Advances in Computational Sciences and Technology
ISSN 0973-6107 Volume 16, Number 1 (2023) pp. 27-34

https://dx.doi.org/10.37622/ACST/16.1.2023.27-34

Detection of COVID-19 and Pneumonia using Chest X-ray Scans with Deep Learning

**Nomula Divya, Vulpala Akshara, Valgot Shreya,
Vundavelly Gayathri and Nirupama Varma**

Department of Computer Science & Engineering
G. Narayanamma Institute of Technology and Science, Hyderabad, India

Abstract

COVID-19, a brand-new coronavirus that causes pneumonia, recently appeared. Since then, COVID-19 has been responsible for several infections and fatalities across the globe. One way to stop the spread of this virus is to isolate those who are afflicted. Radiologists can diagnose COVID-19 in hospitals with the help of X-rays, which give radiologists a precise image of the lungs. However, a person's X-rays typically contain hundreds of slides, and using these scans to diagnose COVID-19 Pneumonia can cause delays in hospitals. A thorough investigation is required on the effectiveness of an automated diagnosis system for identifying COVID-19 patients to stop the coronavirus from spreading. By using four different pre-trained architectures, including VGG16, ResNet50, InceptionV3, and CNN-11, which were trained and tested on X-ray images of COVID-19, Normal, and Viral Pneumonia, this proposal seeks to address this issue. The outcomes from the three classes (Normal, COVID-19, and Pneumonia) were presented.

Keywords— VGG16, CNN-11, ResNet 50, Inception V3.

I. INTRODUCTION

Currently, human-to-human transmission is the primary method of coronavirus transmission, and COVID-19 vaccination options are scarce. According to reports, the best strategy to stop the spread of the coronavirus is to quickly identify it in big crowds and then isolate the afflicted people. Therefore, routine COVID-19 tests are required to identify and isolate infections in patients. RT-PCR (transcription-polymerase chain reaction) assays are primarily used in hospitals to identify patients with COVID-19.
Computed tomography (CT) pictures and X-rays are other diagnostic methods for identifying COVID-19. On X-ray or CT pictures of the lung organs, medical

professionals search for indications of COVID-19 deformations. For accurate determination of the kind of pneumonia, this process needs some time.
Computed tomography (CT) scans and X-ray images are two more diagnostic tools for determining COVID-19. Medical practitioners look for signs of COVID-19 deformations in the lungs through CT or X-ray images. This method takes some time to accurately determine the type of pneumonia. The ImageNet dataset contains 1000 items. The performance findings demonstrate that these models can reach object-level accuracy that is nearly human-level.
High classification performance in the classification of medical images is another outcome of these models. Therefore, our suggested method helps radiologists identify COVID-19 infection from these X-rays quickly and accurately.

II. RELATED WORK

Satya P. Singh[1]., This research discussed the evolution of the convolutional neural network (CNN) and its architectures. Deep learning models are being used more frequently in the medical field as a result of the quick development of graphics processing, machine learning technology, and the accessibility of medical imaging data. This study provides a concise mathematical description of the 3D CNN as well as details on the pre-processing techniques required for medical images before they are fed to 3D CNNs. They also by tracing the development of the 3D CNN to its machine-learning origins. They examined important studies on 3D CNNs (and their variants) used in medical imaging analysis in several fields, including classification, segmentation, detection, and localization.
Yassine Meraihi[2]., This document provides a summary of over 160 ML-based strategies created to tackle COVID-19. They are examined and divided into two groups: techniques based on supervised learning and those based on deep learning. The ML algorithm used in each category is described, along with a list of the applied parameters. According to the results, 79% of cases include deep learning, 65% of which are based on convolutional neural networks (CNN), and 17% employ specialized CNN. They claim that only 16% of the methodologies they studied used supervised learning.
Dina M. Ibrahim[3]., This study focuses on the diagnosis of pneumonia, covid-19 Lung cancer chest diseases. Deep learning architectures are applied to a dataset that contains a combination of both X-Rays and CT scans. Since CT scans are helpful even before symptoms appeared and X-rays are least effective in the early stages. They considered four architectures those are ResNet152V2 + Bidirectional GRU (Bi-GRU), ResNet152V2 + Gated Recurrent Unit (GRU), ResNet-152-V2, and VGG19-CNN. Results indicate that the VGG19+CNN model performs better than the other three proposed models and achieved 98.05% accuracy,98.05% recall, and 98.43% precision.
Sajib Sarker[4]., The main aim of this paper is to propose an automated method for diagnosing covid 19, Viral Pneumonia, and Bacterial Pneumonia from X-ray images. The dataset consists of 1140 images out of which 240 COVID-19 images, 300 Normal images, 300 Viral Pneumonia images, and 300 Bacterial Pneumonia images. Here they used eight architectures those are InceptionResNetV2, VGG19, DenseNet201, ResNet50V2, MobileNet, Xception, and InceptionV3. Out of all DenseNet201 gives

