International Journal of Innovative Research in Computer Science and Technology (IJIRCST)
ISSN (online): 2347-5552, Volume-11, Issue-6, November 2023
https://doi.org/10.55524/ijircst.2023.11.6.2
Article ID IRP1412, Pages 6-12
www.ijircst.org

Evaluating Median Accuracy of ResNet50 and VGG16 Models in COVID-19 Detection

Mallisetty Siva Mahesh[1], Kattamuri B N Ayyappa[2], Maddela Murali[3], Mididoddi Surendra Babu[4], and Nagababu Pachhala[5]

[1,2,3,4] Student, Department of Information Technology, Vasireddy Venkatadri Institute of Technology, Guntur, Nambur, India
[5]Assistant Professor, Department of Information Technology, Vasireddy Venkatadri Institute of Technology, Guntur, Nambur, India

Correspondence should be addressed to Mallisetty Siva Mahesh: sivamaheshmallisetty789@gmail.com

Received 7 October 2023; Revised 23 October 2023; Accepted 4 November 2023

ABSTRACT- The increasing number of Covid-19 cases and the lack of reliable, quick-to-use testing tools herald a new era in X-ray analysis employing deep learning methods. The Covid-19 virus's emergence poses a threat to human existence. Therefore, it will take time to develop a quick and accurate method of identifying the Covid-19 virus in patients. The reference method is the conventional RT-PCR technique. The goal of this research is to create an automated system that uses CNN models such as Resnet50, VGG16, and Grad-CAM to analyze X-ray images in order to provide a reliable and effective method of diagnosing Covid-19 infection. The created models use image processing techniques to pre-process the X-ray picture. Afterwards, deep learning is used to classify the images after they have been segmented and transformed. The CNN model that is being used provides strong classification accuracy and shows the location in the lung where the disease is attacked, even for a normal person, we can anticipate the likelihood of where the COVID may affect them. Our model utilizes a convolution neural network that is trained on the standard COVID-19 Radiography Dataset.

KEYWORDS- CNNResnet50,VGG16, Grad-CAM, Covid.

I. INTRODUCTION

Most of the world has already noticed the virus's existence by the early spring of 2020. The World Health Organization (WHO) formally classified the novel virus as a pandemic by the end of March 2020 [1]. According to WHO figures, as of April 30, 2021, the virus had impacted over 157 million people, and over 3 million people had died as a result [2].
Although corona viruses are not a novel occurrence, SARS-CoV-2 is not your typical corona virus [3]. This virus most likely came from an animal reservoir at some point [4]. Treatment for COVID-19 is administered differently than for other corona virus infections [5]. The disease is still not well understood, and as the pandemic spreads, so does our knowledge of it [6]. The COVID-19 virus is known to cause fever, coughing, exhaustion, sore throats, and body aches. Worldwide, there have been several reports of taste and smell loss. Rarer yet more severe instances could see patients struggle to breathe, have a high fever, chills, exhaustion, aches in their muscles, or possibly pass away [7].
Models of artificial intelligence (AI) offer a possible remedy [8]. With its remarkable accuracy, the deep learning method has been popular and successful in applications involving the classification of medical images. Recent developments in deep learning technology have aided in the creation of intelligent diagnostic tools that help medical professionals make well-informed judgments regarding the health of their patients. For example, Lopez et al.'s study [9] addressed the problem of classifying skin lesions, specifically the early detection of melanoma.

II. LITERATURE SURVEY

Several studies have used deep learning on chest X-ray images to identify COVID-19 patients. Here, we focus only on those that are directly relevant to our proposal.
Ioannis et al. [1] evaluate the effectiveness of the most recent CNN architectures—VGG19, Grad-CAM, Inception-ResNetv2, Xception, and MobileNet v2—that have been employed in medical image categorization in recent years. Because deep learning is effective at identifying a variety of anomalies in small medical picture data sets, the author employs it [2]. They made use of 1,442 patient X-ray data sets, comprising 504 healthy individuals, 224 cases of proven COVID-19 illness, and 714 cases each of bacterial and viral pneumonia. Based on the results, it can be concluded that MobileNet-v2 and VGG19 provide the most accurate classification among the remaining CNNs. MobileNet-v2 performs better in terms of sensitivity and specificity (reaching 99.10 and 97.09 percent, respectively), but VGG19 exceeds the other techniques in terms of accuracy (reaching 98.75 percent). The issue of the limited number of COVID-19 test kits available in public hospitals was resolved by Narin and associates. suggests using an automatic detection system as an additional quick diagnostic tool to stop COVID-19 from spreading and putting strain on healthcare facilities. The author presented three CNN-based models (Inception-ResNetV2, Grad-CAM, and ResNetV2) employing a total of 100 chest X-ray

