JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A CNN-Based Framework for Pneumonia Detection Among COVID-19 Patients from CTR Images

Deepak Neeboriyra
Digital Communication
Department of Electronic Communication Engineering
Shriram college of engineering and management, Banmore
Gwalior, India
deepakneeboriya003@gmail.com

Ashish Dubey
Electronic Communication Engineering
Shriram college of engineering and management, Banmore
Gwalior, India
adsrcem@gmail.com

Abstract*— COVID-19 (Coronavirus Disease) is a highly contagious illness. It's caused by a coronavirus that's just been found. COVID-19 has been tough to detect & treat at an early stage all over the globe. Overcrowding in hospitals is becoming a significant problem due to patients presenting signs of COVID. New pathways and opportunities for sickness diagnosis have been opened up thanks to the contribution of Deep Learning (DL) to big data medical research. To stop the spread of COVID-19, scientists need to develop a classifier that can tell the difference between corona-positive X-ray images that are positive and negative. As a result, we conducted a literature analysis, gathered data on research resources, and mapped out potential future research areas. This study compares and discusses the many ways of evaluating diagnostic procedures. Tables have been used to demonstrate the research done by different writers. Predictions made with 100% average accuracy, 99.1 percentage sensitivity and 100% specificity using the &e model were accurate. Aside from reducing the amount of time spent training, the model also improves accuracy.

Index Terms: **COVID-19, Corona Virus, Pneumonia, Chest X-Ray, VGG16, EfficientNETB5**

I. INTRODUCTION

Like a pandemic, COVID-19 spread quickly and extensively throughout the globe and profoundly influenced the sociopolitical environment and healthcare systems. WHO declared the COVID-19 epidemic a global health emergency on January 30th, 2020. New coronavirus has caused epidemics in several nations throughout the globe because it is more easily transmitted than SARS and MERS. Over 12 million individuals have been infected & 2.8 million people have died from COVID-19 pneumonia1 [1]. Even though COVID-19 has a lower casualty rate than SARS (9.6 percent) & MERS (34.4 percent), its high pathogenicity is noteworthy.

Clinical manifestations may range from asymptomatic carriers to patients who need ventilator support, both of which can result in admission to an intensive care unit (ICU) and an increased risk of death [2][3]. Nasopharyngeal swab RT-PCR test is a diagnostic test considered standard for confirmation of illness [4]. A test is helpful; nonetheless, a minuscule number of false-negative outcomes have been recorded.

To stop the spread of COVID-19, which causes respiratory system inflammation, the only viable method is to conduct a brief examination of the community and isolate those afflicted. A positive Nucleic Acid Testing (NAT) result utilizing Reverse Transcription Polymerase Chain Reaction (RT-PCR) is now the gold standard for detecting COVID-19. Specificity is reasonable, but the sensitivity is modest in this experiment [5]. Since radiological imaging of patients might identify impacted lung characteristics even when RT-PCR test results are negative, chest CT has become an unreliable diagnostic tool for COVID-19 symptoms. For radiologists, chest CT scans provide visual evidence of coronavirus infection and the size of lesions, allowing them to observe changes in disease progression accurately. This prompted the development of CT imaging, and the experiment findings show that chest radiography pictures are more accurate in detecting COVID-19 infected individuals than other diagnostic procedures. It is critical to recognize COVID-19 presumed cases accurately and quickly to implement quarantine and corrective measures as soon as possible [6].

Artificial Intelligence (AI) techniques such as machine learning, image processing & pattern recognition are being used to detect & anticipate COVID-19 infections & offer an appropriate reaction to reduce the virus's transmission and effect. Deep Learning (DL) for COVID-19 identification & lesion segmentation was suggested by Wang et al. [7]. They pinpointed the location of the lungs using a deep neural network that had already been trained to predict the likelihood of COVID-19 contamination. It is possible to detect COVID-19 using a computer-aided system that uses the multi-view representation ML technique. They used numerous CT scans to train a unified latent representation for diagnosis by extracting various features from each picture [8]. For training multiple elements of CT image features of a COVID-19 lung, they described latent representations for training numerous CT image features. Using an AGGAN-based model named CovidGAN, Waheed et al. [9] devised a method for generating synthetic CXR images. To identify the coronavirus, they improved the convolutional network model & used DL (Deep Learning) method on CXR pictures. Using CovidGANgenerated synthetic pictures, they could increase the performance of typical CNN.