the highest accuracy i.e., recall: 92.81%, accuracy: 92.54%, and precision: 93.05%.

III. METHODOLOGY

The working of the project can be understood and analyzed through the following steps:
Step-1: Data collecting and cleaning Step-2: Data preprocessing.
Step-3: Bulid Train and validation set
Step-4: Data augmentation
Step-5: Define a convolutional neural network Step-6: Train model.
Step-7: Visualizing model performance.
Step-8:Comparative study of the models.

A. Data Collection

Data for covid-19 class was collected from the COVID-QU-Ex Dataset, the Italian Society of Medical and Interventional Radiology (SIRM) COVID-19 DATABASE, the Novel Corona Virus 2019 Dataset, and the COVID-CT dataset.
Whereas the data for the pneumonia class was collected from "The Radiology Society of North America (RSNA)". Finally, the dataset consisted of

- 3600 Images of COVID-19
- 6000 Images of normal
- 941 Images of Pneumonia

Since the data was containing images from multiple datasets, there were some inconsistencies in the final dataset. So to make them consistent Gray-Scaling and image resizing to 299x299 are done.

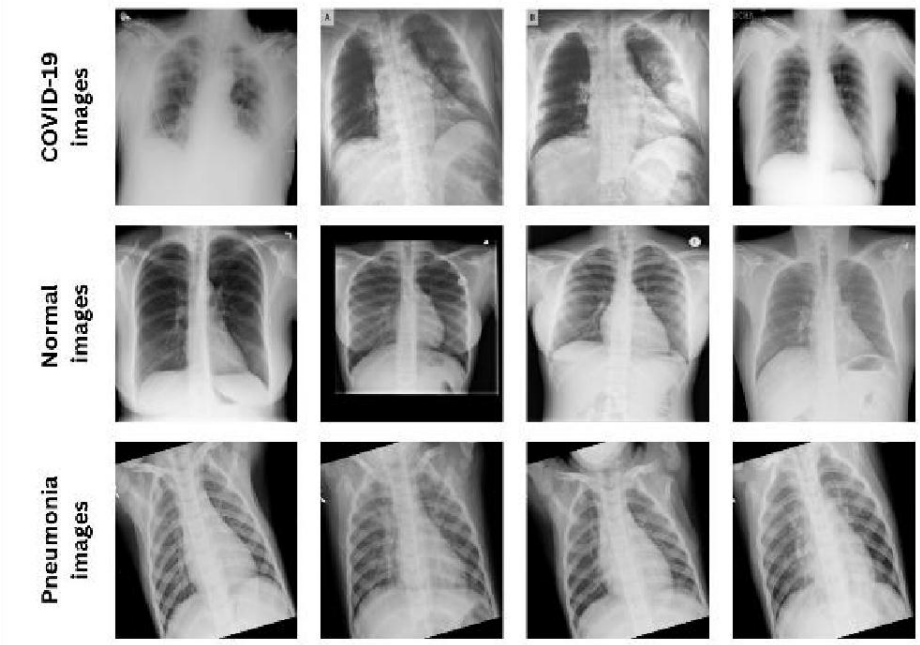

Fig. 1. Several representative chest X-ray pictures from the collected datasets with COVID-19, Pneumonia, and Normal patients

B. Data Preprocessing

An innovative technique called contrast-limited adaptive histogram equalization (CLAHE) uses an optimized image processing filter chain that includes a noise reduction filter, an image sharpening filter, and a CLAHE filter to process 2D xray pictures. The final Images after Preprocessing are shown in Fig.2.

