Interpretable COVID-19 Classification leveraging Ensemble Neural Network and Explainable AI

by

Shakib Mahmud Dipto
17101153
Irfana Afifa
18101316
Sumya Kabir
17101298

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
June 2020

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:

Shakib Mahmud Dipto
17101153

Irfana Afifa
18101316

Sumya Kabir
17101298

Approval

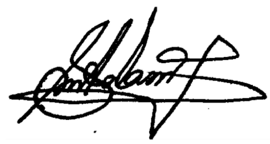
The thesis/project titled "Interpretable COVID-19 Classificationleveraging Ensemble Neural Network and Explainable AI" submitted by

1. Shakib Mahmud Dipto (17101153)
2. Irfana Afifa (18101316)
3. Sumya Kabir (17101298)

Of Spring, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on August 23, 2015.

Examining Committee:

Supervisor:
(Member)



Md. Ashraful Alam, PhD
Assistant Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:
(Member)

Md. Tanzim Reza
Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Md. Golam Rabiul, PhD
Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD
Chairperson and Associate Professor
Department of Computer Science and Engineering
Brac University

Abstract

COVID-19 which is also none as Corona Virus Disease is first discovered in a city of China named Wuhan at December 2019 and it has been announced as a global pandemic at the middle of 2020. SARS-CoV-2 virus COVID-19 and that can also act as a trigger to cause respiratory tract infection ranging from mild to deadly. According to experts, this virus may also infect the upper respiratory system, which includes the sinuses, nose, and throat, as well as the lower respiratory system, which includes the windpipe and lungs. The disease can infect other people via respiratory droplets and coming near to the COVID-19 infected people as well as touching those objects of surfaces which are the virus contaminated. Nowadays, millions and millions of people across the globe are suffering from this disease causing a huge death rate. Even after taking serious precaution measures, the number of patients dealing with this disease and the death toll is still rising at a drastic rate. In this paper, we approach a fast and effective measure to detect COVID-19 using CT scan images. First, we collected data and classified using VGG16, VGG19, EfficientNetB0, ResNet50 and ResNet152. Form our result; we got an accuracy rate of 85.33% from VGG16, 87.86% from VGG19 and 82.35% from ResNet101. Then we formed an ensemble model with these best three classifiers and achieved a best overall accuracy rate of 90.89% from COV19EXAI V1 and 91.82% from COV19EXAI V2. Finally, we integrated XAI in our model to achieve a better understand of our classification.

Keywords: COVID-19, Corona Virus, CV19AA, Deep Neural Network, VGG, Inception V3, ResNet, XAI, Ensamble Modeling, VGG16, VGG19, ResNet50, ResNet101.

Acknowledgement

First and foremost, we give thanks to Almighty Allah for allowing us to complete our Thesis work on schedule and without serious setbacks.
That being said, we would like to offer our heartfelt gratitude to Dr. Md. Ashraful Alam, our esteemed instructor and supervisor, for his unwavering support and constant monitoring, which enabled us to complete our project. We also owe a debt of gratitude to our co-supervisor, Md. Tanzim Reza, who has willingly assisted us in this thesis to the best of his ability.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

More than 153 million individuals have been infected with COVID-19 since early December 2019, with more than 3.5 million people dying as a result of the virus. The rate at which the virus is spreading and killing people, on the other hand, is increasing day by day, breaking all previous records. For the consequence, all radiologists throughout the world have had a difficult problem in testing so many patients and delivering test results on time. After a day or two, many people receive their test results. As a result, their health is deteriorating at this time, as the virus is wreaking havoc on their lungs. CT scan pictures of the lungs, on the other hand, reveal the abnormalities caused by COVID-19 infection. Therefore, CT scan data are conveyed sooner than COVID-19 PCR test results. So, in order to achieve the quickest binary classification of COVID-19, we're attempting to develop a system that will identify COVID-19 from infected patients' lungs CT scan pictures by comparing them to CT scan pictures of normal lungs.

1.2 Aims and Objectives

The aim is to prepare a deep learning model which will determine the result of COVID-19 test. Our main objective is to visualize the result of the classification of COVID-19 and understand the underlying features for classification. The Understanding of feature relevance may lead to better performance in the future research.

1.3 History of COVID-19

In December 2019, multiple numbers of patients were admitted to the hospitals in Wuhan city of China having server pneumonia. Later, the doctor discovered that the reason behind this pneumonia is a disease named COVID-19 caused by a virus SARS-CoV-2 (Sever Acute Respiratory Syndrome CoronaVirus 2). This virus is one kind of Baltimore class IV single-strained RNA virus which is considered highly contagious for human race. The virus has an incubation period of 5-12 days and remains infectious in human body for 7-14 days. The virus spreads very easily and sustainably from one person to another through the breathing, coughing, sneezing, talking or even singing of the infected person. The virus produces contaminated droplets and those travels through the air, surface of the skin of the people near around and if someone inhales the virus, it can travel into lungs that can cause new deadly infections. These contaminated droplets are generally airborne can be present in air for a long time and can cause airborne transmission in populated as well as congested indoor places like restaurants, public transport, hospitals, educational institutes or any type of social gatherings infecting hundreds of people at once. This virus can also infect a man who touches his eyes, nose or mouth with the same hand immediately after touching the contaminated surfaces. The CorornaVirus transmission can even occur through other things such as food, wastewater, drinking water, feces, breast milk, from mother to the child during pregnancy, animal disease vector and so on. Loss smell and taste, fever, cough, breathing difficulties are the common symptoms of COVID-19 infection. However, severe symptoms can cause death to the infected person. From the beginning of January 2020, the CoronaVirus begun to spread at a drastic rate infecting thousands of people and spread in many countries other than China. After watching the rapid growth of the transmitted people the World Health Organization (WHO) has announced it as a Public Health Emergency of International Concern in the January 2020. However, soon the death rate due to virus increased at a dramatic rate shocking the entire world and also has spread throughout each and every country worldwide. When millions of COVID-19 cases are recorded throughout the world causing death to thousands of people every single day, the condition has been declared as a global pandemic in March 2020 by WHO. Till early October 2020, the number of COVID-19 patients reached over 34.4 million people worldwide from which more than 1.02 million patients are dead due to the virus, although the real number can be much higher.

1.4 Research Methodology

At first, we have observed several Machine Learning models such as InceptionV3, VGG16, VGG19, ResNet50, ResNet101 and EfficientNet-B0. We have pre-processed the CT-scan image data into a well-defined form of $224 \times 224 \times 3$ in the initial stage. Then we trained and evaluated those machine learning models. After evaluation, we took the best three classifiers and constructed various forms of ensemble with them. Finally, after comparing the performance of the ensembles, we applied Explainable AI to visualize the classification.

1.5 Research Orientation

In chapter 2, we have shown the previous works done by other researcher related with our research area. Then, in chapter 3, we described each and every algorithm, convolution layer and activation functions that we used in our research. Again, in chapter 4, the implementation of our thesis work is described. We also illustrated how our datasets are distributed and after that we go on to describe the pre-processing techniques applied to the used dataset. In chapter 5, we showed the results that we got after implementing the algorithms and later on described the result analysis. Finally, in chapter 6, we provided the conclusion as well as the future work of your research.

Chapter 2

Related Work

Most of the COVID-19 patients suffer from lung complications that include pneumonia or acute respiratory distress syndrome (ARDS). In addition, a large number of COVID-19 infected patients have lost their life due to respiratory failure caused by pneumonia. So, one of the easiest and fastest ways to identify the COVID-19 patients form other people is the CT scan images as well as x-ray images of their lungs. However, neurological complications such as Encephalopathy can be developed among the patients of COVID-19 which is tends to affect elderly people and those who have preexisting medical conditions such as chronic bronchitis, emphysema, heart failure, or diabetes [1]. Meanwhile, the most common indications and patterns of lung variation from the normal one is on CXR in COVID-19 and so as to prepare the clinical network. CXR is a less delicate methodology in the discovery of COVID-19 lung disease contrasted with CT, with an announced baseline CXR affectability of 69% [2]. In another study [3], 260 X-ray images of lungs have been used from the repository of GitHub and Kaggle to ensure the development, validation and testing of that particular proposed model in which half of the images are of COVID-19 patients and the rest are of normal lungs. Furthermore, [4] many researchers have created a new model to detect COVID-19 automatically using raw images of chest x-ray in which both advanced artificial intelligence techniques and radiological images is applied in a combined manner for the accurate detection of the virus as well as for resolving the problem due to the lack of well specialized physicians in remote areas. In paper [5], they developed a COVID-19 diagnostic model utilizing Bayes-SqueezeNet for COVID-19 identification utilizing x-ray pictures.
But a research [6] was set up MODE and CNN are deep learning models that can be used to categorize persons based on whether or not they are affected by COVID-19. CNN, ANN, and ANFIS are that models that are executed for characterizing COVID-19 positive patients form the images of their chest x-rays. In addition, paper [7] focuses on the most common indications and patterns of lung variation from the norm on CXR in COVID-19 so as to prepare the clinical network. CXR is a less delicate methodology in the discovery of COVID-19 lung disease contrasted with CT, with an announced baseline CXR affectability of 69%. To determine the significance of the extracted features for the classification of COVID-19 patients, the researchers used a state-of-the-art Convolutional Neural Network (CNN) called Mobile Net [8] as CNN became successful in mining those significant image features that were discovered in some x-ray images as well as highlights the rarity of the extracted features.

