

A Hybrid FL-Enabled Ensemble Approach for Lung Disease Diagnosis Leveraging Fusion of SWIN Transformer and CNN

by

Asif Hasan Chowdhury

19201128

Md. Fahim Islam

18101501

M Ragib Anjum Riad

18101472

Faiyaz Bin Hashem

18101278

A thesis submitted to the Department of Computer Science and Engineering
in partial fulfillment of the requirements for the degree of
B.Sc. in Computer Science

Department of Computer Science and Engineering
Brac University
September 2022

© 2022. Brac University
All rights reserved.

Declaration

It is hereby declared that

1. The thesis submitted is my/our own original work while completing degree at Brac University.
2. The thesis does not contain material previously published or written by a third party, except where this is appropriately cited through full and accurate referencing.
3. The thesis does not contain material which has been accepted, or submitted, for any other degree or diploma at a university or other institution.
4. We have acknowledged all main sources of help.

Student's Full Name & Signature:

Asif

Asif Hasan Chowdhury
19201128

Fahim

Md. Fahim Islam
18101501

Riad

M Ragib Anjum Riad
18101472

Faiyaz Bin Hashem

Faiyaz Bin Hashem
18101278

Approval

The thesis/project titled “Predicting a T20 Cricket Match Result While The Match is in Progress” submitted by

1. Asif Hasan Chowdhury (19201128)
2. Md. Fahim Islam (18101501)
3. M Ragib Anjum Riad (18101472)
4. Faiyaz Bin Hashem (18101278)

Of Summer, 2022 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on September 22, 2022.

Examining Committee:

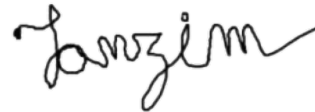
Supervisor:
(Member)



Md. Golam Rabiul Alam, PhD
Professor

Department of Computer Science and Engineering
BRAC University

Co-Supervisor:
(Member)



Md Tanzim Reza
Lecturer

Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi

Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

The study is approved by honourable Advisor Professor Dr. Md. Golam Rabiul Alam and Co-Advisor Lecturer Md Tanzim Reza

Abstract

The significant advancements in computational power create the vast opportunity for using Artificial Intelligence in different applications of healthcare and medical science. **A Hybrid FL-Enabled Ensemble Approach For Lung Disease Diagnosis Leveraging a Combination of SWIN Transformer and CNN** is the combination of cutting-edge technology of AI and Federated Learning. Since, medical specialists and hospitals will have shared data space, based on that data, with the help of Artificial Intelligence and integration of federated learning, we can introduce a secure and distributed system for medical data processing and create an efficient and reliable system. The proposed hybrid model enables the detection of COVID-19 and Pneumonia based on x-ray reports. We will use advanced and the latest available technology that can help to fight against the pandemic that the world has to fight together as a united. We focused on using the latest available CNN models (DenseNet201, Inception V3, VGG 19) and the Transformer model Swin Transformer in order to prepare our hybrid model that can provide a reliable solution as a helping hand for the physician in the medical field. In this thesis, we will discuss how the Federated learning-based Hybrid AI model can improve the accuracy of disease diagnosis and severity prediction of a patient using the real-time continual learning approach and how the integration of federated learning can ensure hybrid model security and keep the authenticity of the information.

Keywords: AI, VGG19, Inception V3, DenseNet201, SWIN Transformer, Federated Learning, Privacy.

Dedication

The Thesis is dedicated to all the people in the world who are suffering from different diseases. May Allah help us all to become his faithful servant. May Allah help us in our entire life time. Thanks to Allah for helping us to find way to accomplish our goal.

Acknowledgement

Firstly, all praise to the Great Allah for whom our thesis has been completed without any major interruption.

Secondly, to our advisor Dr. Md. Golam Rabiul Alam sir, and co-advisor Mr. Md Tanzim Reza sir for their kind support and advice in our work. They helped us whenever we needed help.

And finally to our parents without their throughout support it may not be possible. With their kind support and prayer, we are now on the verge of our graduation.

Table of Contents

Declaration	i
Approval	ii
Ethics Statement	iii
Abstract	iv
Dedication	v
Acknowledgment	vi
Table of Contents	vii
List of Figures	ix
List of Tables	x
Nomenclature	xi
1 Introduction	1
1.1 Research Problem	2
1.2 Research Objective	4
2 Literature Review	6
2.1 Background Studies	6
2.1.1 CNN	6
2.1.2 Convolutional Layers	6
2.1.3 Pooling	7
2.1.4 Sequential Model	7
2.1.5 VGG 19	8
2.1.6 DenseNet 201	8
2.1.7 Inception V3	9
2.1.8 SWIN Transformer	9
2.1.9 Multi-layer Perception	10
2.1.10 Ensemble Model	10
2.1.11 Federated Learning	11
2.2 Related Works	11

3	Methodology	16
3.1	Dataset Collection Process	17
3.2	Dataset Pre-process Step	17
3.3	Classification and Decision Classifier	18
3.4	Brief Work Steps	18
3.5	Classification into 3 Groups	18
3.6	Fusion Model	19
3.7	Federated Learning Centralized Server	19
4	Result and Analysis	20
4.1	VGG-19	20
4.2	DenseNet 201	21
4.3	Inception V3	22
4.4	SWIN Transformer	23
4.5	Fusion Model	25
4.6	Fusion Model Sum Outcome Graph	26
4.7	Fusion Model Average Outcome Graph	28
4.8	Federated Learning Based Outcome	29
4.9	Comparative Analysis	30
4.10	AUC-ROC Outcome	33
4.11	Overview Based on Different CNN Model	36
5	Conclusion	37
	Bibliography	39

List of Figures

2.1	Neural Network	7
2.2	DenseNet 201 Structure	8
2.3	Dense Block Architecture	8
2.4	Hybrid Ensemble Transfer Learning Model	10
2.5	Transfer Learning and Transformer Learning based Fusion Model	11
2.6	Federated Learning	12
3.1	Top Level Work View of Proposed Lung Disease Detection System	16
3.2	Dataset Details	17
3.3	Design Flow	19
4.1	VGG-19 Training Accuracy and Validation Categorical Accuracy	21
4.2	VGG-19 Training Loss and Validation Categorical Loss	21
4.3	DenseNet201 Training Accuracy and Validation Categorical Accuracy	22
4.4	DenseNet201 Training Loss and Validation Categorical Loss	22
4.5	Inception V3 Training Accuracy and Validation Categorical Accuracy	23
4.6	Inception V3 Training Loss and Validation Categorical Loss	23
4.7	SWIN Transformer Training Accuracy and Validation Categorical Accuracy	24
4.8	SWIN Transformer Training Loss and Validation Categorical Loss	24
4.9	Training Accuracy and Validation Accuracy	26
4.10	Training Loss and Validation Loss	26
4.11	Line Model for Training Output	27
4.12	Line Model for Validation Output	27
4.13	Training Accuracy and Validation Accuracy	28
4.14	Training Loss and Validation Loss	28
4.15	Different Hospitals based Training Accuracy and Validation Accuracy	29
4.16	Different Hospitals based Training Loss and Validation Loss	29
4.17	VGG-19 Based Confusion Matrix	30
4.18	DenseNet 201 Model Confusion Matrix	30
4.19	Inception V3 Model Confusion Matrix	31
4.20	SWIN Transformer Model Confusion Matrix	31
4.21	Fusion Model and Federated Learning Based Confusion Matrix	32
4.22	VGG-19	33
4.23	Inception V3	34
4.24	DenseNet 201	34
4.25	SWIN Transformer Model	35
4.26	Fusion Model and Federated Learning Based	35

List of Tables

4.1	Model Comparison Table	36
-----	----------------------------------	----

Nomenclature

The next list describes several symbols & abbreviation that will be later used within the body of the document

AI Artificial Intelligence

CNN Convolutional Neural Network

COVID – 19 Corona Virus Disease 2019

FL Federated Learning

RNN Recurrent Neural Network

SARS Severe Acute Respiratory Syndrome

SML Supervised Machine Learning

SWIN Shifted Window

VGG – 19 Visual Geometry Group

Chapter 1

Introduction

Integrating artificial intelligence with medical science has created a new dimension to the treatment world. Computer-assisted diagnosis can help doctors to sense any forthcoming lethal diseases beforehand. Nowadays, doctors across the globe tend to rely more on artificial intelligence as it is improving swiftly. Humans are proficient in terms of taking complex decisions but when the quantity comes on a larger scale it is wise to take advantage of the computational capabilities of machines.

Artificial intelligence can overcome human limitations. For example, AI can carry out repetitive tasks ensuring accuracy whereas humans can make mistakes out of weariness. Moreover, AI can diagnose diseases from a larger dataset of patients which is quite impossible for physicians to go through every detail with precision. Many diseases can be detected by analyzing images even before symptoms. As a result, doctors get enough time to provide treatment and other clinical assistance. In this thesis, we proposed an ensemble method to detect lung diseases. We focus to achieve preferable accuracy with better performance therefore we build an ensemble method for our thesis. We have used the latest pre-trained models to develop our own hybrid model that can work as a helping hand to medical practitioners.

Needless to say, patient data is very sensitive and any sort of mismanagement may cause privacy violation. Federated learning is a wonderful approach for mitigating the vulnerability of data security. It prevents data exposure while fetching the local models. Thus, patient data does not need to travel through the global server for upgrading the model. We want to build a network where different hospitals will stay connected together and share their effective treatment model for the betterment of the mass people

Hackers usually hack sensitive medical information for evil purposes. They can blackmail any individual using their personal information, a fraud scam with health insurance. Also, hospitals can face lawsuit problems, possible fines, and reputational costs. Sometimes hackers do sell this information in exchange for money.

1.1 Research Problem

SARS is a severe acute respiratory syndrome that is caused by the SARS-associated coronavirus. It was first identified in 2003 during an outbreak in China which was spread in 28 other countries afterward. It can spread by air like cold and influenza. In that outbreak, more than 8000 people were infected and 774 died. The virus that caused this outbreak was known as SARS-CoV. Recently, in December 2019 another coronavirus named SARS-CoV-2 was identified in Wuhan, a city in China that caused a massive pandemic of respiratory illness, called COVID-19. This pandemic has affected 228 countries and territories worldwide. According to the world meter, 608,997,597 COVID-19 cases have been reported and it caused 6,499,213 deaths [23]. According to the reports of the World Health Organization, in 2019 pneumonia is accountable for 14 % of deaths of under 5 years old [24].

The COVID-19 pandemic affected the economy, human life, health sector, and social aspects worldwide. Nearly ten million people are now at great risk of falling into extreme poverty. Many people lost their jobs during the pandemic and were unable to feed their families. Millions of enterprises and businesses shut down. Due to the closure of borders, travel restrictions, and trade restrictions for a long period of time many businessmen, and producers did not get a chance to access markets and sell their products which also disrupted the world food supply chain. 1.6 billion students got out of school.

There are different types of testing methods available for the COVID-19 test. Molecular tests and antibody/antigen tests are mostly used worldwide during the pandemic. To stop the spreading of COVID-19 immersively it is needed to detect more positive cases as early as possible and isolate them from others. But in the low and middle-income countries, we have seen that due to a lack of equipment and testing kits it took time to conduct COVID-19 tests on a vast amount of patients who were already affected which caused the massive spreading of the virus later on.

RT-PCR(Reverse Transcription Polymerase chain reaction) is the most commonly used method to detect COVID-19 which is expensive, less sensitive, and requires specialized medical personnel. But in addition to the detection of COVID-19, to learn about the severity and how much it affected the lungs X-Ray image scanning is crucial.

Many studies have been published in this domain until now but most of the research had used very small datasets to perform the studies. Further studies are required to improve the existing research in this domain.

Another problem that our research is focusing on is data security. Medical data is very sensitive. The leakage of confidential medical data can cause huge problems. Data breaches are malware-based attackers who hack the database for selling the personal information of patients to a third party in exchange for money. Ransomware attackers hack the server or disable the functionality until they get a handsome amount of money. National hackers can be a big threat to the healthcare system of a country as a part of cyber war.

Disease Detection allows doctors and healthcare providers to monitor patients' health conditions. It is important to monitor patients who are admitted to the hospital or are not in a good condition. In the south Asian region, a large number of people are dependent on public hospitals due to their low income. As a result, these hospitals do have more admitted patients than the capacity. For this reason, doctors can not manage to monitor all the patients on a regular basis and emergency patients face the consequences sometimes. An ideal ratio of physician and patient is 1:1000. But according to a World Bank report, in 2019 the doctor and patient ratio in Bangladesh is 0.637:1000, which means for 1000 patients there are only 0.637 doctors present in Bangladesh. It is in the second position from the bottom in South Asia. In Dhaka Medical College and Hospital, the most renowned public hospital in the capital of Bangladesh, the bed capacity is 2,400 whereas 4,000 patients get admitted and take treatment at a time every single day [8].

