



Automatic COVID-19 detection from X-ray images using ensemble learning with convolutional neural network

Amit Kumar Das^{1,2} · Sayantani Ghosh¹ · Samiruddin Thunder² · Rohit Dutta² · Sachin Agarwal² · Amlan Chakrabarti¹

Received: 26 July 2020 / Accepted: 1 March 2021 / Published online: 19 March 2021
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

COVID-19 continues to have catastrophic effects on the lives of human beings throughout the world. To combat this disease it is necessary to screen the affected patients in a fast and inexpensive way. One of the most viable steps towards achieving this goal is through radiological examination, Chest X-Ray being the most easily available and least expensive option. In this paper, we have proposed a Deep Convolutional Neural Network-based solution which can detect the COVID-19 +ve patients using chest X-Ray images. Multiple state-of-the-art CNN models—DenseNet201, Resnet50V2 and Inceptionv3, have been adopted in the proposed work. They have been trained individually to make independent predictions. Then the models are combined, using a new method of weighted average ensembling technique, to predict a class value. To test the efficacy of the solution we have used publicly available chest X-ray images of COVID +ve and –ve cases. 538 images of COVID +ve patients and 468 images of COVID –ve patients have been divided into training, test and validation sets. The proposed approach gave a classification accuracy of 91.62% which is higher than the state-of-the-art CNN models as well the compared benchmark algorithm. We have developed a GUI-based application for public use. This application can be used on any computer by any medical personnel to detect COVID +ve patients using Chest X-Ray images within a few seconds.

Keywords Coronavirus · COVID-19 · Ensemble learning · Deep learning · Convolutional neural network

1 Introduction

Coronavirus, as confirmed by WHO [1], records the first official case in Wuhan, the largest metropolitan area of the Hubei province in China. It has already taken several

thousands of lives till date and has millions of confirmed cases across the world. An epidemic that took the shape of a pandemic, has a catastrophic effect on the health and welfare of the global population. This has caused Severe Acute Respiratory Syndrome coronavirus (SARS-CoV) and the infirmity is known as coronavirus disease 2019 (acronym COVID-19) [2]. This coronavirus belongs to the same family as that of SARS and MERS, but with a more virulent and aggressive nature (2019-nCoV). This contagious infection spreads much faster (through respiratory droplet infection) than other normal flu.

Right now, the majority of tests being used to diagnose COVID-19 are genetic tests known as Reverse Transcription Polymerase Chain Reaction (RT-PCR) [3]. These tests are very accurate. Even if there is only a tiny amount of virus in the patient sample, it can be detected and measured. However, it is worth noting that the PCR test is very complicated, time-consuming, and costly. So not all healthcare facilities have the ability to perform it. Perceiving these limitations, a stand-in approach to detect the disease can be radiography scanning, where chest radiography images can be analyzed to detect the presence of, or

✉ Amit Kumar Das
amitkrdas.kol@gmail.com

Sayantani Ghosh
sayantani31g@gmail.com

Samiruddin Thunder
samird2000@gmail.com

Rohit Dutta
rohitdutta2510@gmail.com

Sachin Agarwal
agarwalsachin518@gmail.com

Amlan Chakrabarti
acakcs@caluniv.ac.in

¹ A. K. Choudhury School of Information Technology,
University of Calcutta, Kolkata, India

² Institute of Engineering and Management, Kolkata, India

the symptoms of the novel coronavirus. Studies show that viruses belonging to this family demonstrate significant manifestation in radiographic images [4–11]. Therefore, it can be said that classification with the help of radiographic images, such as chest X-ray (CXR), can be accurate but at the same time much faster and less expensive than the PCR test. Furthermore, chest X-rays are economical than other radiological tests like CT scans and available in almost every clinic.

