Chest X-ray image-based Covid-19 Detection using Deep Learning Algorithm

G.Shanmukhi Rama
Assistant Professor, Chaitanya Bharathi Institute of Technology, Gandipet, Hyderabad, India

Abstract: **The entire world is dealing with the Coronavirus, which has shattered our lives in innumerable ways. The number of people affected due to the virus is increasing at a pace that has been never witnessed before, leading to a rise in the need for medical facilities and equipment. In the digital revolution we exist in, every field of life is majorly reliant on technology for every daily need of life. Our study focused on developing a deep learning model for the detection of Coronavirus from chest X-ray images.**

To detect Covid-19 data pre-processing would be performed on the raw data of X-ray images. This is the first and the most crucial step as it prepares the data by avoiding the noise and missing values by converting it into a usable format that can directly be used for machine learning models. As the performance of deep learning neural networks often improves with the amount of data available, data augmentation techniques are used to artificially create new training data from existing training data.

Machine learning and Computer Vision techniques are further being implemented for other various operations. One hot encoding is performed as it allows the representation of categorical data to be more expressive since most machine learning algorithms cannot work with categorical data directly. VGG 16, a convolutional neural network architecture, efficient image classification, and localization mechanisms are applied. The accuracy and loss curves are plotted using the metaplot libraries to better the efficient visualization of the model.

***Index Terms*–Covid-19,Convolutional Neural Network, Computer Vision, Chest X-ray images, Machine learning**

I. INTRODUCTION

The entire world is dealing with the Coronavirus, which has shattered our lives in innumerable ways. The number of people affected due to the virus is increasing at a pace that has been never witnessed before, leading to a rise in the need for medical facilities and equipment. There are limited kits for diagnosis, limited hospital beds for admission of such patients, limited personal protective equipment (PPE) for healthcare personnel, and limited ventilators. It is thus important to differentiate which patients with severe acute respiratory illness could have COVID-19 infection to efficiently utilize the limited resources. In this work, we propose using chest X-rays to detect COVID-19 infection in the patients' exhibiting symptoms of SARI. Using our tool one can classify a given X-Ray in one of the four classes: normal, bacterial pneumonia, viral pneumonia, and covid pneumonia. spacing.

II. PROCEDURE FOR PAPER SUBMISSION

The use of X-Ray has several advantages over conventional diagnostic tests: 1. X-ray imaging is much more widespread and cost-effective than conventional diagnostic tests. 2. Transfer of digital X-Ray images does not require any transportation from point of acquisition to the point of analysis, thus making the diagnostic process extremely quick. 3. Unlike CT scans, portable X-Ray machines also enable testing within an isolation ward itself, hence reducing the requirement of additional Personal Protective Equipment (PPE), an extremely scarce and valuable resource in this scenario. It also reduces the risk of hospital-acquired infection for the patients. The main contribution of this work is in proposing a novel deep neural network-based model for highly accurate detection of COVID-19 infection from the chest XRay images of the patients. Further, given the novelty of the virus, many of the radiologists themselves may not be familiar with all the nuances of the infection and may be lacking in the adequate expertise to make a highly accurate diagnosis. Therefore, this automated tool can serve as a guide for those at the forefront of this analysis

III. METHODOLOGIES

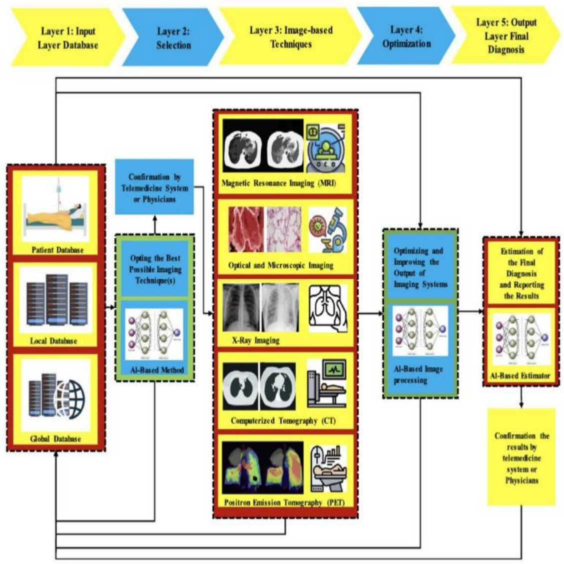

Figure 1.1 Overview of the methodology

Layer 1: Input Layer Database

● Designed for database access

Layer 2: Selection

● the task of adopting the best possible imaging techniques in the light of past experiences

● If physicians confirm the decisions made by this layer, the recommended techniques in the third layer take the required images.