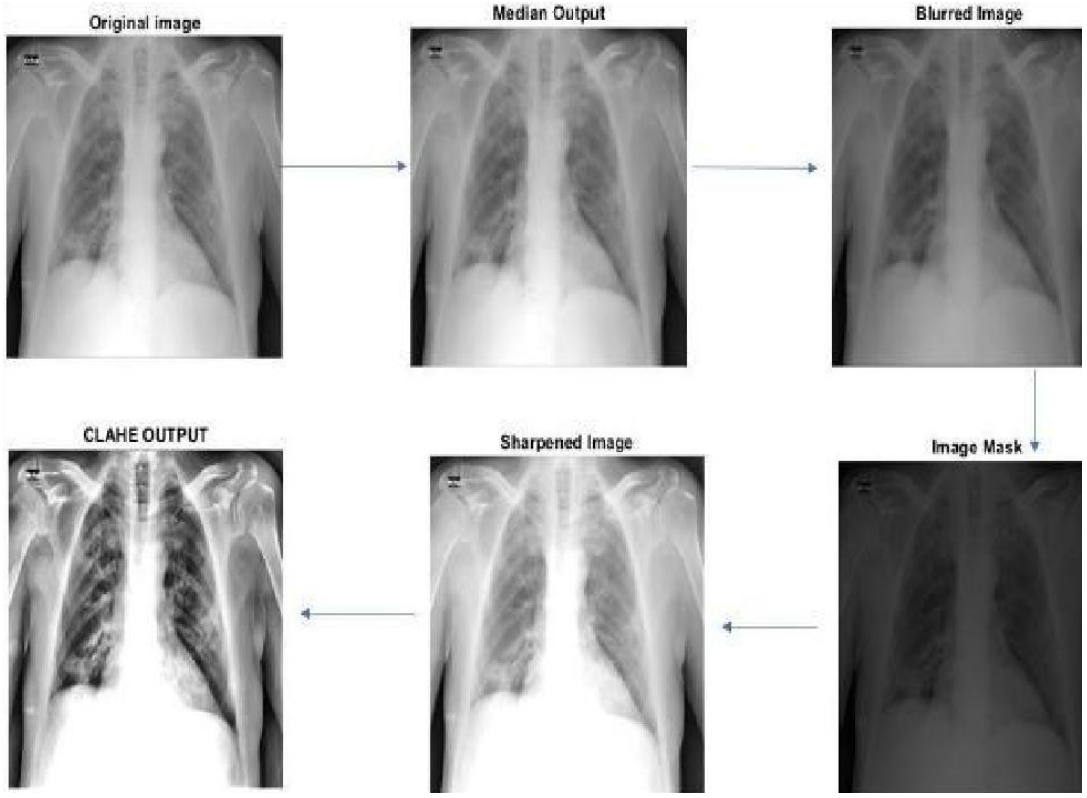

Fig. 2. Final images after data preprocessing.

C. Data Augmentation

Since there is an imbalance in the pneumonia class and covid class compared to normal images, image augmentation is used to increase the sample in the dataset. For the Covid class "Horizontal Flip" Augmentation(3000x2=6000) is done and for Pneumonia, class 5 augmentations are done (941x5=4500(approx)), which are Horizontal flip, 5% Horizontal Translation, 5% Vertical Translation, +15 degree rotation,-15 degree rotation. Some samples of Pneumonia class Images after augmentation are shown in Fig.3.

