A Hybrid Model for Covid-19 Detection using CT-Scans

Nagwa G. Ali[1], Fahad K. El Sheref[2], Mahmoud M. El khouly[3]
Department of Information Technology-Faculty of Computers and Artificial Intelligence, Helwan University, Cairo, Egypt[1, 3]
Department of Information System-Faculty of Computers and Artificial Intelligence, Beni-Suef University, Egypt[2]

***Abstract*—Although some believe it has been wiped out, the coronavirus is striking again. Controlling this epidemic necessitates early detection of coronavirus disease. Computed tomography (CT) scan images allow fast and accurate screening for COVID-19. This study seeks to develop the most precise model for identifying and classifying COVID-19 by developing an automated approach using transfer-learning CNN models as a base. Transfer learning models like VGG16, Resnet50, and Xception are employed in this study. The VGG16 has a 98.39% accuracy, the Resnet50 has a 97.27% accuracy, and the Xception has a 96.6% accuracy; after that, a hybrid model made using the stacking ensemble method has an accuracy of 98.71%. According to the findings, hybrid architecture offers greater accuracy than a single architecture.**

***Keywords*—Covid-19; coronavirus; transfer-learning; CT-scan and ensemble method**

I. Introduction

COVID-19 is caused by the virus SARS-CoV-2. Lung infections can cause anything from a simple cold to potentially fatal illnesses. Symptoms of the respiratory system are usually seen with coronavirus infections. In addition, people can occasionally get minor, self-limiting ailments like the flu that have detrimental effects. Fever, a cough, and trouble breathing might be caused by respiratory issues, exhaustion, and a sore throat [1]-[3].

Huge losses have been caused by the global spread of the COVID-19 pandemic. The urgent issue that health care and medical organizations are addressing is the quick detection of COVID-19 [1]. COVID-19 presents a serious threat to public health and the country's social and economic life because of the quick rise in the number of newly verified and reported cases.

An automated system that can measure the patient's infected zone and assess the progress of infected people by CT scan image analysis and clinical diagnosis is required. Even for skilled medical professionals, diagnosing COVID-19 is challenging [4], [5].

The two most common methods for diagnosing COVID-19 are computed tomography and X-rays. Medical health guidelines advocate chest imaging as a quick and efficient treatment, and it has been acknowledged as the first instrument in viral screening in a series of papers. Chest X-rays and CT scans have promising results despite frequently being examined by qualified radiologists. Computer-aided diagnosis (CAD) is required to reduce error rates while saving time and money [6] since radiologists see many patients daily and the diagnostic procedure takes a long time.

CT scans provide a series of slices of a specific location without including the numerous physiological features, in contrast to traditional X-rays. Compared to traditional X-rays, CT scans offer a considerably more thorough picture of the diagnosis [7].

Deep learning has been used in several studies to analyze radiological images. They were created to address the drawbacks of COVID-19 medical procedures that depend on radiological imaging. CNN is the method that is most effective for identifying the most significant deep learning algorithms. Consequently, the data processing field has shown a lot of interest in deep learning algorithms, especially at CNN [8].

This paper proposes a highly accurate automated computer-aided diagnostic approach for COVID-19 classification. In order to create a hybrid model using stacking ensemble with improved performance on the SARS-COV-2 dataset, a study was done to examine the implementation of three pre-trained convolutional neural network (CNN) models based on transfer learning. This work's main novelty and contributions are as follows:

- Reducing the number of false negative and positive values in the modelling process.
- The base and hybrid models were trained for 100 epochs.
- Comparing the recommended model to that of current DL frameworks in order to determine its effectiveness.
- Comparing the hybrid model's performance to the base-line models.
- Applying the stacking ensemble method in a new way instead of the familiar ways of stacking, which saves more time.

The following parts are organized as follows: The state-of-the-art studies are summarized in Section II, and resources and methods are covered in Section III. Results and analysis are presented in Section IV. The summary and recommendations for future work are presented in Section V.

II. Related Work

Many studies were conducted to diagnose COVID-19 either using CT scan images or using x-rays. This section summarizes the most recent studies, focusing on CT scan images.

Kogilavani, S. V., et al. [8] proposed using many CNN architectures, such as MobileNet, DeseNet121, VGG16, NASNet, Xception, and EfficientNet. The accuracy rates were 96.38%, 97.53%, 97.68%, 89.51%, 92.47%, and 80.19%, respectively.