Layer 3: Image-based techniques

● Magnetic Resonance Imaging (MRI), Computed Tomography Scan (CT scan), positron emission tomography (PET), Optical and Digital Microscopic Imaging Techniques and applications in Pathology, and X-Ray imaging are the techniques that may be used in the process.

Layer 4: Optimization

● dedicated to the optimization and improvement of the images.

Layer 5: Output Layer Final Diagnosis

● reserved for ultimate diagnosis based on the system's saved information and is a layer in which learning algorithms should be done by a CNN method

IV. DATASET DESCRIPTION

The dataset consists of chest X-rays of only Covid and normal patients. A total of 100 COVID-19 chest X-ray images and 100 normal chest X-ray images were obtained, they can be accessed from Kaggle. The ratio used to divide the dataset into training and testing sets was kept at 80:20. For training 160 records were used and 40 were used for testing out of a total of 200.

V. PROPOSED MODEL

The dataset consists of chest X-rays of only Covid and normal patients. A total of 100 COVID-19 chest X-ray images and 100 normal chest X-ray images were obtained, they can be accessed from Kaggle. The ratio used to divide the dataset into training and testing sets was kept at 80:20. For training 160 records were used and 40 were used for testing out of a total of 200.

Exploration of the various models that could be implemented

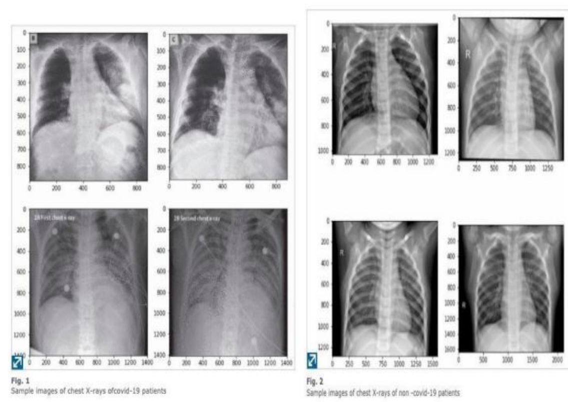

Figure 1.2: Sample X-ray images

In our, project we use the following technologies-Data preprocessing refers to cleaning the data and making it suitable for the model which also increases the accuracy and efficiency of a machine learning model.The X-ray images in the project are converted to RGB channel ordering and resized to 224*224 pixels such that it is ready for our CNN. We then scale pixel intensities in the range [0,1] and convert our data and labels to NumPy array format. In the available data, we divide 80% of the existing data into training data and the rest of the 20% into testing data. The performance of deep learning neural networks often improves with the amount of data available. Therefore, we implement Data Augmentation. Data augmentation is a technique to artificially create new training data from existing training data. This is done by applying domain-specific techniques to examples from the training data that create new and different training Transfer learning makes use of the knowledge gained while solving one problem and applying it to a different but related problem. Using Fine-Tuning we can give the new dataset to fine-tune the pre-trained CNN. Consider that the new dataset is almost similar to the original dataset used for pre-training. Since the new dataset is similar, the same weights can be used for extracting the features from the new dataset. VGG16 (also called OxfordNet) is a convolutional neural network architecture that is 16 layers deep named after the Visual Geometry Group from Oxford, which developed it. We instantiate the VGG16 network with weights pre-trained on ImageNet, leaving off the layer head. From there, we construct a new fully connected layer head and append it on top of VGG16. We then freeze certain weights of VGG16 such that only the layer head will be trained. This completes our fine-tuning setup.

