GUB Journal of Science and Engineering (GUBJSE)
p-ISSN: 2409-0476
vol. 10, no. 01, 2023

COVID-19 Identification System from X-Ray Images of Chest using Deep Neural Network with Transfer Learning

Farhana Akter Sunny[1], Sagufta Sabah Nakshi[1], Mashud Parbhez[1], Muhammad Aminur Rahaman[1]

1 Department of Computer Science Engineering Department, Green University of Bangladesh, Dhaka, Bangladesh.

*Corresponding author's email: farhana@cse.green.edu.bd

Received: 16/03/2023
Accepted: 01/12/2023
Published: 15/07/2024

Data Availability: The data are available on request from the corresponding author.

Competing Interests: 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.

DOI: 10.3329/gubjse.v10i1.74945

Abstract

Recently, the impact of COVID-19 has significantly diminished; however, it has not been completely eradicated. There are still instances where individuals are experiencing suffering due to this life-threatening virus which has a significant impact on health care as well as lifestyles throughout the world. So, early discovery is important to controlling case extension and the death rate. The RT-PCR is known as the true leading diagnosis test; nevertheless, the expense and result times of these tests are long, thus additional quick and accessible diagnostic techniques are required. However, most countries are suffering due to limited testing resources and kits. The unavailability of testing resources, kits, and a rising amount of regular occurrences, caused us to develop a model on Deep Learning which may benefit radiologists as well as doctors for detecting COVID-19 instances using images of chest X-rays. For developing a representation of modality-specific features, a convolutions neural network and a variety of ImageNet pre-trained models are trained and evaluated at the patient level by using different available CXR datasets. We choose 5000 images in total from the dataset collected from Kaggle where we kept 4000 images in case of training and validation, and the remaining 1000 in case of testing. We use four Pre-train Deep CNN Models which are very popular for image calcification. VVG16, VGG19, InceptionV3, and Resnet50 CNN Models we choose to analyze the performance and find the best one among them. In our testing, we get 88.5% testing accuracy on ResNet and 95.10% on InceptionV3 models while VGG19 gives 90.22% accuracy and VGG16 gives the highest 96.10% accuracy. To increase performance accuracy, Transfer Learning knowledge is transmitted and fine-tuned. After applying Transfer Learning in the modified VGG16 we got an accuracy of 97% which is clearly an improvement over the previous VGG16 model.

Keywords: Deep Neural Network, COVID-19 detection, machine learning, transfer learning, Image Dataset.

Highlights

- Utilizing Deep Neural Networks for accurate COVID-19 detection.
- Achieving high testing accuracies up to 97% through Transfer Learning.
- Training and evaluating popular CNN models like VGG16, InceptionV3, and ResNet50.

1 Introduction

According to the World Health Organization, 2020, the disease COVID-19 is brought on by the SARS-CoV-2 coronavirus [1] and according to estimates from the World Health Organization, the disease had spread to almost every country since the first verified case by June of 2021, killing more than four million individuals and causing almost 180 million of the confirmed cases. [2]. The baseline of treatment for COVID-19 involves screening individuals in general medical clinics or hospitals. Hospitals now rely on diagnostic imaging in the case of patients with severe respiratory symptoms because it is simple and quick and enables doctors to detect infections and accompanying symptoms more quickly, even though PCR (polymerase chain reaction) testing for transcription is still employed for the definitive diagnosis. [3].
The first line of prevention is screening people in basic health clinics or hospitals. Patients who are thought to have COVID-19 initially go through an X-ray procedure, and if more information is needed, a CT scan is done [4]. Due to innovation, X-ray images are now generally utilized in hospitals as alternate diagnostic tools to identify COVID-19 and determine the consequences of the virus [5][6]. Doctors use X-rays of the patient's lungs to diagnose patients and search for COVID-19 distortion signs. Due to the rapid danger of transmission, there have been many patients in clinics instantly, making a burden on the diagnosticians. This issue can be fixed by using methods of deep learning, that have advanced impact in recent years with the high speed and accuracy and every amount of information available, and process improvement of deep learning algorithms and models, as evidenced by career-high results in struggle contests [7]. The purpose of deep learning is to develop the multi-hidden level ML technique that is trained with a huge amount of data collected and then used to improve classifying and accuracy rate through understanding more accurate features [8], [9].
Our major contributions to this research are-

1. to add dropout filters to the VGG-16 neural network to decrease overfitting and generate a more generic outcome with lower computing costs.

2. to employ transfer learning in the modified VGG16 model which identifies Covid-19 more accurately in this platform after compared with other CNN models.

The rest of the paper is arranged in the following order- **Section 2** addresses background studies, literature evaluations for this study, and the limitations of previous research. **Section 3** describes the concept and methods for detecting and recognizing the COVID-19 Identification Models. **Section 4** discusses the assessment of the developed system as well as the analysis of the results using performance measurement. **Section 5** concluded the paper by discussing the benefits, limits, and future directions.

2 Literature Review

We investigated and identified various study works where in-depth photographs had been employed together with diverse techniques for different research, maintaining the exceptional training loss in attention. In [10] Khan et al. proposed that CoroNet is a Deep CNN model which can identify COVID-19 through the X-ray pictures of the patient's chest. Their proposed model is developed on the architecture of Xception and the architecture was pre-trained on a dataset named ImageNet and was trained on a dataset created by collecting pneumonia and COVID-19 X-ray(chest) pictures from the two available public resources. Fang et al.[11] suggest research to investigate the sensitivity of CT images as well as the identification techniques of nucleic acid, with the help of the real-time polymerase chain reaction (RT-PCR) for identifying COVID-19. the study includes 51 participants who have traveled recently or are residents in Wuhan, and who have a fever or severe respiratory symptoms, and also an overall 45 years old people. Their conclusion is that chest CT contains a higher sensitivity in comparison with RT-PCR (98% vs. 71%). V. Rajinikanth et al [12] designed and executed a technique for detecting CTSI pneumonia infection. The proposed technique's many steps are depicted in the 2D CTSI, which is derived from the benchmark database. Later, to improve the visibility of the affected lung region, a multi-threshold based on the Otsu function as well as Harmony-Search-Optimization (HSO) is introduced. Watershed segmentation is then used to retrieve the upgraded portion. Finally, the severity rate is calculated using the binary pictures' pixel values of the infected regions.
To correctly categorize CXRs as COVID-19, normal, or bacterial pneumonia, authors used deep learning ensembles that were repeatedly pruned [13]. A number of models were explored to improve classification accuracy, and the best outcomes were combined using a variety of ensemble techniques. However, because the computational overhead of numerous model computations is considerable, this kind of approaches are best suited for short COVID-19 image datasets, and no guarantee that they will keep their accuracy with huge datasets [14], [15]. Rahimzadeh and Attar[16] proposed a training strategy for coping with imbalanced datasets and introduce a deep convolution network that is based on the fusion of ReNet50V2 and Xception to increase accuracy. The suggested network's average accuracy to detect COVID-19 instances is 99.50%,

while the average accuracy for all of the classes is 91.4%.

