Harla J. Eng. Tech. 2022 1(2): 1-15

Harla Journals

Harla Journal of Engineering and Technology

Journal home page: https://journals.ddu.edu.et/index.php/HJET

Harla journal of Engineering and Technology

Developing a Predictive Model for COVID-19 from Chest X-Ray Images Using Deep Learning Techniques

Jarso Gelgelo[1], Jermia Bayisa[2]

[1]Jimma University, Jimma, Ethiopia

[1]Information Science department, Faculty of Computing and Informatics, Jimma University.

[2]Information Science department, Faculty of Computing and Informatics, Jimma University

jarso1gelgelo@gmail.com, jarmbayisa@gmail.com

Abstract

COVID-19 is an outbreak and pandemic disease transmitted through the air and physical contact. This paper aimed to develop an automatic predicting model for COVID-19 from chest X-ray images using deep learning techniques. The techniques that were used in this study were image preprocessing and data augmentation. Two pre-trained Convolutional Neural networks (VGG16 and ResNet50) and CNN proposed model were selected to carry out 2-class prediction tasks using chest X-ray images. 80% of the chest X-ray images were used for training, while twenty percent (20%) were used to evaluate the model. The retrieved features are fitted into a neural network with 500 epochs, an 80/20 splitting ratio, and a learning rate of 0.001. The convolutional neural network model achieved with the ResNet50 of 98.16% average training accuracy compared to VGG16 with 93.65% and the proposed convolutional neural network classifier with 73.85%. The experimental result showed that the overall ResNet50 classifier yielded the highest performance evaluation of 95.4% accuracy compared to VGG16 with 93.08% and the Convolutional Neural network proposed model classifier with 55%. Future research will focus on the issue of the image number; the larger the number of images, the better the model can be trained from scratch.

Keyword: *Chest X-Ray, COVID-19, Deep techniques, pre-trained model*

1. Introduction

COVID-19 is a communicable disease that causes illness in the respiratory system in humans. It is the new virus that is impacting the whole world badly [1]. It is spreading primarily through contact with the person. It was initially discovered in December 2019 in the Chinese city of Wuhan, where it originated from animals and rapidly spread over the world [2]. In March 2020, the World Health Organization (WHO) declared the COVID-19 outbreak a pandemic. It is now a cause of a large number of deaths

---

**Corresponding author: Jarso Gelgelo, jarso1gelgelo@gmail.com,*
DOI:
 Manuscript received August, 2022; revised November, 2022; accepted December, 2022.

across the World [3]. Now, free COVID-19 vaccinations have been developed for the diagnosis of this disease, and clinical trials and research is being carried out for the development of drug and vaccine. It spreads through the air and through physical contact with an infected individual, such as hand-to-hand contact [4]. This virus spreads by respiratory droplets generated from the cough or sneeze of a COVID-19 patient. COVID-19 symptoms can be very mild to severe. Some people have no symptoms are fever, dry cough, and breathing problem. Symptoms such as weariness, sore throat, muscle discomfort, and a loss of taste or smell can also be experienced in coronavirus patients in other ways [5].

Many approaches have been recommended to diagnose COVID-19, from a range of medical imaging modalities, blood tests (CBCs), and Polymerase Chain Reaction (PCR). The WHO claims that, "all diagnoses of Coronavirus disease must be confirmed by Reverse-Transcription Polymerase Chain Reaction (RT-PCR) [4] [6]." However, this method of testing with RT-PCR is highly time-consuming, and this issue leads to the people for risky the person with COVID-19. Hence, first, medical imaging is carried out for the primary classification of COVID-19, then the RT-PCR test is worked to help the medical experts (physician) in making final accurate diagnosis. Chest X-ray (CXR) is one of medical imaging processing, which is used to diagnose COVID-19 [7] [4]. COVID-19 is a class of pneumonia case and attack the lung. In the diagnosis and monitoring of COVID-19 patients, chest X-ray (CXR) imaging is become crucial [7]. Machine learning methods can be used to analyze and classify medical (chest X-ray) images in a variety of medical imaging processes. It is one of the statistical tools to explore the data [8].

Deep Learning is a subclass of Machine Learning that allows computers to tackle problems that would otherwise be impossible to design directly. It enables computer models with numerous processing layers to learn several levels of concept representations of data. Despite the fact that a principled method for training deep networks has been available since the 1980s [9]. It discovers complex structure in large data sets by using the back-propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer [9]. Nowadays, Deep Learning is used everywhere, from spam filters to advertisements recommendation systems, from face recognition to self-driving cars, from predicting stock prices to generating music [10]. Deep Learning, has changed the expectations in many applications of artificial intelligence in data processing by reaching human-level accuracies in many tasks, including medical image analysis [11] [12].

The purpose of this study was to build a model that can predicting whether the medical image processing contains a COVID-19 suggested or non-COVID based training and testing dataset. Working on medical image processing is primarily the main task of dataset collection of COVID-19 because the COVID-19 dataset is scarce which is very much complicated to acquire. Most of the researchers focused on definitive work like filtering, segmentation, feature selection [13]. This study established a model which can accomplish all the fundamental and major necessary tasks to find a COVID-19 from chest X-ray images. The study was proposed a method that helps in the predicting of the COVID-19 without any human assistance, based on deep learning technique.