A convolutional neural network (CNN) is a type of multilayer neural network containing two or more hidden layers. The hidden layers mainly perform two different kinds of functions: convolution and pooling. A convolution filter is useful to solve complex problems and generate feature maps. In the case of human activity data, for a particular time window,a continuous data stream of input activity data can be connected to a single hidden unit. Several features can be extracted using convolution. But a large set of features can lead to overfitting. Dimensions of the feature sets are reduced through the pooling mechanism.

Transfer learning is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task.Transfer Learning differs from traditional Machine Learning in that it is the use of pre-trained models that have been used for another task to jump-start the development process on a new task or problem. Transfer learning involves the concepts of a domain and a task. A domain DD consists of a feature space XX and a marginal probability distribution P(X)P(X) over the feature space, where X=x1,⋯ ,xn∈XX=x1,⋯ ,xn∈X. For document classification with a bag-of-words representation, XX is the space of all document representations, xxii isthe ii-the term vector corresponding to some document and XX is the sample of documents used for training.

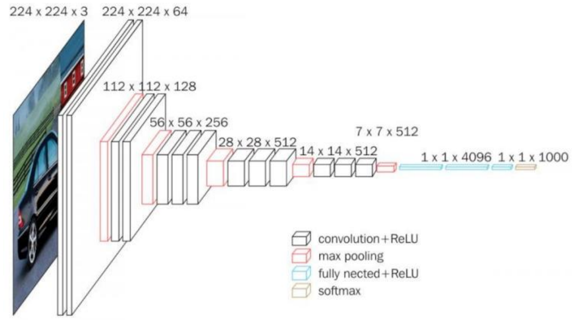

Figure 2.1: Transfer learning architecture

Existing solutions Chest X-Ray (CXR) is one of the important, non-invasive clinical adjuncts that play an essential role in the preliminary investigation of different pulmonary abnormalities. It can act as an alternative screening modality for the detection of nCOVID-19 or validate the related diagnosis, where the CXR images are interpreted by expert radiologists to look for infectious lesions associated with nCOVID-19. The earlier studies reveal that the infected patients exhibit distinct visual characteristics in CXR images. These characteristics typicallyinclude multi-focal, bilateral ground-glass opacities and patchy reticular (or reticulonodular) opacities in non-ICU patients, while dense pulmonary consolidations in ICU patients. However, the manual interpretation of these

Related Works Chest radiography is widely used for the detection and classification of Pneumonia and other pulmonary diseases. In the context of COVID-19 research, a closer look at the literature showed increased use of CXR images over CT scans due to potentially more data available from various sources. However, accurate annotation and analysis of radiography images require a radiology expert which requires significant expertise and processing time. To identify underlying features from radiography images for diagnostic analysis, a series of recent studies showed promising results using state-of-the-art computational and deep learning algorithms. We review the current literature on COVID-19 CXR image analysis using deep learning methods. Deep learning architectures are exceptionally utilized for the determination of pneumonia since 2016 the most explored deep learning strategies are VGG16, Inception_V2, and decision tree. We have picked these threetechniques because of the high outcome and accuracies they offer. Convolutional neural network architecture is one of the most popular and effective approaches in the diagnosis of COVD-19 from digitized images. Several reviews have been carried out to highlight recent contributions in assembling a dataset to train models for COVID-19 detection. In, a database of 190 COVID-19, 1345 viral Pneumonia, and 1341 normal chest X-ray images was introduced. Training and validation on four different pre-trained networks, namely, Resnet18, DenseNet201, AlexNet, and SqueezeNet for the classification of two different schemes (normal and COVID-19 Pneumonia; normal, viral Pneumonia, and COVID-19 Pneumonia). The classification accuracy for both schemes was 98.3% and 96.7% respectively. The sensitivity, specificity, and precision values were also reported. Having reviewed the related work, it is evident that despite the success of deep learning in the detection of Covid-19 from CXR and CT images, dealing with class imbalance, effectivefinetuning and validation of the models have not been explored. In this research, we aimed to extend the development of automated multi-class classification models based on chest X-ray images. For that, we created a balanced dataset and implemented an efficient and lightweight deep learning pipeline. We developed a Generative Adversarial Network (GAN) to generate synthetic COVID19 data and finally, we fine-tuned and optimized the hyper-parameters to improve the performance of the model.