Maghdid et al.[17] offered Artificial Intelligence techniques which can assist radiologists in promptly diagnosing COVID-19 patients using X-ray as well as CT scan pictures. AlexNet (modified pre-trained) was their network's backbone. In the binary classification system, an upgraded CNN accomplishes accuracy rate of 94.1%, whereas a pre-trained model gets an accuracy rate of 98%. In [18] the authors used a cutting-edge CNN model dubbed MobileNet to classify 3905 X-Ray pictures into six distinct groups. In addition, 455 COVID-19 CXR pictures are added. They obtained 99.18% accuracy in case of binary classification tests and 87.66% accuracy in case of the tasks of seven class classification. In [19], Loey et al. proposed a Generative Adversarial Network or GAN along with deep TL (transfer learning) for the purpose of detecting COVID-19 with tyhe help of CXR images. They use GAN to create additional pictures because of the limited CXR images. Approximately, 307 pictures were gathered in four different categories, including normal, COVID-19, viral pneumonia, and bacterial pneumonia. For training purposes as well as the validation stage with GAN they have used 90% of their dataset, while 10% is maintained in the case of testing. As transfer learning, pre-trained AlexNet, GoogleNet, and Resnet18 networks are employed. On a two-class (GoogleNet), three-class (AlexNet), as well as four-class (GoogleNet) classification issue, the proposed model achieved an accuracy of 100%, 85.3%, and 80.6%, respectively. Zhang et al.[20] offered a classification technique based on the deep TL method to classify cervical cells as normal or pathological. To test the overall performance, a dataset named HEMLBC (private) and a dataset named Herlev (public) are used. On the Herlev and HEMLBC datasets, they obtain an accuracy of 98.3% and 98.6%, respectively. According to Abbas et al., [21], the previously built Transfer, CNN, Decompose, and Compose networks (DeTraC-Net) use ResNet-50 which is a pre-trained model, as TL for categorizing COVID-19 chest X-ray data from normal as well as severe acute respiratory syndrome cases. The accuracy of their suggested approach is 95.12%.

For detecting COVID-19 more accurately another researh shows that VGG16 provides 80% accuracy among VGG16, VGG19, ResNet, DenseNet, and InceptionV3 in [22]. Authors claimed that performance has been improved by image resizing, augmentation methods. To detect COVID-19 Duong et al. proposed a deep neural network approach using EfficientNet and MixNet [23]. To increase the accuracy authors utilized ImageNet, AdvProp, and NS learning techniques and they got 96% accuracy in EfficientNet-B3. Although their accuracy was very satisfying but authors didn't propose any approach for reducing over-fitting as well as computation cost. Authors have collect approximate 3800 images which was insufficient [24][25]. There are eight classifiers were used along with deep transfer learning. Authors modified VGG16 with 18 frozen layers and found an outstanding accuracy. With a low dataset that results over-fitting and counteracts generalization.
In some research, datasets were collected from the patient, and social media was not ample for computer vision. The authors did not remove other artifacts, such as text and medical device traces on chest X-rays. Some researchers used very few datasets and some used low-resolution images. There is no assurance that they will maintain their accuracy with huge datasets. Some research improved the accuracy but in terms of the robustness, the performance was very poor. At the same their compuatational cost was high.

3 Methodology

We used CXR images from normal patients and patients with COVID-19 to evaluate the efficiency of our suggested technique. We trained the different CNN models, which were pre-trained on the ImageNet database, and that ImageNet Database was pre-trained with many classes using transfer learning (TL) as well as fine-tuning. TL shortens the training period and reduces generalization errors for a neural network model. For assessment, it utilizes the weight with the highest validation accuracy. These settings have been fine-tuned using hyper-tuning.

3.1 Proposed Methodology

We have presented a deep learning-based COVID-19 identification method, in which the program employs an algorithm of deep learning to determine if the images of the patient's lung are COVID-19 or normal. Figure 1 shows the overall architecture of our proposed methodology. Firstly, from the Dataset[26], we have used a total of 5000 chest X-ray images, where 2500 are COVID-19 and 2500 are Normal images. The data collection has been split into three sections. The training set is utilized for model training and learning, as well as parameter optimization. The validation set is used to evaluate the model during training, optimize it, and fine-tune its parameters. Our model's final findings are validated using test sets. The CNN model was performed using Python in the Keras package, which was running on TensorFlow. We began by downloading the data from Kaggle. Then, in our project's home directory, we created a directory called "Dataset." We established two folders in the dataset directory, "Covid" and "Normal" each with associated pictures. After that, we

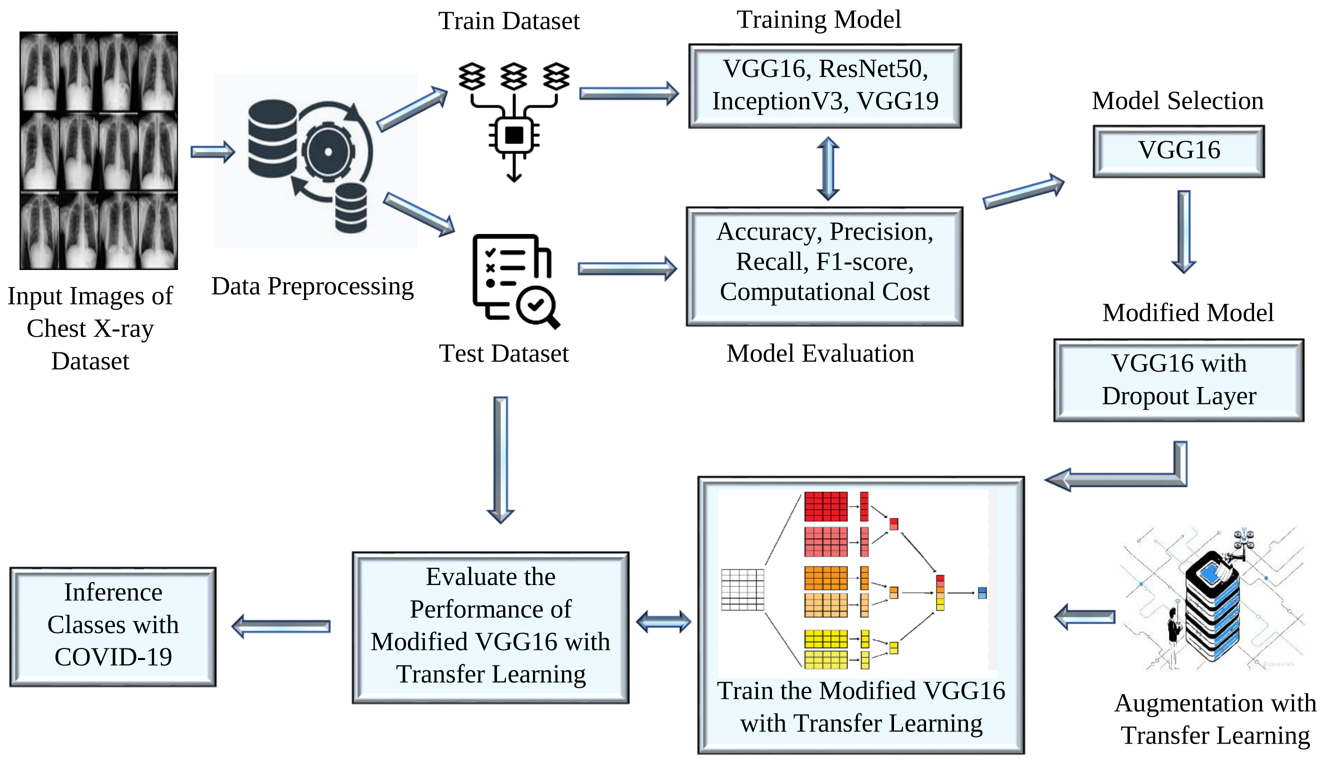


Fig 1. Proposed screening structure of COVID-19 detection based-on Deep learning

copied all of the pictures to the dataset directory and utilized them for running the models. Because all of the images are PNG files, the first thing we do is decode them to RGB grid pixels using a tensorflow keras ImageDataGenerator instance. For training, we utilize four models: VGG16, VGG19, Inception V3, and ResNet50. After data pre-processing, to create training and testing sets, we divided the dataset. Then train the models with a training set and evaluate or measure the performance based on different parameters. For COVID-19 X-ray image binary classification, we used four pre-trained CNN models. After training, testing, and validating all the models we will choose the best model VGG16. After that, we modified the VGG16 model by adding a dropout layer to reduce the overfitting which is shown in Figure 2. After that, we trained the modified VGG16 with transfer learning to enhance the accuracy of the result. The updated VGG16 deep transfer learning models are used to perform binary and multi-class classification using COVID-19 X-ray data. The model performance exhibits impressive outcomes and is straightforward to deploy.