At present, the statistics of the death rate of COVID-19 disease are increasing daily. Because of the lack of a method to quick detection/classification this infectious disease [14]. All over the world, a 5.06 million of people died of this disease in 2020 [15]. For doctors and health professionals around the globe and local too quick and accurate detection of the virus is a major challenge to reduce the death rate caused by COVID-19 [16]. Because of the inaccessibility of diagnosis systems everywhere, the test of COVID-19 is currently a difficult task that is causing panic fear. Because of the limited availability of COVID-19 testing kits, it have to depend on different determination measures. The medical practitioner (physician, doctor, specialist) are uses frequently chest X-ray images to analyze or diagnoses different diseases and it's possible to use chest X-rays to test for COVID-19 without the dedicated test kits based the feature of COVID-19 finding according to WHO are putted. The limitation of using chest x-ray examination requires a radiology master and takes huge time [17]. Therefore, developing an automatic analysis system is essential to save medical professionals and patients life valuable time.

Deep learning techniques are hidden when extracting the features of an image that are not visible in the actual image. So, convolutional neural networks very helpful to classification and training of features and which had widely implemented by the scientific community [18]. A deep learning classification technique is used to classify chest X-ray images. To improve convolutional neural network accuracy, the significant issue is taking a more important number of images for training the network and increasing the number of iterations [19]. AI (deep learning) today learn complex feature from data and predict the trained model which is used for classification (prediction).

Still as estimated more than 521 million people confirmed cases with COVID-19 and more than 6.2 million people deaths with COVID-19 worldwide [20]. In Ethiopia, this pandemic disease COVID-19 is estimated to be more than 471 thousand confirmed cases in this disease and more than seven (7) thousand is deaths with COVID-19 [21]. This disease is contagious and it can spread through by touching and air from an infected person to person to the normal person. It can be needed to diagnosis at an early stage diagnosis. However, it is more and more of a serious problem, particularly in developing countries, including Ethiopia. The basic difficulties are diagnosis COVID-19 manually screening, lack of accuracy (sensitivity) of availability method and absence of radiologist. Therefore, this study is developing the automatic predicting model for COVID-19 using deep learning techniques. These automatic systems was provided timely assistance to radiologists in diagnosis (inform to suggested disease whether non-COVID or suggested COVID-19).

2. Methods

From the deep learning techniques convolutional neural network architecture were selected because of the convolutional neural network prototype is examined, as it performs well in a small number of images with very low computational resources like CPU and GPU. The model of proposed contains a total of 8 layers within ReLU in a hidden layer with an activation function to provide non-linearity during network training. After the first two fully connected layers, dropout is utilized. It was utilized to reduce the problem of over-fitting. VGG16 is a deep convolutional neural network that contains of sixteen (16) layers and architecture of ResNet is prominent for its use of residual blocks and have forty eight (48) layers. In this paper VGG16 and ResNet50 pretrained model was selected because of the VGG16 is simpler while Resnet50 is complexity.