DESIGN OF THE PROPOSED SYSTEM

Figure 3.1: Block diagram

Module Description

1 Extraction of COVID and normal dataset We create a dataset containing two folders, in which one has sampled X-Ray images of normal patients. We have taken around 100 sampled X-ray images of Normal patients. Then we create another folder in which we put the X-Ray images of coronavirus patients. Data analysis is implemented here. After creating two folders we will merge the images and set the labels. Then we will split that into training and testing sets and create a VGG model that will predict our data.

2 **Preprocessing and cleaning the data**the preprocessing of the text data is an essential step as it makes the raw text ready for mining, i.e., it becomes easier to extract information from the text and apply machine learning algorithms to it. The objective of this step is to clean noise and ensure the data is ready to be fed to the Convolutional Neural Network. Here, we clean the data, perform data augmentation and one-hot encoding on the dataset. Each image in the assembled dataset is resized to 224x224 pixels to reduce computation time and to maintain consistency throughout our dataset. Additionally, to account for the large variability of the image appearance (brightness and contrast), depending on the acquisition source, radiation dose, etc an image normalization stage has been applied. This stage normalizes and scales the pixel intensities to a range of [0, 255].

3 Extract features to analyze preprocessed data, it needs to be converted into features. The epochs and batch size are initialized as per the requirements for our learning mechanism. The labels and the features are extracted further.

4 **Model Building We instantiate the VGG16** network with weights pre-trained on ImageNet, leaving off the FC layer head. From there, we construct a new fully connected layer head consisting of POOL => FC = SOFTMAX layers and append it on top of VGG16. 20 We then freeze the CONV weights of VGG16 such that only the FC layer head will be trained. This completes our fine-tuning setup.

5 **Gaining Insights from the results**the model built in the previous module is evaluated against accuracy metrics. The sensitivity and specificity and other attributes are derived through the confusion matrix.

6 **Visualization of results**the insights from the results are represented visually by plotting graphs for loss and accuracy for testing and training data.