In addition, the deep learning based strategy is also successfully for the detection of coronavirus contaminated patient utilizing X-ray images [9]. In paper [10], to detect the presence of CoronaVirus, a three-dimensional deep learning method is presented by using the images of chest CT scan images because CT may share certain comparable imaging highlights between COVID-19 and other sorts of pneumonia, hence, making it troublesome to differentiate. Again, in the study [11], both CT scan and x-ray images are used for analyzing lungs for detecting COVID-19 infection. The research provides an alternative rapid screening strategy that searches for visual signs in COVID-19 patients' chest radiography imaging.
In contrast, computed tomography images are utilized to pre-train the deep learning system as well as to train and assess the deep learning system's performance [12]. This approach is capable of distinguishing high-risk individuals from low-risk individuals and also provides a handy tool for quickly screening COVID-19 infection. However, in the findings [13] a model based on 600 patients from 18 laboratories was utilized and verified using 10 fold cross validation and train-test split techniques. The model also computes precision, F1-score, re-call, AUC, and accuracy scores to measure prediction performance, and 6 distinct deep learning models are employed to create this model. The model can also be used for validating the initial laboratory findings and clinical prediction studies.
On the other hand, pre-trained deep learning models ResNet and Inception is used for the automatic classification of white blood cell [14]. Initially, they proposed three consecutive pre-processing calculations named as color distortion, bounding box dostortion and image flipping mirroring. After that, Inception and ResNet architectures are used in recognizing white blood cell with hierarchy topological feature extraction. In the research paper [15], researchers used VGG architecture for the successful identification and prediction of lungs cancer. Each of the 16 layers of VGG performs such functions that consists convolutional functions where all the features of an image are extracted and those extracted features are passed on to the cancer nodule to identify the probability and predict the possibility of cancerous and non-cancerous in lung cancer. Their research shows that VGG is the most efficient architecture in finding the most accurate result compared to the others.
COVID-19 is identified utilizing a huge dataset of chest X-rays and an ensemble of deep convolutional neural networks based on EfficientNet [16]. This research offers a CNN-based architecture with a front-end pre-trained EfficientNet for extraction and model snapshots for detecting COVID-19 infected individuals from a chest x-ray data set. Furthermore, the research provides an ensemble of their suggested architecture to forecast outcomes by taking into account the judgments of many radiologists for the final prediction in order to create an effective assessment system.. Furthermore, the paper [17] is based on the use of EfficientNet with differential privacy practice in the diagnosis of Coron-aVirus utilizing radiological images.They described their system using the Grad-Cam method, in which the areas of the pictures EfficientNet-B0 model leverages the characteristics in CXR pictures for the diagnosis of COVID-19. Also, they applied DP approximation in EffiecntNet, thus many people are motivated to apply AI based models in healthcare services while providing data security..Again, according to the study paper [18] ,the researchers devised an algorithm that utilizes the architecture EfficientNet extended and dubbed it K-EfficientNet for the goal of accurately detecting COVID-19 using deep learning image classifier and K-COVID chest x-ray images5 dataset. With the help of

the proposed algorithm K-EfficientNet, they basically resized the images and increased the size at the time of the training process to get their defined target. In the article [19], the researchers proposed a model which is based on the EfficientNet CNN named as "Efficient-CovidNet" which is also an approach based on deep learning for detecting COVID-19 from the chest x-ray images. For the detection of COVID-19, EfficientNetB1 has been used because it has acquired more sensitivity and PPV comparing to other models. Transfer lessons and convolutional neural networks are utilized to automatically detect COVID-19 in radiographic imagery [20]. At first, they collected 1437 x-ray images including 224 confirmed COVID-19 infection images, 700 confirmed bacterial pneumonia images, and 504 normal conditions images from the available X-ray images on public medical repositories. This data set is used for the purpose of training and testing the CNNs. Again, a transfer learning–BiLSTM network architecture based on Convolutional Neural Network has been approach for detecting COVID-19 infection [21]. Here, the researchers have proposed mainly two deep learning architectures for the detecting COVID-19 infected patients automatically using chest CT scan images. The first one is lung segmentation also known as processing of the CT scan images and this thing is automatically performed with ANN (Artificial Neural Networks). Again, the second architecture which is proposed is basically one king of hybrid layer which includes a Bidirectional Long Short-Term Memories layer and the job of this layer is to consider temporal properties. AlexNet architecture is present in both architectures which mean that the suggested method is one kind of transfer learning method. The study [22], proposed COVID-19 identification from CT scan images of the chest using classic data augmentation algorithms and conditional Generative Adversarial Nets (CGAN) based on a deep transfer learning model. The main purpose is to collect all prospective COVID-19 photographs that exist as of the publication of this research and utilize traditional data augmentations in conjunction with Generative Adversarial Nets to obtain more picture data to help in COVID-19 identification.
The study [23] on the other hand, is based on the use of an explainable deep learning architecture for the aim of differential diagnosis of COVID-19 infection using chest x-ray images. They used image enhancement, image segmentation as well as a modified version of stacked ensemble model which is consist of four CNN based learner, namely, VGG-16, ResNet-50, DenseNet-121, and DenseNet-169. In this article, explain-ability is incorporated using Grad-CAM visualization for acquiring more trust for AI in medical system. Again, in the article [24] a prospective framework is used for the purpose of the COVID-19 risk prediction depending on XAI (Explainable Artificial Intelligence). The researchers used one kind of two-step procedure for the purpose of the non-clinical prediction of the infection risk of COVID-19. At first, the primary risk of COVID-19 is detected carefully by considering selected parameters that are associated with COVID-19 infection symptoms. After, the result of the first step is generated, an optional prediction system is also provided by analyzing the chest x-ray images in the second step of the explainable AI based prospective framework that they have proposed. In addition, in the study [25], the researchers proposed and developed XAI techniques for COVID-19 classification models, and also compared them. The findings show that quantitative and qualitative visualizations can help clinicians understand and make better decisions by providing more granular information from the learnt XAI models' outcomes. Furthermore, in the paper [26], the researchers presented an explainable DNN-based methodology for au-

tomated identification of COVID-19 symptoms from chest x-ray images. They have proposed 'DeepCOVIDExplainer,' a novel diagnosis methodology based on neural ensemble technique, as a first step toward an AI-based clinical aid tool for COVID-19 diagnosis. The 'DeepCOVIDExplainer' pipeline begins with histogram equalization, filtering, and not so sharp masking of CXR pictures, then trains DenseNet, ResNet, and VGGNet architectures in a transfer learning (TL) environment, resulting in model samples. They presented the 'DeepCOVIDExplainer' in this study to use explainable COVID-19 prediction based on CXR pictures. A deep CNN architecture based approach is provided in [27] in order to identify COVID-19 infected individuals utilizing the chest x-ray pictures. Many, state-of-the-art CNN models, such as, DenseNet201, Resnet50V2 and InceptionV3, is combined together and use in the proposed model. The models are individually trained and after that weighted average ensembling technique is applied to combine these models to make prediction of a class value. In addition, in publication [28], the researchers employed an ensemble deep learning network to identify COVID-19 from CT scan pictures.The model also employs transfer learning to generate model parameters, and it has pretrained three deep CNN models: AlexNet, GoogleNet, and ResNet.In addition, EDL-COVID, an ensemble classifier, is generated using relative majority voting. Finally, the ensemble classifier was evaluated with three other component classifiers in terms of accuracy, sensitivity, specificity, F value, and Matthew's correlation coefficient.

Chapter 3

Research Methodology

3.1 Our Methodology

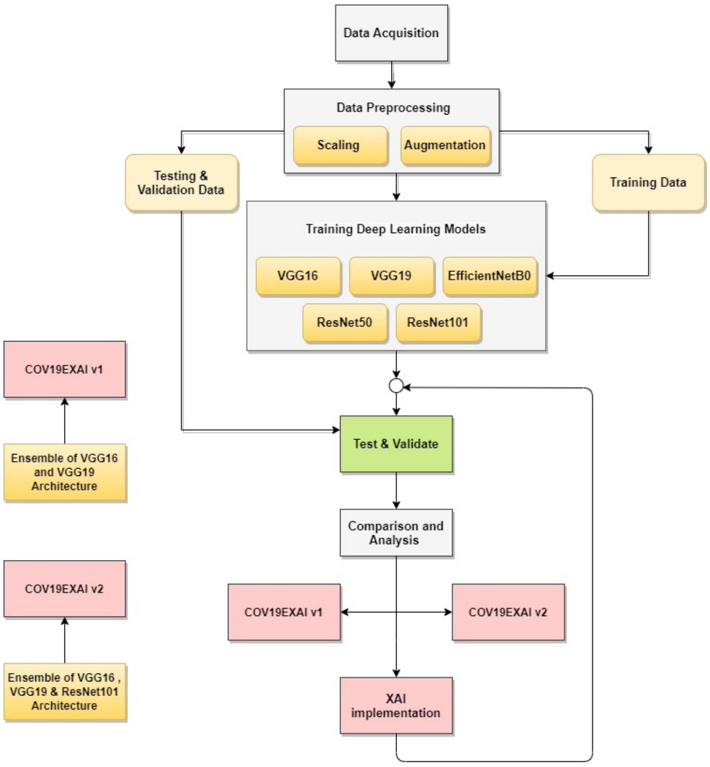


Figure 3.1: WorkFlow Diagram

An overview of each and every step that we have taken in order to train and validate our models is shown through the work flow diagram in figure 3.1. Firstly, we gathered our data and performed data preprocessing that includes data scaling and augmentation. Then, we trained our dataset via training deep learning model (VGG16, VGG19, EfficientNetB0, ResNet50 and ResNet101). After this, the dataset was tested and validated. In the next step, we used the models that we have created, COV19EXAI V1 (Ensemble of VGG16 and VGG19 architecture) and COV19EXAI V2 (Ensemble of VGG16, VGG19 and ResNet101 architecture), to train the data set and then tested and validated the data to find accuracy.