COVID-19 patients will be evaluated for signs of pneumonia so that they may be removed from other patients & provided life-saving care. The purpose of this study is to do this. In this work, we show that robust models may achieve an accuracy of up to 90 percent in distinct test populations while preserving acceptable specificity in non-COVID-19 linked pneumonia & representing adequate generalizability to patient groups or unknown sites. EfficientNetB15 model is used to categorize CXR pictures of COVID-19 patients & pneumonia patients and also forecast COVID-19 patients with pneumonia.

II. LITERATURE REVIEW

The primary goal of this study [10] is to identify COVID-19 in infected people by analyzing CT scan pictures from various sources, including those with pneumonia and healthy individuals. An image's contrast is improved using a fuzzy normalized histogram, which is the basis for the process. GGO, crazy paving and consolidation have been used to demonstrate the algorithm's usefulness on CT lung images for COVID-19 patients. Based on results from clinical trials, it's safe to say that 81.89% of the 254 individuals had characteristics in both lungs, with 9.5% on the left lung & 10.24% on the right. A large portion of the right lower lobe was involved (79.53 percent).

The methods given in this research for illness diagnosis include DNN depends upon fractal imaging characteristics & CNN, which utilizes lung imaging directly. With 93.2 percent accuracy and 96.1 percent sensitivity, CNN exceeds the DNN technique with 83.4 percent precision & 86 percent sensitivity. During the segmentation step, a CNN architecture is used to identify contaminated tissue in the lung images. According to the data, the data shows nearly 4,444 contaminated places may be spotted with an accuracy rate of 83.84 percent. Monitoring and regulating the expansion of the patient's protected region is based on this discovery [11].

Using 3D CT volumes, [7] designed a weakly-supervised DL system for COVID-19 classification & lesion location. Therefore, a pre-trained UNet was utilized to segment the lungs of every patient, and the resultant 3D lung area was put into 3D DNN (Deep Neural Network) to estimate the probability that COVID-19 will be infectious. COVID-19 lesions were then found utilizing the activation areas and unsupervised connected components of the classification network. CT volumes were utilized for training 499 & testing 131, respectively. Our approach achieved AUCs of 0.959 and 0.976 for ROC and PR. To identify COVID-positive and COVID-negative, the method attained 0.901 accuracy, 0.840 positive predictive value, & 0.982 highly high negative predictive value when the probability threshold was set at 0.5. The algorithm processed a single patient's CT volume in only 1.93 seconds while running on a dedicated GPU. A chest CT scan using our weakly-supervised DL model can reliably estimate COVID-19 infection probability & locate lesion sites without the requirement for annotation.

In this research, an efficient DCNN termed "DeepChest" has been developed for Pneumonia & COVID-19 detection on CXR images [12]. There are just a few convolutional layers, a few max-pooling layers, & few training iterations in "DeepChest" likened to contemporary methods & state-of-the-art DCNN. They used 7512 chest X-ray pictures in our experiments to test the feasibility of the suggested technique. There is a total accuracy of 96%, 98% for COVID-19 detection, & 98% for Pneumonia detection using the suggested method, which is an improvement over the previous method. Pneumonia and COVID-19 may be accurately detected in CXR pictures using the method given in this article as a diagnostic tool.

Their results suggest that DL techniques that have been trained on the various multinational cohort of 1280 patients can identify parietal pleura or lung parenchyma also then identify COVID-19 pneumonia with up to 90.8 percent accuracy, 84.0 percent sensitivity, and 93.0 percent specificity when evaluated on the independent test set of 1337 patients (which was not part of training and validation). As typical controls, chest CTs from oncology, emergency, & pneumonia-related reasons were employed. In 140 patients with diverse (non-COVID-19) pneumonia confirmed in the lab, the false-positive rate was 10 percent. As well as identifying CT scans with COVID-19-associated pneumonia with high specificity, AI-based algorithms can distinguish pneumonia not linked with COVIDs in various patient demographics [13].

