X-Ray Classification to Detect COVID-19 Using Ensemble Model

by

Ishmam Ahmed Solaiman
19341012
Tasnim Islam Sanjana
19341011
Samila Sobhan
17141018
Tanzila Sultana Maria
17141004

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

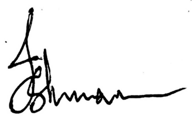
Department of Computer Science and Engineering
Brac University
June 2021

Declaration

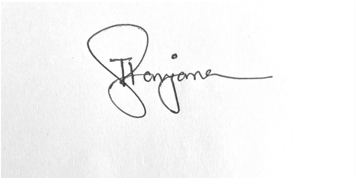
It is hereby declared that

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.

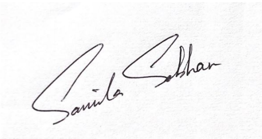
Student's Full Name & Signature:



Ishmam Ahmed Solaiman
19341012



Tasnim Islam Sanjana
19341011



Samila Sobhan
17141018

Tanzila Sultana Maria
17141004

Approval

The thesis/project titled "X-Ray Classification to Detect COVID-19 Using Ensemble Model" submitted by

1. Ishmam AHmed Solaiman (19341012)
2. Tasnim Islam Sanjana (19341011)
3. Samila Sobhan (17141018)
4. Tanzila Sultana Maria (17141004)

Of Spring, 2021 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on June 2, 2021.

Examining Committee:

Supervisor:
(Member)

Md. Khalilur Rhaman

Md Khalilur Rahman, PhD
Associate Professor
Department Of Computer Science And Engineering
BRAC University

Thesis Coordinator:
(Member)

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

Head of Department:
(Chair)

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

Ethics Statement (Optional)

This is optional, if you don't have an ethics statement then omit this page

Abstract

Diagnosis with X-Rays and other forms of medical images has soared to new heights as an alternative visual Covid infection detector. Radiographic images, primarily CT scans and X-Rays images play massive roles in assisting radiologists to detect and analyse severe medical conditions. Computer-Aided Diagnosis (CAD) systems are used successfully to detect diseases such as tuberculosis, pneumonia and other common diseases from chest X-ray images. CNNs have been widely adopted by many studies and achieved laudible results in the field of medical image diagnosis, having attained state-of-art performance by training on labeled data.This paper aims to propose an Ensemble model using a combination of deep CNN architectures, which are Xception, InceptionResnetV2, VGG19, DenseNet-201 and NasNetLarge, that can aid in the diagnosis of various diseases using image processing and artificial intelligence algorithms to quickly and accurately identify COVID-19 and other coronary diseases from X-Rays to stop the rapid transmission of the virus. In our experiment, we have used classifiers for the Xception model, VGG19, and Inception-Resnet model. We have compiled a CXR dataset from various open datasets. The compiled dataset was lacking 1000 images for viral pneumonia in comparison with Covid-19 and Normal CXRs, We used image augmentation and focal loss to compensate for the unbalanced data and introduce more variation. After implementing the focal loss function, we were able to get better results. Moreover, we implemented transfer learning on these models using ImageNet weights. Finally, we obtained a training accuracy of 92% to 94% across all models. Our Accuracy of the Ensemble Model was 96.25%.

Keywords: COVID-19; Pneumonia; Coronavirus; Deep Learning; X-Rays; Convolutional Neural Network; Ensemble model; Transfer learning; CAD

Dedication (Optional)

A dedication is the expression of friendly connection or thanks by the author towards another person. It can occupy one or multiple lines depending on its importance. You can remove this page if you want.

Acknowledgement

Firstly, all praise to Allah who bestowed upon us the opportunity to complete our thesis without any major interruption, particularly in these trying times. It has been an insightful journey where we could accumulate all sorts of knowledge and experience.
Secondly, our heartiest gratitude to our supervisor, Mr Md. Khalilur Rahman sir for his diligent support and guidance and being forbearing throughout the course of our work.
Finally to our parents without whom, every step would not have been as effortless. By the virtue of their endless sacrifices and prayers we are able to complete our graduation.
A moment to appreciate BRAC University and the authority for being the Alma Mater that gave us the necessary time, uninhibited support when we lacked direction and the resources to successfully reach our thesis completion.

Table of Contents

List of Figures

List of Tables

Nomenclature

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

α Alpha

γ Gamma

CNN Convolutional Neural Network

$COVID-19$ Coronavirus Disease 2019

$CRNN$ Convolutional Recurrent Neural Network

CXR Chest X-Rays

$DCNN$ Deep Convolutional Neural Network

GPU Graphics Processing Unit

$GradCAM$ Gradient-Weighted Class Activation Mapping

NN Neural Network

Chapter 1

Introduction

The Novel Coronavirus 2019 (COVID-19) was formulated in Wuhan of the Hubei province of China and spread drastically all over the world, risking millions of lives and the world economy. The World Health Organisation proclaimed the virus as a global pandemic on the 11th of March, 2020 and it took over Bangladesh on the 8th of March. The coronavirus is a highly contagious one, transmitted through the form of droplets from an infected person while sneezing or coughing. It can also be transmitted from touching contaminated surfaces and then the eyes, mouth, or nose. Some of the most common symptoms are fever, dry cough, experiencing breathing difficulties, sore throat, fatigue and losing the sense of smell and taste. A COVID-19 patient can carry the virus up to two weeks from the appearance of any of the symptoms. There are also many cases surrounding asymptomatic patients who unknowingly carry and spread the virus, affecting others. This is why the transmission of the virus is almost impossible to curb, making it a lethal disease with a high fatality rate.

1.1 Motivation

With the appalling second wave and the growing number of cases, timely detection and diagnosis of COVID-19 are essential and demanding. The real-time Reverse Transcription Polymerase Chain Reaction (RT-PCR) is the definitive test used for COVID-19 diagnosis but is not sensitive enough. It is unable to cater to the increasing number of patients every day. The process is not only time-consuming but also prone to error in times of emergencies. The biggest problem radiologists are facing now is dealing with false-negative results. Many people are unable to afford to take the test. Accurate diagnosis is essential for COVID-19 patients because they need to be quarantined to prevent the rapid spread of the virus. A modern way to detect diseases in extreme times and that too in an efficient, prompt way, must be adapted. An effective method of diagnosis with minimum variance is by implementing deep learning models on medical images. Detecting abnormalities and diagnosing severe conditions using medical images have had notable success such as detecting lung cancer and breast cancer in comparison to traditional analog techniques. Medical images display essential features such as complicated organ positions and tissue structure which are imperative for diagnosis. The development in graphic processing cards (GPU) hardware and deep learning techniques allow automatic detection from Chest X-Ray images with high rates of accuracy. Nevertheless, the use of X-Rays is

not entirely explored to its full potential. In a developing country like Bangladesh, with finite medical equipment and resources, the supremacy of disease detection using medical imaging does not reach out to the percentage of the population with limited means. Since Bangladesh is a disease-prone nation and the aggravating ratio of doctors to patients being 5.26:10000, providing immediate proper care is certainly not a privilege. Considering the spike in daily COVID cases, discrepancies in diagnosis are also highly unaffordable.
Radiological images are useful in the diagnosis and assessing the after-effects of COVID-19, for example, pneumonia. As many patients experience pneumonia as an after-effect, radiological examinations are necessary for follow-up and to track the recovery process. There are some detection systems available that utilize Chest Computed Tomography (CT) scans which have outperformed the RT-PCR test results. But these systems are expensive to install and their routine burdens radiologists, hence making them vaguely popular in developing countries. The need to recognize and successfully interpret COVID-19 features on Chest X-Rays is increasing. Keeping that in mind, X-Rays maintain the good potential to be a cost-effective approach to the aforementioned issues. In retrospect, there is a lack of widespread use of X-Rays based detection systems in diagnosis [19]. There are several machine learning and deep learning techniques designed to identify chest anomalies from X-Rays [12]. Deep learning is a subset of machine learning and deep learning techniques are artificial neural networks, processing data, focussing on automatic feature extraction and image classification. The biggest hurdle researchers face with developing deep learning-based diagnoses is that there are very limited open and available COVID-19 datasets. The ever-changing structure of the virus, coupled with the increasing number of patients makes it difficult to collect data.