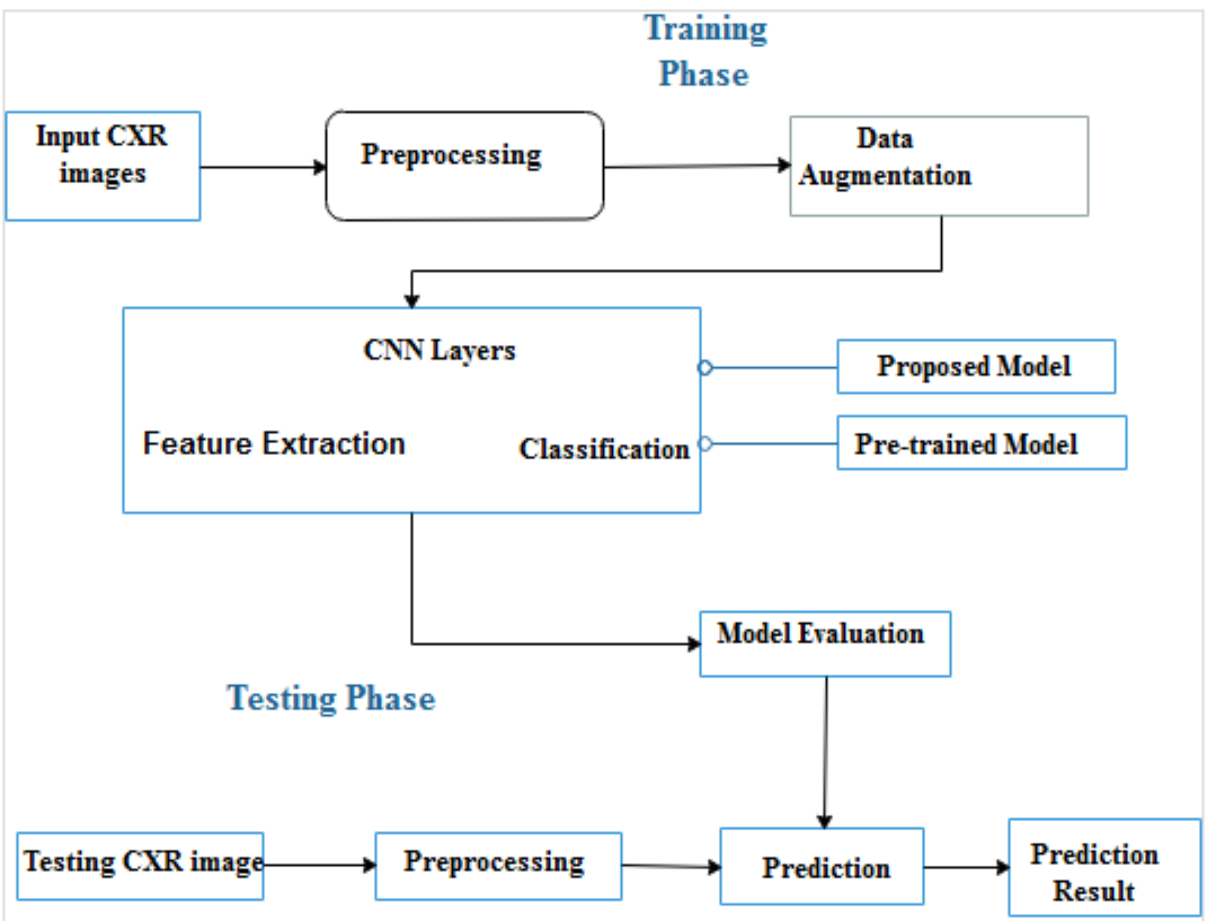

Fig.1: Proposed Block Diagram

Here are two key parts to the Covid-19 predicting process training and testing phase. During the network's training steps, the CNN algorithm accepts size images from input chest X-ray images, by employing data augmentation techniques to train the model, extra training data is produced. Within CNN layers, valuable features extracted and classification from input chest X-ray images was conducted using the extracted features. Then the testing step is performed by giving unseen images during the training to the predictive model.

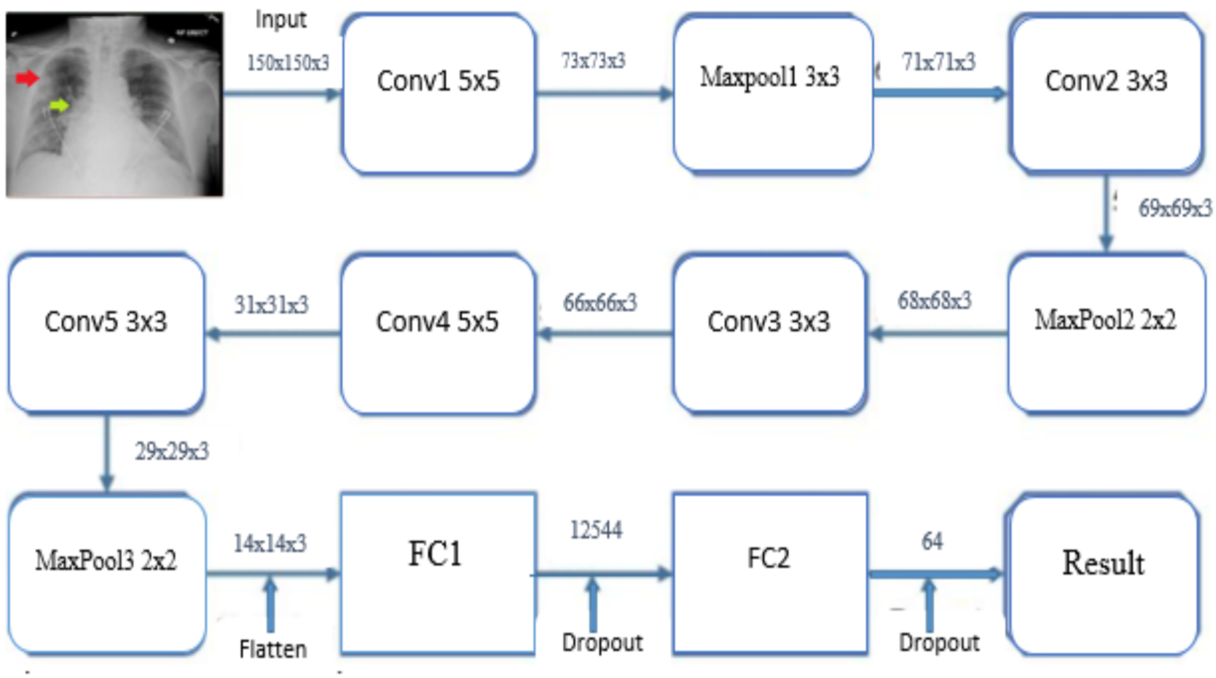

Fig. 2: Proposed forecasting of CNN Architecture

By using input image size (R1), receptive field size (F), stride (S), and amount of zero paddings (P) that can compute the spatial size of the output volume in each layer [22]. The equation that follows is employed.

$$Output\ size = \frac{R1-F+2P}{S} + 1 \quad (1)$$

Where: R_1 is the resolution of the input image, F is filter size, P is the amount of zero paddings, and S is the stride.

A. Input Layer: The Convolutional Neural Network model's input layer takes images with a resolution of 150x150x3 and two separate classes (suggested COVID-19 and non-COVID-19).