Aged people in our home also need support and proper monitoring as they bear many health issues. In most cases, they can not express their health conditions to their surroundings which lead their health in danger. Patients with blood pressure and diabetes problems also need proper healthcare and monitoring to understand their current condition and health risk. In many cases, the patients get a severe heart attack and even death for not understanding the health risk on time and late admission to the hospital. According to reports, in the USA 40% of the whole population are suffering from chronic diseases whereas 7 out of 10 death cases happen because of chronic disease and heart attack. Moreover among all the existing diseases that modern doctors are treating 46% of them are chronic diseases [17].

According to the reports, the protected health information of nearly 50 million Americans got hacked in 2021. The number of reported cases of healthcare data breaches increased by 19% in 2021. Among all healthcare data breaches, hacking accounted for 35% cases in 2016, which increased to 74% in 2021. Moreover, 5% of the cyber-attacks are just for fun whereas 91% of hacking purposes are financially driven. A very important point to be noted, 50% of those hospitals whose devices are interconnected with a common internet connection are most vulnerable to cyber-attacks [22].

1.2 Research Objective

The main objective of this research is to develop a system to detect lung diseases using the Deep CNN model to prevent severe conditions and accidental deaths. The research will also focus on the privacy of the patients and ensure data security using the federated learning approach.

The study will focus on how the existing transfer learning approaches can be improved with the integration of the Swin Transformer model. In the traditional ML model, we need to train the model from scratch which requires a large amount of data and expensive computational power. Our research aims to use transfer learning models which are already pre-trained so that we can use them as our starting point and reach our goal with less computational power and better accuracy.

Our research aims to develop an Fusion Deep CNN model which will average the result of traditional transfer learning models like VGG-19, Inception v3, and DenseNet201 and integrate with a famous transformer model, SWIN Transformer to ensure better accuracy and prediction. One of the objectives of this research is to prevent lung patients from Acute respiratory distress syndrome (ARDS) which is a life-threatening condition where the lungs can not help to provide enough oxygen to other body organs.

In this research, our goal is to use a federated learning approach for better accuracy, lower latency, less power consumption, and ensuring data security. Federated Learning is a collaborative machine learning approach without training centralized data. In conventional machine learning approaches, we require centralized training data on a central machine. It can cost data privacy, accuracy, and high power consumption.

Our goal is to establish a model that enables our hospital devices to learn collaboratively a shared prediction model without sharing any training data using the Federated learning approach. Hospitals' local devices will download our multi-level classification model, collect datasets from patients, train the datasets using the model and summarize the output as an update. Then, these updates from the local devices will be sent to the central cloud or central server. After that, these updates will be averaged to improve the shared model. This improved collaborative model can be used by the hospitals' local devices to detect the diseases and predict the current risk factor of the patients with better accuracy, less power consumption, and low latency.

Another aim we are focusing on is to ensure data security of the patient's data. As our model will work based on confidential datasets of patients, it is important to ensure data privacy. In federated learning, the training dataset is kept on the local devices. So, it does not require any data pool which keeps the data safe. But according to researchers, there are some shortcomings in federated learning which we are aiming to solve.

Federated learning relies on a central server. So, if the central malfunctions it will give inaccurate model updates. Also, the data of the datasets can be tampered by someone inside the organization. To prevent data tampering we are aiming to keep the data within the hospital's servers. This will keep the local model secured also the data that will be generated through patients' medical tests

- Lung disease detection using the Deep CNN model to analyze the severe conditions of patients.
- Improve existing Transfer Learning models by building a new Fusion model that combines Transfer Learning and Transformer Learning.
- Integration of Shifted Window (SWIN) Transformer model for better accuracy and detection.
- Utilization of federated learning models for ensuring data security, low latency, less power consumption, and better accuracy.

Chapter 2

Literature Review

2.1 Background Studies

2.1.1 CNN

CNN stands for Convolutional Neural Network, a class of Deep Neural Networks that is most commonly applied to visual image datasets. It does not work with matrix multiplication like other neural networks. Unlike other neural networks, it works with a special algorithm called convolution. The main role of the CNN model is to reduce a visual image into a form that is easier to process, without losing specific features that are important for ensuring a good prediction.

We know that an image is a matrix of pixel values. An RGB image has three planes whereas a grayscale image is represented as a single plane. In the convolution technique, we take a kernel/filter which is a 3×3 matrix, and apply it to the image to get a convolved feature as an output. After that, the convolved feature is passed to the next layer.

2.1.2 Convolutional Layers

CNN is a composition of multiple layers of artificial neurons where artificial neurons are some functions that calculate the weighted sum of multiple inputs and outputs activation values. When we input an image to a ConvNet it generates several activation functions and passes it to the next layer.

The first layer extracts some basic features. The second layer extracts some intermediate features. That's how much we move deeper or increase the layers it can identify more complex features like faces and objects. Based on the activation map the final convolution layers output a set of scores between 0 and 1 which specifies the 'class' of the image.

2.1.3 Pooling

The pooling layer is responsible for reducing the size of convolved features. It decreases the requirement of computational power to process the data by reducing the dimensions. There are two types of pooling. One is max pooling. Another one is average pooling. In max pooling, when an image is covered by the kernel/filter it finds the maximum value of the pixel. It reduces the dimension along with discarding the noisy activation functions. On the other hand, average pooling outputs the average of all values from the image covered by kernels. Max pooling performs better than average pooling in terms of dimension reduction.

2.1.4 Sequential Model

The sequential model is one of the fundamental models of the Deep Learning World. Machine learning models that input and output data in sequences are known as sequential models. Sequential models are preferable to handle sequential data. An important thing about sequential models is that data we are working on are not independently distributed. Sequential models are quite popular for speech recognition, voice recognition and natural language processing.

The Sequential model is the process of generating a sequence of values from a set of input data. In this processed X-ray dataset of COVID-19, Pneumonia and normal points are dependent on different points of the dataset. DNA sequences and meteorological data can be good examples of sequential data. Datasets can have sequential forms. There are datasets of sounds and sentences that rely on one another. Recurrent Neural Network and Long Short-Term Memory are the examples of sequential model.

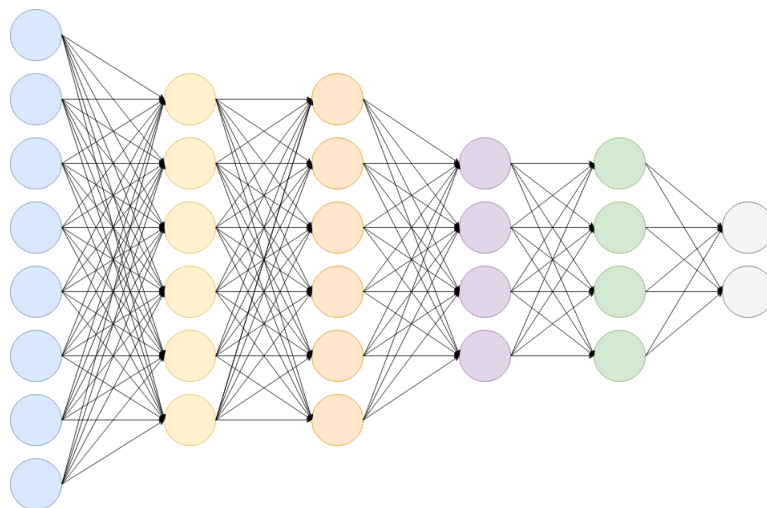


Figure 2.1: Neural Network

2.1.5 VGG 19

VGG 19 is the latest version of the Visual Geometry Group model series. This model series is the successor of the AlexNet. This model consists of 19 layers. Out of 19 layers, 16 layers are Convolutional layers and 3 fully connected layers and 5 MaxPool layers, and 1 SoftMax layer.

In order to categorize the photos into 1000 object categories, Simonyan and Zisserman (2014) presented the VGG19. There are numerous 3 x 3 filters used by each convolutional layer. Because each convolutional layer uses numerous 3 x 3 filters, it is a highly well-liked technique for classifying images.

2.1.6 DenseNet 201

A typical convolutional neural network is started with an input image and runs through the network to get a predicted label. The output of the previous convolutional layer is used by the subsequent convolutional layer, which receives the input image from the previous layer and constructs an output feature map.

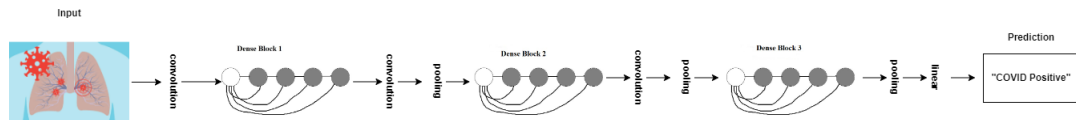


Figure 2.2: DenseNet 201 Structure

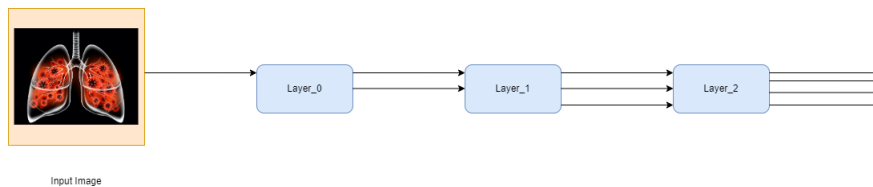


Figure 2.3: Dense Block Architecture

But, In a DenseNet architecture, All layers are densely connected. That means an inter-layer connection exists between each layer. Moreover, L connections exist between L levels, one between each layer and the layer below it. So, There are $L(L+1) / 2$ direct links in the network. The feature maps of all layers before it are utilized as inputs for each layer, and its own feature maps are used as inputs into all levels after it [5]. The DenseNet architecture's dense connectivity can be represented as:

$$x(l) = H(l)([x(0), x(0), \dots, x(l - 1)]) \quad (2.1)$$

2.1.7 Inception V3

The third generation of Inception convolutional neural network designs is known as Inception-v3. Among other improvements, the Inception-v3 convolutional neural network architecture makes use of Label Smoothing, Factorized 7 x 7 convolutions, and the addition of an auxiliary classifier to move label information lower down the network (along with the use of batch normalization for layers in the sidehead).

The Inception architecture is built to function successfully even when memory and compute resources are severely limited. Though the architectural simplicity of VGGNet is appealing, it comes with a considerable computational cost when assessing the network. As Inception is lower while higher-performing successors, It is feasible to utilize the e Inception networks in big-data scenarios.

The architecture of an Inception-v3 is progressively built, step-by-step.

1. Factorized Convolutions: This decreases the number of parameters used in a network, which lowers computational efficiency. It also monitors the effectiveness of the system.
2. Smaller convolutions: This causes training to go more quickly by substituting smaller convolutions for larger ones. Say a 5 X 5 filter has 25 parameters; replacing it with two 3 X 3 filters results in only 18 ($3*3 + 3*3$) parameters [14].

2.1.8 SWIN Transformer

Swin Transformer is stated as Shifted Windows Transformer. This is basically a hierarchical Transformer that is computed with shifted windows. To address the challenges of differences between two domains, such as large variations in the scale of visual entities and the high resolution of pixels in images compared to words in the text, this hierarchical Transformer or Swin Transformer was proposed.

In Swin Transformer architecture, first of all, it splits an RGB input image into non-overlapping patches using modules like ViT(Vision Transformer), Each of the split patches are called “token” which features are set as a concatenation of raw RGB pixel values. For our research work, we use a patch size of (2 X 2), so the feature dimension was $2 \times 2 \times 3 = 12$. Stage 1 then projects this raw-valued feature to any dimension(denoted as C) by applying a linear embedding layer to it. Then patch tokens have some self-attention modified Swin Transformer blocks put to them. This Transfer blocks maintained the number of tokens ($H/2 \times W/2$), The number of tokens is decreased via patch merging layers as the network becomes deeper in order to create a hierarchical representation. The 4C-dimensional concatenated features are applied to a linear layer in the first patch merging layer, which concatenates the features of each set of 2 X 2 neighboring patches. This results in a multiple of $2^2 = 4$ (2 downsampling of resolution) token reduction, and the output dimension is set to 2C. The resolution was then maintained at $(H/4 \times W/4)$ after the application of Swin Transformer blocks.

