT images, sometimes a combination of both. Previously, a hand-crafted feature generation technique is used where the heuristic approach was used a lot. Prior to COVID-19, P. Rajpurkar et al.[69], describe a model named as CheXNet where they use a deep neural network to classify 14 kinds of Pneumonia disease. Their model uses over 100,000 frontal view X-ray images with 14 diseases to develop their model. It was a very promising research work that can greatly influence the study of COVID-19 research using ML. L. Wang et al.[70] develop a machine learning model, they name it as COVID-Net using 13,975 CXR images. They had a total of 8,066 patient cases who have no pneumonia i.e. normal cases and 5,538 patient cases who have non-COVID19 pneumonia. They had only 266 COVID-19 patient cases which makes their dataset have data imbalance issues. S. Minace et al.[71], develop their model using 5000 Chest X-rays. They used the transfer learning technique to use the pre-trained weights in their model. The main disadvantage of this kind of transfer learning weight is that the original deep learning network did not actually use a medical dataset, more especially the COVID-19 dataset, so the weights may lead to bias prediction to some extent. A.I. Khan et al.[72], named their model as CoroNet which consist of 310 normal images, 330 Pneumonia Bacterial, 327 Pneumonia Viral, and only

284 COVID-19 sample images. Their research was done in early April 2020 and hence their model was built with such a lower number of sample cases.

2.3.3 COVID-19 Detection using CT images

Most of the literature review we already had completed helped a lot when we were designing the model using Chest Computed Tomography(CT) images only. Here we were need to research on current development on CT image related studies.

H. Panwar et. al [73], used a combination of CXR and CT images. Their research work showed that CT images were also very efficient for designing the model. Later on many research work[74-78] came up as the pandemic progresses where they used CT images to design their prognosis model. We research on what were the shortcomings of those research work and in this regard we found out that EfficientNet[79] performs better than other deep neural network on CT image dataset. EfficientNet is a very efficient high-performance convolutional neural network architecture and scaling method that uniformly scales all dimensions of depth, width and resolution using a compound coefficient. Where benchmark deep neural models arbitrarily scale these factors, the EfficientNet scaling method uniformly scales network width, depth, and resolution with a set of fixed scaling coefficients. The effectiveness of model scaling is based on the baseline network which is playing a key role in high performance. The baseline network builds up with a deep neural architecture search using the AutoML MNAS framework[80], which optimizes both accuracy and efficiency.

2.3.4 COVID-19 Detection using combination of Chest CT and CXR images

As we already gathered a lot of knowledge by developing our different work on COVID-19, we now moved to think of optimizing the model more efficiently and focused on most advanced deep neural network model that come up very recently. The transformer deep neural network models had already proven architecture for the problem related to Natural Language Processing (NLP). While investigating more on their application on different domain we found out that they had great potential for image related application as well, like object detection, segmentation, classification etc.[81-89]. They are generally called as 'Vision Transformer'. While reviewing on these Vision transformer research works, we also investigated problems related to data imbalance issues. The research articles [90-95], shows how data imbalance issues are very crucial while designing machine learning models. The medical dataset, in many cases lacks positive sample cases and if we designed the model with less number of positive samples compared to the bigger number of negative samples then due to data imbalance the weightage values for detection became faulty and do not perform well on test dataset or unknown dataset.

Then we implemented our model using vision transformer, then we extend our research review more on the very recent advancement of this field, especially in the image detection and classification of challenging dataset. Then we study and try to develop our own classification model using "A ConvNet for the 2020s"[96]. Which is the most recent research in this area where advance ResNet and CNN was used.

CHAPTER III - SURVIVAL PREDICTION MODEL FOR RENAL TRANSPLANTATION

Our in this published research work[97], we mainly focused on how to speed up the deep neural network. We used parallelizing in every stage of the development of our parallel implementation model. Automated recommender system designed by artificial intelligence and machine learning method saves lot of time as it is very efficient and easy to implement. It not only improves the medical system but saves a lot of human life. Medical prognosis has become an emerging area in health care research where several reliable prognostic models based on survival analysis procedures have been useful to a variety of domains, with different degrees of success [98]. With the invent of new bioinformatics technology and medical engineering the success of organ transplantation model and survival improve a lot. In this paper, an effective and accurate Parallel Phase Neural network model (PHNNM) has been proposed for the prediction of the long-term survival of liver patients who undergo liver transplantation (LT). Any medical prognosis system needs a lot of data analytics systems and historical observational data from medical hospitals and research laboratories. Efficient and accurate medical prognosis systems is always in demand but most of the current systems rely on serial implementations which lacks performance and speeds.

3.1 Methodology

A. Data Source

Our dataset was collected from Scientific Registry of Transplant Recipients(SRTR), USA which collected the data from United Network for Organ Sharing (UNOS) registry, USA.

B. Data Preprocessing and final Data Description

After receiving the dataset, a whole lot of preprocessing need to be done to use for a data analytics model. We first corrected the missing values and adjusted the outlier's data. The continuous features are transformed using the MaxMinScaler(), the numerical features are scaled by z-score normalization and the categorial features are transformed using One-hot encoding to scale to the range [0, 1]. Keep in mind that now both our categorical features, numerical features and continuous features are all in the range [0, 1]. Then we'll concatenate categorical features, numerical features with continuous features using Python's NumPy's hstack function. Then the data is then split into training and test data. The final data consists of the following things,

1) The total number of instances used = 8272
2) The total number of attributes used = 61

There are 29 categorical attribute and 10 numerical attribute, 22 continuous attribute was used in this research work.

These are as follows,

i) Categorial: 29 ('DON_GENDER', 'DON_RACE_SRTR', 'DON_ETHNICITY_SRTR', 'DON_ABO', 'DON_ANTI_CMV', 'DON_ANTI_HCV', 'DON_INOTROP_AGENT_GE3', 'DON_NON_HR_BEAT', 'CAN_GENDER', 'CAN_RACE_SRTR', 'CAN_ETHNICITY_SRTR', 'CAN_ABO', 'REC_TX_EXTRA_VESSEL', 'REC_CMV_IGG', 'REC_CMV_IGM', 'REC_CMV_STAT', 'REC_HBV_ANTIBODY', 'REC_HBV_SURF_ANTIGEN', 'REC_HCV_STAT', 'REC_LIFE_SUPPORT', 'REC_PORTAL_VEIN',

'REC_TIPSS', 'CAN_DRUG_TREAT_HYPERTEN',
'CAN_PERIPH_VASC', 'CAN_DRUG_TREAT_COPD',
'CAN_PULM_EMBOL', 'CAN_BACTERIA_PERIT',
'CAN_PORTAL_VEIN', 'CAN_TIPSS')

ii) Numerical: 10 ('TARGET_VALUE', 'DON_HGT_CM', 'DON_WGT_KG', 'DON_CREAT', 'DON_SGOT', 'REC_HGT_CM', 'REC_WGT_KG', 'REC_BMI', 'CAN_HGT_CM', 'CAN_WGT_KG')

iii) Continuous: 22 ('DON_AGE', 'DON_HIST_DIAB', 'DON_HIST_HYPERTEN', 'DON_A1', 'DON_A2', 'DON_B1', 'DON_B2', 'DON_DR1', 'DON_DR2', 'DON_HIGH_CREAT', 'DON_HTN', 'DON_COD_DON_STROKE', 'REC_DGN', 'DON_ORG_SHARED', 'REC_VENTILATOR', 'REC_LIFE_SUPPORT_OTHER', 'REC_DGN2', 'REC_AGE_AT_TX', 'CAN_DGN', 'CAN_DIAB_TY', 'CAN_AGE_AT_LISTING', 'CAN_PREV_LI')

iv) As we eliminate all missing attribute values so there is no missing value for attribute we fed to the system.

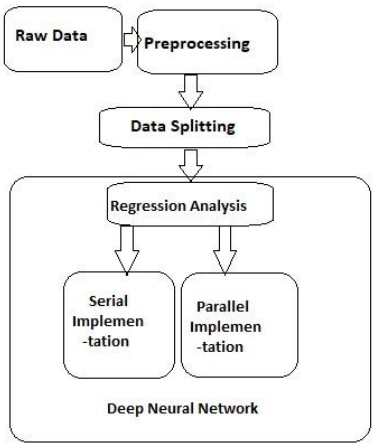

Figure 3.1 *Block Diagram of Proposed Approach*

C. Model Description

Keras[99] is an Application Programming Interface(API) mainly used for designing neural network. The model runs on top of TensorFlow In this paper, a neural network is built in Keras to solve the regression issue, i.e. one where our dependent variable ,Survival of the patient(y) is in numerical format and we are trying to predict the quantity of y with as much accuracy as possible. We done Random forest regression analysis. We train the neural network using the 135 input variables, along with 4 hidden layers of 32, 16, 8 and 4 neurons respectively, and finally using the linear activation function to process the output. Here this paper uses Mean Squared Error (MSE) and Mean Absolute Error (MAE) as our loss functions. Validation split set to 0.2, that is 80% training data and 20% test data. We have specified 200 epochs for our model. This means that we are essentially training our model over 200 forward and backward passes, with the expectation that our loss will decrease with each epoch, meaning that our model is predicting the value of y more accurately as we continue to train the model. parallel implementation of our model using CUDA also introduced in our implementation[100].

3.2 Results

Both the training and validation loss decrease in an exponential fashion as the number of epochs is increased, suggesting that the model gains a high degree of accuracy as our epochs is increased and then reached the saturation point.

A. Serial Implementation Result

Table 3.1 Serial Performance

Mean Squared Error(MSE)	Mean Absolute Error(MAE)	Test Variance Value
12.62	2.71	0.11

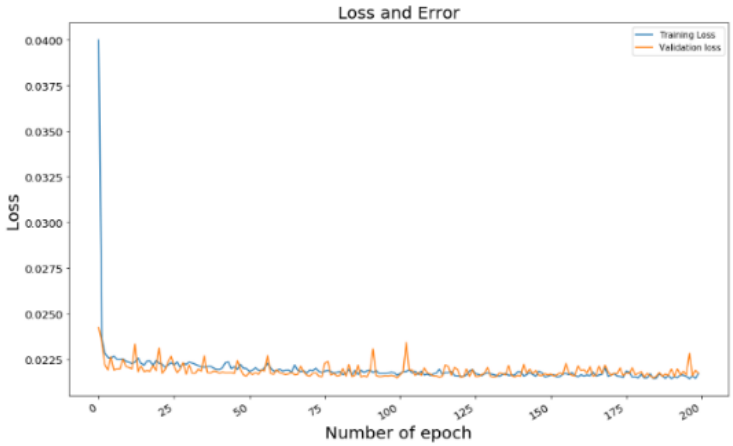

Figure 3.2 *Training and Validation Loss over increasing epoch.*

B. Serial result using Random Forest Regression

Random Forest Regressors(RFR) is the way for splitting criterion to measure the quality of a split. Supported criteria are “MSE” for the mean squared error, which is equal to variance reduction as feature selection criterion, and “Mean Absolute

Error" for the mean absolute error. We have applied same parameter for random forest regressor.

Table 3.2 Random Forest Result

Coefficient of determination R^2 of the prediction	Mean Squared Error(MSE)	Test Variance value
0.8693909563710034	0.02	0.06

C. Parallel implementation Result

The computational time improves a lot in parallel implementation, through MSE and MAE nearly same with the serial implementation.

Table 3.3 Illustrate Runtime with Different Batch Size and epoch

No. of Epoch	No. of CPU Core	Runtime				
		Batch Size = 16	Batch Size =32	Batch Size =64	Batch Size =128	Batch Size=256
50	1	74.098	42.022	23.251	15.090	10.059
	4	72.270	38.715	22.787	13.490	10.013
	8	71.276	38.154	21.932	13.435	9.641
	12	71.184	38.768	21.795	13.417	9.450
	16	70.686	37.587	21.269	12.932	9.290
	20	70.498	37.580	20.999	12.814	9.184
200	1	716.048	376.748	205.262	124.256	85.568
	4	705.525	370.102	198.557	122.077	82.892
	8	703.625	367.999	197.342	118.512	82.779
	12	697.447	366.161	195.877	117.644	77.748

	16	696.682	364.096	195.787	114.996	77.399
	20	692.742	361.660	194.612	114.445	73.100
500	1	716.048	376.748	205.262	124.256	85.568
	4	705.525	370.102	198.557	122.077	82.892
	8	703.625	367.999	197.342	118.512	82.779
	12	697.447	366.161	195.877	117.644	77.748
	16	696.682	364.096	195.787	114.996	77.399
	20	692.742	61.660	194.612	114.445	73.100

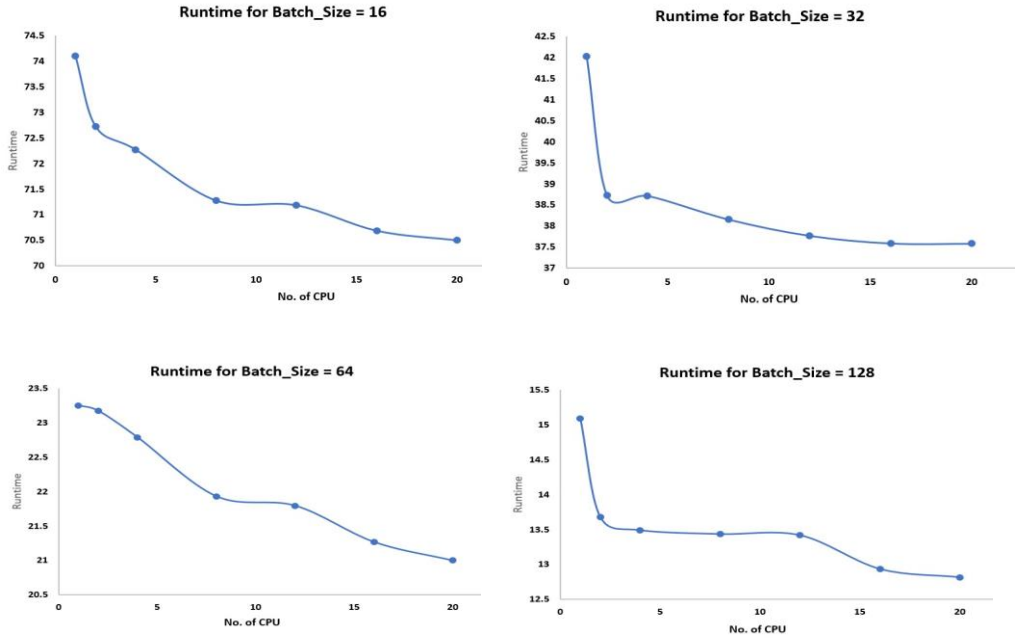

Figure 3.3 *Runtime over Epoch = 50 with different batch sizes*

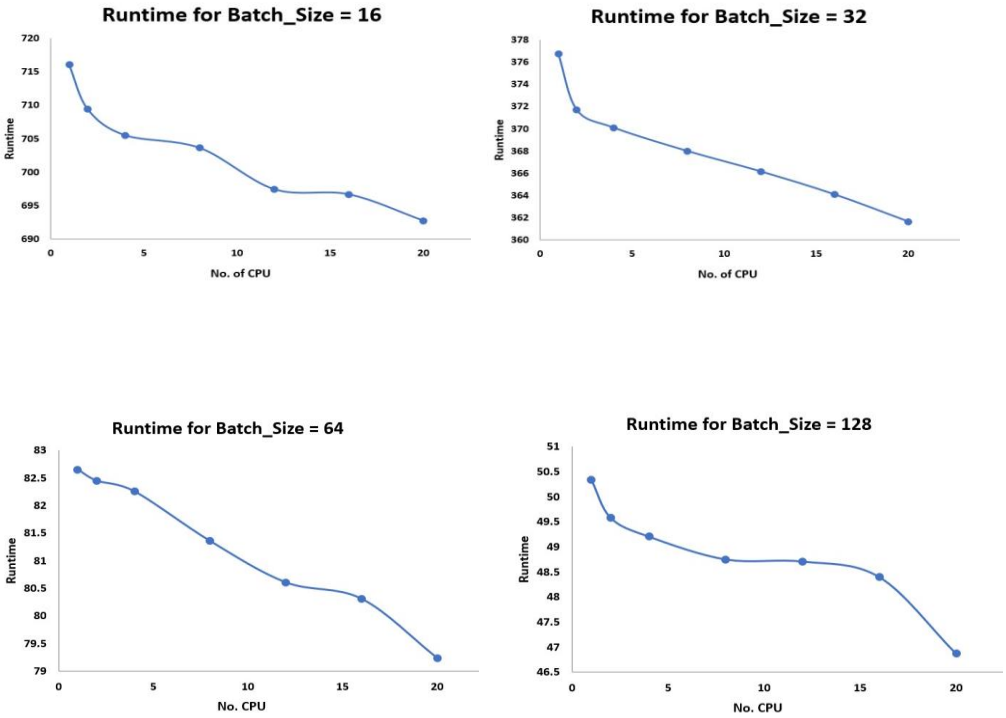

Figure 3.4 Runtime over Epoch = 200 with different batch sizes.