B. Convolutional layer: The first convolutional layer taken filters size input images 150x150x3 by using 32 kernels with a size of 5x5x3 with a stride of two pixels. The second convolutional layer takes as input the output of the first convolutional layer and filters it by using 32 kernels of size 3x3x32. The left convolutional layers are connected to each other without intervening pooling layer. With 64 kernels of size 3x3x64 the third convolutional layer taken. The fourth convolutional layer with 64 kernels of size 5x5x64 and the fifth convolutional layer also has 64 kernels of size 3x3x64. ReLU is used activated function to make non-linearity.

C. Pooling layer: was taken 3x3 filter size at the first convolutional layer with stride 1 and taken 2x2 at second output of the convolutional layer of second max-pooling layer with stride 1 and taken 2x2 at third out of convolutional layer with stride 2 three max-pooling layers and hence the number of parameters in these layers is 0.

D. Feature Extraction: The important features that are used to classify the images are extracted by the convolutional neural network algorithm. During training convolutional neural network the network learns the features from the input image. Features are extracted by convolution layers of convolutional neural network. These layers have a series of learnable kernels which aims to extract features from the input image.

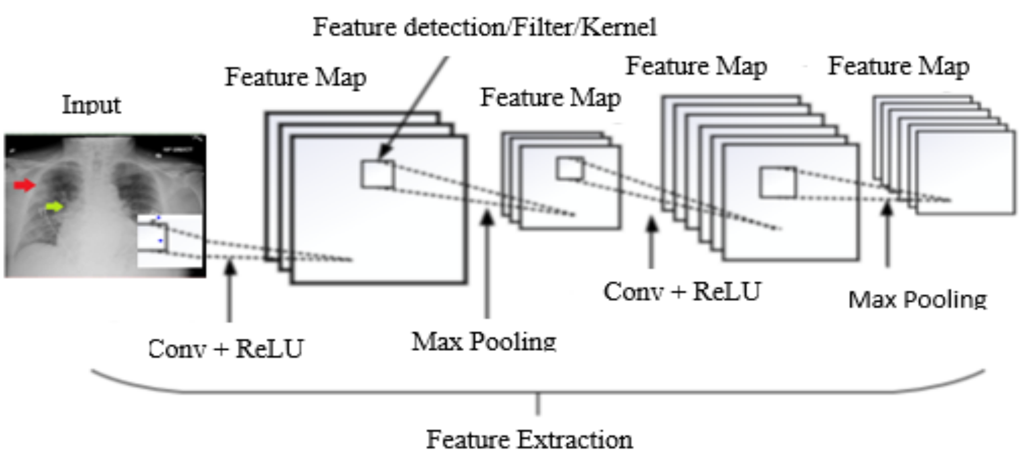

Fig. 3: Feature Extraction

E. Classification of Forecasting Proposed Model: It have a total of three (3) fully connected layers with the output layer. The layers used to categorize the input image based on the features extracted by the convolution layers. Each value of the vector is representing a probability that a certain feature (grayscale color in the dataset) belongs to a class.

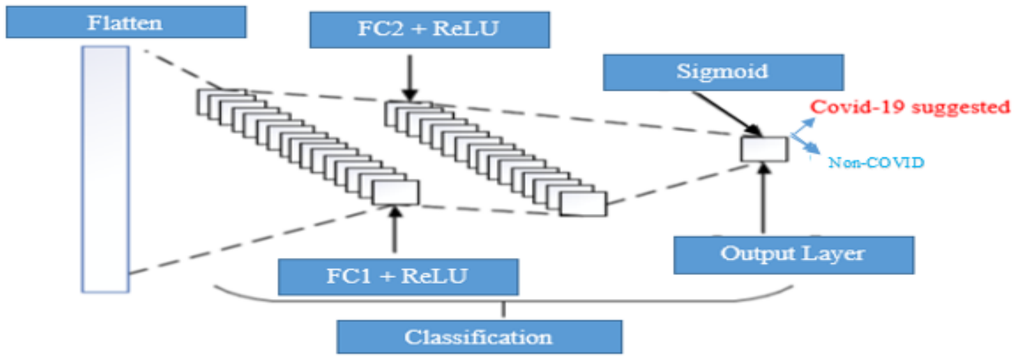

Fig. 4: Classification of the proposed forecasting model

3. Experimental Parameter Setting and Results

A. Image processing

In CXR image analysis, preprocessing was done to enhance and remove noises from the image and to extract essential information for further image analysis. It includes contrast enhancement, edge sharpening, and noise suppression. Used to prepare all needed information for feature extraction.

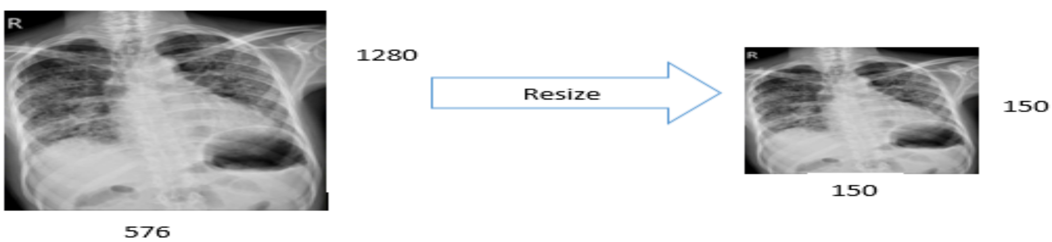

Fig. 5: Resized image

Data augmentation was applied during preprocessing steps. Horizontal and vertical flipping, zooming, shear range and rotation techniques were performed on the collected chest x-ray images of the dataset.

