# COVID-19 PNEUMONIA PREDICTION USING CHEST X-RAYS

*Padmini Ramesh, Sidhant Naveria*

Institute of Systems Science, National University of Singapore, Singapore 119615

## ABSTRACT

Corona virus disease 2019 (COVID-19) is an infectious disease caused by SARS-CoV-2, a virus closely related to the SARS virus. On 30 December 2019, Dr.Li saw a patient's report which showed a positive result with a high confidence level for SARS corona virus tests.[1].Globally the total number of people infected by COVID -19 has spiked up to 3,588,773 . The total number of deaths have accounted to about 247503 as reported on 6 th of May 2020.[2].Chest X-rays and CT-scans are considered as routine imaging tool for COVID-19 diagnosis.Everyday clinicians are exposed to reading high volumes of images during each shift.Tiredness,distractions are few factors that might lead them to miss important details in an image.This is when the idea of using an automated image analysis tools comes into picture.Therefore , a model that distinguishes between normal and Covid-19 chest x-ray images was developed.Three different classification models such as Residual Attention network, Resnet-18 and Inception v3 were used and their performances were compared.

***Index Terms***— COVID-19, CHEST X-RAYS , RESIDUAL ATTENTION NETWORK,RESNET 18 ,INCEPTION V3.

## 1. INTRODUCTION

During late December 2019, an outburst of an unidentified corona virus occurred in Wuhan, a city in China which further escalated globally as a menace resulting into a global pandemic.The disease outbreak was later named as Covid-19 by the World Health Organisation (WHO). COVID-19 usually poses as an acute respiratory infection and person to person transmission may occur through contact or droplets of infected people in the vicinity.Sometimes, the virus may occur asymptomatically and the person act as a carrier in spreading the disease. The reaction of the virus includes causing febrile illness or respiratory infection, clinical,pathological or radiological evidences of pneumonia, gastrointestinal(nausea, diarrhoea), muscluoskeletal and neurologic disorders. Diagnosis of Covid-19 was carried out procedures such as all-time reverse transcription polymerase chain reaction (RT-PCR) , Chest x-rays and ct-scans which showed evidences of bilateral pneumonia with lung consolidqation and ground glass opacities.

Chest X-ray have proven to be an efficient tool for an early stage COVID-19 diagnosis in places that lack resources for RT-PCR. It has achieved a sensitivity of about 69 percent. Few abnormal features present in CXRs that contribute towards the presence of COVID-19 are as follows:

- Ground glass densities.
- Bilateral lower lobe consolidations.
- Peripheral air space opacities
- Diffused air spaces.

Ever since the pandemic has gained momentum, doctors have continued to rely on portable CXR as they are readily available and has reduced the capability of infection spread.As the number of X-rays taken has become larger and larger everyday , it has become difficult for the radiologists to analyse the images and deduce accurate diagnosis. Hence the need for an automatic image analysis tool arouse.[3] This project focuses on the development and implementation of three different classification models : Residual attention networks,Resnet-18 , Inception V3 and comparison of their performance metrics. Initially , the chest x-rays images were subject to a detection model namely detectron2 , which detects the lung part present in x-rays and crops them out. The cropped images are then fed into the classification model in order to categorize normal and covid-19 affected lungs.

The following sections describe the published work on the architecture of the model, the experimentation carried out and the results obtained .

## 2. RELATED WORK

### 2.1. Detectron

The work carried out has been divided in to two parts : detection and classification. Detectron2 was developed by Facebook AI research(FAIR) in order to provide state of the art codebase for object detection research.[4].It includes the implementation of object detection algorithms such as:

- Mask RCNN
- Faster RCNN

- RPN
- Fast RCNN
- R-FCN
- RetinaNet

The backbone of above-mentioned architectures are as follows:

