# Deep Learning App to Classify COVID-19 X-Rays

Syed Mohammad Waqas Ali
*Avionics Engineering Department*
*College of Aeronautical Engineering, NUST*
Risalpur, Pakistan
smwaqasali1999@gmail.com

***Abstract*—The first case of the COVID 19 Pandemic was confirmed in Wuhan, China on December 31st. Since then, it has drastically affected the course of this world. The healthcare systems were pushed to their limits since more and more people required screening. Reverse Transcription Polymerase chain reaction (RT-PCR) was the definitive test for the detection of the virus, but the need of better and faster results was increasing, and therefore medical advancement in COVID 19 Testing was required. Chest X-ray based disease classification has emerged as an alternative method to help diagnosis of COVID 19. Some of the symptoms of COVID 19 includes cough, fever, shortness of breath, and tiredness [1]. The same symptoms are seen in a person who is affected with pneumonia. Hence, it creates a barrier for radiologists for an accurate detection of COVID 19 through CXRs. In this survey paper, we have discussed what features are there in a CXR that show a possibility of COVID 19 pneumonia in a patient. Furthermore, convolutional neural networks and their types are discussed. Pre-trained models such as VGG-16, ResNet50, ResNet18, Inception V3 are discussed in detail. This survey paper discusses their architecture and how prior researchers have used them to obtain their conclusions.**

***Index Terms*—COVID 19, CXR, convolutional neural networks, VGG-16, ResNet, Inception V3**

## I. Introduction

COVID-19 is a virus that causes severe respiratory illnesses ranging from the ordinary cold to life-threatening infections such as Severe Acute Respiratory Syndrome (SARS) and Middle East Respiratory Syndrome (MERS) (MERS).

The principal symptoms of COVID-19, according to WHO reports, are similar to those of the ordinary flu: fever, weariness, dry cough, shortness of breath, pains, and sore throat [25] [26]. The similarity of COVID-19 symptoms to flu symptoms makes early diagnosis of the coronavirus problematic. The coronavirus, like other viruses and bacteria, has been discovered to cause pneumonia in some people, and the treatment for coronavirus-induced pneumonia differs from that for other types of pneumonia.

One of the quickest ways to diagnose patients is to use radiography and radiology images to detect the condition. Early research revealed distinct abnormalities in the chest radiographs of COVID-19-infected patients.

Advances in artificial intelligence have enabled the implementation of algorithms powered by convolutional neural networks to detect COVID 19 through Chest X-rays. These algorithms learn from the CXR images that are already available from different public sources and improve their accuracy throughout time. It saves a great amount of time for the doctors after which the patient with COVID 19 is confirmed through PCR testing, and then dealt accordingly.

## II. Literature Review

### A. Features that show COVID 19 in a CXR

It is observed that COVID 19 patients who show clinical symptoms, are more likely to show an abnormal CXR [2]. Recent study shows that the main findings in a CXR image shows radiographic features that include consolidation, ground-glass opacities (GGO), pulmonary nodules and reticular–nodular opacities, with basal, peripheral, and bilateral predominance [3].

Figure 1 shows a chest X-ray showing a simple case of lung tissue involvement with right infrahilar reticular–nodular opacity. On the other hand, Figure 2 shows a more severe case of lung tissue involvement. This CXR shows right lower zone lung consolidation and diffuse bilateral airspace reticular–nodular opacities, which are more prominently visible at the boundary parts of lower zones. Figure 3 shows a CXR which exhibits a severe case of lung tissue involvement. This is caused by diffuse bilateral airspace reticular–nodular opacities that are more prominent on boundary parts of the lower zones, and ground glass opacity in both lungs which are dominant in mid-zones and on the lower zones. Figure 4 shows a normal CXR with clear lungs [4].

For Example, Figure 5 shows four different patients with positive COVID-19. Figure 5A shows a chest X-ray that showed bilateral peripheral mid and lower zonal air space consolidation opacities (arrows). Figure 5B shows chest X-ray that showed bilateral peripheral zonal air space consolidation opacities along the periphery of both lungs (arrows). Figure 5C shows a chest X-ray which showed right peripheral mid and lower zonal air space consolidation opacities (arrows), smaller patches of consolidation are seen on the side (arrows). Figure 5D shows chest X-ray that showed bilateral air space consolidation opacities (arrows) of peripheral midzonal distribution in the right lung and neither peripheral nor perihilar mid and lower zonal distribution in the left lung (arrows).

According to research conducted to describe the main chest radiological features (CXR) of COVID-19 and correlate them with clinical outcome [5], the results that were obtained which shows variable features in CXRs are shown in Table 1, and their distribution in Table 2.

Fig. 1. CXR of a COVID 19 patient showing lung tissue involvement.