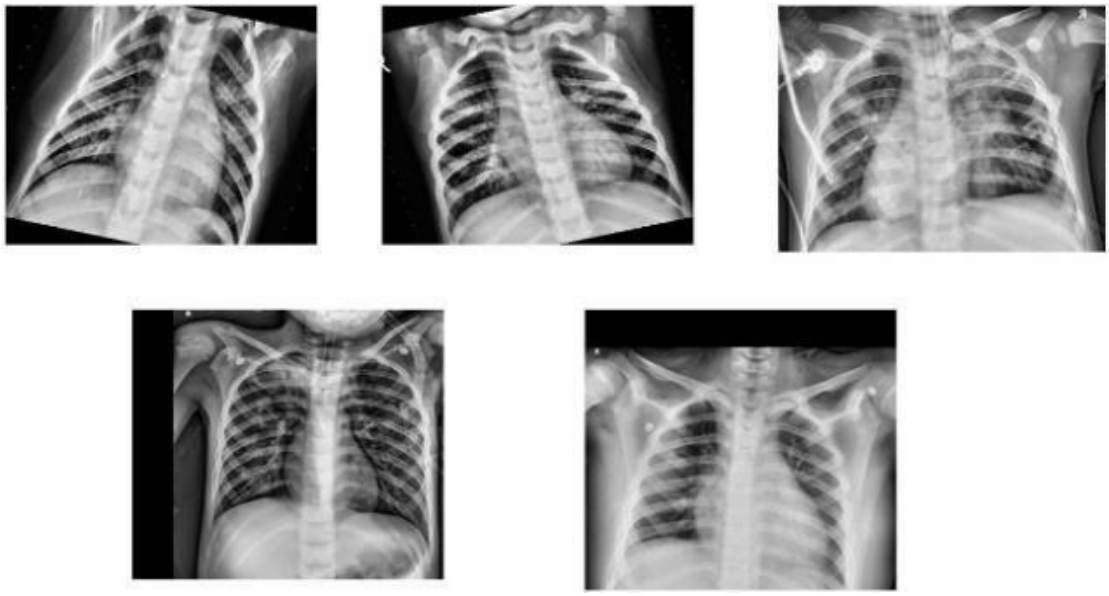

Fig. 3. Pneumonia class augmentation

D. Define a convolutional Network

Once the data is preprocessed the data is fed to the deep learning models and a comparative study will be performed based on the results. Implemented the following Deep Learning Models and compared the accuracies. Following were the models implemented

- CNN-11 Layer Architecture
- RESNET-50
- InceptionV3
- VGG16

IV. RESULTS AND CONCLUSION

E. Results

In our proposed system, each model is performed with different training and testing ratios to observe which ratio gives the best accuracy for all the models. It is observed that the 70:30 train and test ratio resulted in higher accuracy for all the models. The accuracy achieved for CNN-11 is 80.12%, Inception V3 is 90.3%, VGG-16 is 82.51%, and ResNet-50 is 83.1%.

A comparative study is done between the algorithms without the implementation of image preprocessing and with the implementation of image preprocessing. Table 1 shows that without Image pre-processing the accuracies are all around 6575% and with Pre-processing, the proposed model was able to learn to feature much faster and achieved high results.

Below are the accuracy graphs of an individual model with a 70:30 train and test ratio.

TABLE 1. COMPARISON OF ALGORITHMS

Schemes	Model	Accuracy	Number of layers	Time is taken to saturate for 100 epochs
Without Image Preprocessing	CNN-11	70.04%	11	2 Hrs
	RESNET 50	65%	50	5 Hrs
	InceptionV3	**75.32%**	48	5 Hrs
	VGG16	68.59%	16	3 Hrs
With Image Pre-processing	CNN-11	80.12%	11	2 Hrs
	RESNET 50	83.1%	50	4 Hrs
	InceptionV3	**90.3%**	48	4 Hrs
	VGG16	82.51%	16	2.5 Hrs

