Automatic detection of covid-19 disease using deep-convolution neural networks

Aravindan G
Dept. of Computer Science and Engineering Global Academy of Technology Bengaluru, India
aravindan.spark@gmail.com

Dr. Bhagyashri R Hanji
Dept. of Computer Science and Engineering Global Academy of Technology Bengaluru, India
bhagyashri@gat.ac.in

Ms.Sushmitha S
Dept. of Computer Science and Engineering Global Academy of Technology Bengaluru, India
sushmitha@gat.ac.in

Dr. R Kanagavalli
Dept. of Computer Science and Engineering Global Academy of Technology Bengaluru, India
kanagavalli@gat.ac.in

Abstract— The 2019 novel coronavirus (COVID-19), with a starting point in Wuhan, China, has spread rapidly among people living in other countries and is approaching approximately 4.5 million cases worldwide according to the statistics of the European Centre for Disease Prevention and Control. There are a limited number of COVID-19 test kits available in hospitals due to the increasing number of cases daily. Therefore, it is necessary to implement an automatic detection system as a quick alternative diagnosis option to prevent COVID-19 from spreading among people. In this study, three different convolutional neural network-based models (ResNet50, InceptionV3, and InceptionResNetV2) have been proposed for the detection of Coronavirus Pneumonia infected patients using chest X-ray radiographs. ROC analyses and confusion matrices by these three models are given and analyzed using 5-fold cross-validation. Considering the performance results obtained, it is seen that the pre-trained ResNet50 model provides the highest classification performance with 98% accuracy among the other two proposed models (97% accuracy for InceptionV3 and 87% accuracy for Inception-ResNetV2).
The result is based on the data available in the repository of GitHub, Kaggle, and Open-i as per their validated X-ray images.

Keywords— Covid-19, Coronavirus, Convolutional Neural Networks, Deep Learning methods, X-rays, CT scans, Pre-trained Models.

I. INTRODUCTION

The 2019 Novel Coronavirus (COVID-19) pandemic appeared in Wuhan, China in December 2019 and has become a serious public health problem worldwide since then. The virus that caused COVID-19 was called a severe acute respiratory syndrome coronavirus, also named SARS-CoV-2. Coronavirus (CoV) is a large family of viruses that cause diseases resulting from colds such as the Middle East Respiratory Syndrome (MERS-CoV) and Severe Acute Respiratory Syndrome (SARS-CoV). Coronavirus disease (COVID-19) is a new species that was discovered in 2019 and has not been previously identified in humans. Coronaviruses are zoonotic due to contamination from animals to humans. There are studies that the SARS-CoV virus is contaminated from musk cats to humans, and the MERS-CoV virus is contaminated from dromedary to humans. COVID-19 virus is presumed to be contaminated. from bats to

humans. Respiratory transmission of the disease from person to a person caused the rapid spread of the epidemic.

While COVID-19 causes milder symptoms in about 82 percent of cases, the others are severe or critical. Coronavirus cases the total number is approximately 170.47 million with 35L deaths and 1.8 million recovered patients as of today (01-06-2021). While 95% of the infected patients survive the disease slightly, 5% of the rest have a serious or critical condition.

Signs of infection include respiratory symptoms, fever, cough, and dyspnoea. In more serious cases, the infection can cause pneumonia, severe acute respiratory syndrome, septic shock, multi-organ failure, and death. It has been determined that men are more infected than women and that there is no death in children between the ages of 0-9. Respiratory rates of cases with COVID-19 pneumonia are faster compared to healthy people. Even in many developed countries, the health system has come to the point of collapse due to the increasing demand for intensive care units simultaneously. Intensive care units are filled with patients who get worse with COVID-19 pneumonia. The distribution of COVID-19 cases seen worldwide to date is shown in Figure 1.

