IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper  Volume 11, Spl Iss 5, 2022

# Covid-19 Detection in X-Ray Images UsingDeepLearningAlgorithms

Anusha Marouthu[1], SireeshaMoturi[2],Srikanth Vemuru[3]

[1]Associate Professor, Department of CSE, Koneru Lakshmaiah Education Foundation (KLEF), Vaddeswaram, Green fields, Guntur, Andhra Pradesh, India -522302

[2]Assoc. Professor, Department of CSE, Narasaraopeta Engineering College, Narasaraopeta, Guntur (D.T) A.P, India

[3]Professor, Department of CSE, Koneru Lakshmaiah Education Foundation (KLEF), Vaddeswaram, Green fields, Guntur, Andhra Pradesh, India -522302

**DOI : 10.48047/IJFANS/V11/Splis5/43**

***ABSTRACT-TheCOVID-19 pandemic had a variety of effects on global health, the global economy, and global lifestyle. Thus, it is essential to identify viruses early in order to treat patients more effectively. With the help of deep learning methods such convolution neural networks (CNN), VGG16, and VGG19, thisresearchwillexamine the detectionof COVIDvirusinx-raypictures. The actual diagnosis test, called RT-PCR for reversetranscription polymerase chain reaction, is quite expensive andtakes a long time to get results. Thus, additional sophisticatedtesting and diagnostic instruments are required. Inspired by therecent research that is used to detect the COVID-19 presence inthe X-ray images, this research uses deep learning methods andalgorithms to evaluate these images and classify them as covid positive and covid negative cases respectively. The proposed approach includes the preprocessing of the x-ray images which includes removing of their relevant surroundings and bias producing results. After the preprocessing stage, training the classification model under the transfer learning scheme, and outputs are analysed and interpreted through visualization. In this approach, we achieved the accuracy of 95% using the CNN model.***

***KEYWORDS:DeepLearning,SupervisedLearning,Convolutional Neural Networks, VGG16, VGG19 Preprocessing,classificationmodel,Transfer Learning.***

## I. INTRODUCTION

COVID-19 is a virus which arised in wuhan, china in February of 2020. Although there was no greater spread in initial days. Later on, the spread of virus globally has increased vigorously and great fear arised due to no cure was invented. Later on, the practitioners performed many tests todetect the virus in people. Some of the tests are detectingthrough analysis of blood samples, figuring the symptoms ofvirus in people and RT-PCR test. These tests are time takingprocess and highly expensive and some are limited to the laboratory only. Many researches have been performed ondetection of COVID-19 from the chest x-ray images. Thecorona virus has similar virus pathologies to the pneumonia bacteria [1]. RT-PCR test is the main test that takes bloodsamples and detects the corona virus [2]. Many others studied the correlation with chest CT and chest x-ray images [3]. There are many artificial intelligence researches that use deep learning methods that can be applied to medical images processing.

Artificiallearningisthesupersetofmachinelearninganddeeplearning.Nowadays, there are many applications of the deep learning like drive-less cars, healthcare [4,6], chat bots [5], entertainment, etc.

In deep learning, the computer models acquire theknowledge of how to identify images, text and sound. Largedatasets with labeled data and many neural networks aremainlyusedtotrainthesemodels.Inthisresearch,weused The supervised learning approach suits for our predictive analysis. The model is trained using the labeleddata of our dataset which contains as covid, normal and viral pneumonia. These labels help to handle new data and classify the image into respective classes [7].

The main purpose of this approach is to early detection of the corona virus which helps in providing better treatment as early as possible to the people.



