Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

International Journal of Intelligent Networks 1 (2020) 112–118

Contents lists available at ScienceDirect

International Journal of Intelligent Networks

journal homepage: www.keaipublishing.com/en/journals/international-journal-of-intelligent-networks

Detection and identification of COVID -19 based on chest medical image by using convolutional neural networks

Jothi V. Pranav [a,*], R. Anand [b], T. Shanthi [b], K. Manju [b], S. Veni [c], S. Nagarjun [d]

[a] School of Electrical and Electronics Engineering, Nanyang Technological University, Singapore
[b] Sona Signal and Image Processing Research Center(SIPRO), Department of ECE, Sona College of Technology, Tamil Nadu, India
[c] Department of ECE, Amrita Vishwa Vidyapeetham, Amrita School of Engineering, Coimbatore, India
[d] Department of ECE, Sona College of Technology, Salem, Tamil Nadu, India

ARTICLE INFO

Keywords:
Deep learning
COVID-19
Convolutional neural network
X-ray images
Machine learning: image processing

ABSTRACT

Covid-19 pandemic has caused major out-break all around the world. This pandemic out-break requires lot of testing, which is a tedious process. Deep learning is a successful method that has evolved in image category in the past few years. In this work to detects the presence of coronavirus by using deep learning approach. Here, convolutional neural networks with specific focus on to classify Covid-19 chest radiography images. The database comprises Covid-19, normal and viral pneumonia chest X-ray images with 800 different samples under each class. We evaluated the model on 500 images and the networks has achieved a sensitivity rate of 95% and specificity rate of 97%. The DenseNet121 Architecture performed slightly better, compared to other state of art networks. The performance achieved by the method proposed is very encouraging and the accuracy rates can be improved further with larger datasets. Apart from sensitivity and specificity rates, the proposed model is also compared on receiver operating characteristic (ROC), and area under the curve (AUC) of each model. The model is implemented on the TensorFlow framework with the datasets that are publicly available for research community.

1. Introduction

Last few months, we have come across a deadly virus named as coronavirus which is affecting many people worldwide. In this paper, we propose an algorithm for identifying the presence of coronavirus based on chest X-ray medical images. In this paper, the deep learning algorithms [1,2] are employed for disease identification. This work has been carried out with a focus to identify presence of Covid-19 virus and classify the chest X-ray medical images as healthy, pneumonia, covid19 [3]. We have approached the problem of identifying Covid-19 by utilizing deep learning architectures like VGGNet19 and DenseNet121.

The convolutional neural network with its promising accuracy in image classification problem have attracted many researchers towards it. Apurva A. Desai [4] has used a neural network type of classification and compared it with other machine learning approaches. This literature shows that neural network-based approach gives a better result when compared to other approaches. Apurva A. Desai has done this comparison for optical character recognition. Hidden Markov Model and fuzzy logic algorithm has been used on Urdu scripts by Muhammad Imran Razzak [5] for character recognition. The approach was evaluated on Nastaliq and Nasakh script-based languages. With around 26 time-invariant features and statistical features, the approach provided 87.4% and 74.1% of accuracy for Nastaliq and Nasakh scripts respectively. An effective optical character recognition system to recognize and identify the Persian character using Support Vector Machine (SVM) was proposed by Abdelhak Boukharouba [6]. Using chain code algorithm, a transition table in vertical and horizontal direction is extracted. The extracted feature is used for classification using an SVM classifier. Parshuram M. Kamble [7] have used a Feed Froward Artificial Neural Network (FFANN) for Marathi character recognition. Yang, Xu Lijia [8], combined statistical and structural features to improve the accuracy when compared to other features. Additionally, they have utilized 184 epochs in CNN for increasing the accuracy. Amit Choudhary [9], describes binarization feature extraction for English characters which increases the prediction rate in a Back Propagation Neural Network (BPNN).

Several researchers have geared up their research for the recently arrived pandemic Covid-19. The author Chaolin Huang [10] and his colleagues have reported that, 41 infected people are admitted to the

* Corresponding author.
E-mail addresses: pranavjo001@e.ntu.edu.sg (J.V. Pranav), anand.r@sonatech.ac.in (R. Anand), shanthi@sonatech.ac.in (T. Shanthi), manju.kandaswamy@gmail.com (K. Manju), s_veni@cb.amrita.edu (S. Veni), nagaarjun2512000@gmail.com (S. Nagarjun).

https://doi.org/10.1016/j.ijin.2020.12.002
Received 1 November 2020; Received in revised form 9 December 2020; Accepted 9 December 2020