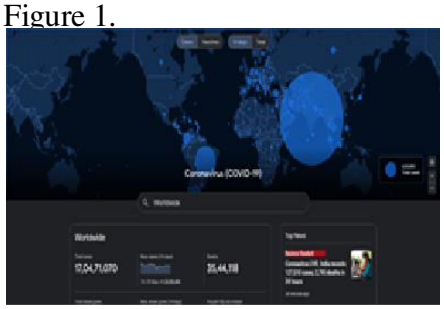

Figure 1: Total number of Covid-19 cases Worldwide (dated: 01-06-2021)

II. EASE OF USE

Instead of the patients waiting to get positive virus tests, diagnoses now include everyone who reveals the prominent pneumonia pattern of chest scan COVID-19. Through this method, authorities will be able to isolate and treat patients more quickly. Even if death does not occur in COVID-19, some patients survive with permanent lung damage.

According to the World Health Organization, COVID-19 also opens holes in the lungs like SARS, giving them a "honeycomb-like appearance". Computed Tomography (CT) scan of the chest is one of the methods used to diagnose pneumonia. Artificial Intelligence (AI) based automated CT image analysis tools for the detection, quantification, and monitoring of coronavirus and to distinguish patients with coronavirus from disease-free have been developed. This process is aimed to establish an early screening model to distinguish COVID-19 pneumonia and Influenza-A viral pneumonia from healthy cases using pulmonary CT images and deep learning techniques.

In Shuai et a, study, based on the COVID-19 radiographic changes from CT images, they have developed a deep learning method that can extract the graphical features of COVID-19 to provide a clinical diagnosis before pathogenic testing and thus save critical time for the disease diagnosis.

X-ray machines are used to scan the affected body such as fractures, bone dislocations, lung infections, pneumonia, and tumors. CT scanning is a kind of advanced X-ray machine that examines the very soft structure of the active body part and clearer images of the inner soft tissues and organs. Using X-ray is a faster, easier, cheaper, and less harmful method than CT. Failure to promptly recognize and treat COVID-19 pneumonia may lead to an increase in mortality.

In this study, we have proposed an automatic prediction of COVID-19 using deep convolution neural network-based pre-trained transfer models and Chest X-ray images. For this purpose, we have used

ResNet50, InceptionV3, and Inception-ResNetV2 pre-trained models to obtain a higher prediction accuracy for a small X-ray dataset. The novelty of this paper is summarized as follows:

i) The proposed models have an end-to-end structure without manual feature extraction and selection methods.
ii) We show that ResNet50 is an effective pre-trained model among the other two pre-trained models.
iii) Chest X-ray images are the best tool for the detection of COVID-19.
iv) The pre-trained models have been shown to yield very high results in the small dataset (50 COVID-19 vs. 50 Normal).

II. MATERIALS AND METHODS

In this study, chest X-ray images of 50 COVID-19 patients have been obtained from the open-source GitHub repository shared by Dr. Joseph Cohen [20]. This repository is consisting chest X-ray / CT images of main patients with Acute Respiratory Distress Syndrome (ARDS), COVID-19, Middle East Respiratory Syndrome (MERS), pneumonia, Severe Acute Respiratory Syndrome (SARS). In addition, 50 normal chest X-ray images were selected from the Kaggle repository called "Chest X-Ray Images (Pneumonia)". The analysis has been based on a created dataset with chest X-ray images of 50 normal and 50 COVID-19 patients (100 images in total). All images in this dataset were resized to 224x224 pixel size.
In Figure 2 and Figure 3, representative chest X-ray images of normal and COVID-19 patients are given, respectively.