pictures (50 health images and 50 COVID-19 images) to identify patients with corona virus pneumonia. Given the excellent performance outcomes, we can draw the conclusion that the pretrained ResNet50 model outperforms the other two suggested models, achieving 98% accuracy (Grad-CAM earns 97% accuracy, and Inception-ResNetV2 only achieves 87% accuracy).

Ozturk et al. [3] suggest a new methodology that uses raw radiography images to automatically detect COVID-19 in order to accurately detect the virus and help address the shortage of specialized doctors in distant regions. The real-time object detection system (YOLO), which is composed of 17 convolutional layers, uses the DarkNet model among the suggested models as its classifier. The authors filter at the layer level in a different way. Offering an accurate diagnosis for both a two-class classification (COVID/no-findings) and a multi-class classification (COVID/no-findings/pneumonia) is the aim of this method. 97.08% of the binary classification and 87.02% of the multi-classification have accurate classifications.

III. METHODOLOGY

In this study, we train, assess, and test three popular pre-trained deep learning architectures for the purpose of classifying images from chest radiography into two classes: Normal chest X-ray and COVID-19. Deep learning models that are widely employed are VGG16, ResNet50, and Grad-CAM.

A broad overview of the study's methodology is given in Figure 1, which details the deep learning approach that is suggested based on a straightforward standard pipeline, including the pre-processing of chest images and the subsequent deep learning-derived classification model. A deep model is trained following the pre-processing of the data. For example, we will fine-tune the model by unfreezing some of the top layers of the frozen model, which is used for feature extraction, and then training the newly added component (the fully-connected classifier in our experiment) as well as the unfreeze top layers.

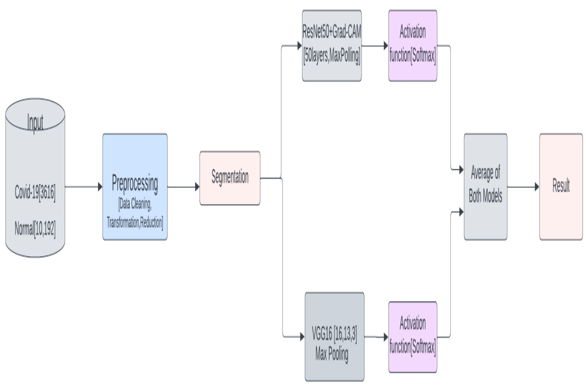


Figure 1: Overview for Covid-19 and Non-Covid-19 Chest X-Ray Images Classification

IV. SYSTEM IMPLEMENTATION

A. Dataset

A.1 Creation

Since preparing the dataset is the initial step in using deep learning, we begin there. We examine a patient's lung health using a chest X-ray image as COVID-19 targets the epithelial cells lining our airways.

Rather of using computer tomography scans, we use chest X-ray pictures in this work to refine the three suggested classification models. X-rays are far quicker, less expensive, require fewer radiation dosages for the patient, and are more readily available than CT scans, which require greater radiation exposure and take longer. Furthermore, portable X-ray devices can be tested in isolation units, which lowers the need for personal protective equipment and lowers the danger of hospital infections.

Moreover, the PCR approach can be practically substituted with chest X-ray image analysis. From the disease's discovery to the identification of high-risk patients for isolation and priority, as well as selective testing to find false-negative PCR cases, they can offer a range of support. However, because the majority of viral pneumonia cases are identical and overlap, it takes a lot of time and effort for radiologists and medical professionals to visually distinguish essential facts. Using deep learning models may provide a precise answer.