The only perceived challenge in CXR-based detection of COVID-19 patients is that trained doctors may not be available all the time, especially in remote areas. Also, the radiological manifestations related to COVID-19 are new and unfamiliar with many experts not having past experience with COVID-19 positive patient CXRs. So, we have proposed a simple and inexpensive deep learning-based technique to classify COVID-19 +ve and -ve cases using CXR images. Using this technique a near-accurate detection of COVID-19 positive patients can be done in a few seconds. As a part of this research, we have also contributed a tool that can be used to detect COVID-19 positive patients. Even in the absence of a radiologist or if there is any difference in opinions of the doctors, this deep learning-based tool will always give an opinion without the need for human intervention. In this article, with the data available from open sources we have shown the efficacy of the proposed tool in terms of classification accuracy and sensitivity. It has also been compared with an existing benchmark work.

Few works have used only the individual deep learning techniques with CXR images [4, 12, 13] to make COVID-19 +ve and -ve prediction. One work has attempted to develop a custom network [5]. But our work is focused on using multiple state-of-the-art deep learning models and then ensembling them to achieve better accuracy. It is based on the simple philosophy that an ensemble of multiple models provides better performance compared to individual models [14].

The remaining part of the paper is organized as follows:

- Section 2 lays down the related research, their approaches and their methodology.
- Section 3 contains details about our proposed technique along with some context around the state-of-the-art models that we have used.
- In Section 4 we have discussed the experimental set-up used in our research.
- Section 5 presents the experimental results including classification accuracy, sensitivity and F1-score obtained from the proposed work.
- Finally in Sect. 6, the paper has been concluded with a summary of the outcome of our research.

2 Related works

Computer vision has found prominent usage in medical diagnosis. It is useful in the medical fields that require visual checks like dermatology. Computer vision is used as a tool to diagnose whether a skin abnormality is an early potential indicator of skin cancer. It is also used to detect issues within the body, especially in the tissues, blood vessels, joints, etc. It is used in ophthalmology to diagnose diseases like diabetic retinopathy, thus helping to prevent blindness. It has also shown great success in surgeries as well as therapies. Computer vision solutions use various types of medical imaging e.g., Computed Tomography scan (CT scan), Magnetic resonance imaging (MRI), Positron Emission Tomography scan (PET scan), ultrasound and Chest X-Ray (CXR) images.

Studies show that medical images help in improving the analysis of the presence of viruses in the lungs. In multiple works, deep learning-based techniques have been developed to identify pneumonia [15, 16], different classes of thoracic diseases [8, 9, 17], skin cancer [18], haemorrhage classification [19], etc. from medical images. Some of these works have given promising results with relatively simple architecture [8].

In a work by Singh et al. [20], a convolutional neural network (CNN) model has been used to identify COVID-19 patients with the help of CT scan images. There are several more research works to detect the presence of the COVID-19 virus in the human lungs with the help of a CT scan [11, 20–23]. In the work by Yan et al. [11], a multi-task, self-supervised AI model has been developed for the diagnosing of the COVID-19 virus in human lungs with the help of CT scan images, with an accuracy of 89%. Automatic segmentation and quantification of the lungs are done by Shan et. al. [21]. Li et al. [22] describe a fully automatic framework to detect coronavirus affected lungs from chest CT scan images and differentiated it from other lung diseases. However, Ng et al. [24] and Huang et al. [25] have concluded that CXR images are better than any other means in the detection of COVID-19 because of their promising results along with the availability of CXR machines and their low maintenance cost.

There have been multiple works done by researchers in the area of COVID-19 patient detection using CXR images [4, 5, 7, 12, 26–32]. In one such work by Makris et al. [4], transfer learning has been used with the Inception-v3 network to classify normal, pneumonia and COVID-19 patients using CXR images. Another work by Mangal et al. [6] has used the DenseNet with ChexNet architecture to segregate normal subjects, bacterial and viral pneumonia patients and COVID-19 patients. Rahimzadeh et al. [7] has used a concatenated Xception and ResNet50V2. Another

work by Xu et al. [26] has used ResNet to detect viral pneumonia and COVID-19 patients. Unavailability of a large number of image data of COVID-19 +ve patients is a challenge faced by most researchers working in this area. Development of COVIDGAN for the generation of data artificially has been done in a work by Waheed et al.[33]—which in turn will help in improved COVID-19 detection.