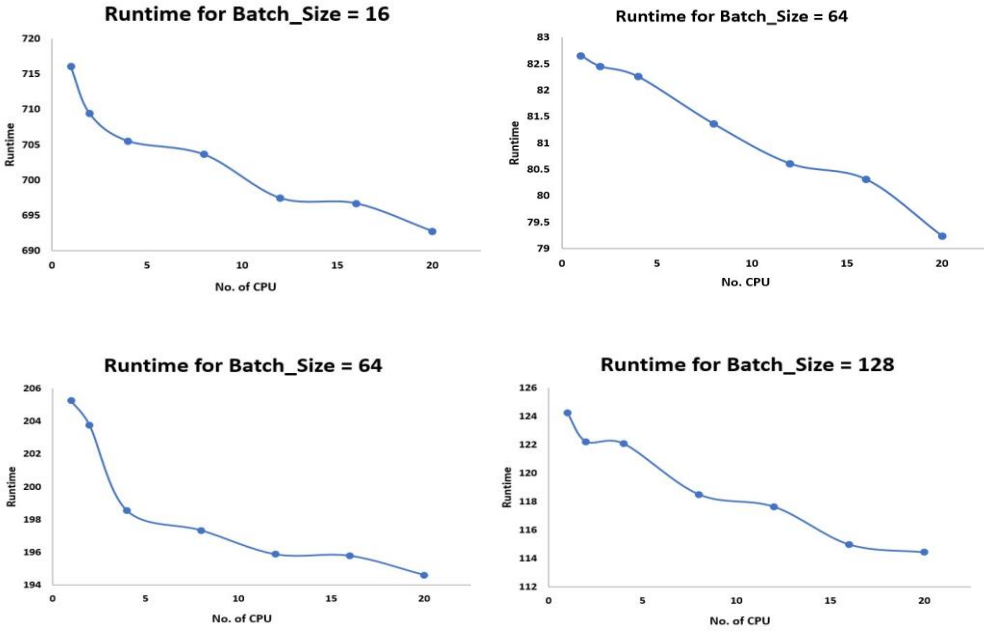

Figure 3.5 *Runtime over Epoch = 500 with different batch sizes*

The above figures, 3.5, 3.6, 3.7 illustrate how runtime decreases over the increase of batch size as well as increase of number of epochs. These show the inversely proportional relationship as we increase batch size the runtime decreases. At higher values of CPU, we received an approximate 2x decrease in computation time. This is what we expected, as we were spreading the work over 20 cores.

Table 3.4 Illustrates Speedup with batch size and epochs

#Epoch	#CPU			Speedup		
		Batch Size = 16	Batch Size = 32	Batch Size = 64	Batch Size = 128	Batch Size = 256
	2	1.018872531	1.085209345	1.003374684	1.103234533	1.002865208
	4	1.02528877	1.085413561	1.020361035	1.118619812	1.004560947
	8	1.039585507	1.101388052	1.060166207	1.123239456	1.043324605
50	12	1.040931751	1.112649311	1.066792729	1.124698449	1.058879504
	16	1.048264296	1.118003005	1.093202151	1.166864258	1.082835449
	20	1.051063833	1.11820142	1.107243846	1.177675261	1.095267443
	2	1.034900764	1.012438642	1.002435047	1.015201915	1.252254939
	4	1.038565488	1.012529357	1.004807716	1.022954555	1.329733023
	8	1.039412193	1.014973298	1.015902556	1.032491296	1.364403162
200	12	1.041884447	1.021144269	1.025328093	1.033360859	1.37813022
	16	1.059586139	1.043212523	1.029110537	1.040007666	1.397060562
	20	1.059803539	1.04350392	1.043024557	1.073935364	1.430054563
	2	1.009392495	1.013562475	1.007332236	1.016686305	1.011191557
	4	1.014914844	1.017957405	1.033769142	1.017843871	1.032286682
	8	1.017655584	1.023773591	1.040136045	1.048462678	1.033693981
500	12	1.026670401	1.028914117	1.047915642	1.056204112	1.10058289
	16	1.027796694	1.034749502	1.048394931	1.080526567	1.105545257
	20	1.033643117	1.041720056	1.054728064	1.085719938	1.170564799

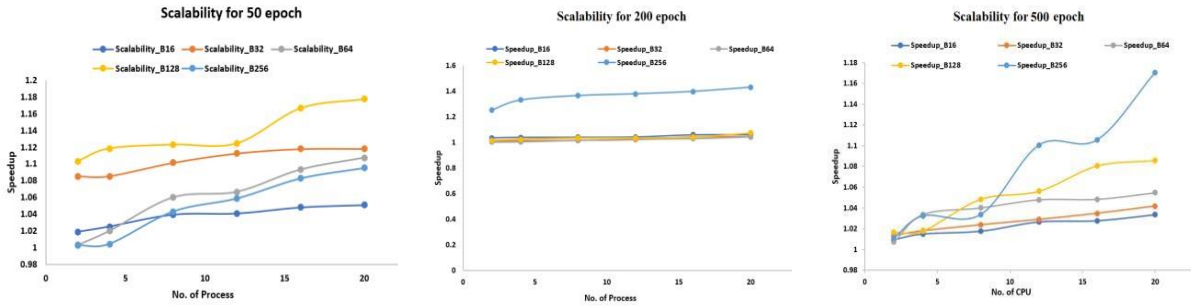

Figure 3.6 *The speedup/scalability over different batch sizes and epochs.*

Figure 17 shows the scalability of our proposed parallelism technique on CPU algorithm. Based on the result we see that speedup is increasing with the increase of number of processing elements. This technique is validated on several number of epochs and the effect of the increasing in speedup is envisioned similarly that proves that the system is scalable. In case of epoch 200, the performance doesn't make so differences except Batch size = 256.

Table 3.5 Illustrates efficiency over different batch sizes and epochs

#Epoch	#CPU	Efficiency				
		Batch Size = 16	Batch Size = 32	Batch Size = 64	Batch Size = 128	Batch Size = 256
50	2	0.509436266	0.542604673	0.501687342	0.551617267	0.501432604
	4	0.256322193	0.27135339	0.255090259	0.279654953	0.251140237
	8	0.129948188	0.137673507	0.132520776	0.140404932	0.130415576
	12	0.086744313	0.092720776	0.088899394	0.093724871	0.088239959
	16	0.065516519	0.069875188	0.068325134	0.072929016	0.067677216
	20	0.052553192	0.055910071	0.055362192	0.058883763	0.054763372
200	2	0.517450382	0.506219321	0.501217523	0.507600958	0.62612747
	4	0.259641372	0.253132339	0.251201929	0.255738639	0.332433256
	8	0.129926524	0.126871662	0.12698782	0.129061412	0.170550395

	12	0.086823704	0.085095356	0.085444008	0.086113405	0.114844185
	16	0.066224134	0.065200783	0.064319409	0.065000479	0.087316285
	20	0.052990177	0.052175196	0.052151228	0.053696768	0.071502728
500	2	0.509436266	0.542604673	0.501687342	0.551617267	0.501432604
	4	0.256322193	0.27135339	0.255090259	0.279654953	0.251140237
	8	0.129948188	0.137673507	0.132520776	0.140404932	0.130415576
	12	0.086744313	0.092720776	0.088899394	0.093724871	0.088239959
	16	0.065516519	0.069875188	0.068325134	0.072929016	0.067677216
	20	0.052553192	0.055910071	0.055362192	0.058883763	0.054763372

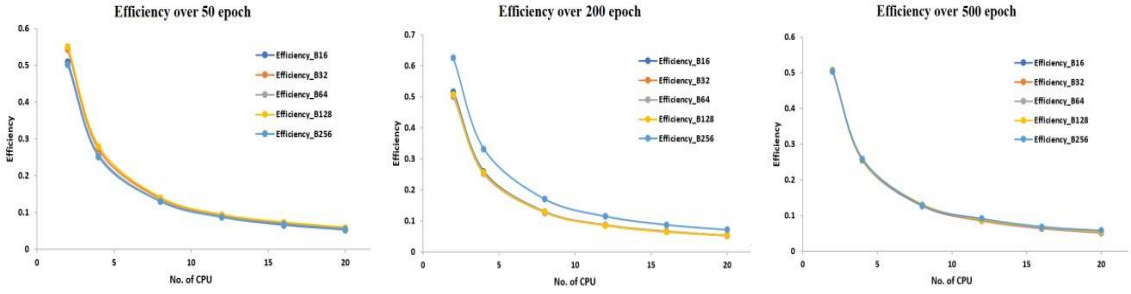

Figure 3.7 *Illustrates the efficiency with increase batch size and epochs.*

CHAPTER IV - MEDICAL PROGNOSIS SYSTEM FOR BREAST CANCER

This chapter will describe our work related to medical prognosis. Initially, we tried to investigate the dataset with different clustering techniques which fall under the unsupervised machine learning methodology. Then we developed the full-fledged medical prognosis system for breast cancer. The research works presented in [101,102] are divided into two portions, one is clustering analysis and another one is the medical prognosis system design. The following sections will describe both the approaches.

4.1 Unsupervised learning exploration using breast cancer dataset

Cancer is a group of diseases involving abnormal cell growth with the potential to invade or spread to other parts of the body.

According to US Federal Government's principal agency, The National Cancer Institute (NCI), Cancer is the name given to a collection of related diseases. In all types of cancer, some of the body's cells begin to divide without stopping and spread into surrounding tissues.

Cancer is the second leading cause of death in the United States, exceeded only by heart disease. One of every four deaths in the United States is due to cancer.

Breast cancer is one of the major kinds of cancer that develops from breast tissue. Signs of breast cancer may include a lump in the breast, a change in breast shape, dimpling of the skin, fluid coming from the nipple, a newly inverted nipple, or a red or scaly patch of skin.

According to the US's leading national public health institute Centers for Disease Control and Prevention (CDC), Breast cancer is the second most common cancer among women in the United States.

Clustering techniques can generally be divided into two areas, one is hard clustering where the dataset can be clearly divided into two or more parts. We have a clear viewpoint whether a data point belongs to a particular cluster or not. For soft clustering, instead of setting each data point into a distinct cluster, a probability or likelihood of that data point to be in those clusters is given.

4.1.1 Methodology

A. Dataset Description and Pre-Processing

Table 4.1 Details of Attributes of Breast Cancer Data

No. of Attributes	Domain
1. Sample code number	Id Number
2. Clump Thickness	1-10
3. Uniformity of Cell Size	1-10
4. Uniformity of Cell Shape	1-10
5. Marginal Adhesion	1-10
6.Single Epithelial Cell Size	1-10
7.Bare Nuclei	1-10
8. Bland Chromatin	1-10
9.Normal Nucleoli	1-10
10. Mitoses	1-10
11. Class	2 for benign, 4 for malignant

We preprocess the data using normal preprocessing techniques like eliminating the null values, using soft-encoding we set the parameter values within the range of 1-10 etc. The data is downloaded in .csv format. In the csv extension, we delete all the 16 rows that had the missing values. After that the csv extension was converted to a Weka friendly file

which is Attribute Relation Data Format (ARFF)extension. Then the data goes through a series of preprocessing steps.

No. of Instances in the dataset before Preprocessing = 699

No. of Instances in the dataset after Preprocessing = 683

Then we use Weka software to further normalize the dataset using min-max normalization method so that all the feature values are in [0, 1]. Clustering is an unsupervised learning method, so only feature values are used for clustering. This means we do not normalize the category label which is the last column in the dataset.

First, we remove the ID number. Then we use Weka to further preprocess the dataset. Then we investigate the dataset to understand correlation among the datapoints using Hopkins Statistic Index to understand whether the dataset has good clustering tendency or not. Then we perform different sort of clustering technique both using Weka Software Tools and R programing language.

4.1.2 Results

In our Breast Cancer data, we got Hopkins Statistic Index = 0.7997314 which indicates the dataset is highly clustered.

K-means Clustering Analysis Results:

Table 4.2 Cluster Instances after K-means

Cluster Instances		
Benign	402	59%
Malignant	281	41%

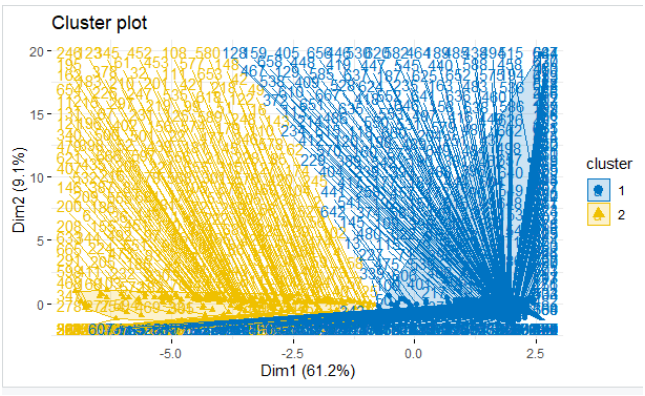

Figure 4.1 *K-means cluster visualization*

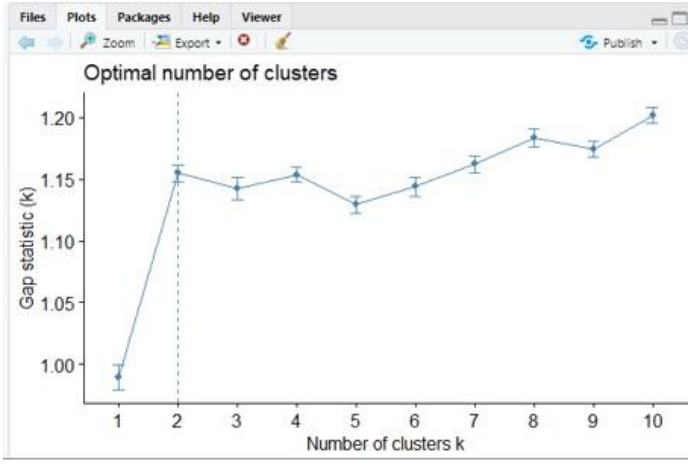

Figure 4.2 *Box plot clearly illustrate the dataset has two optimal clusters.*

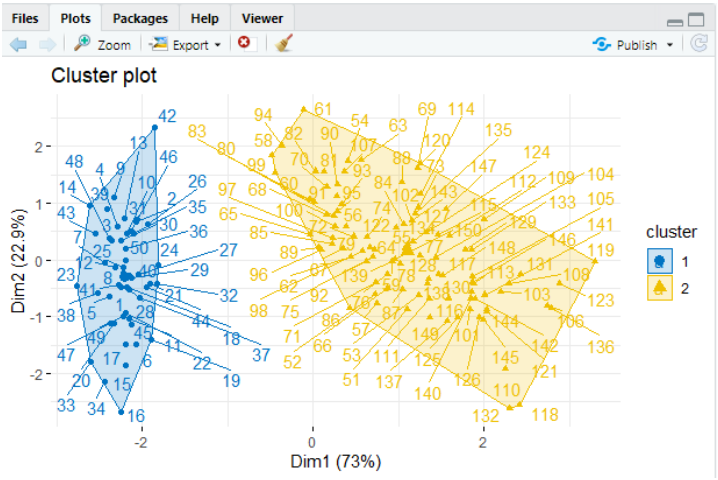

Figure 4.3 *K-means clustering with varying parameter values*

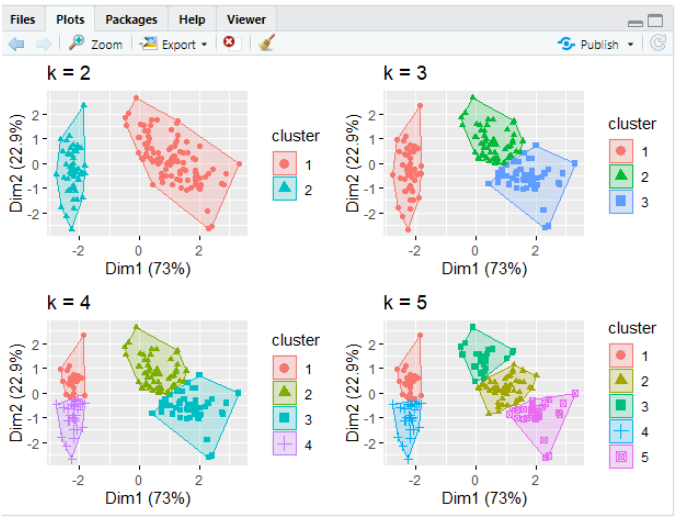

Figure 4.4 *K-means for different values of k*

K-medoids clustering or Partitioning Around Medoids (PAM) results:

An alternative to k-means clustering is the K-medoids clustering or Partitioning Around Medoids(PAM) [103], which is less sensitive to outliers compared to k-means clustering methods.

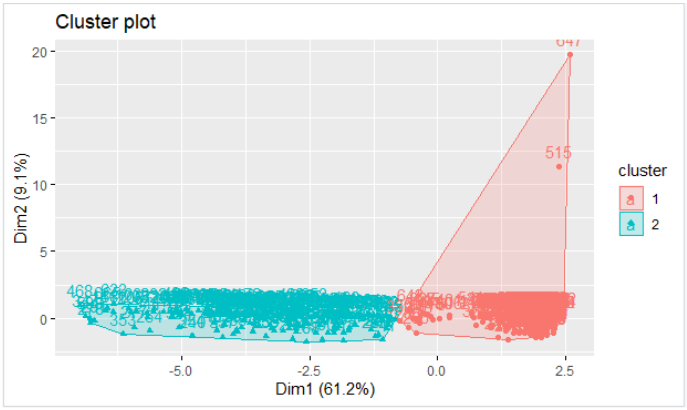

Figure 4.5 *K-medoids clustering or Partitioning Around Medoids(PAM)*

Silhouette Analysis Using R on Breast Cancer Dataset:

Silhouette analysis allows you to calculate how similar each observation is with the cluster it is assigned relative to other clusters. This metric (silhouette width) ranges from -1 to 1 for each observation in your data and can be interpreted as follows:

i) Values close to 1 suggest that the observation is well matched to the assigned cluster.

ii) Values close to 0 suggest that the observation is borderline matched between two clusters.

iii) Values close to -1 suggest that the observations may be assigned to the wrong cluster.

In our research we got the metric (silhouette width) value = 0.57 which clearly shows that the Breast Cancer Dataset well matched to the assigned cluster.