Ebenezer Jangam et al. [9] developed a stacked ensemble method using VGG-19, ResNet-101, DenseNet-169, and WideResNet-50-2 as four pre-trained computer vision models. With the help of three separate chest CT scans, the suggested model was trained and assessed. The suggested technique has an accuracy of 0.8473, 0.99, and 0.935 on these datasets.

Rohit Kundu et al. [10] presented a bagging ensemble method of three transfer learning models, including Inception v3, ResNet 34, and DenseNet 201, which has been utilized to improve the performance of the individual models. The proposed model had a 97.81% accuracy rate.

A deep learning model, called truncated VGG16, was created by Mukul Singh et al. [11] to screen COVID-19 CT scans. For feature extraction from CT scan pictures, the VGG16 architecture has been fine-tuned; principal component analysis (PCA) is also used to choose features. For the final classification, four distinct classifiers are compared. On 208 tested images, the best classifier was a bagging ensemble with SVM, which obtained an F1 score of 95.3%, an accuracy of 95.7%, a precision of 95.8%, and an area under the curve (AUC) of 0.958.

For the diagnosis of COVID-19 from CT scans, Bejoy Abraham et al. [12] suggested using the kernel support vector machine with features extracted from five distinct CNNs (MobilenetV2, Darknet53, Shufflenet, Xception, and EfficientnetB0). The technique achieved 0.916 accuracy, a 0.91 F-score, and 0.917 sensitivity.

Nirmala Devi Kathamuthu, et al. [13] suggested applying a variety of foundation models, including VGG16, VGG19, Inception V3, Densenet 121, Xception, and Resnet50; the best model of these models attained an accuracy of 98% using ct-scan images. The difference between this model and ours is the data augmentation technique used, the training parameters, and the fact that we finally built a hybrid model using the stacking ensemble method.

Gifani et al. [14] employed fifteen pre-trained convolutional neural networks (CNNs) architectures: EfficientNets (B0-B5), NasNetLarge, NasNetMobile, ResNet-50, SeResnet 50, Xception, DenseNet121, ResNext50, and Inception ResNet v2. Thus, to further improve recognition performance, they created an ensemble approach based on majority voting on the ideal concoction of deep transfer learning outputs. The experimental results show that the majority voting of five deep transfer learning models with EfficientNetB0, EfficientNetB3, EfficientNet5, Inception ResNet v2, and Xception outperform individual transfer learning structures in terms of precision (0.857), recall (0.854), and accuracy (0.85) metrics for diagnosing COVID-19 from CT scans.

Horry, Michael J., et al. [15] developed a conceptual transfer learning framework to support COVID-19 detection with image categorization utilising deep learning models for X-ray, ultrasound, and CT scans, including Resnet50, Inception V3, Xception, VGG16/VGG19, InceptionResNet, DenseNet, and NASNetLarge. They settled on the VGG19 model, which was then adjusted with the proper parameters to reach extremely high levels of COVID-19 detection versus pneumonia and normal in all three types of lung images, with a precision of up to 86% for X-rays, 100% for ultrasounds, and 84% for CT scans.

The authors of [16] propose an augmented CNN to identify COVID-19 on CT scan and X-ray chest images and to differentiate COVID-19 patients from non-COVID-19 cases. While using these augmented images to train CNN, a classification accuracy of 98.97 percent for X-ray images and 95.38 percent for CT scan images was achieved.

Tanvir Mahmud et al. [17] presented the hybrid neural network CovTANet for the early prediction and diagnosis of COVID-19 using CT scans. A segmentation network was used to predict the lesions with a 95.8% accuracy rate.

A transfer learning strategy was employed by Chun Li et al. [18] to propose a way of training the model using a few CT images. The evaluation accuracy of this method was 87.6% for COVID-19 severity using pre-trained ChexNet.

In order to detect COVID-19 from a CT scan, Varan Singh Rohila et al. [19] offer a DCNN model known as ReCOV-101 that makes use of ResNet-101 as its basis. To increase the dataset, data augmentation, transfer learning, and the 94.9 percent accurate "skip connection" method are employed. From the previous review, further work needs to be done to increase the reliability and accuracy of COVID-19 detection using CT scan images.