B. Hyper-parameter Configuration

Hyper-parameters are configurations that used before applied the training process of the external to the deep learning algorithm.

Table 1: hyper parameters configuration used for forecasting the proposed model

Optimization algorithms	Learning rate	Loss function	Activation function	Number of epochs	Batch size
Adam	0.001	BCE	sigmoid	10 to 500	32

C. Experimental Results

The proposed Convolutional Neural network model is used to analyze the chest x-ray. There are eight (8) layers in total, five (5) convolutions, and three (3) dense layers model have. An experiments are applied by the proposed Convolutional Neural network model through with changing various ratio of training and testing dataset, through with various learning rates, and through with various activation functions.

Table 2: Experiments of training and testing dataset ratios

Training Split	**Test Split**	**Accuracy**			**Loss**		
		Training	**Validation**	**test**	**Training**	**Validation**	**test**
70%	30%	0.685	0.53	0.747	0.604	3.382	0.22
80%	20%	**0.7317**	**0.8549**	**0.8874**	**0.54**	**0.34**	**0.32**
90%	10%	0.703	0.566	0.836	0.361	2.401	0.162

From above Table 2 shown that, 8:2 ratio training and testing higher than two methods of split ration.

Table 3: Experiments learning rate

Learning Rate	Accuracy			Loss		
	Training	Validation	Test	Training	Validation	Test
0.1	69.6%	51.4%	74.7%	57.7%	49.1%	50.4%
0.01	69.15%	55.42%	74%	55.7%	45.5%	17.2%
0.001	**73.17%**	**85.49%**	**88.79%**	**52.84%**	**34.72%**	**32.54%**

As shown above Table 3, 0.001 learning rate were give training accuracy and less loss accuracy when we compare another two methods. It conclude that 0.001 slowly training the model but it was given good accuracy and optimal.

Table 4: Experimental r*esults of the proposed model by using activation functions*

Activation Function	Accuracy			Loss		
	Training	Validation	Test	Training	Validation	Test
SoftMax	48.82%	50.00%	88.74%	68.70%	66.44%	32.54%
Sigmoid	**73.17%**	**85.49%**	**88.74%**	**52.84%**	**34.72%**	**32.54%**

After experimental, the average training accuracy of the proposed convolutional neural network model were 73.17%, and average training loss which 54.72%.