Apart from using individual state-of-the-art deep learning models, there has been one work by Wang et al. [5] which has developed a custom architecture termed as COVIDNet architecture for the classification of COVID-19 patients, healthy subjects and pneumonia patients. This custom network, designed using a lightweight projection-expansion-projection-extension (PEPX) design pattern, has demonstrated a classification accuracy of 94%—a result that outperforms laboratory testing.

As can be observed, most of the works related to COVID-19 detection from CXR images have utilized individual deep learning models e.g., DenseNet, ResNet, Xception, etc. [12, 27, 28]. None of the works have tried to combine the models to multiply their capability of classification. Various works done on Ensemble Learning with Deep Neural Networks show that ensembling learning methods are superior in prediction than an individual model and also helps in preventing overfitting [34]. In [35], a weighted average of the output probabilities has been introduced as a method for ensembling. It is found to be better than the unweighted average. In another work Cheng et al. [36], the relative performance of the different ensemble methods with Convolutional Neural Networks like unweighted average, majority voting, Bayes Optimal Classifier, and Super Learner has been compared.

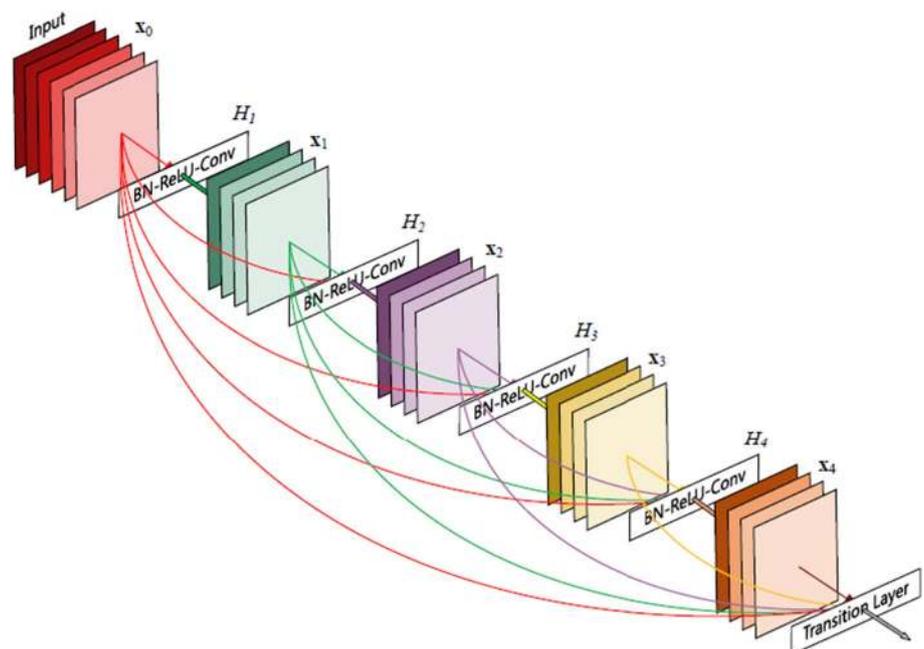
In this research, we have proposed a new method to ensemble three state-of-the-art CNN models—DenseNet201, Resnet50V2, and Inceptionv3 to classify COVID-19 +ve patients from CXR images.

3 Proposed approach

In real life, we always prefer to come up with a medical diagnosis based on multiple medical expert views. The combined opinion of the medical experts helps in reaching a more reliable conclusion. Following the same philosophy, multiple benchmark CNN models have been adopted in our proposed work. They have been trained individually to make independent predictions. Then the models are combined, using a new method of weighted average ensembling technique, to predict a class value. This newly proposed ensembling method is expected to make the prediction more robust. Our proposed work comprises of three pre-trained CNN models—DenseNet201 [37], Resnet50V2 [38] and Inceptionv3 [39]. The biggest advantage of Dense Convolution Network or DenseNet, shown in Fig. 1, is that it requires comparatively fewer parameters than similar types of traditional CNN. An additional reason to choose DenseNet is that each layer takes the feature maps of all preceding layers as inputs. This helps to strengthen feature propagation and encourages feature reuse.