Limitations of the previous studies include:

- The difficulty of sharing medical data is data privacy.
- The datasets that are currently accessible are incorrect, unclear, noisy, and incomplete.
- Most studies in the literature are conducted using datasets from different internet sources.

III. METHOD

The method used to achieve the objectives of the study is briefly described in this section. The diagram of the suggested method is shown in Fig. 1. Each model was trained individually on the training set; after that, the models were concatenated using stacking ensemble learning and fine-tuned using the training set. Fig. 3 and Fig. 4, which are subfigures of Fig. 1, explain the proposed model in more detail.

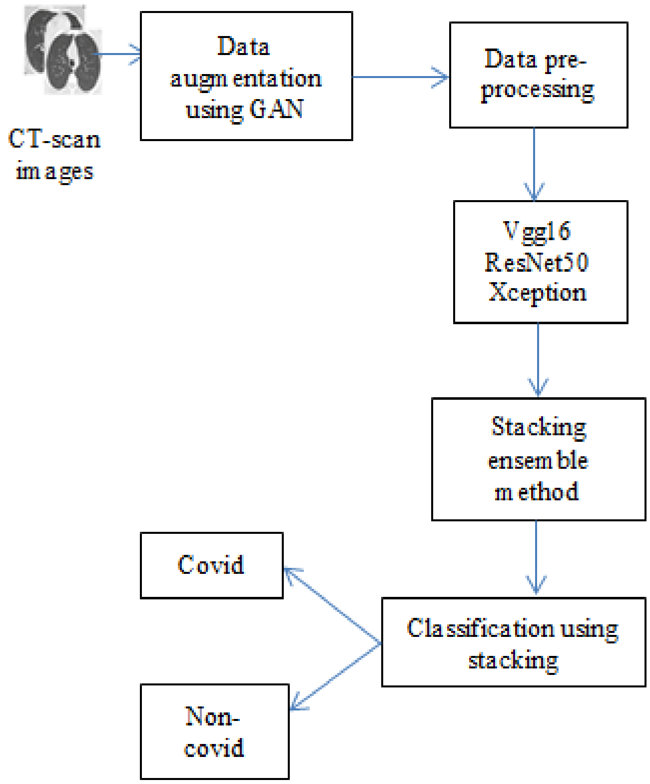


Fig. 1. Proposed model architecture.

A. Dataset

The SARS-CoV-2 CT scan dataset [20] was used in the proposed work to identify COVID-19 cases. This dataset is the most global because it contains more CT scan images that are clear and free of noise, and it is also used in more studies. The dataset consists of 2482 CT scan images, 1252 positive for COVID-19 (+ve) and 1230 negative for COVID-19 (-ve). These data were compiled from information provided by actual patients who went to hospitals in Sao Paulo, Brazil. This dataset aims to advance research and develop artificial intelligence (AI) tools that can detect SARS-CoV-2 or COVID-19 infection by looking at CT images.

B. Data Augmentation

A form of neural network design called a generative adversarial network (GAN) has much potential for artificial intelligence. The min-max two-player zero-sum game is an important factor in GAN. In this game one player profited by the same amount from the other's loss. The actors in this scenario are two separate GAN networks known as the discriminator and generator.

The primary goal of the discriminator, indicated as D, is to determine whether a sample is real or fake[21]. A fake sample of an image is produced by the generator, known as G, in contrast, in order to deceive the discriminator.

The discriminator determines the probability that a particular sample is real. The probability value will probably be greater for a real sample. A probability value that is very close to 0 indicates fake samples. When the discriminator is no longer able to distinguish between a real and fake sample and the probability value is close to 0.5, the generator can have an ideal answer [21]**.** Fig. 2 shows a sample of COVID and non-COVID images generated by GAN.