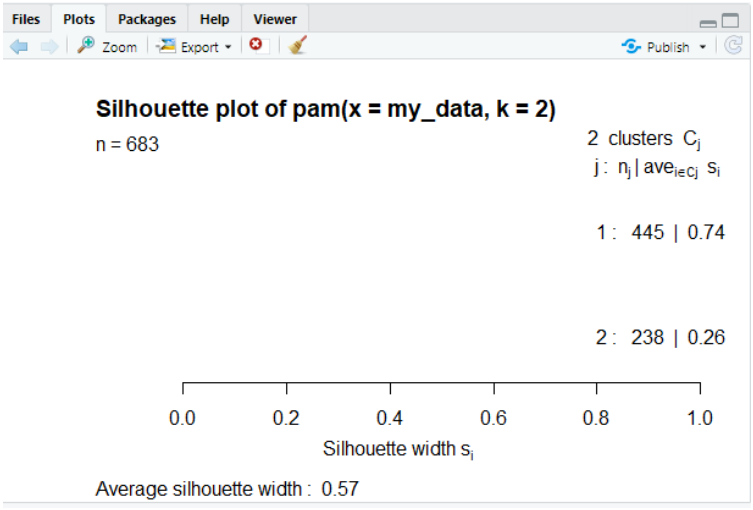

Figure 4.6 *Silhouette Plot of K-medoids or PAM*

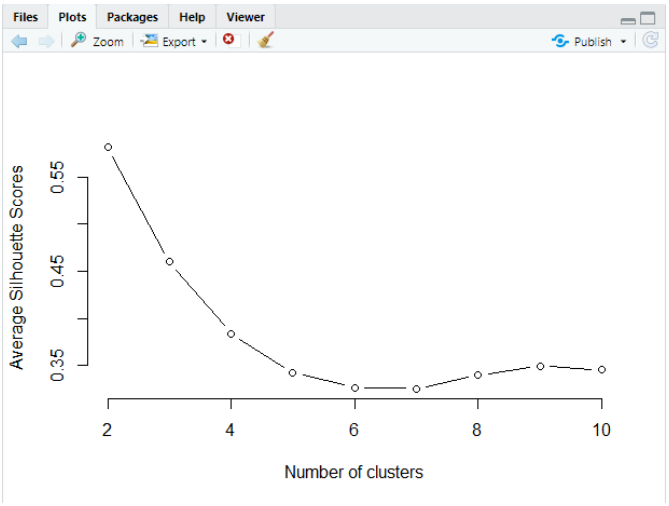

Figure 4.7 *Average Silhouette Score plot of K-means clustering.*

Average Silhouette Score plot of K-means clustering where K varies from 2 to 10 and the Highest Silhouette value is 0.582 which shows that the Breast Cancer Dataset well matched to the assigned cluster, when cluster size =2.

4.2 Medical Prognosis System Design for Breast Cancer

This section would elaborate the work[102] for designing medical prognosis system design for breast cancer.

4.2.1 Methodology

Weka Software tool

Waikato Environment for Knowledge Analysis (Weka), developed at the University of Waikato, New Zealand. It is free software licensed under the GNU General Public License, and the companion software to the book "Data Mining: Practical Machine Learning Tools and Techniques". It is mainly used for Machine learning, Data Analysis and Data Visualization purposes.

The Weka version that is used in analyzing this breast cancer data and predict whether a patient has breast cancer or not is 3.8.4. The following discussions shows how the methods was used in Weka.

We use the same dataset by which we work for clustering technique, so this section will not describe again the dataset and preprocessing part also the clustering technique, rather in this section will describe more about supervised methods and the whole process of designing the model. The different classification methodologies we have used are as follows,

A. Random Forest(RF)

One of the most interesting supervised learning technique is the Random Forest[104]. It actually leverage the power of decision trees and it ensembles multiple decision tress and chose the optimum value for decision making. It also unlock the potential of bagging method which also combine multiple learning methods to optimize the overall results. Random forest essentially constructs a number of decision trees and merges them to get a more accurate, robust and balanced prediction. The parameters used in the random forest method in Weka are:

i. The size of each bag. (i.e., P)
ii. Calculate the out of error bag. (i.e., O)
iii. Number of iterations or number of trees in random forest. (i.e., I)
iv. Minimum number of instances. (i.e., M)
v. Minimum variance of split. (i.e., V)
vi. Seed for random number generator. (i.e., S)
vii. Depth of the tree.

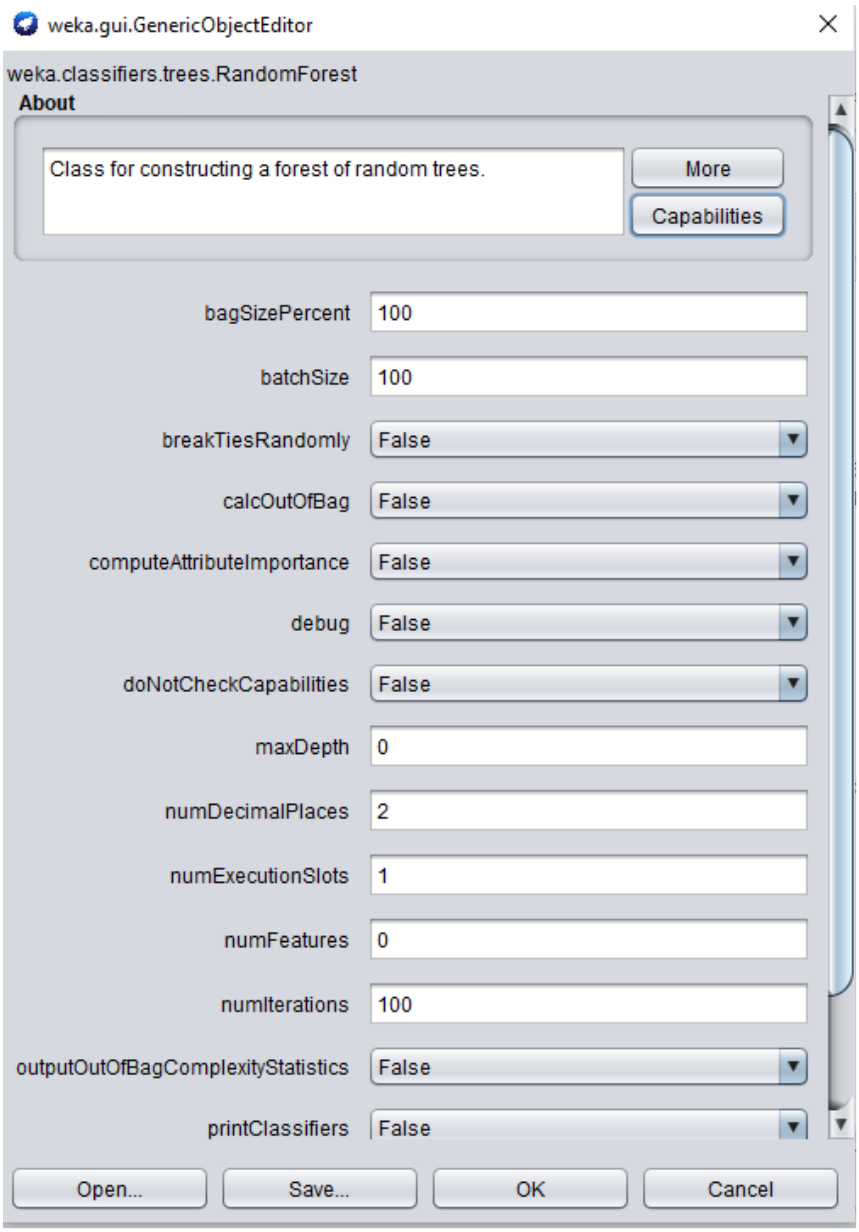

Figure 4.8 *Weka interface for parameter tuning for Random Forest(RF)*

B. Support Vector Machine(SVM)

Support Vector Machine(SVM)[104], is a very powerful and widely used supervised learning technique where the algorithm try to find a hyperplane in an N dimensional space. Here N is the total number of the features used in that specific dataset. Hence, in this approach SVM able to distinctly segregate the datapoints, most of the cases very

successfully. SVM can perform well on linear classification problems as well as non-linear classification problem. The Kernel function that was used in Weka for our dataset is Radial Basis Function(RBF). RBF mostly used Euclidian Distance most of the cases to calculate the distance of the datapoint to the hyperplane where mostly source or center of the datapoint is considered.

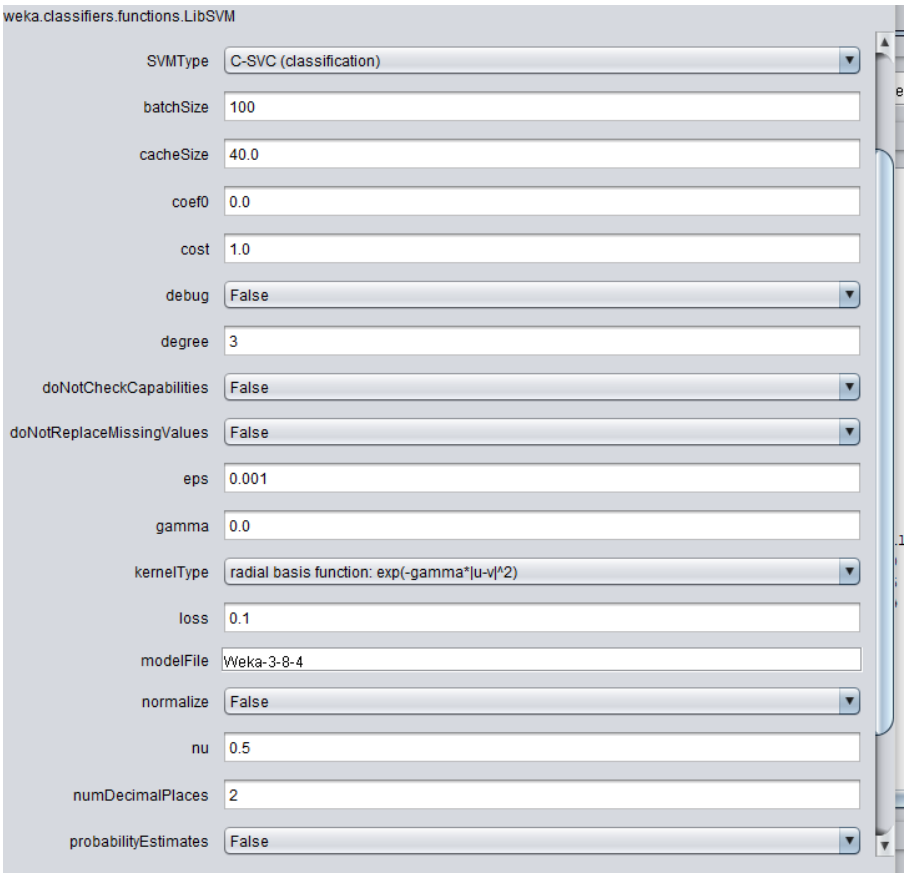

Figure 4.9 *Weka interface for Support vector Machine(SVM)*

C. Multi-layer perceptron(MLP)

Multilayer perceptron is a deep artificial neural network, and it is composed of more than one perceptron. They are made of an input layer, one or more hidden layers and an output layer that makes a prediction about the input. Each layer is made up of units. The input to the network corresponds to the attributes measured for each training tuple. Multilayer

perceptron is often applied to supervised learning problems. Fig. 4.2.1. shows the multilayer feedforward neural network for our Breast Cancer Dataset. The input layer consists of 10 attributes values, 2 hidden layers, 1st hidden layer having 4 nodes and 2nd hidden layer having 6 nodes and output layer have 2 nodes for 2 class, 2 for benign and 4 for malignant. We also heuristically change the number of hidden layers and other parameters to get the optimum parameter settings.

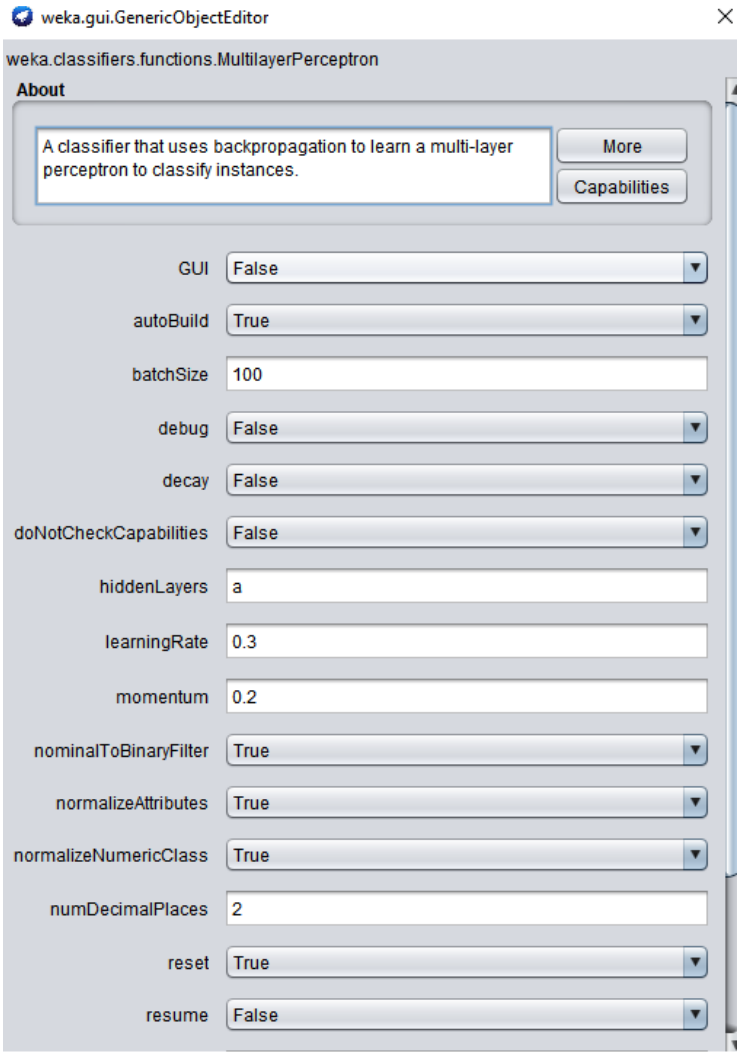

Figure 4.10 *Weka interface for multi-layer perceptron*

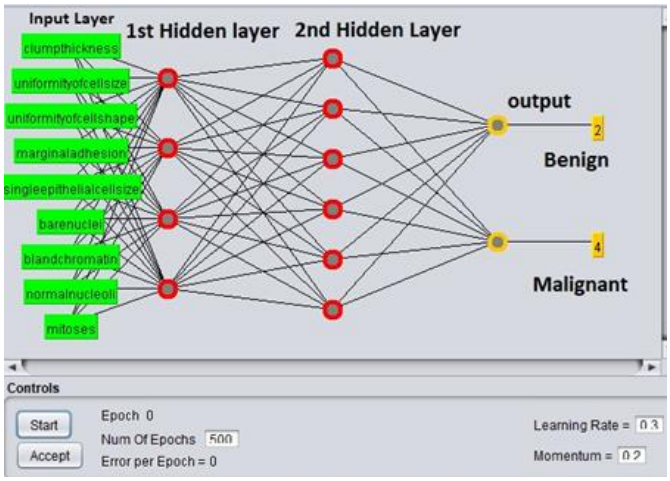

Figure 4.11 *Multilayer Feedforward Neural Network for our Breast Cancer Dataset.*

The following are some of the parameters that was used in Weka on our dataset.

i. Learning rate. (i.e., L)

ii. Momentum rate. (i.e., M)

iii. Number of epochs.(i.e., N)

iv. Percentage size of validation set.(i.e., V)

v. The value use to seed the random number generator. (i.e., S)

vi. The number of consecutive increases of error allowed for validation testing before training terminates.(i.e., E)

The system also performs cross-validation and parameter optimization.

D. C4.5 Algorithm or J48

C4.5 is a specialized decision tree classifier in the domain of supervised learning algorithms which are mostly used in data mining domain and have a huge potential to use decision tree algorithms to make proper classification decisions.

C4.5 was invented by J.R. Quinlan[105] in 1993. Then onwards it was used in many domains of decision making problem solving and classification tasks. J48 is named coined for open-source Java implementation of C4.5 algorithm in Weka machine learning tool. The following figure 4.2.2 shows the classification in Weka system using J48 algorithm.

4.2.2 Results

The results were very promising for this work. This work received best paper award in an International Conference held in 2021.

After heuristically run-through many types of cross validation the optimum value settle with 10 fold cross-validation. The following table shows the results with all the techniques discussed above.

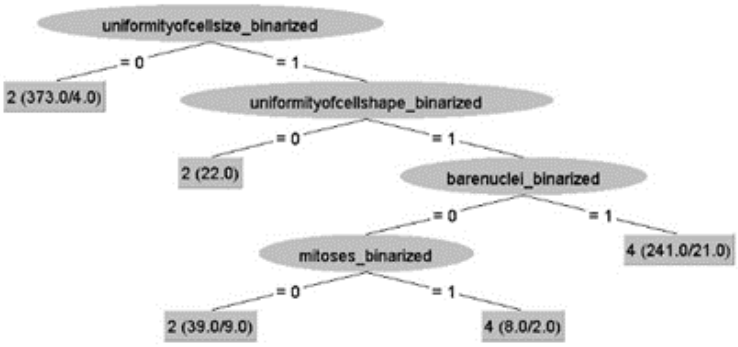

Figure 4.12 *Generating Decision Tree using J48 for Breast cancer dataset*

Table 4.3 Comparing All Methods using 10-Fold Cross Validation.

