Deep Transfer Learning Platforms for SARS-CoV-19 Diagnostics based on Human Lungs CT Scan Imaging

Krishna Kumar Joshi[1*], Kamlesh Gupta[2] and Jitendra Agrawal[3]

[1] SoIT, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, India
[2] Rustam Ji Institute of Technology, Tekanpur, Gwalior, Madhya Pradesh, India
[3] SoIT, Rajiv Gandhi Proudyogiki Vishwavidyalaya, Bhopal, Madhya Pradesh, India
[1*]krishnakjoshi@gmail.com
[2]kamlesh_rjitbsf@yahoo.co.in
[3]jitendra@rgpv.ac.in

Abstract. COVID-19 is one of the most serious diseases caused by the SARS coronavirus. This is a fatal disease, and it becomes difficult to save the life of the infected person because it progresses very rapidly and starts in the lungs and causing damage to the entire body. With the help of advanced machine learning techniques, this disease can be detected early and with very little probability of error. In this paper, we have developed two deep transfer learning models, VGG 16 and MobileNetv2, to detect COVID-19. Both models were applied to datasets based on human lung CT scan images, and their performance was evaluated. To analyze and compare their performance, we have used several performance metrics such as Prevalence, Null Error Rate, False DR, Negative PV, False OR, LR ratio (+), LR ratio (-), CSI, Accuracy, FM Index, BM, Diagnostic Ratio, MK and Critical Success Index, that have not been used in previous papers.

Keywords: Transfer Learning, Deep Learning, CT Scan Imaging, SARS-CoV-19, Critical Success Index, Diagnostic Ratio, False Omission Rate

1 Introduction

SARS coronavirus is a deadly disease that has caused a lot of deaths across the world. Initially, it was complicated to identify coronavirus correctly because its symptoms were similar to fever, cold, and cough. That is why it was challenging to identify it in the initial stages. The most effective and reliable technique to detect coronavirus is a lung CT scan because coronavirus mainly affects the lungs and damages them very quickly, which can even lead to death. Machine learning techniques are used to effectively diagnose many diseases and medical applications, and it has also been used to effectively diagnose coronavirus. Although these works were quite effective, there were many shortcomings in them. Our aim in this presented approach is to minimize those shortcomings and develop an effective technique.

If we talk about India, then the number of confirmed cases here till April 2024 has reached 45,035,393, out of which 533,570 people have died, and the rest have been saved. If we talk about China, the number of cases there till April 2024 was 503,302, out of which 5272 people died and the rest were saved. The most effective way to prevent it is to detect it in its early stages. But many times, this was not possible because the methods of detection were not so accurate, and also took a lot of time.

The scope of Machine Learning techniques is expanding in medical diagnosis as they are capable of giving very accurate and fast results. All these special features of machine learning techniques inspired us to use these techniques for effective, fast, accurate, and correct diagnosis of COVID-19 so that COVID-19 patients can be classified correctly and quickly using our approach.

R. Nagariya et al. (eds.), *Proceedings of the International Conference on Sustainable Business Practices and Innovative Models (ICSBPIM-2025)*, Advances in Economics, Business and Management Research 349,
https://doi.org/10.2991/978-94-6463-872-1_18

We have proposed two approaches based on MobileNetV2 and VGG 16, respectively, and then compared them on the scale of various performance parameters.

2 Background Study and Motivation

Machine Learning technology is helping in various fields today due to its effectiveness. The foundation of machine learning, its methodologies, and various applications are covered in [1] and [2]. Machine learning with security can be helpful to achieve some more complicated goals as discussed in [3]. These techniques are helpful in various medical diagnoses, such as Breast cancer [4], heart disease acute prediction [5], melanoma disease prediction [6], Classification of Brain tumor through Efficient B3 model [7], through IFAS approach [8] and applying U-Net on 3D MRI scan images [9]. Some other notable work on VOID-19 regarding long COVID sequelae based on the clinical data is provided in [10-11]. Various approaches for the prediction of COVID discussed in [12][13][14], mostly based on Transfer learning, Internet of Things, Deep Learning, and AI. A detailed review is provided in [15]. The impact of alcohol consumption on COVID-19 is discussed in [20].

After studying many recent research papers, many limitations have been identified. Some of the limitations or gaps in the previous approaches are identified as:

- Most of the work is done on small datasets [16]. These datasets contain very few images; we cannot ensure that the applied model or approach will provide the best results, as its scope is limited to a very small number of images.
- Very few image pre-processing techniques applied in the previous approaches provided low-quality results [17].
- In most of the approaches, we found that the work is done in older CNNs such as AlexNet or ResNet, which have few limitations [18].
- In most of the approaches, we have identified that basic classifiers such as SVM Classifiers or KNN classifiers are used. These classifiers are good for text-type datasets only [19].
- Very few parameters were taken for the validation of the work.

The main objective is the diagnosis of COVID-19 while considering its increasing impact on fast and accurate results of machine learning and its derived techniques. Since many approaches have already been used to diagnose COVID-19, it is also our objective to understand their shortcomings and remove them in the new approach, with the help of various existing feature extraction algorithms, such as histogram analysis, Histogram Equalization, denoising, convolutional networks, etc. Applying the various existing algorithms used for classification, such as the confusion matrix, for detection and classification, thus enhancing and improving the accuracy by using the hybrid techniques such as deep learning. And finally, compared the performance of the approaches themselves.