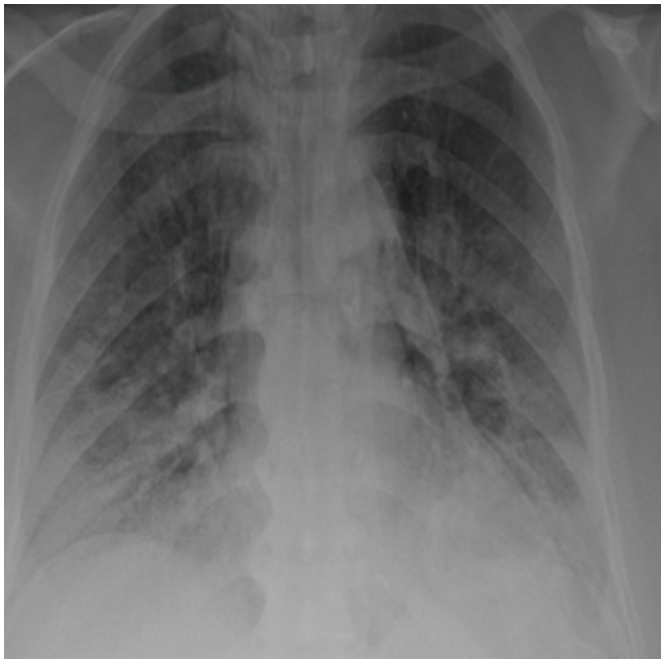

Fig. 2. CXR of a COVID 19 patient showing a more severe case of lung tissue involvement than that of Figure 1. The consolidation and diffuse bilateral airspace reticular–nodular opacities can be seen at the edges of lower parts on the right lung.

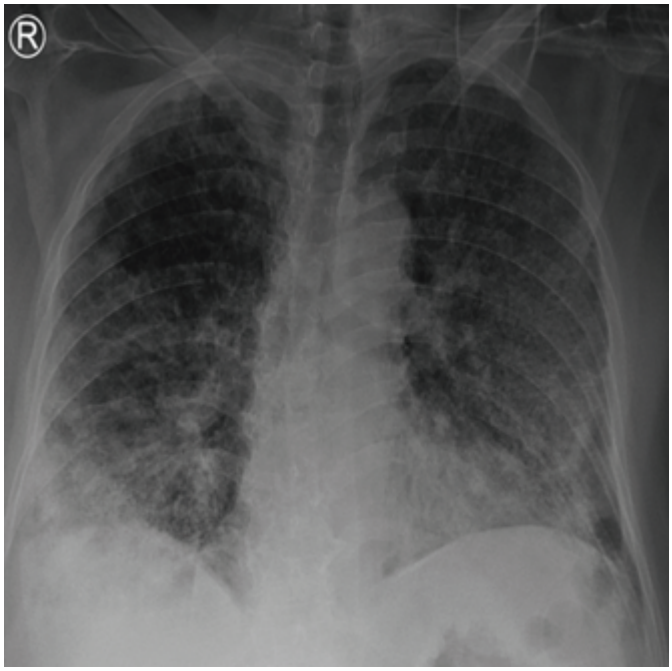

Fig. 3. CXR showing a more severe case of lung tissue involvement. The CXR shows diffuse bilateral airspace reticular–nodular opacities on the boundary parts of the lower zones, and ground glass opacities in the middle of both lungs.

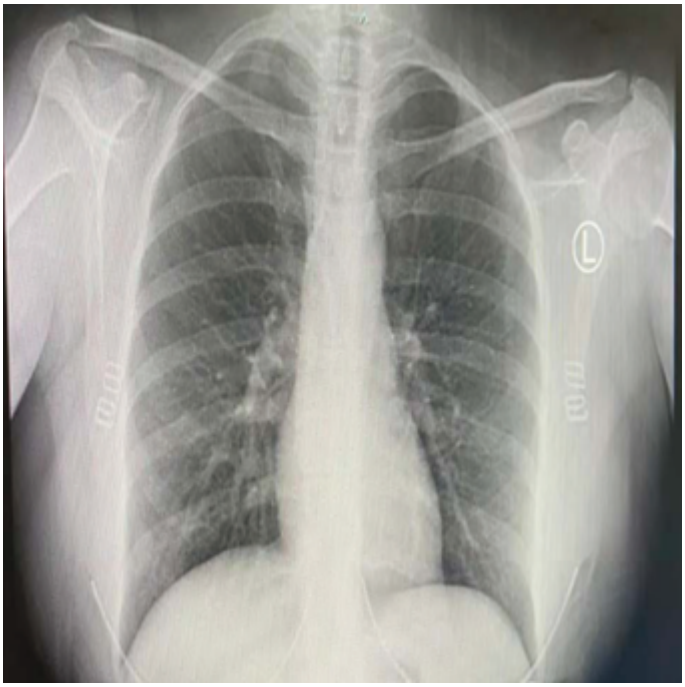

Fig. 4. Normal CXR with clear lungs, and no lung tissue involvement.

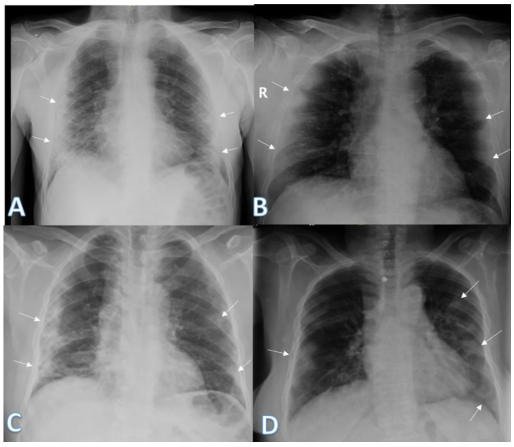


Fig. 5. CXR of COVID 19 patients showing different radiological features such as patchy or diffuse reticular–nodular opacities and consolidation, with basal, peripheral, and bilateral predominance.