IJFANS International Journal of Food And Nutritional Sciences
Official Publication of International Association of Food and Nutrition Scientists

ISSN PRINT 2319 1775 Online 2320 7876
Research paper  Volume 11, Spl Iss 5, 2022

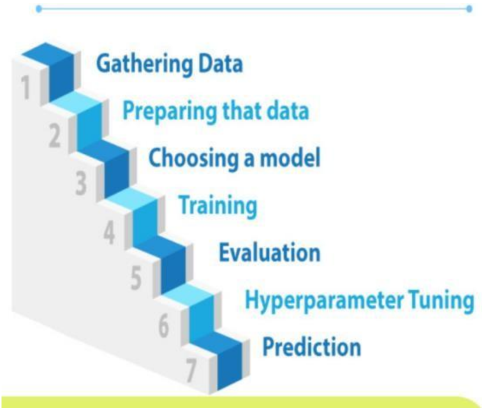



Fig.1.StepsInvolvedInDeepLearning

## II. LITERATURESURVEY

For identifying whether corona virus is present ornot, we need the data which contains different chest x-rayimages of variouspeople which containthevirus affectedimages and normal images which do not contain virus.Thisdataiscrucialinpredictingtheimagetowhichclassitbelongsto.

Thereweremanyproposedsystemswhoresearchedon detecting COVID-19 from chest x-ray images. The deep feature along with support vector machine (SVM) based methodology is recommended in this research for the use in corona virus patient detection utilizing X-ray pictures. In this paper they have achieved the accuracy of 93%. [8] In this paper, they worked on a dataset of 5000 Chest X-ray andperformed transfer learning and trained the model with four networks CNN, ResNet18, Aqueeze Net, and DenseNet-121.In this paper they achieved the accuracy of 90%. [9] In thispaper, they used two-stage approach to detect covid-19.Infirst stage they implemented the ResNet50 architecture andin second stage they used the ResNet101 architecture. In thispaper theyachievedthe accuracyof95%.

The proposed system deals with such a system, which trains the computer-based tasks to detectcovid-19 and classify the results using VGG16 [10],VGG19 and Convolutional Neural Network(CNN)[12] algorithms.

ISSN PRINT 2319 1775 Online 2320 7876
Research paper © 2012 IJFANS. All Rights Reserved, Volume 11, Spl Iss 5, 2022

After pre-processing, the data is chopped into training, testing and validation sets. In this approach we used 3 deep learning models which are CNN, VGG16 and VGG19. Each model is build and trained with the trained data andvalidated against validation data. The evaluation of the three models gave the prediction accuracy more than 90%.

## III. PROPOSEDMETHODOLOGY

Our approach contains of three different tests toassess the execution of the three different models. We usedsame dataset for our three models. The dataset contains thethree directories of covid images, normal images and viralimages.

### *A. Dataset*

We used the dataset that is available at Medical Imaging Data bank of theValencia region BIMCV, 2020 [13]. Our dataset contains total of 7107 images which contain 3617 covid images, 2290 normal images and 1200 viral images.

### *B. Imagepreprocessing*

Our dataset contains images with the different sizes respectively. So, the images have to be preprocessed before training the model [14]. The images are loaded initially andeach image is resized to the width and height of (100,100)respectively with three channels specified. After resizing theimages we are converting them to gray scale for efficientusage of the image and it also helps in elimination of thecomplexities during computational time and simplifies thealgorithms. After that, we rescaled the images array between 0-255 and normalize with a1/255 factor and target for values between 0 and 1.

### *C. ImageSeparation*

For the image classification task, we separated our images into 80 percent training data and 20 percent fortesting and validation data [15]. The training data contains 5685 images and the training and validation images contain 711 and 710imagesrespectively.

### *D. ChoosingTheModel*

Model selection is the process of choosing between different deep learning approaches like CNN [17], ResNet, VGG 16, VGG 19, U-Net, SVM, Random Forest, K-means etc [11].Some of the models used in this work are Convolution neural networks (CNN),VGG 16 and VGG 19.

A deep learning system called convolution neural networks (CNN) can analyze an input image, rate various objects or elements within the image, and differentiate between them. Convolutional, pooling, and fully connected layers make-up a convolution neural network.