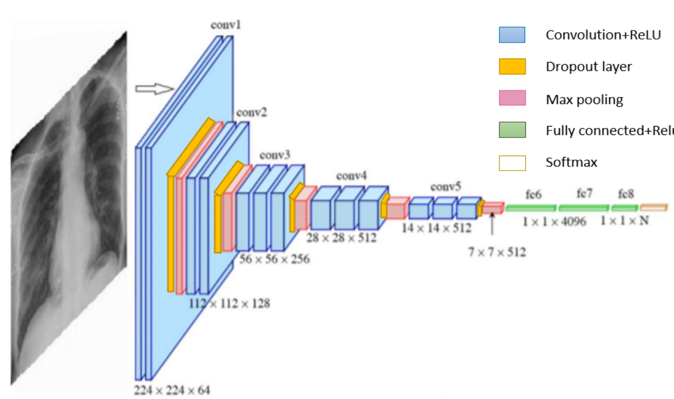


Fig 2. The Structure of Modified VGG16 model

3.2 CNN Models

COVID-19 has been effectively diagnosed using several deep learning networks [27][28]. CNN is the most widely used technique for COVID-19 illness classification, segmentation, and prediction. We have trained the following four models.

1. VGG16
2. ResNet50
3. VGG19 and
4. Inceptionv3.

3.2.1 VGG16

The VGG16 network has 16 layers, with convolutional layers (13) with 3×3 filters and 2×2 max-pooling layers layered on top. The ReLU activation function is employed between these levels. The majority of the network's parameters are then stored in three completely connected layers.

3.2.2 ResNet50

ResNet50 [29] has developed as a revolutionary deep neural network (DNN) model addressing computer vision issues has emerged. Identity shortcut connections — connections that skip one or more layers – have been used to alleviate this issue in ResNet50 models with up to 150 layers. This creates a detour for slopes to travel over while remaining unchanged. We employed the 50-layer ResNet-50 as the model's fundamental architecture.

3.2.3 Inceptionv3

Inception-v3 [30], a deep pre-trained CNN with 48 layers, is a network version that was trained using a million or more images from the ImageNet dataset. To categorize photos into 1000 different item categories, this network has previously received training.

3.2.4 VGG19

VGG-19 is a variation of VGG architectures that has 19 linked layers and routinely outperforms other models. Due to the model's highly linked convolutional and fully connected layers, it is possible to extract features more effectively and to downsample data before classification using the SoftMax activation function by using Max pooling rather than average pooling.

4 Result Analysis

4.1 Data Pre-processing

The "preprocess input" function in Keras alters the image representation to meet the model's requirements. The model's performance is then enhanced by using a variety of data augmentation methods on the training sets. For this, Keras "ImageDataGenerator" API is used.
The photos are periodically altered and delivered to the system during training time in this experiment, using the "rotation range" and "fill mode" data augmentation techniques. The photos are rescaled initially in this procedure. The rotation range was then set to 15 degrees. After that, we adjusted the sheer scale to 10% and the width shift ranges and height shift ranges to 10%. Afterward when we set the zooming level to 20% and ensured that the horizontal flips and vertical flips were true. We also integrate each pre-trained model's pre-process inputs.

4.2 Data Training and Testing

The experiments using our selected CNN models were conducted using benchmark dataset [26]. They are scanning chest X-ray pictures with COVID-19 and Normal classifications, etc. From the Dataset, we have used a total of 5000 chest X-ray images, where 2500 are COVID-19 and 2500 are Normal images. The information gathered is separated into three sections. The model is trained and learned from, and the characteristics are changed, using the training set.
Following the model's training, the validation set is utilized to verify, improve, and fine-tune its parameters. Our model's concluding results were verified using test sets. First, we randomly selected 1000 (20%) from each group's photos for testing, with the remaining 4000 (80%) divided into training and validation. These photos are pre-processed before being used to train CNN models.
We trained our proposed models using Adam, an SGD optimizer, with a preliminary rate of learning of 1e-3 to 1e-6. Here, at first we have considered the Adam optimizer for all four models. But later it was observed that InceptionV3 and ResNet50 work better with Softmax and SGD optimizers accordingly than the Adam Optimizer. Hence, table 1 shows the optimizes with whom the models perform better We choose to use 32, 64, and 128 as our batch sizes. If the validation accuracy does not increase after every five iterations, the reduced learning rate is a factor of 2. The training is terminated early if accuracy doesn't rise for 20 subsequent epochs. The network is trained in this manner for 100 epochs.
For evaluation, we take the weight with the highest validation accuracy. Hyper-tuning was used to correct these

Table 1. Train, Testing and Validate using different learning Values:

CNN Model	Input Size	Batch Size	Optimizer	NO of Epochs
VGG16	224x224x3	32	Adam	100
Inception V3	299x299x3	32	SoftMax	100
ResNet50	224x224x3	32	SGD	100
VGG19	224x224x3	32	Adam	100

values.

4.3 Performance Analysis

We trained the dataset many times using different CNN Models and settings to get the best accuracy while minimizing loss. The more we train our algorithm and provide more datasets, the better the outcomes. We worked with a good number of CXR pictures in this project. We attempted to achieve the best outcomes and production possible. We also complete our overall system performance study. Using the pooled datasets, we assessed the accuracy, computational cost, and loss and obtained excellent results.

4.3.1 Performance evaluation metrics

To evaluate COVID patient categorization, We apply the models discussed in the previous section to the X-ray image datasets, modifying the models during the training phase to improve their accuracy. We describe three outcomes that are characteristic of CNNs for each model:
Model accuracy curve: The accuracy curves shows for the training and validation models that how well the model generalizes or trains.
Model loss curve: Overfitting is shown by the difference between training and validation accuracy. The network's learning direction and training process is shown by the model loss curve. The network may still be able to learn more with training, according to a considerable discrepancy between the training and validation curves.
Confusion matrix: a table that describes the classifier's performance in a set of test data after knowing the true values is called a confusion matrix. Every confusion matrix is connected with four basic elements [31].

1. True Positives [TP]: These are situations in which we forecasted "yes" and the patients had the illness.

2. True Negatives [TN]: We forecasted "no," and they are not contaminated.

3. False Positives [FP]: We predict a "yes" for the condition, but the patients do not have the ailment. This mistake is occasionally referred to as a Type I error.

4. False Negatives [FN]: Even though the proposed model predicts "no," patients have the illness. This is known as a Type II mistake. It shows crucial prediction data, making it easier to interpret and get meaningful experiment patterns.

To appreciate the performance of our model here we have used some parameters such as Accuracy (ACC) **[Eq-1]**, Precision (P) **[Eq-2]**, Recall (R) **[Eq-3]**, F1-score (F1) **[Eq-4]**, Sensitivity (SN) **[Eq-5]** and Specificity (SP) **[Eq-6]**.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$F1 - score = 2x\frac{PxR}{P + R} \tag{4}$$

$$Sensitivity = \frac{TP}{Total} \tag{5}$$

$$Specificity = \frac{TP}{Total} \tag{6}$$

4.3.2 Performance measurements

We set up a comparable setting to measure the performance of each model based on the above measurement criteria. The same dataset with 2500 Covid and 2500 Non-Covid pictures was used, with 64 batch sizes and an initial learning rate of 1e-3. A total of 100 training epochs were used for each model. We compared all of our training models and the measurement results of each CNN model are shown in the figures 3 to 10.