III. PROPOSED METHODOLOGY

Problem Statement:

The global spread of COVID-19 is accelerating exponentially, yet its transmission mechanism is not fully understood. 2–8 percent of persons infected with the virus develop a quickly deteriorating and frequently deadly pneumonia [14][15][16] despite having no symptoms. It is hard to determine mortality, prevalence, & transmission dynamics of SARS-CoV-2 infection due to unique obstacles like peak infectiousness occurring just before or just after the onset of symptoms and poorly understood multi-organ pathophysiology with lung predominance & lethality. Global healthcare systems have been stretched by the fast spread due to a lack of essential protective equipment and competent practitioners, partly due to varying availability of point-of-care diagnostic technologies, including RT-PCR. Even as quicker RT-PCR testing options become more widely accessible, difficulties persist, including delays in processing, high false-negative rates, and variations in test methodology.

Dataset:

Database repository kaggle[1] has been cited as the source of the dataset.

Diagnoses of COVID-19 are now made using RT-PCR. An early diagnosis of COVID-19 may be made using CXR pictures, which are readily accessible and give images for diagnosis promptly.

CXR (Pneumonia & Covid-19) Each of the three folders in the dataset (train, test, and validation) is subdivided into three subfolders (COVID19, PNEUMONIA, NORMAL). The DataSet comprises 6432 x-ray pictures, of which 10% of the total photos, 10% of the validation images, and 70% of the training images make up the test data.

Preprocessing:

1. The BGR colour scheme was swapped out for GRAY in the picture.
2. After resizing to 224 by 224 pixels, this picture,
3. The scaled picture was, after that, subjected to an adaptive threshold.
4. Once again, the photos were shifted from Grey to RGB colour

[1] *https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia*

The data were split into three sets for training, testing, and validation: 70% data for training, 10% for validation, and 10% for testing before being fed into the EfficientNetB5 model.

Proposed methodology

Previously, a CNN-based VGG16 model was used, which has some drawbacks, such as

- Heavy-duty version
- Training time should be increased.
- The disappearing gradient issue.
- VGG is an excellent example of a deep network that may have a more significant test error and generalize less than a simpler network.

Using a compound coefficient, Efficientnet scales all depth/width/resolution dimensions in a convolutional neural network architecture. A set of preset scaling coefficients is used to consistently scale network width, depth, & resolution in an efficientnet scaling approach instead of variable scaling. As a simple example, if we wish to double the number of computing resources available, we can expand network depth (n), width (n), & picture size (n) by using constant coefficients given by the initial tiny model in a small grid search. Efficientnet employs a compound coefficient to scale network breadth, depth, and resolution logically. "

Existing CNN architectures, like Mobile Net & ResNet, may reap the benefits of the compound scaling technique. Compound scaling only enhances the predictive performance of networks by reproducing the convolutional processes and network structure of the baseline network. For optimal results, a resilient baseline network is required.

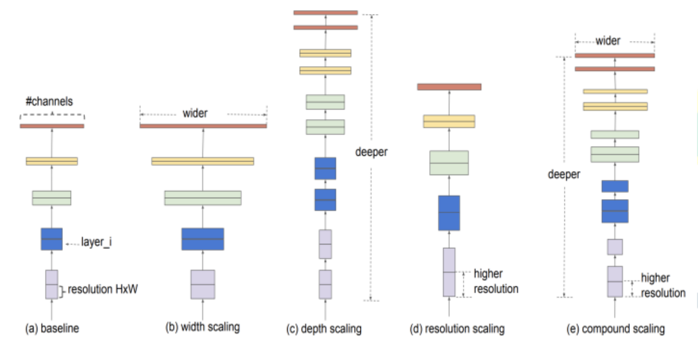


Figure 2. **Model Scaling.** (a) is a baseline network example; (b)-(d) are conventional scaling that only increases one dimension of netw width, depth, or resolution. (e) is our proposed compound scaling method that uniformly scales all three dimensions with a fixed ra

Figure1: Flow diagram of the EfficientNET model

In this network, MBConv serves as the primary construction component, and squeeze-and-excitation optimization is applied on top of it. MobileNet v2's inverted residual blocks are identical to MBConv's inverted residual blocks. A convolutional block's beginning and end are connected via these short-cut connections. Convolutions with 1x1 dimensions are used to expand feature maps of input activation first. Following this, 3x3 Depth-wise & Point-wise convolutions reduce the no. of channels in the final feature map. The broader layers are connected to the narrower ones between the skip connections. Several DL features, like AveragePooling2D, dense, flatten, dropout, and Image-DataGenerator were enabled to utilize the model.