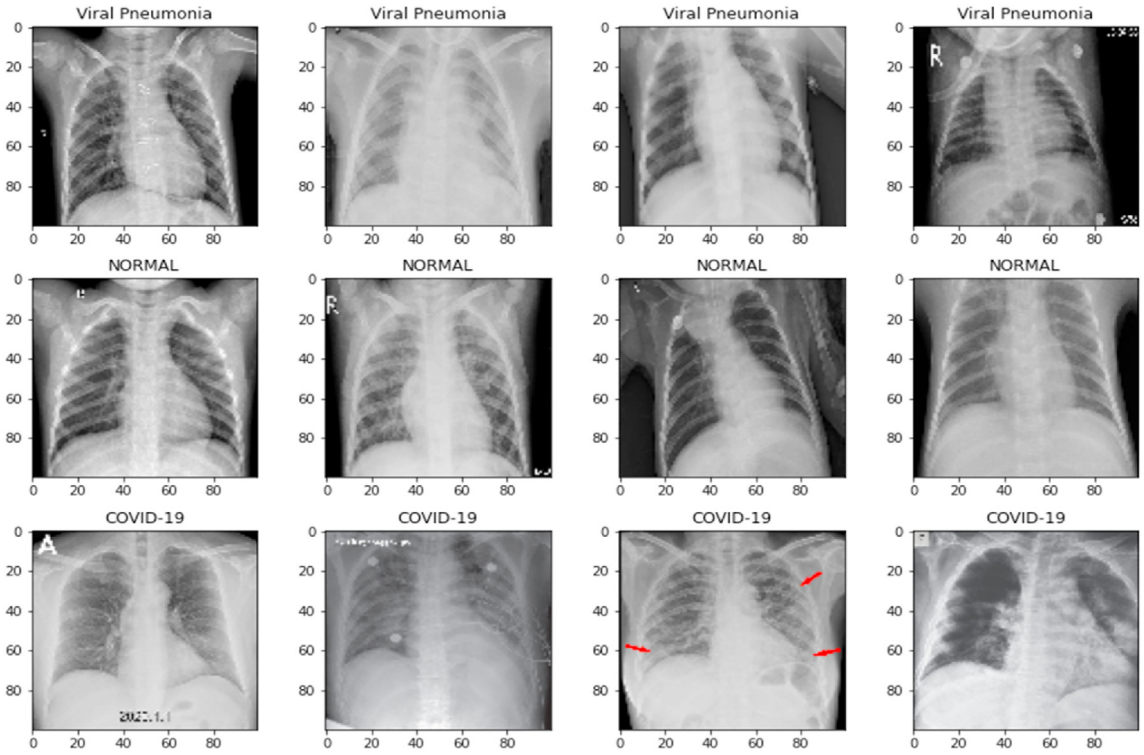


Fig. 1. Sample images of chest X-Ray medical image dataset.

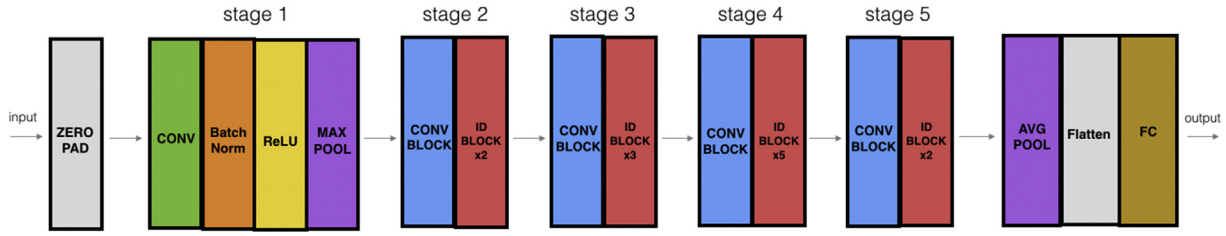


Fig. 2. ResNet-50 architecture.