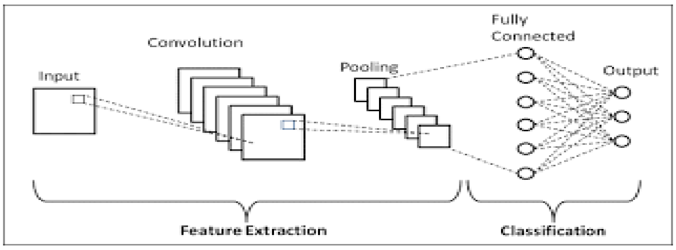



Fig.2.ArchitectureofCNN

ISSN PRINT 2319 1775 Online 2320 7876
Research paper  Volume 11, Spl Iss 5, 2022

TherearetwomainpartsinCNNarchitecture:

a) *FeatureExtraction*is aconvolutiontool thatextracts and detects the various patterns of theimageforanalysis.
b) *Classification*includesfullyconnectedlayerthatutilizes the output from the convolution processand predicts the class of the image based on thefeatureextracted inpreviousstages.

TherearedifferentlayersinCNN,whichare

*Convolutinal Layer:* This layer performs a mathematicaloperationscalledconvolutionbetweentheinputandafilterofaspecificsizeMxMinordertoextractfeaturesfromtheinputimages.

*PoolingLayer:*Thisisusedtoreducecomputationalcosts,thisisutilizedtolowerthesizeoftheconvolvedfeaturemap.

*Fully Connected Layer:* In order to connect the neuronsbetween two layers, this layer used. The hyper parameters used in Convolution layers have a 3x3 kernel size with the same padding. In Max pooling layers, where the pool size is 2x2, the dropout rate is 0.1 in the first two expansion and contraction blocks, 0.2 in the third and fourth, and 0.3 in the fifth contraction block. Kernel size, strides, and padding for transposed convolution layers are all 2x2.

A typical deep Convolutional Neural Network (CNN) design with numerous layers is Visual Geometry Group, also known as VGG. There are a total of 13 convolution ans 3 fully connected layers in VGG16. For difficult classificationtasks, the incredibly deep VGG19 has been trained on anenormous variety of pictures. The VGG19 consists of 19 layers, of which 3 are fully linked layers and 16 are convolutional layers.

*E. EVALUATINGTHE MODEL*

After loading the dataset and performing the image processing to convert each image to array, the data is splitinto subsets with 80% training data and 20% testing data. After that, the model is build with "Adam"optimizer with categorical cross entropyloss. Adamoptimizer is used to update the learning rate for each network respectively. Here, we used the categorical cross entropy forconverting true labels to one hot encoded. These true labels are transformed into integer encoded. In our approach they are converted into 0, 1, and 2 since our problem is a 3-classproblem.

Since we used image Data generator to send the data to the model, we are using the model. fit generator method to fit the model by keeping an eye on a particular models parameter in this situation, Model Checkpoint enables us to save the model. Only when the model's current validation lossis less than its previous validation loss will the model be stored to disc. Evaluation on the test data gives the accuracy of 93% and loss of 0.457 for CNN model. Similarly, for VGG16 weachieved the accuracy of 87% and loss of 0.20 and for VGG19model we achieved the accuracy of 91% and loss of 0.1690.After that we plotted the Training and Validation Acccuracy graphs and Training and Validation Loss graphs for respective models.