3 Materials Used and Methodology

This section covers information about the datasets, image and data pre-processing, machine learning techniques used, deep learning models, model formation, including compilation methods, various augmentation techniques, training and testing procedures, and performance parameters with their descriptions and mathematical formulas.

Our problem statement is as follows: To develop an approach that can effectively and accurately diagnose coronavirus and correctly classify COVID and non-COVID cases with the help of advanced ML techniques using a dataset based on human lung CT scan images. For this, we have specially collected datasets from trusted public sources and combined them to prepare a large dataset, and then designed an effective model so that the results are highly accurate

In the proposed work, we have tried to overcome all or most of the limitations or research gaps mentioned. Our research methodology focuses on all of the research gaps, and we divided it stepwise to cover all of these remedies:

- To study existing pre-processing techniques on the dataset of COVID-19.
- To study existing feature extraction algorithms on the COVID-19 chest dataset.
- To study existing classification algorithms for detecting COVID-19.
- To enhance the performance using various advanced techniques.
- To provide the analytical results where feasible.
- To compare the proposed techniques' results with existing techniques in previous works.
- To give conclusions as well as recommendations for further enhancement.

We have prepared our dataset by merging the two publicly available datasets (shown in Table 1).

Table 1. Information Regarding the Dataset

Dataset Prepared From	Dataset Contains	Dataset Contains Total images	Total COVID Patients Images	Total Non-COVID Patient
SARS-CoV-2 CT Scan Dataset And UCSD-AI4H COVID CT scan dataset	Chest CT Scan Images of Both COVID-19 and non-COVID-19 Patients	3227	1601	1626

We proposed two different approaches here. Figure 1 shows the complete approach for both processes. The only difference between the two approaches is that in the first

approach, we used MobilenetV2 CNN, and in the second approach, we used VGG 16 CNN for applying Transfer Learning.

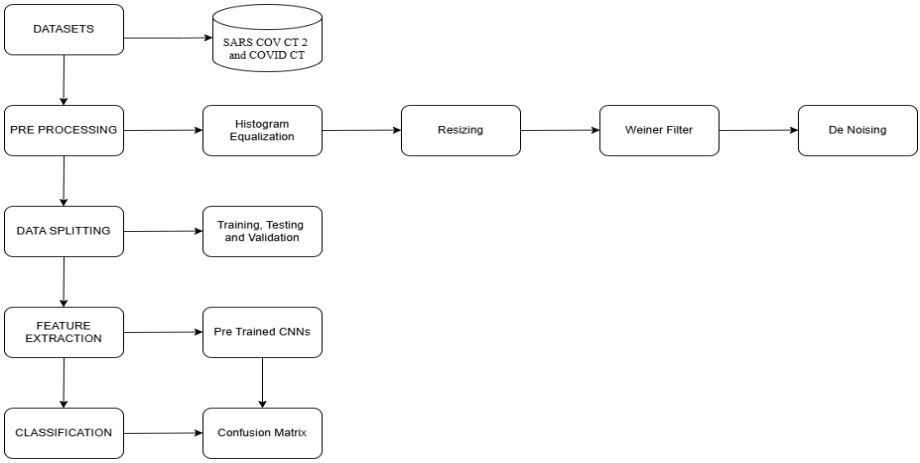

Figure 1. Flow Structure for the proposed approaches

Initially, we applied some of the pre-processing techniques. Firstly, we have resized these images and then applied the histogram equalization using CLAHE on these images to enhance the basic contrast of these images, as they have less contrast. We converted these images into grayscale levels so that better results can be achieved. We have also removed the noise from these images to obtain better quality. Then we break the dataset into three parts for training, testing, and validation**.** We have applied various techniques of data augmentation to prevent the model from overfitting. Information about the augmentation parameters is shown below (Table 2).

Table 2. Augmentation Methods and Their Parametric Values

Augmentation Achieved By	Values Considered
Rotating the Images with	20 Degree
Zooming the image by	20%
Shearing the Image By	20%
Shifting the Images with	20%
Flipped the Images	Horizontally

Optimization of the model is achieved using various optimizers and performance matrices (Table 3). We have taken Binary Cross Entropy as a loss parameter. Learning Rate is another important parameter that influences the stability of the convergence and speed. It is issued to calculate the total number of steps so that an optimized model can be defined by adjusting the weight and bias. Metrics are the performance parameters.

An important metric is accuracy. Epochs are simply one complete pass. We need to define the number of epochs for our model to prevent it from underfitting or overfitting.

Table 3. Model Parameters and Their Parametric Values Taken

Model Compilation Parameters	Values Considered
Optimizer Selected	Adam
Loss Parameter Selected	BCE
Learning Rate Considered	e-4
Metrics Considered	Accuracy
Batch Sizes	15
Epochs Taken	80/50

Then, in the first approach, we applied Mobilenetv2, which has a total of 53 layers which including the input layer, conv 2D layer, bottleneck layer, average pooling layer, and lastly output layer. The input layer takes less input, while the output layer gives output. The conv 2D layer performs the work of convolution, while the average pooling layer does the work of averaging. An important layer in MobileNetV2 is the bottleneck layer, which performs less matrix multiplication and reduces the number of parameters so that the depth of the model can be increased, as well as its processing speed can be increased (Figure 2).