In Stage 3 and Stage 4, respectively, this process was done twice. Together, these steps result in a hierarchical representation [19].

2.1.9 Multi-layer Perception

MLP algorithm inherits the microstructure of the human brain but to a limited extent. MLP contains three layers input layer, hidden layer, and output layer. The hidden layer is the middle one which can be split into multiple layers. The algorithm starts with a random weight value between -1 to 1. Here, weights behave like a catalyst for the model. Subsequently, the weighted sum is forwarded to the activation function. To find out the most accurate weights which would reduce the cost function, back-propagation is applied. The iteration goes on and updates the weights till the gradient of input and output converge.

As per our paper, we determine the disease of a patient with the severity of the illness with the help of a parameter named risk_factor. So, we need to classify the disease and the risk_factor at the same time for a particular patient.

2.1.10 Ensemble Model

The ensemble method creates a hybrid model with higher predicted accuracy than a single algorithm by combining many machine learning models. We created models for VGG19, Inception-v3, DenseNet, and SwinTransformer using our dataset for each. Then, by averaging VGG19, Inception-v3, and DenseNet, we combined them into a single hybrid model using the ensemble approach. Then, using the average ensemble method, we built another hybrid model utilizing the first hybrid model and the Swin Transformer model. In comparison to the individual models of VGG19, Inception-v3, DenseNet201, and Swin-Transformer, the final model yields better accurate and predictive outcomes according to the ensemble approach. Below is the final Ensemble model:

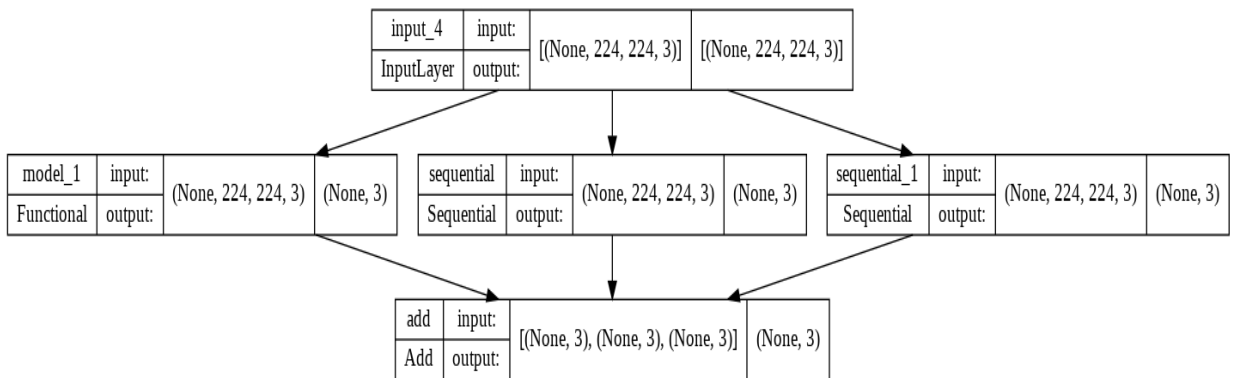


Figure 2.4: Hybrid Ensemble Transfer Learning Model

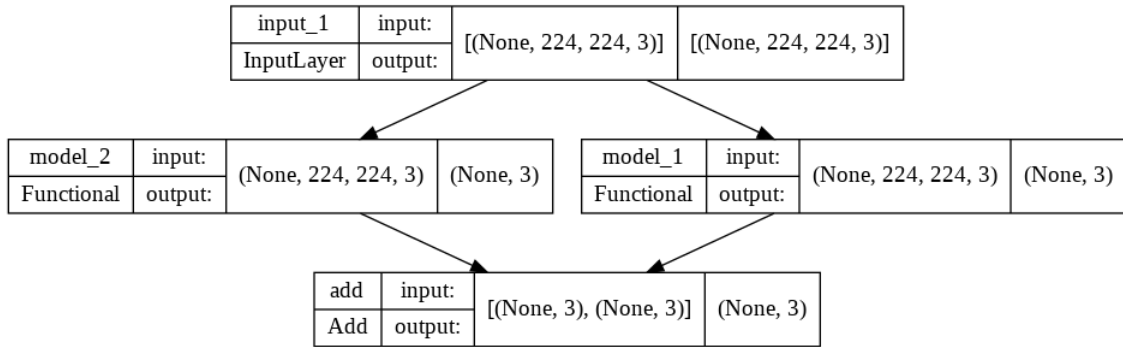


Figure 2.5: Transfer Learning and Transformer Learning based Fusion Model

2.1.11 Federated Learning

Federated Learning: Federated Learning is a decentralized system of machine learning models. Generally, in machine learning, we train our data that is gathered from various edge devices, including smartphones, computers, and other gadgets. After that, we took all the data together in a centralized server and make a machine learning model. But, this process of taking data from each device is a big threat to privacy. The sensitive data could be revealed to every one of the central server administrative. For this Federated learning is used to train the central model on any decentralized data. In this system, the central model will have a copy of every local model those are connected with that central server. And the user-inputted data can be used by the local models to train and gradually learn, improving over time. Also, the devices can transfer their trained local model copy back to the central server. By this, data privacy will be ensured.

2.2 Related Works

Firstly, we get to know about the current threats that health care systems are facing, from [11]. Sometimes it is difficult to detect these types of threats or attacks. These attackers can be divided into many types according to their working strategy and background. Data breaches are malware-based attackers who basically hack the data for selling them to others in exchange for money. Ransomware attackers typically disable the server or the functionality of computer monitoring systems until they get a handsome amount of money that can be threatening to patients. Moreover, evil employees can turn up against their own organizations due to some past records and be a threat to MCPS. We want to address these issues and come up with a better system. When our system will be functional it will have good security features that will give people reliance over the system and reduce the chance of hacking and data leakage.

In addition, in paper [11], Social and national hackers can be a threat to the MCPS to exploit the health care network system of other countries as a part of cyber war. DOS attacks are the most common types of attacks in MCPS which can retrieve patients' data and deliver it to the wrong hands. Another common type of attack is UDP (User Datagram Protocol). It targets sending large amounts of data pack-

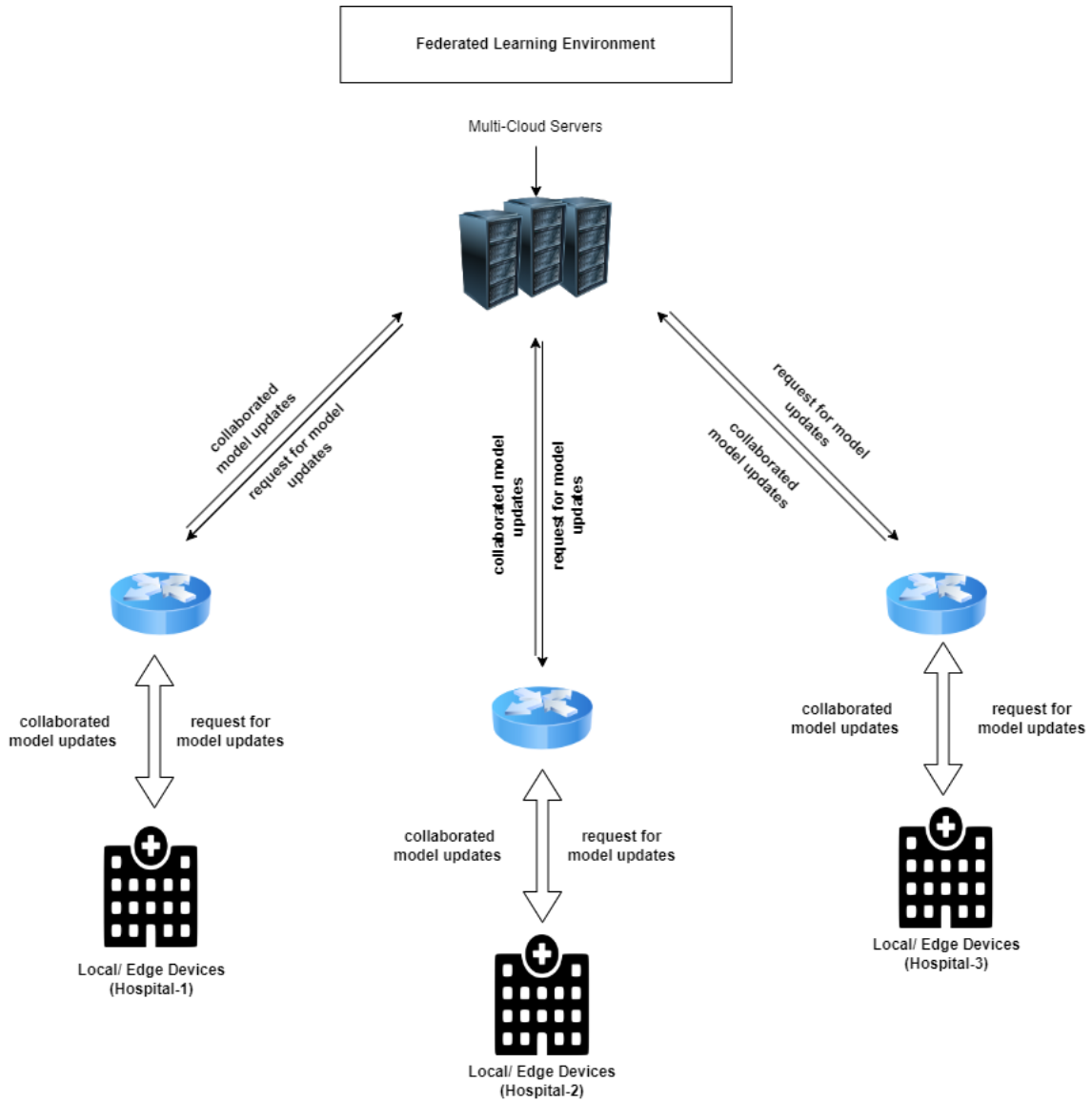


Figure 2.6: Federated Learning

ets to the server until the firewall of the system gets exhausted. Internet control of message protocol (ICMP) is another popular type of attack where hackers send large amounts of ICMP requests which exploit the available bandwidth and block authorized access to the server.

According to Schneble et al. in paper [12], an Intrusion Detection System (IDS) using a federated learning system can be a solution to the MCPS cyber-attacks. Basically, federated learning is a distributive machine learning algorithm that improves its current model by learning from data and sending the updates to the cloud using encrypted communication. To detect attacks, federated learning goes through an organized design process. The first step is called the clustering of patients. The process starts with registering the patients with the server. The patients are clustered into groups formulated based on their history so that the detection accuracy can get better and the training time can be minimized at the same time. Network limitation has a significant impact on the model, for example, if there is a tiny group

of patients, the model will encounter a shortage of data to learn and get updated. So, it is recommended to have a larger group of patients to establish a model which would be more efficient.

In addition, Schneble et al. further explain in paper [12] that training and updating the model comes into view. In this process, the server has to go through a lot of computation and it stores the updated federated model in the server afterward. If any patients fail to send data to the server for calculation, the iteration will be decreased by one to fetch the missing data. Once the model is updated after calculation, it is delivered to all clustered mobile devices and the method continues until the model converges, which means the result does not differ remarkably between iterations. The mobile devices with the updated model can be either in training mode or in testing mode. When the device is in training mode, it contributes to the global model of the server by making predictions and delivering updates. On the other hand, when the device is in testing mode, it only makes predictions. Testing mode saves communication expenses as it does not submit anything to the server.

Sometimes, the system is configured by limiting mobile devices used per iteration to reduce the whole computation costs without affecting respective training time or detection accuracy since a lesser amount of devices indicate less traffic overhead between the devices and the server. The final design process is attack detection where the model is supposed to detect normal and abnormal behavior. Here, correlated factors are connected by nodes. For instance, pulse rate and oxygen saturation must be changed accordingly. In abnormal behavior, the readings are out of the anticipated range or unexpected in terms of relationships among nodes. Data modification and data injection may create abnormalities in the model. Data modification may manipulate node values to lower than usual. Additionally, data injection may corrupt the correlations among the features. If an abnormality is found, an alert is produced for the medical staff to intercede.