1.2 Research Objective

Our work is based on relatively more Covid-19 data than any other papers we have come across, elaborated in the Dataset segment. Furthermore, we have trained our models using transfer learning. Transfer learning is a process by which the knowledge of a network, pre-trained initially with data, is used to perform a differently related task, using fresh data [14]. Transfer learning has proven to enhance performance in a time-effective manner [17]. It also produces better results when the size of the dataset is small[17], as in the case of COVID-19 datasets, since the disease is still fresh and the volume of data is low. We used image augmentation techniques to compensate for this by providing random augmentation of the images as they are fed into the models for training. This greatly improves the variation of training images lowering the potential for over-fitting the dataset.
We used 5 different feature extraction networks with a custom classification network to produce 5 X-ray classification models. We used a fully connected 2048 dense layer with a 10% dropout rate followed by a 1024 unit dense layer with a 20% dropout rate, both having a Relu activation function. Finally, the output layer is a 3 unit dense layer with a softmax activation function for the classification network. Along with that, we also used a GradCAM (Gradient-weighted Class Activation Mapping) to show the heatmap of the infections in the Chest X-ray images.

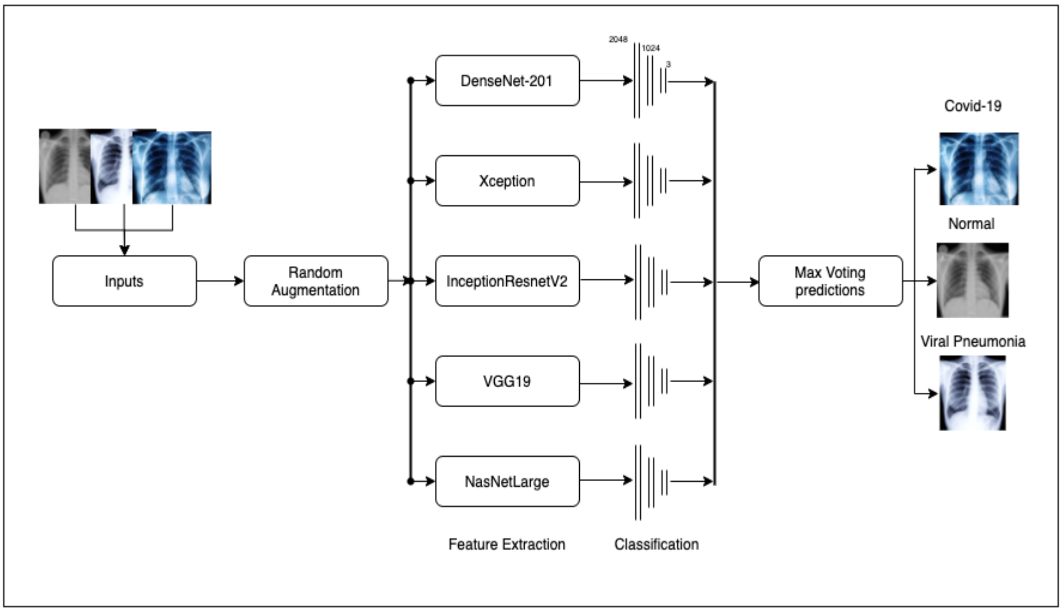


Figure 1.1: Model Overview

Chapter 2

Literature Review

A Convolutional Neural Network (CNN) is a type of deep neural network. CNNs are widely used in the field of image processing, classification, image segmentation and signal processing. There have been numerous studies that prove detection and diagnosis implementing CNNs are quicker and successful, especially in detecting pneumonia, tuberculosis, and lung cancer [20] [8][7] [3]. CNNs have made ground-breaking results in umpteen visualization tasks and computer-aided diagnosis (CAD). CADs help with the initial screening of images and attempt to reduce the load on radiologists. While the CT scan provides accurate diagnosis, X-Rays are more favored as they are comparatively inexpensive and easier to comprehend, with ample, scalable applications and extensively used in diagnosis and monitoring diseases. Datasets available on the public domain that contained labeled CXR images allowed researchers to apply deep learning algorithms for segmentation purposes, anatomical structure detection, detection of suspicious region anomalies and computer-aided diagnosis (CADx). Prior research includes thoracic disease identification and localization, lung regional segmentation and disease report generation. X-Rays portray crucial features such as textures and tissue structures which yield fruitful results in diagnosing lung diseases, in particular pneumonia, tuberculosis and lung cancer.

CNN is basically used to extract a feature map out of images and the corresponding branch structure. Wang et al. [11] connect X-Rays and segmentation based on deep learning to detect lesions. An instance segmentation algorithm is applied to bone segmentation to segment and label the clavicles and ribs automatically. 180 CXRs were randomly selected with the assistance of digital radiography (DR) machine at Hospital of Traditional Chinese Medicine.The basic network framework of Mask R-CNN, an improved structure of the Faster R CNN, for automatic segmentation and annotation method was implemented. The feature map is extracted by the basic network, followed by the candidate regions being screened by RPN. Lastly, the segmentation, classification, and mask tasks of image targets are completed by 3 branch structures. Compared with traditional manual labeling, automatic labeling has great significance for the auxiliary diagnosis and treatment of computers. Automatic labeling is believed to provide good prospects for the future application of AI in medical imaging. This paper is the first to propose an instance segmentation algorithm that solves the problem of automatic segmentation annotation in medical images.

In [3], the experiment was carried out on a vast dataset, implementing basic augmentation techniques to prevent overfitting, yet making optimal use of the large dataset. GoogleNet, Inception V3, and ResidualNet architectures were implemented. GoogleNet, architecture with sufficient complexity, achieves significant, random classification accuracy when labeling normal and abnormal. The results determined that further fine-tuning architectures are potentially effective at increasing performance but would not change the robust results significantly. Symmetry appears to be a salient feature of normal CXR images detected by the model. Although this model is not yet ready for clinical adoption, it promises a future functional classification network.

The authors in [12] propose an automatic COVID-19 classification model, where they have used both COVID and non-COVID-19 image datasets and implemented HRNet for feature extraction purposes. Initially, the model was trained for 25 epochs for each fold, associated with a 0.005 learning rate and a customized dice coefficient loss function. The input image size was 512×512 pixels and was grayscale. The results surpass existing models in terms of accuracy, specificity, sensitivity, and other evaluation metrics. HRNet avoids the loss of small target information in the feature map since the convolutions are parallelly connected and also for the high-resolution feature representation. A segmented COVID dataset containing 910 images was used for training, with ten-fold cross-validation. By implementing the K-fold algorithm, 1 fold was used for testing while the remaining 9 folds were used for training.

The pre-trained Vgg16 and ResNet-101 CNNs were compared with each other to analyze lung images. Images were classified into 2 classes - normal and abnormal and achieved a 82% success rate. Since the performance was relatively low, a different approach was implemented to measure accuracy. If the classification result was in the top 3 decisions determined by the network, the process was considered successful with a 90% success rate. Smaller network structures that provide higher performances for Chest X-ray chest classification were thus investigated. This model succeeded in detecting diseases using only the X-ray image without any prior knowledge about the patient's history. Three CNNs were examined comparatively increasing the number of layers. As for a future study, the network can be retrained to recognize the abnormalities in MR images. The size of the input images was reduced, sacrificing performance in order to reduce the training time [10].

In the attempt to improve the performance of deep learning models on small-sized datasets, transfer learning is currently the most popular strategy to be employed. Transfer learning empowers a deep learning model to adequately learn from a small dataset by transferring learned features from another deep learning model that recently learned from a similar, but larger sized dataset. An automatic deep learning-based method using X-rays to predict COVID-19 was proposed by Narin et al in (2020)[16]. The method used three deep CNN architectures. They used a dataset that consisted of 50 X-ray images of COVID-19 patients and 50 normal X-ray images and all the images were resized to 224×224.To overcome the issue of the predetermined number of dataset, the authors utilized transfer learning models. The dataset was divided into two parts: 80% for training and 20% for testing. The developed deep CNN was based on pre-trained models (ResNet50, InceptionV3, and Inception-

ResNetV2) and allowed the authors to differentiate COVID-19 from normal X-ray images. Transfer learning with the K-fold method was used as a cross-validation method with a k 1/4 [14]. The final results showed a convincing accuracy of 96.78%.

In [6] the authors strive to configure transfer learning from CheXNet to assimilate mammogram data. Their findings show the best configuration only employs the first two dense blocks from the original CheXNet model. The optimal number of layers in the last used block is also fewer than compared to the original model, i.e. 6 layers out of 12. A better procecure to search for hyperparameter, for instance, grid search and random search might be able to discover a more ideal configuration as opposed to the trial-and-error approach that is employed in this research.