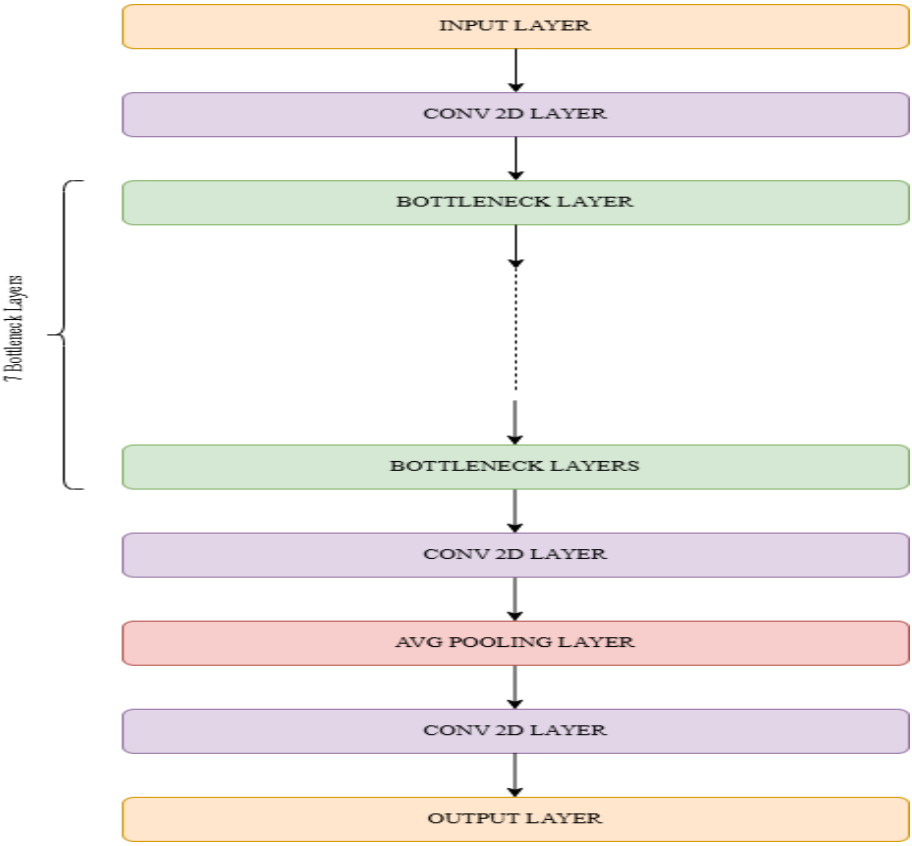

Figure 2. Layered Architecture of MobilenetV2

In the second approach, we applied VGG 16. It has a total of 16 layers, which are the combination of 2 Conv Layers consisting of 64 filters, 2 Conv layers consisting of 128 filters, 3 Conv layers consisting of 256 filters, 6 Conv layers consisting of 512 filters, and three FCC layers of 4096 nodes.