- ResNeXt
- ResNet
- Feature Pyramid Network
- VGG 16

### 2.2. Residual Attention Network

Residual Attention Network was unrolled by SenseTime, Tsinghua University, Chinese University of Hong Kong (CUHK), and Beijing University of Posts and Telecommunications. In this architecture, multiple attention layers are stacked up to generate attention aware features. It includes two types of branches:Mask branch and trunk branch.Trunk branch is an upper branch in attention layer used for feature extraction and Mask branch uses bottom-up top-down structure to learn the same-size mask M(x).[5]The output of attention module is

$$H_{i,c} = M_{i,c} * T_{i,c} \tag{1}$$

Where i ranges over the spatial location, and c is the channel index from 1 to c. The attention mask can serve as feature selector during forward inference.Figure 1 shows the architecture of residual attention network.1

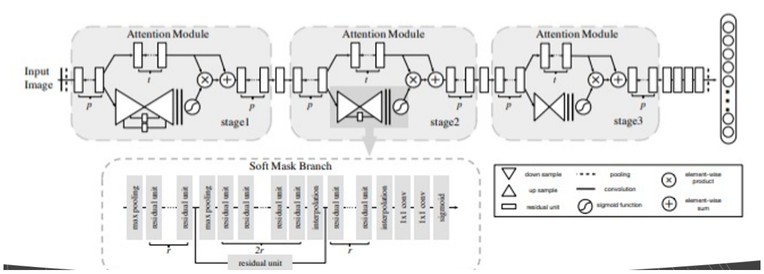


**Fig. 1**. Residual attention network Architecture

### 2.3. Bilateral symmetry of lungs

As reported by the authors of [6], the abnormalities present in the lungs that augment the existence of any disorder are cavitations , consolidations, infiltrates, blunted costophrenic angles, opacities,pleural effusion , ground glass features etc.., Sometimes either of these features are present in smaller or larger quantities and might affect the classifier's performance. Therefore the technique of splitting the lungs into two parts(left and right) and concatenating their features to train the model was suggested by the publishers. This method has been incorporated into this project and has been discussed under the "proposed approach" section.

### 2.4. Residual networks

When the neural network gets deeper, it starts converging posing degradation problem resulting in saturation of accuracy and then degrades rapidly.Resnet networks comes to the rescue which does identity mapping instead of direct mapping using a function H(x). Identity mapping are used directly when the input and output are of same dimension.[7]

$$y = F(x, W_i) + x \tag{2}$$

When the dimensions are different , zero entries are padded with the increased dimension and hence the mapping is represented as follows:

$$y = F(x, W_i) + W_s x \tag{3}$$

Each ResNet block in Resnet 18 is two layer deep and converges faster compared to the plain counterpart of it.Figure 2 is a representation of residual network.2