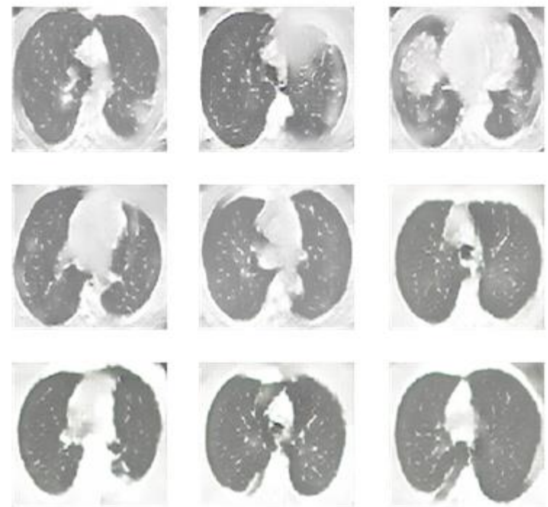


Fig. 2. Sample of COVID and non-COVID images after augmentation.

C. Data Pre-processing

Pre-processing has been done on the CT scan images. The image must fit the input size required to train the deep learning algorithm and make predictions using the data. To test the input size of the network, the data is rescaled. The image size in the proposed system has a dimension of 224 x 224. The data is consequently rescaled to match the supplied dimensions. These images were divided into two groups with an 80/20 split: a training set and a testing set. Table I shows the dataset after augmentation and preprocessing.

TABLE I. SARS COV-2 DATASET AFTER PREPROCESSING

Dataset	Covid	Non-Covid	Total
Training	1239	1239	2478
Testing	311	310	621
Total	1550	1549	3099

D. Classification Phase with Different Deep Learning Models

In order to classify the CT scan images, Xception[22], Resnet 50[23] ,and VGG 16 [22] are used as examples of deep learning architectures. These models are trained through transfer learning. The stacked model is then used to integrate the earlier models, which were each trained for 100 epochs, into one. Complete details are provided below.

E. Transfer Learning

Transfer learning using ImageNet was used to solve the problem of insufficient data. For each model, the weights from the ImageNet training were downloaded. The applied layer training procedure used the feature maps as input. To prevent losing any data during subsequent training rounds, freeze the layers of the model that have already been trained. A fully connected deep neural network was fed with the most recent feature map, which was flattened. The other layers were trained to extract more data from the later convolution layers because they were closer to the output features. As illustrated in Fig. 3, we added three additional layers to the top of each model: namely, dense with an output of 512, a dropout layer, and another dense layer with a sigmoid classifier. For each of the neural networks used in this study, a dropout layer of 0.25 was added to avoid overlapping [24]. The network was trained with a sigmoid classifier using an Adam optimizer for 100 iterations, with a batch size of 32 and a learning rate of 0.00001.

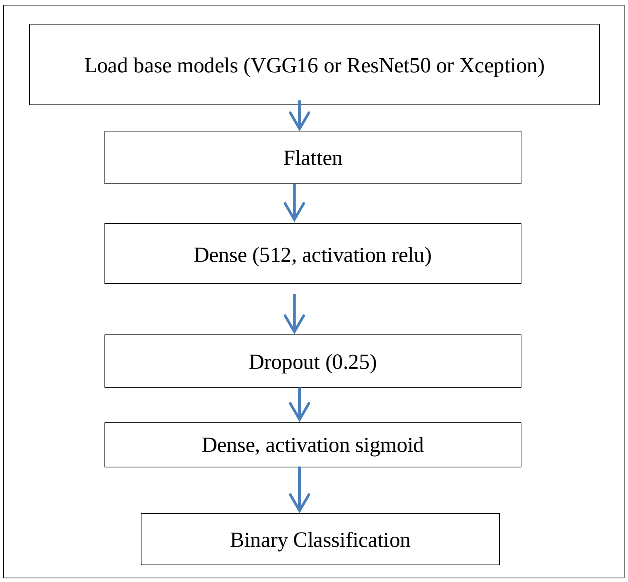


Fig. 3. Proposed transfer learning base models.

F. Stacking Ensemble Method

Ensemble approaches come in a variety of forms, including average, weighted average, boosting, bagging, and stacking. In this paper, stacked generalization is used, which is an ensemble technique that learns how to integrate the predictions from various existing models in the most effective way [25].

This study employs a neural network as a meta-learner when neural networks are used as sub-models. In more detail, the sub-networks can be incorporated into a more extensive multi-headed neural network, which will then figure out how to integrate the predictions from each input sub-model most effectively. It makes it possible to think of the stacking ensemble as a single, enormous model.

All of the loaded models are utilized as a distinct input to the bigger stacking ensemble model, which can be defined just after models are loaded as a list. To prevent the weights from being modified while the new, larger model is being trained, each of the loaded models must have all of its layers marked as not trainable.