Li et al. have discussed the mechanism of federated learning for securing data and overcoming challenges [16]. First, the authors have mentioned the challenges one might face implementing federated learning. Such as expensive communication, system heterogeneity, statistical heterogeneity, and privacy concerns. Then, Li et al. come up with solutions to those challenges. For more communication efficiency, the authors have pointed to multiple methods like local updating, and decentralized training. As model training may differ for devices' hardware specifications, the authors recommend using an asynchronous scheme that applies an optimization algorithm parallelly. Sometimes, data for federated learning are divided among devices non-identically, to resolve this problem meta-learning and multitask learning have been added to federated settings [10]. A framework named MOCHA [6] is used for learning separately but related models for each device. Bonawitz et al. [4] introduced a protocol for individual model updates. In this protocol, the central model will not be able to see the local updates but will observe the aggregated result at the end of every round. This method is an inspiration for our proposed model where we will implement federated learning to overcome privacy leakage.

Shivshan et al. explained how machine intelligence is changing the momentum in

healthcare and medical healthcare systems [9]. The authors have shown a list of algorithms and challenges that might come from processing patients' data. Machine Intelligence can diagnose diseases with more accuracy and find new treatments for sufferers. It can reduce patients' hassle by creating a model for MCPS. A cloud-based machine intelligence model can be implemented to detect diseases faster with the model which would quicken the treatment and reduce expenses [2].

Shivshan et al. discussed healthcare applications that can help doctor to detect diseases by analyzing a large dataset. It enables real-time monitoring of patients. Machine intelligence can predict the progress of patients' health conditions under treatment as well. Doctors can monitor patients' conditions with some actuators and apply machine intelligence-generated suggestions for treatment. In our proposed medical healthcare system, we want to develop a model which can detect the disease and notify the doctor constantly about the patients. So, data collection and processing are too important for our project.

In the paper [9], the authors categorized patients' data. For example, personal data where family history is included, short-term and long-term personal recordings, real-time recording, and unstructured data. Weng et al. [7] proposed a natural language processing algorithm to classify data and put those data into the domain to which they belong to. For transferring medical data, Kocabas et al. [3] conveyed a survey on encryption algorithms. Our approach will be different compared to this since we want to ensure the highest level of security with blockchain.

In this paper [15], The author used the most applicable image classification model CNN and RNN. The author combined CNN-RNN for getting better results in image classification. This study can be used in medical disease detection systems. As medical diagnosis images appear in several shapes and qualities, so we can use the data pre-processing technique from this[15] paper that can work with any type of input shape and quality. Moreover, The combined use of CNN-RNN architecture increases the accuracy of the central diagnosis model of our research.

Battineni et al. in their paper [13] proposed to use of machine learning models in chronic disease diagnosis. They present the trends of using predictive models in the diagnosis and forecast of diseases. Moreover, the studies show there are no specific methods to determine the best result in chronic disease as each method has come with its own advantages and disadvantages. Artificial intelligence(AI) is a technology that represents intelligent behavior using computer knowledge. This technology nowadays comes with more accurate results that authors defined as human intelligence. Furthermore, the paper shares the study that supervised machine learning(SML) is followed by the highest number of studies and easy to make predictive models. Although, it gives important to use unsupervised machine learning with deep learning models that can handle medical datasets with no clear directions. These predictive models also can be used in our research topic as all the medical needs to predict patients' disease in the diagnosis process.

In this paper [1], the author has proposed an approach to predict stroke by integrated machine learning models. The study of this paper shows how clinical data

can be filtered in machine learning models. Clinical data may have some errors or missing data, by performing missing data imputation the data will be able to be used in machine learning models. Also, the feature selection part is highlighted, as CHS datasets have a large number of attributes, information, measurements, etc. the machine learning-based algorithms for the feature selection technique will make the model better than manual selection. They used three algorithms for feature selection: forward feature selection, L1 regularized logistic regression, and “conservative mean” feature selection. After these, the dataset will be able to make a machine learning model for stroke detection. This missing data imputation and feature selection can be added to our research as the medical will need to detect various diseases with various types of datasets.

Kassania et al. proposed a deep CNN approach to detect COVID-19 from X-ray and CT images [18]. The authors tried to get better of the over-fitting issue in deep learning due to the small number of training images by using a transfer learning strategy. Firstly, Kassania et al. applied the image normalization technique to get better visual quality of input images. In the feature extraction step, the authors used a transfer learning strategy to lessen computational resources and accelerate the convergence of the network as their dataset is very limited. Finally, the authors developed a web-based application to assist doctors in detecting COVID-19 by uploading X-ray or CT images. This research contains some limitations such as few training data samples and security assurance. In our paper, we fed our model with a comparatively larger training set while ensuring the privacy of patients. We have developed a fusion model to get more accuracy in an efficient way.

Hemdan et al. introduced a framework of deep learning named COVIDX-Net to diagnose COVID-19 from X-ray images [14]. The COVIDX-Net framework consists of seven DCNNs of different architectures and those are VGG19, DenseNet201, InceptionV3, ResNetV2, InceptionResNetV2, Xception, and MobileNetV2. The authors fed their models with a limited number of data. Hemdan et al. got better results using VGG19 and DenseNet201 whereas the result with InceptionV3 was not satisfactory. This paper shows a comparison among existing DCNN models with limited data.

The existing conventional models have slow computational power and have large sizes. Using Swin Transformer model can increase the computational speed with the size of the image. In this paper Yeonghyeon Gu, Zhegao Piao, Seong Joon Yoo proposed a fusion model that combines Swin transformer blocks and a lightweight U-Net type model that has encoder-decoder structure [20].

In terms of learning global and remote semantic information CNNs have major drawbacks. In the paper, Yun Jiang, Yuan Zhang, Xin Lin, Jinkun Dong, Tongtong Cheng, Jing Liang proposed a new 3D medical picture segmentation model which is named as SwinBTS. This model is a combination of transformer, CNNs and encoder-decoder structure. 3D swin transformer is used for extracting contextual data whereas the convolutional models are used for upsampling and downsampling [21].

Chapter 3

Methodology

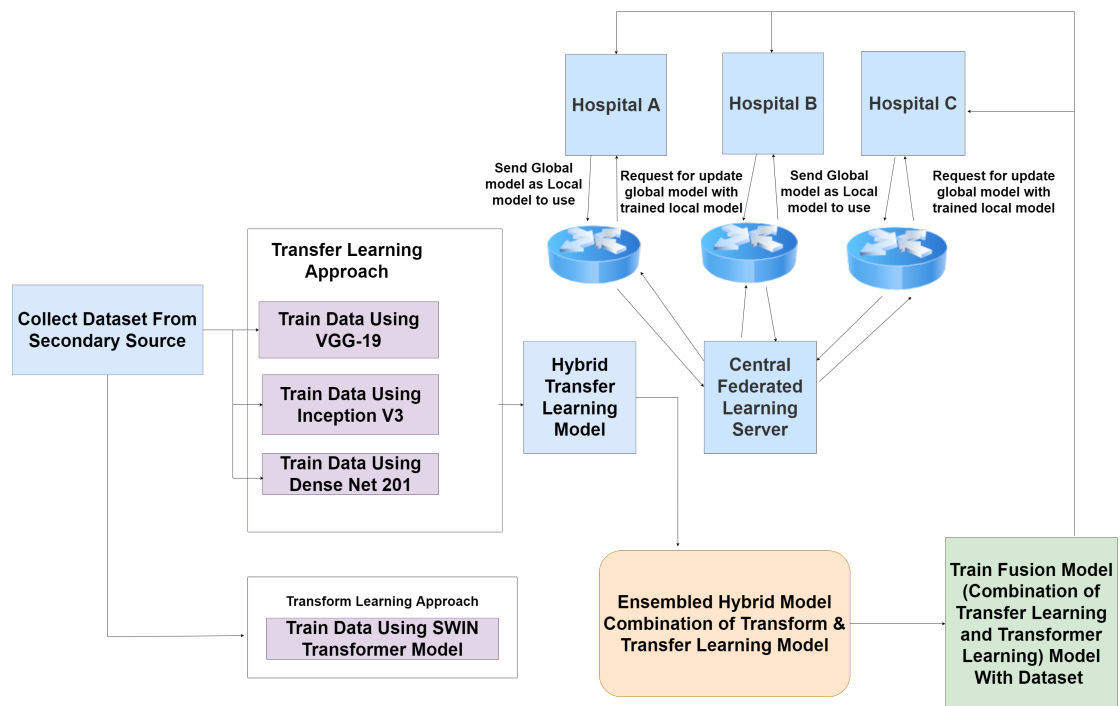


Figure 3.1: Top Level Work View of Proposed Lung Disease Detection System

In Figure 3.1, ext, we have shown our proposed work plan. At first, we will collect the data. Next, we will start to train the existing transfer-based learning models using our dataset. We have planned to work with VGG-19, Inception V3, DenseNet-201 Model. Then we will save the best-trained model individually and Ensemble them together. Next, we will start to train the transformer-based model that we decided to work with the SWIN transformer Model. We will also train this model using our dataset.

Next, we will combine the trained SWIN transformer model with the Transfer learning-based ensemble model to create our own hybrid model. Next, we will train the hybrid model with our available dataset to complete our model training and validation. Moreover, we are using a federated learning approach to secure each individual model that will be held by the hospitals. Hospitals will share the best finding outcome with the global server to ensure better accuracy and outcome.

3.1 Dataset Collection Process

Data collection is the most significant task to start building the CNN model. The initial step of the work plan is to collect data from different secondary sources. We know medical data is sensitive and difficult to manage. We initially looked to find medical data from different hospitals. In most of the cases, we were not able to manage the same disease-related information. Then, we looked at different diseases related papers for the dataset. We get different datasets but still being open source datasets some people alternated the large dataset with wrong files and corrupted files. Thus we have become careful enough during the selection process of the dataset. We have talked to several paper publishers for their datasets. That's how we managed the dataset.

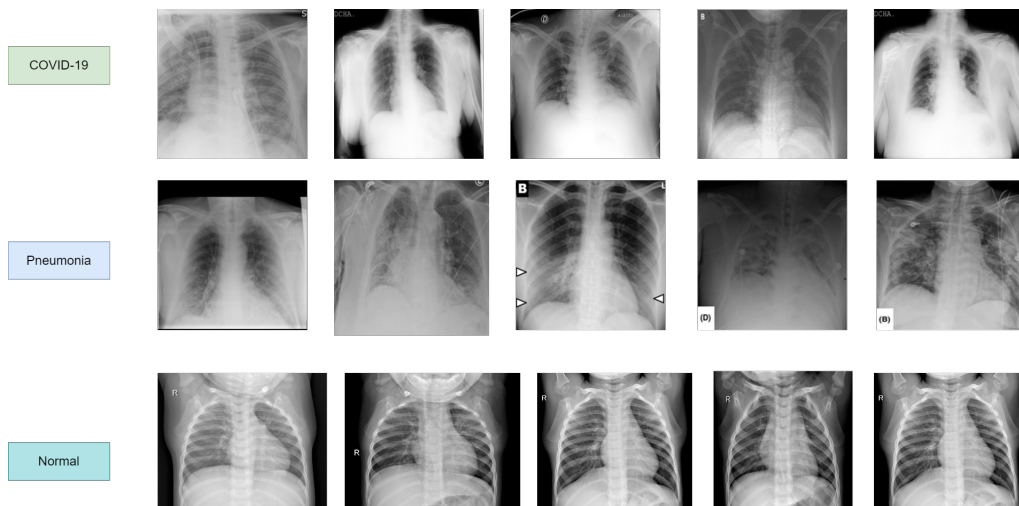


Figure 3.2: Dataset Details

3.2 Dataset Pre-process Step

we processed the image data using the available pre-processing techniques. We have identified corrupted images in the first place. Then, we removed the wrong image data. For example, the X-Ray dataset contains CT-Image data. Then, we figure out the number of images available for the training process and increase the data by the data augmentation process. Next, we used re-scaling, mage-rotation, horizontal-rotation, and zoom-range for the image processing process.

We need to split the dataset into two categories known as training sets and testing sets. The train set will do the task of training the dataset and preparing local models for different hospitals. The test set will do the job of testing the predicted diseases. We followed the convention of the training set(80%) and testing set(20%). We will have more accurate results if we can increase the ratio of the training set.

3.3 Classification and Decision Classifier

As our paper requires multiple predictions, we implement the VGG-19, Inception-V3, and DenseNet 201 for it. In addition, we have used a transformer-based learning model that is the SWIN transformer Model. These algorithms are widely used for classification problems. We build our model with these classification algorithms. Finally, the model will exhibit real-time predictions for each individual. The probabilistic results of the model will help patients and medical practitioners to detect the disease.

3.4 Brief Work Steps

This study seeks to aggregate locally trained models by retrieving them from local servers. The centrally trained model will be sent to all nearby hospitals following the implementation of Federated Learning.