Measurements for Resnet50 Model Figure 3 shows that we have trained the ResNet50 Model with 100 Epoch. The accuracy gets higher respectively with increasing Epochs. we got an overall validation accuracy of 87.5% in this model. In Figure 4, A total of 100 training epochs were used for each model. The networks accurately detect 432 duplicates in the COVID class from the Covid Test data set, but only 68 copies in the Normal class, according to the confusion matrix of the least effective techniques, like ResNet50. The model also correctly identifies 452 samples as normal, but just 48 as COVID from the normal test data set.

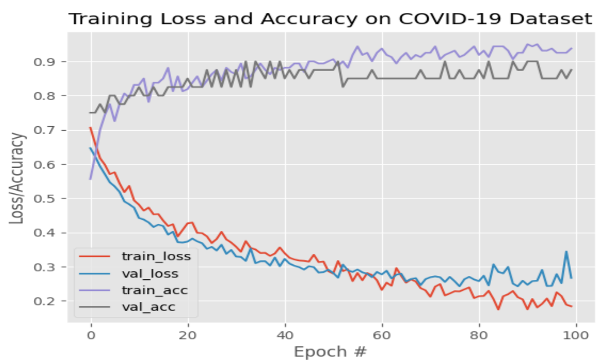
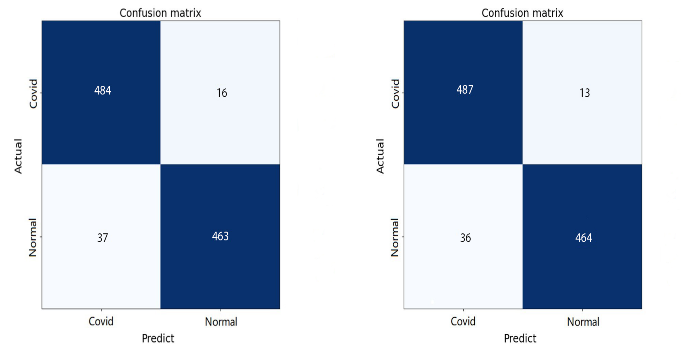
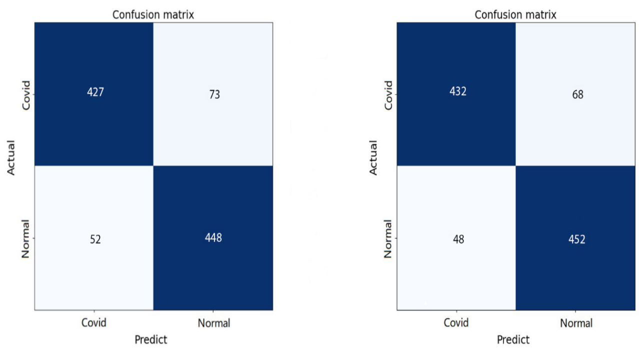


Fig 3. Covid or Normal: Accuracy and Loss of ResNet50 model for on X-ray Dataset



(a) Validation (b) Test

Fig 6. Confusion matrix for InceptionV3 on X-ray Image dataset: COVID-19 or Normal



(a) Validation (b) Test

Fig 4. COVID-19 or Normal in the confusion matrix for ResNet50 on the X-ray dataset

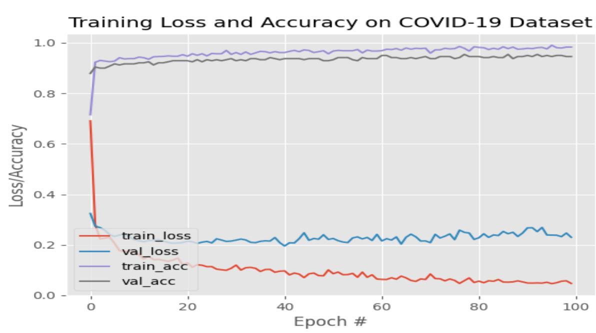


Fig 5. COVID-19 or Normal: Accuracy and Loss of InceptionV3 model on X-ray Image Dataset

Measurements for InceptionV3 Model Figure 5 shows that after training the Covid-19 Model with 100 Epoch we got an accuracy of 94.58%. The accuracy gets higher respectively with increasing Epochs. Also, the loss is reduced very strongly. Figure 6 shows In the InceptionV3 model, From the COVID Test data set, 487 pictures are properly classified as COVID, while 13 images are classified as Normal. From the Normal data set, the model also correctly identifies 468 samples as normal, but only 36 as COVID.

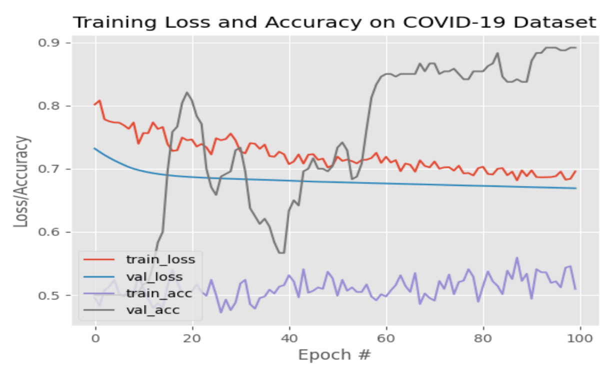
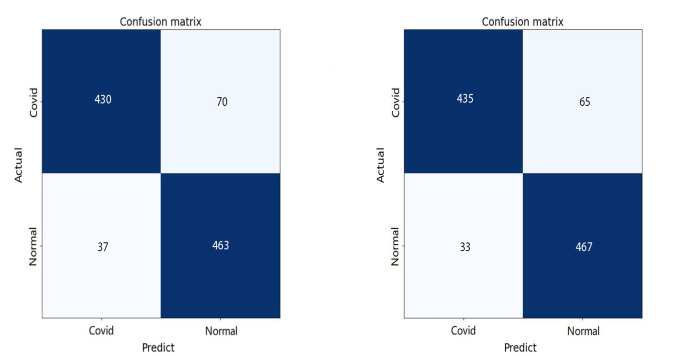


Fig 7. Covid or Normal: Accuracy and Loss of VGG19 model on X-ray Image Dataset



(a) Validation (b) Test

Fig 8. Confusion matrix for VGG19 on X-ray Image dataset: COVID-19 or Normal

Measurements for VGG19 Model Figure 7 shows that after training the Covid-19 Model with 100 Epoch we got an accuracy of 89.17%. The accuracy is getting lower than model InceptionV3. Also, the loss is reduced. Figure 8 Shows that in the VGG19 model, COVID Test data set 435 pictures are properly classified as COVID, whereas 65 images are classified as Normal, and from the Normal data set the model also correctly identifies 467 samples as normal, but only 33 as COVID.

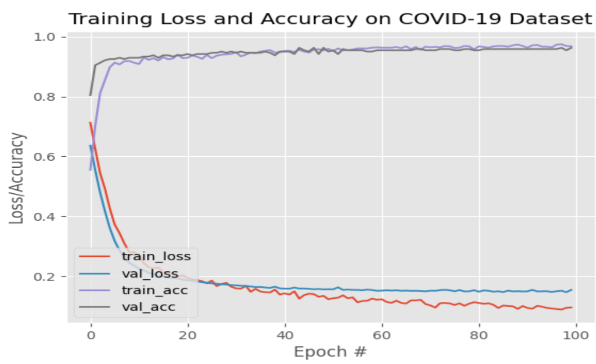
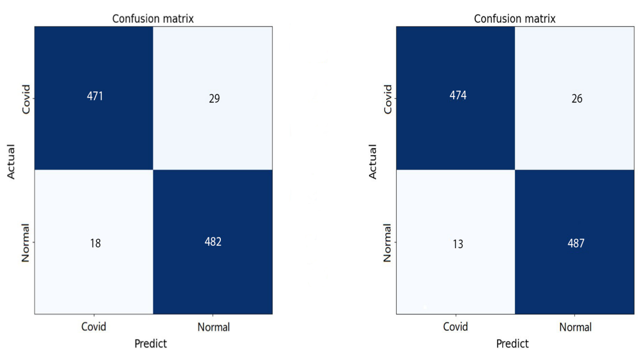


Fig 9. COVID-19 or Normal: Accuracy and Loss of VGG16 model on X-ray Image Dataset



(a) Validation (b) Test

Fig 10. Confusion matrix for VGG16 for COVID or Normal of on X-ray Image dataset

Measurements for VGG16 Model Figure 9 shows that after training the Covid-19 Model with 100 Epoch we got an accuracy of 95.32%. The accuracy gets higher respectively with increasing Epochs. Also, the loss is reduced very strongly.