InceptionV3 is a state-of-the-art model that is pre-trained and is used for transfer learning in this research [5]. This research analysis contributes notably with regards to GAN based synthetic data and four different types of deep learning based models which brought forth state-of-the-art comparable results [13]. The motivation behind using InceptionV3 as transfer learning is because of the lower error rate. The authors discussed how coronavirus can be the real trigger to open the course for rapid integration and installation of Deep Learning in hospitals and medical centers. They review the improvement of deep learning applications in medical image analysis, focusing on pulmonary imaging and giving insights into contributions to COVID-19. [22] Apostolopoulos and Mpesiana in [14] evaluated various state-of-the-art deep architectures on CXR images. VGG19 managed to achieve an accuracy of 98.75% and 93.48% for 2-class and 3-class classification functions respectively on a dataset consisting of 224 COVID-19, 504 normal and 700 pneumonia X-ray images, thus proving to be the best model.

U-nets and Mask RCNNs are used for segmentation tasks to label each pixel of images and are also widely used in medical image classification. However, obtaining successful results are often hindered since Computer Aided Designs (CADs) have stunt development courtesy of the overwhelming absence of labeled data and immense variations in chest X-Rays [3]. Moreover, segmentation plays a crucial role in training a model by getting rid of redundant data on the available image dataset in order for the model to converge primarily on the infected areas. But it has been overlooked in several previous research. Therefore UNet has been tried and tested for segmentation purposes in [12], where they used High-Resolution Network (HR-Net) for feature extraction embedding and the UNet for segmentation purposes.

In [21] the authors talk about the importance of diagnosis with Chest X Rays since the virus has also proven to transmit through asymptomatic patients. They discuss the ease of image diagnosis with the existence of state-of-the-art AI algorithms and easy access to huge data. These models can bridge the gap between diagnosis and result delivery time to simply minutes. The authors suggest that depending on one model can be restrictive since every model has a different method for extracting features from training samples. Thus keeping in mind the sensitivity of the situation and urgent need for correct diagnosis, they suggest an ensemble model that comprises 5 state-of-the-art deep CNN models: VGG19, DenseNet201, ResNet50, ResNet34, and MobilNetV2, to automatically detect COVID-19 in X-Rays. Ensembling models classify new data points from a set of classification models by implementing max voting based on their predictions. The authors plan to increase the prediction accuracy of COVID-19, while attempting to lower the percentage of error and increase robustness by putting together all the strengths of the existing models, using X-ray images collected from Kaggle websites and Github repositories. Their model consists of 2 main techniques: transfer learning and ensembling to be able to architect a robust detection model. However, their experiment is narrowed down to only detecting COVID-19 and nonCOVID-19 cases and our experiment is extended towards detecting viral pneumonia as well. The images were divided for training and validation in the ratio of 80:20. By applying the max voting system their ensemble model results attained a performance accuracy of 99%. The authors are confident that their versatile model has the potential to expand to detecting other chest-related diseases, for example, pneumonia and tuberculosis.

Following the circumstances surrounding insubstantial and restricted medical image datasets and motivated by the success of deep learning and image processing, the present work is going to apply transfer learning techniques that were pre-trained by ImageNet data to overcome lengthy training time and insufficient data. Transfer learning also plays a vital role in upgrading the accuracy of detection.

Chapter 3

Dataset

3.1 Data Description

The cardinal element in deep learning is data. For this experiment, we have accumulated radiography images from several public repositories and classified the images like the following- Viral Pneumonia, Normal and Covid-19. From the dataset by Tawsifur Rahman (COVID-19 Radiography Database) [15][24] we acquired 3616 COVID-19 images, 10192 Normal images from which we examine 3620 and 1345 Pneumonia images. The Chexpert dataset is a large compilation of 224,316 chest X-Ray images of 65,240 patients from Stanford University Medical Centre (CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison) [9]. We have taken 566 Pneumonia cases from this dataset for our research. 176 Pneumonia images were taken from The National Institutes of Health Clinical Centre Chest X-Rays dataset which is the most popular dataset used in the field of medical imaging research and diagnosis. It is the largest available in the public domain containing radiographies of many advanced cases of diseases (NIH Chest X-rays) [4]. As COVID-19 is a fairly new disease even though there were previous cases of coronavirus diseases namely SARS in 2002-2003 and MERS in 2012, datasets are very hard to access from hospitals. Thus, we had to solely rely on publicly available datasets for the course of our experiment. Other diseases and multiclass labels, for example, images containing both pneumonia and some other disease, were eliminated from the NIHCC and Chexpert dataset, focusing only on the aforementioned classes. All the images were read as RGB. Posteroanterior viewing images were only selected to maintain uniformity. After compilation and creation of our dataset, we randomly split 80% of the dataset for training and testing the remaining 20% for validation purposes. The resulting dataset was further split into train and test sets, maintaining a ratio of 80:20 once again. The training set contained 2492 Covid, 1675 Pneumonia, and 2496 normal images whereas the testing set contained 400 images of each class.

Some of the Chest X-Ray Images of COVID-19, Viral Pneumonia, and Normal Patients from our dataset are demostrated below.

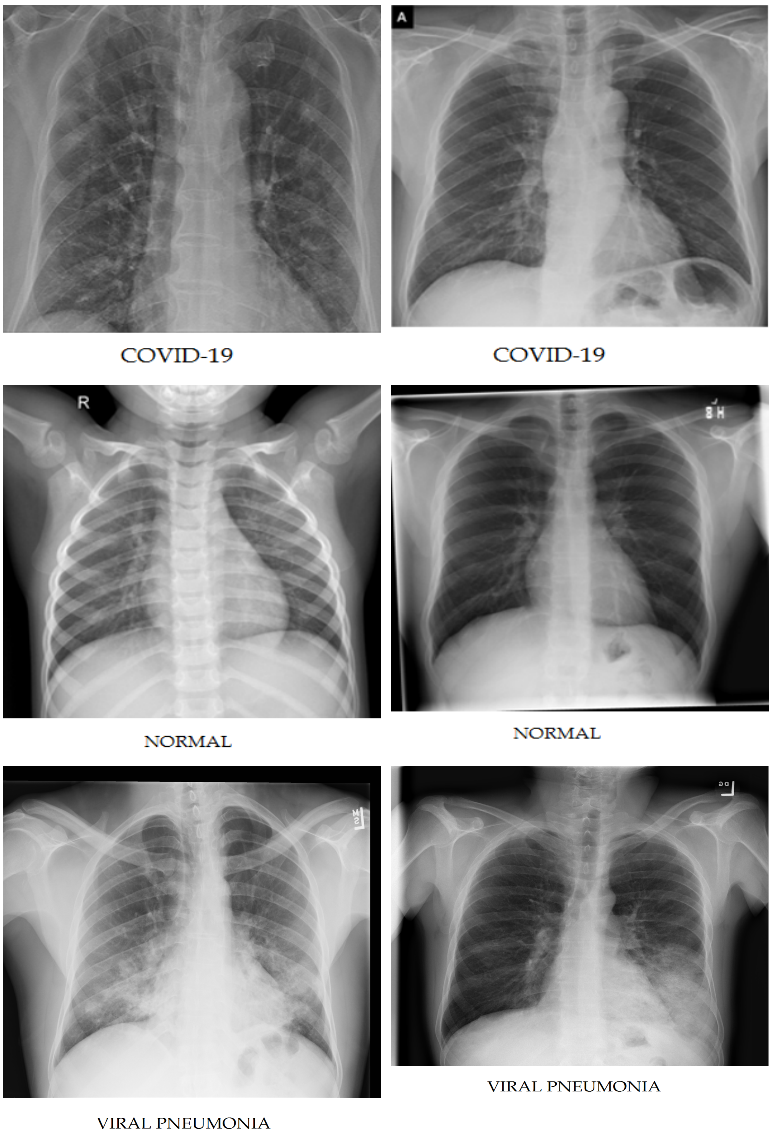


Figure 3.1: Sample Chest X-Ray Images of COVID-19, Viral Pneumonia and Normal Patients

3.2 Data Augmentation

We used ImageDataGenerator from TensorFlow which allows us to perform image augmentation while the data is being fed into the models each epoch. The images were resized to 299 x 299 pixels and augmented over a range of parameters. Firstly all the images are normalized and then a random combination and range of augmentation are applied to each image. This process occurs every epoch producing varied training data each epoch with random augmentation each time. The primary reason to augment our dataset is to increase the size of the dataset, prevent overfitting and add variation. The images were further tuned as shown in the following table:-

Random Augmentation	**Range**
Rotation range	0 - 30
Width Shift Range	0 - 0.2
Shear Range	0 - 0.2
Height Shift Range	0 - 0.2
Zoom Range	0 - 0.2
Channel Shift Range	0 - 0.1

For every epoch that's training, a new image was augmented. For example, each image was rotated a number of times. Even though our dataset was limited, data augmentation allowed us to get reliable training without overfitting.