IMPLEMENTATION

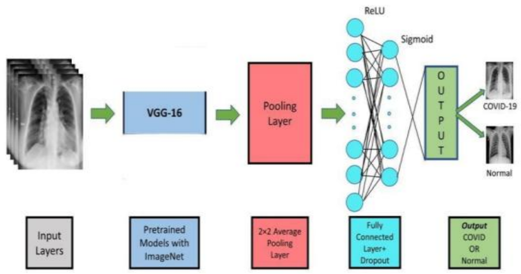

Figure 4.1: Work Flow diagram

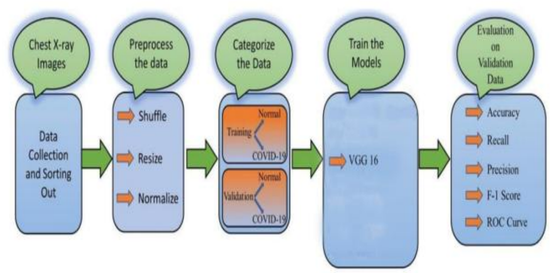

Figure 4.2: System Architecture at a glance

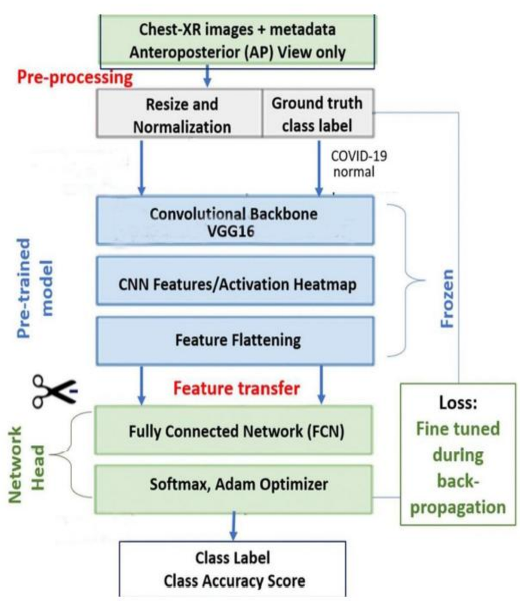

Figure 4.3: The model architecture of the system

Transfer learning stages • We used a pre-trained VGG16 as our feature extractor which was pre-trained on a huge ImageNet dataset, it has learned a good representation of low-level features like spatial, edges, rotation, lighting, shapes, and these features can be shared across to enable the knowledge transfer and act as a feature extractor for new images in different computer vision problems. • The new images have different categories from the source dataset, the pre-trained model is used to extract relevant features from these images based on the principles of transfer learning. • We used TensorFlow, Keras, PyTorch, Scikit-learn, and OpenCV libraries in Python for generating various functionalities of the pipeline.

Pre-trained model backbone and network head removal • We removed the network head or the final layers of the pre-trained model that was initially trained on the ImageNet dataset. • This stage is crucial as the pre-trained model was trained for a different classification task. • The removal of network head removed weights and bias associated with the class score at the predictor layers. • It is then replaced with new untrained layers with the desired number of classes in the new dataset. • We adjusted a network head for the COVID-19 dataset for two labels, (i.e), normal - for healthy patients, COVID-19 - for patients with COVID-19

Fine-tuning • At the initial stage, we froze the weights of the earlier layers of the pre-trained backbone to help us extract the generic low-level descriptors or patterns from the chest X-ray image data. • In the convolutional networks we used, the first few layers learn very simple and generic features that generalize to almost all types of images. As we went higher up,the features are increasingly more specific to the dataset on which the model was trained. • The goal of fine-tuning is to adapt these specialized features to work with the newly fed COVID-19 dataset, rather than overwrite the generic learning. • In the feature extraction experiment, we only trained a few layers on top of a base model. • The weights of the pre-trained network were not updated during training. • At this stage a newly added network head or a classifier is added with the desired number of classes and trained for adapting the weights according to the new patterns and distributions. • One way to increase performance even further is to train the weights of the top layers of the pre-trained model alongside the training of the classifier we added. • This stageforces the weights tobe tuned from generic feature maps to features associated specifically with the dataset

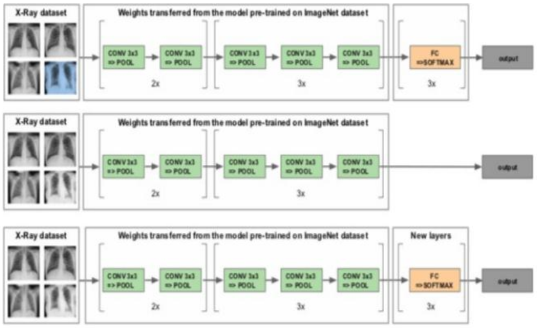

Figure 4.4: Fine-tuning in VGG 16

VI. RESULTS AND DISCUSSIONS

Classification Report We import a classification report from sklearn which builds a text report showing the main classification metrics.7 unique metrics were utilized to assess the proposed method. These metrics are precision, recall, F1 score, support,accuracy, macro avg, and weighted avg

Precision can be seen as a measure of a classifier's exactness. For each class, it is defined as the ratio of true positives to the sum of true and false positives. Said another way, "for all instances classified positive, what percent was Correct?"

$$\text{Precision} = \frac{tp}{tp + fp}$$

The recall is a measure of the classifier's completeness; the ability of a classifier to correctly find all positive instances. For each class, it is defined as the ratio of true positives to the sum of true positives and false negatives. Said another way, "for all instances that were positive, what percent was classified Correctly?"

$$\text{Recall} = \frac{tp}{tp + fn}$$

The F1 score is a weighted harmonic mean of precision and recalls such that the best score is 1.0 and the worst is 0.0. Generally speaking, F1 scores are lower than accuracy measures as they embed precision and recall into theircomputation. As a rule of thumb, the weighted average of F1should beused to compare classifier models, not global accuracy.