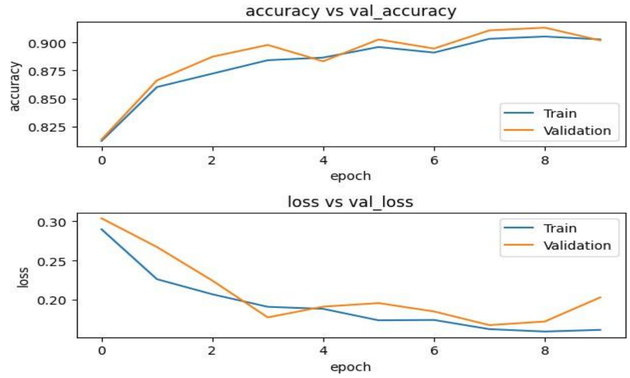
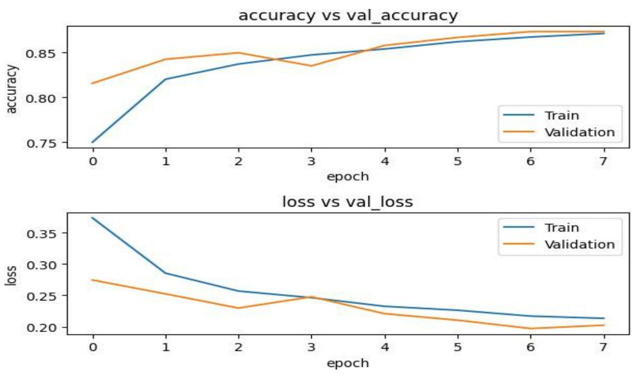



Fig.3.Trainingandvalidationaccuracyandlossgraphsforvgg19model.



ISSN PRINT 2319 1775 Online 2320 7876
Research paper  Volume 11, Spl Iss 5, 2022

Fig.4.Trainingandvalidationaccuracyandlossgraphsforvgg16model.

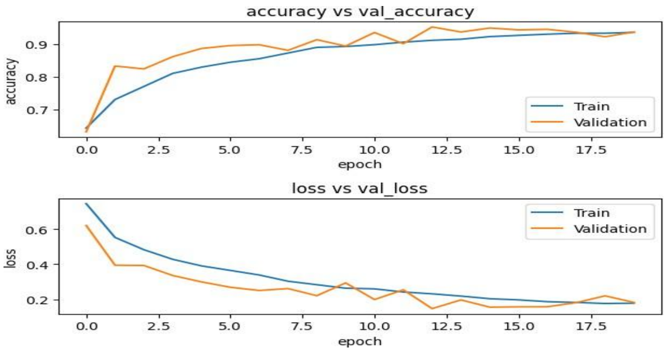



Fig.5.TrainingandValidationAccuracyandLossgraphsforCNNModel

The Fig.2, Fig.3 and Fig.4 represent the Trainnig and Validation Accuracy graphs and Training and Validation Loss graphs. Our model does not seem to be over fitted. From the Training and Validation Loss graphs indicates the good learning rate. After evaluating the model we predicted fewimagesrandomlyonourmodelwhichgavegoodresults.

*F. HARDWAREANDSOFTWARE*

We used Python 3.8.1 to create this application. Using the Keras package and TensorFlow, all models were created. For the majority of the trials, we used Google Co-laboratory. Tensor Processor Units (TPU) were utilized in this example whenever possible, in the absence of a TPU, we used Graphic Processor Units (GPU) in accordance with the Co-laboratory assignment. We used the visual studio code torender the flaskapplication.

ISSN PRINT 2319 1775 Online 2320 7876

*Research paper*  **Volume 11, Spl Iss 5, 2022**

## IV. EXPERIMENTALRESULTS

Testing our models against some random samplesfromtestdatagavethefollowingresultsrespectively.

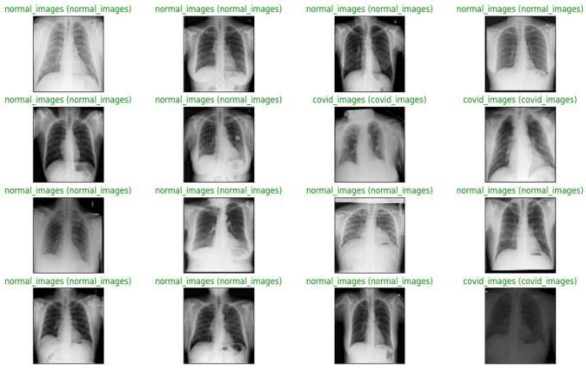



Fig.6.PredictionresultforsomerandomtestimagesusingCNNmodel

Here, figure 11 which are prediction results containtwo different classes which are green and red. The predictions which are correct are represented in green color and the predictions which are not predicted correctly are representedinredcolor.