IV. RESULTS AND DISCUSSIONS

This study has many limitations. To distinguish between COVID-19-associated illness and other conditions, model training was inadequate for patients with positive RT-PCR testing and COVID-19-associated pneumonia on chest CT. RT-PCR is commonly positive, although CT is often negative. However, since viral infectiousness may sometimes precede symptoms, it's possible that combining CT with RT-PCR will provide better results than either method alone. Specific patient subgroups or situations with limited resources, like those examining contact tracing or exposure history, prognosis, triage for resource utilization, and/or isolation compliance, may advantage from delayed RT-PCR or access or accessibility limits, trying to make CT testing more enticing.

We tested the suggested model's accuracy, precision, recall, specificity, and sensitivity using a variety of measures. Parameters of Confusion Matrix (CM) like True Positive (TP), True Negative (TN), False Positive (FP), as well as False Negative (FN) are utilized to assess metrics. Metrics are outlined in the following manner:

The outcomes of the suggested technique to predict pneumonia in COVID-19 patients are discussed in this section. To arrive at the most accurate model, researchers compared their findings.

1. Accuracy

It's the proportion of individuals appropriately classified to the total number of participants. Accuracy is the simplest to grasp.

Accuracy = (TP+TN)/(TP+FP+FN+TN) (1)

2. Precision

It is a percentage of successfully labelled +ve items to all +ve items that our algorithm correctly labelled.

Precision = TP/(TP+FP) (2)

3. Recall (aka Sensitivity)

It is a percentage of people with diabetes accurately classified by our algorithm.

Recall = TP/(TP+FN) (3)

4. F1-score (aka F-Measure / F-Score)

It is the harmonic mean (average) of precision & recall.

F1 Score = 2*(Recall * Precision) / (Recall + Precision) (4)

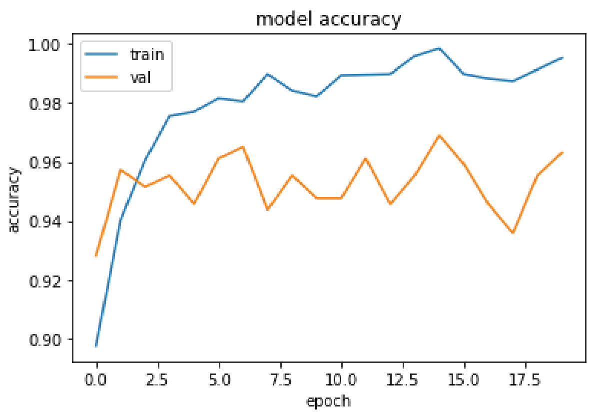


Figure 2: Training accuracy for the EfficientNETB5 model

Figures 2 and 3 illustrate how the proposed EfficientNETB5 model's training accuracy and model loss increase with each subsequent training epoch.

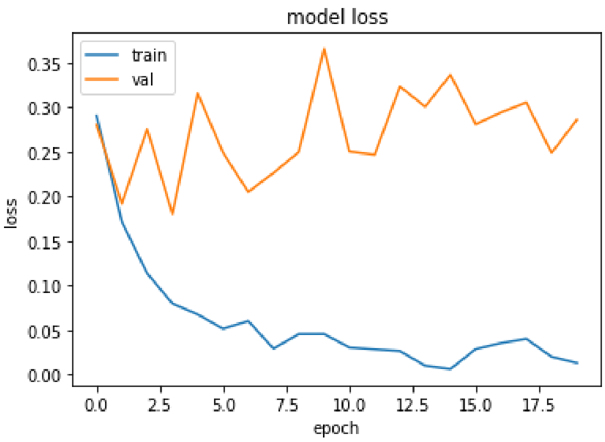


Figure 3: training loss for the EfficientNETB5 model

Fig. 4 depicts the model's training and testing data in a confusion matrix. The CM is often used to characterize the performance of a classification model (or "classifier") on a set of test data for which the actual values are known. While CM is easy to grasp, some of the associated jargon may be perplexing.