Figure 10 Shows, In the Vgg16 model, from the COVID Test data set, 476 pictures are properly classified as COVID, whereas 26 images are classified as Normal, and from the Normal data set the model also correctly identifies 487 samples as normal, but only 13 as COVID.

4.3.3 Comparison, Evaluation and Performance Analysis

Table 2 shows the overall accuracy of the proposed "COVID-19 Identification System from X-Ray Images of the Chest Using Deep Neural Network" systems. We train four existing pre-trained models on the same dataset and compare their results.

Table 2. Performance analysis for Covid/Normal dataset for our trained model

Model	Validation Accuracy (VA)	Testing Accuracy (TA)	Sensitivity (SN)	Specificity (SP)
VGG16	**95.32%**	**96.10%**	95.00%	97.50%
Inception V3	94.58%	95.10%	96.67%	92.50%
VGG19	89.17%	90.22%	85.83%	92.50%
ResNet50	87.50%	88.50%	85.00%	90.00%

We can observe from the above performance in Table 2 that the VGG16 model has greater overall accuracy than the other three models. In this case, the VGG16 model maintained greater than 95% accuracy in both validation and testing. No other model can't reach that much accuracy. So, we choose the VGG-16 model as our primary model to modify with Transfer Learning.

Table 3. Comparison Matrix of Precision, F1 Score and Recall for Covid/Normal dataset.

Model	Dataset	Precision	Recall	F1 Score
VGG16	**Covid**	**0.97**	**0.95**	**0.95**
	Normal	0.95	0.97	0.96
InceptionV3	**Covid**	0.93	0.97	0.95
	Normal	0.97	0.93	0.94
VGG19	**Covid**	0.92	0.86	0.89
	Normal	0.87	0.93	0.90
ResNet50	**Covid**	0.89	0.85	0.87
	Normal	0.86	0.90	0.88

The precision of testing data indicates the precision of the deep learning models. Table 3 shows the value of the greatest accuracy and recall for differentiating COVID-19 instances is 0.97 and 0.95, respectively, achieved by VGG16, followed by InceptionV3 with values of 0.93 and 0.97, respectively.

The VGG19 has the lowest accuracy and recall values of 0.93 and 0.86, respectively, while the ResNet50 has the lowest precision of 0.89 and recall of 0.85. To detect normal instances, VGG16 has the highest precision and recall values of 0.95 and 0.97, while Inception has the highest precision and recall values of 0.97 and 0.93, respectively. VGG19 has the lowest accuracy rate of 0.87. Furthermore, ResNet50 has the lowest value of 0.86 for Recall. Though InceptionV3 has a similar precision to VGG16, VGG16 outperforms it in the "COVID" dataset.

In addition, compared to other parameters recall and F1-score, the VGG16 model consistently and occasionally outperforms them. ResNet50 models scored quite low across all Pre-Training models.

Tables 2 and 3 present the resulting confusion matrices on hidden test datasets to more clearly illustrate how each model performs in classification. Table 2 shows that VGG16 and Inceptionv3 have the highest categorization accuracy. As a result, based on their confusion matrix, it appears that VGG16 can properly detect 474 photos in the COVID class, even though 26 images are classified as Normal class from the COVID test dataset.

So, as discussed above, the VGG-16 model can determine whether an image is COVID or Normal with high accuracy.

4.3.4 Computational Time

As we already indicated, the NVIDIA GeForce GTX 1060ti GPU from our local system is used in this experiment. The Training time per epoch for various CNN models is displayed in Table 4.

Table 4. Training time of four CNN models

Model	Batch size	Training Time(Per epochs)
VGG16	64	103s
InceptionV3	64	107s
VGG19	64	99s
ResNet50	64	124s

So, as discussed above, we chose the VGG-16 model as our primary model. As a result, it is clear to use the VGG16 models to determine the patient's health state using chest X-ray pictures.

4.4 Transfer Learning

The VGG-16 model is our core model, and we use Transfer Learning to increase overall accuracy, specificity, and other conventional metrics. A common deep learning strategy in computer vision systems today is transfer learning (TL). TL allows us to generate an acceptable prediction without having to restart the learning process [32] [33].

4.4.1 Reduce Over Fitting

If data augmentation is not employed during training in the VGG16 model, overfitting issues may arise. There are various techniques for decreasing the problem of overfitting to solve this. First and foremost, the expansion of training datasets may aid in the resolution of this problem. Second, data augmentation techniques such as picture rotation, zooming in or out, and flipping images horizontally or vertically are beneficial. We utilize data augmentation and fine-tune with significant dropout to minimize overfitting.

Image Augmentation To get a better performance we applied image augmentation in both Covid positive and Normal CXR images. Table 5 shows the Parameter that we use to augment the image. Here we get 20 images from 1 image by using different parameters.

Table 5. Initialize The Training Data Augmentation Object

Augmentation Parameter	Value	No Of Images
Rescale	1./255	20
Rotation Range	30	
Width Shift Range	0.3	
Height Shift Range	0.3	
Shear Range	0.3	
Zoom Range	0.3	
Fill Mode	nearest	
Horizontal Flip	TRUE	

Table 5 shows that we utilize the ImageDataGenerator to enhance the variety of data used to train models that rotate, shift, shear, zoom, and flip images. The classifiers are then applied to a 50% dropout to reduce overfitting.

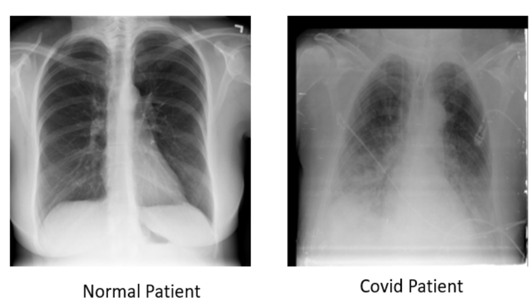


Fig 11. CXR Images of Normal Patient and Covid Patient

Figure 11 showing the Covid-19 NORMAL CXR Image and Covid-19 positive CXR Image. Figure 12 showing the output images of normal patients and covid patients after augmentation was applied. We split 1 image into 20 images in every image augmentation. That gives us 20 different views of one image. Different parameters are used to reduce the overfitting. We train the augmented images again to get better accuracy.

Fine Tuning Kara's VGG16 CNN pre-trained model is uploaded, the convolutional base is frozen, and the top layers are fine-tuned in the transfer learning stage. For feature extraction, a pre-trained model with layers that have been fine-tuned is used. Basic Dense Layer neural networks were previously frozen to preserve Image Net weights during training. To avert overfitting in the model, dropout is used in the layers that are fully linked. It typically hovers around 0.5, and the model has been trained to produce particular measurements. If the measurements are excellent and the overfitting is excellent, reduce the dropout. Increases dropout if overfitting is still severe. In our approach, the drop ratios of 0.3 and 0.2 were used.