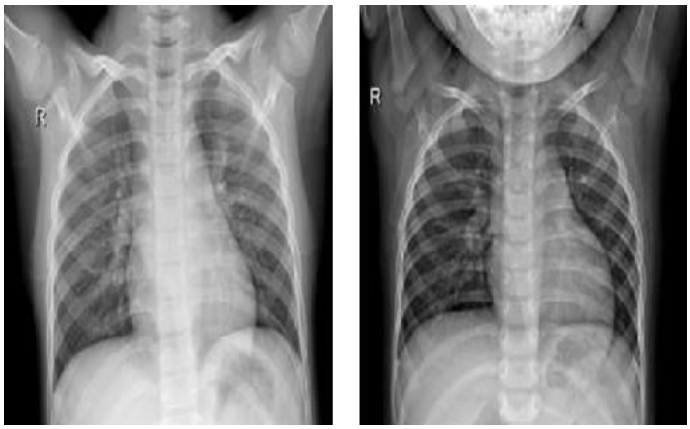

Figure 2: Representative Chest X-Ray images of a Normal Person

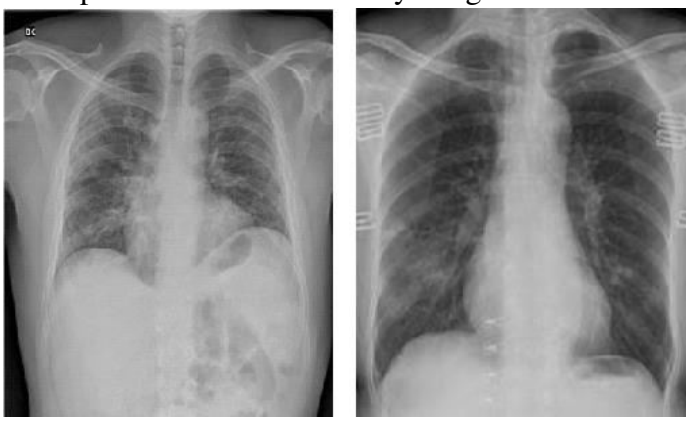

Figure 3: Representative Chest X-Ray images of COVID-19 patients

III. METHEDOLOGY

Deep feature extraction is based on the extraction of features acquired from a pre-trained CNN. The Deep features are extracted from a fully connected layer and feed to the classifier for training purposes. The deep features obtained from each CNN network are used by the SVM classifier. After that, the

classification is performed, and the performance of all classification models is measured. The rice leaf disease identification model based on deep features by the SVM classifier is shown in Figure 4.

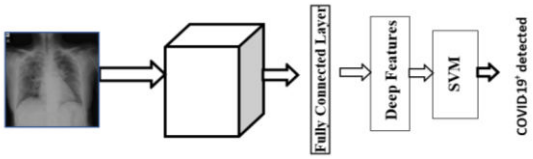

Figure 4: Detection Corona Virus by SVM based on Deep Feature using X-ray images.

In the study, Convolutional Neural Network (CNN) based ResNet50, InceptionV3, and Inception-ResNetV2 models for the classification of COVID-19 Chest X-ray images to normal and COVID-19 classes. In addition, we applied the transfer learning technique that was realized by using ImageNet data to overcome the insufficient data and training time. The schematic representation of conventional CNN including pre-trained ResNet50, InceptionV3, and Inception ResNetV2 models for the prediction of COVID-19 patients and normal were depicted in Figure 5.

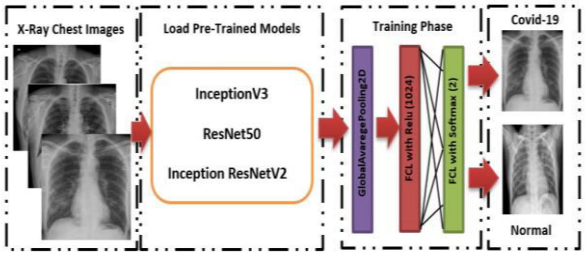

Figure 5: Schematic representation of pre-trained models for the prediction of COVID-19 patients and normal.

The Residual Neural Network (ResNet) model is an improved version of the Convolutional Neural Network (CNN). ResNet adds shortcuts between layers to solve a problem. Thanks to this, it prevents the distortion that occurs as the network gets deeper and more complex. In addition, bottleneck blocks are used to make training faster in the ResNet model. ResNet50 is a 50- layer network trained on the ImageNet dataset. ImageNet is an image database with more than 14 million images belonging to more than 20 thousand categories created for image recognition competitions. InceptionV3 is a kind of convolutional neural network model. It consists of numerous convolution and maximum pooling steps. In the last stage, it contains a fully connected neural network. As with the ResNet50 model, the network is trained with the ImageNet dataset. The model consists of a deep convolutional network using the Inception ResNetV2 architecture that was trained on the ImageNet-2012 dataset. The input to the model is a 299×299 image, and the output is a list of estimated class probabilities.