Our goal in this experiment is to minimize false positives and false negatives by applying three Convolutional Neural Networks (CNNs) to our collection of chest X-ray pictures as part of a deep learning process. There are 5000 photos in the dataset that was created for this project:

A total of 10,000 photos for a normal chest X-ray were chosen from image databases, specifically the “Covid-19 Radiography Dataset” Kaggle repository [19].

From the Covid-19 Radiography Data Set [19], 3653 chest X-ray COVID-19 pictures were gathered. The prepared dataset, which was separated into the Normal and COVID-19 folders, is shown in Table 1 below. Figure 2 illustrates two examples of chest X-ray images taken from our prepared dataset.

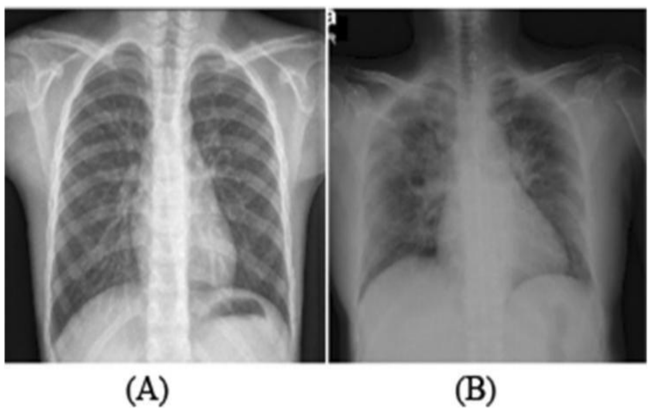


Figure 2: (A) Chest X-Ray Image of a Normal Person (B): Covid-19 Chest X-Ray Image

A.2 Data Preprocessing

Resizing the X-ray pictures is achievable during the data pre-processing phase. This is because different image inputs are needed for different algorithms. Normalizing the photos in accordance with the specified model criteria is necessary. The input photos were initially of varying sizes; after processing, they were all converted to a consistent 224 × 224-pixel size.

B. The Proposed Framework

It might not be able to create a CNN model from scratch to automatically identify COVID-19 from X-ray images due to the lack of free COVID-19 radiography data. We use certain deep learning models to control this issue, and we use the supplied data set to fine-tune three popular pre-trained models.

B.1 COVID-19 Detection using VGG16

The Simonyan et al. team first suggested the VGGNet architecture in 2014 under the title "Very Deep Convolutional Networks for Large-scale Image Recognition" [27]. Three-by-three convolutional layers stacked one on top of the other and increasing depth are the distinguishing features of VGG series networks. A maximum pooling is used to treat reducing the volume size. The VGG16 architecture is composed as following:

- Two Convolutional layers with 64 filters followed by Max pooling layer
- Two Convolutional layers with 128 filters followed by Max pooling layer
- Three Convolutional layers with 256 filters followed by Max pooling layer
- Two stack each with 3 convolutional layers with 512 filters and separated by a max pooling layer. A final Max pooling layer Two fully connected layers with 4096 channels
- Softmax output layer with 1000 classes

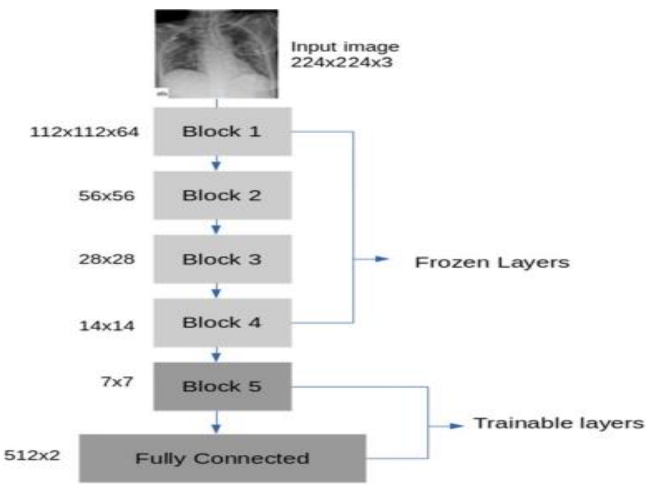


Figure 3: Proposed VGG16 Architecture

B.2 COVID-19 Detection Using ResNet50

Compared to VGG16, ResNet-50 is a deeper CNN with 50 layers. ResNet50's model size is only 102 MB because a global average pool rather than a fully connected layer is used, making the model substantially smaller overall [29]. The residual block learning component of ResNet is unique. This implies that each layer should feed into the one behind it as well as onto levels that are two to three hops away. The following makes up its architecture: A convolutional layer with 64 filters and kernel size of 7 × 7. This is followed by a max pooling layer with a stride size of 2.