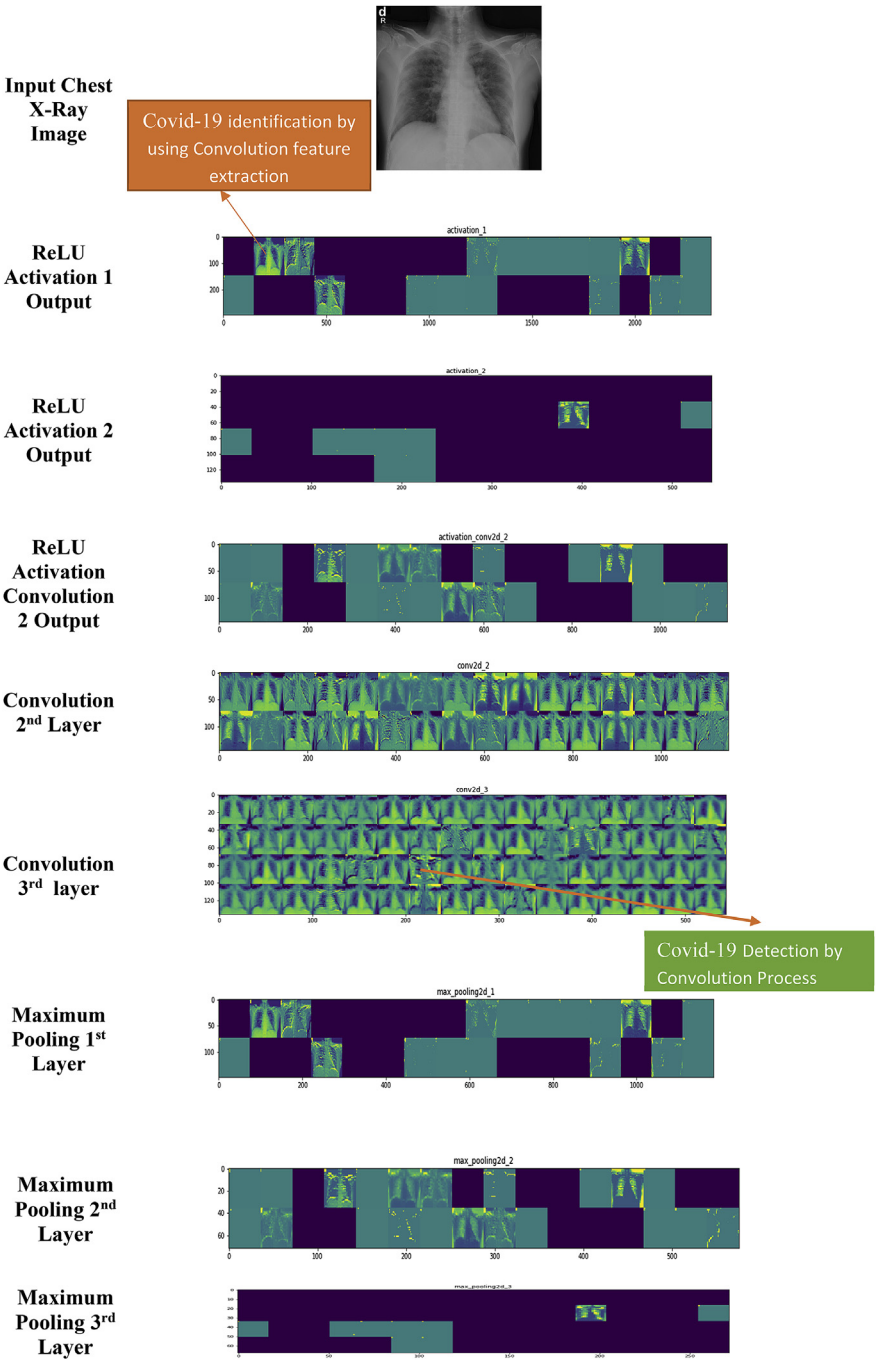
Table 1
Summary of DenseNet121 architecture.

Layers	Output size	DenseNet-121	DenseNet-169	DenseNet-201	DenseNet-264
Convolution	112 × 112	7 × 7 conv, stride 2			
Pooling	56 × 56	3 × 3 max pool, stride 2			
Dense Block (1)	56 × 56	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 6$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 6$
Transition layer (1)	56 × 56	1 × 1 conv			
	28 × 28	2 × 2 average pool, stride 2			
Dense Block (2)	28 × 28	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 12$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 12$
Transition layer (2)	28 × 28	1 × 1 conv			
	14 × 14	2 × 2 average pool, stride 2			
Dense Block (3)	14 × 14	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 24$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 48$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 64$
Transition layer (3)	14 × 14	1 × 1 conv			
	7 × 7	2 × 2 average pool, stride 2			
Dense Block (4)	7 × 7	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 16$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 32$	$\begin{bmatrix} 1\times1 \text{ conv} \\ 3\times3 \text{ conv} \end{bmatrix} \times 48$
Classification layer	1 × 1	7 × 7 global average pool 1000D fully connected, Softmax			

hospital in Wuhan City, which is confirmed to be infected with 2019-covid. Covid-19 [11] has some symptoms like fever, dry cough, malaise and some nonspecific symptoms. The respiratory system in human body is infected slowly. So, we have taken the Chest X-Ray medical images from infected patients as well as normal patients [12,15].

The paper progresses with a detail description of the dataset in section 2. Section 3 talks about the overview of layers present in Convolution neural networks. In Section 4 different CNN architectures with their advantages and disadvantages are discussed in brief. The result and discussions are done in section 5. The conclusion of the paper is given in section 6.

Table 2
Feature extraction to identify COVID 19.