IV. RESULTS AND DISCUSSION

Chest X-ray images have been used for the prediction of coronavirus disease patients (COVID-19). Popular pre-trained models such as ResNet50, InceptionV3, and Inception ResNetV2 have been trained and tested on chest X-ray images.

Training accuracy and loss values for fold-3 of the pre-trained models are given in Figure 6 and Figure 7, respectively. The training stage has been carried out up to the 30th epoch to avoid overfitting for all pre-trained models.

It can be seen from Figure 6 that the highest training accuracy is obtained with the ResNet50 model. InceptionV3 and Inception-ResNetV2 models have similar performance. However, it is seen that ResNet50 shows a fast training process than other models. Although the pre-trained models give very high initial values, the initial values are below 70% due to the low number of data. The training loss values of ResNet50, InceptionV3, and Inception ResNetV2 are shown in Figure 7. When the loss figure is analyzed, it is seen that the loss values decrease in three pre-trained models during the training stage. It can be said that the ResNet 50 model both decreases loss values faster and approaches zero.

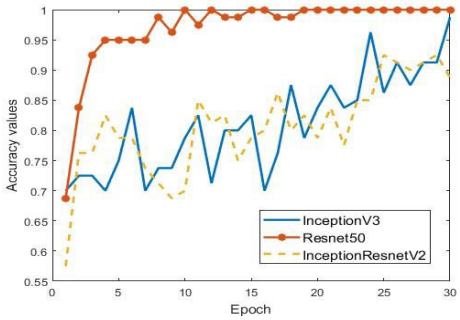

Figure 6: The performance of three pre-trained models.
(Training accuracy values)

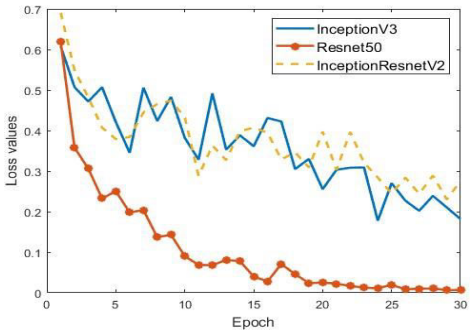

Figure 7: The performance of three pre-trained models.
(Training non-accuracy (loss) values)

In Figure 8, Confusion Matrices of COVID-19 and Normal test results of the models are posted from GitHub again. Firstly, the InceptionV3 pre-trained model classified 10 of the COVID-19 as True Positive for fold3 and classified 10 of the normal as True Negative. Secondly, the ResNet50 model also classified 10 of the COVID-19 as True Positive for fold-3 and classified 10 of the normal as True Negative. Lastly, Inception ResNetV2 classified 10 of the COVID-19 as True Positive for fold3 and classified 9 of the normal as True Negative. Besides the confusion matrix, receiver operating characteristic curve (ROC) plots and areas for each model are given. InceptionV3 and ResNet50 pre-trained models appear to be very high.