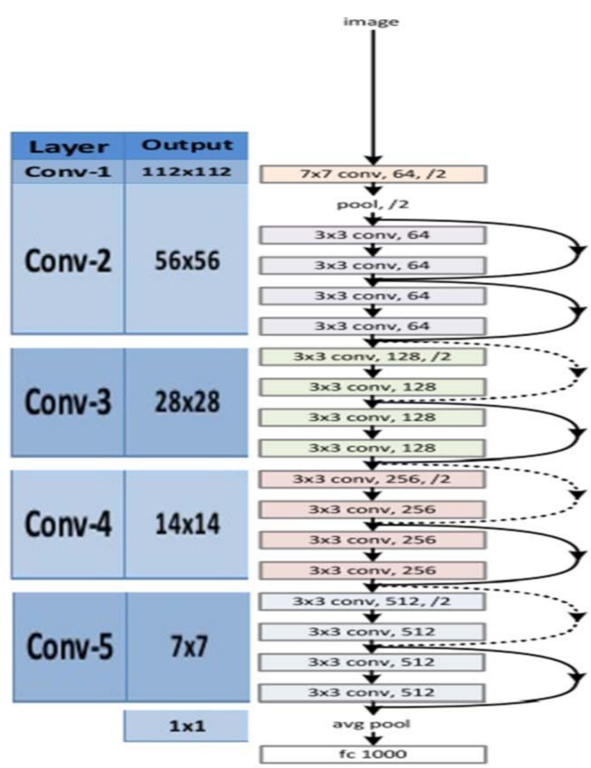


**Fig. 2**. Resnet 18.

### 2.5. Inception V3

Initially , the inception deep convolution architecture was named as Google-Net and later renamed as Inception-v1 . The v1-architecture was refined by adding batch normalization and upgraded into Inception- v2.Later , factorization of convolution layers were added to produce Inception v3. The idea of factorization was inculcated in order to reduce the number of parameters.It also consists of an auxiliary classifier layer. Inception v3 makes use of one auxiliary layer instead of two as in Google-Net.It makes use of efficient grid-size reduction, where two set of feature maps: one obtained from convolution layer with a stride 2 and the other obtained by max pooling.These two feature maps are concatenated and fed into the consecutive inception module.[8].Figure 3 is a depiction of Inception v3 architecture.3

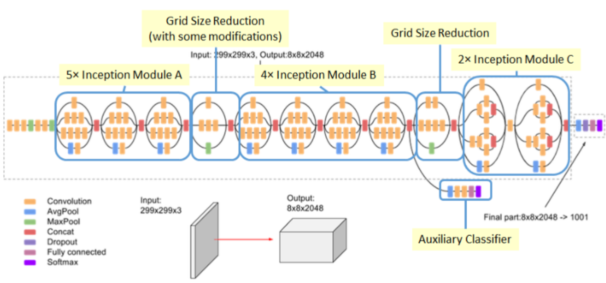


**Fig. 3**. Inception V3 architecture.

## 3. PROPOSED APPROACH

### 3.1. Detection:

Initially,the images of Chest X-rays and the co-ordinate positions of the lungs were provided. This data was trained on Detectron2 and tested on a non-annotated set of images and the position of lungs were obtained. After the detection process, the detected part was cropped and saved for further classification. The dataset description and the training details are as follows:

Total number of chest x-ray images used to train : 26 images.

Annotations provided: Top left x and y co-ordinates of the lungs and its width and height.(Bounding Box co-ordinates)

Total number of chest x-ray images used to test : 25 images.

Learning rate :0.0025

Total number of epochs: 500

Batch size=2

The model puts forward a bounding box over the region where the lungs are present.The region under the bounding box are cropped and saved in order to perform further classification procedure. A spectacle of the detection process can be visualised in figure 4.4

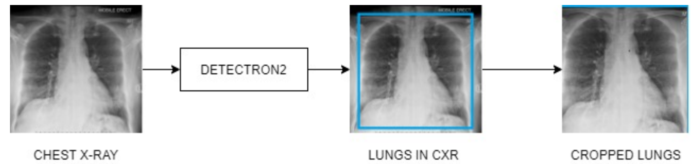


**Fig. 4**. Detection Task.

The experiments for classification were carried out under two stages:

- Experiment 1 : Model was trained on whole lung images
- Experiment 2 : Model was trained by splitting the lungs into left and right parts and concatenating their features.

The process flow of classification task can be visualised in figure 5.5

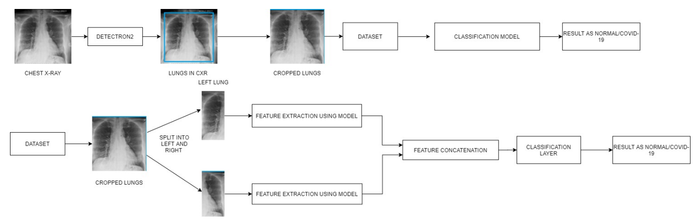


**Fig. 5**. Classification Task for experiment 1 and 2

The following models were explored for the classification task to be accomplished:

- Residual attention network
- ResNet 18
- Inception V3

The dataset split is as stated below:

- Training data : Normal lungs= 200 images , Covid-19 lungs= 198 images
- Testing data : Normal lungs= 23 images Covid-19 images =25 images