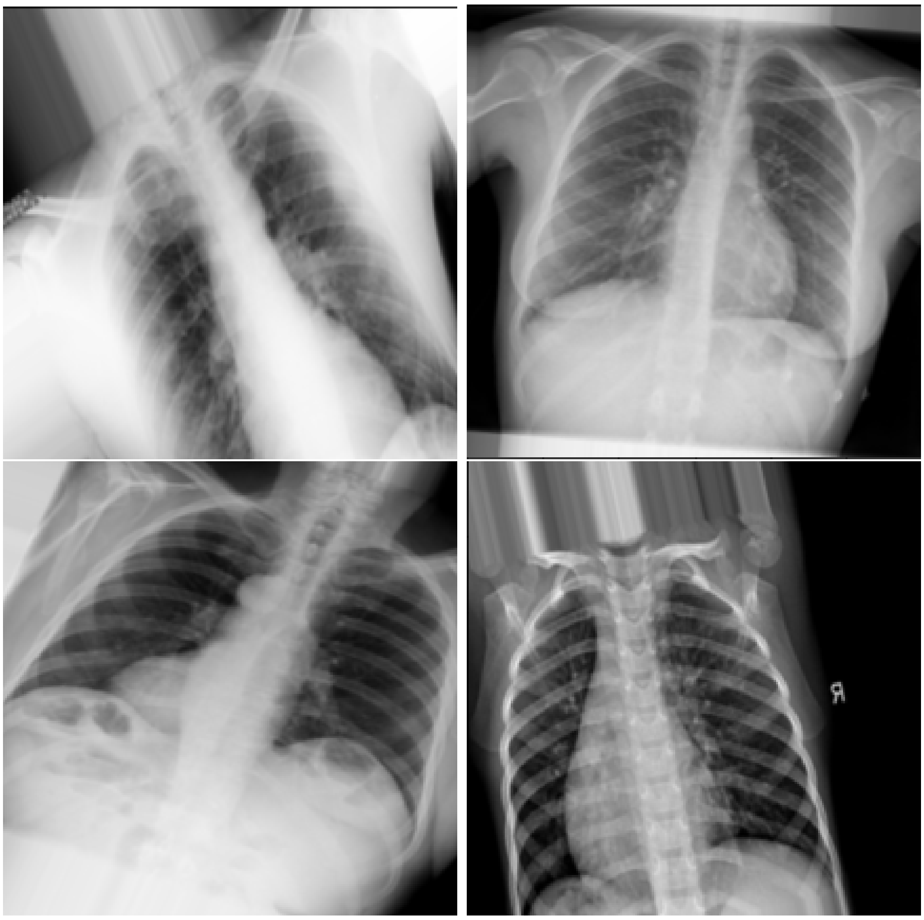


Figure 3.2: CXR Images after Augmentation

Chapter 4

Methodology

4.1 Architectures

For the course of this experiment, we have implemented five CNN architectures for feature extraction- InceptionResnetV2, Densenet201, VGG19, NasNetLarge and Xception. The last layers of all the aforementioned models were removed before our experiment, keeping only the convolutional layers and pooling layers. The structure of our model comprises one of the CNN architectures followed by a global average pooling layer then advances towards a dense layer with 2048 neurons using ReLu activation function and a 10% random dropout rate. Following that is structured another dense layer comprising 1024 neurons, Relu activation function in addition and a dropout of 20%. Lastly, there is a dense layer consisting of 3 neurons for the output class with Softmax activation. Below is an overview of out classification model.

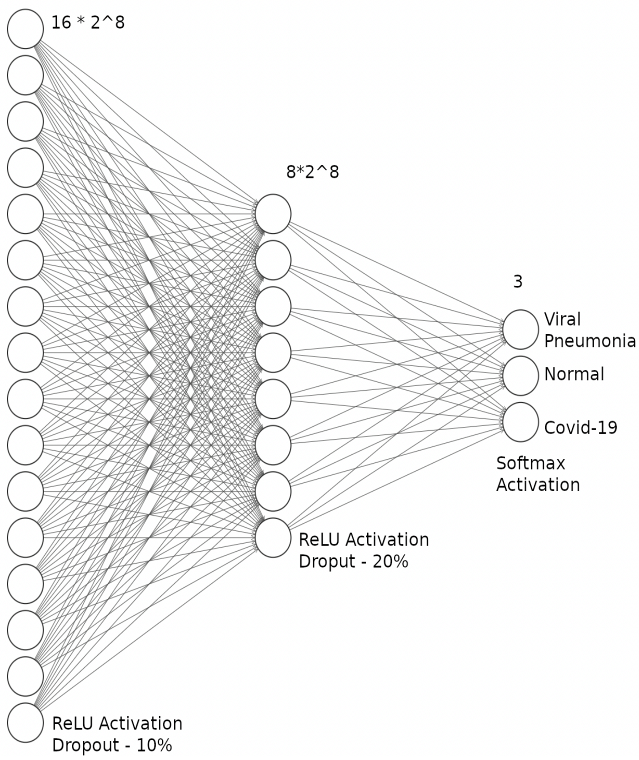


Figure 4.1: Architecture of the Classification Network

In [18], the Coronet architecture is based on Xception with a dropout layer and two fully connected layers at the end. This study accomplished an overall accuracy of 89.6% for 4 classes (Viral Pneumonia, Bacterial Pneumonia, COVID-19 and Normal) while we reached an accuracy as high as 92.53% when we implemented the same architecture. For a ternary classification among COVID-19, Pneumonia and Normal, much likely to our approach, Coronet yielded an accuracy of 95%. On the brighter side, when we implemented Xception with our existing architecture we were able to produce an accuracy of 93.67%. The adversity we faced with this was that the images were not generalised in the right manner as there was an excessive number of cases of False Positive and False Negative. However, with the addition of the layer with 2048 neurons as depicted in the model architecture, the dataset was better graphed and classified, with a lower number of False Positives and False Negatives. There were some oscillations in the results which was due to every epoch creating a newly augmented image. Thus each time the dataset was augmented in a disparate way, there were fluctuations in the results.

4.2 Proposed Model

Compared to other approaches, we present an ensemble deep learning method that will aid to improve deep learning prediction accuracies of COVID-19 and decrease the error-rate of misclassification by combining 5 different models. These models include: InceptionResNetV2, VGG19, NasNetLarge, Xception, and DenseNet201. Shifting from a single model, this approach allows the production of a better predictive performance model.A detailed explanation of the models is mentioned below. For Xception, VGG19, Densenet and InceptionResnetV2 models, adam optimizer and focal loss function were used. However, for Nasnet, the focal loss function did not show promising results. Therefore we switched to adamax optimizer replacing Adam optimizer to see if it worked. After showing unsatisfactory results, we switched to categorical cross-entropy loss function alongside the adamax optimizer.The Focal Loss and the Categorical Crossentropy Functions are defined as:

$$FL(p_t) = -\alpha_t(1 - p_t)^\gamma \log(p_t) \tag{4.1}$$

$$CL = -\sum_{i=1}^{outputsize} y_i \cdot \log(\widehat{y_i}) \tag{4.2}$$

All the models were initialized with imagenet training weights and were trained for 70 epochs where the test report was recorded at the 25th, 50th and the 70th epoch measuring precision, recall and F1-score along with accuracy and loss. For the first 50 epochs the weights for the feature extraction network of each model was frozen during training. After 50 epochs all the layers except batch normalization were unfrozen and trained a further 20 epochs. It can be observed that the accuracy has upgraded as more epochs were run.