	Random Forest (RF)	Support Vector Machine (SVM)	Multi-Layer Perceptron (MLP)	J48
Correctly Classified Instances	97.2182 %	96.4861 %	96.0469 %	93.4114%
Incorrectly Classified Instances	2.7818 %	3.5139 %	3.9531 %	6.5886 %
Kappa Statistics	0.939	0.9241	0.914	0.8564
Mean Absolute Error	0.064	0.0351	0.0404	0.1008
Root Mean square error	0.1615	0.1875	0.179	0.2384
Relative Absolute Error	14.0555 %	7.7211 %	8.8702 %	22.1551 %
Root Relative Square Error	33.8557 %	39.3026 %	37.5253 %	49.9826 %
Total Number of Instances	683	683	683	683

The results showing in the table clearly prove that Random Forest(RF) performs better for our breast cancer dataset. then Support Vector Machine, Multi-layer Perceptron(MLP) and J48 values are rank respectively. The automated ensemble breast cancer prognosis model provides the optimum value for any kinds of test dataset.

a) Results for Random Forest(RF) classifier

After running our data through Weka, the following table and figures were the outputs we got. Table 2 gives the confusion matrix of the 137 instance that was used to test the model. It shows clearly that our model was able to correctly predict 99 instances to be benign out of 102 which were benign. It was able to predict 33 patients who were malignant out of the 35 who were malignant. From Figure 6, 98% of the tuples was

labeled positive by our classifier out of the actual positive (i.e. precision), 97.1% of the positive tuples is the percentage our classifier labeled as positive (i.e. recall), the harmonic mean of precision and recall is 0.975 (i.e. F-measure). This value is very close to 1 which means the models accuracy is better. Again, in Figure 6, the true positive recognition rate is 0.971 (i.e. Sensitivity=TP/P) and the true negative recognition rate is 0.943 (i.e. Specificity=TN/N). Figure 8 and 9 shows the ROC curves for benign (2) and Malignant (4) respectively, the Area Under these Curves i.e. AUROCs, are 0.992. This means our model has almost a perfect accuracy because is close to 1.

Table 4.4 Confusion Matrix for Random Forest(RF)

Actual\Predicted	Benign	Malignant	Total
Benign	99	3	102
Malignant	2	33	35
Total	101	36	137

```
=== Evaluation on test split ===

Time taken to test model on test split: 0.01 seconds

=== Summary ===

Correctly Classified Instances         132               96.3504 %
Incorrectly Classified Instances         5                3.6496 %
Kappa statistic                          0.905
Mean absolute error                      0.0577
Root mean squared error                  0.1681
Relative absolute error                 13.1693 %
Root relative squared error             37.1976 %
Total Number of Instances              137

=== Detailed Accuracy By Class ===

               TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
               0.971    0.057    0.980      0.971    0.975      0.905    0.992     0.997     2
               0.943    0.029    0.917      0.943    0.930      0.905    0.992     0.983     4
Weighted Avg.  0.964    0.050    0.964      0.964    0.964      0.905    0.992     0.993

=== Confusion Matrix ===

  a  b   <-- classified as
 99  3 |  a = 2
  2 33 |  b = 4
```

Figure 4.13 *Weka output figure for Random forest*

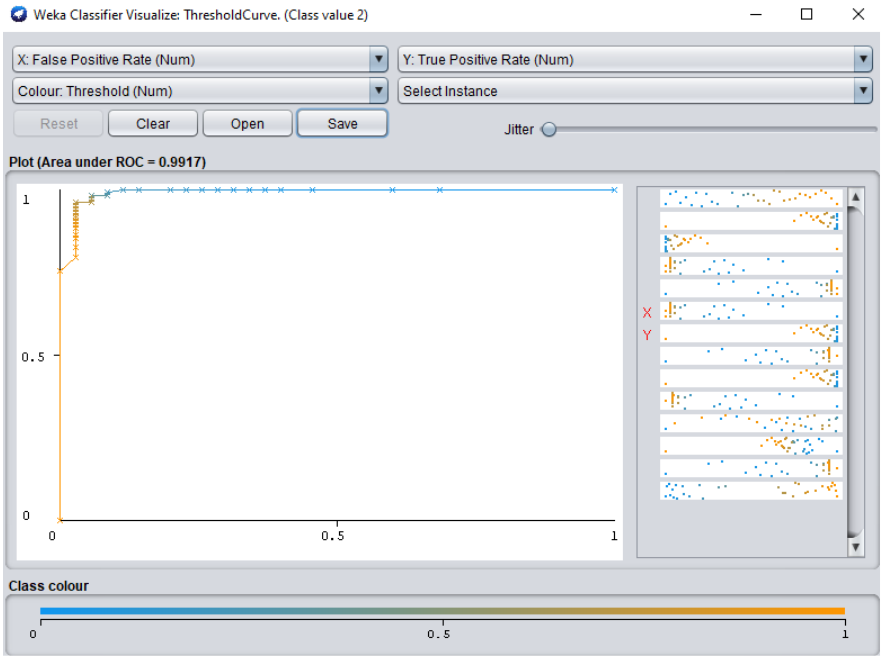

Figure 4.14 *Receiver Operating Characteristic curve(ROC) using RF for class 2*

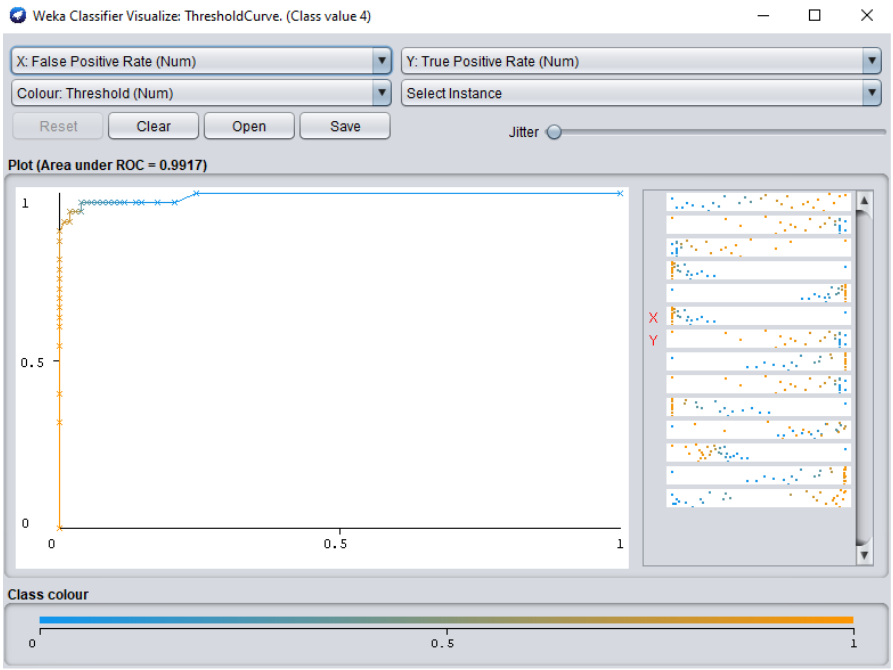

Figure 4.15 *Receiver Operating Characteristic curve(ROC) using RF for class 4*

b) Results for SVM Classifier

When we did run our data through Weka, the following table and figures were the outputs we got. Table 3 gives the confusion matrix of the 137 instance that was used to test the model. It shows vividly that our model was able to correctly predict 97 instances to be benign out of 102 which were benign. It was able to predict 34 patients who were malignant out of the 35 who were malignant. From Figure 9, 99% of the tuples was labeled positive by our classifier out of the actual positive (i.e. precision), 95.1% of the positive tuples is the percentage our classifier labeled as positive (i.e. recall), the harmonic mean of precision and recall is 0.97 (i.e. F-measure). This value is very close to 1 which means the models accuracy is better. Figure 9, the true positive recognition rate is 0.951 (i.e. Sensitivity=TP/P) and the true negative recognition rate is 0.971 (i.e. Specificity=TN/N). Figure 4.2.9 and Figure 4.2.10 exhibits the ROC curves for benign (2) and Malignant (4) respectively, the area under these curves (i.e. AUROCs) are 0.961. This means our model has a better accuracy because is close to 1.

Table 4.5 Confusion Matrix for SVM

Actual\Predicted	Benign	Malignant	Total
Benign	97	5	102
Malignant	1	34	35
Total	98	39	137

```
=== Evaluation on test split ===

Time taken to test model on test split: 0.03 seconds

=== Summary ===

Correctly Classified Instances         131               95.6204 %
Incorrectly Classified Instances         6                4.3796 %
Kappa statistic                          0.889
Mean absolute error                      0.0438
Root mean squared error                  0.2093
Relative absolute error                  9.9894 %
Root relative squared error             46.3027 %
Total Number of Instances              137

=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall   F-Measure  MCC      ROC Area  PRC Area  Class
                 0.951    0.029    0.990      0.951    0.970      0.891    0.961     0.978     2
                 0.971    0.049    0.872      0.971    0.919      0.891    0.961     0.854     4
Weighted Avg.    0.956    0.034    0.960      0.956    0.957      0.891    0.961     0.946

=== Confusion Matrix ===

  a  b   <-- classified as
 97  5 |  a = 2
  1 34 |  b = 4
```

Figure 4.16 *Weka output figure for SVM*

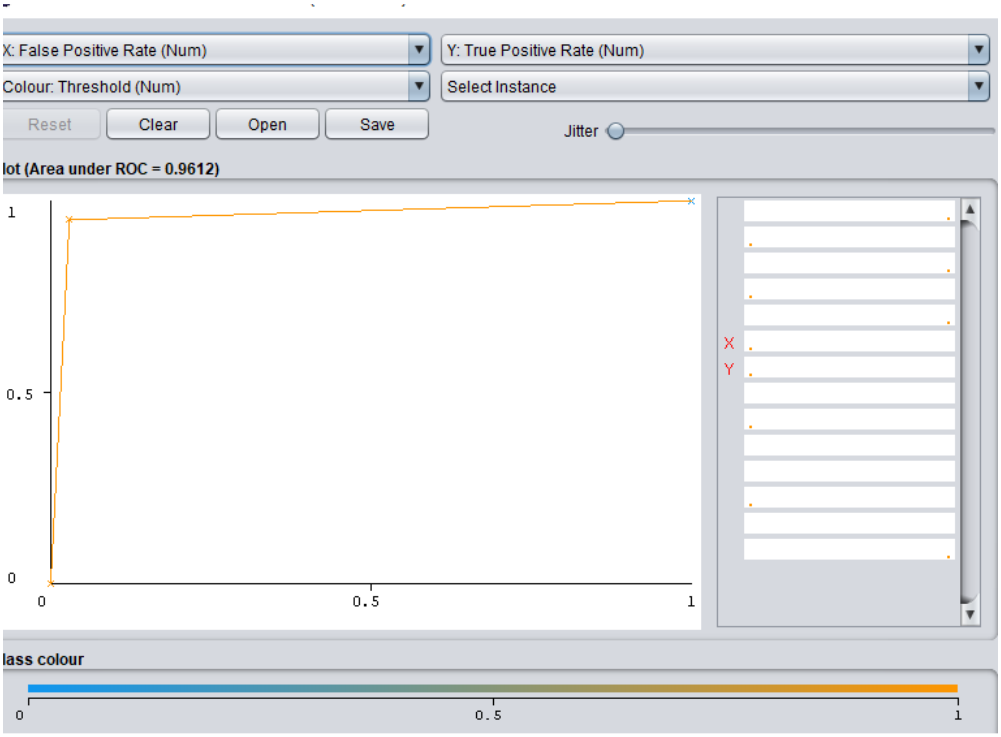

Figure 4.17 *Receiver Operating Characteristic curve(ROC) using SVM for class 2*

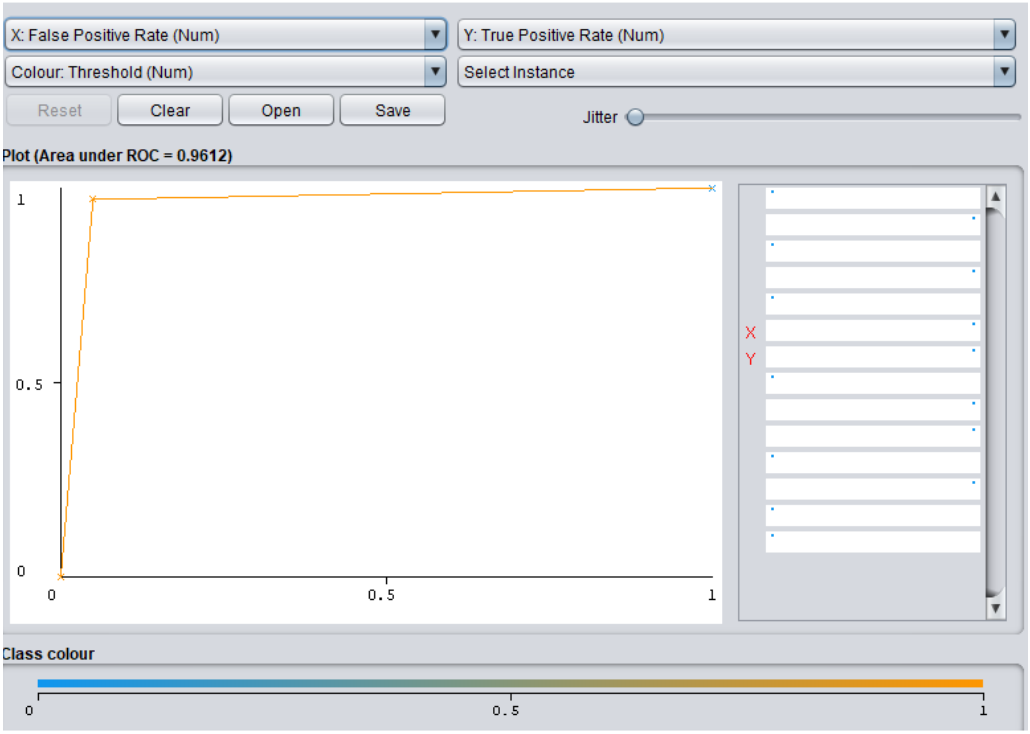

Figure 4.18 *Receiver Operating Characteristic curve(ROC) using SVM for class 4*

c) Results for Multilayer Perceptron (MLP)

Preprocessed breast cancer data set was run through Weka, the following table and figures were the outputs we got. Table 4 gives the confusion matrix of the 137 instances that was used to test the model. It shows that our model was able to correctly predict 99 instances to be benign out of 102 which were benign. It was able to predict 34 patients who were malignant out of the 35 who were malignant. In Figure 12, 99% of the tuples was labeled positive by our classifier out of the actual positive (i.e. precision), 97.1% of the positive tuples is the percentage our classifier labeled as positive (i.e. recall), the harmonic mean of precision and recall is 0.98 (i.e. F-measure). This value is very close to 1 which means the models accuracy is better. Figure 12, the true positive recognition rate is 0.971 (i.e. Sensitivity=TP/P) and the true negative recognition rate is 0.971 (i.e.

Specificity=TN/N). Figure 13 and 14 shows the Roc curves for benign (2) and Malignant (4) respectively, the area under these curves (i.e. AUROCs) are 0.995. This means our model has almost a perfect accuracy because is close to 1.

Table 4.6 Confusion Matrix for MLP

Actual\Predicted	Benign	Malignant	Total
Benign	99	3	102
Malignant	1	34	35
Total	100	37	137

```
=== Evaluation on test split ===

Time taken to test model on test split: 0 seconds

=== Summary ===

Correctly Classified Instances         133               97.0803 %
Incorrectly Classified Instances         4                2.9197 %
Kappa statistic                          0.9247
Mean absolute error                      0.0298
Root mean squared error                  0.1603
Relative absolute error                  6.7976 %
Root relative squared error             35.4649 %
Total Number of Instances              137

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure  MCC      ROC Area  PRC Area  Class
               0.971     0.029     0.990      0.971    0.980      0.925    0.995     0.998     2
               0.971     0.029     0.919      0.971    0.944      0.925    0.995     0.988     4
Weighted Avg.  0.971     0.029     0.972      0.971    0.971      0.925    0.995     0.995

=== Confusion Matrix ===

  a  b   <-- classified as
 99  3 |  a = 2
  1 34 |  b = 4
```

Figure 4.19 *Weka output figure for MLP*

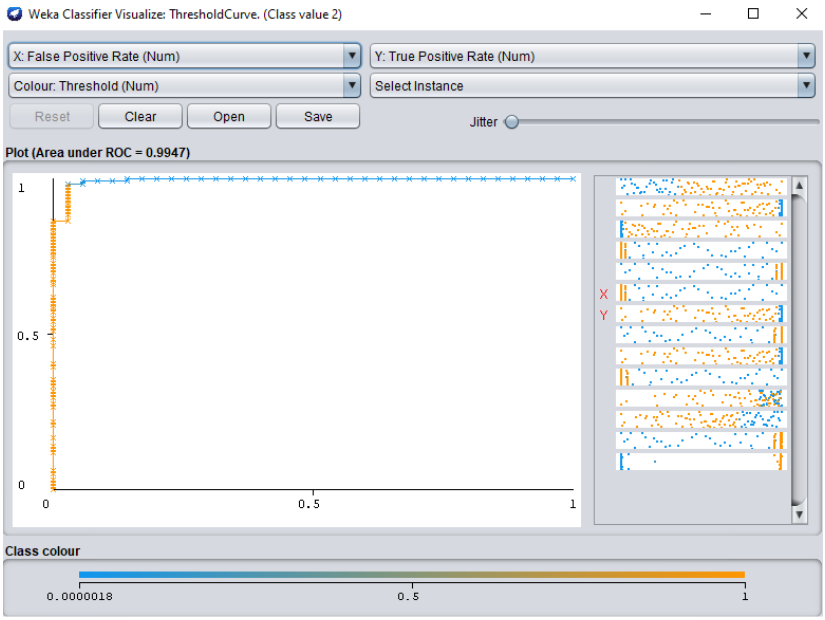

Figure 4.20 *Receiver Operating Characteristic curve(ROC) using MLP for class 2*

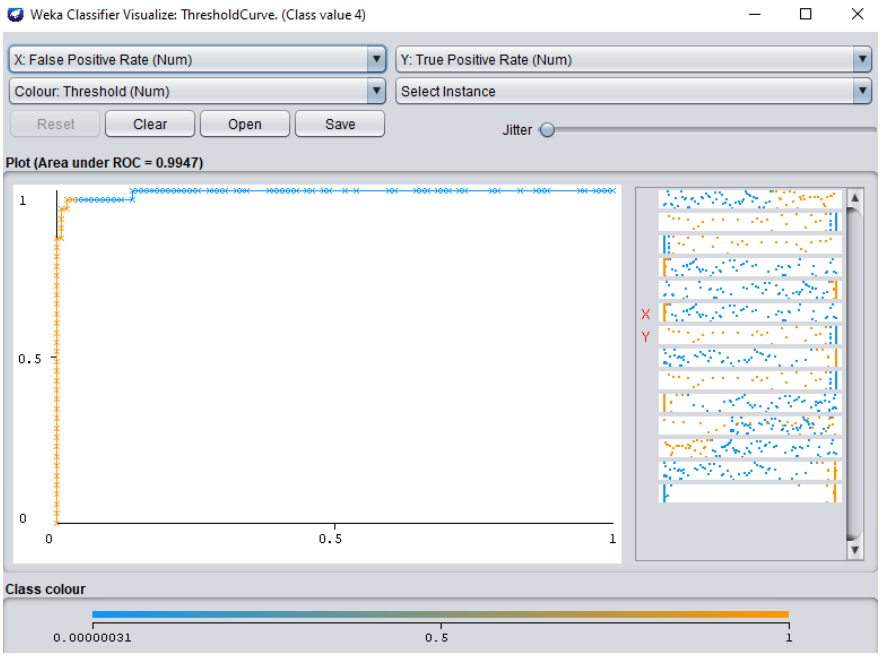

Figure 4.21 *Receiver Operating Characteristic curve(ROC) using MLP for class 4*

CHAPTER V – COVID-19 PROGNOSIS MODEL DESIGN

In this section we would describe about our works related to COVID-19 prognosis model design. Initially we started our work from the very beginning of the pandemic. After gathering most of the research information, dataset and implementation methodologies, we had successfully published one high impact journal article and two in the international conference. The most recent work is yet to be published.