A confusion matrix is used to demonstrate the performance of a categorization system [16]. The output of a classification method is shown and condensedin a confusion matrix. The classification report contains the metrics Precision which is the ratio of true positive to the sum of true and false positives, Recall which is the ratio of true positives and false negatives, F1-Score which is harmonic mean of both recall and precision and Support. The confusion matrices for each model are represented in figures 6, 7and 8.

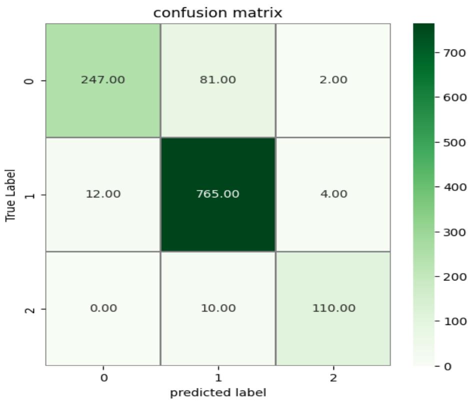



Fig.7. ConfusionmatrixforVGG19Model

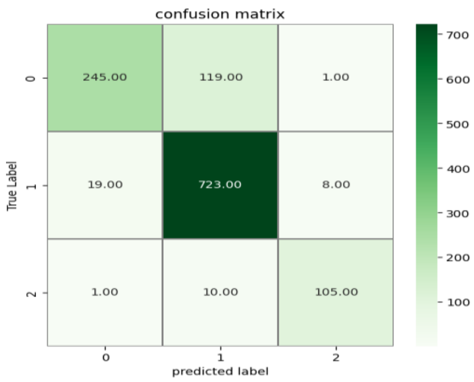
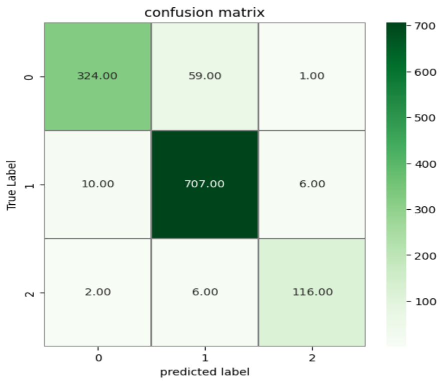
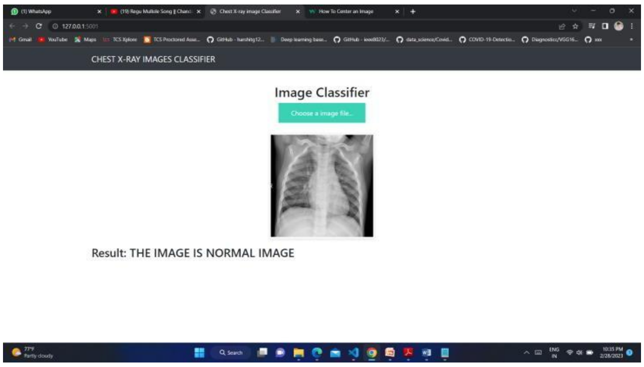



Fig.8.ConfusionmatrixforVGG16 Model.


confusion matrix
324.00
59.00
1.00
10.00
707.00
6.00
2.00
6.00
116.00
True Label
predicted label
0
1
2
700
600
500
400
300
200
100


CHEST X-RAY IMAGES CLASSIFIER
Image Classifier
Result: THE IMAGE IS NORMAL IMAGE

IJFANS INTERNATIONAL JOURNAL OF FOOD AND NUTRITIONAL SCIENCES

ISSN PRINT 2319 1775 Online 2320 7876

Research paper  Volume 11, Spl Iss 5, 2022

Fig.9.ConfusionmatrixforCNNModel.