4.2.1 InceptionResnetV2

InceptionResNetV2 is a CNN architecture that on more than a million images from the ImageNet database. It delivers good performance at a comparatively low computation cost. The technical difference between the Inception model and Inception

ResNet model is that the Inception variants are nonresidual and InceptionResnet variants are residual. This difference indicates that the batch-normalization concept is used only on top of the traditional layer and not above the residual summations[23]. InceptionResNetV2 is naturally 164 layers deep and after adding the 3 layers in our approach it is at 167 layers. The model consists of a total of 55,919,843 parameters of which 1,580,035 are trainable and 54,339,808 are non-trainable. During the first 25 epochs, the training and loss accuracy rested at 0.8837 and 2.1172 respectively. Following running the model for 50 epochs it exhibited a training accuracy of 0.9233 and a training loss of 1.7106. After unfreezing the layers, the non-trainable parameters were made active and a total of 55,919,843 parameters were trained for 70 epochs. The InceptionResnetV2 model achieved a training accuracy of 94.98% and testing accuracy of 92.75%.

Epoch		precision	recall	f1-score	accuracy
	COVID-19	0.96	0.79	0.87	
25	Normal	0.75	0.96	0.84	87.58%
	Viral Pneumonia	0.97	0.87	0.91	
	COVID-19	0.96	0.90	0.93	
50	Normal	0.86	0.96	0.91	92.33%
	Viral Pneumonia	0.98	0.93	0.95	
	COVID-19	0.95	0.93	0.94	
70	Normal	0.89	0.94	0.92	92.75%
	Viral Pneumonia	0.98	0.94	0.96	

Table 4.1: InceptionResnetV2

Figure 4.2: InceptionResNetV2 Accuracy after 25, 50, 70 Epochs

4.2.2 DenseNet201

In DenseNet,proposed by Gao Huang et al [2] and 201 layers deep, each layer inherits additional inputs from all preceding layers and passes on its own feature-maps to all succeeding layers. It has 2 characteristics: simplicity in the training process and exceptionally, parametrically efficient models, due to the potential of feature reuse by various layers. This intensifies the chances of variation in the subsequent layer inputs.

Densenet portrays the best results in terms of accuracy, precision and especially in F1-score compared with the rest of the models. The model achieved a training accuracy of 0.9574 after 25 epochs and increased to 0.9760 following another 25 epochs. DenseNet consists of 20,299,843 parameters of which 1,974,019 were trainable. Unfreezing and training for a total of 70 epochs,the remaining 18,325,824 parameters were activated and the DenseNet model achieved a training accuracy of 96.53% and testing accuracy of 94.83% after training on 20,299,843 in total.

Epoch		precision	recall	f1-score	accuracy
25	COVID-19	0.95	0.94	0.95	93.08%
	Normal	0.88	0.95	0.91	
	Viral Pneumonia	0.98	0.92	0.95	
50	COVID-19	0.97	0.94	0.95	94.92%
	Normal	0.89	0.96	0.92	
	Viral Pneumonia	0.97	0.93	0.95	
70	COVID-19	0.95	0.95	0.95	94.83%
	Normal	0.91	0.94	0.93	
	Viral Pneumonia	0.97	0.93	0.95	

Table 4.2: DenseNet201

Figure 4.3: DenseNet201 Accuracy after 25, 50, 70 Epochs

4.2.3 NasnetLarge

NasNet, which is Neural Architectural Search (NAS) Network, was manufactured by the Google ML team. It's architecture depends on reinforcement learning. NASNetLarge has been trained on over a million images from the Imagenet database and has the capability to classify images into 1000 class categories. NASNet-Large consists of 89,065,813 parameters, 4,140,931 trainable and 84,924,882 non-trainable. It is a CNN architecture with an image input size of 331 x 331. The parts of the architecture incorporate a Controller Recurrent Neural Network (CRNN) and a CNN block.

NASNet includes two sorts of cells: A normal cell that returns a feature map of the same dimension and reduced cell that returns a feature map where the height and width of the said feature map is reduced by a factor. Categorical loss was implemented in our paper For the case of NasNet instead of focal loss. And instead of using Adam optimizer as an optimizer, we used Adamax. After the first 25 epochs, normal class accuracy was a little less; the training accuracy amounted to 0.9222 and training loss of 0.2049. After 50 epochs, training accuracy improved to 0.9457 and loss fell to 0.1492 with improvement of f1-score of all classes being above 92%. At 70 epochs, after activating the non-trainable parameters and running on 89,065,813, the model achieved a training accuracy of 95.11% and a testing accuracy of 94.42%.

Epoch		precision	recall	f1-score	accuracy
25	COVID-19	0.94	0.87	0.90	91.58%
	Normal	0.82	0.94	0.87	
	Viral Pneumonia	0.97	0.90	0.93	
50	COVID-19	0.97	0.93	0.95	93.00%
	Normal	0.89	0.96	0.92	
	Viral Pneumonia	0.97	0.94	0.96	
70	COVID-19	0.95	0.93	0.94	94.42%
	Normal	0.89	0.94	0.91	
	Viral Pneumonia	0.96	0.93	0.95	

Table 4.3: NasNetLarge

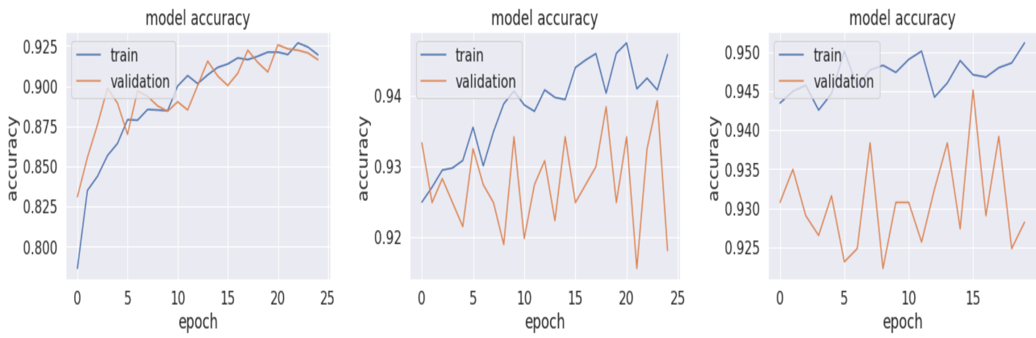


Figure 4.4: NasNetLarge Accuracy after 25, 50, 70 Epochs

4.2.4 Xception

The Xception is a CNN architecture with 71 layers and is an extension of the Inception model proposed by Francois Chollet in [1]. Xception is known to outperform Inception v3 on the ImageNet dataset.This architecture reestablishes the inception module with depthwise separable convolutions operations, in which the convolutions are not only in a depthwise manner but also as a pointwise one. It has 22,970,923 parameters in total, among which 2,105,347 were trainable and consists of depth-wise convolution layers which are independent instead of the conventional convolution layers. It takes into account the mapping of spatial correlations and cross-channel correlations which can be decoupled in CNN feature maps in their entirety. Another approach to utilize a pre-trained model is to train not only a new classifier but also fine-tune higher convolutional layers of the pre-trained model that are responsible for significant feature extraction.

For the first 25 epochs, the model was successful in achieving 0.9084 training accuracy and 1.6731 training loss.The model was initialized with Imagenet training weights. The accuracy improved to 0.9409 after the second 25 epochs. Non-trainable 20,865,576 parameters were made trainable and a training accuracy of 94.92% was obtained and 93.67% accuracy on testing at 70 epochs.

Epoch		precision	recall	f1-score	accuracy
25	COVID-19	0.92	0.90	0.91	90.33%
	Normal	0.84	0.90	0.87	
	Viral Pneumonia	0.95	0.91	0.93	
50	COVID-19	0.95	0.92	0.93	92.25%
	Normal	0.86	0.96	0.91	
	Viral Pneumonia	0.99	0.91	0.95	
70	COVID-19	0.96	0.91	0.93	93.67%
	Normal	0.88	0.94	0.91	
	Viral Pneumonia	0.96	0.94	0.95	

Table 4.4: Xception

Figure 4.5: Xception Accuracy after 25, 50, 70 Epochs

4.2.5 VGG19

VGG19 is a CNN architecture that is a descendant of VGG-16 with 19 weight layers (16 convolutional and 3 dense) and is used as a pre-processing model. Compared with traditional CNNs, it has been improved in network depth. It utilizes a substituting structure of different convolutional layers and non-linear activation layers.