3.2 Used Architectures

3.2.1 VGG

VGG16 is a convolutionary variant of the neural network suggested in [15]. VGG (Visual Geometry Group) is one type of CNN architecture which is mostly used for many identification and detection purposes. This architecture consists primarily of many convolution layers with the activation function of ReLU, max- pooling layers, entirely linked layers with the activation function of ReLU and the activation function of softmax in the final output layer. The network input is a $224 \times 224 \times 3$ RGB image. This picture is transferred with kernels of dimension 3×3 across multiple convolutionary layers. The filters are of dimension 2×2 and are added with a phase of 2 in the max-pooling layers.

VGG16

VGG16 is considered to be one of the most excellent vision model architecture till date. This model is actually consisted of convolutions layers, max pooling layers, and fully connected layers. There are a total of 16 layers with 5 blocks and each block with a max pooling layer [29].

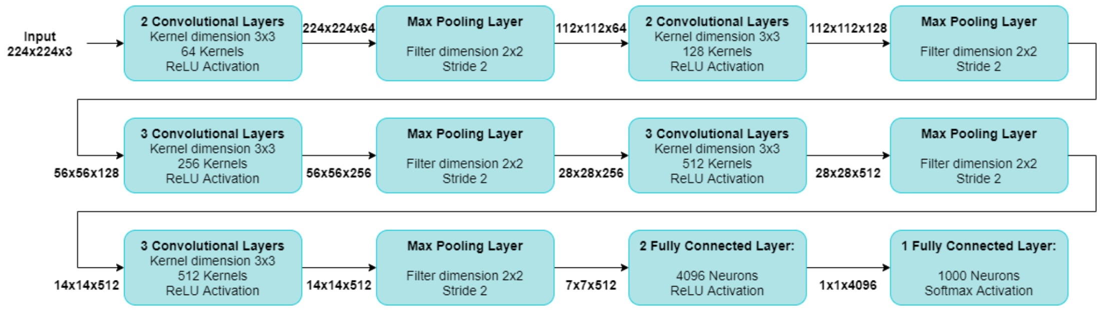


Figure 3.2: The detailed architecture of VGG16

VGG19

VGG19 is also a CNN algorithm which is 19 layers deep. It is basically a combination of convolution layers, fully connected layers, MaxPool layers and SoftMax layer. It also has approximately 144 million parameters [30].

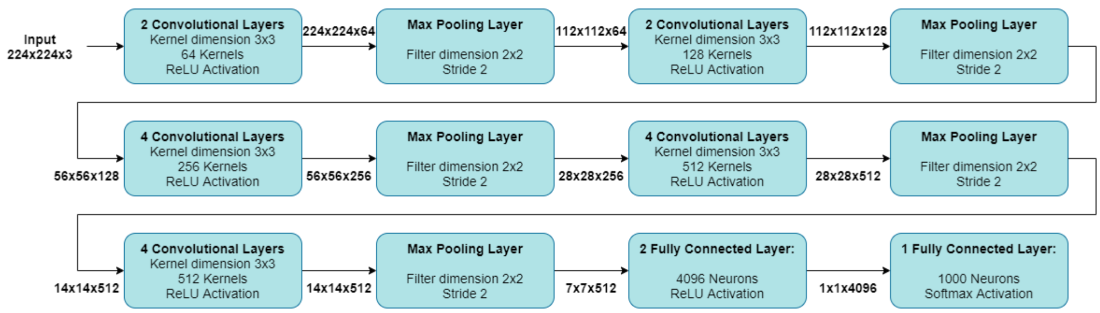


Figure 3.3: The detailed architecture of VGG19

3.2.2 ResNet

ResNet [31] which is also known as residual neural network is one kind of artificial neural network that builds on constructs known from pyramidal cells in the cerebral cortex. ResNet that works an Ensemble of Smaller Networks is very popular for solving problems like vanishing gradient with the help of its too much deep network. The fundamental feature of ResNet is that it allows training extremely deep neural networks with 150+layers successfully.

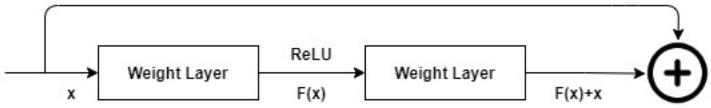


Figure 3.4: Residual Learning

ResNet50

ResNet-50 helps one to effectively train incredibly deep neural networks because it avoids the issue of the disappearing gradient. The concept of skip connections was first introduced in this model. In a model like this, along with the immediate layer, the output of a layer may be transferred to a distant layer. It means that all the layers are operating at the same power. In this model, there are 5 stages, each stage consisting of different convolution layer combinations [31].

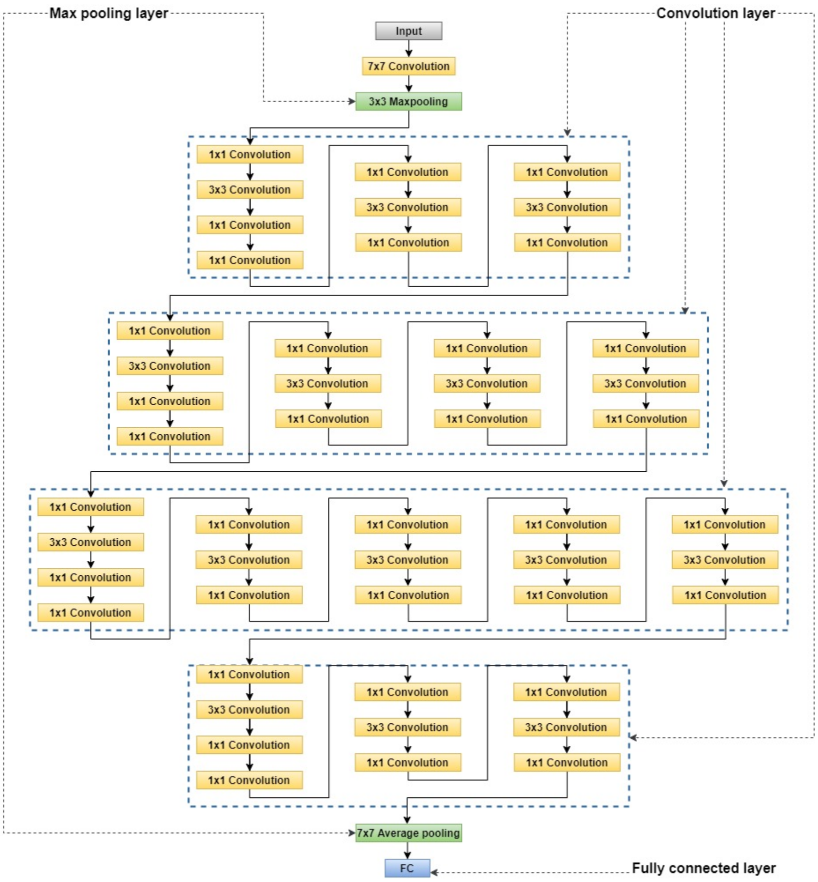


Figure 3.5: Internal Architecture of ResNet50

ResNet101

ResNet-101 is also a CNN model which is 101 layers deep. A pretrained version of the network trained on more than a million images from the ImageNet database can be easily loaded using ResNet-101 and this pretrained network can classify images into 1000 object categories [31].

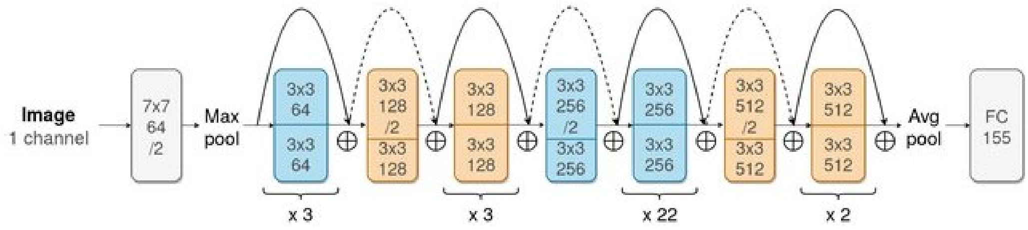


Figure 3.6: Internal Architecture of ResNet101

3.2.3 EfficientNet

EfficientNet is used as a simple and highly effective compound scaling method which helps to easily scale up a baseline ConvNet to any target resource constraints in a more principled way, while maintaining model efficiency [29]. EfficientNet is widely considered as one of the most powerful convolutional neural network architectures as well as a scaling method that uses a compound coefficient to uniformly scale network width, depth, and resolution in a principled way. This architecture is basically based on the baseline network developed by the neural architecture search using the AutoML MNAS framework. The compound scaling method allows EfficientNet models to be scaled in such a way that it achieves state-of-the-art accuracy with an order of magnitude fewer parameters and FLOPS, on ImageNet and other commonly used transfer learning datasets.