## 4. EXPERIMENTATION AND RESULTS

### 4.1. Residual Attention Network

#### *4.1.1. Experiment 1:*

The model was initially trained for about 40 epochs with the learning rate 3e-02 and the best model weights were saved. Again, the model was retrained using the best weights for 30 epochs with the learning rate of 3e-03. The process was repeated by reducing the learning rate and retraining the model with the best weights saved at each point. The following parameters were used : Optimiser: SGD Weight decay: 0.0003 Momentum: 0.8 Table 1 shows the confusion matrix for this experiment.1

**Table 1**. Confusion Matrix for Residual attention network (exp 1)

| Class names | Covid-19 | Normal |
|---|---|---|
| Covid-19 | 24 | 5 |
| Normal | 3 | 16 |

#### 4.1.2. *Experiment 2:*

The model was trained by splitting the images into two parts: left and right and their features concatenated. There are two branches in the model and each of them is a residual network. The left and the right parts were trained on the left and right branch respectively to obtain features. Their features were then concatenated and fed to the classifier layer. The remaining training process is like experiment 1. Table 2 shows the confusion matrix for this experiment.2

**Table 2**. Confusion Matrix for Residual attention network (exp 2)

| Class names | Covid-19 | Normal |
|---|---|---|
| Covid-19 | 22 | 3 |
| Normal | 1 | 22 |

### 4.2. ResNet 18

#### 4.2.1. *Experiment 1:*

A pre-trained Resnet 18 model was used with the final classifier layer modified to the number of classes according to this problem that is two. Initially the model was trained for about 50 epochs with a learning rate of 1e-3 and the best model weights were saved. The model was re-trained using this best weights with a learning rate of 1e-4 for about 50 epochs. The process was repeated and approach is same as that followed for residual attention network.Table 3 puts down the confusion matrix for this experiment.3

**Table 3**. Confusion Matrix for Resnet 18(exp 1)

| Class names | Covid-19 | Normal |
|---|---|---|
| Covid-19 | 19 | 6 |
| Normal | 8 | 15 |

#### 4.2.2. *Experiment 2:*

The pre-trained model was used by removing its last layer as a feature extractor that could extract features from the left and right parts respectively and were concatenated and fed in to the classification layer. The model was then fine-tuned and trained for about 50 epochs with a learning rate of 1e-2.Table 4 is the confusion matrix for this experiment.4

**Table 4**. Confusion Matrix for Resnet 18 (exp 2)

| Class names | Covid-19 | Normal |
|---|---|---|
| Covid-19 | 20 | 5 |
| Normal | 5 | 18 |

### 4.3. Inception V3

#### 4.3.1. *Experiment 1:*

A pre-trained Inception V3 model without the auxiliary classifier layer and a modified final classification layer according to the number of classes.The model was initially trained for 50 epochs with a learning rate of 1e-2 and the remaining training approach was similar to that used in experiment 1 of other models.Table 5 puts forward the confusion matrix for this experiment.5

**Table 5**. Confusion Matrix for Inception V3 (exp 1)

| Class names | Covid-19 | Normal |
|---|---|---|
| Covid-19 | 15 | 10 |
| Normal | 11 | 12 |

#### 4.3.2. *Experiment 2:*

The pre-trained model was fine-tuned in a way similar to 3.3.2 without the auxiliary classification layer.Table 6 shows the confusion matrix for this experiment.6

**Table 6**. Confusion Matrix for Inception V3 (exp 2)

| Class names | Covid-19 | Normal |
|---|---|---|
| Covid-19 | 24 | 1 |
| Normal | 15 | 8 |

### 4.4. Experimental results

After testing the models, following metrics were compared to find the one that gives the best performance:

- Accuracy
- Precision

- Recall
- F1 score
- ROC score and Curve

Table 7 and 8 has recorded the above mentioned metrices.7 8

**Table 7**. The performance comparison for experiment 1

| Model | Residual attention | Resnet 18 | Inception v3 |
|---|---|---|---|
| Acccuracy | 83.3 | 70.83 | 56.25 |
| Precision | 0.88888 | 0.6522 | 0.5217 |
| Recall | 0.8275 | 0.7143 | 0.5455 |
| F1-score | 0.857 | 0.6818 | 0.5333 |