Experimental 1: Graphical Result of predicting proposed model

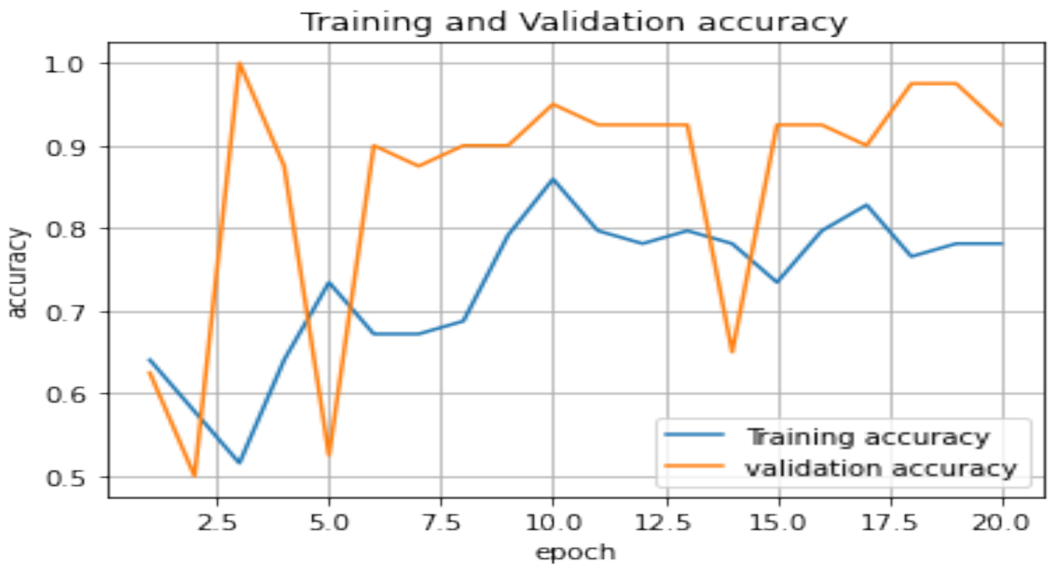

Fig. 6: Training and validation accuracy of the proposed model

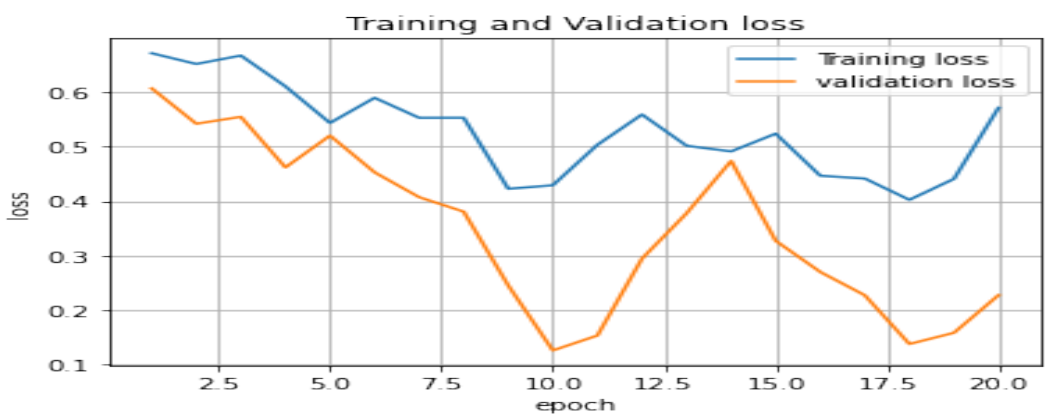

Fig. 7: Training and validation loss of proposed model

As depict above Fig. 5 graph of the training and validation accuracy of the model at the beginning, the accuracy of the training start at 63%, while accuracy of validation 62% and the accuracy of validation increase between 2.5 and 5.0 epochs, also higher after epochs 5.0, and accuracy of training was increase at start epochs 2.8, the average accuracy training and validation was 73.17% and 85.49% respectively.

The average loss of training, validation, and test as depict above fig.6 graph were 32.54%, 52.84%, and 34.72% respectively. It concluded that there is much gap between the validation and training as shown fig.5 and 6 which shown problem of the overfitting.

Experimental 2: Graphical Results of VGG16

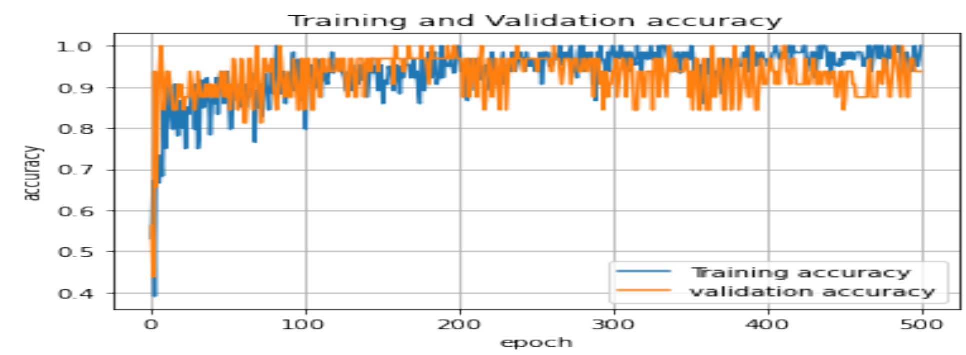

Fig. 8: Training and validation accuracy for VGG16

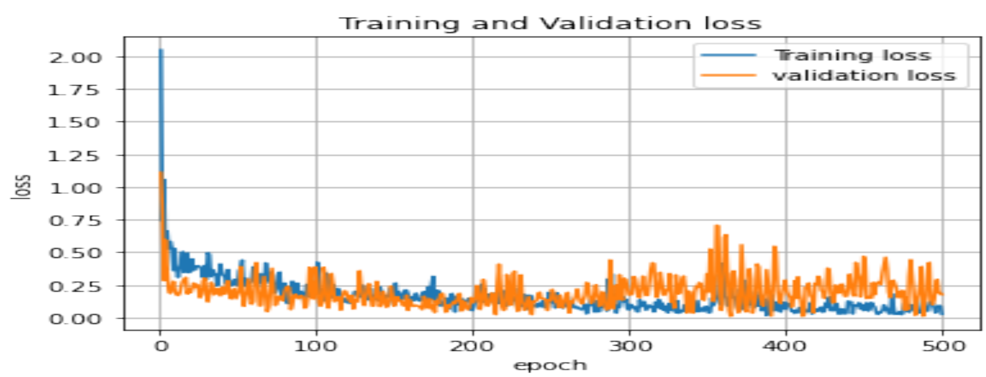

Fig. 9: Training and validation loss for the VGG16

As shown above fig.8 and 9 graph result of VGG16, the model was trained with 500 epochs which reached highest average accuracy, validation and test were 93.65%, 92.69%, and 98.4% respectively. The Average loss 16.06% training, 19.56% validation, and 5.17% test and it shown in **table 5** below.

Table 5: The Average Accuracy and loss of VGG16

Metrics	Average Accuracy			average Loss		
	Training	Validation	Test	Training	Validation	Test
Value	93.65%	92.69%	98.40%	16.06%	19.56%	5.17%