5.1 COVID-19 Prognosis Model Design using Chest X-Ray images

In this research work, we had developed a Deep Learning Method (DLM) to detect COVID-19 using chest X-ray (CXR) images. Radiographic images are instantly accessible and can be used efficiently for COVID-19 detection contrasted to other costly and slow pathological tests. We used a dataset of 10,040 samples, of which 2143 had COVID-19, 3674 had pneumonia (but not COVID-19), and 4223 were normal (no COVID-19 or pneumonia). Our model had a detection accuracy of 96.43% and a sensitivity of 93.68%. The area under the ROC curve was 99% for COVID-19, 97% for pneumonia (but not COVID-19 positive), and 98% for normal cases. In conclusion, ML approaches may be used for rapid analysis of CXR images and thus enable radiologists to filter potential candidates in a time-effective manner to detect COVID-19.

5.1.1 Methodology

A. Dataset Development and Preprocessing

The dataset collection started from the mid of April as some open source platform including Github, Kaggle etc.[106-109] provide the Chest X-Ray(CXR) images of the COVID-19 positive patients. COVID-19 is kind of a complex pneumonia so we try to think of if other kinds of pneumonia detection can easily possible through chest X-Ray images then if we have enough number of COVID-19 positive CXR then we could develop the deep learning model for the design of the prognosis system.

The dataset has three types of image files,

i. CXR images of COVID-19 infected patients
ii. CXR images of patients with pneumonia other than COVID-19
iii. CXR images of normal human

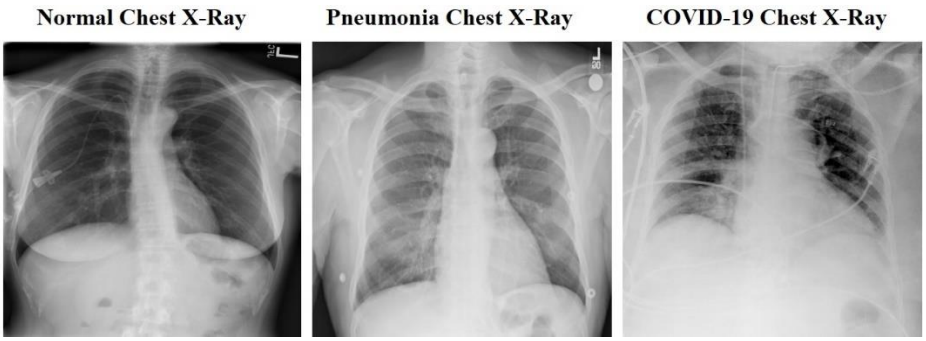

Figure 5.1 *Three categories of Chest X-Ray(CXR) images*

This figures shows three categories of the images that we have collected. We make training dataset, validation dataset and test dataset. Though there are some patches visually seen to identify and detect COVID-19 but it could result to human error. When we pass this images though a deep learning prognosis system model, it undergoes deep layers in the neural network which feature engineering the underlining characteristic

which are not visually seen by human eye and can more efficiently make more accurate detection from different kinds of CXR images.

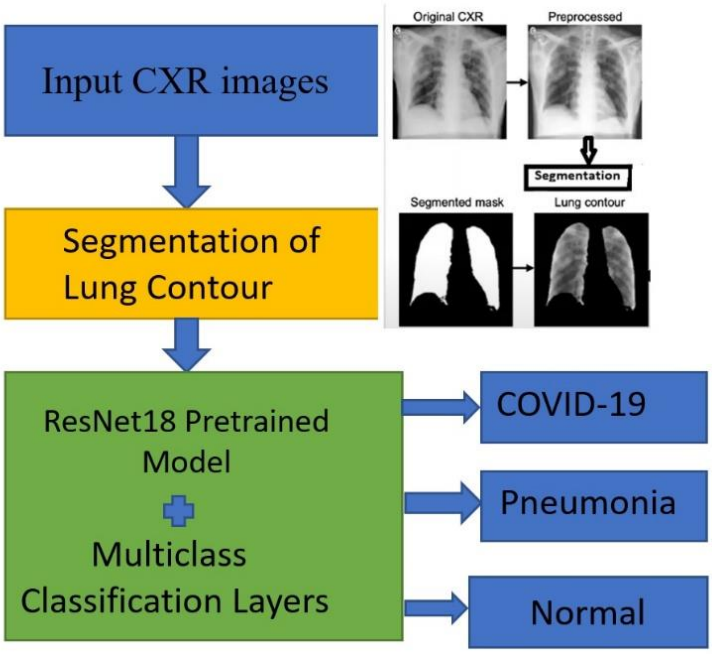

Figure 5.2 *A Block Diagram Representation of the Deep Learning Model (DML)*

The overall methodology is illustrated in this Figure 5.2 The steps of our procedure that we had followed are mention as follows,

i. First we collected the CXR dataset from open source repositories from the internet.

ii. Pre-process those images using different preprocessing technique like, deleting unlabeled images, histogram normalization etc.

iii. Segment the lung contour of the images needs pulmonary contour masks[110,111] and efficient segmentation algorithm which can use those masks to segment the segmented lung from the images. Here we use FC-DenseNet103 semantic segmentation algorithm[112] for the segmentation purpose.

iv. For designing the deep neural network model, there are basically two main section of deep layers, the first part is known as feature generation layer, where we use ResNet18 pretrained weightages to speed up the training process and leverage the transfer learning technique. The second part is the classification section where we design the layers according to multi-class classification requirements for our dataset.

v. Then the multi-class classification layer classifies the images into three sections, COVID-19, Pneumonia and normal classes.

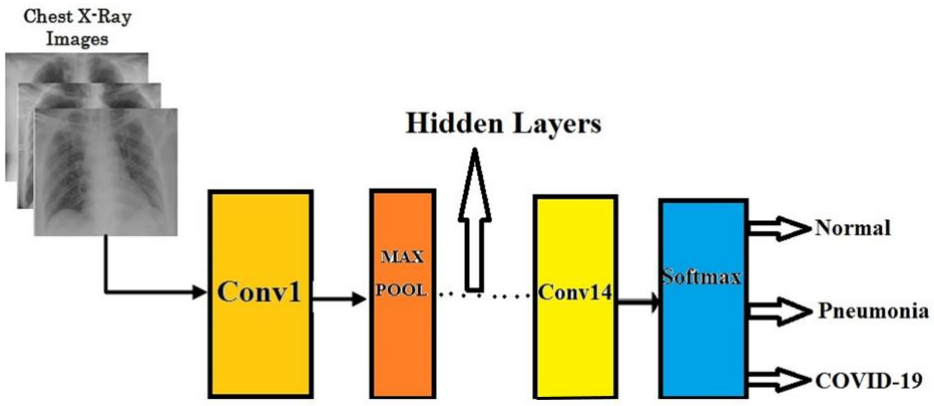

Figure 5.3 *The Architecture of proposed Deep Learning Model (DML)*

The Figure 5.3 illustrate the architecture of the proposed Deep Learning Model(DML), which consist of 46 layers, out of which, 14 convolution layers, 23 hidden layers, 4 max-pooling layers, 2 average pooling layers, 2 dropout layers, and one SoftMax layer. In this research work, we utilize the trained weightages of ResNet18 deep learning model architecture which was trained on ImageNet[113], that is the benchmark standard for deep learning community as it is the biggest variety of image dataset and the total number of image is 3.2 million which makes the deep learning architecture more robustness. As

ResNet18 already trained on this gigantic dataset, so when our model used the weightages it saves a lot of training time and complexity become more simpler. The transfer learning of the feature deep layer framework is very helpful and currently widely used standard in research community. After the feature generation architecture is completed, we just need the classification layer to classify the images accordingly.

5.1.2 Results

The terminology used to evaluate the results is given in the Figure 5.1.4.

	Based on the Gold Standard		
	Disease Present	**Disease Absent**	**Total**
Predicted Model Positive	True positive (*TP*)	False positive (*FP*)	*TP* + *FP*
Predicted Model Negative	False negative (*FN*)	True negative (*TN*)	*FN* + *TN*
Total	*TP* + *FN*	*FP* + *TN*	*TP* + *FP* + *FN* + *TN*

$$Sensitivity\ or\ Recall = \frac{TP}{(TP+FN)}$$

$$Specificity = \frac{TN}{(FP+TN)}$$

$$Positive\ Predictive\ Value\ or\ Precision = \frac{TP}{(TP+FP)}$$

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

$$The\ weighted\ average\ of\ precision\ and\ recall, known\ as\ F1\ Score = \frac{2(Recall * Precision)}{(Recall + Precision)}$$

$$False\ Positive\ Fraction = \frac{FP}{(FP+TN)}$$

$$True\ Positive\ Fraction = \frac{TP}{(TP+FN)}$$

Figure 5.4 *Terminology used to evaluate the performance of classification*

Table 5.1 COVID-19 dataset details

Data	COVID-19	Pneumonia	Normal	Total
Before Preprocessing	2143	3674	4223	10040
After Preprocessing	3535	6072	6967	16574

The dataset that we had used to design our deep learning model is very well balanced and collection of large sample data size. The total number of sample is 10,040. Out of which 2143 had COVID-19, 3674 had pneumonia and 4223 were normal patients data. We used PyTorch's[114] augmentation library. The performed operations include horizontal, vertical and rotational flips operation. By using this augmentation processes the image sample size was increased by 65%.

After the augmentation procedure was completed, we used 80 % data sample for training, 20 % data sample for testing. 10 % data were used for validation purpose.

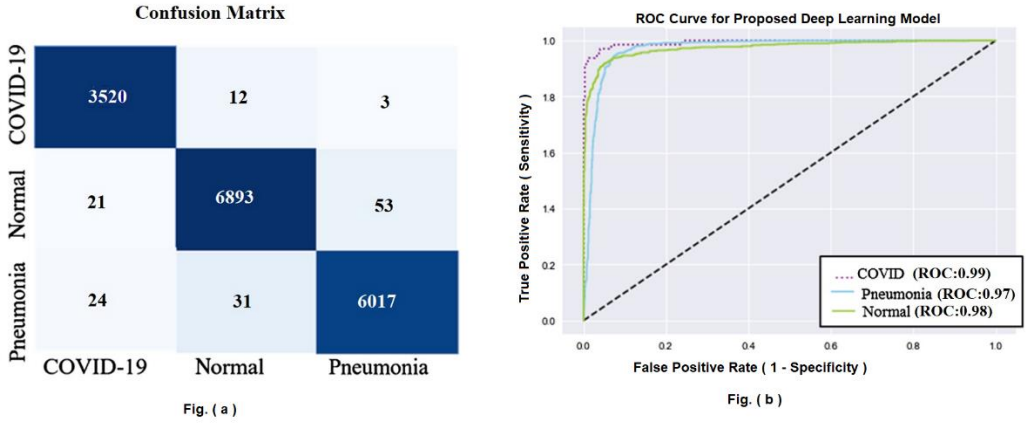

Figure 5.5 *The Confusion matrix and Receiver Operating Characteristic(ROC) curve*

Figure 5.5 illustrate the confusion matrix and Receiver Operating Characteristic(ROC) curve for our proposed Deep Learning Prognosis model. From the confusion matrix it is very much evident that 3535 COVID-19 cases, 3520 were correctly detected, 12 images were detected as normal cases, and 3 as pneumonia cases. Out of 6967 normal cases, 6893 were detected correctly, 21 were detected as COVID-19, and 53 were detected as pneumonia. Moreover, out of 6062 pneumonia cases, 6017 were detected correctly, but 24 were detected as COVID-19 and 31 as normal cases. The Receiver Operating Characteristic(ROC) curve clearly shows the high prediction accuracy of our model.

5.2 Ensemble Deep Learning Model design for COVID-19 Detection

This section would describe about the ensemble model that we had designed for the detection of COVID-19. Ensemble models help to get the better accuracy as it decreases the chance of getting errors from a single model, rather it optimizes the parameter values which helps the model to be more efficient and accurate. It also helps to avoid overfitting hence ensures higher consistency. Ensemble model also ensure all the local bias is optimize and variance values are neutral to the model. This Ensemble deep learning model provides accuracy of 93.56% and a sensitivity of 91.24% and an F1 score was 0.91.

5.2.1 Methodology

As we already collected the data and performed data pre-processing operations and augmentation operation, here in this work we used the same dataset with little bit of modification. In this model we used a dataset of 6641 samples, of which 1684 had COVID-19, 1845 had pneumonia, and 3112 were healthy individuals.

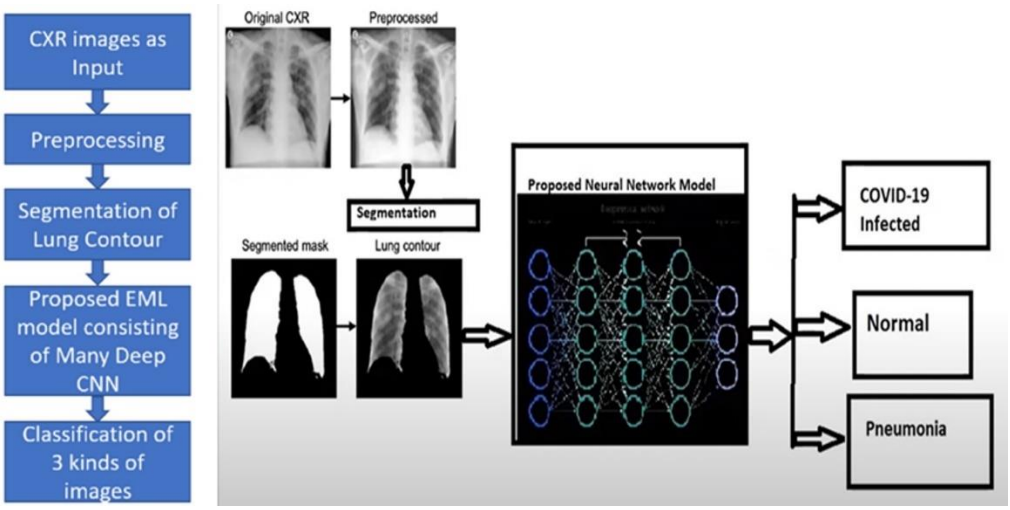

Figure 5.6 *A Block Diagram Representation of the Ensemble Deep Learning Model*

From the figure 5.6 the procedure that we have followed in designing this model is described as follows,

i. Image data samples were collected

ii. Image data samples were preproceed

iii. Segmentation is perform to separate out the lung contour of the images. Here we used a medical image segmentation algorithm, named as U-Net[115] which is widely used for medical image segmentation purposes.

iv. Then the segmented images were fed to ensemble network for classification. Here in this ensemble model also used for classifying there kinds of images. One the COVID-19 positive cases, another is other kinds of Pneumonia and the rest is healthy patient data.