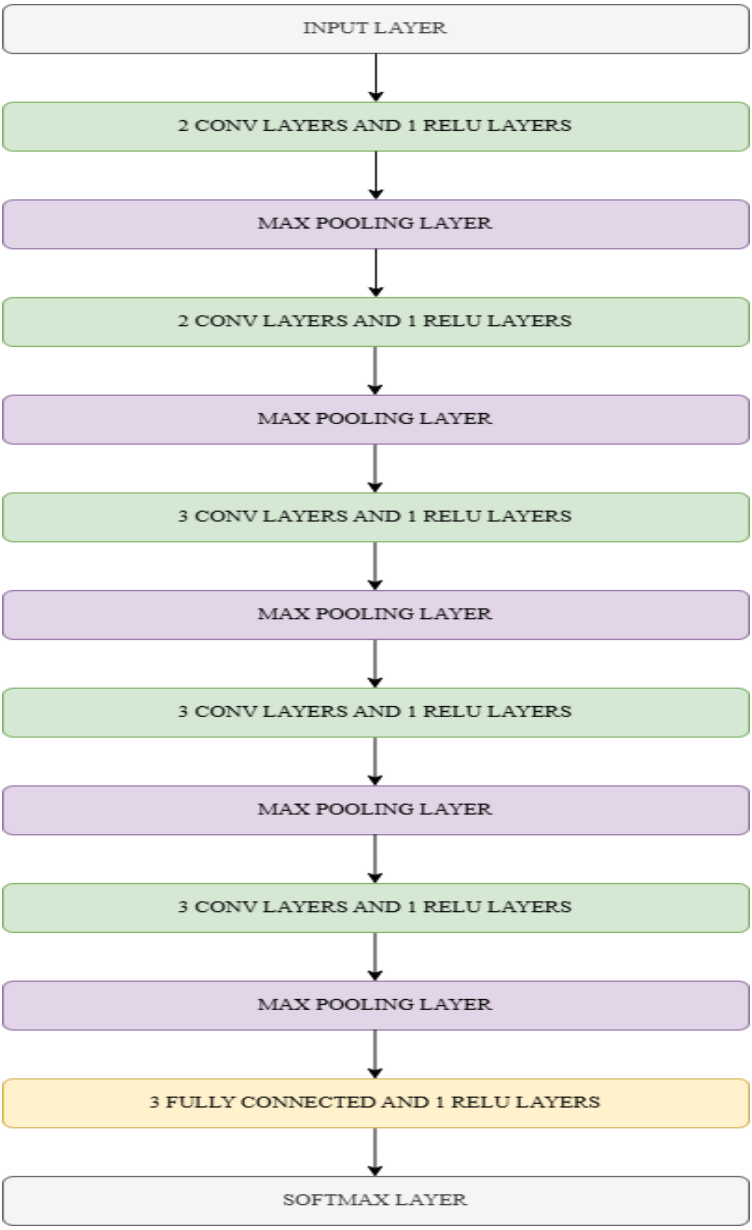

Figure 3. Layered Architecture of VGG 16

4 Training and Testing of the Models

The MobilenetV2-based model (Approach 1) is trained for 80 epochs. One can easily see the difference, which shows that training loss is reduced from the 1st to the 80th epoch, and training accuracy is increased (Figures 4 and 5).

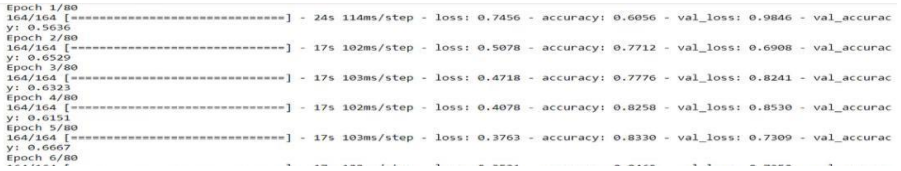

Figure 4. Status of Performance during 1st Epoch

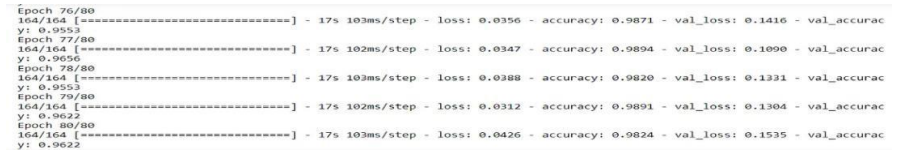

Figure 5. Status of Performance during 80th Epoch

The VGG 16-based model (Approach 2) is trained for 50 epochs. One can easily see the difference, which shows that training loss is reduced from the 1st to 50th epoch and training accuracy is increased (Figure 6 and Figure 7).

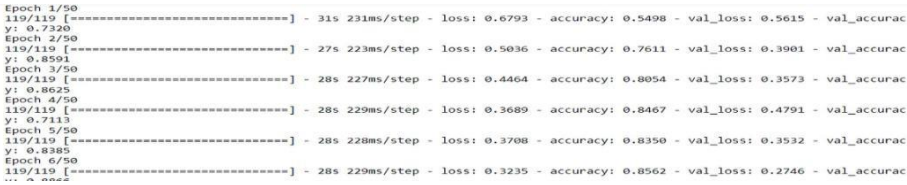

Figure 6. Status of Performance during 1st Epoch

```
y: 0.9519
Epoch 45/50
119/119 [==============================] - 29s 234ms/step - loss: 0.0163 - accuracy: 0.9946 - val_loss: 0.2185 - val_accurac
y: 0.9416
Epoch 46/50
119/119 [==============================] - 29s 233ms/step - loss: 0.0404 - accuracy: 0.9855 - val_loss: 0.1434 - val_accurac
y: 0.9656
Epoch 47/50
119/119 [==============================] - 29s 236ms/step - loss: 0.0416 - accuracy: 0.9813 - val_loss: 0.2366 - val_accurac
y: 0.9038
Epoch 48/50
119/119 [==============================] - 29s 236ms/step - loss: 0.0176 - accuracy: 0.9950 - val_loss: 0.1645 - val_accurac
y: 0.9450
Epoch 49/50
119/119 [==============================] - 29s 236ms/step - loss: 0.0421 - accuracy: 0.9843 - val_loss: 0.1504 - val_accurac
y: 0.9381
Epoch 50/50
119/119 [==============================] - 29s 233ms/step - loss: 0.0225 - accuracy: 0.9958 - val_loss: 0.1099 - val_accurac
y: 0.9725
```

Figure 7. Status of Performance during 50th Epoch

We have plotted the matrix of confusion regarding the testing of our models, and based on it, a classification report is plotted.

In the first approach of MobileNetV2-based, for plotting the confusion matrix, we used 149 support images of Covid-negative patients and 174 images of Covid-positive patients (Figure 8).