ResNet50V2, shown in Fig. 3, is a contemporary convolutional network which addresses the vanishing or exploding gradient problems by the use of “residual blocks” in the architecture. In a residual network, multiple residual blocks

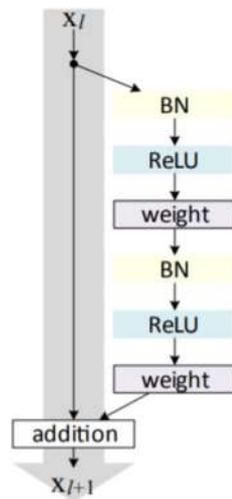
Fig. 1 DenseNet architecture [37]



stacked up one after another. Each residual block, shown in Fig. 2, is formed of short-cut connections skipping one or more layers. Resnet50V2 uses the pre-activation of weight layers.

The Inception model, shown in Fig. 5, is a powerful model that can accomplish very high accuracy to extract the features and classify images based on those features. Inceptionv3 has a 48-layer deep architecture consisting of 11 inception modules. Each inception module, shown in Fig. 4, consists of convolution filters, pooling layers, and ReLU (Rectified Linear Unit)

Fig. 2 Residual block [38]



activation function. As a means of regularization, before a fully connected layer, a dropout of 0.6 is added.

One salient point in our proposed approach is that we have used a new weighted average based ensembling method to combine individual CNN models. In this method, if one model, say ResNet50V2, is performing better than the other two models i.e., having lower validation error, it is assigned a higher weight so that its contribution in deciding the class value is higher. Assuming the accuracy percentage of the i -th model as a_i , the validation error is $(100 - a_i)$. We define a factor k_i as below:

$$d_i = 100 - a_i \tag{1}$$

$$D = \sum d_i^2 \tag{2}$$

$$k_i = \frac{d_i^2}{D}, \tag{3}$$

Weight of the i -th network is defined as:

$$w_i = \frac{1}{\sum \frac{1}{k_i^2}}, \tag{4}$$

Let us assume that the output from the neural networks is of the form $[y_0, y_1]$, where y_0 denotes the probability of Class 0 and y_1 denotes the probability of Class 1. Let the predictions

Fig. 3 ResNet50V2 architecture [38]