Algorithm 1 The Proposed Algorithm in VGG16 Model

procedure OUR PROCEDURE
layer ← pretrainmodel.layer
True ← pretrainmodel.trainable
False ← settrainable
0 ← count
top:
if `layer<0` **then return** false
1 ← count
loop:
if layer(name) = (layername) **then**
True ← settrainable
goto loop
close;
goto top
if settrainable: **then**
True ← layer.trainable **else** False ← layer.trainable **close**

Table 6. Fine-Tuning based on VGG16 Pre-training

Layer (Type)	Output Shape	Param #
block3_conv1 (Conv2D)	(None, 37, 37, 256)	295168
block3_conv2 (Conv2D)	(None, 37, 37, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 18, 18, 256)	0
block4_conv1 (Conv2D)	(None, 18, 18, 512)	1180160
block4_conv2 (Conv2D)	(None, 18, 18, 512)	2359808
block4_conv3 (Conv2D)	(None, 18, 18, 512)	2359808
block4_pool (MaxPooling2D)	(None, 9, 9, 512)	0
block5_conv1 (Conv2D)	(None, 9, 9, 512)	2359808
block5_conv2 (Conv2D)	(None, 9, 9, 512)	2359808
block5_conv3 (Conv2D)	(None, 9, 9, 512)	2359808
block5_pool (MaxPooling2D)	(None, 4, 4, 512)	0

Table 7. Accuracy, Sensitivity and Specificity of Modified VGG16 Model for 100 EPOCH and 200 EPOCH

Model	Validation Accuracy (VA)	Testing Accuracy (TA)	Sensitivity (SN)	Specificity (SP)
Modified VGG16 with 100 Epoch	96.00%	**97.10%**	97.00%	95.00%
Modified VGG16 with 200 Epoch	**96.50%**	96.70%	**98.00%**	**95.00%**

The proposed algorithm searches through a pre-trained model's layers repeatedly until it locates a layer with an identifiable name. It sets the trainable attribute of that layer based on the value of set-trainable when the layer

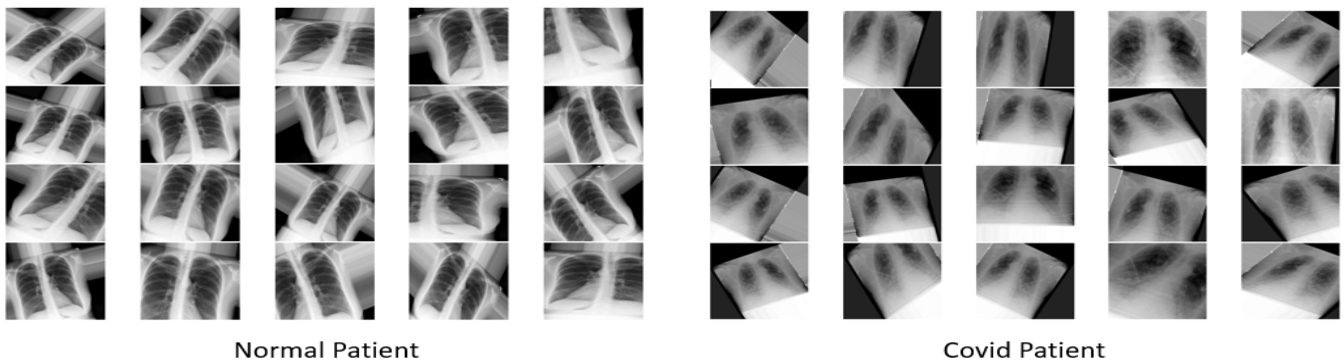


Fig 12. Output Images of Normal Patient and Covid Patient After Augmentation

with the provided name has been located. The layer turns trainable if set-trainable is true; otherwise, it stays untouched. The count variable is used by the method to track the number of layers processed. The program terminates when all of the layers have been evaluated or if the layer index goes negative.

In Table 6, the summary of the VGG16 model, one layer is chosen to be unfrozen. Change the layer name once it has been selected to enable training of the classifier layer while the other layers are frozen. We choose a layer at random from the top of the model since levels before it is more general. For VGG16, for example, the block4 conv1 layer is used. Table 6 depicts fine-tuning based on VGG16 pre-training. After completing all of the essential TL stages, we trained the VGG16 model in the same scenario with enhanced EPOCH and obtained an overall superior performance with an accuracy of more than 96.50 %.

Table 8. Precision, Recall and F1-score of Modified VGG16 Model for 100 EPOCH and 200 EPOCH

Model	Dataset	Precision	Recall	F1 Score
Modified VGG16 with 100 Epoch	**Covid**	0.95	0.97	0.96
	Normal	0.97	0.95	0.96
Modified VGG16 with 200 Epoch	**Covid**	0.96	0.98	0.97
	Normal	0.98	0.96	0.96

4.5 Performance Measurement with Modified Model

After applying different TL we train the same dataset in the Modified VGG16 Model again for 50, 100, and 200 EPOCH. For 50 EPOCH in VGG16, we get 93.82% validation and 94.34% testing accuracy. But we get improved performance when we trained the dataset for 100 and 200 EPOCH.

Table 9. Comparison of the VGG16 Model Before and After Applying Transfer Learning Approach

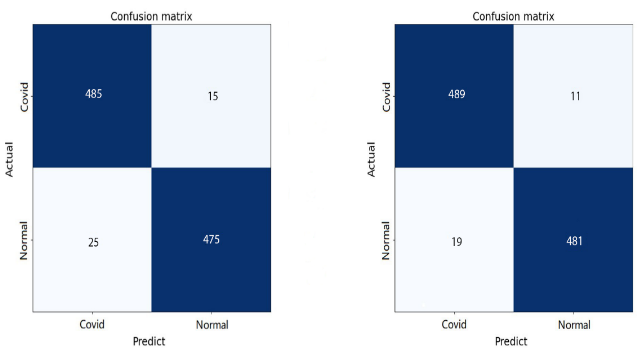
Model	Validation Accuracy (VA)	Testing Accuracy (TA)	Sensitivity (SN)	Specificity (SP)
Modified VGG16 (Before Applying TL) with 100 Epoch	95.32%	96.10%	95.00%	97.50%
Modified VGG16 (After Applying TL) with 100 Epoch	**96.00%**	**97.10%**	**97.00%**	95.00%
Modified VGG16 (After Applying TL) with 200 Epoch	**96.50%**	**96.70%**	**98.00%**	95.00%

Table 7 is showing When were applying TL in modified VGG16 with 100 Epoch and 200 Epoch then we get the Validation Accuracy (VA), Testing accuracy (TA), Sensitivity (SN), and Specificity (SP) of the Modified VGG16 Model. Here in this model, we had got a validation accuracy of 96.00%, Testing Accuracy of 97.10%, a sensitivity of 97.00%, and Specificity of 95.00% for 100 Epoch, and a validation accuracy of 96.50%, testing accuracy of 96.70%, a sensitivity of 98.00%, and specificity of 95.00% for 200 Epoch which is better than the previous modified VGG16 model without Transfer Learning.