VGG19 has 20,554,819 parameters which includes 529,411 trainable parameters.Hence, the network has learned rich feature representations for a wide range of images.The training accuracy and loss accuracy is 0.8906 and 1.9317 respectively after 25 epochs and 0.9220 and 1.4485 after 50 epochs. After unfreezing the layers, activating the remaining 20,025,408 parameters and training for 70 epochs, the VGG19 model achieved a training accuracy of 92.73% and testing accuracy of 91.92%. Even though

VGG19 takes time to learn, they are utilized in image classifications because of their good accuracy results.

Epoch		precision	recall	f1-score	accuracy
25	COVID-19	0.90	0.92	0.91	89.83%
	Normal	0.86	0.92	0.89	
	Viral Pneumonia	0.96	0.88	0.92	
50	COVID-19	0.91	0.95	0.93	93.25%
	Normal	0.90	0.92	0.91	
	Viral Pneumonia	0.98	0.93	0.95	
70	COVID-19	0.90	0.92	0.91	91.92%
	Normal	0.87	0.92	0.89	
	Viral Pneumonia	0.98	0.91	0.94	

Table 4.5: VGG19

Figure 4.6: VGG19 Accuracy after 25, 50, 70 Epochs

4.3 Transfer Learning

In our experiment, we also implemented transfer learning on these models using ImageNet weights. Transfer learning is a widespread Machine Learning technique which presumes utilizing an prevailing, trained Neural Network, that has been engineered for one task, as the core foundation for another task. Transfer learning is favoured as it removes the necessity of training vast amounts of data for completing a task since the basic features required to train a model are imported from previously accomplished analyses. The most prominent challenge associated with transfer learning is to retain the existing knowledge in the model while adapting the model to new tasks as it leads to the problem of the number of layers or parameters required to be re-trained to achieve optimal results. The primary steps of transfer learning involves finding the sustainable pre-trained model, secondly, replacing the ultimate layer of the model consistent with the amount of output layers for the upcoming task and eventually, resume training the model with fresh data and fine-tuning the model till the accuracy converges towards a higher and acceptable value. To begin with, the models were initialised with pre-trained Imagenet weights. For the first 50 epochs, we freezed the feature extraction layers of the model meaning that the trainable weights will not be updated. We kept the batch normalization layer on inference mode and trained the classifier. During inference mode, the layer normalizes the current batch using a moving average of the mean and standard deviation, rather than using the mean and variance of the current batch. The moving mean and moving variance are non-trainable variables that are updated each time the layer is called in training mode. For the next 20 epochs, we unfreezed the feature extraction layers allowing the weights to be updated and fine tuned the upper convolutional layers. The batch normalization layer was switched to training mode during which the layer normalizes the current batch using the mean and variance of the current batch of inputs.

4.4 Ensembling

There are several ways to perform ensembling on the trained model. The methods include linear averaging, bagging, boosting, max voting etc.The Ensemble model has two types of averaging results from the base learners - Linear average and Weighted average. We implemented Ensembling of models which is a standard approach in Applied Machine Learning to make sure that the foremost stable and absolute best prediction is formed. Generally, ensemble learning involves training quite one network on an equivalent dataset, then using each of the trained models to form a prediction before combining the predictions in some way to configure a final outcome or prediction.
After taking into account all the test predictions of the 5 models used, we implemented a max voting system and a linear averaging system. A max voting system is where each of the multiple models used will predict and vote for a particular class. The image will be then classified as the class with the maximum number of votes since most of the models predicted the image as that corresponding class. Linear averaging was achieved by taking the average of the possibilities predicted by the individual models. Comparing between max voting and averaging, max voting gave better results. Results show that the f1 score for all the 3 classes are all good, especially for COVID-19 which has the highest f1 score.

Chapter 5

Results

The performance of each model was evaluated based on the precision, recall and f-1 score metrics, as shown in the previous section. The training and testing accuracy can be seen in the Table 5.1.

Models	Training Accuracy	Testing Accuracy
InceptionResNetV2	94.98%	92.75%
DenseNet201	96.53%	94.83%
NasNetLarge	95.11%	94.42%
Xception	94.92%	93.67%
VGG19	92.73%	91.92%

Table 5.1: Final Accuracy after 70 Epochs

From the confusion matrices of the five classification models after a total of 70 epochs shown below, we can observe that DenseNet201 has identified 379 COVID images correctly and identified 16 as Normal cases and 5 as Pneumonia. Densenet also classified 378 Normal classes and 373 Viral Pneumonia correctly.

On the other hand, InceptionResNetV2 classified 376 Viral Pneumonia cases without fail, which is an increment compared to DenseNet201, while the detection of other classes fall a little behind. VGG19 and NasNetLarge have performed similarly to DenseNet and InceptionResNetV2, however, NasNetLarge detected 22 COVID images as Normal and 25 Viral Pneumonia images wrongly as Normal. VGG19 detected 368 COVID, 367 Normal and 362 Pneumonia images correctly, mistaking 29 COVID images as Normal. Lastly, the Xception model has identified 377 images correctly in both the Normal and Viral Pneumonia classes, with 363 correctly identified COVID-19 images and falling a little back with identifying 33 COVID images as Normal images. Therefore we can conclude that DenseNet201 has out-performed all the other classifiers in terms of both correct class detection and f-1 scores for all the classes: 0.95 for COVID, 0.93 for Normal cases and 0.95 for Viral Pneumonia. Also, we have observed that all of the models have the lowest f1-score for Normal images among the three disease classes. Low image quality and not enough pre-processing might have affected the results.