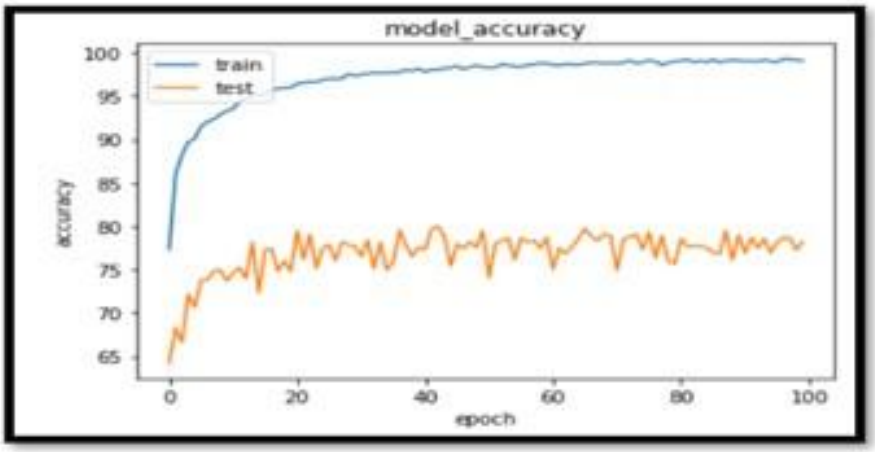

Fig. 4. Accuracy graph of CNN-11 Model

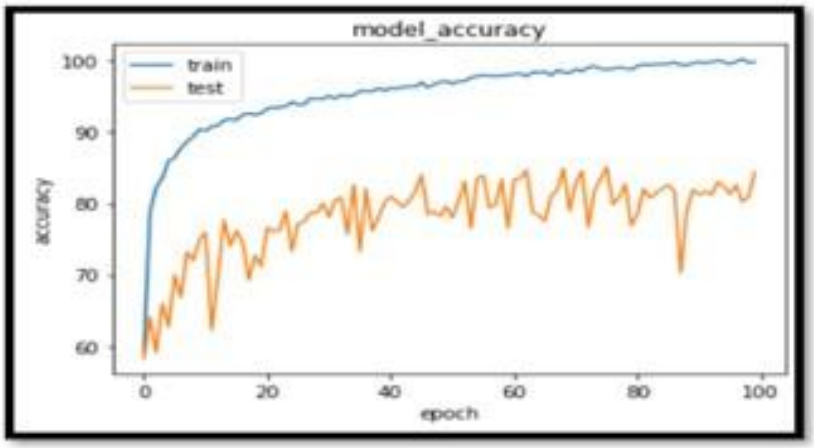

Fig. 5. Accuracy graph of VGG-16Model

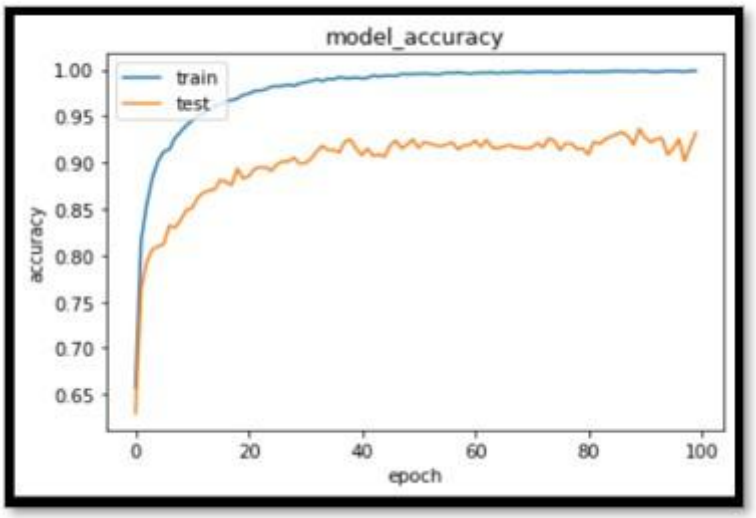

Fig. 6. Accuracy graph of Inception V3 Model

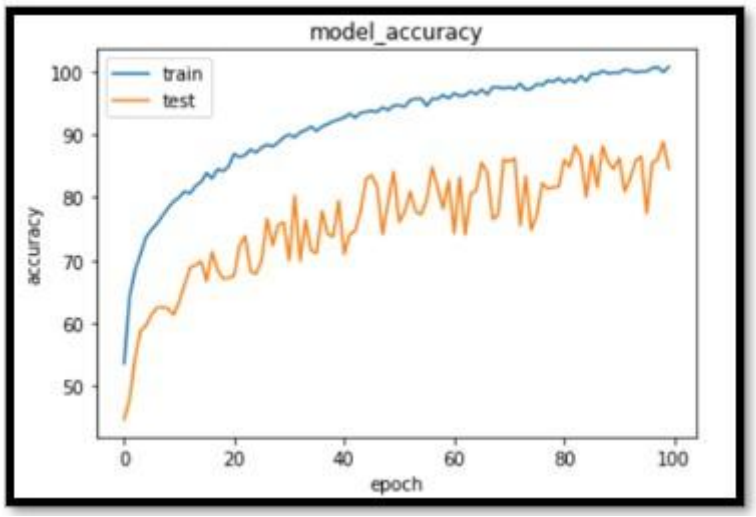

Fig. 7. Accuracy graph of ResNet-50 Model

F. Conclusion and Future Enhancements

In the modern era, the Deep Learning model for diagnosing COVID-19 and pneumonia from X-ray studies is very significant since it is a reliable way to determine whether a person has covid or pneumonia or not. This model is very wellliked since it provides quicker results for bigger datasets, which aids in the diagnosis of lung-related problems. Quick and automated Identifying COVID-19 can serve as one of the crucial measures for preventing the transmission of the virus the unaffected communities. In light of this, this study has put a proposal for precisely an approach, namely based on four different

pre-trained models (VGG16, ResNet50, InceptionV3, CNN-11) to detect COVID-19 patients automatically. Based on the 3-class classification, the provided models could successfully identify COVID-19-infected cases. With a 90.3% accuracy rate, the three-class classification showed that Inception V3 is the best model for classifying COVID-19infected cases.