2. Dataset description

The dataset used in this paper combines two different datasets. The total data for training and testing was around 3700. For training the model, 2200 images have been used and for validation 1500 images were used. The data is split between three different classes Covid-19, Normal and Viral pneumonia. which is shown in Fig. 1. Here, we used our learning rate is 0.0001, number of epochs is 20, and batch size will be 50 for tuning convolutional neural networks. Data augmentation (by rotation, zoom, flipping, shifting) is performed to increase the dataset while training the model. ImageNet mean value (in RGB order) is used to during augmentation. The data was obtained from IEEE802.3 collected by Joseph Paul Cohen, which contains data of two classes including Covid-19 images. Remaining data are obtained from Covid-19 radiography database. There are 219 Covid-19 positive images, 1341 normal images, 1345 viral pneumonia images. Three different architectures like, VGGNet19 [13], and DenseNet121 [14], ResNet50 were used for diagnosis which is shown in Fig. 2. The following section shows, overview of convolution layer, pooling, ReLU (Rectified Linear Unit) layers.

3. Overview of layers present in convolution neural networks

Convolution neural network (CNN) is like a regular neural network with several layers stacked in between the input and output layer. The

CNN is comprised of alternative layers like convolution layer, pooling layer, activation layers etc [17].

The convolution layer (conv layer) [23] is named since it performs mathematical convolution operation between the input matrix and filters. The dimension of the feature maps (W × W) produced by these conv layers depends on the dimension of the input matrix(M × M), filter size (F), padding (P), stride(S) and number of filters (K). The dimension can be computed with the equation no. 1.

$$W = \frac{M - F + 2ZP}{S} + 1 \quad (1)$$

Pooling layer [21] generally used for lessening the aspect of the feature maps. This decreases the number of learnable parameters to a greater extent [22]. The pooling function can be either max pooling or average pooling. The max pooling chooses the maximum value from the selected window and the average pooling chooses the average value from the selected window. Activation layers are inserted between the layers to introduce nonlinearity in between the layers [25]. The ReLU function is preferred in most of the networks as it is simple and efficient to implement.

4. Deep learning architectures

4.1. VGGNet19 architecture

Image classification and detection is one of the difficult tasks when it comes to real time applications. The deep learning approach with the state-of-the-art predefined architectures eases this task. One such architecture is VGGNet19 (Visual Geometry Group). The input image size to the architecture is [224 × 224]. VGGNet19 has 19 layers that includes 16 convolutional layers,2 fully connected layers and a softmax layer [24], [25] The first two layers are conv layers with [3 × 3] filters with stride of 1. This outputs 64 feature maps with a size of [224 × 224], which is then passed to the max pooling layer with a size of [2 × 2] and stride of 2. This will reduce the size to [112 × 112 × 64]. Then the feature maps are passed through a couple of conv layers. The number of filters used at this stage is 128 and the dimension is of [112 × 112]. Then a pooling layer is added that results in feature map of dimension [56 × 56]. Then it is passed through 2 conv layers with 256 filters and a pooling layer, this reduces the size to [28 × 28] with 256 filters. Then it is passed through 4 conv layers with 512 filters and then pooling layer, that outputs a feature map of size [14 × 14]. Further it is passed through 4 conv layers. The size of the last conv layer is [7 × 7] with 512 filters. Now, they are connected to a two fully connected layers with size of 4096. The FC layer is then connected to a softmax layer of size 1000. The softmax layer will classify the image based on the threshold and the features stored in the neurons.

4.2. DenseNet121 architecture

DenseNet121 [16,22] is a simple and efficient network that is constructed by stacking dense block and pooling layers alternatively. Table 1 shows the summary of the architecture. In DenseNet121 architecture a feed-forward connection is established from each layer to every other layer directly. A regular convolutional network with N layers has N connections between each subsequent layer, whereas the DenseNet121 has N(N+1)/2 straight connections. The DenseNet121 adopts a shorter connection between the initial and final layers that results in notable improvement in the accuracy and efficiency of the network. As the input progress through the subsequent layers, the feature maps from each layer is concatenated with the previous one. This enables the access of feature

