Journal Homepage: -www.journalijar.com

INTERNATIONAL JOURNAL OF ADVANCED RESEARCH (IJAR)

Article DOI:10.21474/IJAR01/19972
DOI URL: http://dx.doi.org/10.21474/IJAR01/19972

RESEARCH ARTICLE

COVID-19 DETECTION USING CHEST X-RAYS WITH VGG-16 DEEP LEARNING MODEL

Chinmay Muralidharan
Manipal School of Information Science AI/ML Dept Manipal, Karnataka India.

Manuscript Info

Manuscript History
Received: 19 September 2024
Final Accepted: 27 October 2024
Published: November 2024

Key words:-
Covid-19 Detection, Vgg-16, Chest X-Ray, Deep Learning, CNN, Medical Imaging

Abstract

This study demonstrates a deep learning approach using VGG-16 convolutional neural network to detect COVID-19 from chest X-ray images. The model achieves a high classification accuracy of 93%, with balanced precision, recall, and F1-score values, demonstrating its potential as an efficient tool for assisting in the diagnosis of COVID-19. The results show that VGG-16 can generalize without significant overfitting, making it an effective method for COVID detection.

Introduction:-

One ofthe most important ways to diagnose COVID-19 is to use radiological images, including X-ray and computed tomography(CT) scan. Chest imaging is a quick and easy procedure which is recommended by medical and health protocols as the first tool in screening during epidemics [1, 2]. In comparison to RT-PCR, images from CT scan have a high sensitivity in diagnosing and detecting cases with COVID19, but have low specificity.Astudy was done inWuhan, China, where among the diagnosed patients, consolidation and ground-glass opacities (GGO) were not observed in CT scan in 14% of the cases, that is, 14% of definitive cases of COVID-19 were misdiagnosed as completely healthy based on CT scan tests[3]. Additionally, CT scanner rays cancause problems in patients requiring multiple CT scans during the course of the disease.

It is recommended by the American College of Radiology to not use CT scans as the first line of diagnosis due to issues such as transmission of disease while using the device along high cost, creating serious hurdles for the healthcare systems and the patient. Hence, it is advised that, when there is a requirement of medical imaging CXR radiography should be preferred over CT scan [4].

X-Ray imaging makes a far more broad and cost-effective method compared to conventional diagnostic tests. Unlike CT scans, X-Ray imaging does not need rare and expensive equipment. This saves a significant amount in the running costs. Also, there are portable CXR machines available which can be used in isolated rooms reducing the risk of infection.

Workflow

Dataset is taken from a Kaggle repository, which is open source for several datasets. After collection of dataset, the CXR images are pre-processed using data augmentation. The pre-processed data is thendivided into Training, Validation and Test dataset in 80%, 10% and 10% ratio respectively.

Data augmentation is a method to synthetically increase the amount of data to generate various possible combinations from the original dataset. It helps in reducing data overfitting problems and creating variability in data.

Corresponding Author:-Chinmay Muralidharan
Address:-Manipal School of Information Science AI/ML Dept Manipal, Karnataka India.

The data augmentation techniques used for our dataset are as follows Zoom, Shear range, Horizontal Flip, Rescaling=1/255, Pixel size= 224X224.

Google Collab platform is used to build and train the model. The model used is VGG-16 which is a convolutional neural network that has 16 layers. After training the model, it is tested using the CXR images. CXR images can then be fed into the model to give results. The model predicts a binary outcome: COIVD-19 or normal.