TABLE I
COVID 19 RADIOLOGICAL FEATURES FOUND IN THE RESEARCH.

| COVID-19 Radiological features | Percentage (%) |
|---|---|
| Reticular–nodular opacities | 66.6 |
| Ground glass opacities | 62.8 |
| Consolidation | 57.7 |
| Vascular congestion signs | 39.3 |
| Cardiomegaly | 29.9 |
| Nodules | 23.5 |
| Pleural effusion | 16.7 |
| Pneumothorax | 2.4 |

### *B. Artificial Intelligence in Medicine*

Becoming a doctor is part of a process. A doctor trains himself throughout medical school, they observe and practice surgery repetitively and learn from mistakes to make themselves accurate at what they do. A similar process is seen when an AI algorithm is trained to do its job. In order to make an effective AI Algorithm, it is provided with a lot of data which is structured which means that each data point has a label that is recognizable to the algorithm. The Algorithm is fed with enough data points and data labels so that it is able to differentiate between different test labels. Hence, a test data, which the answers are already known is input to check the performance of the AI Algorithm. After which the according to the results, the performance can be improved by changing the algorithm accordingly and train it again, just like a doctor [6].

TABLE II
DISTRIBUTION OF COVID 19 RADIOLOGICAL FEATURES FOUND IN THE RESEARCH

| Distribution of the Features on CXRs | Percentage (%) |
|---|---|
| Peripheral | 57.7 |
| Perihilar | 20.7 |
| Diffuse | 41 |
| Basal predominance | 58.5 |
| Superior predominance | 13.1 |
| Right lung | 58 |
| Left lung | 42 |
| Bilateral | 69.2 |


Fig. 6. An algorithm that learns the basic anatomy of a hand and can rebuild where a missing digit should be is shown in the image above. Hand X-rays are used as input, and the output is a trace of where missing hand parts should be. The hand outline is the model in this case, which may be generated and applied to other photos. This could help doctors to see where a limb should be reconstructed or where a prosthetic should be placed.

The increasing amount of computation power paired with massive amounts of data being produced by the medical industry daily can be used to produce AI solutions that will cover an extensive part of medical industry problems relating to time consuming tests, and misdiagnosis.

For example, recently in 2018 researchers at Seoul National University Hospital and College of Medicine developed and validated a deep learning–based automatic detection algorithm (DLAD) for malignant pulmonary nodules (cancer cells) on chest radiographs and to compare its performance with physicians including thoracic radiologists [7]. The algorithm's performance was compared to multiple physician's detection abilities at the observer performance test on the same images and the algorithm outperformed 17 of 18 doctors, and it enhanced physicians' performances when used as a second reader.

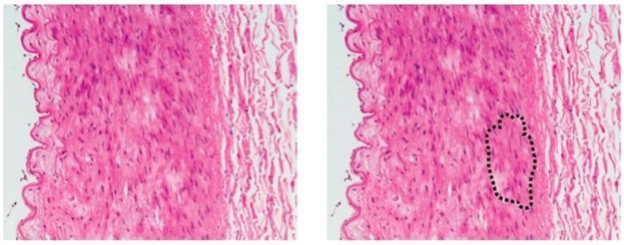

Fig. 7. The left panel shows the image fed into an algorithm. The right panel shows a region of potentially dangerous cells, as identified by an algorithm, that a physician should look at more closely.

One more example comes from Google AI Healthcare. In 2018, AI researchers at Google AI Healthcare created an algorithm called LYNA [8]. They used histology slides to or detect metastatic breast cancer in sentinel lymph node biopsies. Surprisingly, this algorithm could identify suspicious regions undistinguishable to the human eye in the biopsy samples given, being correct 99% of the time.

Both DLAD and LYNA are the prime examples of algorithms that are aiding the physicians in diagnosing by classifying healthy and diseased samples by showing them the features of the radiographic image that corresponds to

the specific disease and exemplify the potential strengths of algorithms in medicine.

## III. Image Recognition Models

### A. 2D Sequential CNN Models

2D Sequential CNN Models are one class of architectures that are available in the deep learning literature. They are used to analyze multidimensional data such as images and are classified as a special class of feedforward neural networks. The input images are transformed into a matrix which is then processed by the many layers of the CNN Model. CNNs are actually the regularized versions of multilayer perceptrons which means that each neuron in one layer is fully connected to each neuron of the next layer. They conserve memory relative to multilayer perceptrons by sharing parameters and using sparse connections. The model contains several convolutional and pooling layers.

*1) Convolutional Layer:* In a CNN, the input to the algorithm is a tensor which has a particular shape given by:

$$\text{(Number of inputs)} \times \text{(Input height)} \times \text{(Input width)} \times \text{(Input channels)}$$

When the input is passed through a convolutional layer, it creates an image processing kernel which contains a number of filters, after which the tensor input becomes a feature map (i.e. motifs). The feature map is also called an activation map. It has the following dimensions:

$$\text{(Number of inputs)} \times \text{(Feature map height)} \times \text{(Feature map width)} \times \text{(Feature map channels)}$$

The Convolutional Layer convolve (set of dot products) the input and pass the output to the next layer, and each convolutional layer processes the data only for its respective field. This layer determines the features in the patterns in the input image, and is very useful in analyzing multi-dimensional data (e.g. images) [9].

