Using Chest Scans for the Diagnosis of COVID-19

By

Reem Fida Muhammad

274784-MS(CS)-2018

Supervisor:
Dr. Khawar Khurshid

School of Electrical Engineering and Computer Science

NATIONAL UNIVERSITY OF SCIENCES AND TECHNOLOGY, ISLAMABAD

DECEMBER 2020

Approval

It is certified that the contents and form of the thesis entitled "Using chest scans for the diagnosis of Covid-19" submitted by REEM MUHAMMAD have been found satisfactory for the requirement of the degree

Advisor: Dr. Khawar Khurshid

Signature: ____________________

Date: **03-Dec-2020**

Committee Member 1:Dr. Salman Abdul Ghafoor

Signature: ____________________

Date: **03-Dec-2020**

Committee Member 2:Dr. Ahmad Salman

Signature: ____________________

Date: **03-Dec-2020**

Signature: ____________________

Date: ____________________

THESIS ACCEPTANCE CERTIFICATE

Certified that final copy of MS/MPhil thesis entitled "Using chest scans for the diagnosis of Covid-19" written by REEM MUHAMMAD, (Registration No 00000274784), of SEECS has been vetted by the undersigned, found complete in all respects as per NUST Statutes/Regulations, is free of plagiarism, errors and mistakes and is accepted as partial fulfillment for award of MS/M Phil degree. It is further certified that necessary amendments as pointed out by GEC members of the scholar have also been incorporated in the said thesis.

Signature: ____________

Name of Advisor: Dr. Khawar Khurshid

Date: 03-Dec-2020

Signature (HOD): ____________

Date: ____________

Signature (Dean/Principal): ____________

Date: ____________

Dedication

The whole work is dedicated to my parents, siblings and friends. Their support and motivation helped me to achieve this accomplishment

Certificate of Originality

I hereby declare that this submission titled "Using chest scans for the diagnosis of Covid-19" is my own work. To the best of my knowledge it contains no materials previously published or written by another person, nor material which to a substantial extent has been accepted for the award of any degree or diploma at NUST SEECS or at any other educational institute, except where due acknowledgement has been made in the thesis. Any contribution made to the research by others, with whom I have worked at NUST SEECS or elsewhere, is explicitly acknowledged in the thesis. I also declare that the intellectual content of this thesis is the product of my own work, except for the assistance from others in the project's design and conception or in style, presentation and linguistics, which has been acknowledged. I also verified the originality of contents through plagiarism software.

Student Name: REEM MUHAMMAD

Student Signature: _______________

Acknowledgement

I was able to accomplish my work only because of the guidance from Allah. I am also very grateful to many individuals, without their help I may not be able to achieve this task. I am thankful to my supervisor Dr Khawar Khurshid for helping me during my work with his support and wisdom.

I would like to show my respect and love to my parents, who were always there to support me in a difficult time and I would also like to express my gratefulness towards my siblings and friends for their knowledge and advice, which I needed the most to complete my work.

Table of Contents

List of Figures

List of Tables

Abstract

COVID-19 is a viral infection that affects the human respiratory system. Since the victims of COVID-19 are increasing globally, this contagious disease is characterized as a pandemic by the World Health Organization (WHO). Laboratory testing is not considered effective for COVID-19 patients due to an increase in their absolute number. In this study, we proposed XCov-Net which is a deep learning model that can diagnose patients suffering from corona using chest X-ray scans. XCov-Net model is formed by the combination of the Xception model and XCov block. The proposed model performs multiclass and binary classification. The accuracy obtained in multiclass classification is 0.948 and 0.987 for four classes (COVID-19 cases, Normal cases, cases of Pneumonia caused by bacteria, cases of Pneumonia caused by a virus) and three classes (COVID-19 cases, Normal cases, Pneumonia cases) respectively. For two class classification, the model performed outstandingly good with accuracy recorded as 1.

Chapter 1

Introduction

The novel Coronavirus, also known as COVID-19, is a contagious disease that has affected the entire world. It was recently found in China which has spread nearly into every country. To present, Coronavirus cases are still soaring around the world with well over 4 million people infected due to which it has been declared as a pandemic by WHO[1]. Around 937,391 people had lost their lives to COVID-19. Some countries managed to gain control over the virus but others are still struggling.

COVID-19 belongs to a family of coronaviruses which is termed as RNA virus that infects human respiratory system. It is believed that the virus of such type has never occurred before among humans thus making them first of their kind. It is also believed that the virus may have passed on from animal to humans. This contagious disease spreads primarily through droplets generated when an infected person coughs, sneezes or speaks. The common symptoms are dry cough, fever, shortness in breath and tiredness. Some other less common symptoms include aches and pains, nasal congestion, headache, conjunctivitis, sore throat, diarrhea, loss of taste or smell, a rash on skin or discoloration of fingers or toes[2]. There is no proper cure for the disease. The only way for not getting infected is by maintaining social distance and washing your hands continuously before touching your mouth, eyes, or nose.

1.1 Problem Statement

Clinical testing called RT-PCR (real-time polymerase chain reaction) is the way to test a person for COVID-19 where the specimen for the test is collected from either nasopharyngeal swab or oropharyngeal swab. The drawback with this method is that kit for RT-PCR test is costly and takes a lot of time. The process of testing with RT-PCR, from sample collection to result declaration, is a time-consuming process keeping in view the pandemic. Even though the number of cases has been dropped, yet there are ongoing active cases. The process can take around 24 hours or even days[3]. Increasing number of COVID-19 patients may cause a shortage of kits, as a result, testing a patient for COVID-19 becomes a difficult task. It has also been found that RT-PCR may produce false-negative results[4].

If the person suffering from COVID-19 is reported as positive with in minimum amount of time will then be treated immediately by the health experts on the other hand waiting for the report for days may cause the spread of the disease quickly. To overcome the problem of Coronavirus, some other way is required to test the patients for COVID-19 which takes less time and produce accurate result.

1.2 Questions

The dissertation tries to answer the following questions:

1. How to perform COVID-19 diagnosis using proposed model?
2. How well the proposed model distinguishes between COVID-19 and Pneumonia?
3. Comparison of the proposed model with previous work.
4. How the proposed model will deal with the pandemic?
5. How health experts will be benefitted by the proposed model?

1.3 Goal of the Study

Researchers are working tirelessly to find ways to test and generate results in minimum time and cost. Researchers have found several alternative to RT-PCR test which includes LAMP, Antibody test, CT-scans, Antigen test, Chest X-ray scans etc. For Chest CT-scans or X-ray, Radiological imaging can be used to diagnose the disease by detecting lung opacities caused due to COVID-19[3]. Since is it difficult for the radiologists to analyse and diagnose the disease due to increased COVID-19 patients, a computer-aided system is required to detect and analyse the disease automatically. The main purpose of the study is to aid health experts in pandemic by providing them with the results within few seconds so that the health experts can spend more time in making decision depending on the severity of the disease.

1.4 Advantages and Areas of Application

1.4.1 Advantages

The advantages of the dissertation are mentioned below:

1. Improve the analysis of disease
2. Help in reducing misclassification error
3. Help in grading the disease
4. Help health experts to spend more time in decision making

1.4.2 Areas of Application

Areas of Application for the study are mentioned below:

1. Can be used in clinics
2. A source of education and research

Chapter 2

Review of the Literature

This Chapter discusses the AI based methods and how they are used in medical diagnosis. Many countries are still greatly effected with the novel Coronavirus while some other countries reported that the cases of the disease have been dropped but still the COVID-19 is not completely eliminated. Different methods to deal with COVID-19 have been described in the literature. Researchers are working tirelessly in finding an efficient way to test the patients and are also working effortlessly to find a way to cure the disease.

2.1 Background Concepts

In the past few decades, advancement in AI based approaches have resulted in increased demand of automatic applications in the medical field for the diagnosis of disease[5]. These AI applications helps in improving the diagnostic accuracy. Many such applications have been developed which are used mostly in clinics and hospitals at US.

AI applications are being extensively used in Image processing for detection, classification and segmentation of objects with in the images. Several such applications performs high level tasks such as predicting disease. Researchers mainly focuses in the development of the automatic systems for such tasks that are time consuming for the health experts to perform analysis, thus helping the health experts in better treatment of the patients. Pathologists, Radiologists are greatly benefitted from these AI based applications.

2.2 Machine Learning

With the increase in large amount of data, a tool is required to analyze these data and derive knowledge from the huge amount of data and Machine Learning can be used for this task. Machine learning is a part of AI that helps in the development of an automatic system in which the computers learn automatically about the given problem in hand without being explicitly programmed. Machine Learning can be used in all sort of complex tasks.

ML applications require hand crafted features to work. Feature selection is an important step in Machine Learning where important features are selected for model to learn and to converge smoothly. There are several ML applications which has been used in medical diagnosis. Recently many ML applications have been developed in drug industry where the system may identify the patients who are more likely to get benefit from the drugs. Machine Learning applications are divided into three main categories, Supervised, unsupervised and reinforcement learning ML algorithms.

Supervised Learning is way in which labels are given to the model and the model tries to find the association between input and output. The output from the supervised model can predict different categories where the model is actually trying to classify and this is known as 'Classification' where as if the output of the model is some scalar value then it is called 'Regression'. In Unsupervised learning, the model does not require labels to train. Reinforcement learning does not have fixed dataset, the model learns based on experience.

ML applications work on three basic steps, dataset, features and models. The dataset is the large amount of data from which the knowledge has to be derived. The features are part of these data that helps in learning process and model is the representation of the phenomenon that a ML application has learnt.

2.2.1 Classification

If the data has multiple classes or categories, the model will try to group similar data under one category by learning the relation between inputs and outputs. The model tries to create a decision boundaries among different classes. If the given data falls within the decision boundary then that class will be assigned to the input data.