This method has the advantage of immediately giving the meta-learner the outputs of the sub-models. This new model will employ a different input head for each input layer from each sub-model. This indicates that any input data must be delivered to the model in multiples of k, where k is the total number of input models, which in this case are 3.

The results from all the models can then be combined. Here, a straightforward concatenation merge was employed. Then, a hidden layer is built to interpret this "input" for the meta-learner, and an output layer is defined to produce its own probabilistic prediction. The base model's training parameters were used to train the stack model as well. Fig. 4 shows the architecture of the stacked generalization.

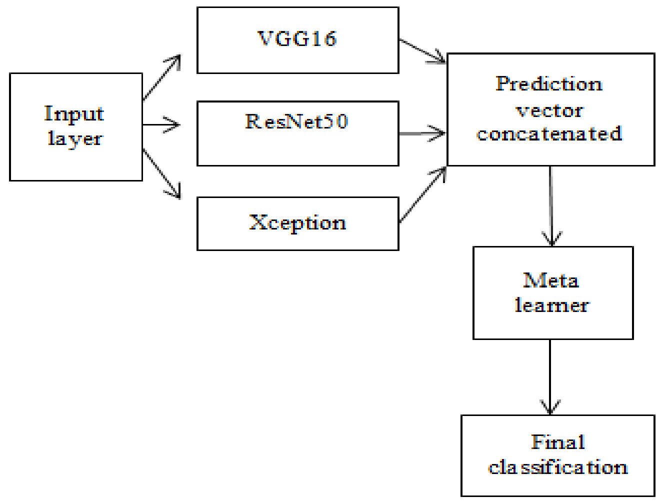


Fig. 4. Stacked generalization architecture.

IV. Results

This section discusses the results of both base and hybrid models using different evaluation measures.

A. Performance Evaluation Measures

A number of metrics, such as accuracy, precision, recall, and F1-score, can be used to evaluate the model's performance.

Accuracy: The ratio of correctly anticipated observations to all observations is the easiest and most obvious performance statistic. Given in the equation below:

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

where

- True positives (TP) are instances in which we made a prediction that someone had the disease and they actually did.
- True Negatives (TN): As expected, they are free of the illness.
- False positives (FP): Although we expected them to have the illness, they don't.
- False negatives (FN): Despite our prediction, they are infected.

The ratio of correctly predicted positive observations to all positively expected observations is known as precision. Given in the equation below:

$$\text{Precision} = TP/TP+FP \tag{2}$$

Recall is the percentage of accurately predicted positive observations among all observations in the current class. This is known as sensitivity. Given in the equation below:

$$\text{Recall} = TP/TP+FN \quad (3)$$

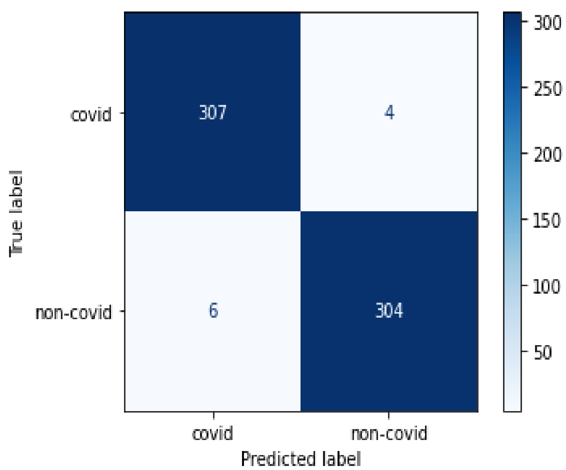
The mean of recall and precision is known as the F1 score. It is provided in the equation below:

$$\text{F1Score}=2*(\text{Recall}*\text{Precision})/(\text{Recall}+\text{Precision}) \quad (4)$$