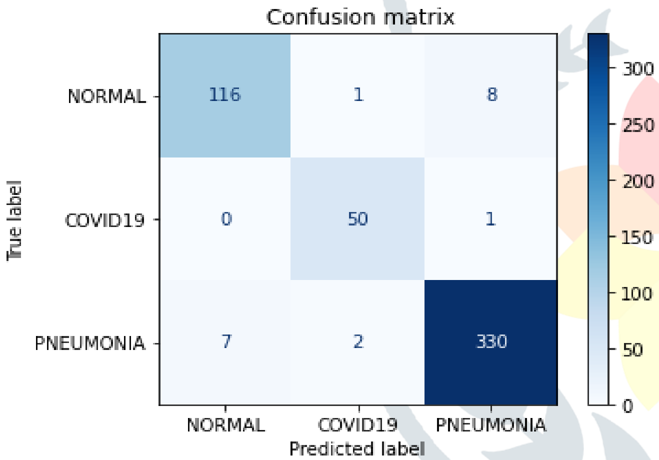
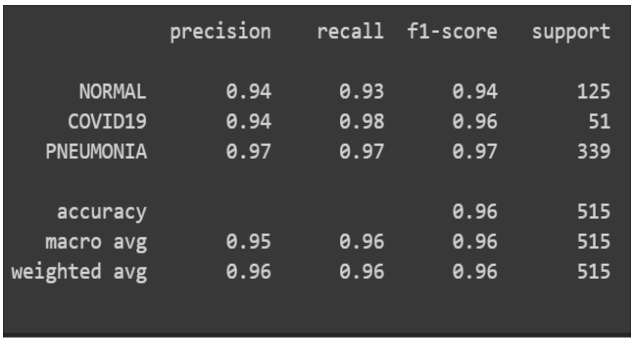


Figure 4: Confusion matrix of EfficientNETB5 model

From figure 5, it can be noted that the result is obtained in 3 classes, normal patients, patients suffering from Covid-19 & pneumonia patients. The model is efficient in predicting the less true negative values from these classes, which leads the model to gain more accurate classification results.



	precision	recall	f1-score	support
NORMAL	0.94	0.93	0.94	125
COVID19	0.94	0.98	0.96	51
PNEUMONIA	0.97	0.97	0.97	339
accuracy			0.96	515
macro avg	0.95	0.96	0.96	515
weighted avg	0.96	0.96	0.96	515

Figure 5: Classification report of the EfficientNETB5 model

Fig. 5 is the classification report of a model, which shows the values of all the three classes achieved by the proposed model. The proposed model has achieved a remarkable output in all the performance measuring metrics.

Table 1: Values of Training and Validation: Accuracy & Loss for Proposed Model

Model	Train Acc	Val Acc	Train Loss	Val Loss
EfficientNetB5	1.00	0.9631	9.9755e-04	0.285

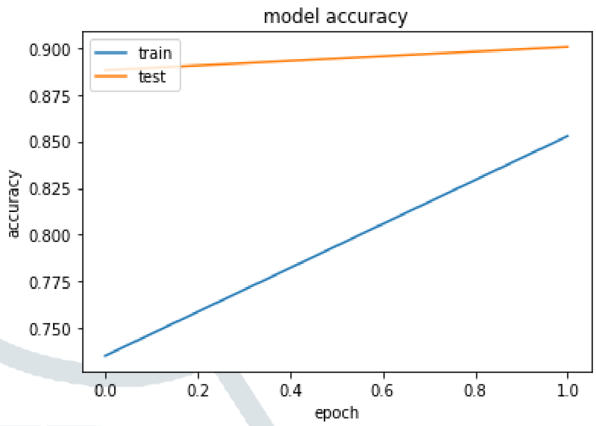


Figure 6: training accuracy for the VGG16 model

To validate (or "testing") the generalization capacity of your model or for "early stopping," as shown in Table 1 and Figures, the term "test" or "testing" generally refers to the accuracy of the validation, i.e., the accuracy you compute on the data set you do not use for training.

The total of mistakes produced in training or validation sets is called loss. If a model performs poorly after each optimization iteration, the loss value is a good indicator of this. Algorithms' performance is measured using an accuracy metric that may be easily understood.