After applying TL in the modified VGG16 model, In 13,

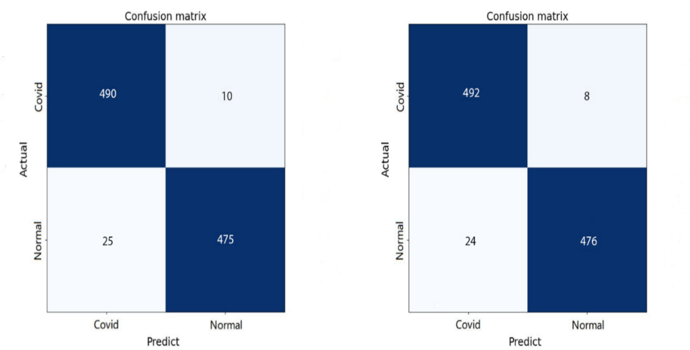
Table 10. Before and After Applying TL Approach: Comparison Matrix of Precision, Recall, and F1 Score

Model	Dataset	Precision	Recall	F1 Score
Modified VGG16 (Before Applying TL) with 100 Epoch	**Covid**	0.97	0.95	0.95
	Normal	0.95	0.97	0.96
Modified VGG16 (After Applying TL) with 100 Epoch	**Covid**	0.95	0.97	0.96
	Normal	0.97	0.95	0.96
Modified VGG16 (After Applying TL) with 200 Epoch	**Covid**	0.96	0.98	0.97
	Normal	0.98	0.96	0.96



(a) Validation (b) Test

Fig 13. Confusion matrix for Modified VGG16 on X-ray dataset: COVID-19/Normal for 100 Epoch



(a) Validation (b) Test

Fig 15. Covid or Normal: Confusion matrix for Modified VGG16 on X-ray Image dataset for 200 Epoch

we can see the confusion matrix of the best-performing algorithms. we can see that the networks accurately detect 489 copies in the COVID class from the Covid Test data set, but only 11 copies in the Normal class. The model also correctly identifies 481 samples as normal, but just 48 as COVID from the normal test data set.

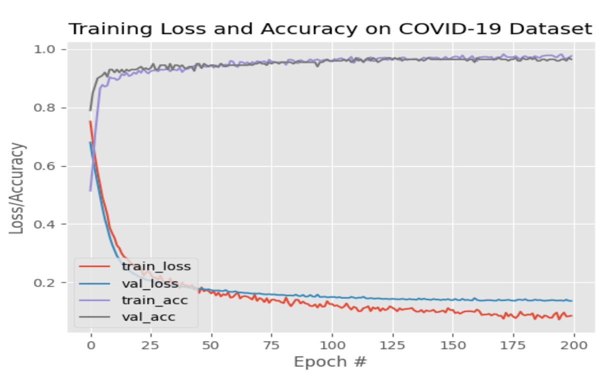


Fig 14. Covid or Normal: Accuracy and Loss of Modified VGG16 model on same X-ray Image Dataset for 200 Epoch

The confusion matrix of the highest-performing algorithms is shown in Figure 15. From the Covid Test data set, we can see that the networks correctly recognize 492 photos in the COVID class but only 8 images in the Normal class. From the normal test data set, the model properly detects 476 photos as normal, but only 24 images as COVID.

4.6 Before and After Applying Transfer Learning: Performance Analysis and Comparison of Modified VGG16 Model

From Table 9 we can see that when we train the modified VGG16 model for 100 Epoch, it provides better accuracy in both sectors. Previously we got 95.32 % validation accuracy and 96.10% test accuracy. **After applying TL in modified VGG16 we get 96.00 % validation accuracy and 97.10% test accuracy which is clearly an improvement over the previous VGG16 Model.**

Again, when we train the modified VGG16 model for 200 Epoch, it provide the best validation accuracy but testing accuracy decreased anyhow. So, increasing the number of epochs neither affected that much overall performance in our modified VGG16 Model.

5 Conclusion

Our COVID-19 identification method works for COVID and Normal pictures in CXR format. We can also adapt the system to work with different CNN models. We intend to rebuild the entire deep-learning version from the ground up as soon as possible. This research contributes to adding dropout filters to the VGG-16 neural network to decrease overfitting and generate a more generic outcome with lower computing costs. Moreover, we employed and modified a VGG16-based neural network and applied transfer learning to identify Covid-19 more accurately in this platform. We believe that a new model and additional flexibility on our part will create a more realistic scenario. By exposing them to the loads, we may be able to open up new chances to use this belonging for future work, and this may communicate new opportunities to move one step further in addition to increasing COVID-19 detecting identification in the future.

5.1 Discussion

Due to its extremely infectious nature, the coronavirus infection (Covid-19) threatens the lives of billions of people. According to WHO, Both the overall infection rate and the mortality rate are rising significantly. This viral infection causes inflammation in the lungs of infected individuals. Therefore, a chest X-ray is one method that could be used to find these inflammations.
Using chest X-ray pictures, we were able to separate COVID-19 individuals from normal and COVID cases. We used four distinct CNN models, and their performance was assessed using several performance indicators. The collected findings show that the VGG16 model is the ideal model for this job. Then we modified and fine-tuned the VGG-16 model and applied transfer learning to get better accuracy and minimize the computation time while achieving a high accuracy rate.

5.2 Limitations and Future Work

We tried to design and improve Deep learning models with the best accuracy. To do so, we have faced many obstacles in many aspects. The following are the main limitation of our research:

- we aimed to work with COVID, Normal CXR images along with other viral Pneumonia. But we were only able to work with COVID X-ray and Normal X-ray images.
- We have collected the dataset only from online sources instead of the local hospitals and medical intuition.

In our research work, we presented a basic CNN design that outperforms CNNs in both the training and testing stages. To create CNN models, data from other sources must be included. To achieve this aim, we will develop a new model in our future work that can identify CXR pictures from the COVID, Normal, and Pneumonia Datasets. In addition, to improve accuracy, we will train our model with a huge quantity of data and add more dropout layers, as well as reduce overfitting.

Acknowledgements

This work was supported financially in part by the Department of CSE, GUB.

References

[1] C. Sohrabi, Z. Alsafi, N. O'neill, *et al.*, "World health organization declares global emergency: A review of the 2019 novel coronavirus (covid-19)," *International journal of surgery*, vol. 76, pp. 71–76, 2020.

[2] W. H. Organization *et al.*, *Weekly epidemiological update on covid-19–29 june 2021. 2021.*

[3] T. L. Sandri, J. Inoue, J. Geiger, *et al.*, "Complementary methods for sars-cov-2 diagnosis in times of material shortage," *Scientific reports*, vol. 11, no. 1, pp. 1–8, 2021.

[4] G. Bargshady, X. Zhou, P. D. Barua, R. Gururajan, Y. Li, and U. R. Acharya, "Application of cyclegan and transfer learning techniques for automated detection of covid-19 using x-ray images," *Pattern Recognition Letters*, vol. 153, pp. 67–74, 2022, ISSN: 0167-8655. DOI: `https://doi.org/10.1016/j.patrec.2021.11.020`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S0167%20865521004128`.

[5] N. Narayan Das, N. Kumar, M. Kaur, V. Kumar, and D. Singh, "Automated deep transfer learning-based approach for detection of covid-19 infection in chest x-rays," *IRBM*, vol. 43, no. 2, pp. 114–119, 2022, ISSN: 1959-0318. DOI: `https://doi.org/10.1016/j.irbm.2020.07.001`. [Online]. Available: `https://www.sciencedirect.com/science/article/pii/S1959031820301172`.

[6] M. Elsharkawy, A. Sharafeldeen, F. Taher, *et al.*, "Early assessment of lung function in coronavirus patients using invariant markers from chest x-rays images," *Scientific Reports*, vol. 11, no. 1, pp. 1–11, 2021.

[7] B. E. Bejnordi, M. Veta, P. J. Van Diest, *et al.*, "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer," *Jama*, vol. 318, no. 22, pp. 2199–2210, 2017.

[8] J. Amin, M. Sharif, M. Yasmin, and S. L. Fernandes, "Big data analysis for brain tumor detection: Deep convolutional neural networks," *Future Generation Computer Systems*, vol. 87, pp. 290–297, 2018.

[9] L. A. Pastur-Romay, F. Cedrón, A. Pazos, and A. B. Porto-Pazos, "Deep artificial neural networks and neuromorphic chips for big data analysis: Pharmaceutical and bioinformatics applications," *International journal of molecular sciences*, vol. 17, no. 8, p. 1313, 2016.