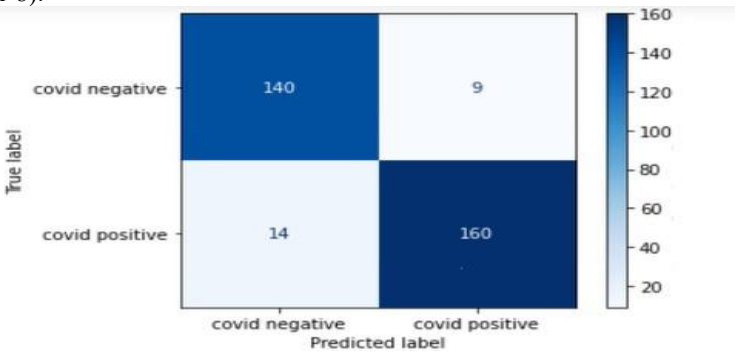

Figure 8. A confusion matrix for the MobilenetV2-based approach.

The Classification report in Figure 9, displaying the values of performance benchmarks based on the following values on TP, TN, FP, and FN:

True (-) = 140,
False (+) = 9,
False (-) = 14 and
True (+) = 166

	precision	recall	f1-score	support
Covid-(-)	0.91	0.94	0.92	149
Covid-(+)	0.95	0.92	0.93	174
accuracy			0.93	323
macro avg	0.93	0.93	0.93	323
weighted avg	0.93	0.93	0.93	323

Figure 9. Classification Report (Testing) contains performance parameter values

Similarly, for the second approach based on VGG 16, the matrix of confusion is plotted (Figure 10).

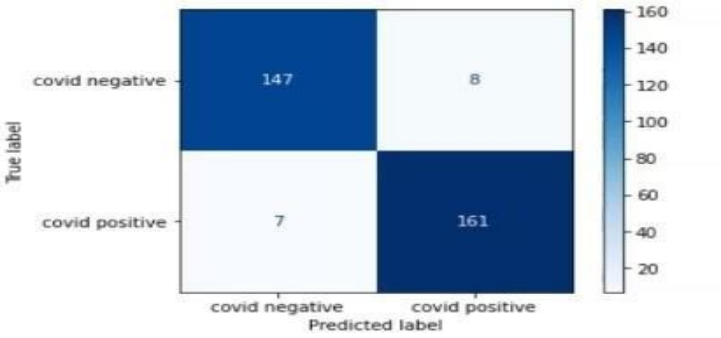

Figure 10. A confusion matrix for the VGG 16-based approach

The Classification report in Figure 9, displaying the values of performance benchmarks based on the following values on TP, TN, FP, and FN:

True (-) = 147
False (+) = 8
False (-) = 7 and
True (+) = 161

	precision	recall	f1-score	support
Covid-(-)	0.95	0.95	0.95	155
Covid-(+)	0.95	0.96	0.96	168
accuracy			0.95	323
macro avg	0.95	0.95	0.95	323
weighted avg	0.95	0.95	0.95	323

Figure 11. Classification Report (Testing) contains performance parameters

5 Comparative Analysis

5.1 Training Performance Comparisons

In the first approach (MobilenetV2-based), we have 80 epochs, while in the second approach (VGG 16-based), we have 50 epochs. So, the comparison of the performance on the first and last epochs is shown in Table 4 below:

Table 4. Model Parameters and Their Parametric Values Compared in First and Last Epochs

Approach	Epochs	Training Loss	Training Accuracy
MobilenetV2 Based	1st	0.7456	0.6056
	80th	0.0426	0.9824
VGG 16 Based	1st	0.6793	0.5498
	50th	0.0225	0.9958

The MobilenetV2-based approach has a loss of 0.7456 and an accuracy of 0.6056 in its first epoch during training, while the loss drops to 0.0426 and the accuracy increases to 0.9824 in its last epoch (80th epoch). The VGG 16-based approach has a loss of 0.6793 and an accuracy of 0.5498 in its first epoch, and by the 50th epoch, the loss decreases to 0.0225 with an accuracy rising to 0.9958. Therefore, we can say the VGG 16-based model performs better during training in terms of both accuracy and loss compared to the MobilenetV2-based approach, even with fewer epochs.

5.2 Testing Performance Comparisons

For performance comparison, during testing, we used the classification report, which is plotted based on the respective confusion matrices. The performance comparison of both models using the fundamental performance parameters based on the classification report is shown here (Table 5).

Table 5. Models and Their Parametric Values Taken

Approach	Type of Patient	Parameters Considered			
		Precision	Recall	F1- Score	Support
MobilenetV2 Based	COVID	0.91	0.94	0.92	149
	Non-COVID	0.95	0.92	0.93	174
VGG 16 Based	COVID	0.95	0.95	0.95	155
	Non-COVID	0.95	0.96	0.96	168

The MobilenetV2-based approach acquired a precision of 0.910, while recall of 0.94, and F1 Score of 0.92 when considering the COVID (-) or normal patients with the support value of 149, while acquired a precision of 0.95, with Recall of 0.92, and F1 Score of 0.93as well when considering the COVID (+) or normal patients with the support value of 174. The VGG 16-based approach acquired a precision of 0.95, with a Recall of 0.95, and F1 Score of 0.95 when considering COVID (-) or normal patients with the support value of 155, while acquired a precision of 0.95, with Recall of 0.96, and F1 Score of 0.96 as well, when considering COVID (+) or normal patients with the support value of 168. Thus, based on these parametric values, it is clear that the VGG 16-based approach is performing better as compared to the MobilenetV2-based approach during testing.

Table 6 is used to compare the accuracy of our MobilenetV2-based approach and VGG 16-based approach for 323 support values during testing these approaches.