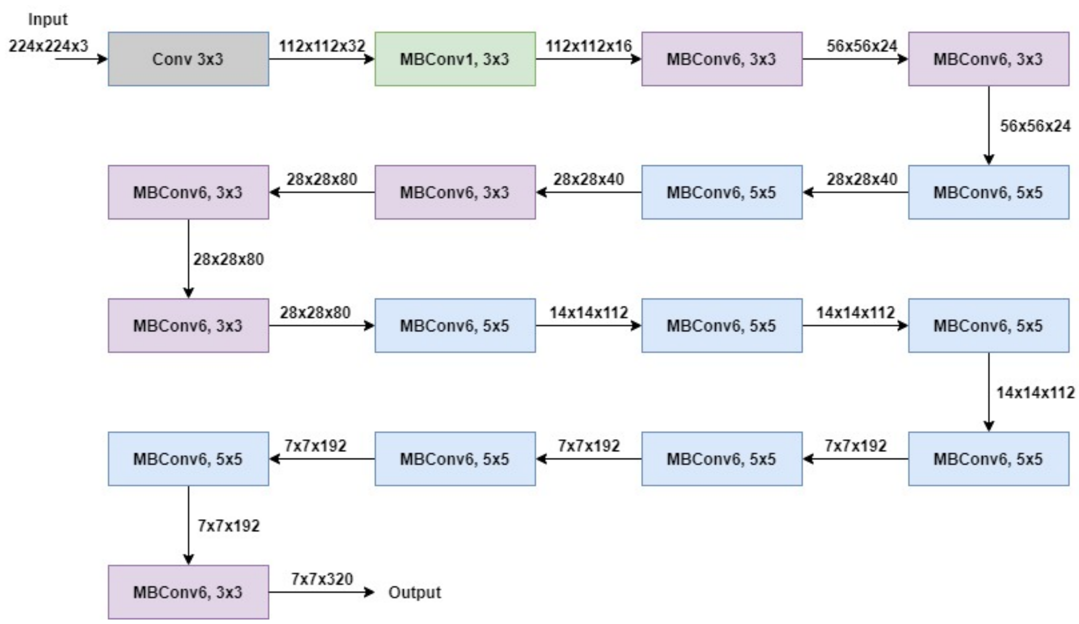


Figure 3.7: Internal Architecture of EfficientNet

3.3 Convolutional layer

Convolutional Neural Network's main building element is a convolutional layer. This network works well with two-dimensional data. It consists of convolutional filters that convert 2D into 3D and has excellent performance. Model learning is extremely quick.CNNs are multi-layer perceptrons that have been biologically inspired. They have a proclivity for recognizing visual patterns in raw picture pixels. Using multiple layer neurons and shared weights in each convolutional layer, these deep networks look at tiny areas of the input picture called receptive fields[32][33].

3.3.1 Activation function

The most important units in the structure of a neural network are net inputs, which are processed and changed into an output result known as unit activation by using an activation function [34]. A network layer's output values can be anything between $-\infty$ to ∞.

Rectified Linear Unit(ReLU)

In neural networks, ReLU is a non-linear activation function that is commonly utilized[35]. Because not all neurons are stimulated at the same time, ReLU is more efficient than other functions.The weights and biases are not updated during the back-propagation stage in neural network training in some circumstances because the gradient value is zero.

$$f(x) = max(0, x)$$

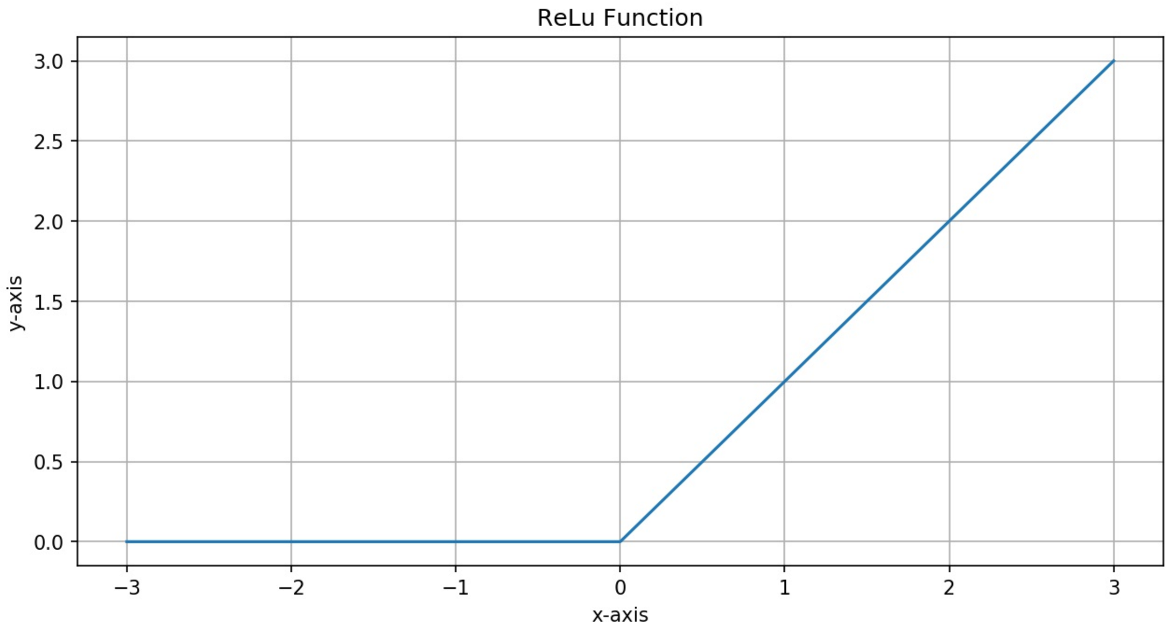


Figure 3.8: Plotted Graph of ReLU Function

Softmax

Multiple sigmoid curves are assembled to obtain the Softmax function. Because a sigmoid function provides values in the range of 0 to 1, we may use them as probabilities of a certain class's data points[35] [36]. Unlike sigmoid functions, which are utilized for binary classification, the Softmax function may be utilized to solve issues with many classes.The probability is returned by the function for each data point in each particular class.

$$f(x) = \frac{e^{x_i}}{\sum_{n=1}^{k} e^{x_i}}$$

Here x_i is the output value of i-th neuron given i= 1,2,3. . . , k and the bigger the value of x_i, the greater the probability we will get for that instance

3.3.2 Max Pooling Layer

Max pooling is a sample-based filter which picks the image's brightest pixels and its overall purpose is to down-sample an input configuration (image, hidden-layer output matrix, etc.) by diminishing its multiplicity and permitting assumptions about the features included within the scrapped subregions. It is highly useful for darkening the image's backdrop and lightening the image's pixels. It works by applying a maximum filter to (typically) non-overlapping subregions of the initial representation. It also lessens the computational effort by minimizing the number of parameters to learn and gives the underlying depiction fundamental linearity.

3.4 Optimizer: Adam

Adam is a new gradient-based optimizer for stochastic fitness function that is intended at machine learning applications with big datasets and/or steeply parameter spaces [37]. The advantages of two recently popular optimization methods (the ability of AdaGrad to deal with sparse gradients, and the ability of RMSProp to deal with non-stationary objectives) are combined in this method. This approach is easy to accomplish and consumes little memory/processing resources.

3.5 Ensemble Modeling

Ensemble modeling [30] [38] is a multilayer way to create predictions based on a collection of several varied models. The inputs and outputs for each model are kept the same form for the averaging layer to modify. More significantly, this sort of modeling minimizes errors, such as misleading negative and positive predictions. Nonetheless, there is a stipulation for using ensemble mode, which is that the models being ensembled must be architecturally distinct.
In a word, an ensemble model is a unified model that combines numerous models to forecast the occurrence more accurately and with improved feature analysis.

Averaging Layer

In our proposed "COV19EXAI v1" model, we have used the fusion of VGG16 and VGG19 architectures and for "COV19EXAI v2", we have used the fusion of VGG16,

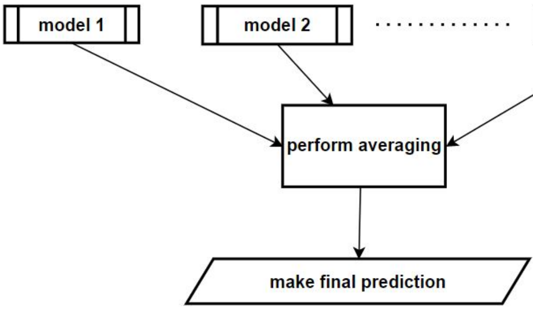


Figure 3.9: Basic structure of an ensemble model

VGG19 and ResNet101 architectures. Here, in the averaging layer, the layer itself will take the output probability from different architectures having same output dimension, as input. Furthermore, it will make an average of those multiple probabilities and estimate the average probability for the two labels: COVID positive and COVID negative.