Most hospitals don't want to share their patient data for privacy issues. So, In our research methodology, we do not take each hospital's datasets, for those hospitals who are willing to connect in this system, we give them a model which is already trained with some test datasets, that model is called a global model. This global model will send to each connected hospital and give them access to fit and train their dataset in that model to contribute to the improvement of accuracy, that model sent to each hospital is called a local model. After successfully fitting and retraining local model will be sent to a central server.

The server will take the top 80% models based on their accuracy and test whether the model is most accurate than the previous global model if the global model will be overridden by the most accurate local model. The CNN and Swin Transformation algorithms are the foundation of the model. After using the hybrid ensemble CNN model with the VGG19, Inception-v3, and DenseNet algorithms, Swin Transformation was included, and the primary model was developed.

Above Figure 4.1 is the whole proposed methodology of our thesis. First of all X-Ray Data is being collected to make a global model. The Dataset is being preprocessed for training in our predictive model, there we have to remove some corrupted image data also and augmentation is being applied.

3.5 Classification into 3 Groups

According to the dataset we had worked with two types of diseases COVID-19 and Pneumonia. So we classified those two diseases into groups and another normal healthy lung condition. There can be more than 2 classifications of lung diseases, but the dataset should be Chest X-Ray or CT image. It is necessary for the model

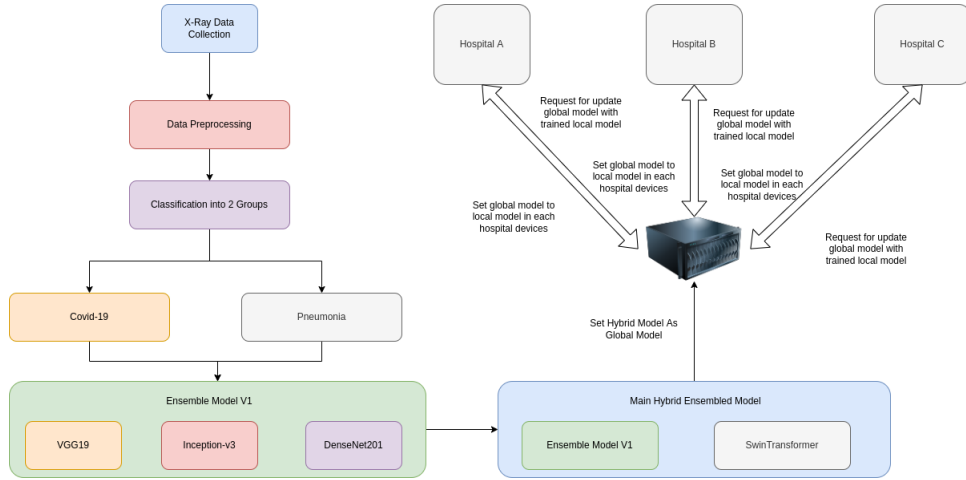


Figure 3.3: Design Flow

efficiency that the model always receive the same format of data. In this Classification, we have split the dataset into two parts also. One is training and another one is testing. The training part contains 80% of the data used to train the model and the Test contains 20% of the data to test the accuracy of the model.

3.6 Fusion Model

The main part of our model we used hybrid model V1 of VGG19, Inception-V3, and DenseNet201. All the data is being trained separately in VGG19, Inception-V3, and DenseNet201. After the Ensemble process, the accuracy will increase and make the model more reliable. Furthermore, adding Swin-Transformer in the main Fusion model based on the transfer learning model and transformer model will have the leverage to combine transfer learning and transformer-based learning that will ensure novel factors for the model.

3.7 Federated Learning Centralized Server

The main Hybrid trained model will be our initial global model of FL integrated central server. Then the server will send the global to each hospital's local devices. After that, the hospitals will have their local model in their local devices and continue to work with that. If any of the hospitals have enough datasets to train or fit again in the local model they can fit into that and make a request to the central server to update that model with the global model. The Central server will check and take the top 80% model with better accuracy and update the global model with a local model which model's predictive performance is better than previous global and other local models. This work will be done in a loop whenever any hospital will request to update the global model.

Chapter 4

Result and Analysis

Disease Prediction is one of the most sophisticated examples of advanced computational ability. Now it is possible to analyze and detect diseases based on CT-Image and X-Ray images. Thanks to the advancement of Artificial Intelligence. There are several AI-based models that can do the job of prediction.

We have used some cutting-edge technology to predict the difference between COVID-19, Pneumonia, and normal X-Ray Images. We have used VGG-19, Inception V3, DenseNet 201, and SWIN Transformer to create our model that can provide reliability to medical practitioners. We have tried to come up with the best approach that can help the medical sector from our computer science field. Moreover, we compared different outcomes that helped us in the way of making the hybrid model more advanced compared to the existing disease detection model with much reliability.

4.1 VGG-19

VGG-19 is the latest pre-trained model from VGG net architecture. It is the updated version of VGG-16. The size of each layer is now 47 which was 41 before in VGG-16. Also, it has variants of filter sizes 64, 128, 256, and 512. In our VGG-19 model, we have added 3 batch normalization layers along with two dense layer sizes of 128 and 64.

Moreover, we also made the middle layers trainable false so that we can avoid overfitting issues. Also, we added a dropout layer with a value of 0.5 to make sure that our model is safe from the overfitting problem. In addition, we have used image sizes of (224, 224,3) for our overall processing steps. We have been careful during the choice of size considering our processing unit capability and required time to complete the training without any issue.

In addition, during training time we have calculated steps per epoch and the number of epochs based on the available number of images that also keep our model safe from overtraining. Moreover, With all these careful steps, we have found 94.4% of validation accuracy during our training time.

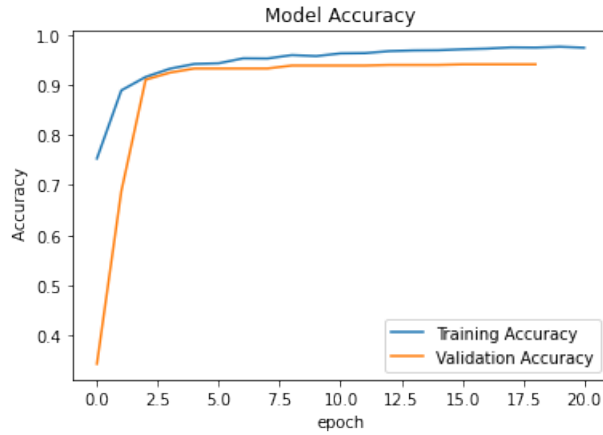


Figure 4.1: VGG-19 Training Accuracy and Validation Categorical Accuracy

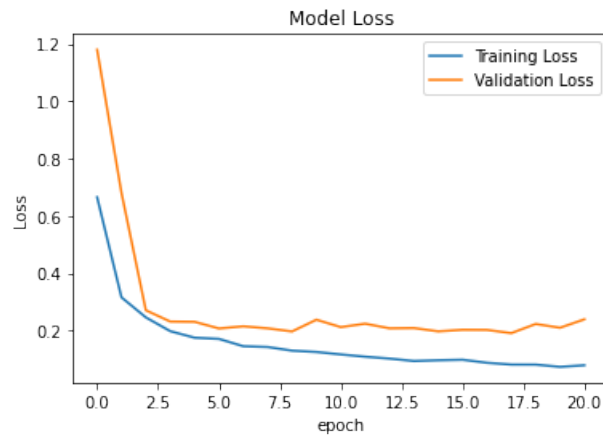


Figure 4.2: VGG-19 Training Loss and Validation Categorical Loss

4.2 DenseNet 201

Now, we start to work with another important convolutional neural network model known DenseNet 201. It is one of the latest neural network architectures available that helped us to make our model even more reliable consisting of 201 layers. We kept the image size (224, 224, 3) during our model training procedure. Along with this, to avoid over-training we have made the internal layer trainable to false and added a dropout layer of value 0.5.

Next, we used a sequential model as our backbone architecture to pass the DenseNet 201 layers and make our custom model. We have added the Global Average Pooling 2D layer and a Dense layer with a size of 1024 to complete the process of designing our custom DenseNet 201 model for improved performance considering the fundamental DenseNet 201 model.

In addition, during training time we have calculated steps per epoch and the number of epochs based on the available number of images and batch size count to keep our model safe from overtraining. With all these careful steps, we have found 94.1% of categorical validation accuracy during our training time.

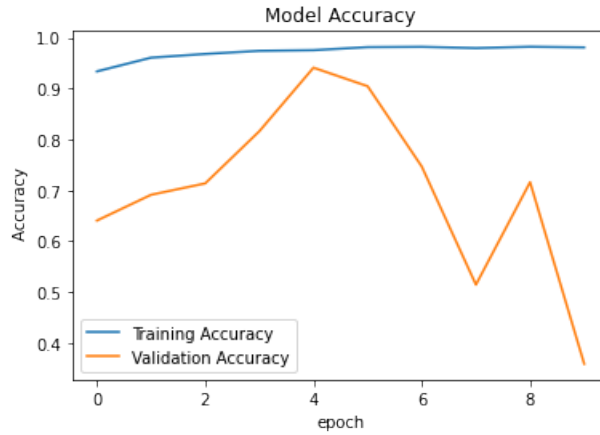


Figure 4.3: DenseNet201 Training Accuracy and Validation Categorical Accuracy

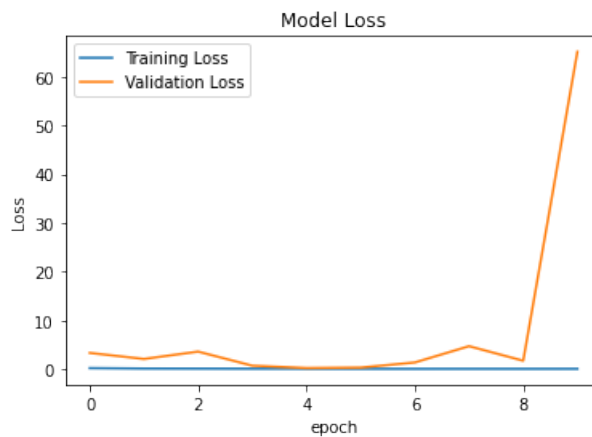


Figure 4.4: DenseNet201 Training Loss and Validation Categorical Loss

4.3 Inception V3

Inception V3 is the third edition of Google’s Inception Convolutional Neural Network. We have used the latest pre-trained model for our disease detection system. Inception V3 is a parallel processing architecture. The default input image size is (299 * 299 * 3). However, we used (224, 224, 3) like our previously used model VGG-19, DenseNet201.

We have used Sequential Model as our backbone architecture during the implementation of the Inception V3 model. We have added the Global Average Pooling 2D layer and a Dense layer size of 1024 during the development of our custom Inception V3 model. We have added a dropout layer with a value of 0.5 to avoid over-fitting problems.

Furthermore, we have calculated steps per epoch that is 200 and the number of epochs is 30 based on the image count of more than six thousand and batch size of 25. We have fine-tuned the training structure keeping in mind our hardware limitation and avoiding over-training. With all these careful steps, we have found 94.5% of validation accuracy during our training time.

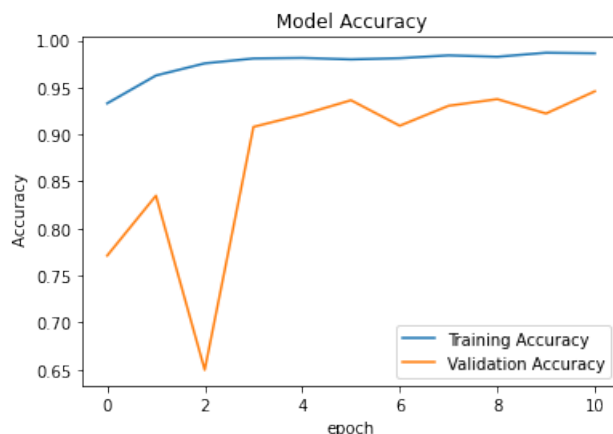


Figure 4.5: Inception V3 Training Accuracy and Validation Categorical Accuracy

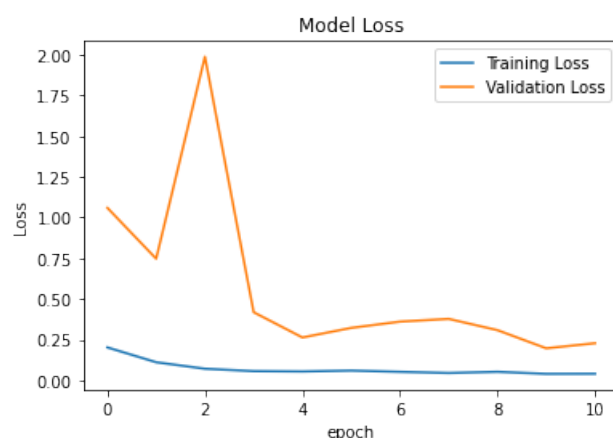


Figure 4.6: Inception V3 Training Loss and Validation Categorical Loss

4.4 SWIN Transformer

SWIN Transformer is the CNN architecture with the origin branch of the transformer-based learning approach. It is one of the most prominent architectures developed by Microsoft. The full form of SWIN is Shifted Window. In this process, we can reach the pixel-level image detail of an image.