$$F_1 = \left(\frac{2}{\text{recall}^{-1} + \text{precision}^{-1}}\right) = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

Support is the number of actual occurrences of the class in the specified dataset. Imbalanced support in the training data may indicate structural weaknesses in the reported scores of the classifier and could indicate the need for stratified sampling or rebalancing. Support doesn't change between models t instead diagnoses the evaluation process.

	precision	recall	f1-score	support
Covid	0.95	1.00	0.98	20
Normal	1.00	0.95	0.97	20
accuracy			0.97	40
macro avg	0.98	0.97	0.97	40
weighted avg	0.98	0.97	0.97	40

Figure 5.1: Classification Report

A Confusion matrix is an N x N matrix used for evaluating the performance of a classification model, where N is the number of target classes. The matrix compares the actual target values with those predicted by the machine learning model. This gives us a holistic view of how well our classification model is performing and what kinds of errors it is making. Figure 5.2 represents the confusion matrix of VGG16 with ADAM Optimizer. The matrix contains the following values: TP, FP, FN, and TN

True Positive (TP)

● The predicted value matches the actual value ● The actual value was positive and the model predicted a positive value

True Negative (TN) ● The predicted value matches the actual value ● The actual value was negative and the model predicted a negative value

False Positive (FP) – Type 1 error ● The predicted value was falsely predicted ● The actual value was negative but the model predicted a positive value ● Also known as the Type 1 error

False Negative (FN) – Type 2 error ● The predicted value was falsely predicted ● The actual value was positive but the model predicted a negative value ● Also known as the Type 2 error

The accuracy of a test is its ability to differentiate the positive and normal cases correctly. To estimate the accuracy of a test, we should calculate the proportion of true positive and true negative in all evaluated cases. Mathematically, this can be stated as:

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

The sensitivity of a test is its ability to determine the positive cases correctly. To estimate it, we should calculate the proportion of true positive in positive cases. Mathematically, this can be stated as:

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

The specificity of a test is its ability to determine the normal cases correctly. To estimate it, we should calculate the proportion of true negatives in normal cases. Mathematically, this can be stated as:

$$\text{Specificity}=\frac{TN}{TN+FP}$$

```
[[20  0]
 [ 1 19]]
acc: 0.9750
sensitivity: 1.0000
specificity: 0.9500
```

Loss and Accuracy Curves :

The classification results are represented using a training curve and validation curve. From the training dataset, the training curve is calculated which represents how appropriately the prototype is learning, while the validation curve which is determined from the hold-out validation dataset is used for representing how properly the model is generalizing

Figures 5.3 and 5.4 represent the training and validation curve of VGG16 with ADAM optimizer

Loss Curve

It is noticed that in the train data loss curve is gradually decreasing from epoch 1 to the last epoch and the loss is approximately equal to 0.16 for epoch 10. The same is noticed for the validation data loss curve and the loss value for the last epoch is 0.13 approx.

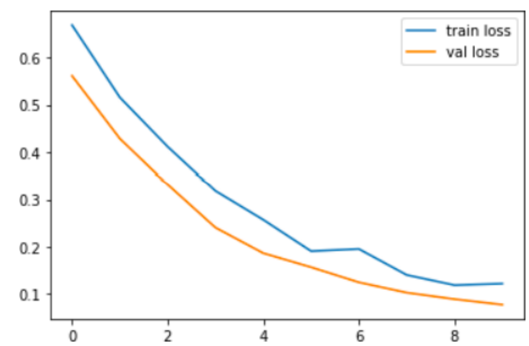

Figure 5.3 Model loss on each epoch for train vs test data

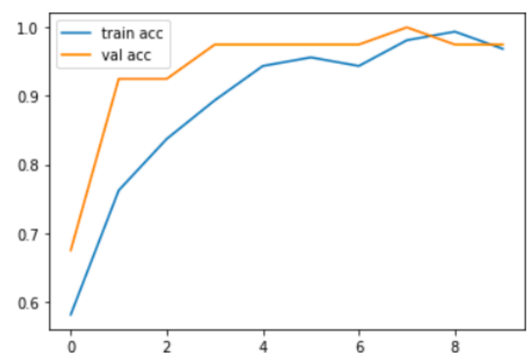

Figure 5.4 Model accuracy on each epoch for train vs test data

It can be noticed that the train data accuracy curve is gradually raising from epoch 1 to epoch 10 and then it converges to a value around 96.98%. Similarly, the validation accuracy curve is also gradually raising from epoch 1 to epoch 10 and reaches a value of approx. 96.55% for epoch 10.