- Then, a convolutional layer with 64 filters and a kernel size of 1 * 1, followed by a second convolutional layer with 64 filters and a kernel size of 3 * 3. Then, we have another convolutional layer with 256 filters and a kernel size of 1 * 1. These 3 layers are replicated in total 3 time and 9 layers are obtained at this stage.
- Next, 3 convolutional layers, the first one is with 128 filters and a kernel size of 1 * 1, the second one is with 128 filters and kernel size of 3 * 3, and the third one is with 512 filters and a kernel size of 1 * 1. These layers are replicated 4 time to give us 12 layers at this stage.
- Afterwards, we have convolutional layer with 256 filters and a kernel size of 1 * 1, and two others with 256, 1024 filters and a kernel size of 3 * 3, 1 * 1. This is replicated 6 time to give us totally 18 layers.
- Then, we have a convolutional layer with 512 filters and a kernel size of 1 * 1, with two others with 512, 2048 and a kernel size of 1 * 1, 3 * 3. This is replicated 3 times to give us totally 9 layers.

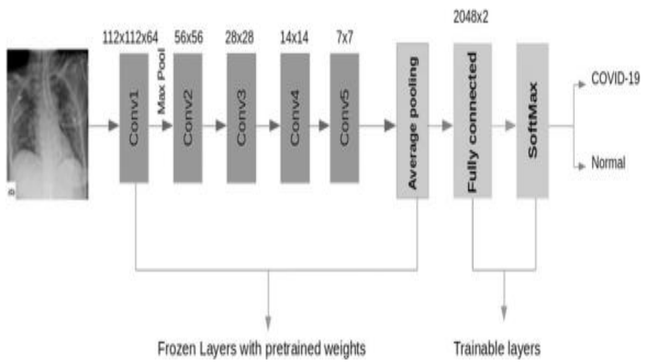


Figure 4: Proposed Resnet50 Architecture

Finally, we apply an average pooling and finish it with a fully connected layer (with 1000 nodes) and then a softmax function to give us 1 layer as a final stage.

V. EXPERIMENTAL RESULTS

The Rectified Linear Unit activation function activates each of the buried levels. The form of the input is [224, 224, 3]. We adjust every model over a span of 25 epochs. We utilize ADAM as the loss function, set the learning rate to 0.0001, and batch size to 32. We made the decision to apply a cross-entropy loss function to every model.

The dataset, which is explained in Section 3.1, is split into two groups: training (80%) and validation (20%). The former is used for training, while the latter is utilized to test the assessment at the end. The performance is evaluated using the following six performance criteria: F1 score, recall, accuracy, specificity, sensitivity, and precision. These outcomes were attained.

A. Training Results

In Figures 6 and 7 the training outcomes of the suggested models are documented and displayed as graphs. The blue curve represents training, and the orange curve represents validation.

Each image highlights the fact that the training accuracy rate for all three models is as high as 97% and the training loss is lowered to 0.1 based on the results. This may indicate successful classification outcomes, particularly in the area of medical diagnosis.

B. Performance Criteria

By employing the a fore mentioned criteria, we may utilize a variety of performance metrics to evaluate a classification model's effectiveness, including sensitivity, accuracy, specificity, recall, precision, and F1 score.

A model could be assessed using the two criteria of specificity and sensitivity. In the field of health, they are in fact commonly employed [18].

B.1 Definition of the Terms

To assess the performance of this classifier, we should distinguish four types of elements that are classified for the desired class: TP (True Positive), TN (True Negative), FP (False Positive), and FN (False Negative).

- **TP:** It's when the model correctly predicts the positive class. Here, positive class refers to a patient suffering from COVID 19.
- **TN:** It's when the model correctly predicts the negative class. Here, negative class refers to a patient NOT suffering from COVID 19.
- **FP (Type 1 Error):** It's when the model incorrectly predicts the positive class. Predicted that a patient suffering from COVID-19 but it's wrong.
- **FN (Type 2 Error):** It's when the model incorrectly predicts the negative class. Predicted that a patient NOT suffering from COVID-19 but it's wrong.