Table 6 Testing Accuracy Comparison of both proposed approaches

Approach	Accuracy	Support
MobilenetV2 Based	0.93	323
VGG 16 Based	0.95	323

The accuracy of our MobilenetV2-based approach is 0.93, and the VGG 16-based approach is 0.95 for 323 support values during testing these approaches. This accuracy comparison indicates that the VGG 16-based model is better compared to the MobilenetV2-based model, as it is predicting more accurately.

5.3 Performance Measurement

The Performance of the proposed models is calculated and compared based on the mentioned parameters. We have taken various numbers of different parameters so that we can check the feasibility on various aspects. These parameters are described and defined below:

- Accuracy:

  Accuracy is used to determine how accurately our model is predicting the given input data values.

It can be formalized as:

$$\text{Accuracy} = \frac{(TN+TP)}{(TN+TP+FN+FP)} \quad (1)$$

- Error Rate:

  Error Rate is used to determine how wrong our model is in predicting the given input data values. Mathematically, it can be formalized as:

$$\text{Err Rate} = \frac{(FP+FN)}{(TN+TP+FN+FP)} \quad (2)$$

- Null Error Rate:

  It can be obtained by dividing the wrong predictions by all predictions.

$$\text{Null Err Rate} = \frac{(FP+TN)}{(TN+TP+FN+FP)} \quad (3)$$

- Predicted Positive Value or Precision:

  It can be obtained by dividing the True Positive occurrences from all the Positive occurrences

$$\text{Predicted Val}(+) = \frac{TP}{(TP+FP)} \quad (4)$$

- Predicted Negative Value:

  It can be obtained by dividing the True Negative occurrences from all the possible Negative occurrences (True Negative and False Negative).

$$\text{Predicted Val}(-) = \frac{TN}{(TN+FN)} \quad (5)$$

- True Positive Rate or Recall:

  It can be obtained by dividing the True Positive occurrences from the possible True Positive and False Negative occurrences.

$$\text{True Rate}(+)\ (\text{Recall}) = \frac{TP}{(TP+FN)} \quad (6)$$

- True Negative Rate or Specificity:

It can be obtained by dividing the True Negative occurrences from the possible True (-) and False (+) occurrences.

$$\text{True Rate(-)} = \frac{\text{TN}}{(\text{TN+FP})} \tag{7}$$

- Balanced Accuracy:

It can be obtained by finding the average of the TPR and TNR.

$$\text{Balanced Accuracy} = \frac{(\text{True Rate(+)} + \text{True Rate(-)})}{2} \tag{8}$$

- False Positive Rate:

It can be obtained by subtracting the TNR from one.

$$\text{False Rate(+)} = 1 - \text{True Rate(-)} \tag{9}$$

- False Negative Rate:

It can be obtained by subtracting the TPR from one.

$$\text{False Rate(-)} = 1 - \text{True Rate(+)} \tag{10}$$

- False Discovery Rate:

It can be obtained by subtracting the PPV from one.

$$\text{False Discovery Rate} = 1 - \text{Predicted Val(+)} \tag{11}$$

- F1- Score:

It can be obtained using the following formulae:

$$\text{F1- Score} = 2 * \frac{(\text{Predicted Val(+)} * \text{True Rate(+)})}{(\text{Predicted Val(+)} + \text{True Rate(+)})} \tag{12}$$

- False Omission Rate:

It can be obtained by dividing the False Negative occurrences by all the possible Negative occurrences (True Negative and False Negative).

$$\text{FOR} = \frac{\text{FN}}{(\text{TN+FN})} \tag{13}$$

- Prevalence:

Prevalence, also known as variability, is used to find out the number of positive items within a dataset. Its usefulness is even greater in binary classification because it can be used to estimate the proportion of something that is related to positive items.

$$\text{Prevalence} = \frac{(\text{FN+TP})}{(\text{TN+TP+FN+FP})} \tag{14}$$

- Positive LR:

It is the ratio of TPR and FPR.

$$\text{LR(+)} = \frac{\text{True Rate(+)}}{\text{False Rate(+)}} \tag{15}$$

- Negative LR:

It the ratio of FNR and TNR.

$$\text{LR(-)} = \frac{\text{False Rate(-)}}{\text{True Rate(-)}} \tag{16}$$

- Diagnostic Ratio:

It is the ratio of PLR and NLR.

$$DR = \frac{LR(+)}{LR(-)} \quad (17)$$

- FM Index:

It can be obtained using the below formulae:

$$FM\ Index = \sqrt{Predicted\ Val(+) * True\ Rate(+)} \quad (18)$$

- CS Index:

The Critical Success Index provides the correct predicted positive occurrence. It can be obtained using the following formulae:

$$CS\ Index = \frac{TP}{(TP+FN+FP)} \quad (19)$$

Table 7. Performance Comparison of Proposed Models on Various Performance Parameters