Experimental 3: Graphical Result of ResNet50

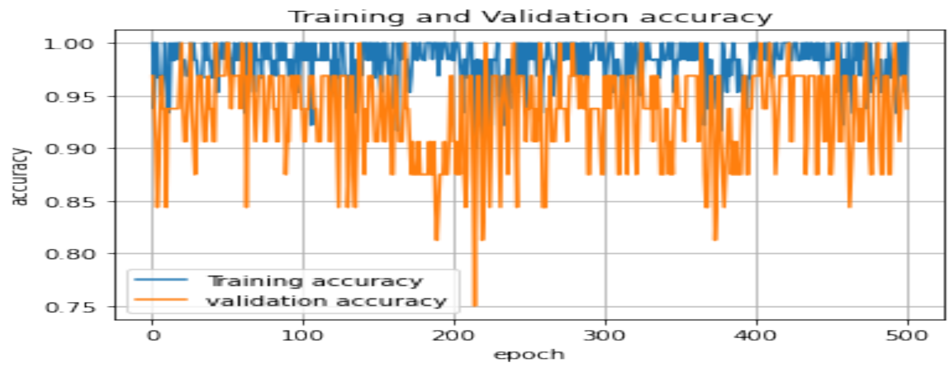

Fig. 10: Training and validation accuracy of ResNet50 model

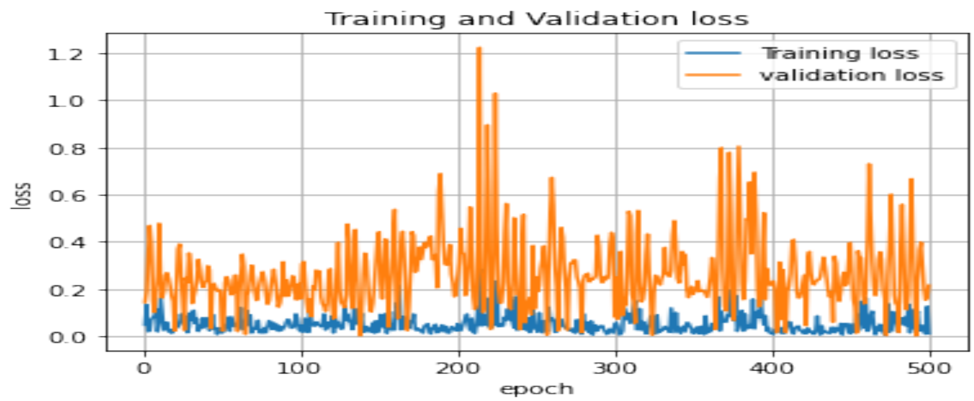

Fig. 11: Training and validation loss of ResNet50

As shown above fig.10 and 11 graph result of ResNet50, the model was trained with 500 epochs which reached highest average accuracy, validation and test were 98.16%, 93.30%, and 99.46% respectively. The Average loss training, validation, and test were 5.1% training, 26.18%, and 2.2% and it shown in **table 6** below.

Table 6: Average accuracy and loss of ResNet50

Metrics	average accuracy			average loss		
	Training	Validation	Test	Training	Validation	Test
Value	98.16%	93.30%	99.46%	5.1%	26.18%	2.2%

General, several studies were done which is related with this paper, Hamdan *et al.* [25] in (2020) conducted research on Convolutional Neural Network with different classifiers and COVIDX-Net and achieved a less interesting performance in accuracy (90.0%) in comparing with this study with ResNet50 with 95.4%. The result difference could be due to the increment in hyper parameter and features that fully express the real working environment and the current work values difference. Since a more number of images features can increase the system performance. The current work conducted by ResNet50 accuracy performance measurement registered 95.4% respectively which are less specified by the former work.

The study by Hamdan *et al*. [25] COVIDX-Net, a comparative comparison of seven deep-learning models for the diagnosis of COVID-19. VGG19, DenseNet201, ResNetV2, InceptionV3, ResNetV2, Xception, and MobileNetV2 were tested with a binary data set of 50 X-ray samples (25 healthy and 25 COVID-19). With VGG19 and DenseNet201, the study attained the greatest accuracy of 90%. The study does, however, suffer from the drawback of a small data set.

For the identification of COVID-19, Sethy and Behra [26] developed a deep learning-based technique, ResNet50 with support vector machine (SVM). For deep features, they used a pre-trained Convolutional Neural Network. After that, the collected deep features were used to classify the data using a support vector machine (SVM) classifier. In comparison to this study, the accuracy gained from ResNet50 with SVM was 95.38 percent, but the overall accuracy of ResNet50 was 95.4 percent.

Narin *et al*. [27] examined the results of three deep networks, ResNet-50, Inception-V3, and Inception-V2, in classifying COVID-19 in chest X-ray images. his study also compared this study of three Convolutional Neural network algorithms, namely proposed model, ResNet50, and VGG16, and after training all networks, ResNet50 achieved the highest evaluation with an accuracy of over 95.4 percent, while this study also compared this study of three Convolutional Neural network algorithms, namely proposed model, ResNet50, and VGG16, and after training all networks, ResNet50 achieved the highest evaluation with an accuracy of over 95.4 percent, which is less than Narin *et al* [27] work, the difference is parameter of hyper setting, area, machine they used and the Convolutional Neural network architectures types and size.

However, challenges are still experience related to identify COVID-19 disease is dataset of the images because of deep learning has used large number of data and failed to learn more complex features in small dataset. Developing model for the COVID-19 predicting is a strategy, method, or system for running or regulating a COVID-19 new outbreak disease predicting process using highly automated means, such as electronic equipment, with the goal of minimizing human interaction. Developing model for COVID-19 disease diagnosis tools provide a number of advantages in the fight against this new epidemic illness. Measurements and features enable radiologists and doctors make a diagnosis, and it aids radiologists and doctors in their diagnosis method for accuracy and efficiency. Because computer-assisted diagnosis is often faster than manual screening, and because it can assist them in making decisions. Furthermore, the manual screening procedure requires less human intervention.