*2) Pooling Layer:* The Pooling layer reduces the spatial dimensions of output volume. It does so by reducing the number of feature maps, and network parameters. It combines the output of neuron clusters at one layer into a single neuron in the successive layer. Global Pooling is a pooling layer which acts on all the neurons of the feature map. The two types of pooling layers which are popularly used are max pooling layer, and average pooling layer. The task of an average pooling layer is to take the average value of each local cluster of neurons in the feature map, while a max pooling layer takes the maximum value. Furthermore, pooling layer helps improve the generalization of the model by reducing overfitting [10]. The output of the pooling layer is then a combination of features which are invariant to translational shifts and distortions [11].

*3) Dropout:* One of the main problems in neural networks is overfitting. It is because a fully connected layer occupies most of the parameters. Dropout is a method that reduces overfitting [12]. It introduces regularization within a network, which in turn improves generalization. It does so by avoiding training all nodes on all training data, therefore decreasing overfitting. This method also improves the training speed of the algorithm.

*4) Fully Connected Layer:* After passing the input through the convolutional and pooling layers, the final classification of the input is done through the fully connected layer. The neurons inside the fully connected layer have connections with all the activations in the previous layer. Hence, it outputs nonlinear transformed output via an activation function. As discussed above, it works on the features from all stages and produce a nonlinear set of classification features. The rectified linear unit (ReLU) was used in this step as it helps in overcoming the vanishing gradient problem [13].

TABLE III
The results from the 2D Sequential CNN [4]

| **Dataset** | **Accuracy** | **Senstivity** | **Specificity** | **F1-Score** | Precision |
|---|---|---|---|---|---|
| Public Dataset | 96.1% | 92.8% | 97.8% | 94.2% | 95.7% |
| Fused Dataset | 93.7% | 85.7% | 99.7% | 92.1% | 99.5% |
| Public Dataset for training and Local Dataset for testing | 79% | 62.8% | 97.8% | 76.2% | 97.1% |
| Fused Dataset for training and Local Dataset for testing | 89.3% | 80.4% | 99.7% | 89% | 99.7% |

### B. Pre-trained Models

The rapid development in Computer Vision, and Image Recognition has introduced the concept if Transfer Learning, which is a machine learning method where the application of knowledge obtained from a model used in one task, can be reused as a foundation point for another task. In simple words, Transfer learning allows the use of pre-trained models which are already trained on huge and quality data sets. They are more accurate and they reduce the need of training a new model from scratch therefore reducing time and cost which makes it an effective method to be used in image recognition.

The pre-trained models discussed above are state-of-the-art models, and they are used widely in image recognition industry.

*1) Evaluation Metrics:* There are different metrics which can be used for evaluating the performance of classification models, such as classification accuracy, sensitivity, specificity, precision, and F1-score as defined in the Equations $(1)$-$(5)$.

$$\text{Accuracy} = \frac{t_p + t_n}{t_p + f_n + f_p + t_n} \quad (1)$$

$$\text{Sensitivity} = \frac{t_p}{t_p + f_n} \quad (2)$$

$$\text{Specificity} = \frac{t_n}{t_n + f_p} \quad (3)$$

$$\text{F1score} = \frac{2 \times t_p}{2 \times t_p + f_p + f_n} \quad (4)$$

$$\text{Precision} = \frac{t_p}{t_p + f_p} \quad (5)$$

where $t_p$: true positive, denotes participants that have been appropriately classified into a predetermined (positive) category. $f_n$: false negative, which denotes subjects who were

incorrectly categorised in the other (negative) class. $f_p$: false positive, which denotes patients who have been misclassified into a predetermined (positive) category. $t_n$: true negative, denotes participants that were accurately classified in the other (negative) category.

### C. VGG-16

VGG-16 [14] is the one of the most popular pre-trained models that is used in the image recognition industry for image classification. It was developed by a research team (Visual Graphics Group) in University of Oxford. It was first introduced in 2014 and has been refining and improving its performance since.

The data set used to train VGG is ImageNet. ImageNet is data set which contains over 15 million high resolution labelled images. It has over 22,000 categories. The images can further be classified into training images, validation images, and testing images. Therefore, these images can be used for all the purposes of training, validating, and testing a CNN. ImageNet consists of variable resolution. These images are down-sampled to a fixed resolution of $256 \times 256$ to train the VGG-16 algorithm. In Total, there are 13 Convolutional Layers, 5 Pooling Layers, and 3 fully Connected Layers in the architecture of VGG as shown in Figure 8.

The input is a $224 \times 224$ RGB image. The image is passed through a stack of convolutional layers. The Conv1 layer which contains two convolutional layers (Conv1-1, and Conv1-2) followed by a max pooling layer. The filters (64 filters in both Conv1-1, and Conv1-2) used in the convolutional layer have a very small receptive field $(3 \times 3)$. Then the image filters are doubled to 128, and the image is then passed through the second convolutional layers which has 2 more convolutional layers (Conv2-1, and Conv2-2) inside them, and then again, the layer is succeeded by a max pooling layer. In a similar way, Convolutional Layer 3 which contains three sub convolutional layers (The filters are again doubled to 256) and a max pooling layer, followed by Convolutional Layer 4 which also contains three sub convolutional layer (The filters are doubled to 512), also followed by a max pooling layer, followed again by a Convolutional Layer 5 which is same as Convolutional Layer 4.