**Table 8**. The performance comparison for experiment 2

| Model | Residual attention | Resnet 18 | Inception v3 |
|---|---|---|---|
| Acccuracy | 91.6 | 79.17 | 66.67 |
| Precision | 0.9595 | 0.7826 | 0.3478 |
| Recall | 0.88 | 0.7826 | 0.8889 |
| F1-score | 0.928 | 0.7826 | 0.5000 |

The following figures show the ROC curves with a threshold of 0.5 for all the three models for the corresponding experiments implemented:6 7 8 9 10 11

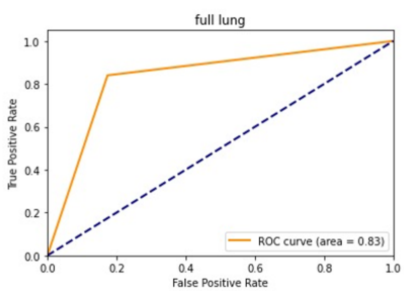


**Fig. 6**. ROC curve for residual attention network for experiment 1

An instance of how the system works can be visualised below:

- Initially an image is sent for prediction.
- The output is returned in a text file and it contains the filename of the image and the probability that the image might have covid-19 infection.
- The naming convention of the file is the current timestamp that is ,the time when the image was sent for prediction.12

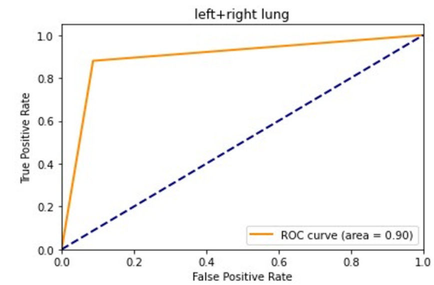


**Fig. 7**. ROC curve for residual attention network for experiment 2

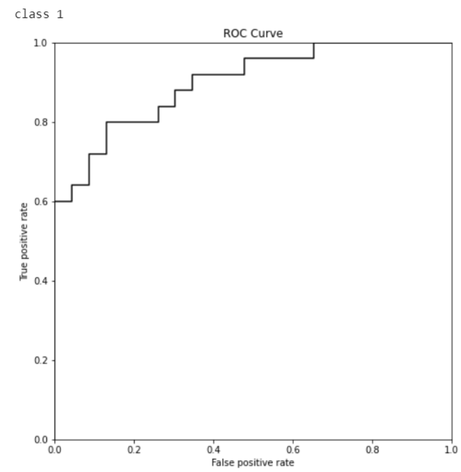


**Fig. 8**. ROC curve Resnet18 for experiment 1

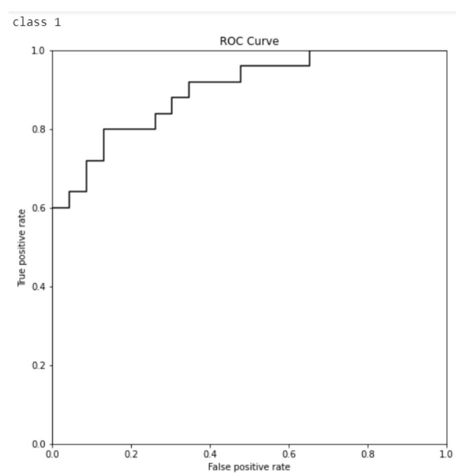


**Fig. 9**. ROC curve Resnet18 for experiment 2

## 5. CONCLUSIONS

From the metrics displayed in table 7 and table 8 , it can be observed that the accuracy obtained from experiment 2 ,that is model trained by splitting the input into two parts and con-