This study achieved good results compared to the others which have smaller datasets. This model will be helpful in the medical field to detect and diagnose COVID-19 in a much faster way. Enhancing the performance of the model can be achieved with the help of the expansion of the training dataset. This model needs scientific testing but with higher performance to make a way into the world for the diagnosis of COVID-19 and pneumonia.

REFERENCES

[1]. Satya P. Singh Parasuraman & Wang, Lipo& Gupta, Sukrit&Goli, Haveesh& & Gulyas, Balazs. Padmanabhan, "3D Deep Learning on Medical Images: A Review", 2020.

[2]. Yassine Meraihi,Asma Benmessaoud Gabis,Seyedali Mirjalili,·Amar Ra mdaneCherif,Fawaz E. Alsaadi. "Machine LearningBased Research for COVID-19 Detection, Diagnosis, and Prediction: A Survey". SN Computer Science (2022) 3:286

[3]. Dina M. Ibrahim, Nada M. Elshennawy, Amany M. Sarhan. Deep-chest: Multi-classification deep learning model for diagnosing COVID-19, pneumonia, and lung cancer chest diseases. Computers in Biology and Medicine 132 (2021) 104348

[4]. SajibSarker, Ling Tan, Wenjie Ma, Shanshan Rong, Osibo Benjamin Kwapong, "Multi-Classification Network for Identifying COVID-19 Cases Using Deep Convolutional Neural Networks". Journal on Internet of Things, Vol.3, No.2, pp. 39-51, DOI:10.32604/jiot.2021.014877, 2021

[5]. Zunair, Hasib & Rahman, Aimon& Mohammed, Nabeel & Cohen, Joseph. "Uniformizing Techniques to Process CT Scans with 3D CNNs for Tuberculosis Prediction". 10.1007/978-3-030-59354-4_15, 2020. [6]. Serte S, Demirel H. "Deep learning for diagnosis of COVID-19 using 3D CT scans". Comput Biol Med. 2021 May;132:104306. doi 10.1016/j.compbiomed.2021.104306. Epub 2021 Mar 10. PMID:33780867; PMCID: PMC7943389.