This transformer learning technique divided the image into different patches before sending it for training. Like the existing CNN model, the SWIN transformer is a large encoder decoder block that processes the input data.

SWIN Transformer is the general purpose backbone of computer vision. The shifting window technique brings astonishing efficiency by limiting self-attention computation to non-overlapping local windows while also allowing the cross-window connection. We have used a patch size of (2,2) and a number of attention heads of 8. Moreover, we have used a window size of 7 with a shift size of 1. In our training structure, we have maintained the image dimension is (224, 224, 3).

Furthermore, we have calculated steps per epoch that is 200 and the number of epochs is 10 based on the image count of more than six thousand and batch size of

25. We have fine-tuned the training structure keeping in mind our hardware limitation and avoiding over-training. With all these careful steps, we have found 82.5% of validation accuracy during our training time.

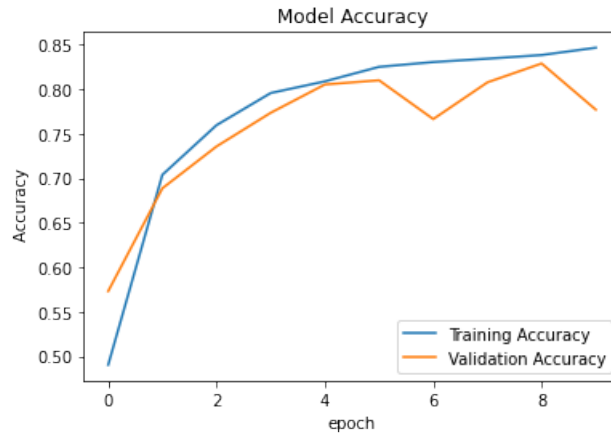


Figure 4.7: SWIN Transformer Training Accuracy and Validation Categorical Accuracy

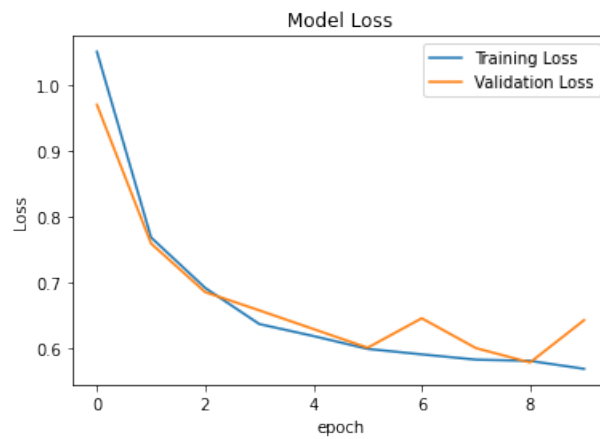


Figure 4.8: SWIN Transformer Training Loss and Validation Categorical Loss

4.5 Fusion Model

We have individually trained and tested our discussed models. Now in order to create the hybrid model. The hybrid model is the combination of transfer learning-based models (VGG-19, Inception V3, DenseNet 201) and the Transformer Model (SWIN Transformer). We have developed this hybrid model.

We have used the ensemble technique to combine the entire model and build one single unique model that will work as the backbone of our disease detection system. We are hopeful that this system that we have worked with COVID19, Pneumonia, and Normal X-ray image will also provide significant outcomes if this model is going to use for any other detection system development.

Furthermore, we have used previous training configurations in order to maintain the proper collaboration of the different models and ensure the best throughput of the hybrid model. We have used (224, 224, 3) image size with a dropout layer value of 0.5. Next, we have maintained a proper training structure consisting of the number of epochs, steps per epoch, and batch size. With all these careful steps, we have found 97.0% of validation accuracy during our training time by combining techniques that sum the weight of all four models. Next, we found 94.0% of validation accuracy by combining techniques that average the weight of all four models.

4.6 Fusion Model Sum Outcome Graph

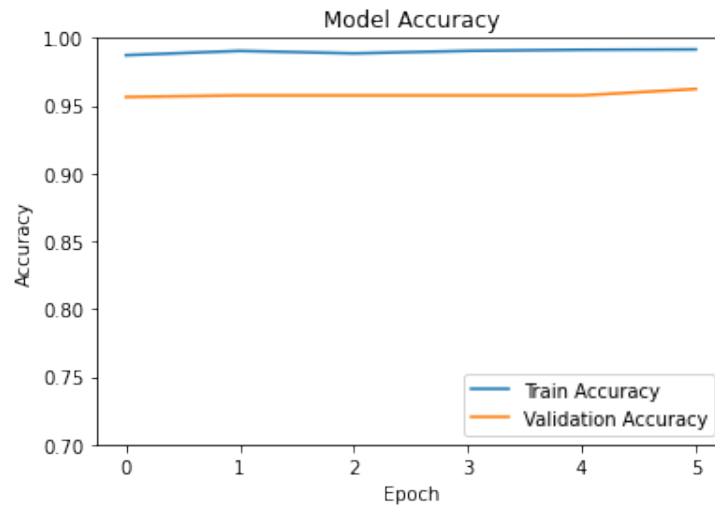


Figure 4.9: Training Accuracy and Validation Accuracy

We have trained our dataset using Fusion Model. We have found training accuracy 99% approximately and categorical validation accuracy of 96% approximately.

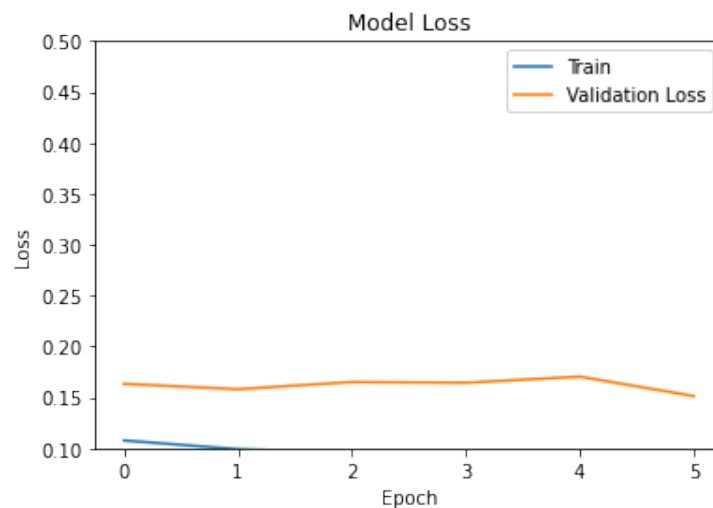


Figure 4.10: Training Loss and Validation Loss

We have trained our dataset using Fusion Model. We have found training loss approximately less than 0.10% and categorical validation loss of less than 0.14% approximately.

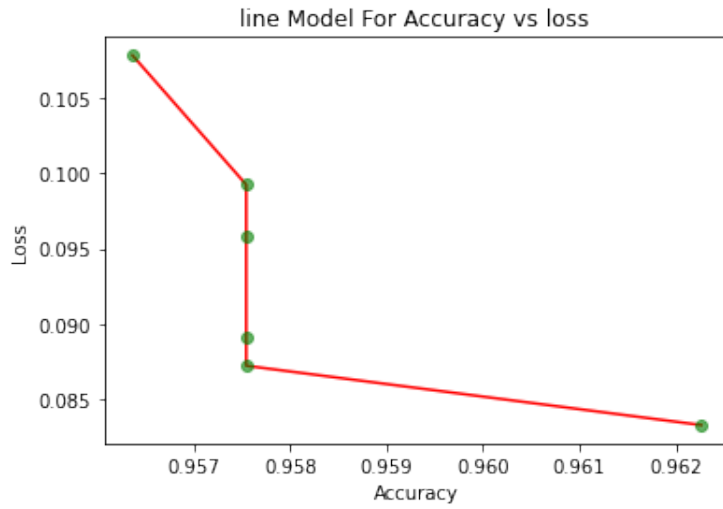


Figure 4.11: Line Model for Training Output

We have trained our dataset using Fusion Model. We have found training accuracy 99% approximately and categorical validation accuracy of 96% approximately.

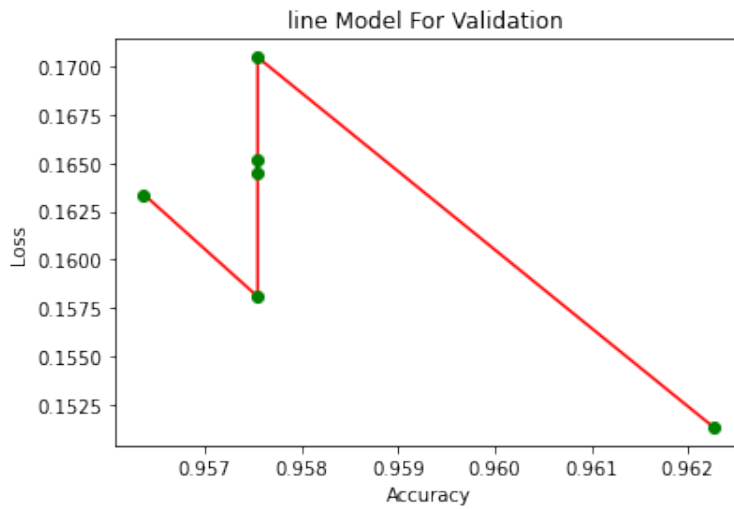


Figure 4.12: Line Model for Validation Output

We have trained our dataset using Fusion Model. We have found training loss approximately less than 0.10% and categorical validation loss of less than 0.14% approximately.

4.7 Fusion Model Average Outcome Graph

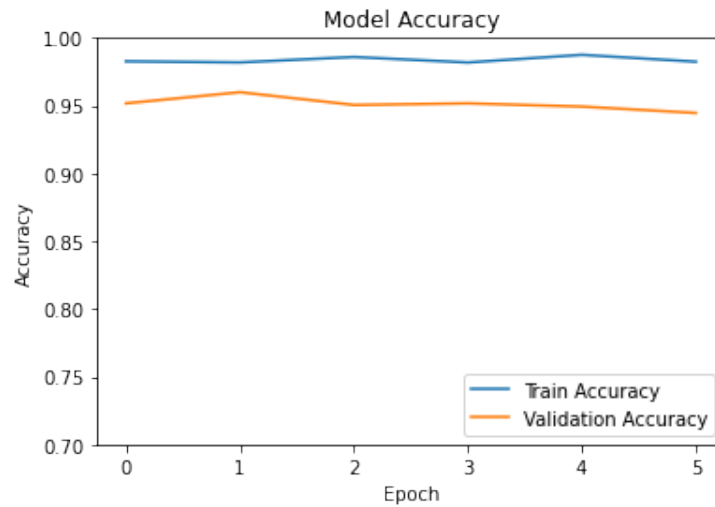


Figure 4.13: Training Accuracy and Validation Accuracy

We have trained our dataset using Fusion Model (Average Approach). We have found training accuracy 98% approximately and categorical validation accuracy of 94% approximately.

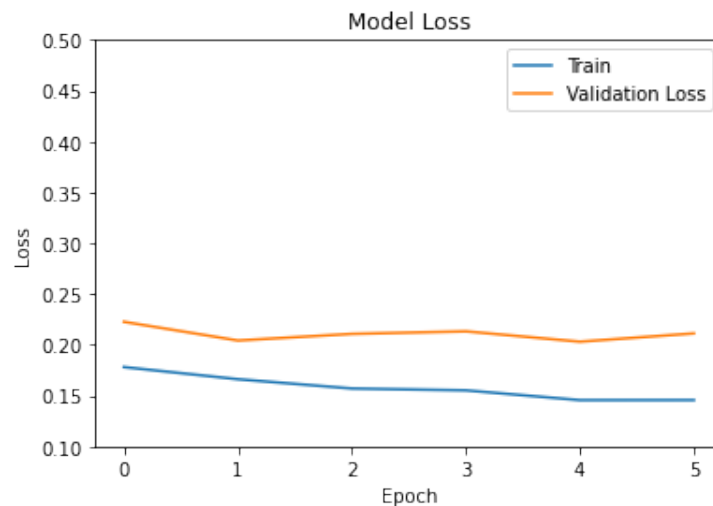


Figure 4.14: Training Loss and Validation Loss

We have trained our dataset using Fusion Model (Average Approach). We have found a training loss of approximately less than 0.18% and a categorical validation loss of less than 0.25% approximately.