Output Layer

In this layer, depending on the output probability of the averaging layer our proposed "COV19EXAI v1" and "COV19EXAI v2" models predict either the patient to be COVID positive or negative based on given CT-scan input images to the models of that individual.
For example, if we get an prediction matrix where the probability of first index is greater than the probability of second index from the output layer. Then, the class will be COVID negative.

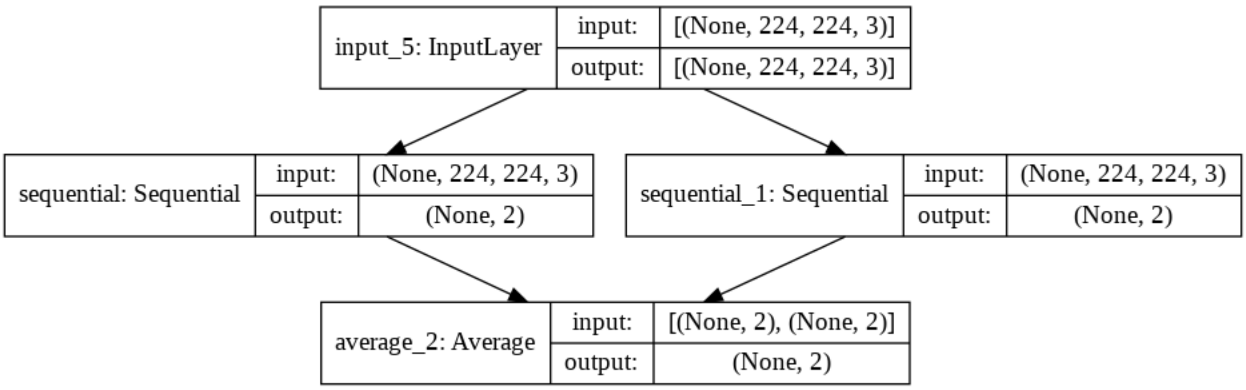


Figure 3.10: Output layers of the ensemble model hierarchy for COV19EXAI v1

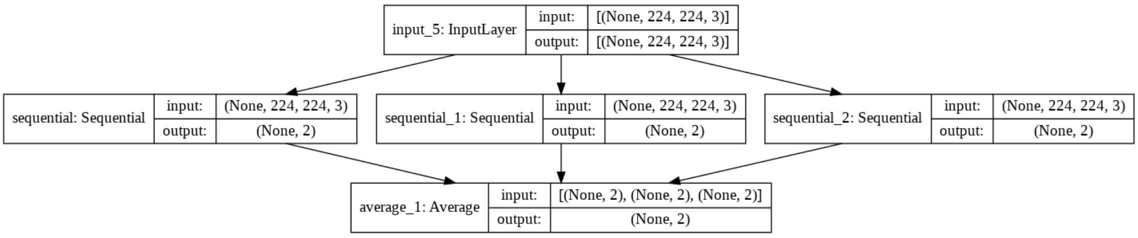


Figure 3.11: Output layers of the ensemble model hierarchy for COV19EXAI v2

3.6 Explainable Artificial Intelligence (XAI)

Explainable Artificial Intelligence (XAI) [39] creates a suite of machine learning techniques that produces more explainable models, while maintaining a high level of learning performance. Through adapting deep learning techniques to obtain explainable attributes, it allows supervised care. It also contains procedures for acquiring more hierarchical, generalizable, and explanatory representations, and other pattern inference tools for asserting an understandable paradigm with any black - box testing model. Client happiness, influencing the multispectral images goal attainment, credibility analysis, and consistency are all areas where XAI thrives. In other terms, Explainable AI (XAI) is artificial intelligence that permits learners to acknowledge the consequences of the procedure. It is a collection of tools and regulatory frameworks that help individuals in comprehending and deciphering machine learning model projections. Feature attributions for model predictions in AutoML Tables and AI Platform, and visually investigate model behavior using the What-If Tool can also be generated by the help of XAI.

Chapter 4

Implementation

4.1 Dataset

4.1.1 Source

COVID-CT: Constructed in [40]
COVID-19 image data collection: Constructed in [41]

4.1.2 Data Sample

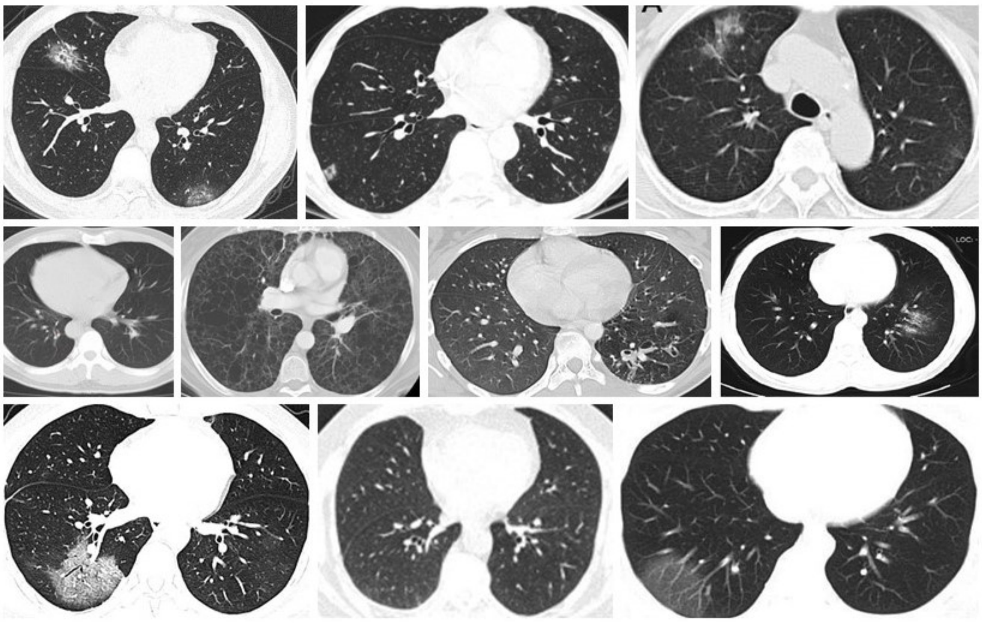


Figure 4.1: Sample Data of the Dataset

In figure 4.1, we can see some data samples of the CT-scan images of lungs.

4.1.3 Data Classification

We divided the dataset into 7:2:1 in our study, where 80% of the data was used to train the model, 20% was used as a test set and the remainder of the 10% was used to validate our models.

Training Set

The stage during which labeled example data with the responses or output labels are given to the machine learning algorithm process.

Testing Set

In certain instances, as the algorithm iterates to enhance performance, a series of examples is used for real-world research, it may learn special characteristics of the training set. For an unseen test collection, good results will improve confidence that the algorithm can have right answers in the real world.

Validation Set

A sample of data is utilized to do an unbiased evaluation of a model fit on the training dataset while tuning model hyperparameters. The evaluation becomes increasingly biased when competency on the validation dataset is included in the model setup.

4.2 Data Pre-processing

4.2.1 Image Resize

Our model is primarily designed to achieve the most positive diagnosis of COVID-19 CT-scan images. That is why, as demonstrated, we categorized and trained certain models separately. The dataset contains CT-scan images of patients who are diagnosed with COVID-19 and who are not. As we are training the network, every image is resized in a fixed size of 224×224. Therefore, Scikit Image [42], TensorFlow [43], and Caffe frameworks [44] will be used for this purpose. Before modeling, the ImageDataGenerator [45] [46] class in Keras is used to scale pixel values in our picture dataset. This class aggregates our picture dataset during creation, verification, or evaluation, then restores photos to the algorithm in batches and does scaling operations as needed when asked. This gives a strong and logical way for scaling visual data when modeling using neural networks. The Image Data Generator can handle a range of various feature selection methods as well as percentage of pixel scaling approaches. The Image Data Generator class, which largely uses the mean derived on the training dataset as feature-wise centering, allows a reference to leveling. Before regression on the training sample, statistics must be calibrated.

4.2.2 Normalization and Scaling Images

Principal component analysis (pca) [47] [48] is a multivariate technique for analyzing a data table in which observations are characterized by a number of quantitative dependent variables that are all inter-related. The eigen flat fields of a set of flat fields are computed using principal component analysis (PCA). The most relevant eigen flat fields are then linearly combined to normalize each CT-scan image projection. Experiments show that the suggested dynamic flat field correction reduces systematic errors in projection intensity normalization by a substantial amount when compared to the standard flat field correction.The Image Data Generator class of Keras will be used for this purpose. Normalization is the process of reducing data to a scale of 0-1 scale. This may be obtained by converting the re-scale input to a ratio that can be multiplied by each pixel. This may be accomplished by converting the re-scale parameter to a ratio that can be multiplied by each pixel to get the desired range.