4. Conclusion

The proposed CNN model has attained an average training accuracy of 73% with 20 epochs, 0.001 learning rate, and sigmoid activation function using a chest x-ray image. This outcome is modified when the model is trained using pre-trained models such as VGG16 and ResNet50, which climbed to an average training accuracy of 93.63% and 98.16% with 500 epochs with 0.001 learning rate, respectively. Furthermore, overall performance evaluation outcomes show that ResNet50 provides an accuracy of 95.4%; this architecture outperforms the other architectures chosen in this study. The Convolutional Neural Network model employing ResNet50 and VGG16 was archived with the maximum accuracy with 500 epochs within a small dataset for this study.

References

[1] F. Saiz and . I. Barandiaran, "COVID-19 Detection in Chest X-ray Images using a Deep Learning Approach," *DOI: 10.9781/ijimai.2020.04.003,* 2020.

[2] W. J. Chan, S. Yuan and K. H. Kok, "A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating," *The Lancet,* vol. 395. no.10223, 2020.

[3] A. Afifi, N. Hafsa, A. Mona , A. Alhumam and S. Alsalman, "An Ensemble of Global and Local-Attention Based Convolutional Neural Networks for COVID-19 Diagnosis on Chest X-ray Images," *Journal of Symmetry,* 2021.

[4] A. Shoeibi and M. Khodatars, "Automated Detection and Forecasting of COVID-19 using Deep Learning Techniques: A Review," 2020.

[5] World Health Organization, "Use of chest imaging in COVID-19," 2020.

[6] C. Sohrabi, . Z. Alsafi, N. ONeill, M. Khan and A. Kerw, "World health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19)," *International Journal of Surgery,* 2020.

[7] A. E. Hassanien, L. N. Mahdy and K. A. Ezzat, "Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine," *medRxiv,* 2020.

[8] E. Goldstein and D. Keidar, "COVID-19 Classification of X-ray Images Using Deep Neural Networks: doi: https://doi.org/10.1101/2020.10.01.20204073," pp. 2-27, 2020.

[9] Y. LeCun and Y. Bengio, "Deep learning," *doi:10.1038/nature14539,* 2015.

[10] D. M. L. Nguyen, "Detect COVID-19 images using Deep Learning," 2020.

[11] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. . M. Swetter, H. M. Blau and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature,* vol. 542, 2017.

[12] M. A. Mazurowski, M. Buda, A. Saha and M. R. Bashir, "Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on mri," *Journal of magnetic resonance imaging,* Vols. vol. 49, no. 4, 2019.

[13] L. Orioli, M. P. Hermans, J. P. Thissen, D. Maiter, B. Vandeleene and C. J. Yombi, "Covid-19 in diabetic patients: related risks and specifics of management," *in Annales D'endocrinologie,* 2020.

[14] A. Remuzzi and G. Remuzzi, "Covid-19 and italy: what next?," *The Lancet,* 2020.

[15] "World Health Organization:The true death toll of Covid-19," 2020. [Online]. Available: https://www.who.int/data/stories/the-true-death-toll-of-covid-19-estimating-global. [Accessed Monday Octomber 2021].

[16] N. A. Alam and M. Ahsan, "COVID-19 Detection from Chest X-ray Images Using Feature Fusion and Deep Learning," *https://www.mdpi.com/journal/sensors,* pp.2-26, 2021.

[17] P. K. Sethy and S. K. Behera, "Detection of coronavirus Disease (COVID-19) based on Deep Features," *doi:10.20944/preprints202003.0300.*pp.1-9, 2020.

[18] K. . K. Lella and A. PJA, "Automatic COVID-19 disease diagnosis using 1D convolutional neural network and augmentation with human respiratory sound based on parameters: cough, breath, and voice: This document is trying to connect to:," *AIMS Public Health,* vol. 8, Issue 2, pp. 2-25, 2021.

[19] S. Asif, Y. Wenhui, H. Jin, Y. Tao and S. Jinha, "Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Networks," *doi: https://doi.org/10.1101/2020.05.01.20088211,* pp.1-7, 2020.

[20] "Worldometer," [Online]. Available: https://www.worldometers.info/coronavirus/countries-where-has-spread/. [Accessed 8 May 2021].

[21] "Ethiopia Covid-19 case," [Online]. Available: https://www.worldometers.info/coronavirus/country/ethiopia/. [Accessed 8 May 2021].

[22] F. Li, J. Justin and Y. Serena, " Convolutional Neural Networks for Visual Recognition," Stanford University, Spring 2018. [Online]. Available: Available: http://cs231n.stanford.edu/index.html. [Accessed 2 August 2021].

[23] A. Krizhevsky, I. Sutskever and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems,* 2012.

[24] P. Josh and G. Adam, "Deep Learning A Practitioner's Approach, Sebastopol: O'Reilly Media," 2017.

[25] E. E-D. Hemdan, M. A. Shouman and M. E. Karar, "COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images," *Institute of Electrical and Electronics Engineers (IEEE),* 2020.

[26] P. K. Sethy and S. K. Behera, "Detection of Coronavirus Disease (COVID-19) Based on Deep Features," 2020 .

[27] A. Narin, C. Kaya and Z. Pamuk, "Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks," 2021.