Let us first define the performance criteria used to evaluate the performance of the pre-trained models we used.

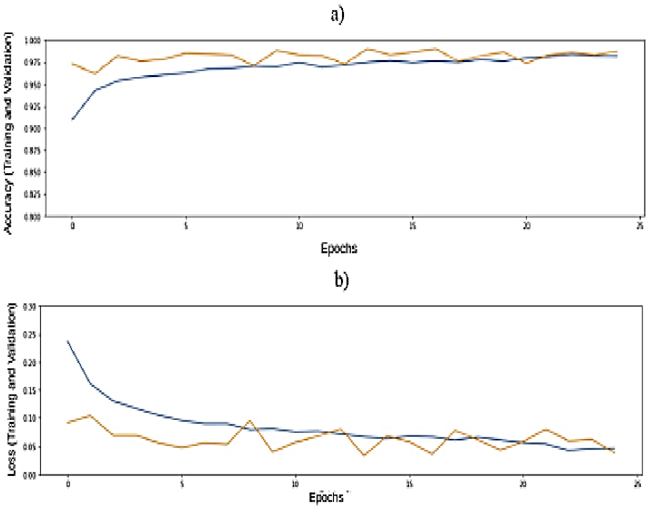


Figure 7: Plots of (a) Training and Validation Accuracy and (b) Training and Validation Loss by Using Training Epochs-VGG16

- Classification accuracy = TP+ TN / (TP + TN + FP + FN): The accuracy is defined as the rate of correctly classified images.
- Sensitivity = TP / (FN + TP): Measures how the model detects events in the positive category. Therefore, given that COVID-19 is a positive category, sensitivity can quantify how much X-ray images are correctly predicted as COVID-19.

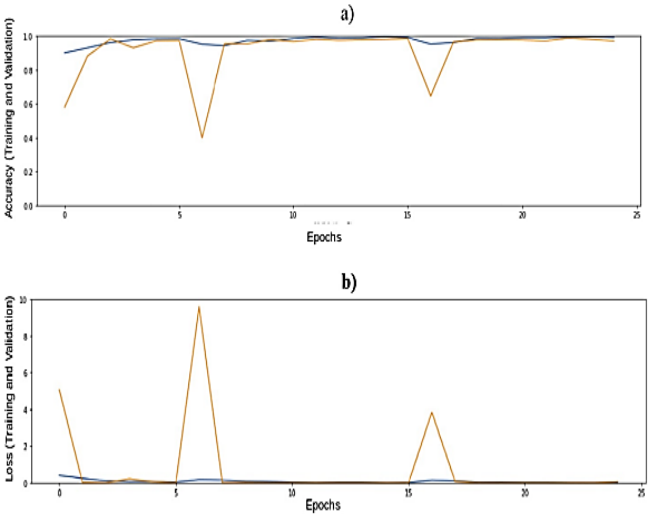


Figure 8: Plots of (a) Training and Validation Accuracy and (B) Training and Validation Loss by Using Training Epochs-Resnet50

- Specificity = TN/ (FP + TN): Specificity determines the proportion of actual negatives which are correctly detected.
- Precision= TP/(TP + FP): It is the proportion of the number of correctly classified positive categories to the number of predicted positive categories. In other words, precision is the response of the question: among all patients predicted as positive how many are really infected by COVID-19. The Precision should be high.
- Recall = TP / (FN + TP): The recall rate is the proportion of the number of correctly classified positive subjects to the number of positive subjects. The aim is to have it as high as possible.
- F1 score = 2 * (precision * recall)/ (precision + recall): To compare two models with high recall but low precision, or with low recall but high precision, is not an easy task. F1-Score is generally used, to make this comparison feasible. It enables the measuring of precision and recall at the same time. In practice, we replace the Arithmetic Mean by the Harmonic Mean. The result is that we further penalize the extreme values.