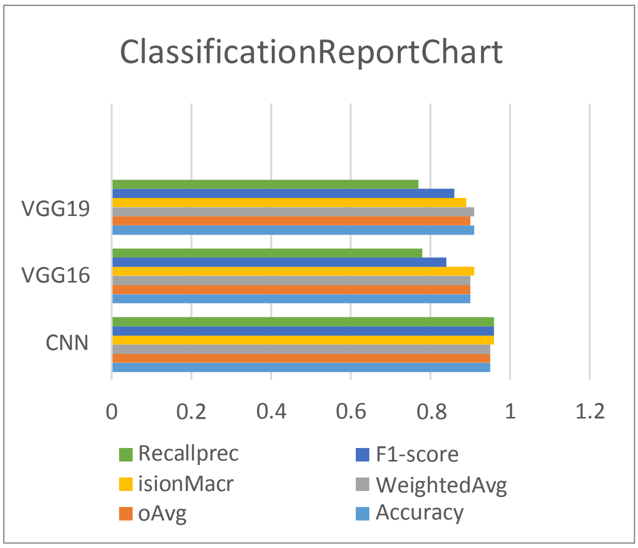



Fig.10.Classification report chart for CNN,VGG 16 and VGG 19.

The fig.9 represents that the accuracy and precision for the CNN model is more. Although vgg 16 and vgg 19 models gave good results, they are not as efficient as CNN model.

## V. OUTPUTSCREENS

The figures 10, 11 and 12 represent the output results for predicting covid image, normal image and viral image. The output screen contains an option to choose a file and a predict button to predict the image.

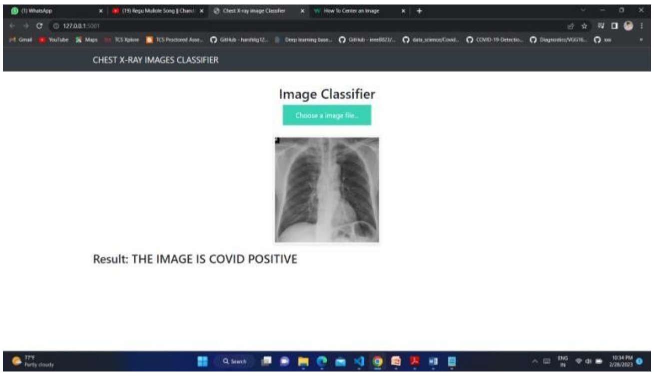



Fig.11.Resultofpredictionforcovidimage

ISSN PRINT 2319 1775 Online 2320 7876
Research paper  Volume 11, Spl Iss 5, 2022

Fig.12.Result of prediction for normal image

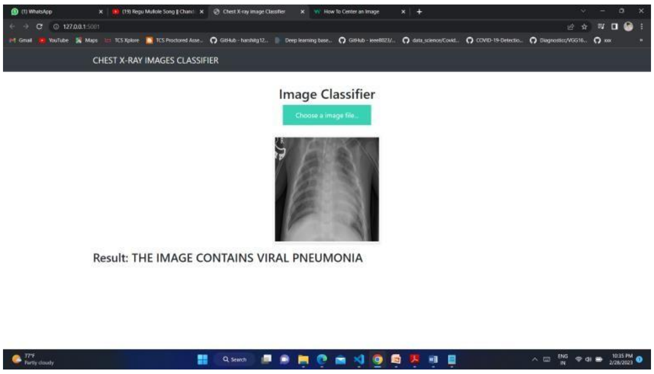



Fig.13.Result prediction for viral pneumonia image

## VI. CONCLUSION

This method demonstrates how existing models can be useful for a variety of applications, especially when it is taken into account that the VGG models are not more productive. Additionally, it demonstrates how model bias can be caused by image noise. The majority of metrics indicate that segmented pictures are less effective at classifying COVID illness. A further investigation reveals that even though measures are improved, these models are still reliant on obvious diseases throughout the lungs, which are prominent indicator of COVID. As a result, true correct models mustfocus on lung sections for classification. In this situation, segmentation is required to less enthis bias and produce credible results. For the findings shown, transfer learning was essential. As demonstrated, this method's classification models require 20 to 30 epochs to converge. We have achieved the average accuracy of 90% to identify COVID-19 disease in chest x-ray images with respective classes COVID images, Normal images and Viral pneumonia images.