Performance Parameter	Computed Value (Approach 1)	Computer Value (Approach 2)
Accuracy	0.929	0.954
Error Rate	0.071	0.046
Null Error Rate	0.461	0.479
Positive Predicted Val	0.949	0.953
Negative Predicted Val	0.909	0.955
TPR	0.922	0.958
TNR	0.939	0.948
Balanced Accuracy	0.931	0.953
FPR	0.061	0.052
FNR	0.778	0.417
FDR	0.515	0.474
F1 - Score	0.935	0.955
FOR	0.090	0.045
Prevalence	0.557	0.520
PLR	15.24	18.52
NLR	0.082	0.043
Diagnostic Ratio	186	430
FM Index	0.935	0.955
Critical Success Index	0.878	0.915

The MobilenetV2-based approach acquired the values as 0.929, 0.071, 0.461, 0.949, 0.909,0.922,0.939, 0.931, 0.061, 0.778, 0.515, 0.935, 0.090, 0.557, 15.24, 0.082, 186, 0.935 and 0.878 for Accuracy, Error Rate, Null Error Rate, Positive Predicted Values, Negative Predicted Values, True Positive Rate, True Negative Rate, Balance Accuracy, False Positive Rate, False Negative Rate, False Discovery Rate, F1- Score, False Omission Rate, Prevalence, PLR, NLR, Diagnostic Ratio, FM Index and Critical Success

Index, respectively. The VGG 16-based approach acquired the values as 0.954, 0.046, 0.479, 0.953, 0.955,0.958,0.948, 0.953, 0.052, 0.417, 0.474, 0.955, 0.045, 0.520, 18.52, 0.043, 430, 0.955 and 0.915 for Accuracy, Error Rate, Null Error Rate, Positive Predicted Values, Negative Predicted Values, True Positive Rate, True Negative Rate, Balance Accuracy, False Positive Rate, False Negative Rate, False Discovery Rate, F1-Score, FOR, Prevalence, PLR, NLR, DR, FM Index and Critical Success Index, respectively. From the calculated values of these parameters, it is clear that the VGG 16-based approach performed better in every case.

5.4 Discussion

Here, we proposed two approaches on a dataset created by combining two collections of Chest CT scan images, featuring both COVID-19-infected individuals and healthy persons. In the first approach, we used MobileNetV2 CNN. The second approach is based on VGG 16 CNN. While implementing the first approach, we felt that its performance could be improved, which led us to develop the second approach, replacing MobileNetV2 with VGG-16. To demonstrate our results, we compared the performance of both approaches using common metrics like accuracy, precision, F1 score, recall, etc. We observed that in both training and testing phases, the second approach (based on VGG-16) outperformed the first.

But we do not limit our reasoning to just these performance parameters; but we also compared both models based on many other performance parameters like Error Rate, Null Error Rate, Positive Predicted Values, Negative Predicted Values, PLR, NLR, Diagnostic Ratio, FM Index, and Critical Success Index. After confirming that the VGG 16-based approach is better than the MobileNetV2-based approach, we compared our VGG 16-based approach with existing methods based on various performance metrics. The proposed approach has the highest recall, at 98%, while the highest recall among other approaches was 94%, which is lower than our model. Our proposed approach has the highest accuracy, at 95%, while the maximum accuracy among existing methods was 98.8 (Table 6.5), which exceeds our model's accuracy. Our proposed model achieved the highest precision, at 95%, while the highest precision among other approaches was 93.6 (Table 6.7), which is lower than our model. Our proposed model achieved the highest F1-score, at 96%, compared to the maximum recall of 94% by other approaches, showing our model's superior performance. Based on these results and the performance demonstrated by the proposed approach, it is clear that the VGG 16-based model outperforms the MobileNetV2-based model. Additionally, the comparative results with other existing models indicate that our model is better.

6 Conclusion

Here, we propose two deep transfer approaches for classifying COVID-19, based on a dataset prepared by combining two collections of Chest CT from both COVID-19-infected individuals and healthy individuals. In the first approach, we utilize MobileNetv2 CNN, while the second approach relies on the VGG 16 CNN. During the implementation of the first approach, we found that the model's performance could be improved, which led us to develop the second approach, where we substituted MobileNetV2 with VGG 16. To support our claim, we compared the performance of both approaches using popular parameters and others. We observed that in both training and testing cases, the performance of the second approach (based on VGG 16) is superior to that of the first. However, we did not limit our analysis to these performance parameters; instead, we also evaluated both models based on several additional metrics, including Error Rate, Null Error Rate, Positive Predicted Values, Negative Predicted Values, PLR, NLR, Diagnostic Ratio, FM Index, and Critical Success Index. Our findings indicate that the VGG 16-based model is better than the MobileNetV2-based model.

With these results and the performance shown by our model, we can conclude that the VGG 16-based model is better than that of our first proposed model (MobileNetV2-based). In addition, the comparative results with other existing models show that our model performed better in terms of many performance parameters, while it needs to improve its performance on accuracy metrics. In the future, we will work to enhance the accuracy of our model using some advanced augmentation techniques and larger datasets, and add some new transformers to produce some kind of hybrid model.