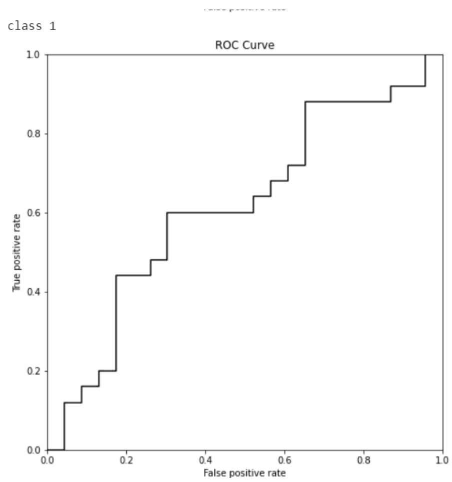


**Fig. 10**. ROC curve Incpetion v3 for experiment 1

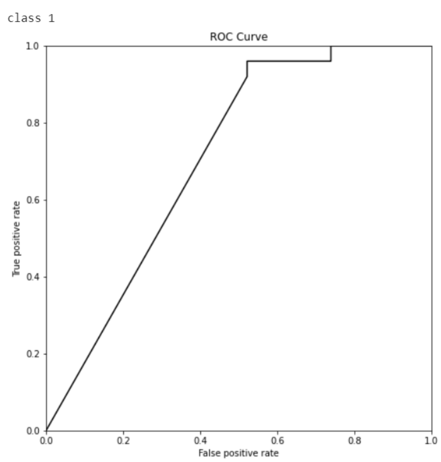


**Fig. 11**. ROC curve Incpetion v3 for experiment 2

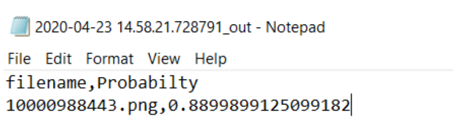


**Fig. 12**. Example of prediction

catenation have proven to give better performance than the one trained using whole images. It can also be seen that the metrics obtained from residual attention has been better than the other two models. Therefore after completion of all experiments it can be concluded that, residual attention network has demonstrated better efficiency as compared to the other models from both the experiments.

## 6. REFERENCES

[1] "Corona virus disease 2019 , from wikipedia the free encyclopedia," .

[2] "World health organisation situation report 107," .

[3] Adam Jacobi, Michael Chung, Adam Bernheim, and Corey Eber, "Portable chest x-ray in coronavirus disease-19 (covid-19): A pictorial review," *Clinical Imaging*, 2020.

[4] Yuxin Wu, Alexander Kirillov, Francisco Massa, Wan-Yen Lo, and Ross Girshick, "Detectron2," https://github.com/facebookresearch/detectron2, 2019.

[5] Fei Wang, Mengqing Jiang, Chen Qian, Shuo Yang, Cheng Li, Honggang Zhang, Xiaogang Wang, and Xiaoou Tang, "Residual attention network for image classification," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2017, pp. 3156–3164.

[6] KC Santosh and Sameer Antani, "Automated chest x-ray screening: Can lung region symmetry help detect pulmonary abnormalities?," *IEEE transactions on medical imaging*, vol. 37, no. 5, pp. 1168–1177, 2017.

[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*, 2016, pp. 770–778.

[8] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alexander A Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Thirty-first AAAI conference on artificial intelligence*, 2017.

## 7.CONTRIBUTIONS

| TEAM MEMBERS | WORK DONE |
| --- | --- |
| SIDHANT NAVERIA | RESEARCH ON PREVIOUS PNEUMONIA DETECTION AND CLASSIFICATION PAPERS, FAMILIARISATION WITH MEDICAL IMAGE ANALYSIS USING U-NET FOR IMAGE SEGMENTATION, CLASSIFICATION OF X-RAY USING DENSENET, IMPLEMENTATION OF OBJECT DETECTION MODELS, RESIDUAL ATTENTION NETWORK MODEL. |
| PADMINI RAMESH | RESEARCH ON PREVIOUS PNEUMONIA DETECTION AND CLASSIFICATION PAPERS, FAMILIARISATION WITH MEDICAL IMAGES USING U-NET FOR IMAGE SEGMENTATION, IMPLEMENTATION OF TWO LAYER CONVOLUTIONAL CLASSIFICATION MODEL TO CHECK THE BILATERAL SYMMETRY PERFORMANCE, IMPLEMENATION OF CLASSIFICATION MODELS SUCH AS RESNET 18, INCEPTION V3. |