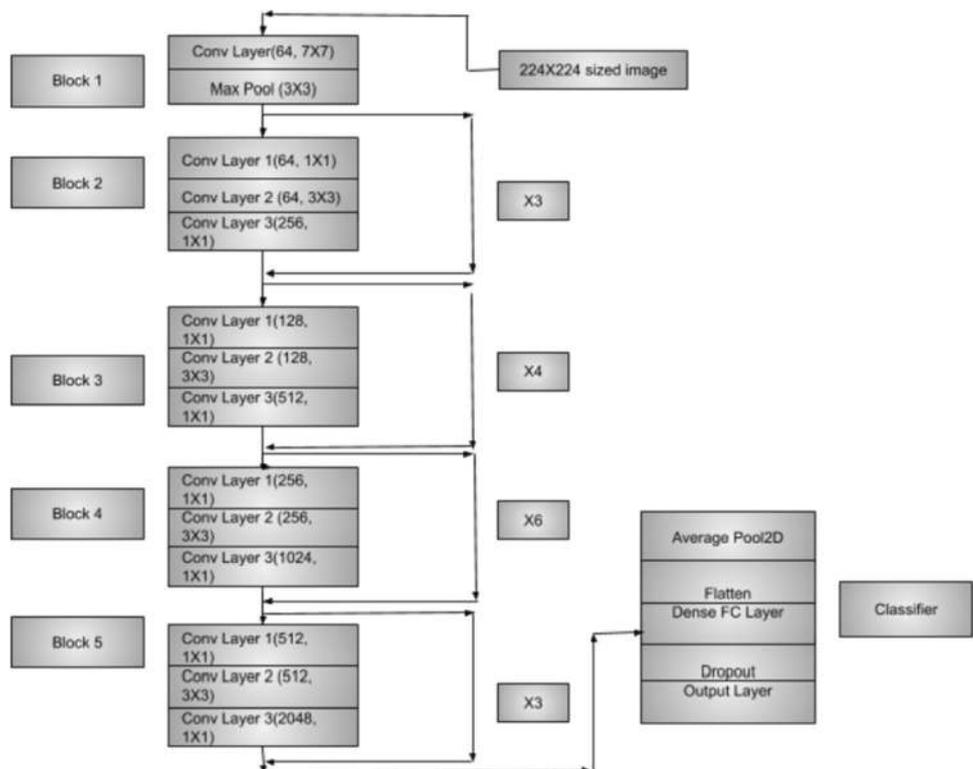


Fig. 4 Inception module [39]

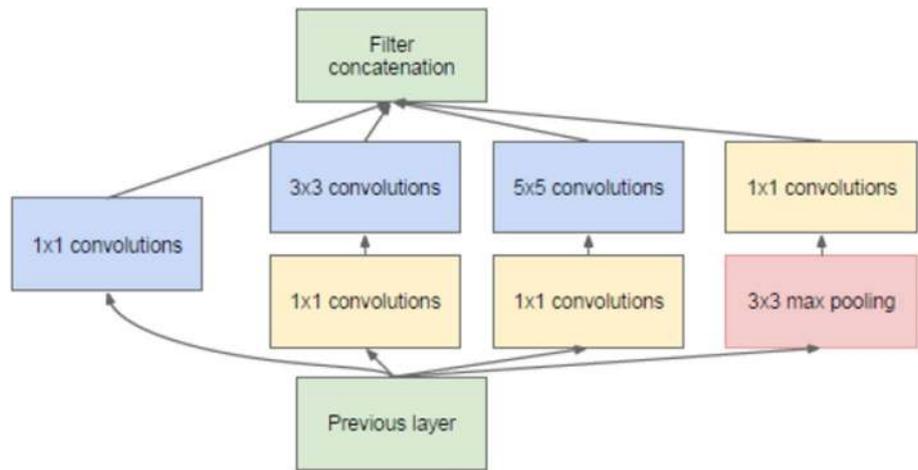
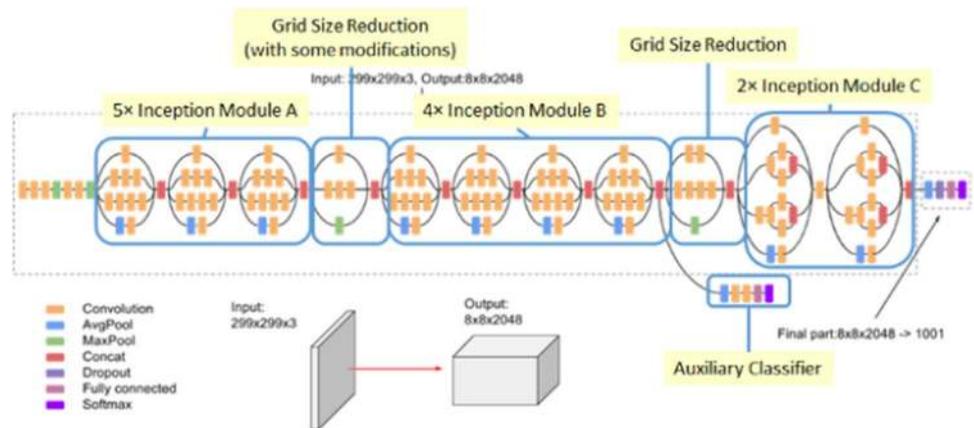


Fig. 5 Inceptionv3 architecture [39]



from the three different models be of the form $[y_{01}, y_{11}]$, $[y_{02}, y_{12}]$, $[y_{03}, y_{13}]$ for models 1, 2 and 3, respectively. Let the weights calculated using the proposed method mentioned in Eq. 4 be $[w_1, w_2, w_3]$ for the models, respectively. Then the weighted probability is calculated as:

$$\text{Average} = \left[\frac{w_1 \times y_{01} + w_2 \times y_{02} + w_3 \times y_{03}}{w_1 + w_2 + w_3}, \frac{w_1 \times y_{11} + w_2 \times y_{12} + w_3 \times y_{13}}{w_1 + w_2 + w_3} \right]$$

into 5 folds. Each time one separate fold is picked up as test data and the remaining folds as training data.