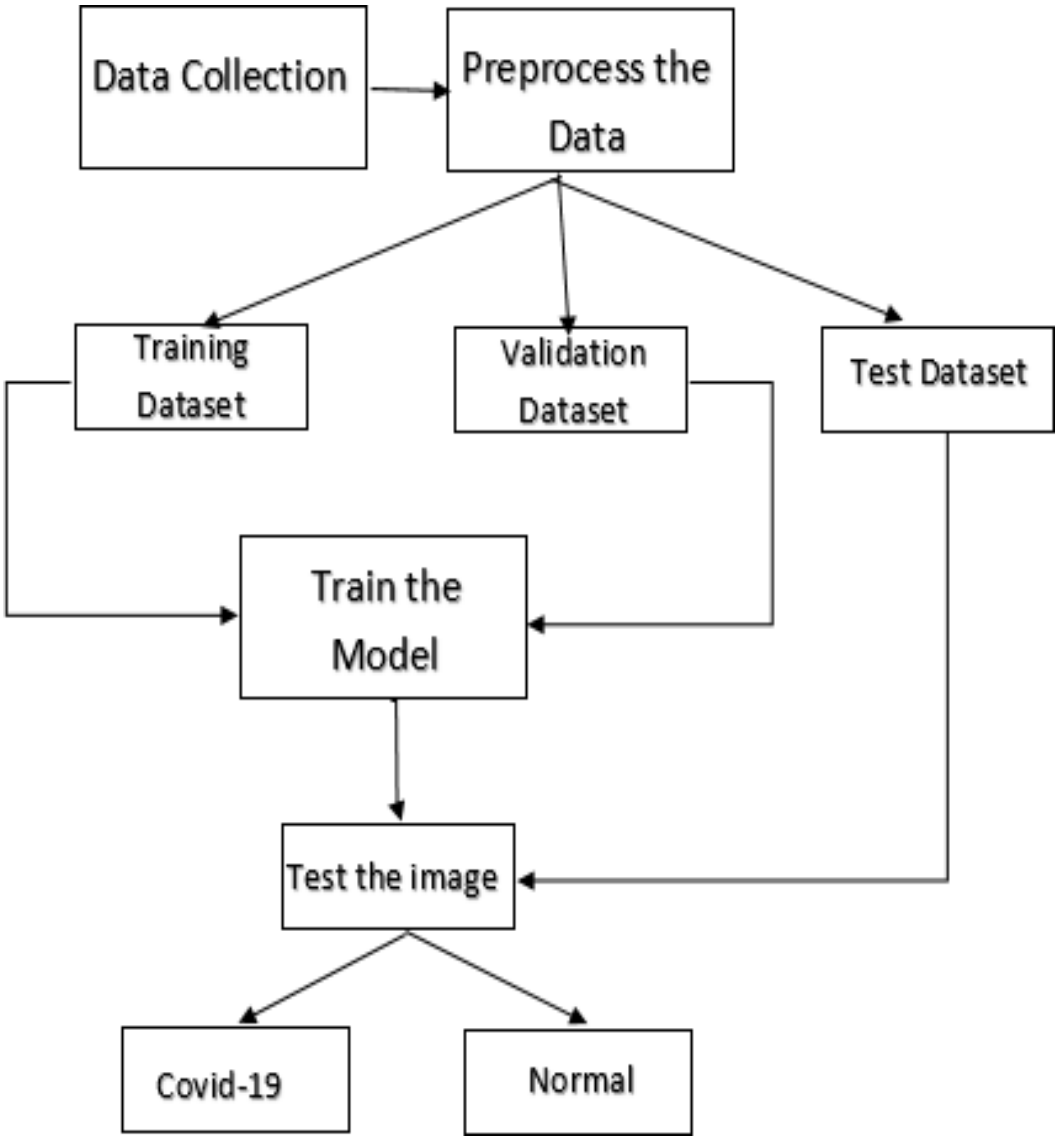

Fig.1: -BlockDiagram of workflow.

Dataset

To train and validate the model, CXR images of normal and Covid -19 positive patients were obtained from Kaggle website. It is a labelled dataset. The dataset contains 10,100 images out of which we took 5500 images of Normal CXR and 5500 images of Covid-19 CXR. The entire dataset is split into 3 parts: training, validation, testing with ratio of 80%, 10% and 10% respectively.