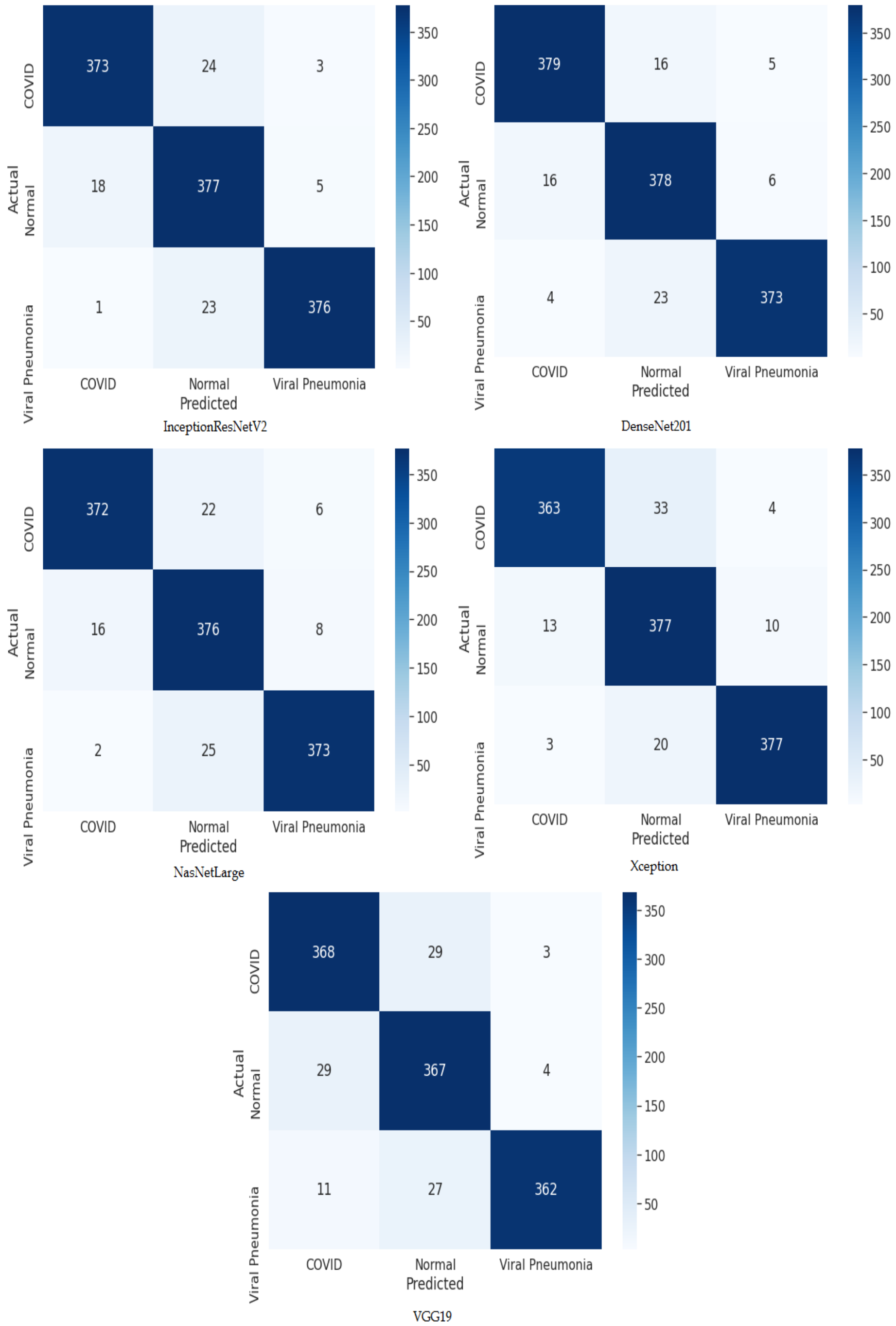


Figure 5.1: Confusion Matrices after 70 Epochs

5.1 Max Voting and Ensemble Linear Average

Max voting is a very commonly used classification scheme where the predictions from the classification models are votes and the majority of the votes are considered as the final prediction. Max Voting identified 386 COVID images accurately and identified only 1 image wrongly as Pneumonia, which is by far the best and most accurate. Furthermore it has been successful in predicting 388 Normal images and 381 Viral Pneumonia images without fail and only mistook one Pneumonia image for a COVID case. The overall performance of the max voting system was outstanding, attaining an accuracy of 96.25%. The f-1 scores for COVID-19, Normal and Viral Pneumonia classes were also very high at 97%, 95% and 97% respectively.

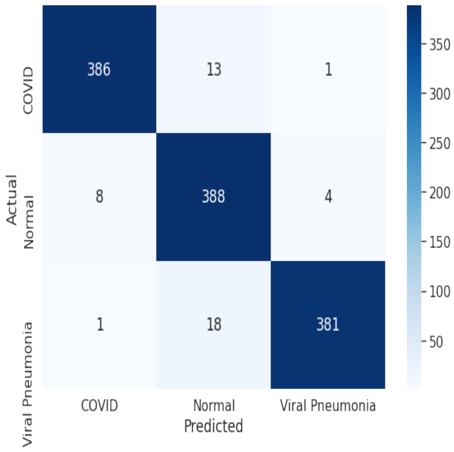


Figure 5.2: Max Voting Confusion Matrix

Alongside max voting, we implemented Ensemble Linear Averaging for final prediction, comprising the prediction from all five of our models to compare with Max Voting results. The accuracy of linear average is at 89.33%, with f-1 score of COVID-19, Normal and Pneumonia at 92%, 84% and 92% respectively, making Max Voting results a clear winner.

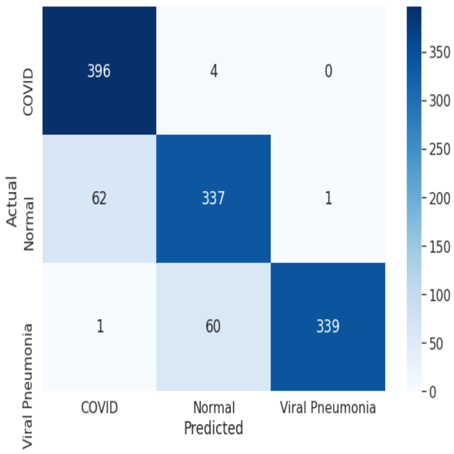


Figure 5.3: Ensembling Linear Average Confusion Matrix

5.2 GradDCAM Results

In order to find out about the COVID-19 detection transparency, we have used the idea of Gradient Class Activation Map (Grad-CAM) based color visualization approach for identifying the regions where the model paid more attention during the classification. The procedure of Grad-CAM provides a visual interpretation for any deeply related neural network and aids with verifying where the model is looking at while predicting. It also allows us to verify whether the model is activating at the correct locations and how well is it actually performing. We have implemented Grad-CAM using Keras and Tensorflow. DenseNet was selected as the model to be used with Grad-CAM because it has the highest average precision, recall and f1-score among the other models and we expected that it give the best results for activation maps as well. Grad-CAM works by taking an image as an input and computes a heat-map by examining the gradient information flowing into the last Convolutional layer or a specific layer of the model. We have selected Conv5_Block32_Concat layer of the DenseNet model to visualize heat-maps. Figure 5.4 demonstrates some sample GradCAM images below.

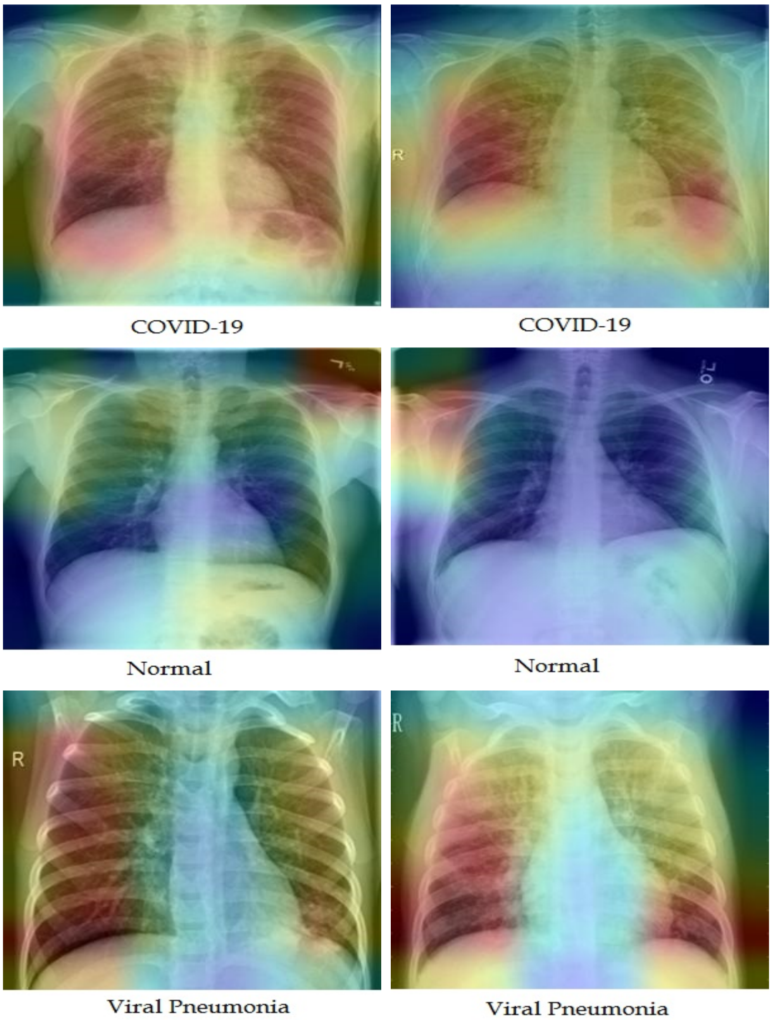


Figure 5.4: GRADCam Images Results of COVID-19, Viral Pneumonia and Normal Patients

Chapter 6

Discussions and Conclusion

6.1 Performance Comparison

While in [21] the authors opt out for binary classification (COVID-19 and Non-COVID-19) and achieved a performance accuracy of 99%, our model is a ternary classification, classifying among Viral Pneumonia, COVID-19 and Normal cases and we achieved an accuracy of 96.25%. Our experiment is different from their approach in terms of the number of classes and the classifiers, wherein they implemented VGG19, ResNet34, ResNet50, MobileNetV2 and DenseNet201.

A deep CNN based solution using Ensemble learning modelled by the authors in [22] to perform a binary classification between COVID-19 and Non-COVID cases had 538 images of COVID positive patients and 468 of negative patients. Three pre-trained models- DenseNet, ResNet50V2 and InceptionV3 were applied. Their approach showed an overall classification accuracy of 95.7% while ours had an accuracy of 96.25% with five pre-trained models and a ternary classification.

In another related study, the authors of [25] proposed a model comprising a 2-stage transfer learning training process and an ensemble learning method. They implemented six pre-trained CNNs - VGG16, ResNet50 , ResNet50-2, DenseNet161, DenseNet169and InceptionV3. 746 CT scan images, inclusive of 349 COVID-19 and 397 Normal cases were used. The model achieved an accuracy of 86.70%, implying our model, with 5 classifiers gravely surpasses said model in terms of accuracy. Moreover it can also be observed that an ensemble model has a better classification accuracy compared to existing models with one or multiple classifiers. In [14] Apostolopoulos et al. successfully obtained an accuracy of 93.48% for a 3-class classification, but falls behind when compared to our Ensemble approach, further proving our point.