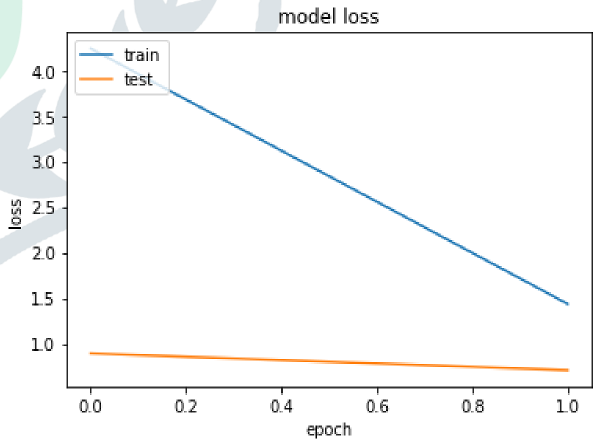


Figure 7: training loss for the VGG16 model

Confusion matrices for the train and test sets are shown in Figures 8 and 9. As shown in Fig. 8, values for incorrectly identified samples are greater. Figures 2, 3, 6, and 7 depict the model's training and testing results as it advances through various epochs. Even while our suggested model is prone to over-fitting, it still provides the best accuracy.

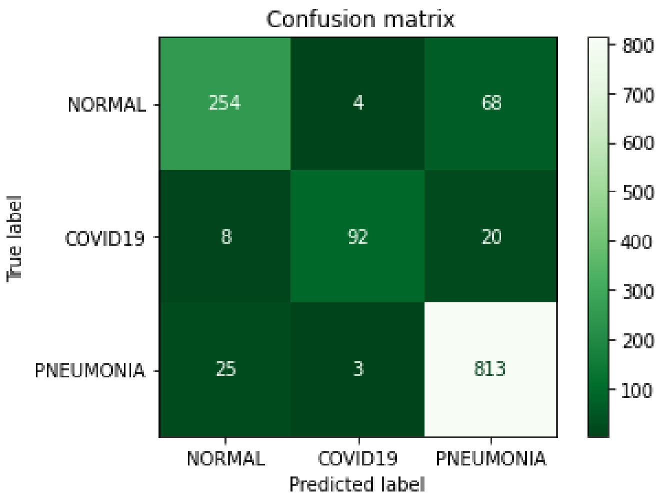


Figure 8: Confusion matrix of VGG16

Total accuracy, specificity, sensitivity, and positive and negative predictive value were measured to see how well the current and new models classified COVID-19 as opposed to any other disease. All-comer/any-indication cohorts and pneumonia cohorts were separately analyzed for COVID-19 false positives.

	precision	recall	f1-score	support
NORMAL	0.89	0.78	0.83	326
COVID19	0.93	0.77	0.84	120
PNEUMONIA	0.90	0.97	0.93	841
accuracy			0.90	1287
macro avg	0.91	0.84	0.87	1287
weighted avg	0.90	0.90	0.90	1287

Figure 9: Classification report of the VGG16model

The classification report in figure 8 shows that the existing model has achieved less output than the proposed model. The precision obtained in this model is 0.89 for normal class, 0.93 for Covid-19 class and 0.90 for pneumonia. In contrast, the proposed model obtained precision values of 0.94, 0.94 and 0.97 for all the three respective classes, which is higher than the previous results.

Table 2: Values of Training & Validation: Accuracy & Loss for Existing Model

Model	Train Acc	Val Acc	Train Loss	Val Loss
VGG16	0.9139	0.9005	0.6349	0.7076

Table 3: Comparison of Proposed and Existing Model in terms of Specificity and Sensitivity

Model	Sensitivity	Specificity
EfficientNETB5	0.991	1.0
Vgg16	0.984	0.92

V. CONCLUSION

The pandemic of Covid-19 is rapidly spreading. As the number of instances grows, it may be necessary to test several cases simultaneously.

The categorization of chest CT for COVID-19 infection can be improved using an AI system built on diverse clinical training data. DL-based AI method to CT imaging cannot be dynamically utilized in diagnosis & screening of COVID-19; nevertheless, it may serve as an objective standard for evaluation of imaging outcomes of COVID-19 and also be useful as a research tool, clinical trial response metric, or potentially as a complementary test tool in particular inadequate populations or for repeated outbreaks settings of research.