2.3 Deep Learning

Deep Learning is a part of Machine Learning and has been extensively used because of the fact that it does not require feature selection but instead it learns the features from the data itself and helps in performing some complex tasks. They are widely being used because they are easier to work with and produces higher accuracy.

Deep learning uses Neural Network. It uses artificial neural network that tries to mimic human brain. Neural network is made up of layers. It basically consists of Input layer, Hidden layer and an Output layer. Deep learning learns the relation between the inputs and the outputs. These layers consists of ‘nodes’. The input layer takes in the data, the hidden layer performs some computation and the output layer give out the desired results. These hidden layers consists of some weights and biases which are updated in order to minimize the loss function so that the model may converge. Introducing many hidden layers in a neural network increases the depth of the neural network thus the word ‘Deep’ is used.

Deep learning models work well with massive amount of data. There are different types of neural networks based on different tasks. When dealing with Images, Convolutional Neural Networks (CNN) are used. Recurrent Neural Network (RNN) is used if to work on Natural Language Processing tasks etc. There are many applications of deep learning like in Robots, Video synthesis, Face recognition and even in Medical field.

2.3.1 Deep Learning in Medical Field

Many applications have been developed for medical diagnosis because of deep learning. Deep learning is helping the health experts in finding more accurate and efficient way of treating patients. These models helps in drug discovery by looking into the medical history of the patient thus helping in better treatment. It can also help to predict the fake claims of patient’s medical insurance. The pathological images have been digitized by using whole slide scanners thus are used in deep learning model. This has helped pathologists in analysis of difficult task such as detection of cancer. Radiological images such as CT, MRI are also being extensively used by these model which helps in assisting the radiologists.

2.4 Previous Work

Various works were on COVID-19 has been done using radiological imaging. Most of the work proposed a deep learning model architecture while others used pre-trained deep learning models. These automatic system helps in detecting and classifying Corona with in few seconds and had shown encouraging results.

Narin et.al.[6] achieved 98% accuracy using ResNet50 model[7]. ResNet is one of the state of the art model which had achieved an outstanding performance in classifying Imagenet dataset. Ozturk et al.[8] proposed a model called DarkCovidNet. The authors were inspired by the DarkNet[9] architecture which is considered as the base architecture of the YOLO algorithm[10] for classification. YOLO (you look only once) is the state of the art object detection algorithm. DarkCovidNet is the modification of DarkNet model. The DarkCovidNet performed binary and multiclass (three class) classification. It has achieved an accuracy of 98.08% and 87.02% on binary and multiclass classification. Yujin et al.[11] proposed a two steps methodology to deal with a limited amount of data. First, the authors performed segmentation using DenseNet[12] and the patches from these segments are passed to ResNet. For the second step, majority voting is performed for classification. This model achieved 88.9% accuracy on multiclass classification. Harsh et al.[13] proposed an architecture that performs binary classification where the authors modified VGG-16 model[14] and called it as nCOVnet. VGG is also one of the state of art model which performed outstandingly well on Imagenet dataset. The nCOVnet model achieved 88.10% accuracy.

Asif et al.[15] proposed a model called Coronet. In this model, multiclass classification (three and four classes) and binary classification were performed. The proposed model is based on Xception model[16]. Xception is an extreme version of Inception model which was developed by Google researchers. The accuracy of the model is 89.6%, 95% and 99% for Four-class, Three-class and Binary Classification. Other than that, Luca et al.[17] proposed two models and both model is based on VGG-16. The first model detects whether the given image is Normal or diseased where if the given images are diseased then the second model detects whether the disease is COVID-19 or another disease. The model performs a binary classification with the accuracy recorded for the first model is 96% and 98% for the second model. Rajarman et al.[18] introduced a method in which the authors trained different pre-trained models and a custom CNN. Three steps which included modality-specific transfer learning, model pruning and ensemble classification were used. The accuracy noted for multiclass (three) classification is 99.01%. Das et al[19] used a methodology that depends on the Xception model. It performs multiclass (three) classification which achieved an accuracy of 97.40%. There have been many works done using chest CT scans[20][21][22] and using lung ultrasound[23] with the help of deep learning method.

Chapter 3

Deep Learning Models

This Chapter discusses different state of the art DL models. These models are based on neural networks and are generally used when dealing with high dimensional data. This Chapters explains AlexNet, VGG, Resnet and Inception models which have achieved great performance on Imagenet dataset.

3.1 DL Models

Deep learning is an emerging field which consists of following basic models.

- Supervised models
- Unsupervised models

Supervised models have both input data and output labels. The model knows the labels and train on the given data. Some of the examples of supervised models are Multilayer Perceptron, Convolutional Neural Network (CNN). In Unsupervised models, labels are not provided to the models. The input is fed into the model and the model tries to make prediction without the labels. Supervised models can further performs two tasks, Classification and Regression.

3.2 Convolutional Neural Network (CNN)

When working with images or dealing with Computer Vision tasks, a DL model called Convolutional Neural Network (CNN) is used. Numerous CNN applications have been developed which helps in object detection, classification and even perform segmentation for example Human face detection, Vehicle identification etc. CNN has been extensively used in medical industry for disease diagnosis.

CNN takes input image, extract features from it and with the help of the features produce desirable output. It is made up of following three basic layer.

- Convolutional Layer: In this layer a filter kernel is convolved over the entire input image thus helping extract features. It results in some weights and bias which are then updated until the model converges.
- Pooling layer: This layer helps to reduce the size of the given input.
- Fully Connected layer: This helps to convert the feature maps into a single vector.

The convolutional layer consist of filter also known as mask/kernel/window which slides on the input that helps to extract specific features from it. The initial layers extract low level feature such as edges, corner and with the increment of more layers, later in the model more complex features are extracted. The input image is usually in the form of RGB, with three channels and the image is basically matrix of numbers. These numbers are the pixels in an image. The filter itself is also a matrix of numbers usually have odd spatial dimension such as (3 x 3) (for simplicity two dimension is considered). This filter is placed over the image and each pixel in the image is multiplied by the corresponding value in the filter and added at the end. The result is the new value in a grid cell like a pixel value in the output feature map. Then the filter slides on the image and the process continues until every pixel is covered in the image. The output of the resultant convolution is called a feature map.

There is also another concept of stride and padding during convolution. When convolution takes place the size of the input is reduced and edge pixels are lost but in order to avoid changing the spatial dimension of the given input and to avoid losing some data a method called padding is done and it has two main types, Reflection and Zero padding. In zero padding, around the edges of the image zero pixels are added thus increasing the dimension of the input where as in reflection padding, the values of the edges are just copied. Stride means how many pixels a filter will slide over. Usually filter with stride two is commonly used. The real world data is really very complex and highly dimensional. In order for the network to converge smoothly non linearity is introduced in the architecture and for this activation functions are used. The most common activation function which is used is ReLU.

There are several different CNN architecture which has performed remarkably excellent for different tasks related to images for example AlexNet, VGGNet, ResNet and GoogLeNet. All these architecture is made up of series of convolutional layer, pooling layer and fully connected layers.

3.2.1 AlexNet

AlexNet architecture was proposed in 2012 and had won ImageNet classification challenge in 2012. It was the first deep neural network for imagenet classification task. Figure 3-1 shows the architecture of AlexNet [24].

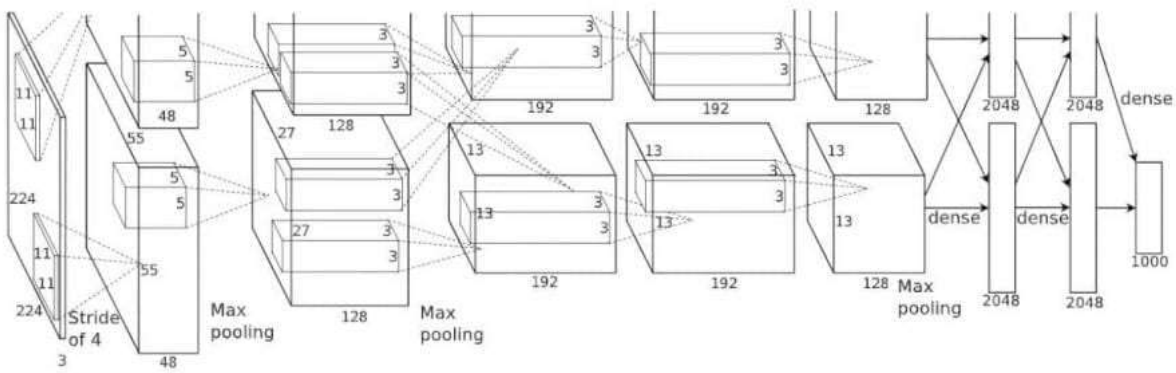

Figure 3-1: AlexNet

AlexNet consists of 8 layers excluding pooling layers. The type of pooling used is Maxpooling. It consists of 5 convolutional layers and 3 fully connected layers. It uses the receptive field of large size (11 x 11) and (5 x 5) in the earlier layers and size (3x3) in the later layers. Each convolutional layer is followed by a ReLU activation function and the last layer uses softmax function to classify 1000 classes. It can be seen in the figure 3-1 that the network is split into two pipeline, the reason is that the network was trained using two GPU's.

3.2.2 VGGNet

VGG architecture was developed by the Oxford Visual Geometry Group and is the state of the art model in 2014.VGG was the improvement of AlexNet architecture. Number of parameters in VGG has been reduced as compared to AlexNet due to the fact it used series of filter with small receptive field of size (3 x 3). On imagenet data, it has recorded 97.2% top 5 test accuracy on Classification task.