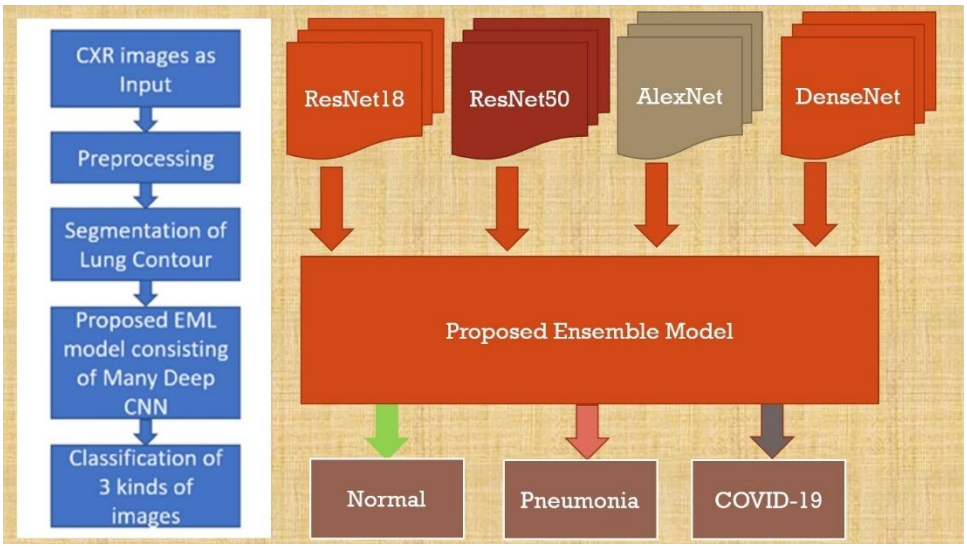

Figure 5.7 *Illustration of different deep learning model used for proposed EML model*

5.2.2 Results

The proposed Ensemble Machine Learning(EML) model had an average detection accuracy of 93.56% and a sensitivity of 91.24% and the F1 score is 0.91. Figure 4 shows the loss curve which is clearly evident the loss decreases and Figure 5 shows the Test and Train accuracy. The Train Accuracy is 97% and Test Accuracy is 92%.

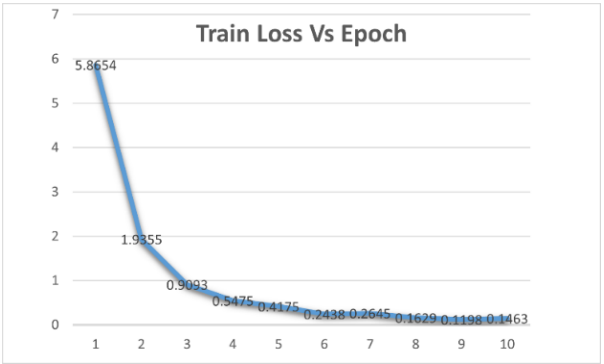

Figure 5.8 *Training loss vs Epoch*

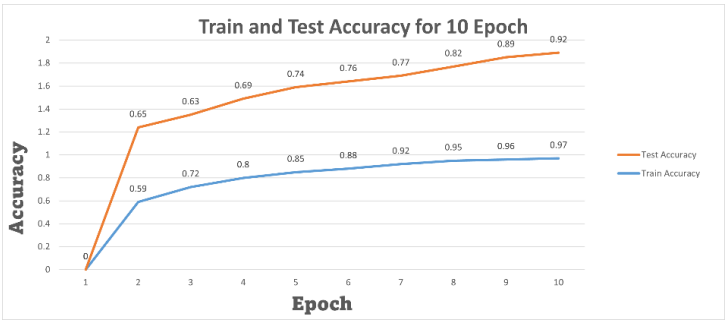

Figure 5.9 *Train and Test Accuracy*

The Figure 5.8 illustrate the training loss vs epoch. As it shows after 10 epoch it converges. Figure 5.9 illustrate the train and test accuracy of the model.

5.3 COVID-19 Prognosis Model Design using Chest CT images

This section will describe our work on COVID-19 prognosis model design using Chest CT images.

5.3.1 Methodology

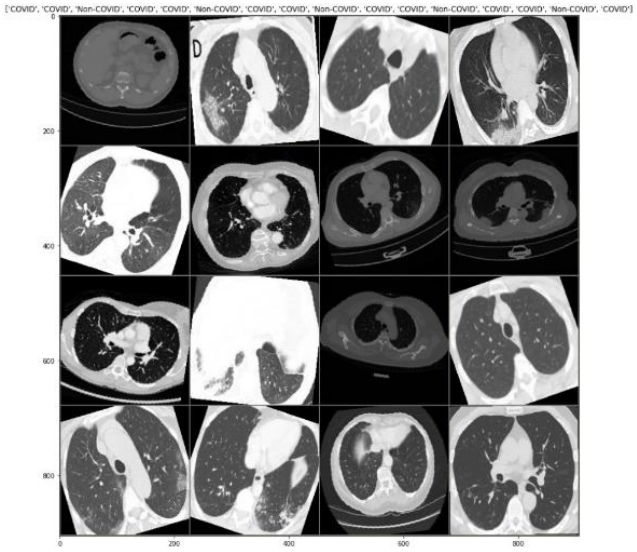

Figure 5.10 *Chest CT images after preprocessing*

Our proposed model used a dataset of 8054 real patient CT scans, of which 5427 had COVID-19 and 4223 were Non-COVID-19 patient images. The dataset is further divided into train, validation, and test sets with the ratio of 7:2:1. Figure 5.10 shows the chest CT images after preprocessing.

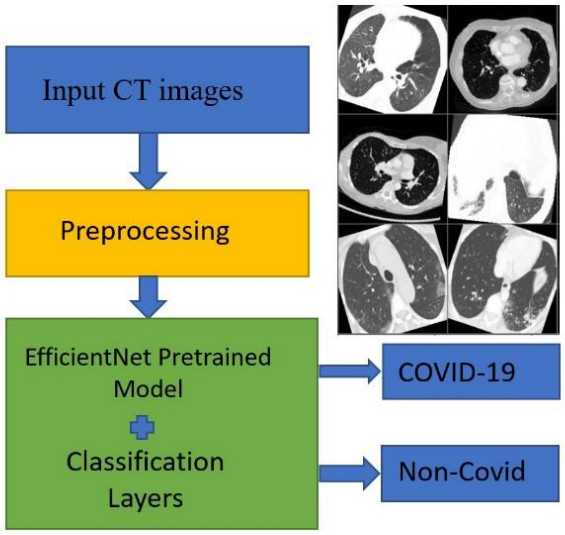

Figure 5.11 *Functional block diagram of the proposed Deep Learning approach*

Figure 5.11 illustrates the functional steps that are followed for the design of the COVID-19 prognosis system using Chest CT images only. The data collection, data preprocessing and augmentation is similar to our previous two implementation but here we use chest CT image and very efficient deep learning architecture, named as EfficientNet[116], a pretrained Model, for the feature generation tasks. Then we added a deep classification layer to classify the CT images into COVID-19 position images and COVID-19 negative images, i.e. Non-COVID-19 images. After execution of the feature generation tasks, the total number of parameters of our dataset is 18,830,011, out of which Trainable params are 1,281,395 and Non-trainable params are 17,548,616.

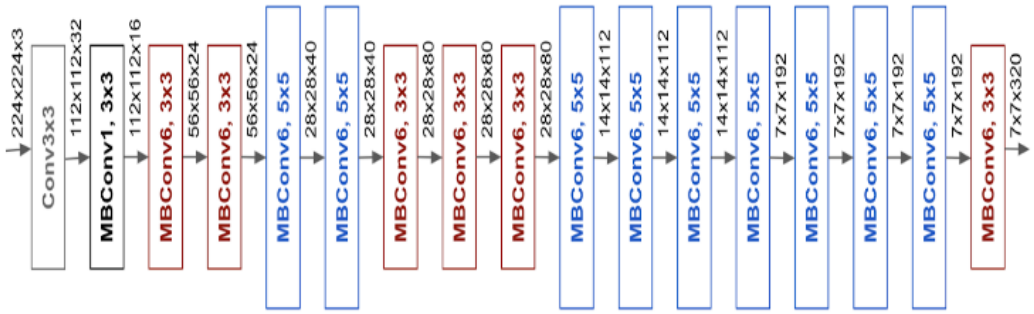

Figure 5.12 *EfficientNet B0 Baseline neural network Architecture[116]*

The architecture of EfficientNet baseline Neural Network is shown in Figure 5.12 EfficientNet is a very efficient high-performance convolutional neural network architecture and scaling method that uniformly scales all dimensions of depth, width and resolution using a compound coefficient. Where benchmark deep neural models arbitrarily scale these factors, the EfficientNet scaling method uniformly scales network width, depth, and resolution with a set of fixed scaling coefficients. The effectiveness of model scaling is based on the baseline network which is playing a key role in high performance. The baseline network builds up with a deep neural architecture search using the AutoML MNAS framework[117], which optimizes both accuracy and efficiency.

The classification layers provide the percentage of the score for COVID-19 patients and Non-COVID-19 patients.

5.3.2 Results

The proposed deep learning model had an average detection accuracy of 91.96% on the test CT images. Sensitivity is 92.24, Specificity is 93.01 and F1 Score is 0.90. The following Figure 5.3.4 shows the Loss Curve. From the graph, it is evident the loss

decreased over the epoch. Figure 5.3.5 shows the accuracy curve. From the graph, it is evident that the accuracy improves over the epoch.

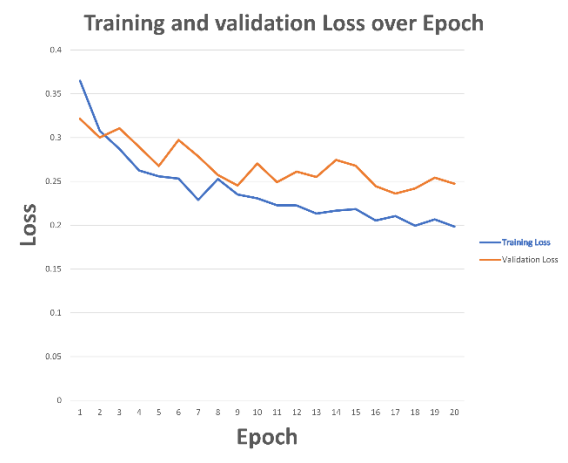

Figure 5.13 *Training and Validation Loss over Epoch*

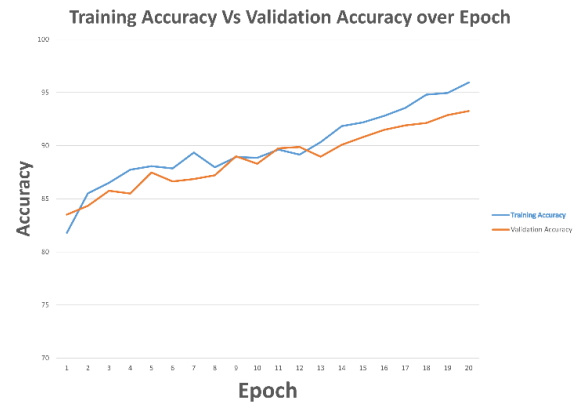

Figure 5.14 *Training Accuracy Vs Validation Accuracy over epoch*

The following Figure 5.15, 5.16, 5.17, 5.18 are the results coming from an unknown test sample and then matching their value for the predicted score obtained using our proposed

model. It is clearly evident from the said figures that the predicted score clearly confirmed the type of the class with very high accuracy.

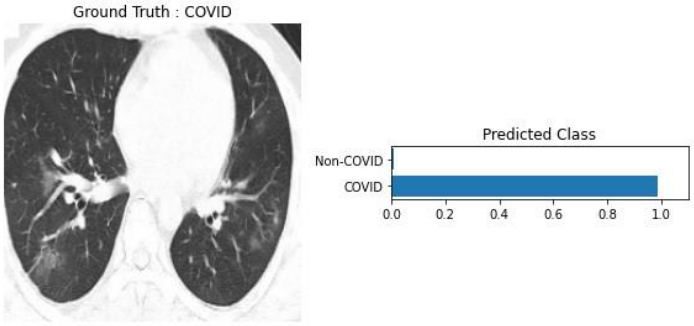

Figure 5.15 *Predicted score matching with the ground truth values.*

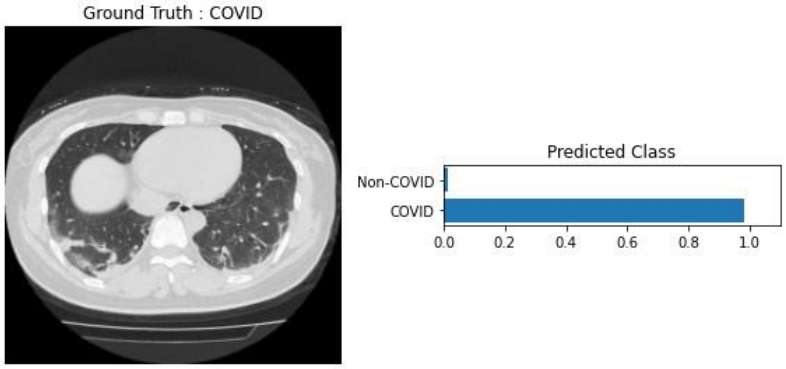

Figure 5.16 *Predicted score matching with the ground truth values.*

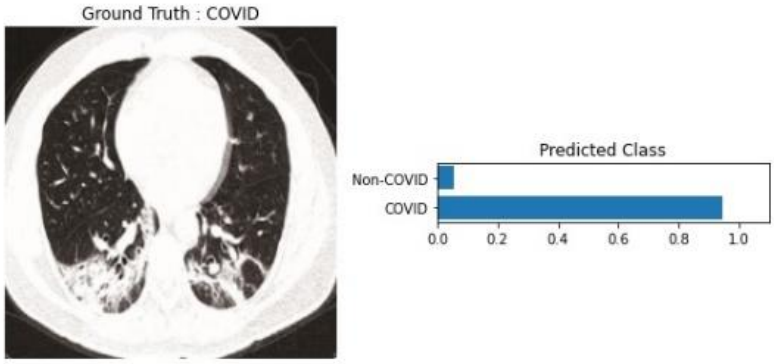

Figure 5.17 *Predicted score matching with the ground truth values.*

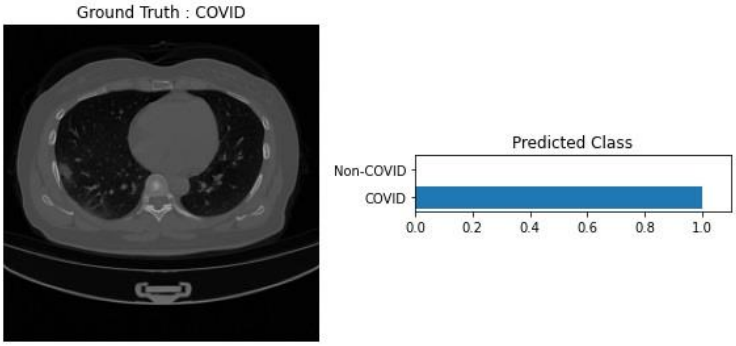

Figure 5.18 *Predicted score matching with the ground truth values.*

5.4 COVID-19 Prognosis Model using combination of Chest CT and CXR images

After successfully completed three diagnosis models on COVID-19 with different approaches we had investigated the possibility to optimize the model in terms of deep learning architecture, data balancing and increasing the strength and size of the sample size. So we combine the CXR and Chest CT images and built a model which was more efficient and accurate.

5.4.1 Methodology

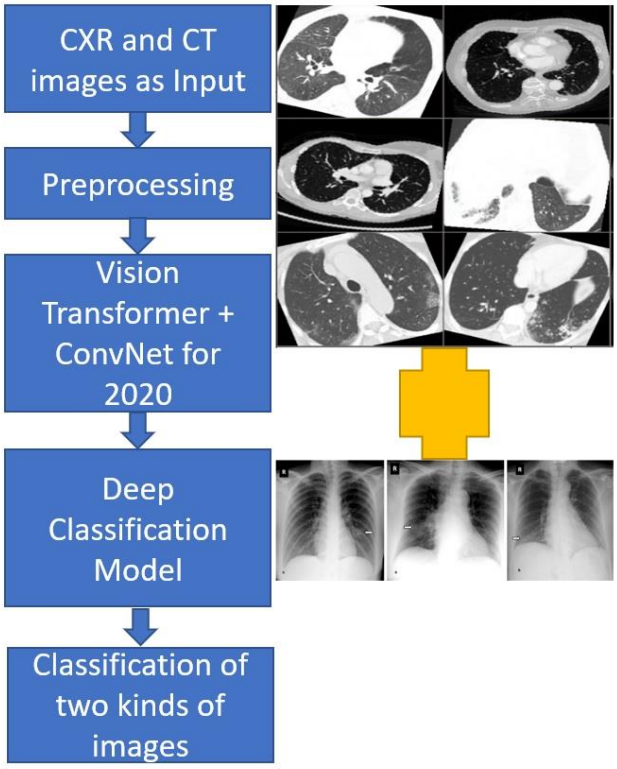

Figure 5.19 *The Functional Flowchart of our Deep Neural Network Model*

The Figure 5.19 illustrate the functional flowchart for our prognosis model. The procedures were followed are as follows,

i. We took CXR and CT both images and used it as an input in three different ways, First only CXR, second one is CT and the third one is the combination of both.

ii. We used vision transformer model like BEiT[118], which is a Bidirectional Encoder representation from Image Transformers, actually a self-supervised vision representation model which was trained in large image dataset ImageNet[61] and optimized the trained parameters. We heuristically try different PyTorch Vision models[119] and found that “A ConvNet for the 2020’s”[96], outperforms every other existing models currently available, so we had

constructed the ‘Feature Generation Architecture” for our diagnosis model using their trained parameter network and implement the classification architecture according to our dataset.

iii. After successful execution we got the output as COVID-19 positive and COVID-19 negative images.

5.4.2 Results

The results that we got in this work was very promising and outperform all existing models including our all previous models.