[10] A. I. Khan, J. L. Shah, and M. M. Bhat, "Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images," *Computer Methods and Programs in Biomedicine*, vol. 196, p. 105 581, 2020.

[11] Y. Fang, H. Zhang, J. Xie, *et al.*, "Sensitivity of chest ct for covid-19: Comparison to rt-pcr," *Radiology*, vol. 296, no. 2, E115–E117, 2020.

[12] V. Rajinikanth, S. Kadry, K. P. Thanaraj, K. Kamalanand, and S. Seo, "Firefly-algorithm supported scheme to detect covid-19 lesion in lung ct scan images using shannon entropy and markov-random-field," *arXiv preprint arXiv:2004.09239*, 2020.

[13] S. Rajaraman, J. Siegelman, P. O. Alderson, L. S. Folio, L. R. Folio, and S. K. Antani, "Iteratively pruned deep learning ensembles for covid-19 detection in chest x-rays," *IEEE Access*, vol. 8, pp. 115 041–115 050, 2020.

[14] M. M. Islam, F. Karray, R. Alhajj, and J. Zeng, "A review on deep learning techniques for the diagnosis of novel coronavirus (covid-19)," *IEEE Access*, vol. 9, pp. 30 551–30 572, 2021.

[15] M. E. Chowdhury, T. Rahman, A. Khandakar, *et al.*, "Can ai help in screening viral and covid-19 pneumonia?" *IEEE Access*, vol. 8, pp. 132 665–132 676, 2020.

[16] M. Rahimzadeh and A. Attar, "A modified deep convolutional neural network for detecting covid-19 and pneumonia from chest x-ray images based on the concatenation of xception and resnet50v2," *Informatics in Medicine Unlocked*, vol. 19, p. 100 360, 2020.

[17] H. S. Maghdid, A. T. Asaad, K. Z. Ghafoor, A. S. Sadiq, S. Mirjalili, and M. K. Khan, "Diagnosing covid-19 pneumonia from x-ray and ct images using deep learning and transfer learning algorithms," in *Multimodal Image Exploitation and Learning 2021*, International Society for Optics and Photonics, vol. 11734, 2021, 117340E.

[18] I. D. Apostolopoulos, S. I. Aznaouridis, and M. A. Tzani, "Extracting possibly representative covid-19 biomarkers from x-ray images with deep learning approach and image data related to pulmonary diseases," *Journal of Medical and Biological Engineering*, p. 1, 2020.

[19] A. Sufian, A. Ghosh, A. S. Sadiq, and F. Smarandache, "A survey on deep transfer learning to edge computing for mitigating the covid-19 pandemic," *Journal of Systems Architecture*, vol. 108, p. 101 830, 2020.

[20] L. Zhang, L. Lu, I. Nogues, R. M. Summers, S. Liu, and J. Yao, "Deeppap: Deep convolutional networks for cervical cell classification," *IEEE journal of biomedical and health informatics*, vol. 21, no. 6, pp. 1633–1643, 2017.

[21] A. Abbas, M. M. Abdelsamea, and M. M. Gaber, "Classification of covid-19 in chest x-ray images using detrac deep convolutional neural network," *Applied Intelligence*, vol. 51, no. 2, pp. 854–864, 2021.

[22] K. Sahinbas and F. O. Catak, "Transfer learning-based convolutional neural network for covid-19 detection with x-ray images," in *Data science for COVID-19*, Elsevier, 2021, pp. 451–466.

[23] L. T. Duong, P. T. Nguyen, L. Iovino, and M. Flammini, "Automatic detection of covid-19 from chest x-ray and lung computed tomography images using deep neural networks and transfer learning," *Applied Soft Computing*, vol. 132, p. 109 851, 2023.

[24] M. M. Taresh, N. Zhu, T. A. A. Ali, A. S. Hameed, and M. L. Mutar, "Transfer learning to detect covid-19 automatically from x-ray images using convolutional neural networks," *International Journal of Biomedical Imaging*, vol. 2021, pp. 1–9, 2021.

[25] S. Asif, Y. Wenhui, K. Amjad, H. Jin, Y. Tao, and S. Jinhai, "Detection of covid-19 from chest x-ray images: Boosting the performance with convolutional neural network and transfer learning," *Expert Systems*, vol. 40, no. 1, e13099, 2023.

[26] "Rahman, t. covid-19 radiography database (2021). https:// www. kaggle. com/ tawsi furra hman/ covid 19- radio graphy- datab ase.."

[27] P. Autee, S. Bagwe, V. Shah, and K. Srivastava, "Stack-net-denvis: A multi-layer perceptron stacked ensembling approach for covid-19 detection using x-ray images," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 4, pp. 1399–1414, 2020.

[28] S. Bandyopadhyay and S. DUTTA, "Associating unemployment with panic attack using stacked-rnn model during covid-19," 2020.

[29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.

[30] "Https://www.kaggle.com/imsparsh/large-categories-image-classifier-inception-v3."

[31] D. S. Kermany, M. Goldbaum, W. Cai, *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.

[32] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.

[33] W. Rawat and Z.Wang, "Deep convolutional neural networks for image classification: Comprehensive review, neural computation" (2017)," vol. 29(9), pp. 2352–2449. 2017.

Farhana Akter Sunny was born in Narsingdi, Dhaka, Bangladesh, in 1992. She received B.Sc. (Hons.) and M.Sc. in Information Technology from Jahangirnagar University. Now she is enrolled to pursue her PhD in the Institute of Information Technology at Jahangirnagar University. At present, she is working as an Assistant Professor, Dept. of Computer Science and Engineering at the Green University of Bangladesh. Previously, she worked as a Lecturer in the Department of Computer Science and Engineering at the European University of Bangladesh until May 2016. Besides her full-time academic teaching, she also worked as a Program Coordinator in the CSE department from 2021 to 2023. She has published many research papers in different International Conferences and Journals. Her research interest areas are Internet of Things, Cyber Security, Semantic Communication, Block Chain Technology, and Machine Learning.

Sagufta Sabah Nakshi was born in Dinajpur, Bangladesh, in 1999. She received her B.Sc. in Computer Science and Engineering from Chittagong University of Engineering and Technology. Now she is enrolled to pursue her M.Sc. in Computer Science and Engineering at Chittagong University of Engineering and Technology. At present, she is working as a Lecturer, Dept. of Computer Science and Engineering at the Green University of Bangladesh from November 2022. Her research interests include Image Processing, BlockChain Technology, Computer Vision, Bioinformatics, and Machine Learning.

Mashud Parbhez was a weekend student of the Department of Computer Science & Engineering, Green University of Bangladesh. He has completed his B.Sc. in 2021. He completed his Diploma in Department of Computer Science & Engineeringin from Feni. Currently, he is working as a Sr. Software Engineer at DevsZone. His current research interests include Computer Vision, Image Processing, Human-Computer Interactio nSystem, and Robotics.

Muhammad Aminur Rahaman was born in Tangail, Bangladesh, in 1981. He received his Ph.D. from the Department of Computer Science & Engineering, University of Dhaka, Bangladesh in 2018. Rahaman has completed his B.Sc. and M.Sc. degree from the Department of Computer Science & Engineering, Islamic University, Kushtia, Bangladesh in 2003 and 2004, respectively. He is a founder Director of Worldgaon (Pvt.)Limited. He is working as an Associate Professor and Chairperson at the Department of Computer Science and Engineering, Green University of Bangladesh, Dhaka, Bangladesh since September 2018 to present. His current research interests include Computer Vision, Image Processing, Human-Computer InteractionSystem, Robotics and Software Engineering. He is a senior member of IEEE.