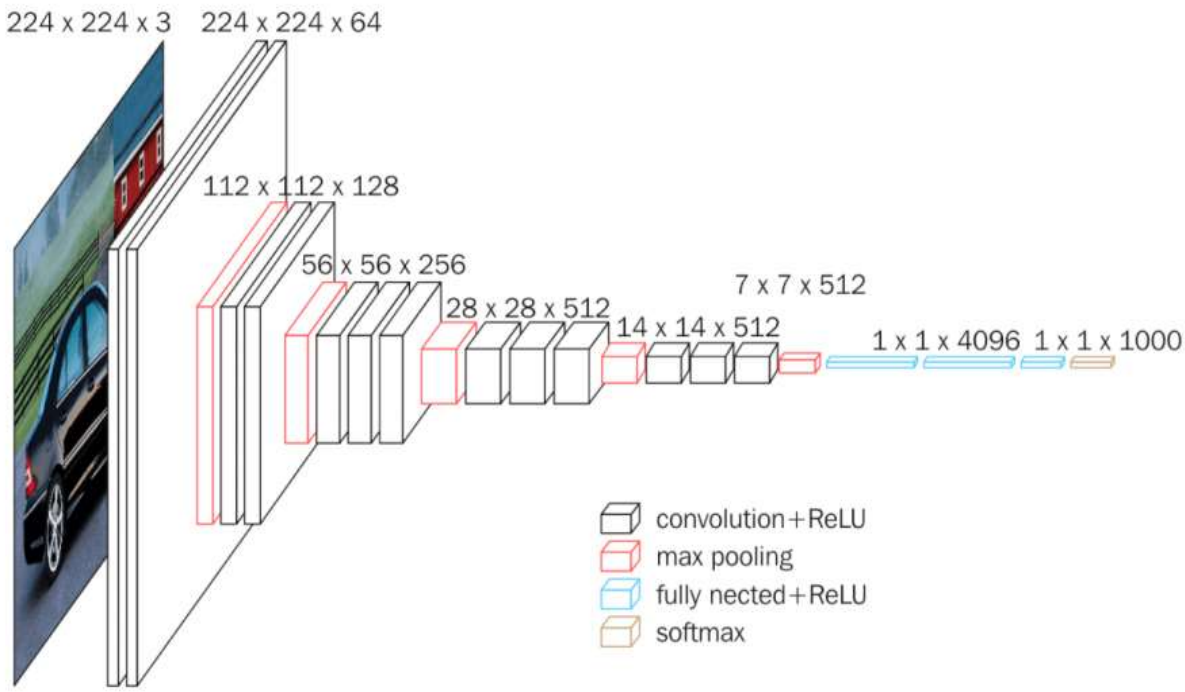

Figure 3-2: VGG model

VGG is actually a 16 layer architecture excluding the pooling layers and the last layer, thus also known as VGG16. It takes the RGB image with input size 256 x 256. Here, Max pooling is and small filter size of (3x3) is used. For nonlinearity ReLU function has been used here. With the increase in the depth of the network, it can be seen that the spatial dimension is decreasing but the depth of the input image is increasing. The reduction in the size of the spatial dimension is because pooling and increase in depth is because of the number of filters used in the convolutional layer. The last layer uses the softmax activation for the classification as the imagenet dataset consists of 1000 classes.VGG19 is a modification of VGG16 where the number of layers have been increased from 16 to 19. VGG is a very simple architecture which is most commonly used in simple classification tasks. Figure 3-2 shows the architecture of VGG model.

3.2.3 ResNet

Deep learning models means the network has many layers which causes the depth of the network to increase but with the increase in the depth the network faces some problems which results in a very bad network with decrease in accuracy. The problems that a deep network usually faces are

- **Exploding/Vanishing Gradient**: With the increase in the layers of the network while updating the weights the, the gradients becomes unstable. With continuous multiplication the value of the gradient may increase to infinitely large value or decrease to infinitely small value thus the weights do not get updated.
- **Network Degradation**: The other problem is that while going deep the performance of the model starts to go downhill. The performance becomes poor thus results in the decrease in the accuracy of the network.

To solve the above mentioned problems a network called Residual Neural Network (ResNet) was formed by the Microsoft.

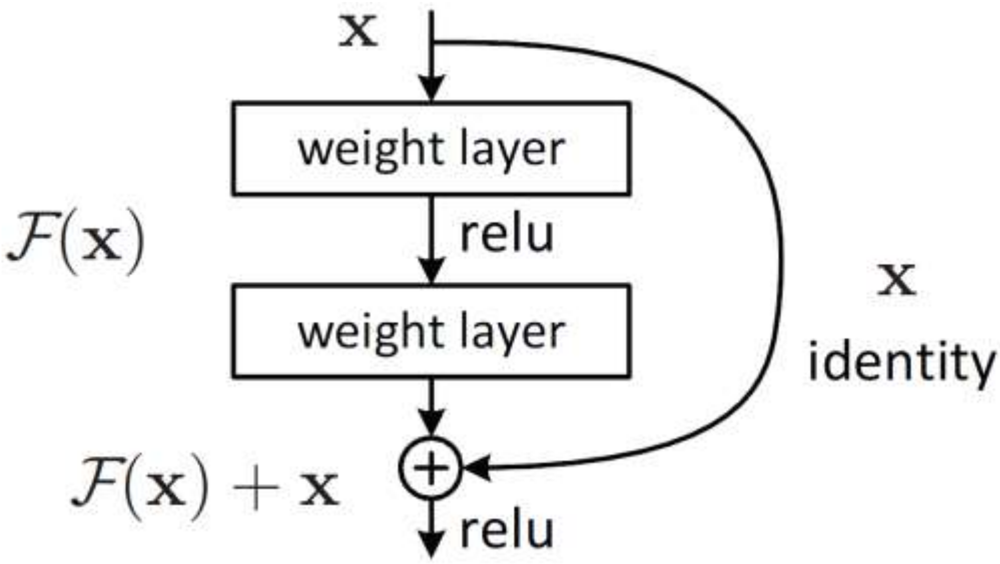

Figure 3-3: Residual Block

The main idea of the ResNet architecture is the use of skip connections after several layers which can be seen in the figure 3-3. The output from the previous layers can be added as it is after couple of next layers. So, with the increase in depth it will avoid the exploding /vanishing gradient problem and also helps to avoid degradation in the performance of the network. These residual blocks are just stacked together throughout the networks thus increasing the depth. Identity skip connection means that the input and the output have the same dimension. Different version of Resnet has been introduced for example resnet50, resnet101, resnet151. The difference among them is just the number of layers.

3.2.4 Inception

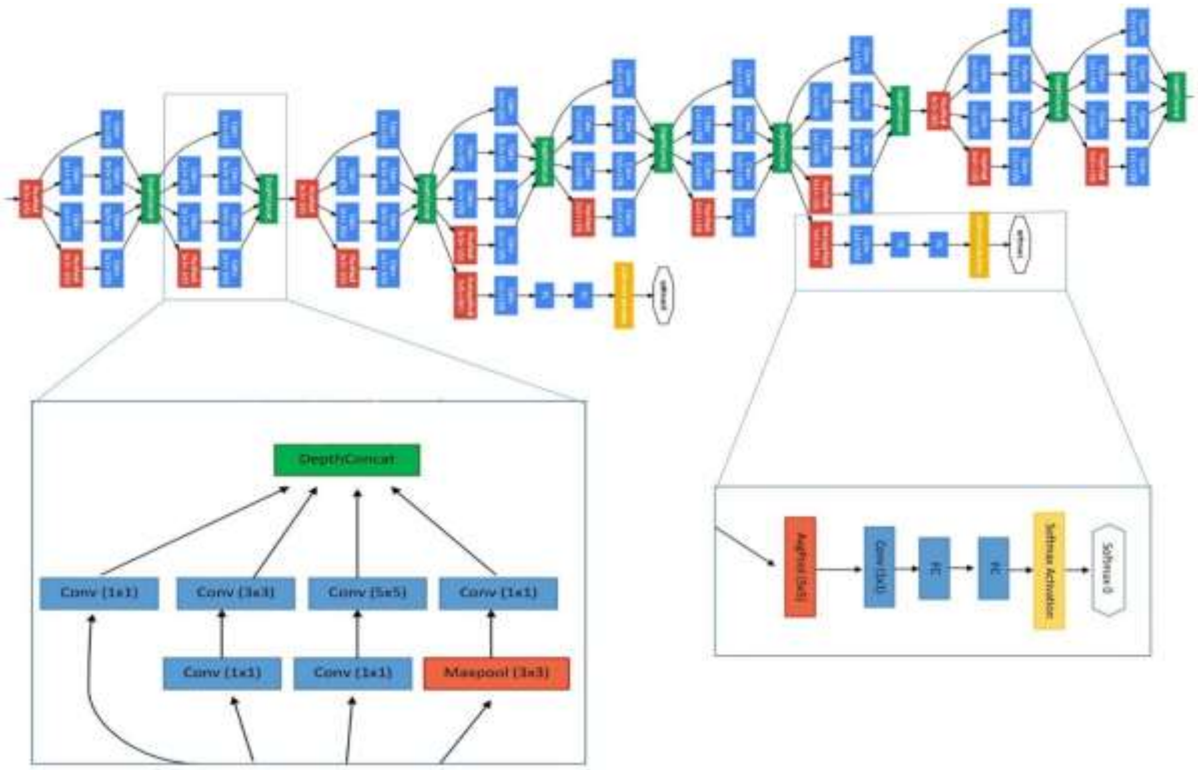

Figure 3-4: GoogLeNet

When working with CNN, it is always thought that going deeper will help in the performance of the network, many such models were formed which focuses on the depth of the network whereas Inception is more complex structure. Inception V1 or GoogLeNet[25] was developed by Google. The main idea is not just to focus on the depth of the network but also to focus on the given input data. It can be seen in different images that the scale of particular object may vary. Suppose an image consists of a bird picture, some images may have all area covered by the bird while some others may have all area covered by the background. In order to deal with this problem different filter size is required. For the data that is globally distributed large filter size is required and vice versa. This problem was resolved by the Google researchers and proposed a module called Inception. The main idea of the Inception module is to use different size of filters in parallel to the input image and then concatenating the output thus increasing the width of the network.