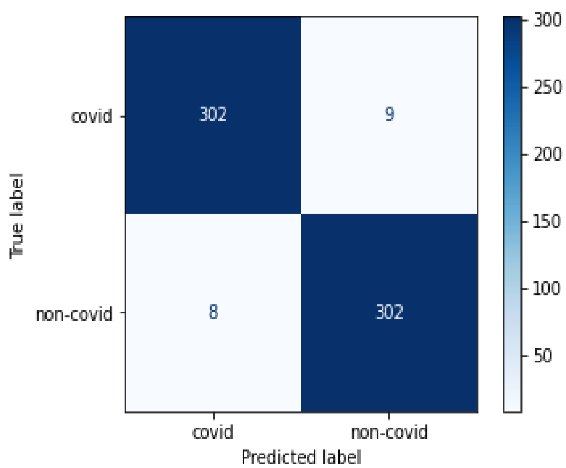
Table II displays the three used models and the hybrid model evaluation measure, while Fig. 5 displays the confusion matrix of the base and hybrid models.

TABLE II. EVALUATION MEASURES OF THE MODELS

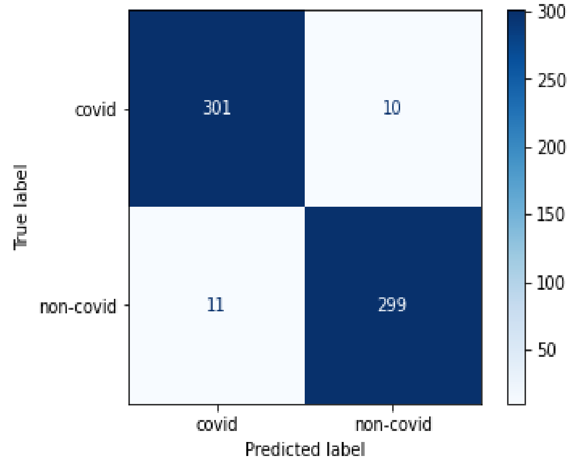
Model	**Accuracy (%)**	**Precision (%)**	**Recall (%)**	**F1-score (%)**
VGG16	98.39	98.7	98.06	98.39
Resnet50	97.26	97.1	97.4	97.26
Xception	96.62	96.76	96.45	96.6
Hybrid	98.71	98.39	99.0	98.7



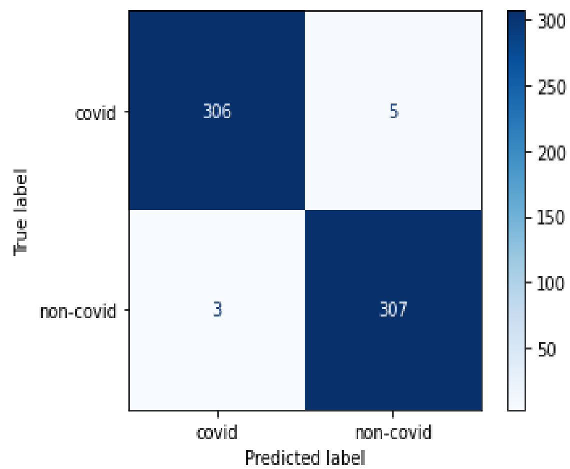
(a) VGG16 predicted 307 as COVID and 304 as non-COVID out of 621 images.



(b) Resnet50 predicted 302 as COVID and 302 as non-COVID out of 621 images.



(c) Xception predicted 301 as COVID and 299 as non-COVID out of 621 images.



(d) The stacked method predicted 306 as COVID and 307 as non-COVID out of 621 images.

Fig. 5. Base and hybrid model confusion matrix.

As shown in Table II and Fig. 5, the hybrid model achieved better results than the single model in terms of accuracy, recall, and f1-score.

V. DISCUSSION

Generally, the outcome is based on testing data, which is comprised of 20% of the entire images from the sars-cov2 dataset. Using the Keras library, the deep model employed in this study was developed. Keras is compatible with deep learning libraries such as TensorFlow and Theano. The model was developed using Google Colab, which enables you to create and run Python code in your browser and offers free access to the GPU.

A. Comparison

The suggested model is compared with other research done using the same dataset (Table III). This comparison shows that although many studies in the literature have combined more deep learning models, these studies are weak compared to ours. The best advantage of our proposed model is that we used GAN as an augmentation technique rather than traditional methods for increasing the number of images used to train the models, trained the model for 100 epochs with the suggested training parameters discussed in Fig. 3, and eliminated overfitting by applying the stacking ensemble method.