4.8 Federated Learning Based Outcome

Here, we calculated the number of times all hospitals' data would fit into a model, and utilizing those models, we created a federated averaging model that replaced the global model. In this case, we gave common rounds a value of 10, and we gave 400 photos and 10 epochs to five hospitals from each institution. However, due to a lack of hardware resources, the model could only run for two epochs. However, based on Figs. 4.15 and 4.16, we determined that if we ever get any good hardware resources, our FL implementation will function.

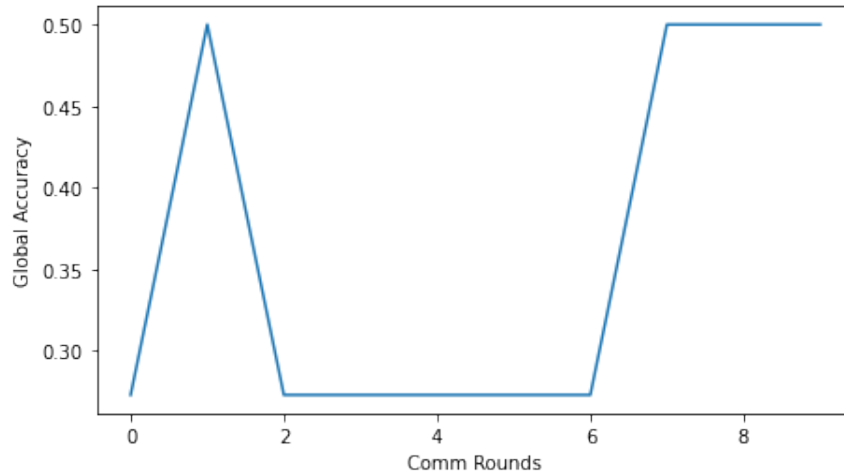


Figure 4.15: Different Hospitals based Training Accuracy and Validation Accuracy

In Figure 4.15 our global model's accuracy was increasing if the common Rounds increase or the hospitals train their local model with their real-time data more time then the model will work more perfectly.

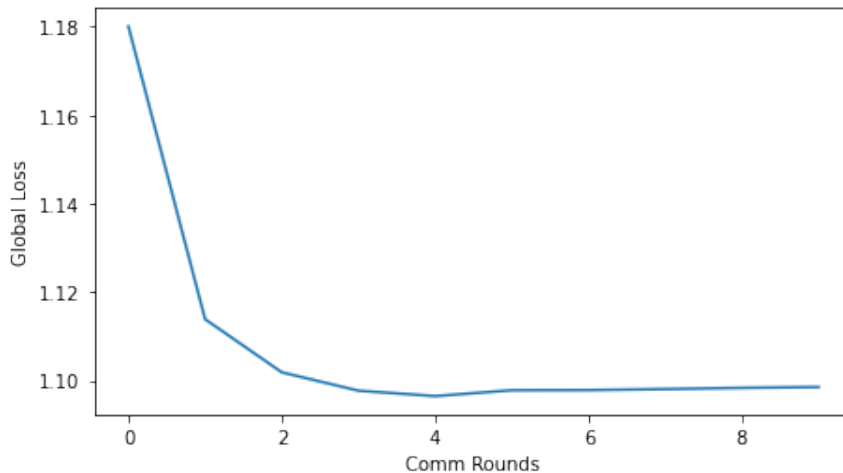


Figure 4.16: Different Hospitals based Training Loss and Validation Loss

In Figure 4.16 our global model's loss also decreasing as the common Rounds increasing.

4.9 Comparative Analysis

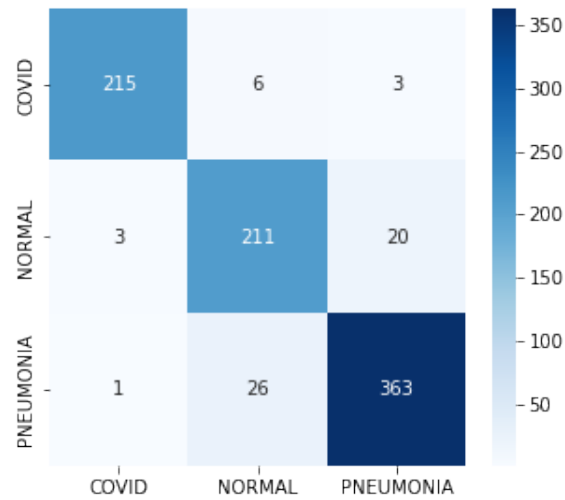


Figure 4.17: VGG-19 Based Confusion Matrix

We have used the VGG-19 model to analyze our test data to identify COVID-19, pneumonia, and normal result. We have found that 251 reports were true detected as COVID-19 while only 3 reports were detected as false normal and 1 report detected false Pneumonia. In addition, this model detected 211 reports as true normal, 6 reports as false COVID-19, and 26 reports as false Pneumonia. Lastly, 363 reports were detected as true Pneumonia, 20 reports were detected as false normal, and only 3 reports were detected as false COVID-19.

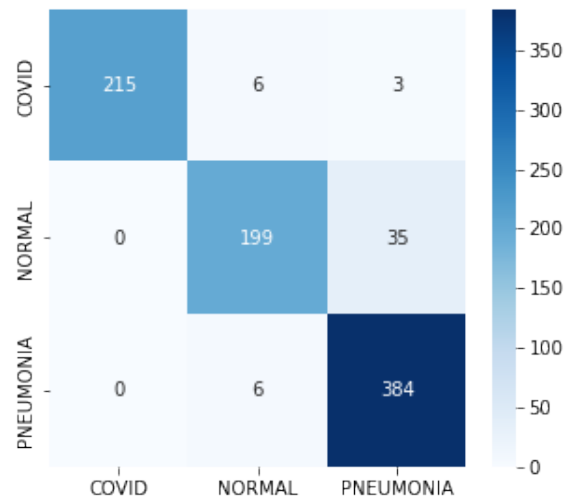


Figure 4.18: DenseNet 201 Model Confusion Matrix

We have used the DenseNet 201 model to analyze our test data to identify COVID-19, pneumonia, and normal result. We have found that 215 reports were true detected as COVID-19 while only 0 reports were detected as false normal and false Pneumonia. In addition, this model detected 199 reports as true normal, 6 reports as false COVID-19, and false Pneumonia. Lastly, 384 reports were detected as true

Pneumonia, 35 reports were detected as false normal, and only 3 reports were detected as false COVID-19.

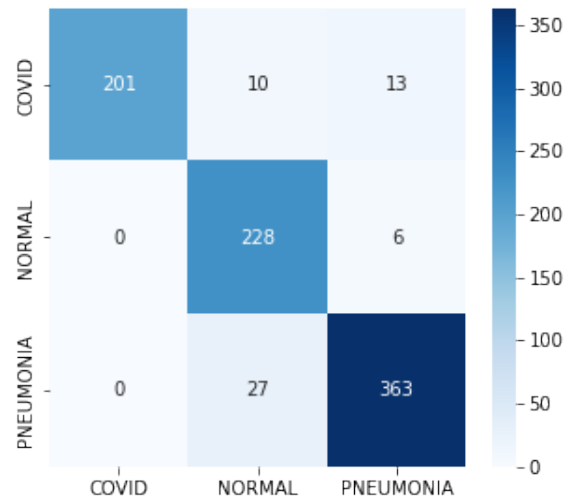


Figure 4.19: Inception V3 Model Confusion Matrix

We have used the Inception V3 model to analyze our test data to identify COVID-19, pneumonia, and normal result. We have found that 201 reports were true detected as COVID-19 while only 0 reports were detected as false normal and false Pneumonia. In addition, this model detected 228 reports as true normal, 10 reports as false COVID-19, and 27 reports as false Pneumonia. Lastly, 363 reports were detected as true Pneumonia, 6 reports were detected as false normal, and only 13 reports were detected as false COVID-19.

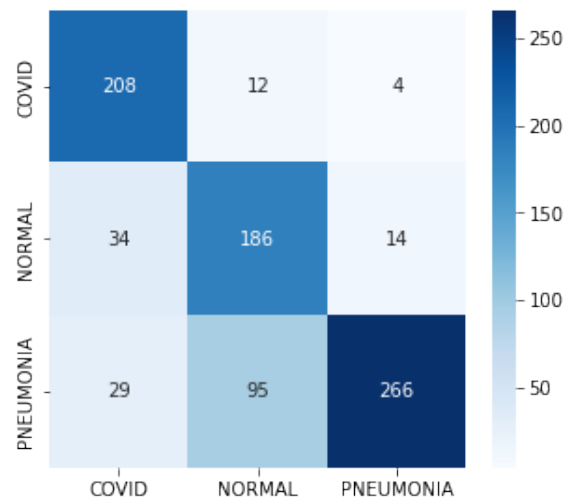


Figure 4.20: SWIN Transformer Model Confusion Matrix

We have used the Transformer learning model to analyze our test data to identify COVID-19, pneumonia, and normal result. We have found that 208 reports were detected as true positive COVID-19 while only 34 reports were detected as false normal and 29 reports were detected as false Pneumonia. In addition, this model

detected 186 reports as true normal, 12 reports as false COVID-19, and 95 reports as false Pneumonia. Lastly, 266 reports were detected as true Pneumonia, 14 reports were detected as false normal, and only 4 reports were detected as false COVID-19.

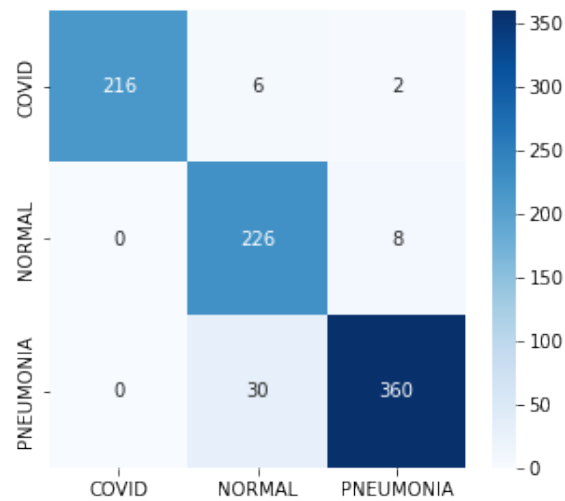


Figure 4.21: Fusion Model and Federated Learning Based Confusion Matrix

We have used the Fusion model to analyze our test data to identify COVID-19, pneumonia, and normal result. We have found that 216 reports were true detected as COVID-19 while only 0 reports were detected as false normal and false Pneumonia. In addition, this model detected 226 reports as true normal, 6 reports as false COVID-19, and 30 reports as false Pneumonia. Lastly, 360 reports were detected as true Pneumonia, 8 reports were detected as false normal, and only 2 reports were detected as false COVID-19.