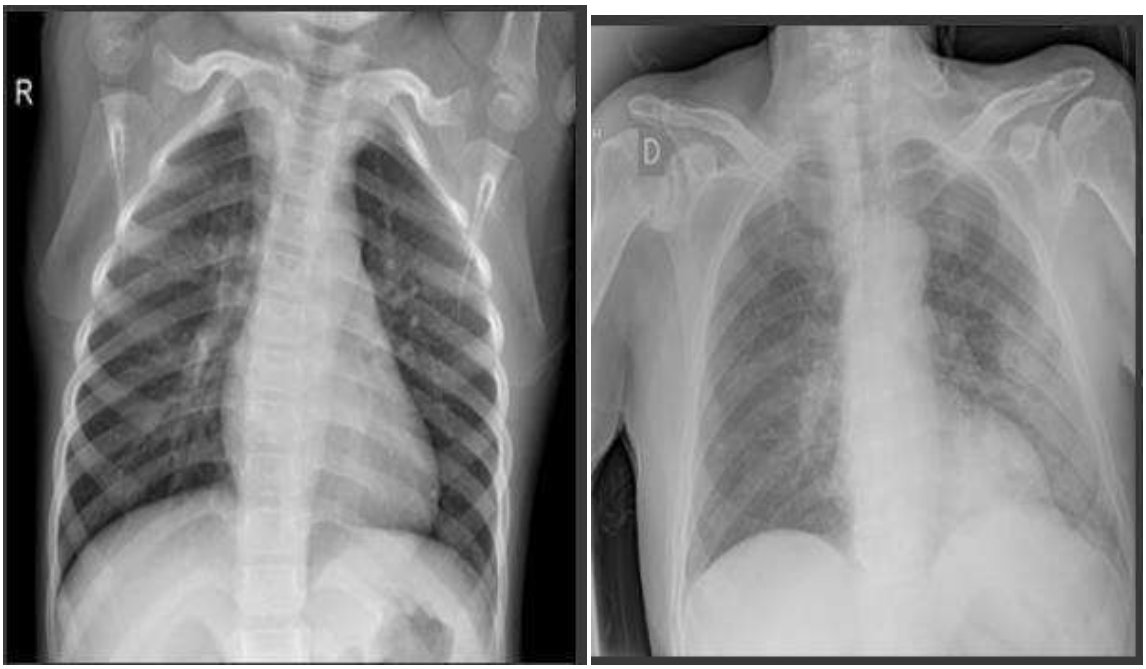

Fig. 2:- ChestX-RayimagesofcovID-19andNormalPatientsfromtheDataset.

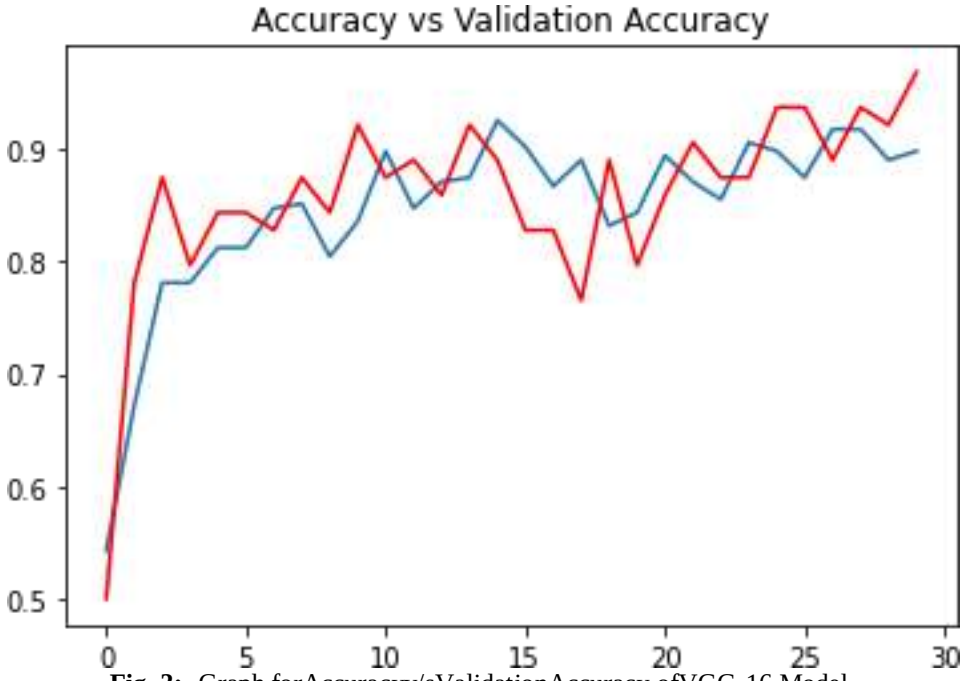

Fig. 3:- Graph forAccuracyv/sValidationAccuracy ofVGG-16 Model.

Results and Discussion:-

The VGG-16 model demonstrates great potential in the automated detection of COVID-19 from CXR. An accuracy of 93%, along with high precision and recall suggests the model's robustness and suitability for clinical applications. The model generalizes well to new data, animportant factor in medical image analysis where overfitting can limit practical utility. Additionally, the use of small 3x3 filters in the VGG-16 architecture appears to effectively capture essential features in the X-ray images, leading to reliable classification results.

VGG-16 network's performance in this study is notable, especially considering that it avoids the excessive computational complexity of larger architectures. The findings suggest that VGG-16 can serve as a dependable tool for preliminary COVID-19 screening, easing the diagnostic burden on healthcare systems.

Vgg-16	Precision	Recall	F1-score	Accuracy
	94%	91%	93%	93%

Fig.4 -TableofParametersforVGG-16: The model has maintained a balanced classification performance, with high precision and recall values, with a robust F1-score of 93% indicating that it is reliably able to distinguish between COVID-19 and normal chest X-rays