The output dimensions are now $7 \times 7$, and the output of the Conv5 is flattened to generate a feature vector. Three Fully Connected (FC) Layers follow the stack of convolutional layers. The first two FC layers generate a feature vector of 1x4096, that is they have 4096 channels each. The third FC layer has 4096 nodes, and it creates 1000 channels for 1000 classes, which is then passed to the activation function. Note that all hidden layers are equipped with the rectification (ReLU) non-linearity. The model is known for its simplicity [15].

Natheer et al. used VGG-16 model to detect COVID 19 with an accuracy of 99% and an F1 Score of 99.1% [4]. They locally collected CXR images from 368 confirmed COVID-19 patients. Data from three freely available datasets was also utilized. The performance was graded in four different ways.

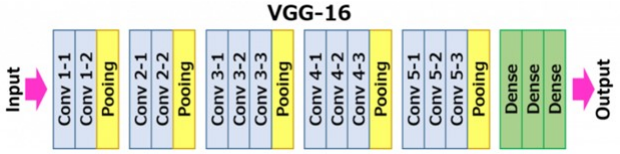


Fig. 8. An intuitive layout of the VGG-16 Model.

TABLE IV
PERFORMANCE EVALUATION METRICS FOR THE VGG-16 MODEL

| Dataset | Accuracy | Senstivity | Specificity | F1-Score | Precision |
|---|---|---|---|---|---|
| Public Dataset | 97.1% | 98.2% | 96.5% | 95.9% | 93.7% |
| Fused Dataset | 98.7% | 99.2% | 98.4% | 98.5% | 97.9% |
| Public Dataset for training and Local Dataset for testing | 97.2% | 97.8% | 96.5% | 97.4% | 97% |
| Fused Dataset for training and Local Dataset for testing | 99% | 99.5% | 98.4% | 99.1% | 98.7% |

For training and testing, the public dataset was used first. Second, data from both public and private sources was used to train and evaluate the models. Third, the model was trained using the public dataset, while the local data were solely used for testing. Finally, the combined data was used for training, while the local dataset was used for testing.

Table 3 shows the values of performance evaluations that resulted from the VGG-16 Model. Note that the model achieved its best accuracy of 99% when the fused data set was used for training, and the local data set was used for testing (Fourth Row of the Table 3).

The corresponding confusion matrices in Table 4 show how much of actual positive cases were predicted to be positive by the model, and how much of the actual negative cases were predicted to be negative by the model. It also shows Type I and Type II errors. Type I error is made by the model when it predicts a case to be negative COVID 19, while its actually positive. Similarly, Type II error is the when the model predicts a case to be positive while its actually negative.

Horry et al. used also used VGG-16 and VGG-19 [18] (contains 19 layers) models to detect COVID 19 with recall and precision rate both equal to 80% [16]. They obtained CXRs from publicly available COVID 19 Image Data Collection [17]. This collection of images is of various size and quality. Since the images were of various sizes and quality, it was first pre-processed before the model would be trained.

They did 2 experiments. The first experiment was of Normal vs COVID 19 and Pneumonia. The data set used for this experiment contained 200 Normal CXRs vs 100 COVID 19 CXRs, and 100 Pneumonia CXRs.

The second experiment was of COVID 19 vs Pneumonia. The data set used for this experiment were 100 pneumonia CXRs vs 100 COVID 19 CXRs.

The results are summarized below in Table 6.

Given the existing unavailability of high-quality COVID-19 X-Ray pictures, VGG16 or VGG19 classifiers were thought to be the most accurate off-the-shelf options for separating COVID-19 and Pneumonia from normal conditions, as well as COVID19 versus Pneumonia. In Experiment 1, both VGG16 and VGG19 did exceptionally well, with an average F1 score

TABLE V
THE CONFUSION MATRICES RESULTING FROM THE VGG 16 MODEL

| | | | |
|---|---|---|---|
| **(a) Public Dataset** | | | |
| | | Predicted Diagnosis | |
| | | Positive | Negative |
| Actual | Positive | 164 | 3 |
| | Negative | 11 | 306 |
| **(b) Fused Public and Local Datasets** | | | |
| | | Predicted Diagnosis | |
| | | Positive | Negative |
| Actual | Positive | 236 | 2 |
| | Negative | 5 | 312 |
| **(c) Public Dataset for Training and Local Dataset for Testing** | | | |
| | | Predicted Diagnosis | |
| | | Positive | Negative |
| Actual | Positive | 360 | 8 |
| | Negative | 11 | 306 |
| **(d) Fused Dataset for Training and Local Dataset for Testing** | | | |
| | | Predicted Diagnosis | |
| | | Positive | Negative |
| Actual | Positive | 366 | 2 |
| | Negative | 5 | 312 |

TABLE VI
THE RESULTS FROM THE EXPERIMENT CONDUCTED BY HORRY ET AL. [16]

| Experiment ID | VGG-16 | | | VGG-19 | | |
|---|---|---|---|---|---|---|
| Performance Parameters | P | R | F1 | P | R | F1 |
| Experiment 1 | 82 | 80 | 80 | 83 | 80 | 80 |
| Experiment 2 | 83 | 81 | 80 | 83 | 81 | 81 |

of 80% in differentiating between normal and COVID-19 or Pneumonia diseases. In experiment 2, VGG19 marginally outperformed VGG16, with a 1% increase in F1 score after an additional 20 training cycles, as evidenced by the original VGG19 training curve for experiment 2. We expect the VGG-19 classifier to outperform the VGG-16 classifier as more high-quality COVID-19 X-Ray image samples become available. For this reason, VGG-19 classifier was recommended over the VGG-16 classifier.