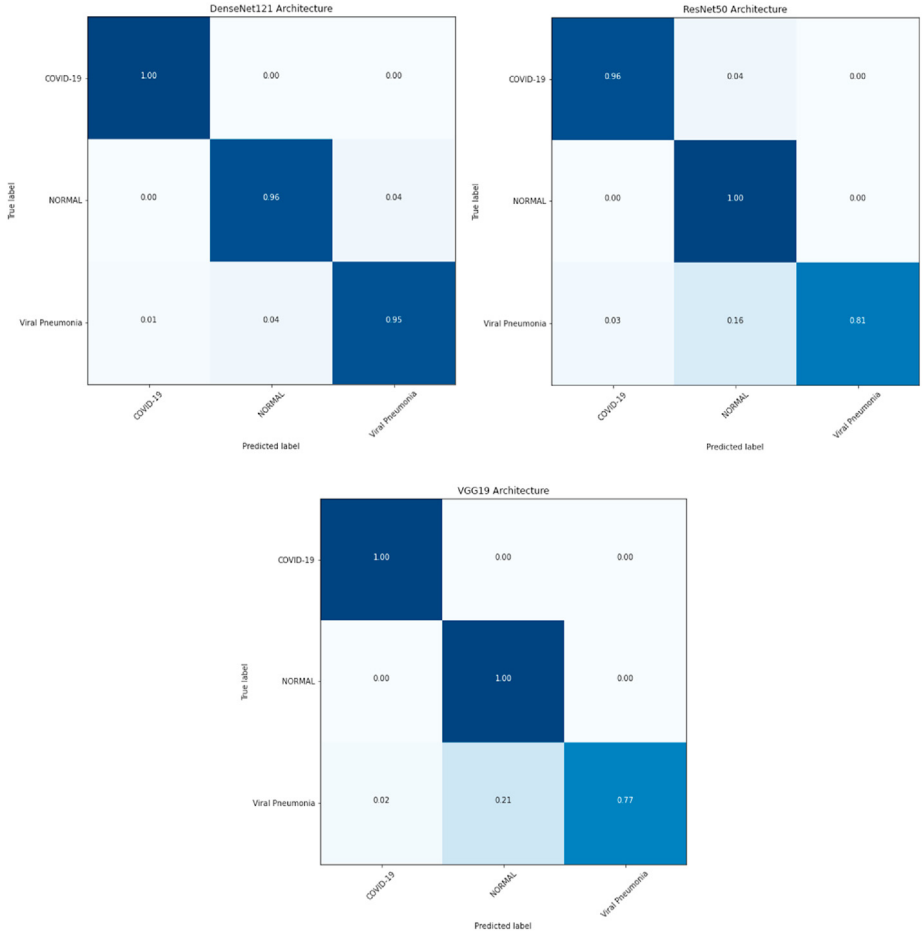


Fig. 3. Confusion matrix of three architectures.

maps of any layer by the subsequent layers. The number of parameters is minimum when compared to a traditional convolutional neural network. The DenseNet121 reduces the vanishing gradient problem and implements feature reusing. The regularization effect of the network reduces the overfitting problem to a greater extent even for a smaller dataset which is shown in Table 2.

4.3. ResNet-50 architecture

ResNet50, or Residual Networks is a popular neural network used in many computer vision tasks. This model has won the ImageNet challenge in 2015. Prior to ResNet50, training a deep neural network was difficult due to the problem of vanishing gradients. ResNet50 uses 150+ layers, which made it more successful than other networks [18]. ResNet50 has outperformed many models on ImageNet task with its depth in number of layers. Deeper the layers, the network must learn more parameters that will lead to the problem of overfitting or underfitting.

5. Results and discussion

The chest X-ray images are used to detect the occurrence of coronavirus [26]. Here, three different algorithms are compared based on the classification performance. DenseNet121 has performed better than VGGNet19 and ResNet50 architecture. In DenseNet121, each layer obtains additional inputs from the previous layers and passes it to the subsequent layers. Thus, each layer receives a 'collective knowledge' from all preceding layers. Which leads to more dense architecture with less parameters to train on, compared to other state of the art architectures. The performance metrices like precision, recall, accuracy, sensitivity and specificity are considered for measuring the performance of the architecture [20].

5.1. Model parameters

The batch size is set to 50, and the epochs was limited to 20 as the accuracy and loss function saturates after 20 epochs. All our implementations are done in TensorFlow with publicly available dataset.

5.2. Evaluation metrics

There are various metrics like testing accuracy, sensitivity, specificity, precision, and F1-score which can be used for evaluating the performance of classification models. The current dataset used to train and test the model is imbalanced with few numbers of covid19 images. Sensitivity and specificity will provide a better insight to compare the performance of different models.