A. Data Samples details

Table 5.2 Chest X-Ray(CXR) image sample details

Sample Name	Train	Validation	Test
COVID-19 Positive	2022	1617	405
COVID-19 Negative	2750	2200	550

Table 5.3 Chest CT image sample details

Sample Name	Train	Validation	Test
COVID-19 Positive	2713	2170	544
COVID-19 Negative	1314	1051	263

Table 5.4 Combined CXR and Chest CT sample details

Sample Name	Train	Validation	Test
COVID-19 Positive	4700	2841	1638
COVID-19 Negative	4034	2376	1409

The data sample that we had used was very large and very well balanced which help our model to be more robust and accurate.

We got the best results using the model convnext_xlarge_in22k [96].

After 3 fold cross validation the results are as follows,

B. Results for CXR

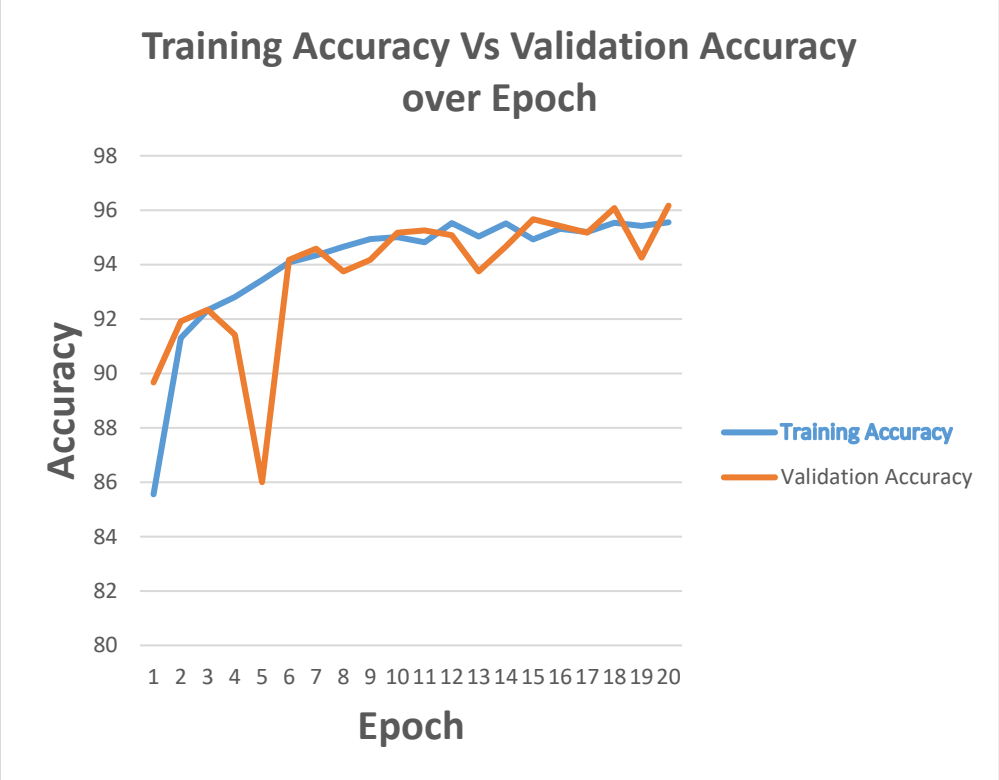

Figure 5.20 *Training accuracy Vs Validation accuracy over epoch*

Figure 5.20 illustrate training accuracy Vs validation accuracy over epoch for CXR images. Here we got training accuracy 95.9376 %, validation accuracy 96.1667 % and testing accuracy 95.1250 %.

C. Results for Chest CT

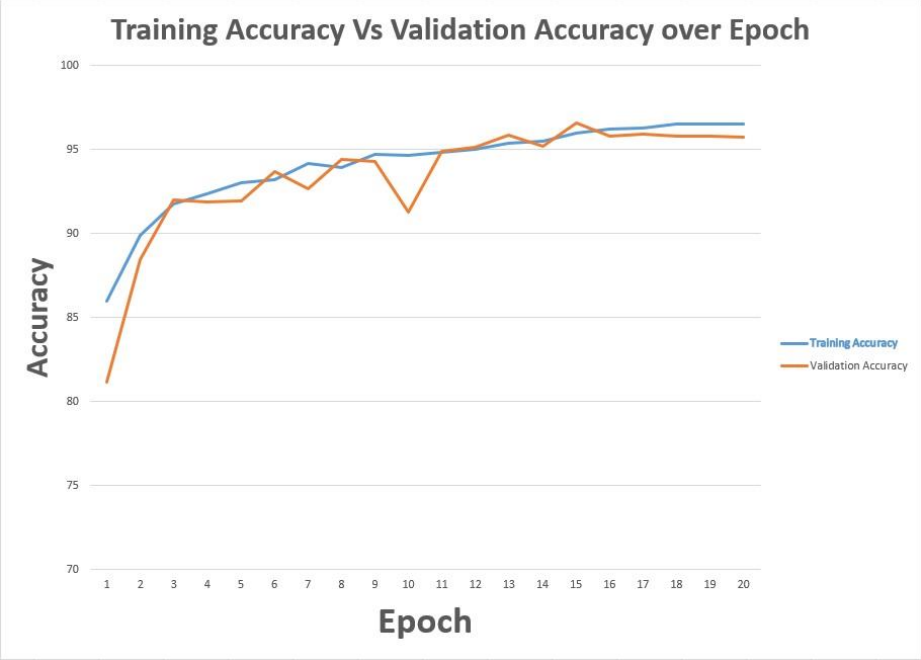

Figure 5.21 *Training accuracy Vs Validation accuracy over epoch*

Figure 5.21 illustrate training accuracy Vs validation accuracy over epoch for Chest CT images. Here we got training accuracy 96.5474 %, validation accuracy 95.7302 % and testing accuracy 97.0588 %.

D. Results for Combined Chest CT images and CXR

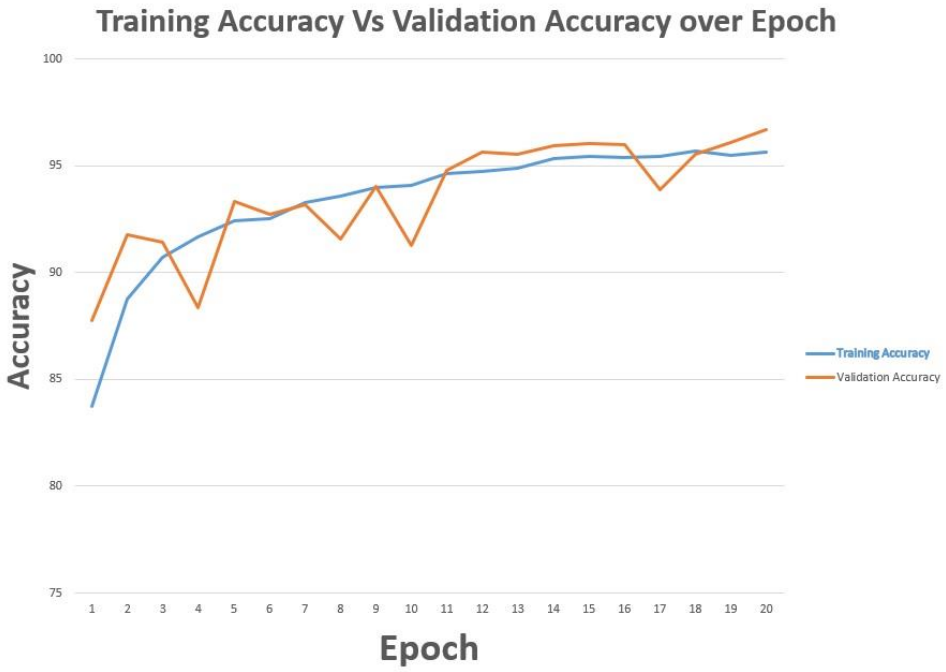

Figure 5.22 *Training accuracy Vs Validation accuracy over epoch*

Figure 5.22 illustrate training accuracy Vs validation accuracy over epoch for Chest CT and CXR combined images. Here we got training accuracy 95.6665 %, validation accuracy 96.6961% and testing accuracy 97.8859 %.

CHAPTER VI – FUTURE WORK AND CONCLUSION

The work that we had completed helped us gain a lot of knowledge for the application of deep learning in the medical data and image analysis. The deep learning model that we have created had potential possibility to apply to other medical disease analysis as well as other fields. The designing of recommender system also helps medical fraternity to take the decision more confidently and appropriately. Many earlier models reported in the literature review section uses a limited number of COVID-19 positive images due to the lack of availability of such data during the early stages of the pandemic. In the present work, a large dataset of CXR and chest CT images were used with balanced positive and negative cases. This adds to the growing body of evidence that deep learning techniques may be used in a clinical setting with confidence for detecting COVID-19 and other disease prognosis model design. One of the recommendations of WHO is to increase testing for COVID-19. Usage of deep learning-based methods would bring down the cost and accelerate the diagnosis process which will enable rapid testing on a global scale.

In our future work, we would investigate vision transformer models in more comprehensive way with this dataset. Transformer architecture performs very well in Natural Language Processing(NLP) tasks and is widely used in NLP domains. But recently many research works comes up with vision applications as well, though Deep Convolutional Neural Network is mostly preferred by most of the researchers in computer vision applications, especially in image processing domains, object detection

domains, etc. We also like to extend our work on our classification layers with more variety of activation functions and loss functions.

REFERENCES

1. Alan M. Turing, Computing machinery and intelligence, in: Parsing the Turing Test, Springer, 2009, pp. 23–65.
2. Stuart J. Russell, (Stuart Jonathan). Artificial Intelligence : a Modern Approach. Upper Saddle River, N.J. :Prentice Hall, 2010.
3. Kevin Warburton, Deep learning and education for sustainability, International Journal of Sustainability in Higher Education 4 (1) (2003) 44–56.
4. Hyeji Kim, Yihan Jiang, Ranvir Rana, Sreeram Kannan, Sewoong Oh, Pramod Viswanath, Communication algorithms via deep learning, preprint, arXiv:1805.09317, 2018.
5. Hoang Nguyen, Le-Minh Kieu, Tao Wen, Chen Cai, Deep learning methods in transportation domain: a review, IET Intelligent Transport Systems 12 (9) (2018) 998–1004.
6. Andreas Kamilaris, Francesc X. Prenafeta-Boldú, Deep learning in agriculture: a survey, Computers and Electronics in Agriculture 147 (2018) 70–90.
7. James B. Heaton, Nick G. Polson, Jan Hendrik Witte, Deep learning for finance: deep portfolios, Applied Stochastic Models in Business and Industry 33 (1) (2017) 3–12.
8. Andre Esteva, Alexandre Robicquet, Bharath Ramsundar, Volodymyr Kuleshov, Mark DePristo, Katherine Chou, Claire Cui, Greg Corrado, Sebastian Thrun, Jeff Dean, A guide to deep learning in healthcare, Nature Medicine 25 (1) (2019) 24–29.
9. KC Santosh, AI-driven tools for coronavirus outbreak: need of active learning and cross-population train/test models on multitudinal/multimodal data, Journal of Medical Systems 44 (5) (2020) 93, https://doi.org/10. 1007/s10916-020-01562-1.
10. Sourodip Ghosh, Ahana Bandyopadhyay, Shreya Sahay, Richik Ghosh, Ishita Kundu, KC Santosh, Colorectal histology tumor detection using ensemble deep neural network, Engineering Applications of Artificial Intelligence 100 (2021) 104202, https://doi.org/10.1016/j.engappai.2021.104202.
11. Somenath Chakraborty, Category identification technique by a semantic feature generation algorithm. In: Deep learning for internet of things infrastructure. CRC Press, pp 129–144.
12. Locatelli, "Nephrology: Main Advances in the Last 40 Years", Journal of Nephrology, Vol. 19, pp 6-11, 2006.
13. Z. Qi et al.,"Application of Polyvinyl Alcohol (PVA) Macro-Encapsulated Isets in Islet Transplantation," 2009 3rd International Conference on Bioinformatics and Biomedical Engineering,Beijing,2009,pp.1-3.
14. A. K. Iyer, G. L. Zenarosa, A. J. Schaefer, C. H. Chang, C. L. Bryce and M. S. Roberts, "A biologically based discrete-event simulation model of liver transplantation in the

United States for pediatric and adult patients," Proceedings of the 2011 Winter Simulation Conference (WSC), Phoenix, AZ, 2011, pp. 1275-1282, doi: 10.1109/WSC.2011.6147848.

15. Alagoz, O., Bryce, C. L., Shechter, S., Schaefer, A., Chang, C.-C. H., Angus, D. C., & Roberts, M. S. (2005). Incorporating Biological Natural History in Simulation Models: Empirical Estimates of the Progression of End-Stage Liver Disease. Medical Decision Making, 25(6), 620–632. https://doi.org/10.1177/0272989X05282719.

16. Alagoz, O., Maillart, L., Schaefer, A., & Roberts, M. (2004). The Optimal Timing of Living-Donor Liver Transplantation. Management Science, 50(10), 1420-1430. Retrieved June 12, 2020, from www.jstor.org/stable/30046184.

17. J. H. Fielder, "Organ transplants," in IEEE Engineering in Medicine and Biology Magazine, vol. 16, no. 2, pp. 84-86, March-April 1997, doi: 10.1109/51.582184.

18. D. Medved, P. Nugues and J. Nilsson, "Predicting the outcome for patients in a heart transplantation queue using deep learning," 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo,2017,pp.74-77,doi: 10.1109/EMBC.2017.8036766.

19. A. Ravikumar, R. Saritha and V. Chandra, "Support vector machine based prognostic analysis of renal transplantations," 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode,2013,pp.1-6.

20. L. Al-Ebbini, A. Oztekin, Z. Sevkli and D. Delen, "Predicting quality of life for lung transplant recipients: A hybrid genetic algorithms-basedmethodology,"2017,IEEEJordan Conference on Applied Electrical Engineering and Computing Technologies(AEECT),Aqaba,2017,pp.1-6.

21. C.G.Raji and S. S. Vinod Chandra, "Long-Term Forecasting the Survival in Liver Transplantation Using Multilayer Perceptron Networks," in IEEE Transactions on Systems, Man, and Cybernetics: Systems, vol. 47, no. 8, pp. 2318-2329, Aug. 2017, doi: 10.1109/TSMC.2017.2661996.

22. S. Muthukumaran and B. Velumani,"A Biclustering Approach for Investigating Patterns for Breast Cancer Attributes," 2014 International Conference on Intelligent Computing Applications, Coimbatore, 2014,pp. 22-26, doi: 10.1109/ICICA.2014.14.

23. J.A. Nepomuceno, A. Troncoso, & J.S. Aguilar-Ruiz, Biclustering of Gene Expression Data by Correlation-Based Scatter Search. BioData Mining 4, 3 (2011). https://doi.org/10.1186/1756-0381-4-3.

24. Doruk Bozdağ, Ashwin S. Kumar, and Umit V. Catalyurek. 2010. Comparative analysis of bi-clustering algorithms. In Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology(BCB'10). Association for Computing Machinery, New York, NY, USA, 265–274. https://doi.org/10.1145/1854776.1854814.

25. V. Sangeetha and J. Preethi, "Cluster analysis of breast cancer data using Genetic Algorithm and Spiking Neural Networks," 2015 IEEE 9th International Conference on

Intelligent Systems and Control (ISCO), Coimbatore, 2015, pp. 1-5, doi:10.1109/ISCO.2015.7282345.

26. T. Padhi and P. Kumar,"Breast Cancer Analysis Using WEKA," 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 2019, pp. 229-232, doi: 10.1109/CONFLUENCE.2019.8776911.

27. D. Verma and N. Mishra, "Analysis and prediction of breast cancer and diabetes disease datasets using data mining classification techniques," 2017 International Conference on Intelligent Sustainable Systems(ICISS), Palladam, 2017, pp. 533-538, doi: 10.1109/ISS1.2017.8389229.

28. L. M. Naeni and A. Salehipour,"Investigating a Breast Cancer Gene Expression Data Using a Novel Clustering Approach," 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Macao, Macao, 2019, pp. 1038-1042, doi:10.1109/IEEM44572.2019.8978949.

29. Taosheng Xu, Ning Su, Rujing Wang and Liangtu Song, "Gene selection for cancer clustering analysis based on expression data," 2015 4th International Conference on Computer Science and Network Technology (ICCSNT), Harbin, 2015, pp. 516-519, doi:10.1109/ICCSNT.2015.7490801.

30. A. L. Fijri and Z. Rustam, "Comparison between Fuzzy Kernel C-Means and Sparse Learning Fuzzy C-Means for Breast Cancer Clustering," 2018 International Conference on Applied Information Technology and Innovation (ICAITI), Padang, Indonesia, 2018, pp. 158-161, doi: 10.1109/ICAITI.2018.8686707.

31. P. H. S. Coutinho and T. P. das Chagas, "Proposal of new hybrid fuzzy clustering algorithms—Application to breast cancer dataset," 2017 IEEE Latin American Conference on Computational Intelligence (LA-CCI), Arequipa, 2017,pp.1-6, doi:10.1109/LA-CCI.2017.8285679.

32. D. Wu, L. Sheng, E. Xu, K. Xing and D. Chen, "Analysis of an ensemble algorithm for clustering cancer data," 2012 IEEE International Conference on Bioinformatics and Biomedicine Workshops, Philadelphia, PA,2012, pp.754-755. doi: 10.1109/BIBMW.2012.6470233.

33. Chen, D., Xing, K., Henson, D., Sheng, L., Schwartz, A. M., & Cheng, X. (2009). Developing prognostic systems of cancer patients by ensemble clustering. Journal of biomedicine & biotechnology, 2009, 632786. https://doi.org/10.1155/2009/63278.