TABLE III. COMPARISON

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
VGG16 [8]	97.68	_____	______	____
VGG19+DenseNet169+ResNet101[9]	94	90	98	94
Inceptionv3+Resnet34+DesNet201[10]	97.81	97.77	97.81	97.77
Five pre-trained CNN models +KSVM[21]	91.6	____	91.7	91
VGG16[22]	98	97.99	97.99	97.90
Proposed model	98.71	98.39	99.0	98.7

VI. CONCLUSION AND FUTURE WORK

In this proposed work, greater transfer learning models such as VGG16, Resnet50, and Xception are trained individually on the training data before being combined to enhance the model's performance and prevent over fitting. Kaggle's collection of CT scan images contains fewer images than necessary, so data augmentation using GAN is used to obtain more images. Performance metrics, including precision, F1-score, and recall, are used to assess the model's performance. The stacked model outperforms the single model in terms of accuracy, coming in at 98.71. By comparing the proposed method with other studies applied to the same dataset, it is obvious that the proposed method gives high recall, accuracy, precision, and the f1-score.

The future work of this study will train the suggested model for more epochs using different preprocessing techniques, apply the suggested model to another dataset, and attempt to use a different optimizer.

Data Availability: The data that support the findings of this study are openly available upon request from the authors.

Funding: The authors received no specific funding for this work.

REFERENCES

[1] M. Sandeep Kumar et al., "Social economic impact of COVID-19 outbreak in India," Int. J. Pervasive Comput. Commun., vol. 16, no. 4, pp. 309–319, 2020, doi: 10.1108/IJPCC-06-2020-0053.

[2] M. Venkatasen et al., "Forecasting of the SARS-CoV-2 epidemic in India using SIR model, flatten curve and herd immunity," J. Ambient Intell. Humaniz. Comput., no. 0123456789, 2020, doi: 10.1007/s12652-020-02641-4.

[3] K. Banu Priya et al., "Pediatric and geriatric immunity network mobile computational model for COVID-19," Int. J. Pervasive Comput. Commun., vol. 16, no. 4, pp. 321–330, 2020, doi: 10.1108/IJPCC-06-2020-0054.

[4] S. K. Mathivanan et al., "Adoption of E-Learning during Lockdown in India," Int. J. Syst. Assur. Eng. Manag., p. 13198, 2021, doi: 10.1007/s13198-021-01072-4.

[5] S. Rajendran and P. Jayagopal, "Accessing Covid19 epidemic outbreak in Tamilnadu and the impact of lockdown through epidemiological models and dynamic systems," Meas. J. Int. Meas. Confed., vol. 169, no. August 2020, p. 108432, 2021, doi: 10.1016/j.measurement.2020.108432.

[6] S. Hasija, P. Akash, M. Bhargav Hemanth, A. Kumar, and S. Sharma, "A novel approach for detection of COVID-19 and Pneumonia using only binary classification from chest CT-scans," Neurosci. Informatics, vol. 2, no. 4, p. 100069, 2022, doi: 10.1016/j.neuri.2022.100069.

[7] P. Silva et al., "COVID-19 detection in CT images with deep learning: A voting-based scheme and cross-datasets analysis," Informatics Med. Unlocked, vol. 20, p. 100427, 2020, doi: 10.1016/j.imu.2020.100427.

[8] S. V. Kogilavani et al., "COVID-19 Detection Based on Lung Ct Scan Using Deep Learning Techniques," Comput. Math. Methods Med., vol. 2022, 2022, doi: 10.1155/2022/7672196.

[9] E. Jangam and C. S. R. Annavarapu, "A stacked ensemble for the detection of COVID-19 with high recall and accuracy," Comput. Biol. Med., vol. 135, no. June, p. 104608, 2021, doi: 10.1016/j.compbiomed.2021.104608.

[10] R. Kundu, P. K. Singh, M. Ferrara, A. Ahmadian, and R. Sarkar, "ET-NET: an ensemble of transfer learning models for prediction of COVID-19 infection through chest CT-scan images," Multimed. Tools Appl., vol. 81, no. 1, pp. 31–50, 2022, doi: 10.1007/s11042-021-11319-8.

[11] M. Singh, S. Bansal, S. Ahuja, R. K. Dubey, B. K. Panigrahi, and N. Dey, "Transfer learning–based ensemble support vector machine model for automated COVID-19 detection using lung computerized tomography scan data," Med. Biol. Eng. Comput., vol. 59, no. 4, pp. 825–839, 2021, doi: 10.1007/s11517-020-02299-2.