5.3. Model sensitivity and specificity

Sensitivity is the proportion of identifying truly positive Covid-19 cases that were classified as positive. Thus, it measures how well the classifier identifies positive cases. Specificity is the proportion of truly negative non Covid-19 cases that were classified as negative. It measures the ability of the classifier to identify the negative cases. The sensitivity and specificity average score of all the three models is 92 % and 90%. This shows that all the models have well performed on the test data, but DenseNet has slightly better results than other two models. DenseNet121 has less parameters for training than other two networks, even though it has more layers. Regularization has been used in all the models to overcome the problem of generalization. ResNet50 and VGGNet19 performed well on validation data without overfitting or underfitting. The below equation gives the formula for sensitivity and specificity, where TP and FN refer to true positive and false negative. TN and FP refer to true negative and false positive based on the confusion matrix obtained for model on test data which is shown in equation 2 & 3.

$$\text{SENSITIVITY} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \tag{2}$$

$$\text{SPECIFICITY} = \frac{\text{TN}}{(\text{TN} + \text{FP})} \tag{3}$$

5.4. Model confusion matrix

Confusion matrix or contingency table is a table that is often used to describe the performance of a classification model on a set of test data for which the true values are known. It is used to compare different models

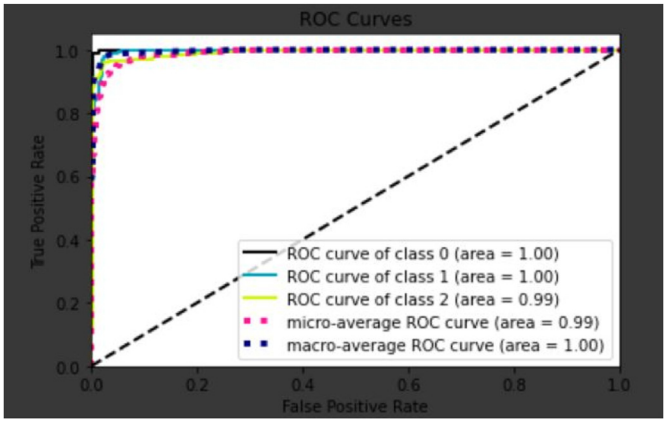


Fig. 4. Shows the ROC curve of DenseNet-121 with micro and macro averages across different classes.

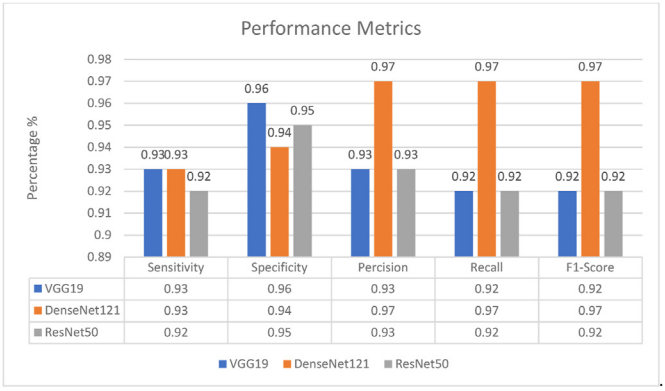


Fig. 5. Performance Metrices for Test data.

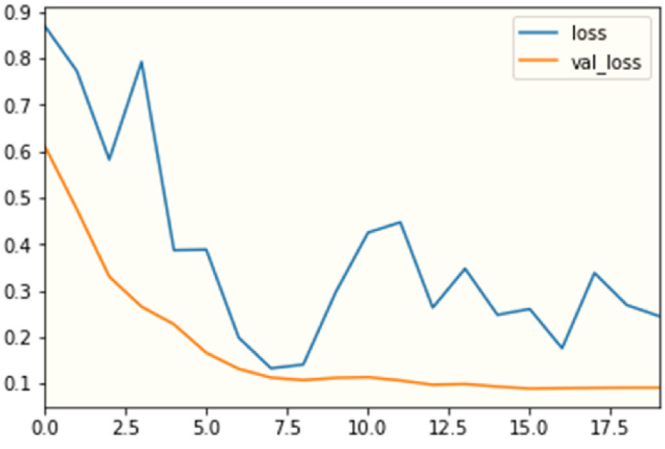


Fig. 6. Training Graph with epoch of 20.