4.2.3 Data Augmentation

To improve the performance of our model, we augment our data by adding a series of random changes to the photos /citechollet2016building. We rotate (90, 180, and 270 degrees) and translate the photographs, as well as flip them horizontally and vertically. The ImageDataGenerator class of Keras will be utilized for this.The mageDataGenerator class in the Keras deep learning neural network framework can fit models using picture data augmentation .Image data augmentation is required to increase the size of the training dataset and improve the model's capacity to generalize. The photographs from the dataset aren't determined analytically. Only upgraded photographs are provided to the model as an option. The random augmentation of images allows for the production of both modified and near-copy images for training purposes. A Data Generator specifies the validation dataset and the test dataset. A separate ImageDataGenerator instance is typically utilized, which may have the same pixel scaling setting as the ImageDataGenerator instance used for the training dataset, but not for data augmentation. The reason for this is that data augmentation is only used to artificially enlarge the training data set in the hopes of improving model performance on an unsegmented dataset. The term "image shifting" refers to the process of moving all of the pixels in a picture in one direction, either horizontally or vertically, as long as the picture dimensions remain the same. The width and height shift range parameters to the Image Data Generator constructor control the amount of horizontal and vertical shift in a sequential manner.

4.3 Architecture Training

Finding the proper weights for the Neural Connections is the first step in training a Neural Network.In order to train neural networks, we divided our dataset into three sections."Train" is the first segment.The data sample that was utilized to fit the model, which was 70%.VGG16, VGG19, ResNet50, ResNet101, and EfficientNetB0 were all ran with it.When we ran these neural network topologies, the batch size was 16 and the epoch was 40."Test" and "validate" make up the rest of the dataset, with test accounting for 20% of the total and validate accounting for 10%.We ran this dataset on our system, which had the following configuration settings are CPU: Intel Core i5 8500, RAM: 16GB, and GPU: NVIDIA GTX1070Ti 8GB.

4.4 Explainable AI Implementation

We used LIME (Local Interpretable Model-Agnostic Explanations) based XAI framework to mark the positive, negative and interruption regions of the images. Using LIME architecture will provide us a better understanding about the 'black box' neural network classification.

Chapter 5

Result and Analysis

5.1 Individual Model

Architecture	val_categorical_accuracy	val_loss
VGG16	85.33%	0.7079
VGG19	87.86%	0.6309
ResNet50	79.44%	1.6673
ResNet101	82.35%	1.3806
EfficientNetB0	80.75%	0.5900

Table 5.1: Comparison of between our used models

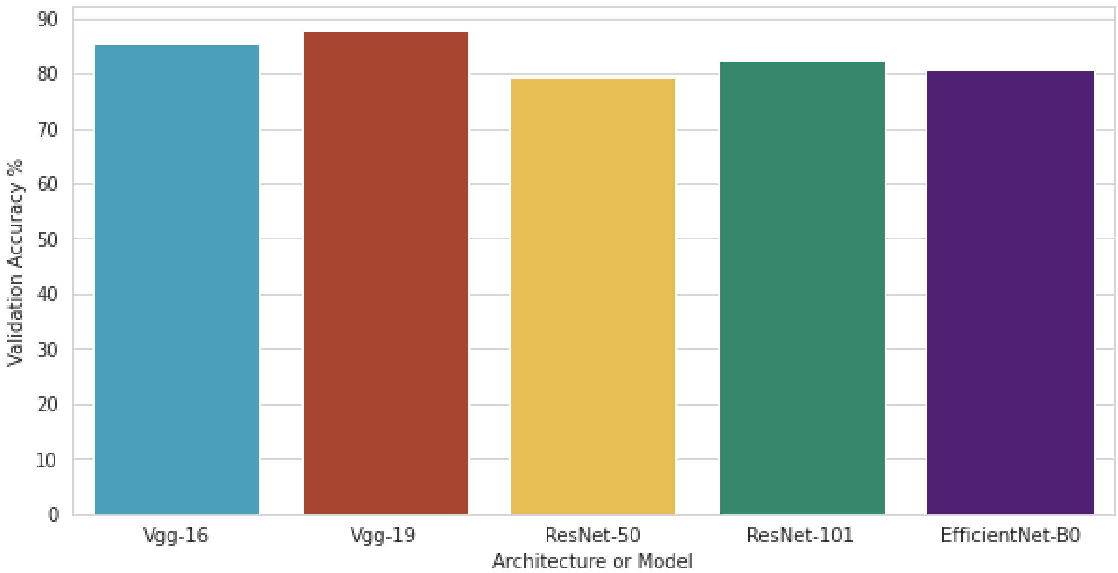


Figure 5.1: Illustration of comparison between our used models using a bar chart

From the val_categorical_accuracy of table 5.1, we can determine that VGG19, VGG16 and ResNet101 shows us 87.86%, 85.33% and 82.35%. So, these are the best three performed architecture. Therefore, we are considering VGG19, VGG16 and ResNet101 for the next level implementation.

5.2 COVID-19 Analysis by Ensemble modelling and Explainable AI version 1 (COV19EXAI v1)

In our "COV19EXAI v1" model, we combined our two best performed architecture which are VGG19 and VGG16 for ensemble modeling.

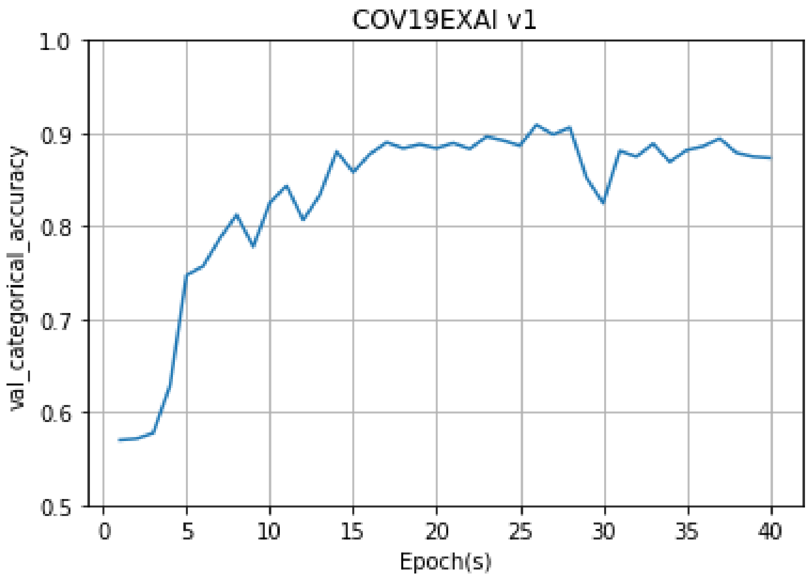


Figure 5.2: Validation accuracy curve of COV19EXAI v1

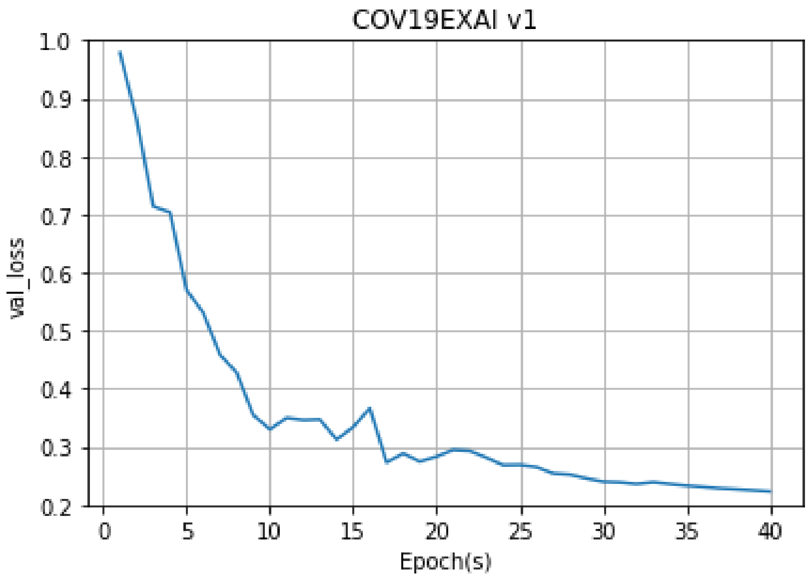


Figure 5.3: Validation loss curve of COV19EXAI v1

From the learning curves of COV19EXAI v1 we can see that we have achived 90.89% validation accuracy.

5.3 COVID-19 Analysis by Ensemble modelling and Explainable AI version 2 (COV19EXAI v2)

In our "COV19EXAI v2" model, we combined our three best performed architecture which are VGG19, VGG16 and ResNet101 for ensemble modeling.