*1) ResNet:* ResNet stands for Residual Network. It is an innovative neural network that was first introduced in 2015 [19]. This model was immensely successful and quickly rose to popularity in the image recognition industry. ResNet comes in a variety of types, each with a different number of layers but the same basic concept. The term Resnet50 refers to a variant that has 50 neural network layers.

Machine learning professionals add extra layers when working with deep convolutional neural networks to tackle an issue in computer vision. These additional layers aid in the faster resolution of complicated problems since the individual layers can be trained for different jobs to produce highly accurate outcomes.

While the number of stacked layers might enhance the model's features, a deeper network can reveal the degradation issue. To put it another way, as the number of layers in a neural network grows, the accuracy levels may become saturated and gradually decline. As a result, the model's performance degrades on both training and testing data.

TABLE VII
THE RESULTS OF RESNET-50 FROM THE EXPERIMENT CONDUCTED BY HORRY ET AL. [16]

| Experiment ID | ResNet-50 | | |
|---|---|---|---|
| Performance Parameters | P | R | F1 |
| Experiment 1 | 70 | 67 | 66 |
| Experiment 2 | 67 | 58 | 51 |

ResNet was built specifically to address this issue. Residual blocks are used in deep residual nets to increase model accuracy. The strength of this form of neural network is the concept of "skip connections,", or "shortcut connections" which is at the heart of the residual blocks.

The first ResNet architecture introduced was the ResNet-34, which involved the introduction of shortcut connections, transitioning a plain network into a residual network. The plain network was influenced by VGG neural networks (VGG-16, VGG-19) in this case, with 3 filters in the convolutional networks. ResNets, on the other hand, have fewer filters and are less sophisticated than VGGNets.

It also followed two simple design rules: each layer had the same number of filters for the same output feature map size, and the number of filters was doubled if the output feature map size was halved to keep the time complexity per layer the same. It was made up of 34 layers, each of which was weighted.


Fig. 9. A picture showing the concept of Skip Connection.

Horry et al. used also used the ResNet50 [20] model to detect COVID 19 with recall and precision rates of 58%, and 67% respectively [16]. Similarly, as discussed above, they obtained CXRs from publicly available COVID 19 Image Data Collection [17]. This collection of images is of various size and quality. Since the images were of various sizes and quality, it was first pre-processed before the model would be trained.

They did 2 experiments. The first experiment was of Normal vs COVID 19 and Pneumonia. The data set used for this experiment contained 200 Normal CXRs vs 100 COVID 19 CXRs, and 100 Pneumonia CXRs.

The second experiment was of COVID 19 vs Pneumonia. The data set used for this experiment were 100 pneumonia CXRs vs 100 COVID 19 CXRs.

The results are summarized below in Table 7.

TABLE VIII
NUMBER OF IMAGES PER CATEGORY IN COVID-XRAY-5K DATASET

| Split | COVID-19 | Non-COVID-19 |
|---|---|---|
| Training Set | 84 (420 After Augmentation) | 2000 |
| Test Set | 100 | 3000 |

TABLE IX
SENSITIVITY AND SPECIFICITY RATES OF RESNET18 MODEL, FOR DIFFERENT THRESHOLD VALUES.

| Threshold | Senstivity | Specificity |
|---|---|---|
| 0.15 | 100% | 72.4% |
| 0.205 | 98% | 90.7% |
| 0.25 | 95% | 92.4% |
| 0.3 | 91% | 95.8% |
| 0.35 | 85% | 98.3% |

ResNet50 performed poorly when compared to the shallower VGG16 and VGG19 networks. Noting our limited sample sizes (200 samples per class in Experiment 1 and 100 samples per class in Experiment 2), as well as the difficulty of training CNNs with tiny data sets as discussed by D'Souza et al [21].

Minaee et al. published a study on COVID-19 prediction in CXR imaging using transfer learning. They used ResNet-50, and ResNet-18 Models. They formed the COVID-X-ray-5k Dataset. Chest X-ray images from two datasets formed the COVID-Xray-5k dataset that contains 2,084 training and 3,100 test images. They contained 184 COVID-19 images all together, and 5000 Non-COVID Images. The number if images oer categoery in COVID-X-ray-5k Dataset is shown in Table 8.

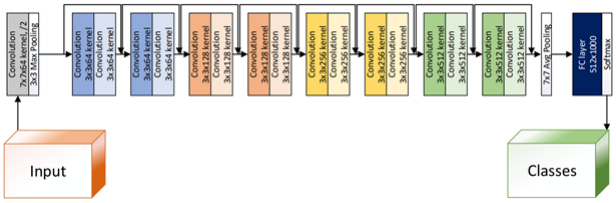


Fig. 10. The architecture of ResNet-18 model [22].

Each model generates a probability score that indicates the likelihood that the image is COVID-19. We may then compare these scores to a threshold to determine whether or not the image is COVID-19. The sensitivity and specificity of each model are calculated using the anticipated labels. We can acquire various sensitivity and specificity rates for each model depending on the value of the cut-off criterion.

The sensitivity and specificity rates for different thresholds are shown in Tables 9 and 10 using ResNet18 and ResNet50. As can be seen, all of these models produce encouraging results, with the top model achieving a sensitivity rate of 98 percent. It can be shown that ResNet18 outperforms ResNet50 by a small margin.