## VII. FUTUREWORKS

The first is semantic segmentation of lung diseases, namely the recognition of consolidation and ground window opacity problems. This is one of the two major topics we have highlighted for further study. In order to summarize data andget over problems with using different picture sources, these on-stage requires extending COVID datasets.

## VIII. REFERENCES

1. Wang D, Hu B, Hu C, et al. Clinical Characteristics of138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan,China. *JAMA*. 2020; 323(11):1061–1069. doi:10.1001/jama.2020.1585
2. SireeshaMoturi, S.N.Tirumala Rao, SrikanthVemuru, “Grey wolf assisted dragonfly-based weighted rule generation for predicting heart disease and breast cancer,” Computerized Medical Imaging and Graphics, Vol. 91, pp. 1-17, June 2021.
3. Mohan, Senthilkumar&Chowdhary, Chiranji. (2019).AnAI-BasedChatbotUsingDeepLearning.10.1201/9780429265020-12.
4. M.Sireesha, S. N. TirumalaRao, SrikanthVemuru, Optimized Feature Extraction and Hybrid Classification Model for Heart Disease and Breast Cancer Prediction,” International Journal of Recent Technology and Engineering, Vol. 7, No. 6, pp. 1754-1772, Mar. 2019.
5. F. Ertam and G. Aydın, "Data classification with deeplearning using Tensorflow," 2017 International Conferenceon Computer Science and Engineering (UBMK), Antalya,Turkey, 2017, pp. 755-758 doi:10.1109/UBMK.2017.8093521.
6. ShervinMinaee, RaheleKafieh, Milan Sonka, ShakibYazdani,GhazalehJamalipourSoufi,Deep-COVID:Predicting COVID-19 from chest X-ray images using deeptransferlearning,MedicalImageAnalysis,Volume65,2020,101794,ISSN, 1361-8415,https://doi.org/10.1016/j.media.2020.101794.
7. GovardhanJain,DeeptiMittal,DakshThakur,Madhup K.Mittal,AdeeplearningapproachtodetectCovid-19coronaviruswithX-Rayimages,BiocyberneticsandBiomedicalEngineering,Volume40,Issue4,2020,Pages1391-1405, https://doi.org/10.1016/j.bbe.2020.08.008.
8. Jiang,Z.-P.;Liu,Y.-Y.;Shao, Z.-E.;Huang,K.-W. An Improved VGG16 Modelfor Pneumonia ImageClassification. Appl.Sci.2021,11,11185.https://doi.org/10.3390/app112311185

ISSN PRINT 2319 1775 Online 2320 7876


9. SakshiIndolia,AnilKumarGoswami,S.P.Mishra,PoojaAsopa,ConceptualUnderstandingofConvolutionalNeuralNetwork-ADeepLearningApproach,ProcediaComputerScience,Volume132,2018,Pages679-688,ISSN1877-0509,https://doi.org/10.1016/j.procs.2018.05.069.
10. MedicalImagingDatabankoftheValenciaregionBIMCV(2020).BIMCV-Covid19–BIMCV. https://bimcv.cipf.es/bimcv-projects/bimcv-covid19/{#}1590859488150-148be708-c3f3.
11. V. Gulati and N. Raheja, "Efficiency Enhancement ofMachineLearningApproachesthroughtheImpactofPreprocessingTechniques,"20216thInternationalConference on Signal Processing, Computing and Control(ISPCC),Solan,India,2021,pp.191-196,doi:10.1109/ISPCC53510.2021.9609474.
12. J. Ker, L. Wang, J. Rao and T. Lim, "Deep LearningApplications in Medical Image Analysis," in IEEE Access,vol.6,pp.9375-9389,2018,doi:10.1109/ACCESS.2017.2788044.