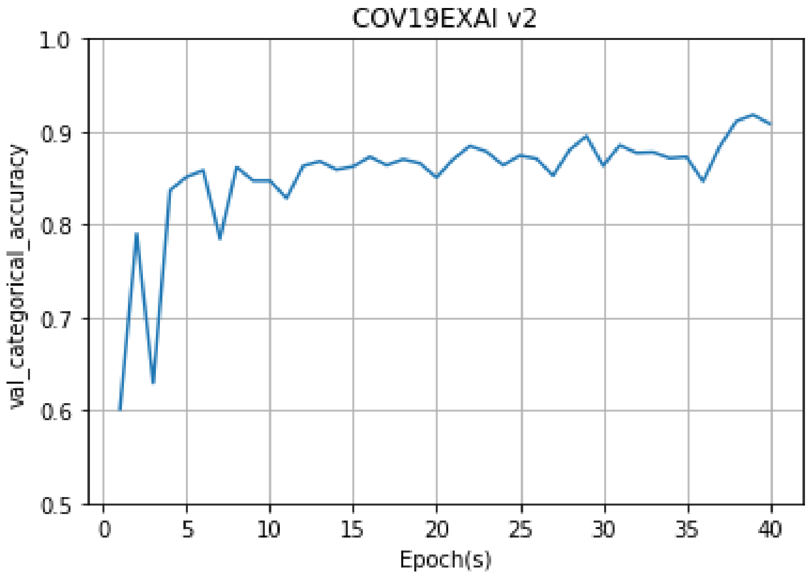


Figure 5.4: Validation accuracy curve of COV19EXAI v2

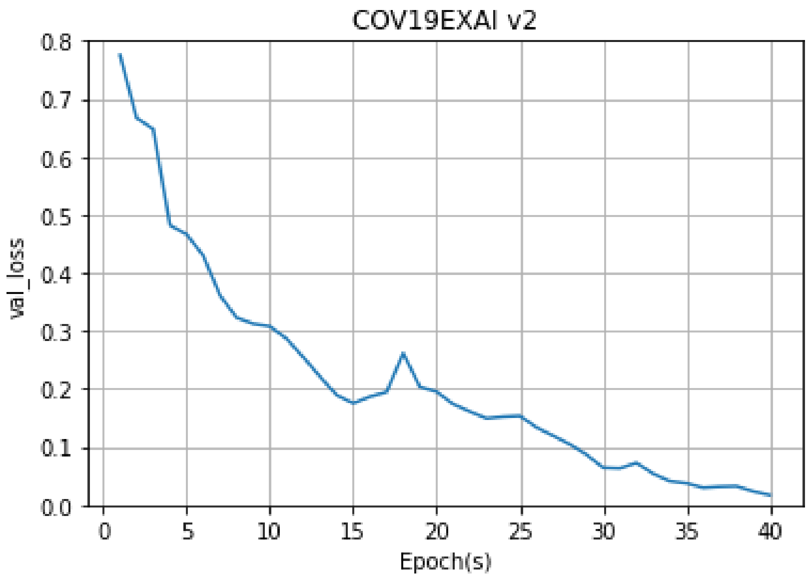


Figure 5.5: Validation loss curve of COV19EXAI v2

From the learning curves of COV19EXAI v2 we can see that we have achieved 91.82% validation accuracy.

5.4 COV19EXI-v1 vs COV19EXI-v2

Model	val_categorical_accuracy	val_loss
COV19EXAI v1	90.89%	0.2652
COV19EXAI v2	91.82%	0.2358

Table 5.2: Comparison between our used models

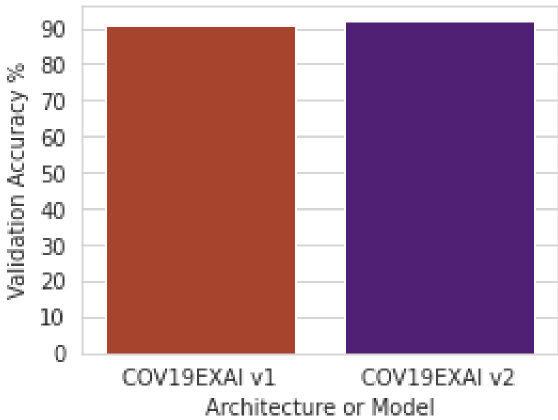


Figure 5.6: Validation loss curve of COV19EXAI v2

From the val_categorical_accuracy of table 5.1, we can determine that "COV19EXAI v2" shows us 91.82% validation accuracy. So, among our two models "COV19EXAI v2" are the best performed model. Therefore, we considered "COV19EXAI v2" for the next level implementation.

5.5 Explainable Artificial Intelligence (XAI)

After implementing Explainable AI on "COV19EXAI v2", we have noticed that there are 3 types of masks represent by 3 different colors Yellow, Red and Green. Here, Yellow borders represent the interpretable regions, Green mask represents the COVID positive responding regions and Red mask represents the COVID negative responding regions.

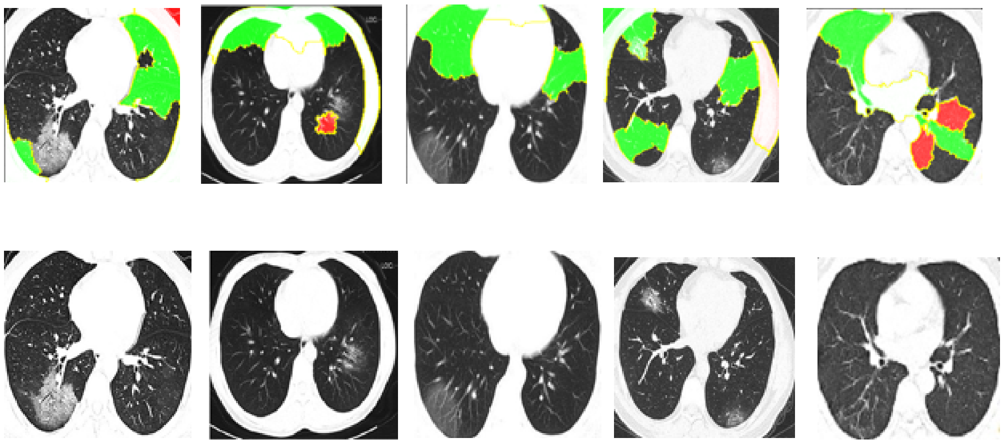


Figure 5.7: Output of XAI for COVID positive cases

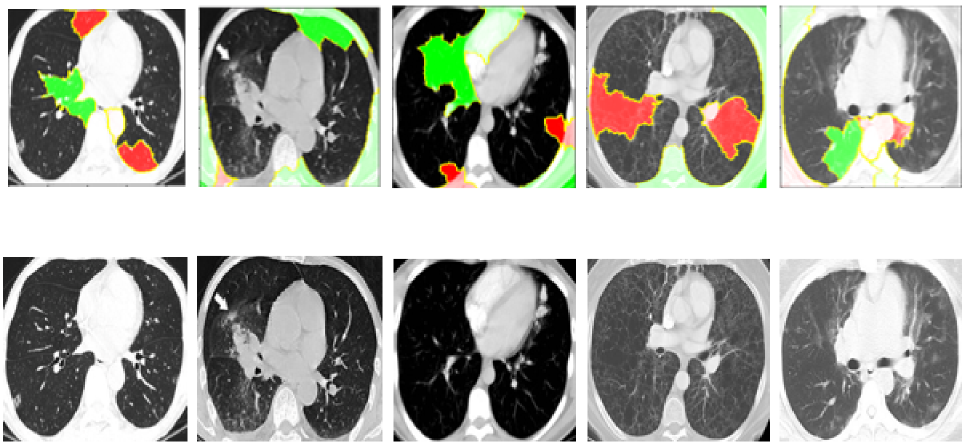


Figure 5.8: Output of XAI for COVID negative cases

We can see that from figure 5.7 and 5.8 that COVID positive region responded mostly mid to the higher portion of the lungs. However, we can also see that there are some wrong classification occurred. Additionally, in the figures, we can see that there are several CT-scan images has masked incorrectly at the very bottom part of the images. From our researched we figured out that, positive level images and the negative level images have some distinct differences at the bottom portion of the image. Therefore, the neural network is interpreting it as a feature for distinguishing COVID and NonCOVID.

Chapter 6

Conclusion and Future Work

6.1 Conclusion and Future Work

In this critical situation of COVID-19 pandemic, the entire world is suffering beyond our imagination and the situation is getting worse every day. The number of COVID19 patients is increasing at a dramatic rate as well as the number of deaths. Millions of people are coming to hospitals every day to test but it takes 10-15 hours to get the results. As there is a huge time required for getting results, some patients who are in severe condition, are not getting the proper treatment in due time and this delay in treatment is costing them their lives. So, to make this condition a little bit improved, our model will be able to detect the CoronaVirus by analyzing the CT-scan images of the patients and also be able to classify the intensity of COVID-19 based on the image data. Moreover, our system will be able to do the binary classification of COVID-19 in a fastest way. In research work, we implemented VGG16, VGG19, EfficientNetB0, ResNet50 and ResNet152 architecture on the dataset to train the data and compared the result with our own architecture COV19EXAI V1 (Ensemble of VGG16 and VGG19 architecture), COV19EXAI V2 (Ensemble of VGG16, VGG19 and ResNet101 architecture) and XAI implantation. In future, we will try to gather more data to improve our model and resolve the wrong classifications which occurred while implementing XAI. Thus, we will be able to build a process that will generate an efficient way to contribute to the COVID-19 crisis.

Bibliography

[1] A. Filatov, P. Sharma, F. Hindi, and P. S. Espinosa, “Neurological complications of coronavirus disease (covid-19): Encephalopathy,” *Cureus*, vol. 12, no. 3, 2020.

[2] A. Jacobi, M. Chung, A. Bernheim, and C. Eber, “Portable chest x-ray in coronavirus disease-19 (covid-19): A pictorial review,” *Clinical Imaging*, 2020.