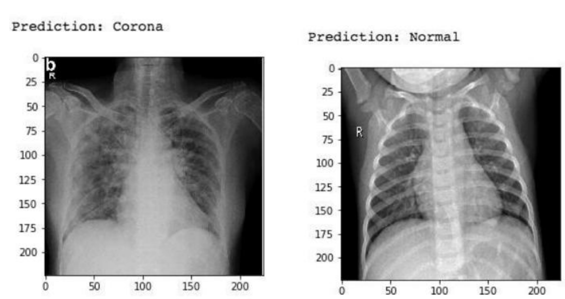

Figure 5.5: Result output X-rays in both scenarios

Fig. 5.5 shows the output of the code we have executed. As we can see, the first image represents the X-Ray of a person infected with Covid and the second image shows the X-Ray of a normal

VII.CONCLUSIONS AND FUTURE WORK

The COVID-19 infection induced by SARS-CoV2 is one of the recent viruses that loomed in late December 2019 in the Wuhan city of South China. This virus causes pneumonia that causes infection to the lung air sacs. RT-PCR is the basic approach used in the identification of COVID-19 infection. The alternative approach that is used in the efficient detection of infection is computed radiography images as RT-PCR is less sensitive for the determination of novel coronavirus at the initial stage. We proposed VGG16, which makes use of an ADAM optimizer for the automatic identification of the COVID-19 X-ray images from other X-ray images. Then, the efficiency of the proposed methodologyhas been enhanced by the application of data preprocessing, augmentation of data, and transfer learning approaches which are utilized to enhance the visual quality of the image as well as to overcome the overfitting problem. Precision, recall, f1score, and accuracy are applied to estimate the efficiency of the models. Since the Coronavirus is turning quite outgrowing and unpredictable with each coming day, there is an alarming need to come up with an efficient diagnosis of the disease. At the same time, since we are working on a medical need, we must ensure that the results areaccurate. Therefore, scope to further increase the accuracy is present. We can also implement approaches with more training data such that the project can be further enhanced. Exploration of newand other convolutional neural network models could provide even better performance. The results of our proposed CNN model performed surprisingly well with small datasets; however, it would be interesting to see its performance with a larger training dataset. Hence, in the future, we will collect a large number of images from various sources and analyze them to get more feasible outcomes. This approach may be helpful for clinical practices and detection of COVID-19 cases to prevent future community transmission

REFERENCES

[1] M. Jamshidi et al., "Artificial Intelligence and COVID-19: Deep Learning Approaches for Diagnosis and Treatment," in IEEE Access, vol. 8, pp. 109581-109595, 2020, DOI: 10.1109 /ACCESS.2020.3001973.

[2] R. Sethi, M. Mehrotra, and D. Sethi, "Deep Learning-based Diagnosis Recommendation for COVID-19 using Chest X-Rays Images," 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2020, pp. 1-4, DOI: 10.1109/ICIRCA48905.2020.9183278.

[3] E. F. Ohata et al., "Automatic detection of COVID-19 infection using chest X-ray images through transfer learning," in IEEE/CAA Journal of Automatica Sinica, vol. 8, no. 1, pp. 239-248, January 2021, DOI: 10.1109/JAS.2020.1003393.

[4] B. K. Umri, M. Wafa Akhyari and K. Kusrini, "Detection of Covid-19 in Chest X-ray Image using CLAHE and Convolutional Neural Network," 2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS), 2020, pp. 1-5, DOI: 10.1109/ICORIS 50180. 2020.9320806