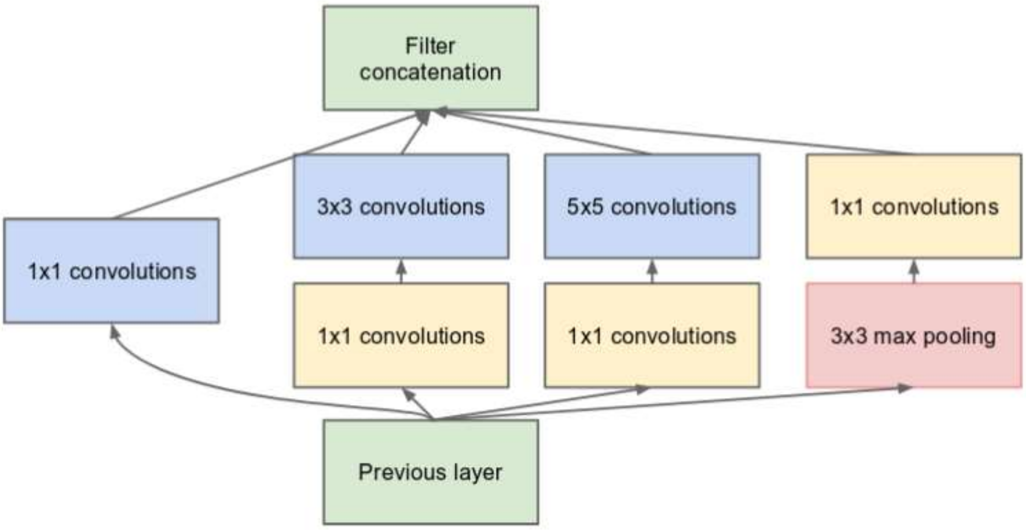

Figure 3-5: Inception V1 block

From the figure 3-5 it can be seen that (3x3), (5x5) and (1x1) filters are applied with (3x3) maxpooling in parallel and then the output from each convolution and pooling layer is then concatenated. Note that before (3x3) and (5x5) convolutions and after pooling, (1x1) convolution is applied. This helps to lessen the computational complexity. Stacking up Inception blocks results in the GoogLeNet as shown in figure 3-4.

GoogLeNet has 9 incpetion block and at the end it uses global average pooling layer. Several different modification have been done on the Inception block which results in Inception V2, Inception V3 and InceptionV4. GoogLeNet has reduced the computational complexity of the network without affecting the accuracy.

Chapter 4

Proposed Model

In this Chapter a brief introduction is given on Xception model and the proposed methodology is explained.

4.1 Motivation

The main contribution of the study is the formation of a deep learning model that can automatically classify patients suffering from corona with the help of chest X-ray images which, in return, will help to overcome the problem caused by RT-PCR and help in fighting against the pandemic by assisting the health experts in analysis of the disease.

4.2 Xception Model

In Chapter 3, different deep learning models have been explained in detail and one of the model which has been discussed is Inception V1. The idea of the Inception V1 model is to apply different filters in parallel thus increasing the width of the network. One of the problem that the Inception V1 model faces is that it is computationally very expensive.

To solve the above mentioned problem a model known as Extreme version of Inception called Xception has been introduced by the Google. It is a very efficient architecture and it does not depend on simple convolutions but rather the idea is to use depthwise separable convolutions. Depthwise separable convolution is depthwise convolution followed by a pointwise convolution.

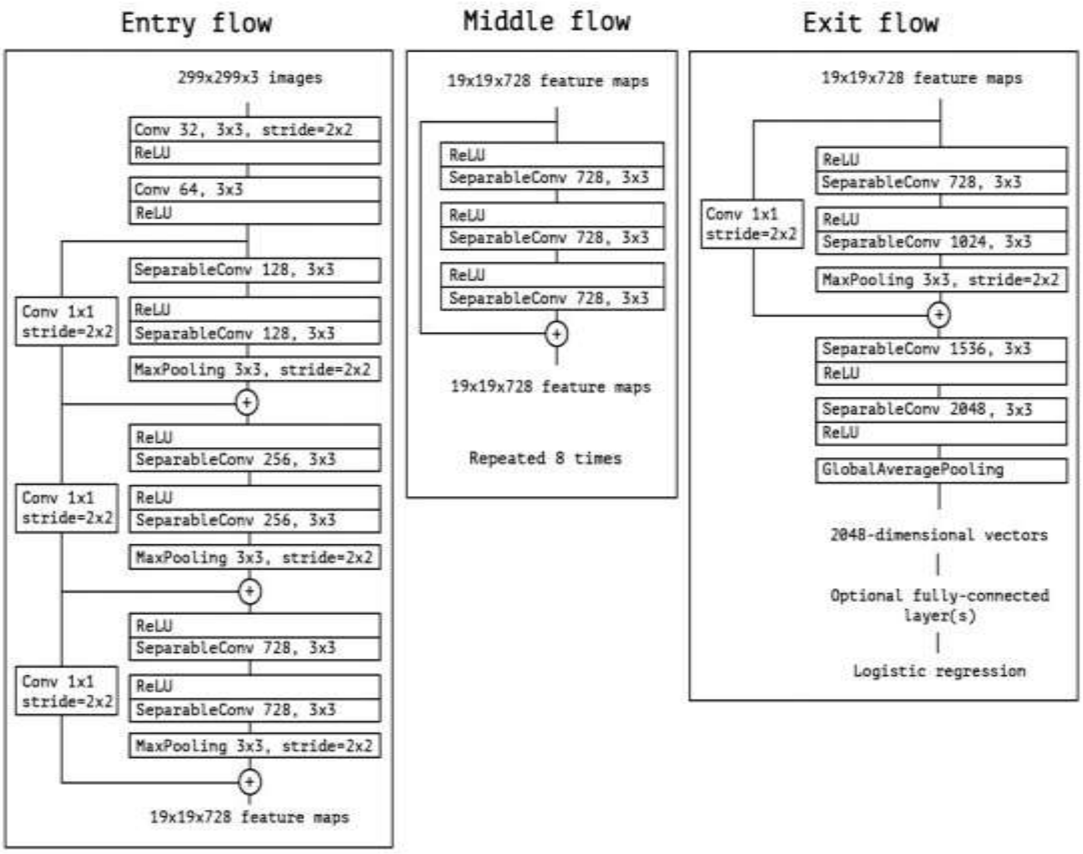

Figure 4-1: Xception Block

The input image is passed into the entry flow section and the into middle flow section. The middle flow is repeated eight times and finally it is passed to the Exit flow. The diagram is shown in figure 4-1. The convolutional and separable convolutional layers are followed by the batch normalization. The Xception model consists of two important concepts

- Depthwise separable convolutions
- Skip connections between convolutional blocks

Because of the depthwise separable convolutions the number of parameters are reduced and the model is far less computationally expensive than Inception models.

4.2.1 Simple Convolutions

In CNN, filters are convolved with the images which extract the features from the image. In the earlier layer, simple features are extracted and in later deeper layers complex features are extracted. The formula for applying a convolution is presented below

$$g\,(x; y; m) = k\,(x; y; c) * f\,(x; y; m) \qquad (1)$$

where g is the resultant filtered image. k is the input image and f is the filter.* represents convolutional operator. Height and width are represented by x and y. The number of channels in the input image k is c and m represents the total number of filters convolved. The depth of the filtered image depends on the number of filters used. The spatial dimension of the resultant image is the same as that of input image considering the input image has been padded. In case of simple convolution, the filter is convolved with the entire image, with all the channels at once. The cost of simple convolution is calculated as

$$g^2 \times f^2 \times m \times c \qquad (2)$$

Here g and f represents the size of output image and filter respectively. m is the number of filters and c represents total number of channels in the input image. This operation costs a lot of computation. With a deeper network, while using simple convolutional, equation (2) proves it to be computationally expensive.

4.2.2 Depthwise Convolutions

The problem associated with the simple convolution has been resolved by using depthwise separable convolution. Instead of applying a single filter to an entire image, each channel of the input image is convolved with the filters.

$$g\,(x; y; c) = k\,(x; y; c) * f\,(x; y) \qquad (3)$$

Here k is the input image and c is the number of channels of the input image. The total number of filters is equal to the total number of channels so there is c number of filters and each f filter is individually convolved with c channels. The spatial resolution remains the same as that of input image if keeping padding in consideration, whereas the depth of the output image will remain the same as that of input image.

4.2.3 Pointwise Convolutions

In order to increase or decrease the depth of image pointwise convolution is used.

$$g(x; y; m) = k(x; y; c) * f(1; 1; m) \qquad (4)$$

Here, the size of the filter f is 1 x 1 whereas m is the total number of 1 x 1 filters used which will be the depth of the resultant output. When we compare the resultant filtered image of both simple and depth wise convolution, both are the same but the number of operations has been reduced. Using such type of convolutions in deeper layers will help to reduce the computational complexity of the model which in turn will help to train the model faster.

The cost of computing depthwise separable convolution is

$$g^2 \text{ x } f^2 \text{ x } c + g^2 \text{ x } c \text{ x } m = c \text{ x } g^2 \text{ x } (f^2 + \text{m}) \qquad (5)$$

Comparing equation (2) with equation (5) it can be seen that equation (5) is computationally less expensive as compared to equation (2) thus proving the fact that with the help of depthwise separable convolution the performance of the Xception model is more efficient than the other Inception models. Figure 4-2 shows an example of depthwise separable convolution [26].