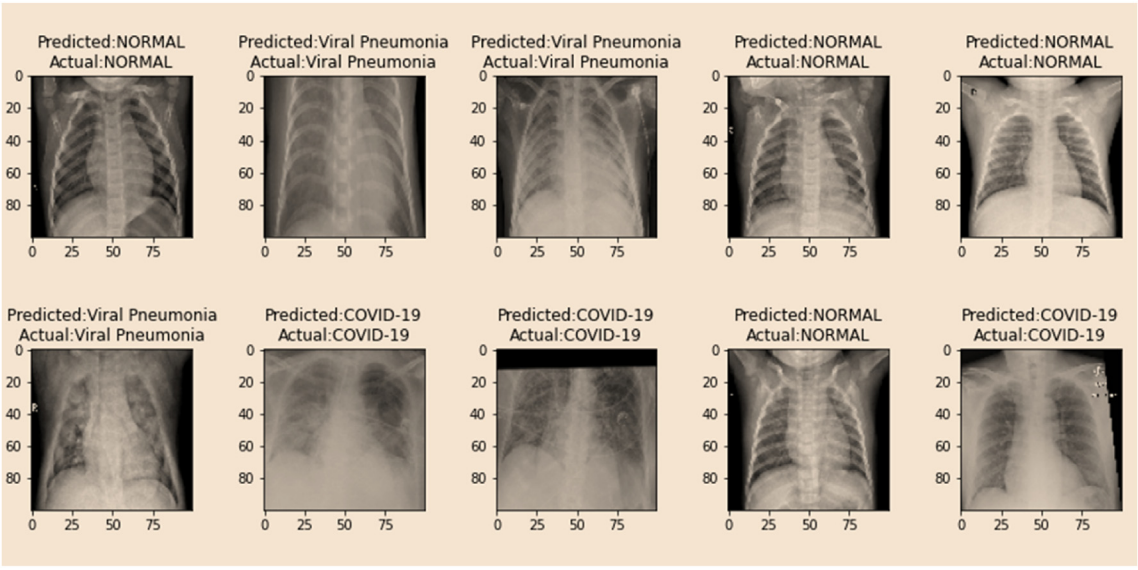


Fig. 7. Predicted result for chest X-Ray images.

based on their performance on the test data. To understand that the model has well generalized for the test data and to tackle the problem of overfitting and underfitting. The confusion matrix shows the way in which your classification model is confused when it makes predictions. A confusion matrix is used to calculate recall, precision, true positive and false positive rates etc. The confusion matrix of three models are shown in Fig. 3. Compared to VGGNet19 and ResNet50 the confusion matrix of DenseNet121 has better results on the test data. It means it predicts the correct labels for Covid-19 and non Covid-19 X-ray images. This shows that the model is well generalized on the training data. Amount of data used for testing is 1/3 of the total Covid-19 data. Fig. 3 shows the confusion matrix of all the three models on the test dataset.

5.5. Model ROC curve

The receiver operating characteristic curve (ROC) is a graph that compares the performance of a classification model with different classification threshold values. It compares the calculated true positive rate and false positive rate from the confusion matrix. AUC stands for 'Area under the ROC curve'. This measures the entire 2-dimensional area underneath the ROC curve from (0,0) to (1,1). The quality of the model, irrespective of the chosen classification threshold is obtained by the AUC [19]. AUC ranges from 0 to 1. A model whose predictions are 100% correct has an AUC of 1. The macro average of VGGNet19 and ResNet50 is 98% and 97%. To overcome the problem of overfitting, L2 regularization technique is used. Fig. 4 shows the ROC of DenseNet121 which is applied on the test data.

Among the 3700 images 2200 images are used for training and 1500 used for validation. Performance matrices are plotted in Fig. 5 & Fig. 6. From the confusion matrix it is evident that the true positive and true negative are high. The hypothesis is considered as to identify the disease. The true positive is to identify the affected person and true negative is to identify a normal patient. False negative is when a corona affected person is identified as normal and false positive is to identify a normal patient as corona affected person. The performance metrices shows that the accuracy rate of DenseNet121 is more compared to other models. For we have obtained an accuracy of 97% for DenseNet121 and for VGGNet19 it is 94%. Specificity and sensitivity are also high for DenseNet121 when compared to the other architectures. Precision obtained for DenseNet121 is 97% which is shown in Fig. 7.

6. Conclusion

A computer vision system for automatic diagnosis of Covid-19 disease using chest radiographic images is proposed in this paper. Chest X-Ray images of healthy individuals and individuals with specific pathological conditions like Covid-19 and pneumonia are considered for training and testing the performance of the model. DenseNet121 has performed better than other two models. Adam optimizer with a learning rate of 0.0001 and decay rate for 20 epochs is used. The predefined weights of ImageNet model are used. DenseNet121 model with 121 layers need fewer parameters for training rather than other two models. The increased number of learning parameter in Resnet and VGGNet19 leads to overfitting. An optimum value for the number of parameters and layers were arrived to achieve better accuracy. The improvement in accuracy is obtained by increasing the data size and by providing ground truth information for inter class images. The accuracy of the model can be further improved by a bigger dataset and fine tuning the network parameters.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] R. Anand, T. Shanthi, R.S. Sabeenian, S. Veni, Real time noisy dataset implementation of optical character identification using CNN, Int. J. Intell. Enterprise 7 (1–3) (2020) 67–80.
[2] R. Anand, T. Shanthi, M.S. Nithish, S. Lakshman, Face recognition and classification using GoogleNET architecture, in: Soft Computing for Problem Solving, Springer, Singapore, 2020, pp. 261–269.
[3] R.S. Sabeenian, M.E. Paramasivam, R. Anand, P.M. Dinesh, Palm-Leaf manuscript character recognition and classification using convolutional neural networks, in: Computing and Network Sustainability, Springer, Singapore, 2019, pp. 397–404.
[4] Apurva A. Desai, Gujarati handwritten numeral optical character reorganization through neural network, Pattern Recogn. 43 (7) (2010) 2582–2589.
[5] Razzak, Muhammad Imran, et al., HMM and fuzzy logic: a hybrid approach for online Urdu script-based languages' character recognition, Knowl. Base Syst. 23 (8) (2010) 914–923.
[6] Abdelhak Boukharouba, Abdelhak Bennia, Novel feature extraction technique for the recognition of handwritten digits, Applied Comput. Infomat. 13 (1) (2017) 19–26.
[7] Yang Yang, Lijia Xu, Cheng Chen, English character recognition based on feature combination, Procedia Eng. 24 (2011) 159–164.