*2) Inception V3:* Inception v3 is a widely-used image recognition model that has been shown to attain greater than

TABLE X
SENSITIVITY AND SPECIFICITY RATES OF RESNET50 MODEL, FOR DIFFERENT THRESHOLD VALUES.

| Threshold | Senstivity | Specificity |
|---|---|---|
| 0.15 | 100% | 78.2% |
| 0.205 | 98% | 89.6% |
| 0.25 | 93% | 94.2% |
| 0.3 | 90% | 97.3% |
| 0.35 | 85% | 98.4% |

78.1% accuracy on the ImageNet dataset. The model is the culmination of many ideas developed by multiple researchers over the years [23]. It is basically a convolutional neural network (CNN) which is 27 layers deep. The inception layer is the core concept of a sparsely connected architecture. The Inception Layer is made up of all of those layers (11 Convolutional Layer, 33% Convolutional Layer, and 55% Convolutional Layer) with their output filter banks concatenated into a single output vector that serves as the input to the following stage. The original inception layer has two important add-ons in addition to the aforementioned layers: $1 \times 1$ Before applying another layer, apply a convolutional layer, which is mostly used for dimensionality reduction; a Max Pooling layer that runs in parallel to the inception layer and provides another alternative.

The author uses the Hebbian principle from human learning to explain the importance of the inception layer's structure which states that neurons that fire together, wire together. The author recommends that when building a following layer in a deep learning model, one should consider the preceding layer's learnings.

Assume that one of the layers in our deep learning model has learned to focus on specific features of a face. The network's following layer would most likely focus on the image's overall face to identify the various things there. Now, in order to accomplish this, the layer must have the necessary filter widths to identify various objects.

The conception layer takes the stage at this point. It allows the internal layers to select which filter size is appropriate for learning the required data. So, even if the size of the face in the image varies (as shown in the images below), the layer recognizes the face appropriately. The first image will most likely require a larger filter size, whilst the second image would require a smaller one.

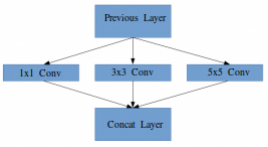


Fig. 11. Idea of an Inception Module

Horry et al. used also used the Inception V3 [24] model to detect COVID 19 with recall and precision rates of 65%, and 68% respectively [16]. Similarly, as discussed above, they

TABLE XI
THE RESULTS FROM THE EXPERIMENT CONDUCTED BY HORRY ET AL. [16]

| Experiment ID | Inception V3 | | |
|---|---|---|---|
| Performance Parameters | P | R | F1 |
| Experiment 1 | 68 | 65 | 61 |
| Experiment 2 | 41 | 50 | 36 |

obtained CXRs from publicly available COVID 19 Image Data Collection [17]. This collection of images is of various size and quality. Since the images were of various sizes and quality, it was first pre-processed before the model would be trained.

They did 2 experiments. The first experiment was of Normal vs COVID 19 and Pneumonia. The data set used for this experiment contained 200 Normal CXRs vs 100 COVID 19 CXRs, and 100 Pneumonia CXRs.

The second experiment was of COVID 19 vs Pneumonia. The data set used for this experiment were 100 pneumonia CXRs vs 100 COVID 19 CXRs.

The results are summarized below in Table 11.

Inception V3 also performed poorly when compared to the shallower VGG16 and VGG19 networks. Noting our limited sample sizes (200 samples per class in Experiment 1 and 100 samples per class in Experiment 2), as well as the difficulty of training CNNs with tiny data sets as discussed by D'Souza et al [21].

## IV. Available Web Applications Relating to Online Detection of COVID 19 through CXRs

In his paper, Gunther et al. showed the developed web application (link: https://toad.li/xray) based on his model based on our model to directly enable users to upload chest X-ray images and detect the presence of COVID-19 within a few seconds [27].

They provide a cutting-edge artificial intelligence-based system for efficient COVID-19 detection, as well as a user-friendly application that has the potential to become a speedy COVID-19 diagnosis method in the near future.

Unfortunately, the link is no longer active.

## V. Aims

The goal of this research is to train a model that can accurately recognize COVID 19 in CXR and then apply that model in a mobile or online application. Given that CXRs are readily available, one of the key aims of this project is to create a web application or a mobile application that can be easily utilized by the general public to detect the existence of viruses. Because the model will be available online, it can also assist doctors in diagnosing the infection.

For this objective, the model created should allow for the development of low-power, small-footprint, and low-latency applications for mobile and web environments.

## VI. Conclusion and Recommendations

Global calamities unite people and inspire new ideas. The current pandemic and its global ramifications should be seen as an opportunity to advance technical solutions that make ordinary living easier. This paper presented a survey-type study showing how radiographic pictures from COVID-19 patients in the hospital may be utilized to train models that can directly identify COVID 19 using a CXR. Researchers that have already created, trained, and tested deep learning artificial intelligence models employing the locally acquired dataset as well as public datasets, both independently and together, were also discussed. The high accuracy results open the door to developing mobile and easy-to-use apps that increase diagnosis accuracy, reduce burden for overworked health providers, and improve healthcare access in understaffed/underequipped locations. This area, as well as the creation and evaluation of multiclass classification models, will be the focus of future research.