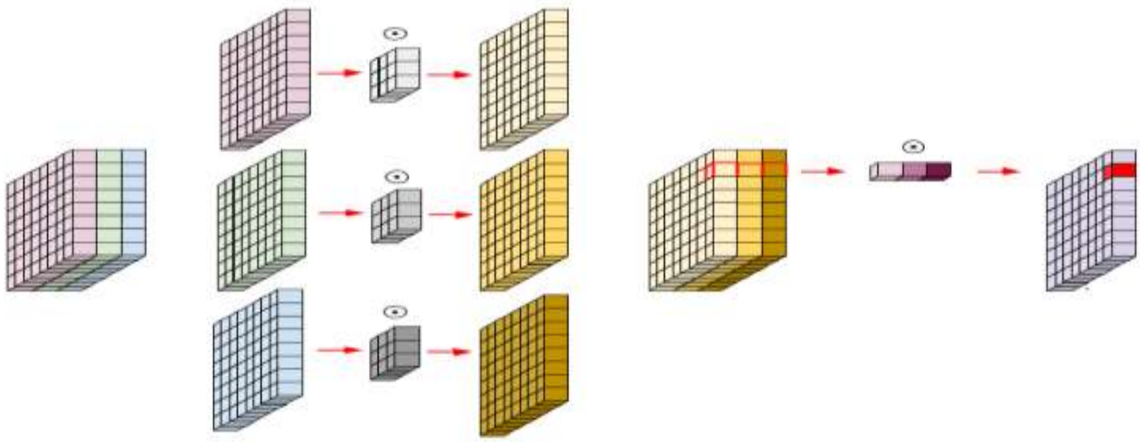

Figure 4-2: Depthwise separable convolution

4.3 Dataset

For the research purpose, we have used publically available data. COVID-19 chest X-ray images were obtained from the Github repository which is collected by Joseph .al[27]. As authors keep on updating their database, the database is a collection of X-ray and CT images of not only COVID-19 but of also other diseases as well. At the time of the study, there were 475 chest X-ray images of COVID-19 patients. For pneumonia and normal images, we used publically available Kaggle dataset[28]. Kaggle repository consists of total 5856 chest X-ray images. 1583 belongs to normal people and 4273 belongs to pneumonia patients. Pneumonia data consists of further two types, pneumonia bacterial and pneumonia viral. It can be seen that the number of pneumonia and normal images are greater than COVID-19 images and this can cause data imbalance issue. To balance the data, 500 images were randomly selected from both pneumonia and normal images. For experimental purpose, the data has been divided into 70% training, 20% validation and 10% testing. Some examples from the dataset are shown in figure 4-3.

Table 4-1: Dataset

Data	Number of images
COVID-19	475
Normal	500
Pneumonia viral	500
Pneumonia bacterial	500

4.4 The Proposed Model

The proposed model is known as XCov-Net which consists of the two parts

- Xception model
- XCov block

XCov-Net is a Convolutional Neural Network which performs classification. Figure 4-4 shows the architecture of the proposed model. The model helps to classify COVID-19 from pneumonia and normal image.

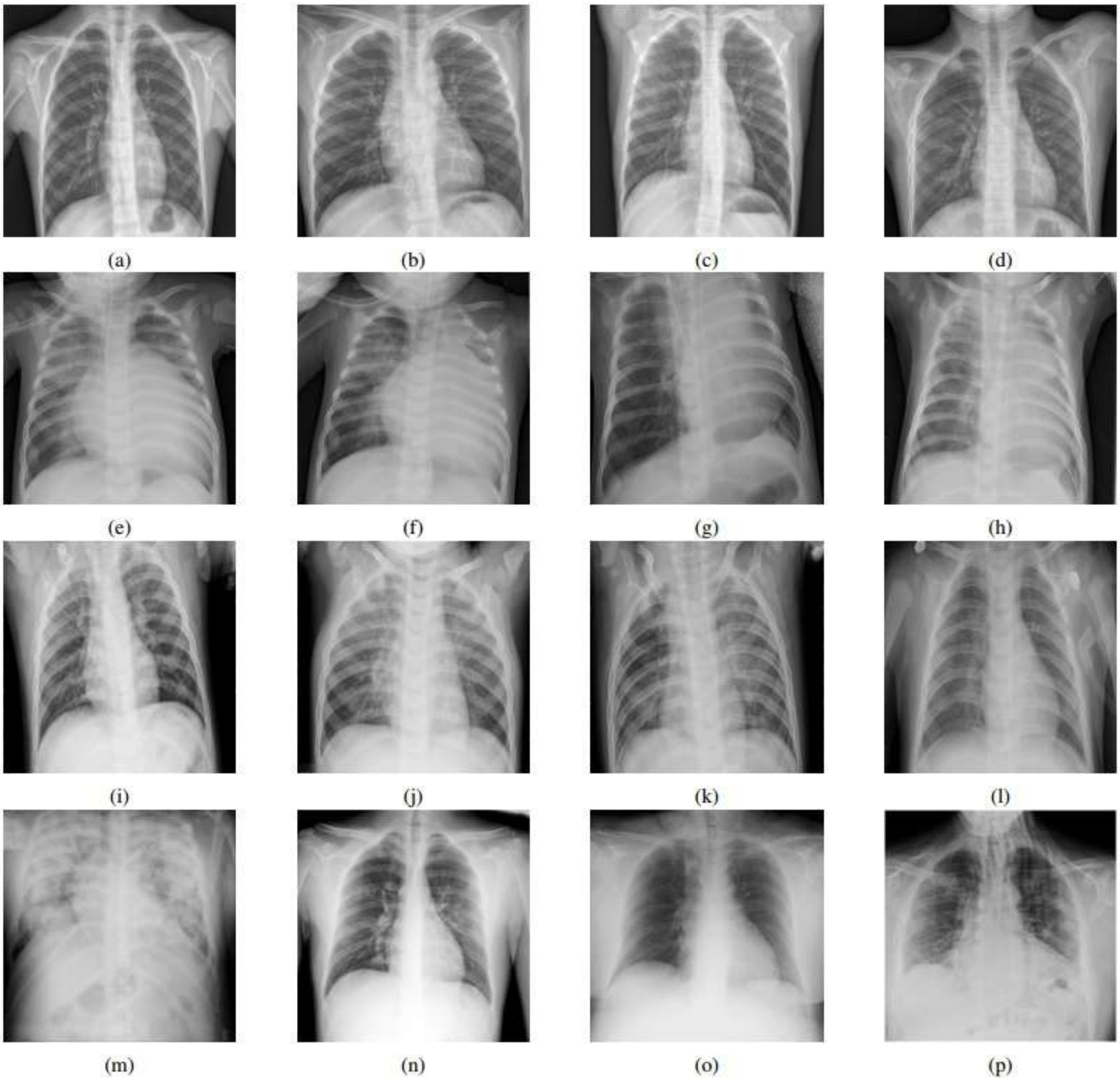

Figure 4-3: Sample of X-ray images from the dataset. (a)-(d) represents Normal X-ray images, (e)-(h) represents X-ray images of Pneumonia caused by bacteria, (i)-(l) shows the X-ray images of Pneumonia caused by virus and (m)-(p) shows COVID-19 X-ray images respectively.

Given input images, it can perform four-class classification, three class classification and binary classification. The proposed model depends on Xception model which is pre-trained on imagenet database[29]. The output of the Xception model is then fed to the XCov block which is then trained end-to-end. Passing the information obtained from the previous task to the new task with dataset different from the previous task is known as Transfer Learning[30].

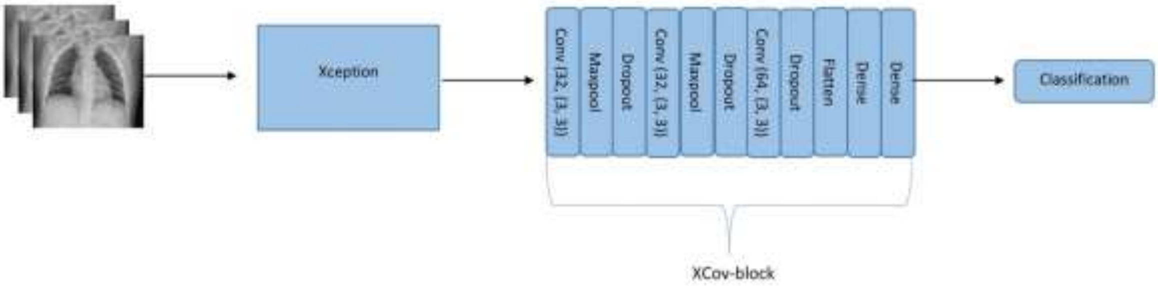

Figure 4-4: XCov-Net

4.4.1 Transfer Learning

When dealing with the dataset of small size, the deep learning model will fail to produce desirable results. Deep learning models require a large amount of data to train. The amount of COVID-19 images is not sufficient to train deep learning model. With an insufficient amount of data, the model may suffer from overfitting. In such cases, a method called Transfer Learning is used.

There are different databases which consist of millions of data, one such database is imagenet. Deep learning models are trained on these large databases and weights of such models are then used. The information obtained from the previous task can be used on the new task. The proposed method uses a pre-trained Xception model and this model was trained on imagenet dataset.

4.4.2 XCov-block

XCov block is a simple CNN which is shown in figure 4-4. It takes the output from the Xception model and performs classification. XCov-block consists of three convolutional layers followed by a dropout layer. Small receptive field of size (3x3) kernels are used in these convolutional layers. The images are resized to (256 x 256 x3) before providing it as an input to the model. Two max-pooling layer of size (2x2) has been used.

For the third convolutional block, there is no need of max-pooling layer as the image size cannot be further reduced. The activation functions used after every convolutional layer is ReLU. The dropout layers help the model to converge without overfitting. The threshold used for the dropout layer is 0.5. Then flatten layer is used to convert the features into a vector. Finally, two fully connected layers are used. The number of units in the last layer depends on the number of classes to classify.