4.10 AUC-ROC Outcome

Area Under Curve(AUC) and Receiver Operator Characteristic(ROC) are both abbreviations for AUC. The probability curve known as the ROC is used to compare the Truth-Per-Rate (TPR) and False-Per-Rate (FPR) at various threshold levels while separating the signal from noise. The measurement capabilities of a classifier's ability to discriminate between classes are known as AUC. The model performs better at differentiating between the positive and negative classes the higher the AUC. True Positive Rate (TPR) is a recall calculated by,

$$TPR = \frac{TP}{TP + FN}$$

False Positive Rate (FPR) is defined as,

$$FPR = \frac{FP}{FP + TN}$$

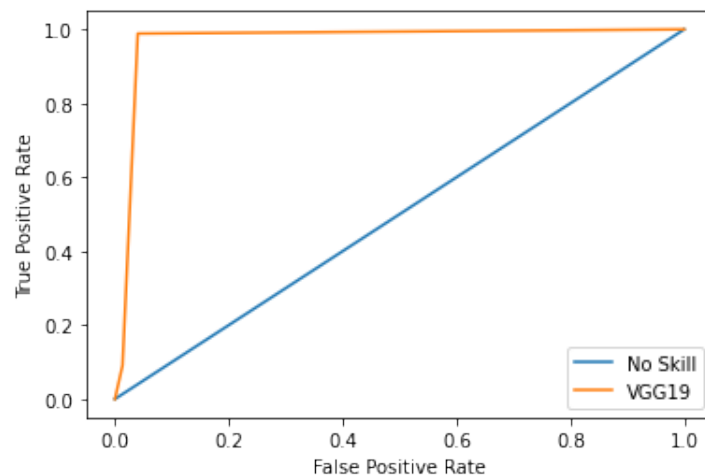


Figure 4.22: VGG-19

Here the VGG19's AUC-ROC curve shows a good Area Under Curve than the NoSkill straight line diagonal. As this VGG19's ROC curve shows lower classification threshold so, this model classifies items as positive mostly

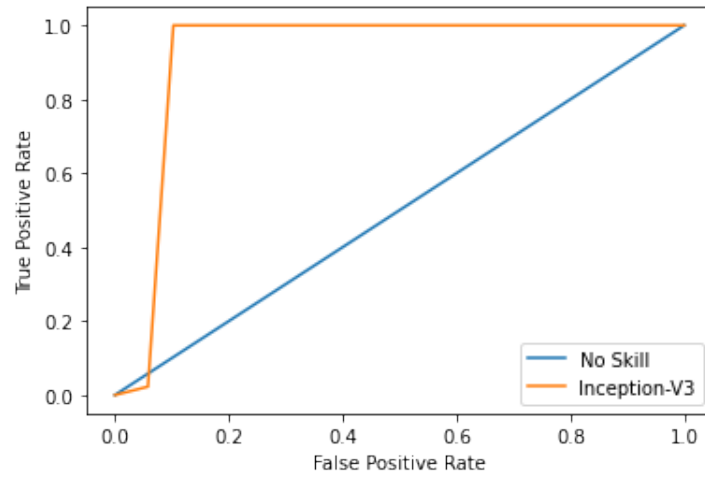


Figure 4.23: Inception V3

Here in Inception-V3 ROC-AUC curve shows less area under the curve than the VGG-19 model. So the Inception-v3 is less accurate than the VGG-19 model. Also, the classification threshold is not as good as the VGG-19 model.

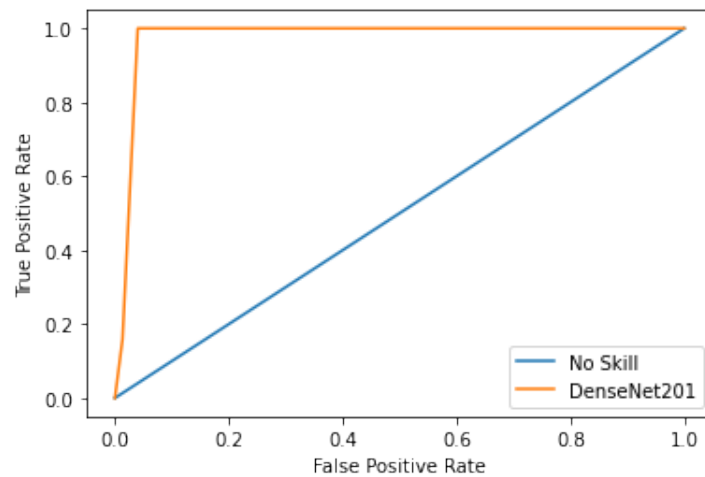


Figure 4.24: DenseNet 201

Here in DenseNet 201, the value of AUC is also well, but initially, the area under the curve got shrinks but the ultimate area is good.

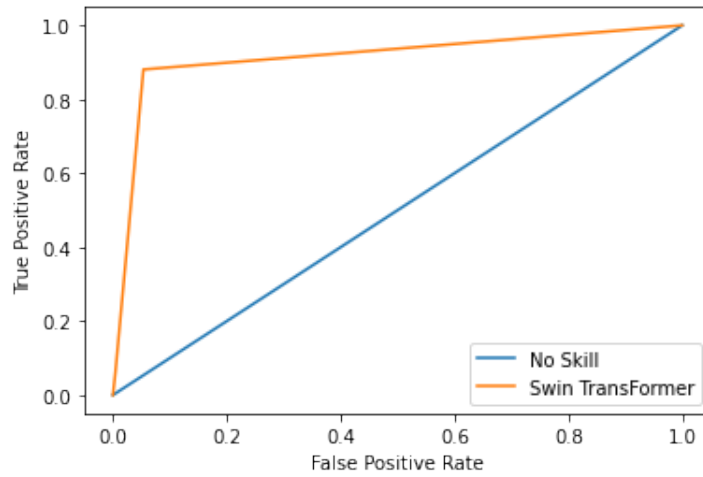


Figure 4.25: SWIN Transformer Model

The SwinTransformer model produces a poor ROC rate. It cannot therefore cover a smaller region beneath the curvature. Swin Transformer classifies this information as false in comparison to all the other individual model classifications. For this reason, we combined various CNN models with a swin transformer to create a fusion model that can categorize things as positive.

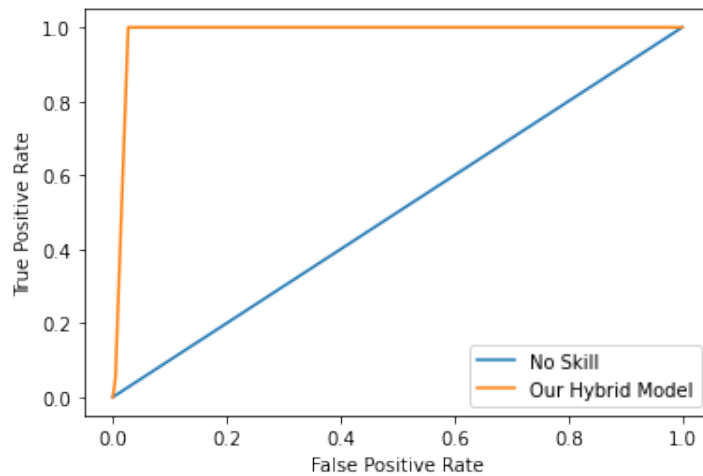


Figure 4.26: Fusion Model and Federated Learning Based

This is the AUC-ROC curve for the Fusion model, which was created by combining all the other models. It has a higher AUC value than the other models did on their own. Fusion models classify more items as positive since their classification thresholds are lower and their area under curves are now significantly better than those of the other separate models.

4.11 Overview Based on Different CNN Model

Model Comparison			
Classifier	Training Time (s) (approx.)	Testing Time (s) (approx.)	Accuracy (%)
VGG-19	14440	4	94.4
Inception V3	15200	2	94.5
DenseNet 201	18120	2	94.1
SWIN Transformer	25650	4	82.5
Fusion Model (Sum)	24122	3	96.24
Fusion Model (Average)	21600	2	94

Table 4.1: Model Comparison Table

Chapter 5

Conclusion

In this research, we have provided brief explanations on **A Hybrid FL-Enabled Ensemble Approach For Lung Disease Diagnosis Leveraging a Combination of SWIN Transformer and CNN**. At first, we discussed our system that is going to be beneficial for mass people. We have used a combined model of transfer learning and transformer learning known as shifted window transformer to make our model reliable. We explained our idea on how to use federated learning to analyze local databases of hospitals and create local models of every different individual dataset and implementation of federated learning to secure the entire procedure. This federated learning implementation process has several advantages along with difficulties but considering the advantages, it has become an important part of the entire project. In the future, we want to work on the concept drift of federated learning to address the limitation of federated learning. Lastly, in order to predict diseases, we are looking forward to using the latest AI models that can help the medical sector people to learn about diseases even better that can help to treat diseases based on vast data and scenarios.

Bibliography

- [1] A. Khosla, Y. Cao, C. C.-Y. Lin, H.-K. Chiu, J. Hu, and H. Lee, “An integrated machine learning approach to stroke prediction,” in *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining*, 2010, pp. 183–192.
- [2] O. Kocabas, T. Soyata, J.-P. Couderc, M. Aktas, J. Xia, and M. Huang, “Assessment of cloud-based health monitoring using homomorphic encryption,” in *2013 IEEE 31st International Conference on Computer Design (ICCD)*, IEEE, 2013, pp. 443–446.
- [3] O. Kocabas, T. Soyata, and M. K. Aktas, “Emerging security mechanisms for medical cyber physical systems,” *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 13, no. 3, pp. 401–416, 2016.
- [4] K. Bonawitz, V. Ivanov, B. Kreuter, *et al.*, “Practical secure aggregation for privacy-preserving machine learning,” in *proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, 2017, pp. 1175–1191.
- [5] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely connected convolutional networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 2261–2269. DOI: 10.1109/CVPR.2017.243.
- [6] V. Smith, C.-K. Chiang, M. Sanjabi, and A. S. Talwalkar, “Federated multi-task learning,” *Advances in neural information processing systems*, vol. 30, 2017.
- [7] W.-H. Weng, K. B. Waghlikar, A. T. McCray, P. Szolovits, and H. C. Chueh, “Medical subdomain classification of clinical notes using a machine learning-based natural language processing approach,” *BMC medical informatics and decision making*, vol. 17, no. 1, pp. 1–13, 2017.
- [8] N. Mostafa, *Critical care medicine: Bangladesh perspective*, Jan. 2018. [Online]. Available: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6549198/>.
- [9] O. R. Shishvan, D.-S. Zois, and T. Soyata, “Machine intelligence in health-care and medical cyber physical systems: A survey,” *IEEE Access*, vol. 6, pp. 46 419–46 494, 2018.
- [10] L. Corinzia, A. Beuret, and J. M. Buhmann, “Variational federated multi-task learning,” *arXiv preprint arXiv:1906.06268*, 2019.

- [11] “Medical cyber physical systems and its issues,” *Procedia Computer Science*, vol. 165, pp. 647–655, 2019, 2nd International Conference on Recent Trends in Advanced Computing ICRTAC -DISRUP - TIV INNOVATION , 2019 November 11-12, 2019, ISSN: 1877-0509. DOI: <https://doi.org/10.1016/j.procs.2020.01.059>.
- [12] W. Schneble and G. Thamilarasu, “Attack detection using federated learning in medical cyber-physical systems,” in *28th International conference on computer communications and networks (icccn)*, 2019, pp. 1–8.
- [13] G. Battineni, G. G. Sagaro, N. Chinatalapudi, and F. Amenta, “Applications of machine learning predictive models in the chronic disease diagnosis,” *Journal of personalized medicine*, vol. 10, no. 2, p. 21, 2020.
- [14] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, “Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images,” *arXiv preprint arXiv:2003.11055*, 2020.
- [15] M. M. Kabir, F. B. Safir, S. Shahen, J. Maua, I. A. B. Awlad, and M. Mridha, “Human abnormality classification using combined cnn-rnn approach,” in *2020 IEEE 17th International Conference on Smart Communities: Improving Quality of Life Using ICT, IoT and AI (HONET)*, IEEE, 2020, pp. 204–208.
- [16] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020. DOI: 10.1109/MSP.2020.2975749.
- [17] *5 reasons to build a health monitoring system for a hospital*, Jul. 2021. [Online]. Available: https://cprimestudios.com/blog/5-reasons-build-health-monitoring-system-hospital?fbclid=IwAR3JWIG6OmjVy_DfpsTU22nbJnPiNrc8qlchDggHme6
- [18] S. H. Kassania, P. H. Kassanib, M. J. Wesolowskic, K. A. Schneidera, and R. Detersa, “Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: A machine learning based approach,” *Biocybernetics and Biomedical Engineering*, vol. 41, no. 3, pp. 867–879, 2021.
- [19] Z. Liu, Y. Lin, Y. Cao, *et al.*, “Swin transformer: Hierarchical vision transformer using shifted windows,” in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2021, pp. 10 012–10 022.
- [20] Y. Gu, Z. Piao, and S. J. Yoo, “Sthardnet: Swin transformer with hardnet for mri segmentation,” *Applied Sciences*, vol. 12, no. 1, p. 468, 2022.
- [21] Y. Jiang, Y. Zhang, X. Lin, J. Dong, T. Cheng, and J. Liang, “Swinbts: A method for 3d multimodal brain tumor segmentation using swin transformer,” *Brain Sciences*, vol. 12, no. 6, p. 797, 2022.
- [22] K. Adams, *Healthcare data breaches by the numbers: 9 stats*. [Online]. Available: <https://www.beckershospitalreview.com/cybersecurity/healthcare-data-breaches-by-the-numbers-9-stats.html>.
- [23] *Coronavirus cases*: [Online]. Available: <https://www.worldometers.info/coronavirus/>.
- [24] *Pneumonia*. [Online]. Available: <https://www.who.int/news-room/fact-sheets/detail/pneumonia>.