B.2 Results: The sensitivity, accuracy, specificity, recall, precision, and F1 score of Resnet50, VGG16, and Grad-CAM are displayed in Table 2. It highlights that, in terms of each of the six performance criteria, the two suggested fine-tuned versions of Grad-CAM and VGG16 perform better than the proposed fine-tuned version of Resnet50. Based on three performance criteria—recall, precision, and F1 score—these two refined versions perform identically, as shown in the final three columns of the final two rows in Table 1.

The adjusted VGG16 model performs marginally better in terms of accuracy and specificity than the refined Grad-CAM. On the other hand, the adjusted Grad-CAM performs more sensitively than VGG16. We can then conclude that overall, the fine-tuned Grad-CAM is the choice that we recommend.

Table 1:

Modified version of	Accuracy	Sensitivity	Specificity	Precision	Recall	F1 Score
Resnet50	97.20 %	98.25 %	97.00 %	97.00 %	96.00 %	97.00 %
VGG16	98.30 %	98.25 %	98.33 %	98.00 %	98.00 %	98.00%

We provide the confusion matrix for every one of the three models that were used. Figures 9, 10 show the confusion matrix of the improved ResNet50, Grad-CAM, and VGG16 models on 1000 test image sets.

According to the confusion matrix for the VGG16 model, only 10 out of 600 COVID-19 images are classified as normal class (false negative), while only 7 out of 400 normal (non-COVID-19) shots are classified as COVID-19 class (false positive).

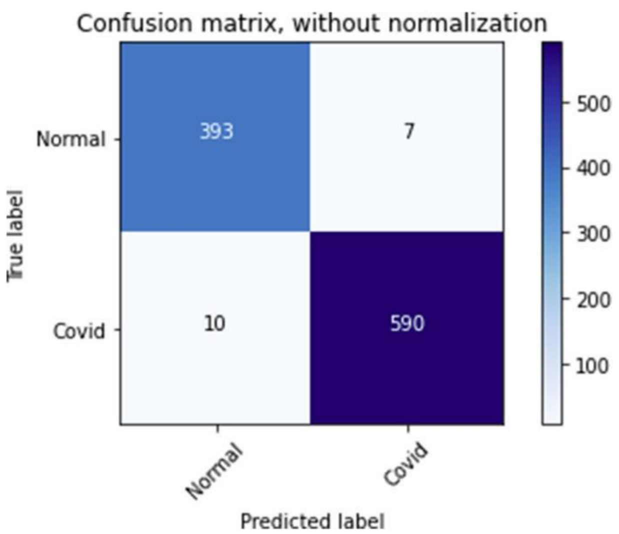


Figure 9: The Confusion Matrix of the Proposed VGG16 Model

For Resnet50, we find 7 false positive cases and 21 false negative ones. The final model's confusion matrix has sixteen false negative cases and just three erroneous positive cases. Our findings demonstrate the possibility of building an accurate CNN model using deep learning models such as VGG16, Resnet50, and Grad-CAM.