34. Breast Cancer Statistics and diagnosis, available at https://www.breastcancer.org/symptoms/diagnosis

35. Diagnosis Screening Details of Breast cancer, available at https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection.html

36. Diagnosis Screening Details of Breast cancer, available at https://www.who.int/cancer/prevention/diagnosis-screening/en/

37. N. Jafarpisheh, N. Nafisi and M. Teshnehlab, "Breast cancer relapse prognosis by classic and modern structures of machine learning algorithms," 2018 6th Iranian Joint Congress on Fuzzy and Intelligent Systems (CFIS), Kerman, 2018, pp. 120-122, doi: 10.1109/CFIS.2018.8336649.

38. Naveen, R. K. Sharma and A. Ramachandran Nair, "Efficient Breast Cancer Prediction Using Ensemble Machine Learning Models," 2019 4th International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), Bangalore, India, 2019, pp. 100-104, doi:10.1109/RTEICT46194.2019.9016968.

39. H. Amitha and I. Selvamani, "A Survey on Automatic Breast Cancer Grading of Histopathological Images," 2018 International Conference on Control, Power, Communication and Computing Technologies (ICCPCCT), Kannur, 2018, pp. 185-189, doi:10.1109/ICCPCCT.2018.8574291.

40. F. Imani, R. Chen, C. Tucker and H. Yang, Random Forest Modeling for Survival Analysis of Cancer Recurrences, 2019 IEEE 15th International Conference on Automation Science and Engineering(CASE),Vancouver, BC, Canada, 2019,pp.399-404, doi:10.1109/COASE.2019.8843271.

41. M. Elnahas, M. Hussein and A. Keshk, Artificial Neural Network as Ensemble Technique Fuser for Improving Classification Accuracy, 2019 Ninth International Conference on Intelligent Computing and Information Systems (ICICIS), Cairo, Egypt, 2019, pp.174-179, doi: 10.1109/ICICIS46948.2019.9014791.

42. N. M. J. Kumari and K. K. V. Krishna, Prognosis of Diseases Using Machine Learning Algorithms: A Survey, 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT), Coimbatore, 2018, pp. 1-9, doi: 10.1109/ICCTCT.2018.8550902.

43. Zhu W, Xie L, Han J, Guo X. The Application of Deep Learning in Cancer Prognosis Prediction. Cancers (Basel). 2020;12(3):603. Published 2020 Mar 5. doi:10.3390/cancers12030603.

44. B. Fu, P. Liu, J. Lin, L. Deng, K. Hu and H. Zheng, Predicting Invasive Disease-Free Survival for Early Stage Breast Cancer Patients Using Follow-Up Clinical Data, in IEEE Transactions on Biomedical Engineering, vol.66, no.7, pp. 2053-2064, July 2019, doi:10.1109/TBME.2018.2882867.

45. H. Lu, H. Wang, S.W. Yoon, A dynamic gradient boosting machine using genetic optimizer for practical breast cancer prognosis, Expert Syst. Appl.116 (2019) 340–350.

46. C. Huang; Y. Wang; X. Li; L. Ren; J. Zhao; Y. Hu; L. Zhang; G. Fan; J. Xu; X. Gu; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020, 395, 497–506.

47. A.S. Fauci; H.C. Lane; R.R. Redfield, COVID-19 Navigating the uncharted. N. Engl. J. Med. 2020, 382, 1268–1269.

48. F. Wu; S. Zhao; B. Yu; Y.M. Chen; W. Wang; Z.G. Song; Y. Hu; Z.W. Tao; J.H. Tian; Y.Y. Pei; et al. A new coronavirus associated with human respiratory disease in China. Nature 2020, 579, 265–269.

49. Gavriatopoulou, M.; Korompoki, E.; Fotiou, D.; Ntanasis-Stathopoulos, I.; Psaltopoulou, T.; Kastritis, E.; Terpos, E.; Dimopoulos, M.A. Organ-specific manifestations of COVID-19 infection. Clin. Exp. Med. 2020, 20, 493–506. https://doi.org/10.1007/s10238-020-00648-x.

50. W. Wang; Y. Xu; R. Gao; R. Lu; K. Han; G. Wu; W. Tan, Detection of SARS-cov-2 in different types of clinical specimens. JAMA 2020, 323, 1843–1844.

51. A. Brihn; J. Chang ; K. OYong; S. Balter; D. Terashita; Z. Rubin; N. Yeganeh, Diagnostic performance of an antigen test with RT-PCR for the detection of SARS-CoV-2 in a hospital setting Los Angeles County, California, June-August 2020. MMWR Morb. Mortal. Wkly. Rep. 2021, 70, 702–706. https://doi.org/10.15585/mmwr.mm7019a3.

52. T. Ai; Z. Yang; H. Hou; C. Zhan; C. Chen; W. Lv; Q. Tao; Z. Sun; L. Xia, Correlation of chest CT and RT-PCR testing for coronavirus disease 2019 (COVID-19) in China: A report of 1014 cases. Radiology 2020, 296, E32–E40. https://doi.org/10.1148/radiol.2020200642.

53. L. Fan; D. Li; H. Xue; L. Zhang; Z. Liu; B. Zhang; L. Zhang; W. Yang; B. Xie; X. Duan; et al. Progress and prospect on imaging diagnosis of COVID-19. Chin. J. Acad. Radiol. 2020, 3, 4–13. https://doi.org/10.1007/s42058-020-00031-5.

54. H.X. Bai; B. Hsieh; Z. Xiong; K. Halsey; J.W. Choi; T.M.L. Tran; I. Pan; L.B. Shi; D.C. Wang; J. Mei; et al. Performance of radiologists in differentiating COVID-19 from viral pneumonia on chest CT. Radiology 2020, 296, 200823. https://doi.org/10.1148/radiol.2020200823.

55. X. He; X. Yang; S. Zhang; J. Zhao; Y. Zhang; E. Xing; P. Xie, Sample-efficient deep learning for COVID-19 diagnosis based on CT scans. Health Inform. 2020. https://doi.org/10.1101/2020.04.13.20063941.

56. L. Wang; Z.Q. Lin; A. Wong; COVID-Net: A tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci. Rep. 2020, 10, 19549. Available online: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7658227/ (accessed 10 December 2021).

57. S.A. Harmon; T. H. Sanford; S. Xu; E.B. Turkbey; H. Roth; Z. Xu; D. Yang; A. Myronenko; V. Anderson; A. Amalou; et al. Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets. Nat. Commun. 2020, 11, 4080. https://doi.org/10.1038/s41467-020-17971-2.

58. D. Das; K.C. Santosh; U. Pal; Truncated inception net: COVID-19 outbreak screening using chest X-rays. Phys. Eng. Sci. Med. 2020, 43, 915–925. https://doi.org/10.1007/s13246-020-00888-x.

59. A. K. Jaiswal; P. Tiwari; S. Kumar; D. Gupta; A. Khanna; J.P.C. Rodrigues; Identifying pneumonia in chest X-rays: A deep learning approach. Measurement 2019, 145, 511–518. https://doi.org/10.1016/j.measurement.2019.05.076.

60. M. Annarumma; S. J. Withey; R. J. Bakewell; E. Pesce; V. Goh; Automated triaging of adult chest radiographs with deep artificial neural networks. Radiology 2019, 291, 272–272. https://doi.org/10.1148/radiol.2019194005.

61. A. Krizhevsky; I. Sutskever; G. E. Hinton; Imagenet classification with deep convolutional neural networks. In Proceedings of the Advances in Neural Information Processing Systems 25: 26th Annual Conference on Neural Information Processing Systems, Lake Tahoe, NV, USA. 3–6 December 2012; pp. 1097–1105.

62. Y. LeCun; Y. Bengio; G. Hinton; Deep learning. Nature 2015, 521, 436–444.

63. Centers for Disease Control and Prevention. Delta variant: What We Know about the Science. 2021. Available online: https://www.cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html?s_cid=11512:cdc%20delta%20variant:sem.ga:p:RG:GM:gen:PTN:FY21 (accessed on 9 December 2021).

64. Centers for Disease Control and Prevention. Omicron Variant: What You Need to Know. 2021. Available online: https://www.cdc.gov/coronavirus/2019-ncov/variants/omicron-variant.html (accessed on 9 December 2021).

65. K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

66. Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. 2017. ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 6 (June 2017), 84–90. https://doi.org/10.1145/3065386.

67. G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 2261-2269, doi: 10.1109/CVPR.2017.243.

68. Xiangyu Zhang, Jianhua Zou, Kaiming He, and Jian Sun. 2016. Accelerating Very Deep Convolutional Networks for Classification and Detection. IEEE Trans. Pattern Anal. Mach. Intell. 38, 10 (October 2016), 1943–1955. https://doi.org/10.1109/TPAMI.2015.2502579.

69. P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, M. P. Lungren, and A. Y. Ng, CheXNet: Radiologist-level pneumonia detection on chest Xrays with deep learning, 2017, arXiv:1711.05225. http://arxiv.org/abs/1711.05225.

70. 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.

71. S. Minaee, R. Kafish, M. Sonka, S. Yazdani, and G. Jamalipour Sou, Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning, Med. Image Anal., vol. 65, Oct. 2020, Art. no. 101794.

72. A. I. Khan, J. L. Shah, and M. M. Bhat, CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images, Comput. Methods Programs Biomed., vol. 196, Nov. 2020, Art. no. 105581.

73. H. Panwar et. al, A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. Chaos Solitons Fractals 140, 110190. https:// doi. org/ 10. 1016/j. chaos. 2020. 110190 (2020).

74. Xuehai He, et al., Sample-efficient deep learning for COVID-19 diagnosis based on CT scans, Medrxiv (2020).

75. X. Wang , X. Deng, Q. Fu, et. al, A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. IEEE Trans Med Imaging 2020;39(8):2615-25.

76. Ozsahin I, Sekeroglu B, Musa MS, Mustapha MT, UzunOzsahin D. Review on Diagnosis of COVID-19 from chest CT images using artificial intelligence, computational and mathematical methods in medicine, Hindawi; 2020.

77. Z. Tang, et. al Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images; 2020, arXiv preprint arXiv:2003.11988.

78. Ying Song, et. al, Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images, IEEE/ACM Trans. Comput. Biol. Bioinform. (2021).

79. M. Tan and Q. V. Le. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv, abs/1905.11946.

80. M. Tan, B. Chen, R. Pang, V. Vasudevan, & Q.V. Le, (2019). MnasNet: Platform-Aware Neural Architecture Search for Mobile. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2815-2823.

81. Hangbo Bao, Li Dong, and Furu Wei. BEiT: BERT pretraining of image transformers. arXiv:2106.08254, 2021.

82. Francesco Pelosin, Saurav Jha, Andrea Torsello, Bogdan Raducanu, Joost van de Weijer, Towards Exemplar-Free Continual Learning in Vision Transformers: an Account of Attention, Functional and Weight Regularization, https://doi.org/10.48550/arXiv.2203.13167.

83. Sayak Paul and Pin-Yu Chen. Vision transformers are robust learners. arXiv preprint arXiv:2105.07581, 2021.

84. Ze Liu, Yutong Lin, Yue Cao, Han Hu, Yixuan Wei, Zheng Zhang, Stephen Lin, and Baining Guo. Swin transformer: Hierarchical vision transformer using shifted windows. In Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021.

85. Hugo Touvron, Matthieu Cord, Matthijs Douze, Francisco Massa, Alexandre Sablayrolles, and Herv'e J'egou. Training data-efficient image transformers & distillation through attention. In ICML, 2021.

86. YuqingWang, Zhaoliang Xu, XinlongWang, Chunhua Shen, Baoshan Cheng, Hao Shen, and Huaxia Xia. End-to-end video instance segmentation with transformers. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021.

87. Tete Xiao, Mannat Singh, Eric Mintun, Trevor Darrell, Piotr Dollar, and Ross B. Girshick. Early convolutions help transformers see better. CoRR, abs/2106.14881, 2021.

88. Jianwei Yang, Chunyuan Li, Pengchuan Zhang, Xiyang Dai, Bin Xiao, Lu Yuan, and Jianfeng Gao. Focal self-attention for local-global interactions in vision transformers. ArXiv, abs/2107.00641, 2021.

89. Pei Yu, Yinpeng Chen, Ying Jin, and Zicheng Liu. Improving vision transformers for incremental learning. CoRR, abs/2112.06103, 2021.

90. S.H. Ebenuwa; M.S. Sharif; M. Alazab; A.A. Nemrat, Variance ranking attributes selection techniques for binary classification problem in imbalance data. IEEE Access 2019, 7, 24649–24666. https://doi.org/10.1109/ACCESS.2019.2899578.

91. M. Khushi; K. Shaukat; T.A. Alam; I.A. Hameed; S. Uddin; S. Luo; X. Yang; M.C. Reyes, A comparative performance analysis of data resampling methods on imbalance medical data. IEEE Access 2021, 9, 109960–109975. https://doi.org/10.1109/ACCESS.2021.3102399.

92. H. Jegierski; S. Saganowski, An "Outside the Box" solution for imbalanced data classification. IEEE Access. 2020, 8, 125191–125209. https://doi.org/10.1109/ACCESS.2020.3007801.

93. T. Alam; C.F. Ahmed; S.A. Zahin; M.A.H. Khan; M.T. Islam, An effective recursive technique for multi-class classification and regression for imbalanced data. IEEE Access 2019, 7, 127615–127630. https://doi.org/10.1109/ACCESS.2019.2939755.

94. J. Maillo; I. Triguero; F. Herrera, Redundancy and complexity metrics for big data classification: Towards smart data. IEEE Access 2020, 8, 87918–87928. https://doi.org/10.1109/ACCESS.2020.2991800.

95. Y. Lu; Y.M. Cheung; Y.Y. Tang, Bayes imbalance impact index: A measure of class imbalanced data set for classification problem. IEEE Trans. Neural Netw. Learn. 2020, 31, 3525–3539. https://doi.org/10.1109/TNNLS.2019.2944962.

96. Zhuang Liu, Hanzi Mao, Chao-Yuan Wu, Christoph Feichtenhofer, Trevor Darrell, Saining Xie, A ConvNet for the 2020s, https://doi.org/10.48550/arXiv.2201.03545.

97. S. Chakraborty and C. Zhang, "Survival Prediction Model of Renal Transplantation using Deep Neural Network," 2020 IEEE 1st International Conference for Convergence in Engineering (ICCE), 2020, pp. 180-183, doi: 10.1109/ICCE50343.2020.9290695.

98. V. Hegde, S. Usmani, "Parallel and Distributed Deep Learning".

99. F. Chollet et. al, (2015). Keras. GitHub. Retrieved from https://github.com/fchollet/keras

100. C. Nardone, “Introduction to GPU Programming,” NVIDIA EMEA.

101. Somenath Chakraborty, Beddhu Murali, Investigate the Correlation of Breast Cancer Dataset using Different Clustering Technique, https://doi.org/10.48550/arXiv.2109.01538.

102. Chakraborty, S., Murali, B. (2022). A Novel Medical Prognosis System for Breast Cancer. In: Mandal, J.K., Buyya, R., De, D. (eds) Proceedings of International Conference on Advanced Computing Applications. Advances in Intelligent Systems and Computing, vol 1406. Springer, Singapore. https://doi.org/10.1007/978-981-16-5207-3_34.

103. L. Kaufman, P. Rousseeuw, Partitioning around medoids (Program PAM), Wiley Series in Probability and Statistics, Wiley (1990).

104. L. Kaufman, & P. J. Rousseeuw, (1990). Finding groups in data: An introduction to cluster analysis. New York: Wiley.

105. J. R. Quinlan, C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.

106. Praveen. Corona Hack-Chest X-Ray-Dataset.

Available online: https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset

107. Cohen, J.P.; Morrison, P.; Dao, L. COVID-19 Image Data Collection. arXiv 2020, arXiv:2003.11597. Available online: https://arxiv.org/abs/2003.11597

108. Paul, C.J. COVID-19 Image Data Collection. 2020. Available online: https://github.com/ieee8023/covid-chestxray-dataset.

109. Paul, M. Kaggle Chest X-ray Images (Pneumonia) Dataset. 2020. Available online: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia

110. B. V. Ginneken; M.B. Stegmann; M. Loog, Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. Med. Image Anal. 2006, 10, 19–40.

111. S. Jaeger; S. Candemir; S. Antani; Y. X. J. Wang; P.X. Lu; G. Thoma; Two public chest X-ray datasets for computer aided screening of pulmonary diseases, Quant. Imag. Med. Surg. 2014, 4, 475.

112. S. Jegou; M. Drozdzal; D. Vazquez; A. Romero; Y. Bengio; The one hundred layers tiramisu: Fully convolutional DenseNets for semantic segmentation. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 1175–1183. https://doi.org/10.1109/CVPRW.2017.156.

113. J. Deng, W. Dong, R. Socher, L. -J. Li, Kai Li and Li Fei-Fei, "ImageNet: A large-scale hierarchical image database," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248-255, doi: 10.1109/CVPR.2009.5206848.

114. A. Paszke; S. Gross; F. Massa; A. Lerer; J. Bradbury; G. Chanan; et al. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In: Advances in Neural Information Processing Systems 32 [Internet]. Curran Associates, Inc.; 2019. p. 8024–35. Available from: http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf

115. Olaf Ronneberger; Philipp Fischer; Thomas Brox,(2015); "U-Net: Convolutional Networks for Biomedical Image Segmentation". arXiv:1505.04597.

116. M. Tan; & Q. V. Le; (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. ArXiv, abs/1905.11946.

117. M. Tan; B. Chen; R. Pang; V. Vasudevan & Q. V. Le; (2019). MnasNet: Platform-Aware Neural Architecture Search for Mobile. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2815-2823.

118. Hangbo Bao, Li Dong, Furu Wei, BEiT: BERT Pre-Training of Image Transformers, available at: https://doi.org/10.48550/arXiv.2106.08254.

119. PyTorch Vision Models. Available at: https://pytorch.org/vision/stable/models.html.