COVID-19-induced pneumonia may be identified using a 2-stage deep residual learning algorithm, including Lung X-ray images. The EfficientNETB5 model performed well in distinctive COVID-19 patients from those with COVID-19-induced pneumonia. Accuracy, sensitivity and specificity were all over 100 percent in the model's predictions for pneumonia. As a result, training loss is reduced while precision is bolstered.

REFERENCES

[1] G. A. Saeed *et al.*, “Correlation between Chest CT Severity Scores and the Clinical Parameters of Adult Patients with COVID-19 Pneumonia,” *Radiol. Res. Pract.*, 2021, doi: 10.1155/2021/6697677.

[2] D. Thomas-Rüddel *et al.*, “"Coronavirus disease 2019" (COVID-19): update für Anästhesisten und Intensivmediziner März 2020.,” *Anaesthesist*, 2020.

[3] V. M. Corman *et al.*, “Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR,” *Eurosurveillance*, 2020, doi: 10.2807/1560-7917.ES.2020.25.3.2000045.

[4] S. A. Bustin and T. Nolan, “Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction,” *Journal of Biomolecular Techniques*. 2004.

[5] A. Kovács, P. Palásti, D. Veréb, B. Bozsik, A. Palkó, and Z. T. Kincses, “The sensitivity and specificity of chest CT in the diagnosis of COVID-19,” *European Radiology*. 2021, doi: 10.1007/s00330-020-07347-x.

[6] C. Zhou, J. Song, S. Zhou, Z. Zhang, and J. Xing, “COVID-19 Detection based on Image Regrouping and ResNet-SVM using Chest X-ray Images,” *IEEE Access*, 2021, doi: 10.1109/ACCESS.2021.3086229.

[7] X. Wang *et al.*, “A Weakly-Supervised Framework for COVID-19 Classification and Lesion Localization from Chest CT,” *IEEE Trans. Med. Imaging*, 2020, doi: 10.1109/TMI.2020.2995965.

[8] H. Kang *et al.*, “Diagnosis of Coronavirus Disease 2019 (COVID-19) with Structured Latent Multi-View Representation Learning,” *IEEE Trans. Med. Imaging*, 2020, doi: 10.1109/TMI.2020.2992546.

[9] A. Waheed, M. Goyal, D. Gupta, A. Khanna, F. Al-Turjman, and P. R. Pinheiro, “CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection,” *IEEE Access*, 2020, doi: 10.1109/ACCESS.2020.2994762.

[10] A. Alzahrani, M. A. A. Bhuiyan, and F. Akhter, “Detecting COVID-19 Pneumonia over Fuzzy Image Enhancement on Computed Tomography Images,” *Comput. Math. Methods Med.*, 2022, doi: 10.1155/2022/1043299.

[11] S. Hassantabar, M. Ahmadi, and A. Sharifi, “Diagnosis and detection of infected tissue of COVID-19 patients based on lung x-ray image using convolutional neural network approaches,” *Chaos, Solitons and Fractals*, 2020, doi: 10.1016/j.chaos.2020.110170.

[12] A. S. Musallam, A. S. Sherif, and M. K. Hussein, “Efficient framework for detecting COVID-19 and pneumonia from chest X-ray using deep convolutional

network," *Egypt. Informatics J.*, 2022, doi: 10.1016/j.eij.2022.01.002.

[13] S. A. Harmon *et al.*, "Artificial intelligence for the detection of COVID-19 pneumonia on chest CT using multinational datasets," *Nat. Commun.*, 2020, doi: 10.1038/s41467-020-17971-2.

[14] S. Hoehl *et al.*, "Evidence of SARS-CoV-2 Infection in Returning Travelers from Wuhan, China," *N. Engl. J. Med.*, 2020, doi: 10.1056/nejmc2001899.

[15] C. C. Lai *et al.*, "Asymptomatic carrier state, acute respiratory disease, and pneumonia due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2): Facts and myths," *Journal of Microbiology, Immunology and Infection*. 2020, doi: 10.1016/j.jmii.2020.02.012.

[16] C. Sohrabi *et al.*, "World Health Organization declares global emergency: A review of the 2019 novel coronavirus (COVID-19)," *International Journal of Surgery*. 2020, doi: 10.1016/j.ijsu.2020.02.034.