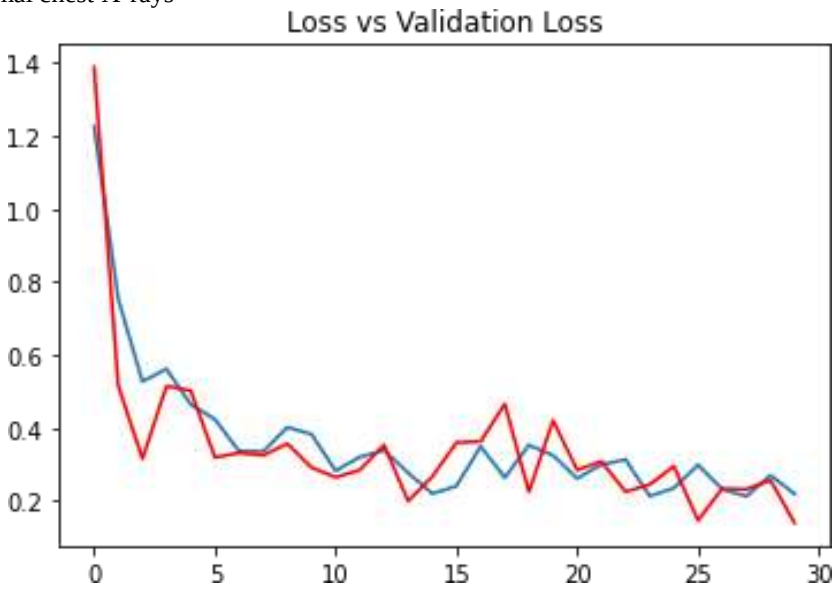

Fig 5:- GraphforLossv/sValidationLossofVGG-16Model.

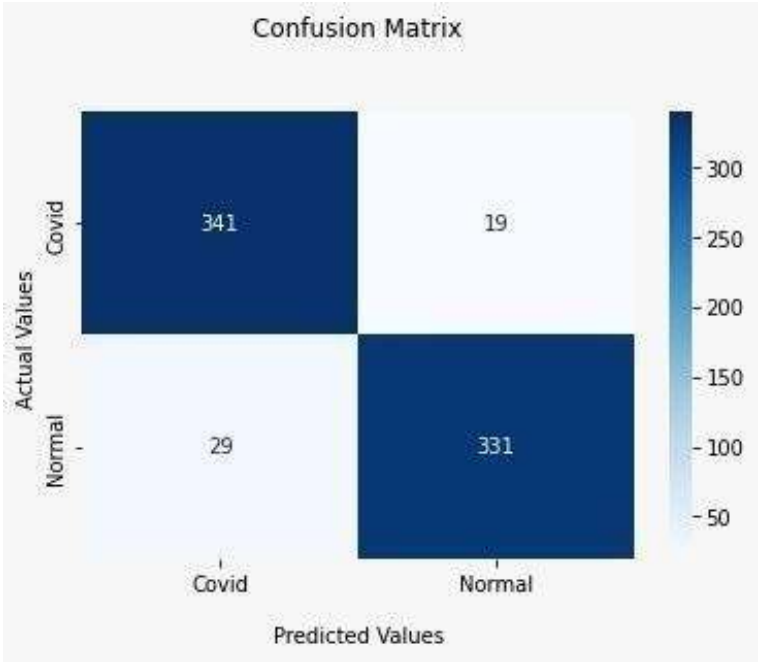

Fig. 6: - ConfusionmatrixforVGG-16: The model shows balanced sensitivity and specificity, with relatively low misclassification rates.

Conclusion:-

To prevent the spread of a global pandemic such as COVID-19, mass testing and early detection are one of the most important steps. To do so factors such as cost, time, accuracy are important to consider. To tackle these issues, a CNN based model is proposed for detecting COVID-19 patients using CXR images. The CNN model is trained with a dataset containing 10,100 CXR images. The images which are used in Dataset are obtained from Kaggle website. Since researchers have proved that CNN has great scope in medical domain the author decided to build a model using CNN technique which will use very limited time, cost and resources for detection of COVID-19 cases. The proposed model will detect the image and classify it as COVID-19 or as normal. With this performance the proposed model can certainly be applied for early detection of COVID-19, helping in decreasing the testing cost and time.

References:-

[1] M.-Y. Ng, E. Y. P. Lee, J. Yang et al., "Imaging profile of the COVID-19 infection: radiologic findings and literature review," Radiology: Cardiothoracic Imaging, vol. 2, no. 1, Article ID e200034, 2020.

[2] S.Wang,B.Kang,J.Maetal.,"AdeeplearningalgorithmusingCTimagescreenforcoronavirus disease (COVID19)." Preprint, 2020

[3] O. Gozes, M. Frid-Adar, H. Greenspan et al., "Rapid AI development cycle for the coronavirus (COVID-19)pandemic: initialresultsforautomateddetection&patientmonitoringusingdeeplearningCT image analysis," 2020. https://arxiv.org/abs/2003.05037.

[4] E. S. Amis, P. F. Butler, K. E. Applegate et al., "American college ofradiologywhite paper on radiationdosein medicine,"JournaloftheAmericanCollege ofRadiology, vol. 4,no.5,pp.272–284, 2007.