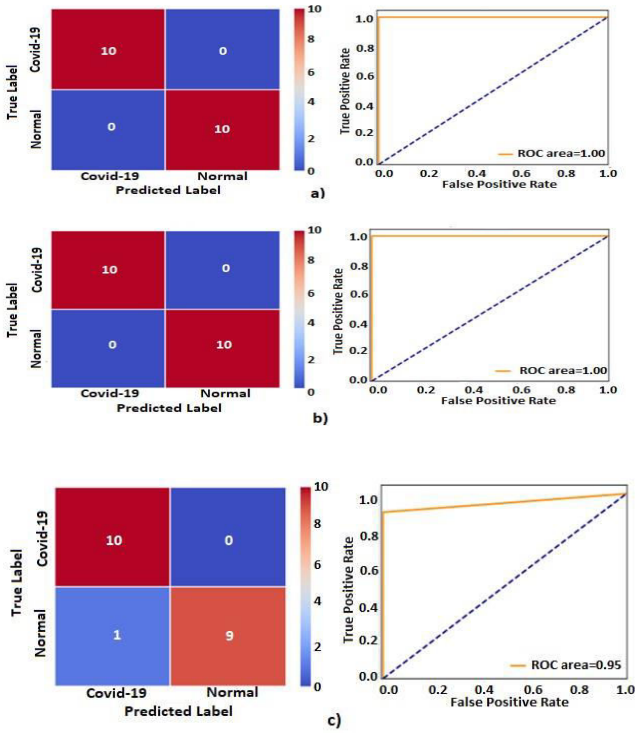

Figure 8: The confusion matrix and ROC plots obtained using pre-trained models results: InceptionV3 b) ResNet50 c) Inception-ResNetV2.

Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are

V. LITERATURE SURVEY

There are very few studies on literature due to the emergence of COVID-19 virus disease. Some of these are as follows:

- The analysis mentioned here is purely based on the X-rays / CT Scans shared on the GitHub repository, the experiment conducted by a group of 8 Students from Zonguldak Bulent Ecevit University, and the statistics available on Google.
- Prabira et al. (A Professor from Xiaowei University) proposed detection of COVID-19 using X-ray images based on deep feature and SVM. He collected X-ray images from GitHub, a repository. They extracted the deep feature of CNN models and fed it to the SVM classifier individually. He has obtained 95.38% of accuracy for ResNet50 & SVM.
- This method is a complete end-to-end system. So, it does not have any feature extraction or selection. Three different pre-trained common models are compared such as ResNet50, InceptionV3, and Inception-ResNetV2.
- Although it is a very new subject and the number of data is limited, the results are quite high.

VI. CONCLUSION

- Early prediction of COVID-19 patients is vital to prevent the spread of the disease to other people. In this study, we proposed a deep transfer learning-based approach using chest X-ray images obtained from COVID-19 patients and normal to predict COVID-19 patients automatically.
- Performance results show that the ResNet50 pre-trained model yielded the highest accuracy of 98% among the three models. In the light of our findings, it is believed that it will help doctors to make decisions in clinical practice due to the high performance.

- To detect COVID-19 at an early stage, this study gives insight into how deep transfer learning methods can be used. In subsequent studies, the classification performance of different CNN models can be tested by increasing the number of images in the dataset.

REFERENCES

[1] Open database of COVID-19 cases with chest X-ray or CT images https://github.com/ieee8023/covid-chestxray-dataset

[2] Chest X-Ray Images (Pneumonia)

[3] https://www.kaggle.com/paultimothymooney/chestxray-pneumonia

[4] www.who.int [Accessed: 13th June 2021]

[5] www.github.com [Accessed: 13th June 2021]

[6] www.kaggle.com [Accessed: 15th June 2021]

[7] https://openi.nlm.nih.gov [Accessed: 16th May 2021]

[8] https://www.nationalgeographic.com/science/2020/02/here-is-what-coronavirus-does-tothe-body/

[9] International committee on taxonomy of viruses (ICTV):

[10] https://talk. ictvonline.org/ [Accessed 14th May 2021]

[11] National health commission of the people's republic of china website - diagnosis and treatment protocols of pneumonia caused by a novel coronavirus (trial version 5). http://www.nhc.gov.cn/yzygj/s7653p/202002/d4b895337e19445f8d728fcaf1e3e13a.shtml

[12] https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

[13] https://biodifferences.com/difference-between-x-ray-and-ct-scan.html

[14] Sethy, P. K., and Behera, S. K. Detection of coronavirus Disease (COVID-19) based on Deep Features.

[15] Yildirim, O., Talo, M., Ay, B., Baloglu, U. B., Aydin, G., and Acharya, U. R. Automated detection of the diabetic subject using pre-trained 2D-CNN models with frequency spectrum images extracted from heart rate signals. Computers in Biology and Medicine, 2019.