References

1. Wang, J., Chen, Y. Deep Transfer Learning. In: Introduction to Transfer Learning. Machine Learning: Foundations, Methodologies, and Applications. Springer, Singapore. https://doi.org/10.1007/978-981-19-7584-4_9 (2023)
2. Huang, X., Jin, G., Ruan, W. Machine Learning Basics. In: Machine Learning Safety. Artificial Intelligence: Foundations, Theory, and Algorithms. Springer, Singapore. https://doi.org/10.1007/978-981-19-6814-3_1 (2023)
3. Gupta, P., Sehgal, N.K., Acken, J.M. Machine Learning Algorithms. In: Introduction to Machine Learning with Security. Synthesis Lectures on Engineering, Science, and Technology. Springer, Cham. https://doi.org/10.1007/978-3-031-59170-9_2 (2025)
4. Mokni, R., Haoues, M. CADNet157 model: fine-tuned ResNet152 model for breast cancer diagnosis from mammography images. *Neural Comput & Applic* **34**, 22023–22046. https://doi.org/10.1007/s00521-022-07648-w (2022)
5. Chakraborty, M. Rule extraction from convolutional neural networks for heart disease prediction. *Biomed. Eng. Lett.* **14**, 649–661. https://doi.org/10.1007/s13534-024-00358-3 (2024)
6. Van Dieren, L., Amar, J.Z., Geurs, N. *et al.* Unveiling the power of convolutional neural networks in melanoma diagnosis. *Eur J Dermatol* **33**, 495–505. https://doi.org/10.1684/ejd.2023.4559 (2023)

7. Das, S., Goswami, R.S. Addressing challenges in accurate brain tumor classification in MRI: a transfer learning approach with EfficientNetB3 and comprehensive model evaluation. *Multimed Tools Appl* (2024). https://doi.org/10.1007/s11042-024-20366-w
8. Kulshreshtha, A., Nagpal, A. IFAS: improved fully automatic segmentation convolutional neural network model along with morphological segmentation for brain tumor detection. *Int. j. inf. tecnol.* **16**, 1517–1525 (2024). https://doi.org/10.1007/s41870-023-01572-5
9. Montaha, S., Azam, S., Rakibul Haque Rafid, A.K.M. *et al.* Brain Tumor Segmentation from 3D MRI Scans Using U-Net. *SN COMPUT. SCI.* **4**, 386 (2023). https://doi.org/10.1007/s42979-023-01854-6
10. Salama, G.M., Mohamed, A. & Abd-Ellah, M.K. Machine learning and deep learning COVID-19 diagnosis system: key achievements, lessons learned, and a transfer learning algorithm. *Soft Comput* **28**, 13715–13742 (2024). https://doi.org/10.1007/s00500-024-10362-5
11. Cordelli, E., Soda, P., Citter, S. *et al.* Machine learning predicts pulmonary Long Covid sequelae using clinical data. *BMC Med Inform Decis Mak* **24**, 359 (2024). https://doi.org/10.1186/s12911-024-02745-3
12. Chaity, T.T., Khan, M.A.R. & Rahman, F. Prediction of impacts and outbreak of COVID-19 on the society using distinct machine learning algorithms. *Iran J Comput Sci* **7**, 155–175 (2024). https://doi.org/10.1007/s42044-023-00166-5
13. Nurjahan, Mahbub-Or-Rashid, M., Satu, M.S. *et al.* Machine learning and deep learning algorithms in detecting COVID-19 utilizing medical images: a comprehensive review. *Iran J Comput Sci* **7**, 699–721 (2024). https://doi.org/10.1007/s42044-024-00190-z
14. Rabby, G., D'Souza, J., Oelen, A. *et al.* Impact of COVID-19 research: a study on predicting influential scholarly documents using machine learning and a domain-independent knowledge graph. *J Biomed Semant* **14**, 18 (2023). https://doi.org/10.1186/s13326-023-00298-4
15. Rabby, G., Berka, P. Multi-class classification of COVID-19 documents using machine learning algorithms. *J Intell Inf Syst* **60**, 571–591 (2023). https://doi.org/10.1007/s10844-022-00768-8
16. Nayak, S.S., Darji, A.D. & Shah, P.K. Machine learning approach for detecting COVID-19 from speech signal using Mel frequency magnitude coefficient. *SIViP* **17**, 3155–3162 (2023). https://doi.org/10.1007/s11760-023-02537-8
17. Nramban Kannan, S.K., Kolla, B.P., Sengan, S. *et al.* Analysis of COVID-19 Datasets Using Statistical Modelling and Machine Learning Techniques to Predict the Disease. *SN COMPUT. SCI.* **5**, 181 (2024). https://doi.org/10.1007/s42979-023-02464-y
18. Ismaeel, N.Q., Mohammed, H.J., Chaloob, I.Z. *et al.* Application of Healthcare Management Technologies for the COVID-19 Pandemic Using Internet of Things and Machine Learning Algorithms. *Wireless Pers Commun* (2023). https://doi.org/10.1007/s11277-023-10663-2
19. Hardy, F., Heyl, J., Tucker, K. *et al.* Estimating nosocomial infection and its outcomes in hospital patients in England with a diagnosis of COVID-19 using machine learning. *Int J Data Sci Anal* (2023). https://doi.org/10.1007/s41060-023-00419-3
20. Rezapour, M., Niazi, M.K.K. & Gurcan, M.N. Machine learning-based analytics of the impact of the Covid-19 pandemic on alcohol consumption habit changes among United States healthcare workers. *Sci Rep* **13**, 6003 (2023). https://doi.org/10.1038/s41598-023-33222-y.