[3] F. M. Salman, S. S. Abu-Naser, E. Alajrami, B. S. Abu-Nasser, and B. A. Alashqar, “Covid-19 detection using artificial intelligence,” 2020.

[4] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, “Automated detection of covid-19 cases using deep neural networks with x-ray images,” *Computers in Biology and Medicine*, p. 103 792, 2020.

[5] F. Ucar and D. Korkmaz, “Covidiagnosis-net: Deep bayes-squeezenet based diagnostic of the coronavirus disease 2019 (covid-19) from x-ray images,” *Medical Hypotheses*, p. 109 761, 2020.

[6] D. Singh, V. Kumar, and M. Kaur, “Classification of covid-19 patients from chest ct images using multi-objective differential evolution–based convolutional neural networks,” *European Journal of Clinical Microbiology & Infectious Diseases*, pp. 1–11, 2020.

[7] A. A. Ardakani, A. R. Kanafi, U. R. Acharya, N. Khadem, and A. Mohammadi, “Application of deep learning technique to manage covid-19 in routine clinical practice using ct images: Results of 10 convolutional neural networks,” *Computers in Biology and Medicine*, p. 103 795, 2020.

[8] I. D. Apostolopoulos, S. I. Aznaouridis, and M. A. Tzani, “Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases,” *Journal of Medical and Biological Engineering*, p. 1, 2020.

[9] P. K. Sethy and S. K. Behera, “Detection of coronavirus disease (covid-19) based on deep features,” *Preprints*, vol. 2020030300, p. 2020, 2020.

[10] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song, *et al.*, “Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct,” *Radiology*, 2020.

[11] H. Panwar, P. Gupta, M. K. Siddiqui, R. Morales-Menendez, and V. Singh, “Application of deep learning for fast detection of covid-19 in x-rays using ncovnet,” *Chaos, Solitons & Fractals*, p. 109 944, 2020.

[12] S. Wang, Y. Zha, W. Li, Q. Wu, X. Li, M. Niu, M. Wang, X. Qiu, H. Li, H. Yu, *et al.*, “A fully automatic deep learning system for covid-19 diagnostic and prognostic analysis,” *European Respiratory Journal*, 2020.

[13] T. B. Alakus and I. Turkoglu, "Comparison of deep learning approaches to predict covid-19 infection," *Chaos, Solitons & Fractals*, vol. 140, p. 110 120, 2020.

[14] M. Habibzadeh, M. Jannesari, Z. Rezaei, H. Baharvand, and M. Totonchi, "Automatic white blood cell classification using pre-trained deep learning models: Resnet and inception," in *Tenth International Conference on Machine Vision (ICMV 2017)*, International Society for Optics and Photonics, vol. 10696, 2018, p. 1 069 612.

[15] T. A. Amma, A. R. Sunny, K. Biji, and M. Mohanan, "Lung cancer identification and prediction based on vgg architecture," *International Journal of Research in Engineering, Science and Management*, vol. 3, no. 7, pp. 88–92, 2020.

[16] N. K. Chowdhury, M. A. Kabir, M. Rahman, N. Rezoana, *et al.*, "Ecovnet: An ensemble of deep convolutional neural networks based on efficientnet to detect covid-19 from chest x-rays," *arXiv preprint arXiv:2009.11850*, 2020.

[17] Z. Müftüoğlu, M. A. Kizrak, and T. Yildlnm, "Differential privacy practice on diagnosis of covid-19 radiology imaging using efficientnet," in *2020 International Conference on INnovations in Intelligent SysTems and Applications (INISTA)*, IEEE, 2020, pp. 1–6.

[18] P. A. K. K. Diallo and Y. Ju, "Accurate detection of covid-19 using k-efficientnet deep learning image classifier and k-covid chest x-ray images dataset," in *2020 IEEE 6th International Conference on Computer and Communications (ICCC)*, IEEE, 2020, pp. 1527–1531.

[19] Y. Chaudhary, M. Mehta, R. Sharma, D. Gupta, A. Khanna, and J. J. Rodrigues, "Efficient-covidnet: Deep learning based covid-19 detection from chest x-ray images," in *2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM)*, IEEE, 2021, pp. 1–6.

[20] I. D. Apostolopoulos and T. A. Mpesiana, "Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, 2020.

[21] M. F. Aslan, M. F. Unlersen, K. Sabanci, and A. Durdu, "Cnn-based transfer learning–bilstm network: A novel approach for covid-19 infection detection," *Applied Soft Computing*, vol. 98, p. 106 912, 2021.

[22] M. Loey, G. Manogaran, and N. E. M. Khalifa, "A deep transfer learning model with classical data augmentation and cgan to detect covid-19 from chest ct radiography digital images," *Neural Computing and Applications*, pp. 1–13, 2020.

[23] R. K. Singh, R. Pandey, and R. N. Babu, "Covidscreen: Explainable deep learning framework for differential diagnosis of covid-19 using chest x-rays," *Neural Computing and Applications*, pp. 1–22, 2021.

[24] V. Sharma, S. Chhatwal, B. Singh, *et al.*, "An explainable artificial intelligence based prospective framework for covid-19 risk prediction," *medRxiv*, 2021.

[25] Q. Ye, J. Xia, and G. Yang, "Explainable ai for covid-19 ct classifiers: An initial comparison study," *arXiv preprint arXiv:2104.14506*, 2021.

[26] M. Rezaul Karim, T. Döhmen, D. Rebholz-Schuhmann, S. Decker, M. Cochez, and O. Beyan, "Deepcovidexplainer: Explainable covid-19 diagnosis based on chest x-ray images," *arXiv e-prints*, arXiv–2004, 2020.

[27] A. K. Das, S. Ghosh, S. Thunder, R. Dutta, S. Agarwal, and A. Chakrabarti, "Automatic covid-19 detection from x-ray images using ensemble learning with convolutional neural network," *Pattern Analysis and Applications*, pp. 1–14, 2021.

[28] T. Zhou, H. Lu, Z. Yang, S. Qiu, B. Huo, and Y. Dong, "The ensemble deep learning model for novel covid-19 on ct images," *Applied Soft Computing*, vol. 98, p. 106 885, 2021.

[29] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks. arxiv 2019," *arXiv preprint arXiv:1905.11946*, 2020.

[30] T. G. Dietterich *et al.*, "Ensemble learning," *The handbook of brain theory and neural networks*, vol. 2, pp. 110–125, 2002.

[31] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[32] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical image analysis using convolutional neural networks: A review," *Journal of medical systems*, vol. 42, no. 11, pp. 1–13, 2018.

[33] M. I. Razzak, S. Naz, and A. Zaib, "Deep learning for medical image processing: Overview, challenges and the future," *Classification in BioApps*, pp. 323–350, 2018.

[34] S. Sharma and S. Sharma, "Activation functions in neural networks," *Towards Data Science*, vol. 6, no. 12, pp. 310–316, 2017.

[35] R. Arora, A. Basu, P. Mianjy, and A. Mukherjee, "Understanding deep neural networks with rectified linear units," *arXiv preprint arXiv:1611.01491*, 2016.

[36] M. Wang, S. Lu, D. Zhu, J. Lin, and Z. Wang, "A high-speed and low-complexity architecture for softmax function in deep learning," in *2018 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS)*, IEEE, 2018, pp. 223–226.

[37] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[38] R. Polikar, "Ensemble learning," in *Ensemble machine learning*, Springer, 2012, pp. 1–34.

[39] D. Gunning, "Explainable artificial intelligence (xai)," *Defense Advanced Research Projects Agency (DARPA), nd Web*, vol. 2, no. 2, 2017.

[40] J. Zhao, Y. Zhang, X. He, and P. Xie, "Covid-ct-dataset: A ct scan dataset about covid-19," *arXiv preprint arXiv:2003.13865*, 2020.

[41] J. P. Cohen, P. Morrison, L. Dao, K. Roth, T. Q. Duong, and M. Ghassemi, "Covid-19 image data collection: Prospective predictions are the future," *arXiv 2006.11988*, 2020. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset.

[42] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, *et al.*, "Scikit-learn: Machine learning in python," *the Journal of machine Learning research*, vol. 12, pp. 2825–2830, 2011.

[43] M. Abadi, P. Barham, J. Chen, Z. Chen, A. Davis, J. Dean, M. Devin, S. Ghemawat, G. Irving, M. Isard, *et al.*, "Tensorflow: A system for large-scale machine learning," in *12th {USENIX} symposium on operating systems design and implementation ({OSDI} 16)*, 2016, pp. 265–283.

[44] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," in *Proceedings of the 22nd ACM international conference on Multimedia*, 2014, pp. 675–678.

[45] A. Gulli and S. Pal, *Deep learning with Keras*. Packt Publishing Ltd, 2017.

[46] N. K. Manaswi, N. K. Manaswi, and S. John, *Deep learning with applications using python*. Springer, 2018.

[47] S. Wold, K. Esbensen, and P. Geladi, "Principal component analysis," *Chemometrics and intelligent laboratory systems*, vol. 2, no. 1-3, pp. 37–52, 1987.

[48] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley interdisciplinary reviews: computational statistics*, vol. 2, no. 4, pp. 433–459, 2010.