The number of parameters in XCov-Net for four class classification is 21,505,083 out of which 21,450,555 are trainable parameters and 54,528 are non-trainable. Table 4-2 shows the summary of the XCov-Net.

Table 4-2: Proposed Model Summary

Sr.no	Layer	Shape	Parameters
1	Xception model	8 x 8 x 2048	20,861,480
2	Conv 2D	8 x 8 x 32	589856
3	MaxPooling 2D	4 x 4 x 32	0
4	Dropout	4 x 4 x 32	0
5	Conv 2D	4 x 4 x 32	9248
6	MaxPooling 2D	2 x 2 x 32	0
7	Dropout	2 x 2 x 32	0
8	Conv 2D	2 x 2 x 64	18496
9	Dropout	2 x 2 x 64	0
10	Flatten	None, 256	0
11	Dense	None, 100	25700
12	Dense	4	404

Chapter 5

Experiment and Results

This Chapter discusses the implementation of the proposed methodology and also discusses the result which shows the performance of the model.

5.1 Keras

Keras is a deep learning framework which is built on top of tensorflow. It is the most commonly used for deep learning tasks and it is very easy to use. Tensorflow is machine learning library that helps to develop and work with the models.

5.2 Implementation

The proposed XCov-Net is implemented in Keras on Kaggle notebook. The model is trained based on three different scenarios. The first model performs four-class classification and the classes it considered are Normal, COVID-19, Pneumonia bacterial and Pneumonia viral. The second scenario is the three-class classification model which was trained to perform classification among Normal, COVID-19. Binary classification is a third scenario in which the model distinguishes between Normal and COVID-19.

Keras makes different deep learning models and pre trained weights easily available to the end user. It helps to save a lot of time by not implementing and training the deep learning models from scratch. The method uses Xception which has been trained on imagenet dataset. The input of size (256 x 256 x 3) is passed with batch size 10 to Xception block. Then the output is then passed to the XCov block. Optimizers help in updating the weights and bias of the network by minimizing the loss function so that the network may converge. The goal of the network is to move towards global minima. There are several different optimizers and the proposed method used an adam optimizer. The learning rate was set to 0.0001. Usually when dealing with classification tasks the most common used loss function is cross-entropy loss. The model was trained end-to-end with 65 epochs using cross-entropy loss function.

5.3 Performance Measures

The most commonly used performance measure for the classification task is the Confusion Matrix. A simple confusion matrix is shown in figure 5-1. Confusion matrix contains True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN) values of actual and predicted classes.

Table 5-1: Confusion Matrix

	Predicted Class		
Actual Class		**Yes**	**No**
	Yes	True Positive	False Negative
	No	False Positive	True Negative

True Positive (TP): The amount the model predicted to be true and is actually true.
True Negative (TN): The amount the model predicted to be false and is actually false.
False Positive (FP): The amount the model predicted to be true and is actually false.
False Negative (FN): The amount the model predicted to be false and is actually true.

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

It is the measure of True Positive to that of all observation in Actual positive class.

Precision = $\frac{TP}{TP+FP}$

It is the measure of True Positive to that of all positive cases in Predicted class.

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

It is the measure of True Positive and True Negative to that of all the observations.

F1 Score = $\frac{2*Recall*Precision}{Recall+Precision}$

It is calculated using Recall and Precision.

5.4 Results

The proposed XCov-Net has been evaluated based on the above-mentioned performance measures. Accuracy shows how well the model predicted correctly. The model has been evaluated on the test data and the performance measures are recorded in the table 5-2. The accuracy obtained for the main model i.e. four-class classification is 0.948. For three-class, the accuracy is recorded as 0.987 and for the binary classification, the accuracy is 1. The three class and binary classification model is a modified version of the main four-class classification model.

Table 5-2: Results

Sr.no	Scenarios	Precision	Sensitivity	Accuracy	F1-Score
1	Four Class	0.949	0.949	0.949	0.948
2	Three Class	0.988	0.987	0.987	0.987
3	Two Class	1	1	1	1

5.5 Graphs

Training a neural network is really a difficult task. There are millions of parameters involved during training , in order to figure out how a network train , we can plot the loss and accuracy of the model at each epoch. This helps in the visualization of how the model performed. The loss and accuracy is calculated at each epoch. The lower the loss means that the predicted output made by the model is almost similar to the actual output. It can be seen in section 5.5 the loss and accuracy curve of the XCov-Net model. Loss and accuracy are usually inversely proportional. In order for the model to converge the loss should decrease and the accuracy should increase. It can be seen that the loss for all the three cases is decreasing. For binary class, it can be noticed from the graph that the model is a good fit. When the gap between the two losses is very low, it means the model is well trained and it can be seen that the losses are decreasing with increase in epochs thus the model converges smoothly. In order for the model to learn, it is trained on certain number of epochs. The number of epochs is a chosen based on the fact that the loss of the model is decreased and the accuracy is increased to such an extent that it may further cannot be improved. The number of epochs is set to 65 as further increasing the number of epochs does not increase the performance of the XCov-Net model.

5.5.1 Four Classes

Accuracy and Loss graph of the Model.

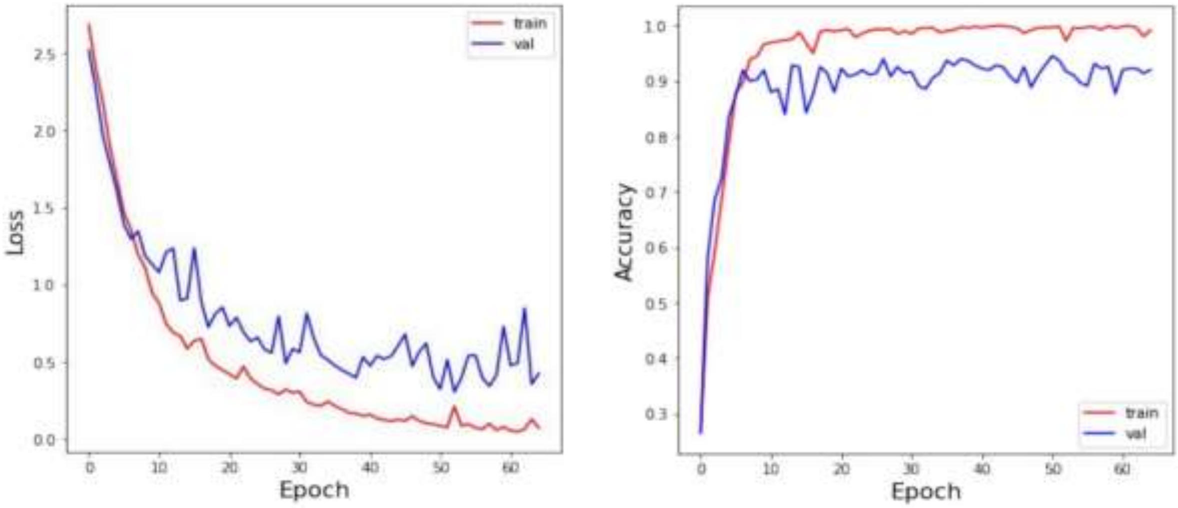

Figure 5-1: Loss and Accuracy graph of four class model

5.5.2 Three Classes

Accuracy and Loss graph of the Model.

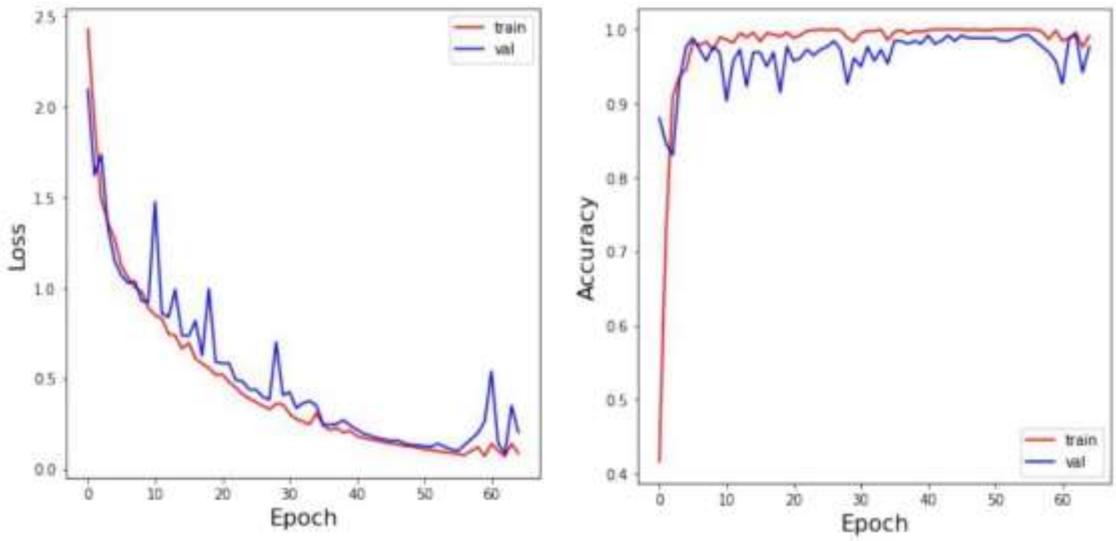

Figure 5-2: Loss and Accuracy graph of three class model

5.5.3 Two Classes

Accuracy and Loss graph of the Model.