[12] B. Abraham and M. S. Nair, "Computer-aided detection of COVID-19 from CT scans using an ensemble of CNNs and KSVM classifier," Signal, Image Video Process., vol. 16, no. 3, pp. 587–594, 2022, doi: 10.1007/s11760-021-01991-6.

[13] N. D. Kathamuthu et al., "A deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications," Adv. Eng. Softw., vol. 175, no. October 2022, p. 103317, 2023, doi: 10.1016/j.advengsoft.2022.103317.

[14] P. gifani, A. Shalbaf, and M. Vafaeezadeh, "Automated detection of COVID-19 using ensemble of transfer learning with deep convolutional neural network based on CT scans," Int. J. Comput. Assist. Radiol. Surg., vol. 16, no. 1, pp. 115–123, 2021, doi: 10.1007/s11548-020-02286-w.

[15] M. J. Horry et al., "COVID-19 Detection through Transfer Learning Using Multimodal Imaging Data," IEEE Access, vol. 8, pp. 149808–149824, 2020, doi: 10.1109/ACCESS.2020.3016780.

[16] D. Giri, K. R. Choo, and S. Ponnusamy, of the Seventh International Conference on Mathematics and Computing. 2021.

[17] T. Mahmud et al., "CovTANet: A Hybrid Tri-Level Attention-Based Network for Lesion Segmentation, Diagnosis, and Severity Prediction of COVID-19 Chest CT Scans," IEEE Trans. Ind. Informatics, vol. 17, no. 9, pp. 6489–6498, 2021, doi: 10.1109/TII.2020.3048391.

[18] C. Li, Y. Yang, H. Liang, and B. Wu, "Transfer learning for establishment of recognition of COVID-19 on CT imaging using small-sized training datasets[Formula presented]," Knowledge-Based Syst., vol. 218, p. 106849, 2021, doi: 10.1016/j.knosys.2021.106849.

[19] V. S. Rohila, N. Gupta, A. Kaul, and D. K. Sharma, "Deep learning assisted COVID-19 detection using full CT-scans," Internet of Things (Netherlands), vol. 14, p. 100377, 2021, doi: 10.1016/j.iot.2021.100377.

[20] E. Soares, P. Angelov, S. Biaso, M. H. Froes, and D. K. Abe, "SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification," medRxiv, p. 2020.04.24.20078584, 2020, [Online]. Available: https://www.medrxiv.org/content/10.1101/2020.04.24.20078584v2%0Ahttps://www.medrxiv.org/content/10.1101/2020.04.24.20078584v2.abstract

[21] H. Alqahtani, M. Kavakli-Thorne, and G. Kumar, "Applications of Generative Adversarial Networks (GANs): An Updated Review," Arch. Comput. Methods Eng., vol. 28, no. 2, pp. 525–552, 2021, doi: 10.1007/s11831-019-09388-y.

[22] B. Nigam, A. Nigam, R. Jain, S. Dodia, N. Arora, and B. Annappa, "COVID-19: Automatic detection from X-ray images by utilizing deep learning methods," Expert Syst. Appl., vol. 176, no. January, p. 114883, 2021, doi: 10.1016/j.eswa.2021.114883.

[23] V. Sangeetha and K. J. R. Prasad, "Syntheses of novel derivatives of 2-acetylfuro[2,3-a]carbazoles, benzo[1,2-b]-1,4-thiazepino[2,3-a]carbazoles and 1-acetyloxycarbazole-2- carbaldehydes," Indian J. Chem. - Sect. B Org. Med. Chem., vol. 45, no. 8, pp. 1951–1954, 2006, doi: 10.1002/chin.200650130.

[24] G. E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, and R. R. Salakhutdinov, "Improving neural networks by preventing co-adaptation of feature detectors," pp. 1–18, 2012, [Online]. Available: http://arxiv.org/abs/1207.0580.

[25] A. I. Naimi and L. B. Balzer, "Stacked generalization: an introduction to super learning," Eur. J. Epidemiol., vol. 33, no. 5, pp. 459–464, 2018, doi: 10.1007/s10654-018-0390-z.