[8] Amit Choudhary, Rahul Rishi, Savita Ahlawat, Off-line handwritten character recognition using features extracted from binarization technique, in: AASRI Procedia, vol. 4, 2013, pp. 306–312.
[9] Rapeeporn Chamchong, Chun Che Fung, Character segmentation from ancient palm leaf manuscripts in Thailand, in: Proceedings of the 2011 Workshop on Historical Document Imaging and Processing, ACM, 2011.
[10] Na Zhu, et al., A novel coronavirus from patients with pneumonia in China, 2019, N. Engl. J. Med. Feb 20;382(8) (2020).
[11] Chaolin Huang, et al., Clinical Features of Patients Infected with 2019 Novel Coronavirus in Wuhan, China, 2020, pp. 497–506. The Lancet 395.10223.
[12] C. Wang, P.W. Horby, F.G. Hayden, G.F. Gao, A novel coronavirus outbreak of global health concern, Lancet 395 (10223) (2020) 470–473.
[13] U. Birajdar, S. Gadhave, S. Chikodikar, S. Dadhich, S. Chiwhane, Detection and classification of diabetic retinopathy using AlexNet architecture of convolutional neural networks, in: Proceeding of International Conference on Computational Science and Applications, Springer, Singapore, 2020, pp. 245–253.
[14] R. Kumar, Adding Binary Search Connections to Improve DenseNet Performance. Available at SSRN 3545071, 2020.
[15] J. Xu, Y. Zhang, D. Miao, Three-way confusion matrix for classification: a measure driven view, Inf. Sci. 507 (2020) 772–794.
[16] F. Wang, H. Zhu, W. Li, K. Li, A hybrid convolution network for serial number recognition on banknotes, Inf. Sci. 512 (2020) 952–963.
[17] J.T. Wight, J.E. Garman, D.R. Hooper, C.T. Robertson, R. Ferber, M.C. Boling, Distance running stride-to-stride variability for sagittal plane joint angles, Sports BioMech. (2020) 1–15.
[18] N.A. Samat, M.N.M. Salleh, H. Ali, The comparison of pooling functions in convolutional neural network for sentiment analysis task, in: International Conference on Soft Computing and Data Mining, Springer, Cham, 2020, January, pp. 202–210.
[19] L. Nanni, A. Lumini, S. Ghidoni, G. Maguolo, Stochastic Activation Function Layers for Convolutional Neural Networks, 2020.
[20] L. Herrmann, C. Schwab, J. Zech, Deep ReLU Neural Network Expression Rates for Data-To-QoI Maps in Bayesian PDE Inversion, 2020.
[21] X. Han, Y. Zhong, L. Cao, L. Zhang, Pre-trained alexnet architecture with pyramid pooling and supervision for high spatial resolution remote sensing image scene classification, Rem. Sens. 9 (8) (2017) 848.
[22] F. Yilmaz, O. Kose, A. Demir, Comparison of two different deep learning architectures on breast cancer, in: 2019 Medical Technologies Congress (TIPTEKNO), IEEE, 2019, October, pp. 1–4.
[23] T. Shanthi, R.S. Sabeenian, R. Anand, Automatic diagnosis of skin diseases using convolution neural network, Microprocess. Microsyst. (2020), 103074.
[24] T. Shanthi, DrR.S. Sabeenian, Modified alexnet architecture for classification of diabetic retinopathy images, Comput. Electr. Eng. 76 (2019) 56–64.
[25] S. Veni, R. Anand, D. Vivek, in: K. Das, J. Bansal, K. Deep, A. Nagar, P. Pathipooranam, R. Naidu (Eds.), Soft Computing for Problem Solving. Advances in Intelligent Systems and Computing, 1057, Springer, Singapore, 2020. https://doi.org/10.1007/978-981-15-0184-5_32.
[26] K. Sanjana, V. Sowmya, E.A. Gopalakrishnan, K.P. Soman, Explainable artificial intelligence for heart rate variability in ECG signal, Healthc. Technol. Lett. 6 (2020 Nov).