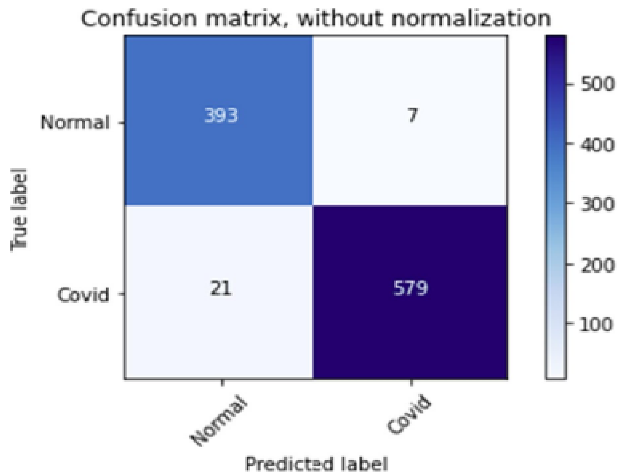
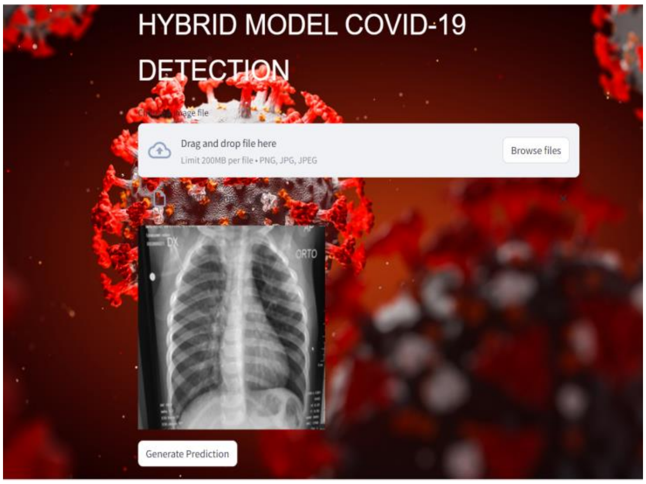
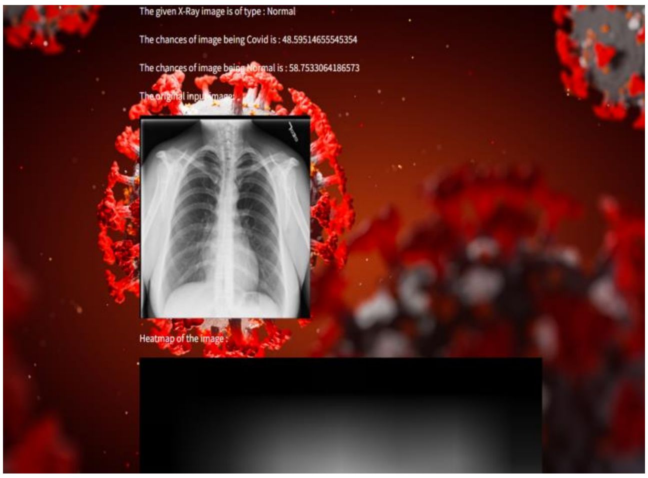


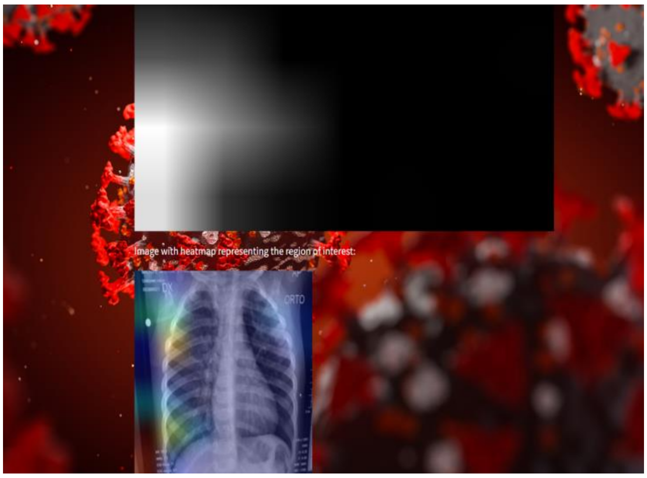
Figure 10: The Confusion Matrix of the Proposed Resnet50 Model

The encouraging results of the deep learning model for detecting COVID-19 in radiographic image detection indicate that in the near future, deep learning will play a greater clinical support role for fighting against this epidemic. Some of the limitations of this study can be overcome by performing a more analysis when we get more available data (symptomatic and asymptomatic patients).

The Webpage i created using streamlit looks like this:







When a X-Ray image is uploaded by the User in the Web page, It automatically displays the original image and below there is a Generate Prediction Button. When the user clicks that button, it displays the result of that X-Ray image as Covid and displays the accuracy. In background the Python file is executed, and it takes the Reference of Model by the help of ".h5" files generated after the training of the VGG16 and Resnet50 models. Finally, it displays the heatmap using the Grad-CAM and it predicts and displays the image with possibility of Effected Lungs which should be take care of. The Heatmap is displayed followed by The Covid Effected Part in the Lungs are Highlighted by RGB colours in X-Ray.

VI. CONCLUSION

To prevent the novel coronavirus from infecting more people, it is imperative to diagnose it as quickly as possible. Parallel to this work, we also build a deep transfer learning-based method that automatically diagnoses COVID-19 infection by combining chest X-ray images of infected and uninfected individuals. The suggested categorization approach for COVID-19 detection assists in pinpointing the lung area where the virus attacks, with an accuracy rate of more than 98%. Our study's findings suggest that, given its excellent overall performance, we believe it is only fitting that it support medical practitioners in making therapeutic decisions. This study offers a comprehensive grasp of the use of deep transfer learning techniques to identify COVID-19.