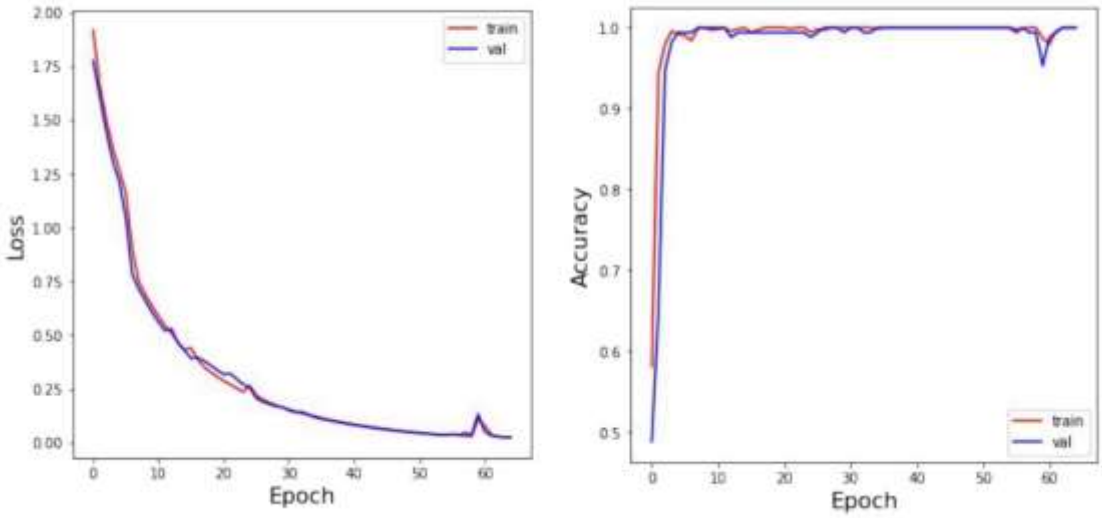

Figure 5-3: Loss and Accuracy graph of two class model

5.6 Confusion Matrix

The confusion matrix is the table that shows well the model performed in performing prediction. The confusion matrix of the model for the three scenarios can be seen in figure 5-4, figure 5-5 and figure 5-6. It shows how well the model made predictions. The confusion matrix of two class model in shows that the model made prediction accurately and there is no any false positive or false negative predictions. With the increase in the number of classes, it can be seen in that model made some false positive and false negative predictions specifically when dealing with images of pneumonia caused by bacteria or by virus. Confusion Matrix of all the three models are shown below.

- Four Class Model

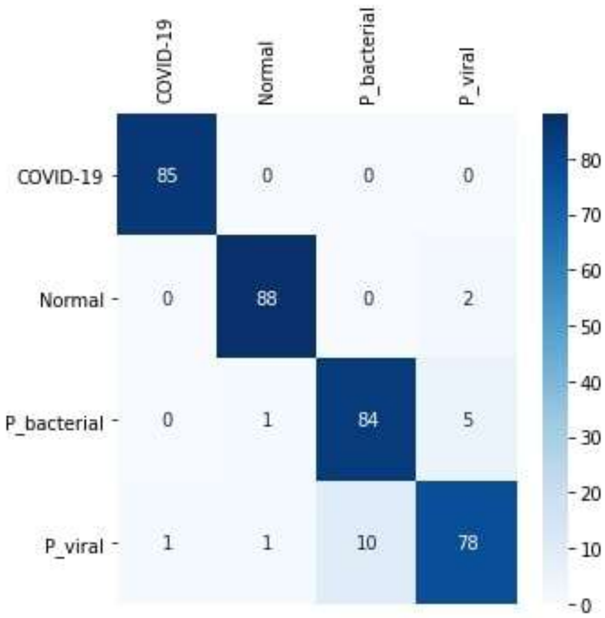

Figure 5-4: Confusion Matrix of four class model

- Three Class Model

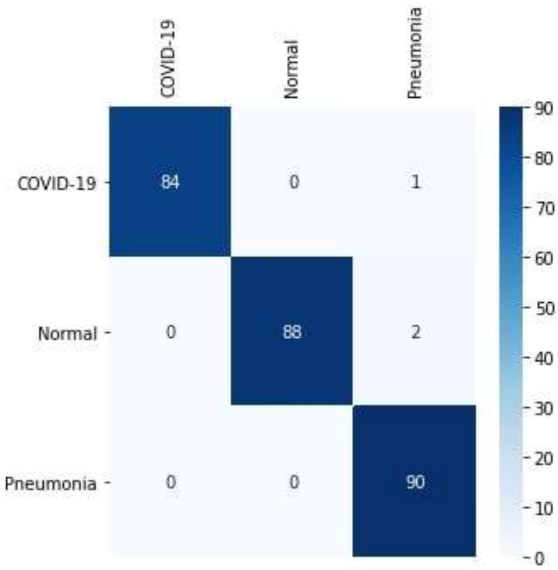

Figure 5-5: Confusion Matrix of three class model

- Two Class Model

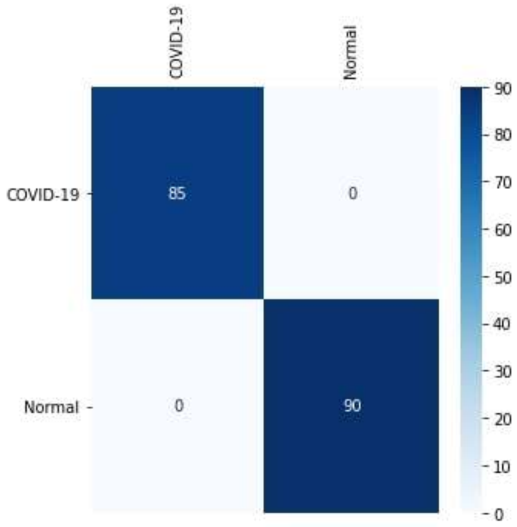

Figure 5-6: Confusion Matrix of two class model

Chapter 6

Discussion and Conclusion

This Chapter shows the test images predicted by the model and compares the proposed model with other models. It also discusses the conclusion and the future work.

6.1 Prediction

Test data is passed to the proposed model to predict the classes.

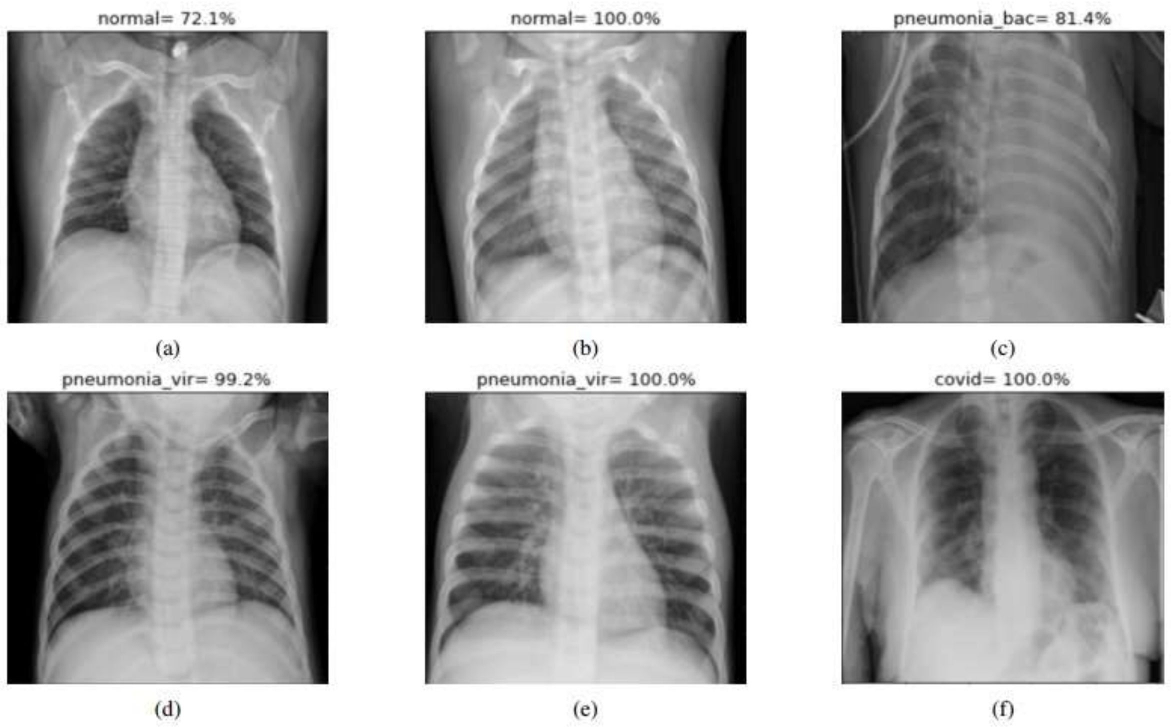

Figure 6-1: Shows the sample of X-ray images correctly predicted by the model along with the predicted probability

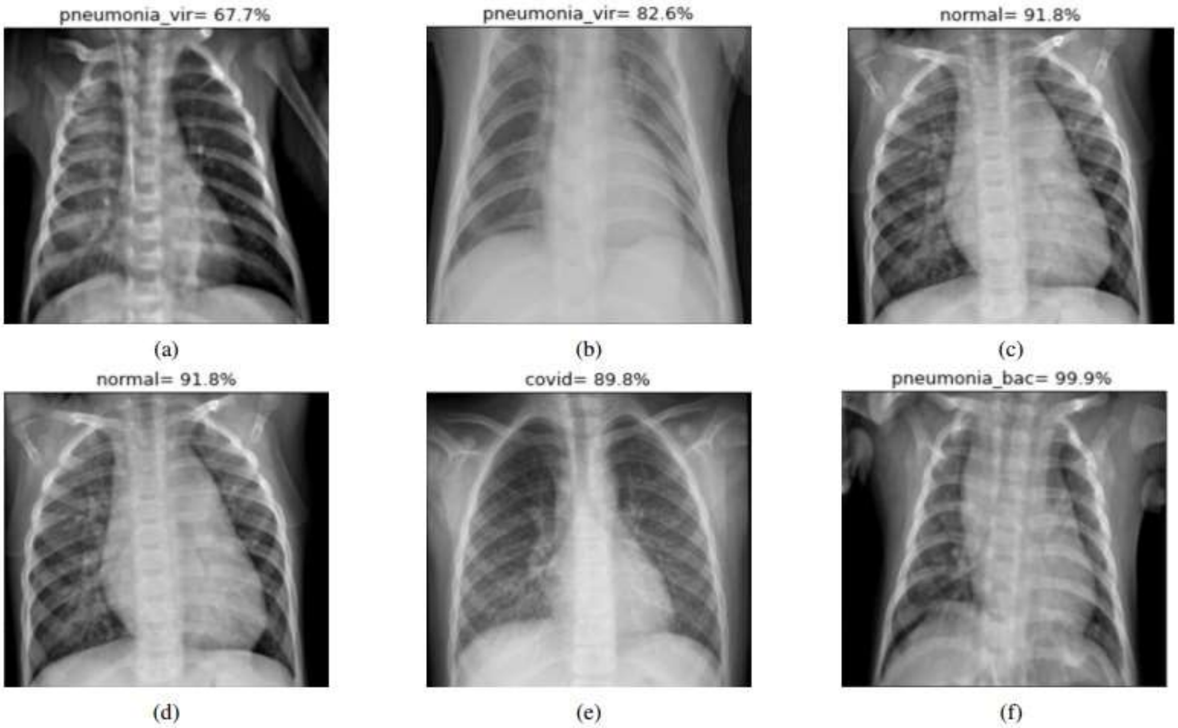

Figure 6-2: Shows the sample of X-ray images incorrectly predicted by the model along with the predicted probability