The paper expansively discusses the importance of using the methodologies discussed in the paper and their types, advantages and disadvantages, the work done before and the work recently going on was discussed and the important points highlighted, this would be very helpful in terms of the knowledge required to successfully complete the project.

## References


[1] Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. Lancet 2020, 395, 497–506.

[2] Samrah, S.M.; Al-Mistarehi, A.H.W.; Ibnian, A.M.; Raffee, L.A.; Momany, S.M.; Al-Ali, M.; Hayajneh, W.A.; Yusef, D.H.; Awad, S.M.; Khassawneh, B.Y. COVID-19 outbreak in Jordan: Epidemiological features, clinical characteristics, and laboratory findings. Ann. Med. Surg. 2020, 57, 103–108.

[3] Hansell DM, Bankier AA, MacMahon H, et al. Fleischner Society: glossary of terms for thoracic imaging. Radiology. 2008;246(3):697–722. doi: 10.1148/radiol.2462070712. [PubMed] [CrossRef] [Google Scholar] Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] N. Khasawneh, M. Fraiwan, L. Fraiwan, B. Khassawneh, and A. Ibnian, "Detection of COVID-19 from Chest X-ray Images Using Deep Convolutional Neural Networks," Sensors, vol. 21, no. 17, p. 5940, Sep. 2021.

[5] Cozzi D, Albanesi M, Cavigli E, et al. Chest X-ray in new Coronavirus Disease 2019 (COVID-19) infection: findings and correlation with clinical outcome. Radiol Med. 2020;125(8):730-737. doi:10.1007/s11547-020-01232-9

[6] Greenfield, D. Artificial Intelligence in Medicine: Applications, Implications, and Limitations. 2019. Available online: https://sitn.hms.harvard.edu/flash/2019/artificial-intelligence-in-medicine-applications-implications-and-limitations/ (accessed on 25 May 2021).

[7] Nam JG, Park S, Hwang EJ, Lee JH, Jin KN, Lim KY, Vu TH, Sohn JH, Hwang S, Goo JM, Park CM. Development and Validation of Deep Learning-based Automatic Detection Algorithm for Malignant Pulmonary Nodules on Chest Radiographs. Radiology. 2019 Jan;290(1):218-228. doi: 10.1148/radiol.2018180237. Epub 2018 Sep 25. PMID: 30251934.

[8] Yun Liu, Timo Kohlberger, Mohammad Norouzi, George E. Dahl, Jenny L. Smith, Arash Mohtashamian, Niels Olson, Lily H. Peng, Jason D. Hipp, Martin C. Stumpe; Artificial Intelligence–Based Breast Cancer Nodal Metastasis Detection: Insights Into the Black Box for Pathologists. Arch Pathol Lab Med 1 July 2019; 143 (7): 859–868. doi: https://doi.org/10.5858/arpa.2018-0147-OA

[9] Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. 2020, 53, 5455–5516.

[10] Scherer, D.; Müller, A.; Behnke, S. Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition. In Artificial Neural Networks–ICANN 2010; Springer: Berlin/Heidelberg, Germany, 2010; pp. 92–101-10.
[11] Ranzato, M.; Huang, F.J.; Boureau, Y.L.; LeCun, Y. Unsupervised Learning ofw Invariant Feature Hierarchies with Applications to Object Recognition. In Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition, Minneapolis, MN, USA, 17–22 June 2007; pp. 1–8.
[12] Srivastava, Nitish; C. Geoffrey Hinton; Alex Krizhevsky; Ilya Sutskever; Ruslan Salakhutdinov (2014). "Dropout: A Simple Way to Prevent Neural Networks from overfitting" (PDF). Journal of Machine Learning Research. 15 (1): 1929–1958.
[13] Nwankpa, C.; Ijomah, W.; Gachagan, A.; Marshall, S. Activation Functions: Comparison of trends in Practice and Research for Deep Learning. arXiv 2018, arXiv:1811.03378.
[14] Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015.
[15] Khan, A.; Sohail, A.; Zahoora, U.; Qureshi, A.S. A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. 2020, 53, 5455–5516.
[16] M. J. Horry, M. Paul, A. Ulhaq, B. Pradhan, M. Saha, and N. Shukla, X-Ray Image Based COVID-19 Detection Using Pre-trained Deep Learning Models, MedRxiv, Cold Spring, NY, USA, 2020.
[17] J. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," arXiv.org, 2020.
[18] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv.org, 2015.
[19] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
[20] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," arXiv.org, 2015.
[21] R. N. D'Souza, P.Y. Huang, and F.C. Yeh, "Structural Analysis and Optimization of Convolutional Neural Networks with a Small Sample Size," Scientific reports, vol. 10, no. 1, pp. 834-834, 2020, doi: 10.1038/s41598-020-57866-2.
[22] He, Kaiming, et al. "Deep residual learning for image recognition." Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.
[23] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.
[24] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," vol. 2016-, ed: IEEE, 2016, pp. 2818-2826.
[25] Zhang W. Imaging changes of severe COVID-19 pneumonia in advanced stage. Intens. Care Med. (2020); 46(5):841–3
[26] Coronavirus and pneumonia. WebMD (2020); https://www.webmd.com/lung/covid-and-pneumonia1
[27] Gunther et al., " COVID-19 Chest X-Ray Image Classification Using Deep Learning," medRxiv 2021.07.15.21260605; doi: https://doi.org/10.1101/2021.07.15.21260605