COVID-19 is a deadly virus that claims millions of lives and threatens the world's medical community. Due to the large number of patients seen in emergency rooms or outside, doctors are frequently under time pressure. With its ability to provide early detection and individualized care, computer-aided analysis holds the potential to save lives.

After efficient training on a small set of images, our improved models outperform the competition in COVID-19 pneumonia classification. We are certain that the proposed CART approach could greatly improve the diagnosis of COVID-19 cases.

This is especially helpful in the event of a pandemic when there are limited medical resources and preventative measures are needed to stop the disease from spreading.

The goal of deep learning research is to create models that can learn these representations from vast volumes of unlabelled data, and to continuously refine reality representations. Some of these representations are based on the latest developments in numerous fields. For example, Ahmad Ali et al.'s research [2] uses deep learning techniques to examine geographical and temporal correlations. Their recommendation is to employ a dynamic deep hybrid spatiotemporal neural network to precisely predict traffic flow throughout the entire city.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

[1] Ai T et al (2020) Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. Radiology: 200642

[2] Ali A, Zhu Y, Zakarya M (2021) A data aggregation-based approach to exploit dynamic spatio-temporal correlations for citywide crowd flows prediction in fog computing.Multimed Tools Appl. https://doi.org/10. 1007/s11042-020-10486-4

[3] Apostolopoulos ID, Mpesiana TA (2020) Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 43(2):635–640

[4] Axell-House DB, Lavingia R, Rafferty M, Clark E, Amirian ES, Chiao EY (2020) The estimation of diagnostic accuracy of tests for COVID-19: A scoping review. J Infect 81(5):681–697

[5] Bloice MD, Roth PM, Holzinger A (2019) Biomedical image augmentation using Augmentor. Bioinformatics 35(21):4522–4524.

[6] Cohen J, Paul et al (2020) Covid-19 image data collection: Prospective predictions are the future. ArXiv preprint arXiv:2006.11988

[7] covid-chest xray-dataset. https://github.com/ieee8023/COVID-chestxray-dataset. Accessed 25 Mar 2020

[8] Eurosurveillance Editorial Team (2020) Note from the editors: World Health Organization declares novel corona virus (2019-nCoV) sixth public health emergency of international concern. Eurosurveillance 25(5): 200131e

[9] Gazzah S, Bencharef O (2020) A Survey on how computer vision can response to urgent need to contribute in COVID-19 pandemics. 2020 International Conference on Intelligent Systems and Vision C (ISCV). IEEE, New York

[10] Globalpulse. Need for greater cooperation between practitioners and the AI community. https://www. unglobalpulse.org/2020/05/need-for-greater-cooperation-between-practitioners-and-the-ai-community/. Accessed 27 May 2020

[11] He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778

[12] Holzinger A et al (2018) Current advances, trends and challenges of machine learning and knowledge extraction: from machine learning to explainable AI. International Cross-Domain Conference for Machine Learning and Knowledge Extraction. Springer, Cham, 2018

[13] Holzinger A et al (2019) Causability and explainability of artificial intelligence in medicine. Wiley Inter discRev Data Min Knowl Discov 9(4):e1312

[14] Isa A. Computational intelligence methods in medical image-based diagnosis of COVID-19 infections. Computational Intelligence Methods in COVID-19: Surveillance, Prevention, Prediction and Diagnosis. Springer, Singapore, pp 251–270

[15] Kallianos K et al (2019) How far have we come? Artificial intelligence for chest radiograph interpretation. Clin Radiol 74(5):338–345
[16] Khan A, Sohail A, Zahoora U, Qureshi AS (2020) A survey of the recent architectures of deep convolutional neural networks. Artif Intell Rev 53(8):5455–5516
[17] Krizhevsky A, Sutskever I, Hinton GE (2017) Imagenet classification with deep convolutional neural networks. Commun ACM 60(6):84–90
[18] Lalkhen AG, McCluskey A (2008) Clinical tests: sensitivity and specificity. Contin Educ Anaesth Crit Care Pain 8(6):221–223