6.2 Discussion

The XCov-Net model can classify corona patient's chest X-ray with performance being remarkably good. There are several other ways to test for corona which includes LAMP, antibody test, antigen test, gene editing, CT scans and Chest X-ray images. Among these tests, CT scans and X-ray images can be used in automatic systems easily and can help in diagnosis within few seconds. The motivation for using X-ray images is that X-ray machines can be found almost in every clinic/ hospital whereas CT scans are not that common and expensive when compared to X-ray machines. The researchers have observed several different findings in radiological images of corona patients which are mentioned below [31].

- Development of Ground-glass opacities which can be bilateral, multifocal subpleural with air bronchograms.
- An irregular interlobular or septal thickening.
- Bronchiolectasis
- Thickening of the adjacent pleura
- Pleural effusion

The above mentioned founding's were observed in the radiological images. These findings can easily be learned by the deep learning models and can help in accurate diagnosis of the disease.

Table: 6-1 Shows the comparison with other models

Classes	Paper	Recall	Precision	F1 Score	Accuracy
Two	Naren et al[6]	0.96	1	0.98	0.98
	Ozturk et al[8]	0.951	0.980	0.965	0.980
	Harsh et al[13]	0.976	0.820	0.891	0.881
	Asif et al[15]	0.993	0.983	0.985	0.99
	Luca et al[17] Model 1	0.96	0.920	0.94	0.96
	Luca et al[17] Model 2	0.87	0.910	0.89	0.98
	XCov-Net	**1**	**1**	**1**	**1**
Three	Ozturk et al[8]	0.853	0.899	0.873	0.870
	Asif et al[15]	0.969	0.95	0.956	0.95
	Das et al[19]	0.970	0.968	0.969	0.974
	XCov-Net	**0.987**	**0.988**	**0.987**	**0.987**
Four	Asif et al[15]	0.899	0.9	0.898	0.896
	XCov-Net	**0.949**	**0.949**	**0.949**	**0.948**

It can be seen from the performance measures in the table 6-1 that the model performed outstandingly well in classifying Normal from COVID-19 patients. The model also performed excellent while detecting COVID-19 from normal and pneumonia. Here, the pneumonia images consist of images from pneumonia caused by bacteria and virus. It can be seen in that four class classification model's performance is relatively low when compared to other scenarios. It can be observed from the confusion matrix of four class classification model in Chapter 5 that the model made incorrect predictions mostly among images of pneumonia. The model accurately detected that the given image is pneumonia image, but it incorrectly evaluated pneumonia bacterial and pneumonia viral images due to which the accuracy of this model is less compared to other models. Figure 6-1 shows some the prediction made on test images which were given as an input to the model and their classes were correctly predicted. Along with the images, the predicted probability in terms of percentage is also displayed by the model. Some of the images were incorrectly predicted by the model and the results of wrong predictions are shown in the figure 6-2

6.3 Conclusion

The XCov-Net can detect COVID-19 cases with greater accuracy and can replace the RT-PCR test. The model can help the radiologists and health experts in the diagnosis of disease and can save a lot of time by providing the results within few seconds which can help the health experts to treat the patients according to the severity of the disease thus helping in less spread of the disease.

The XCov-Net performs classification automatically and the model itself extract features from the images. The accuracy recorded for four class model is 0.948, 0.987 using three class model and 1 using binary class model.

6.4 Future Work

The proposed model can further be used to perform classification among different chest related diseases by using more data. With the increase in data, the performance and robustness of the model can also be increased when dealing with multiclass classification.

References

[1] World Health Organization (who), available:https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports

[2] World Health Organization (who),available:https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a detail/qacoronaviruses::text=symptoms

[3] Times of india, available:https://timesofindia.indiatimes.com/india/whycovid-testing-is-a-slow-process-and-types-of-testsavailable/articleshow/76459365.cms.

[4] Xingzhi Xie, Zheng Zhong, Wei Zhao, Chao Zheng, Fei Wang, and Jun Liu. Chest ct for typical 2019-ncov pneumonia: relationship to negative rt-pcr testing. Radiology, pages 200343–200343, 2020.

[5] Jahanzaib Latif, Chuangbai Xiao, Azhar Imran, and Shanshan Tu. Medical imaging using machine learning and deep learning algorithms: a review. In 2019 2nd International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), pages 1–5. IEEE, 2019.

[6] Ali Narin, Ceren Kaya, and Ziynet Pamuk. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. arXiv preprint arXiv:2003.10849, 2020.

[7] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.

[8] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U Rajendra Acharya. Automated detection of covid- 19 cases using deep neural networks with x-ray images. Computers in Biology and Medicine, page 103792, 2020.

[9] Joseph Redmon and Ali Farhadi. Yolo9000: better, faster, stronger. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 7263–7271, 2017.

[10] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 779–788, 2016.

[11] Yujin Oh, Sangjoon Park, and Jong Chul Ye. Deep learning covid-19 features on cxr using limited training data sets. IEEE Transactions on Medical Imaging, 2020.

[12] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 4700–4708, 2017.

[13] Harsh Panwar, PK Gupta, Mohammad Khubeb Siddiqui, Ruben MoralesMenendez, and Vaishnavi Singh. Application of deep learning for fast detection of covid-19 in x-rays using ncovnet. Chaos, Solitons & Fractals, page 109944, 2020.

[14] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556,2014.

[15] Asif Iqbal Khan, Junaid Latief Shah, and Mohammad Mudasir Bhat. Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images. Computer Methods and Programs in Biomedicine, page 105581, 2020.

[16] François Chollet. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1251–1258, 2017

[17] Luca Brunese, Francesco Mercaldo, Alfonso Reginelli, and Antonella Santone. Explainable deep learning for pulmonary disease and coronavirus covid-19 detection from x-rays. Computer Methods and Programs in Biomedicine, 196:105608, 2020.

[18] Sivaramakrishnan Rajaraman, Jen Siegelman, Philip O Alderson, Lucas S Folio, Les R Folio, and Sameer K Antani. Iteratively pruned deep learning ensembles for covid-19 detection in chest x-rays. arXiv preprint arXiv:2004.08379, 2020.

[19] N Narayan Das, Naresh Kumar, Manjit Kaur, Vijay Kumar, and Dilbag Singh. Automated deep transfer learning-based approach for detection of covid-19 infection in chest x-rays. IRBM, 2020.

[20] Xinggang Wang, Xianbo Deng, Qing Fu, Qiang Zhou, Jiapei Feng, Hui Ma, Wenyu Liu, and Chuansheng Zheng. A weakly-supervised framework for covid-19 classification and lesion localization from chest ct. IEEE Transactions on Medical Imaging, 2020.

[21] Dilbag Singh, Vijay Kumar, and Manjit 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, pages 1–11, 2020.

[22] Shaoping Hu, Yuan Gao, Zhangming Niu, Yinghui Jiang, Lao Li, Xianglu Xiao, Minhao Wang, Evandro Fei Fang, Wade Menpes-Smith, Jun Xia, et al. Weakly supervised deep learning for covid-19 infection detection and classification from ct images. IEEE Access, 8:118869–118883, 2020.

[23] Subhankar Roy, Willi Menapace, Sebastiaan Oei, Ben Luijten, Enrico Fini, Cristiano Saltori, Iris Huijben, Nishith Chennakeshava, Federico Mento, Alessandro Sentelli, et al. Deep learning for classification and localization of covid-19 markers in point-of-care lung ultrasound. IEEE Transactions on Medical Imaging, 2020.

[24] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.

[25] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1–9, 2015.

[26] https://medium.com/@zurister/depth-wise-convolution-and-depth-wise-separable convolution- 37346565d4ec

[27] Joseph Paul Cohen, Paul Morrison, Lan Dao, Karsten Roth, Tim Q Duong, and Marzyeh Ghassemi. Covid-19 image data collection: Prospective predictions are the future. arXiv preprint arXiv:2006.11988, 2020.

[28] Chest x-ray images (normal and pneumonia) https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia.

[29] Jia Deng, Wei Dong, Richard Socher, Li-Jia Li, Kai Li, and Li FeiFei. Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition, pages 248–255. Ieee, 2009.

[30] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10):1345–1359, 2009.

[31] Heshui Shi, Xiaoyu Han, Nanchuan Jiang, Yukun Cao, Osamah Alwalid, Jin Gu, Yanqing Fan, and Chuansheng Zheng. Radiological findings from 81 patients with covid-19 pneumonia in wuhan, china: a descriptive study. The Lancet Infectious Diseases, 2020.