6.2 Limitations

The lack of computer resources was one of the limitations that we had to face, i.e use of cloud computing or distributed learning. The training time could have been reduced. In-depth analysis would have been achievable had we obtained more datasets, which can be a possible extension to our study once more patient data (both symptomatic and asymptomatic patients) becomes available. The perennial pandemic and the lockdown hindered us in getting medical images from hospitals, thus having to rely on public repositories.

There are several scopes of bringing improvement to our work. For example, testing more feature extraction models and combinations of classifier networks are to name a few. Furthermore, We could have implemented segmentation to the Chest X-Rays. Moreover this approach can also be implemented by incorporating a larger dataset to attain a better predictive performance. Some of the adversities faced during experiments were the lack of annotated medical images and classified datasets. Also, more image pre-processing techniques can be applied for better results.

6.3 Future Prospects

Future prospects may include formulating new architectures based on CNN for the detection of COVID-19 alongside other diseases in the medical domain. The aforementioned models can be deployed in Web and Mobile applications, where patients can self diagnose their ailments at their ease, thus saving valuable seconds in dire time. Such applications can also be extended towards hospital IT systems where patients can receive budget-friendly and quick COVID-19 diagnosis alongside the in-action RT-PCR tests. Future directions include to extend the proposed model to risk stratification for survival analysis, anticipating risk status of patients, and predicting hospitalization duration which would be valuable for triaging, patient population management, and individualized care planning.

6.4 Conclusion

Detection of the infamous Coronavirus is more important now than ever with the second wave supported by the ever-evolving nature of the virus variants. As our contribution towards faster diagnosis to curb cases, we propose an ensemble model using five feature extraction state-of-the-art CNN models, training on 2492 COVID images, 1675 Pneumonia images and 2496 Normal images. The testing set consisted of 400 images of each class. Deep learning based recommender systems can be of great help in this scenario when the volume of patients is very high and required radiological expertise is low. Detection of chest diseases from X-ray images is in itself a challenging task thus requires consideration from the research industry. Transfer learning plays a major role in improving the accuracy of detection. Our results prove that an ensemble model surpasses an individual classification model, attaining an accuracy of 96.25% and greater f-1 scores for all the classes.

As the number of patients are increasing all around the world and the symptoms and development of the virus are changing gradually, with the continuous collection of data, we intend to extend the experiment further and upgrade the usability of the model. Our methodology achieved promising outcomes on the assembled dataset and we believe it can be beneficial for radiologists and health experts to gain deeper understandings into critical aspects related to COVID-19 cases.

Such a technique can be sent in far off areas that are blocked off to help analyze respiratory illnesses and save lives. If COVID data were readily available and better documented and annotated it could bear the potential to open several pathways for more data-driven studies in the future. With all that being said, we would also like to thank specialists, medical attendants and all the medical care suppliers who are placing their lives in the front lines to battle the COVID-19 outbreak.

Bibliography

[1] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1251–1258.

[2] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 4700–4708.

[3] C. Tataru, D. Yi, A. Shenoyas, and A. Ma, "Deep learning for abnormality detection in chest x-ray images," in *IEEE Conference on Deep Learning*, 2017.

[4] X. Wang, Y. Peng, L. Lu, Z. Lu, M. Bagheri, and R. M. Summers, "Chestx-ray8: Hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 2097–2106.

[5] Y. Gordienko, P. Gang, J. Hui, W. Zeng, Y. Kochura, O. Alienin, O. Rokovyi, and S. Stirenko, "Deep learning with lung segmentation and bone shadow exclusion techniques for chest x-ray analysis of lung cancer," in *International Conference on Computer Science, Engineering and Education Applications*, Springer, 2018, pp. 638–647.

[6] B. Pardamean, T. W. Cenggoro, R. Rahutomo, A. Budiarto, and E. K. Karuppiah, "Transfer learning from chest x-ray pre-trained convolutional neural network for learning mammogram data," *Procedia Computer Science*, vol. 135, pp. 400–407, 2018.

[7] S. Stirenko, Y. Kochura, O. Alienin, O. Rokovyi, Y. Gordienko, P. Gang, and W. Zeng, "Chest x-ray analysis of tuberculosis by deep learning with segmentation and augmentation," in *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, IEEE, 2018, pp. 422–428.

[8] V. Bhagat and S. Bhaumik, "Data augmentation using generative adversarial networks for pneumonia classification in chest xrays," in *2019 Fifth International Conference on Image Information Processing (ICIIP)*, IEEE, 2019, pp. 574–579.

[9] J. Irvin, P. Rajpurkar, M. Ko, Y. Yu, S. Ciurea-Ilcus, C. Chute, H. Marklund, B. Haghgoo, R. Ball, K. Shpanskaya, *et al.*, "Chexpert: A large chest radiograph dataset with uncertainty labels and expert comparison," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, 2019, pp. 590–597.

[10] E. Kesim, Z. Dokur, and T. Olmez, "X-ray chest image classification by a small-sized convolutional neural network," in *2019 scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT)*, IEEE, 2019, pp. 1–5.

[11] B. Wang, Z. Wu, Z. U. Khan, C. Liu, and M. Zhu, "Deep convolutional neural network with segmentation techniques for chest x-ray analysis," in *2019 14th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, IEEE, 2019, pp. 1212–1216.

[12] S. Ahmed, T. Hossain, O. B. Hoque, S. Sarker, S. Rahman, and F. M. Shah, "Automated covid-19 detection from chest x-ray images: A high resolution network (hrnet) approach," Sep. 2020. DOI: 10.1101/2020.08.26.20182311.

[13] S. Albahli, "Efficient gan-based chest radiographs (cxr) augmentation to diagnose coronavirus disease pneumonia," *International journal of medical sciences*, vol. 17, no. 10, p. 1439, 2020.

[14] 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, Apr. 2020, ISSN: 2662-4737. DOI: 10.1007/s13246-020-00865-4. [Online]. Available: http://dx.doi.org/10.1007/s13246-020-00865-4.

[15] M. E. Chowdhury, T. Rahman, A. Khandakar, R. Mazhar, M. A. Kadir, Z. B. Mahbub, K. R. Islam, M. S. Khan, A. Iqbal, N. Al Emadi, *et al.*, "Can ai help in screening viral and covid-19 pneumonia?" *IEEE Access*, vol. 8, pp. 132 665–132 676, 2020.

[16] K. El Asnaoui and Y. Chawki, "Using x-ray images and deep learning for automated detection of coronavirus disease," *Journal of Biomolecular Structure and Dynamics*, pp. 1–12, 2020.

[17] M. Joseph, "Does imagenet pretraining work for chest radiography images(covid-19)?," Apr. 2020.

[18] A. I. Khan, J. L. Shah, and M. M. Bhat, "Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images," *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105 581, 2020.

[19] Y. Oh, S. Park, and J. C. Ye, "Deep learning covid-19 features on cxr using limited training data sets," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2688–2700, 2020. DOI: 10.1109/TMI.2020.2993291.

[20] R. Sethi, M. Mehrotra, and D. Sethi, "Deep learning based diagnosis recommendation for covid-19 using chest x-rays images," in *2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA)*, 2020, pp. 1–4. DOI: 10.1109/ICIRCA48905.2020.9183278.

[21] B. N. Wiysobunri, H. S. Erden, and B. U. Toreyin, "An ensemble deep learning system for the automatic detection of covid-19 in x-ray images.," 2020.

[22] 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.

[23] S. Hira, A. Bai, and S. Hira, “An automatic approach based on cnn architecture to detect covid-19 disease from chest x-ray images,” *Applied Intelligence*, vol. 51, no. 5, pp. 2864–2889, 2021.

[24] T. Rahman, A. Khandakar, Y. Qiblawey, A. Tahir, S. Kiranyaz, S. B. A. Kashem, M. T. Islam, S. Al Maadeed, S. M. Zughaier, M. S. Khan, *et al.*, “Exploring the effect of image enhancement techniques on covid-19 detection using chest x-ray images,” *Computers in biology and medicine*, vol. 132, p. 104 319, 2021.

[25] J. F. H. Santa Cruz, “An ensemble approach for multi-stage transfer learning models for covid-19 detection from chest ct scans,” *Intelligence-Based Medicine*, vol. 5, p. 100 027, 2021.