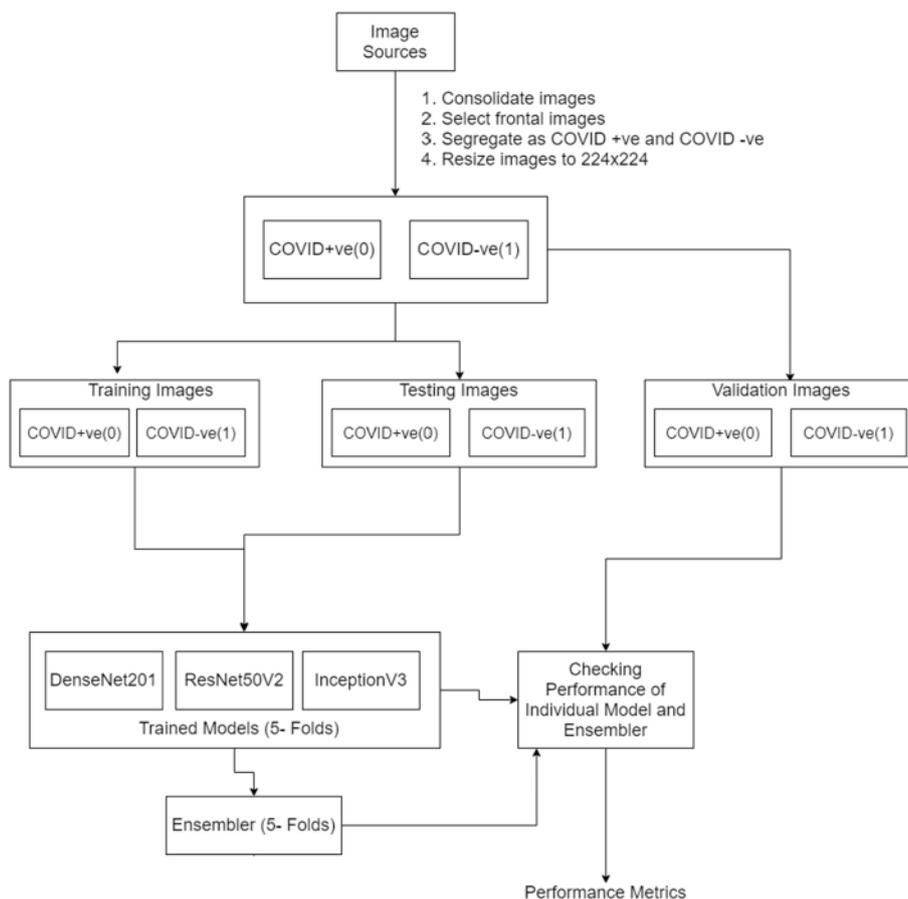
- While dividing the images into training, testing and validation sets, ensure that there is no patient overlap i.e., different images of the same patient is not present

The overall proposed approach, as summarized in Fig. 6, includes:

- Consolidate CXR images for healthy subjects, patients having pneumonia or other bacterial infection and COVID patients from different sources.
- Retain only frontal CXR images.
- Resize images to a uniform size.
- Divide the images into three portions—training, testing and validation datasets. One small portion is retained as validation set to test the efficacy of the trained model while the remaining portion is divided

in multiple sets.

- Train the models - DenseNet201, ResNet50V2, and Inceptionv3 using training set images and do the loss minimization based on the test set images. Calculate the weights of the 5-fold cross validation based on the test set.
- Run the trained models on the validation set images and select class label value 0 or 1 based on weighted average ensembling of the 3 models.

Fig. 6 Proposed approach

4 Materials and methods

4.1 Dataset generation

For this research work, we have collected the images from different open sources [40–46]. These open-source public datasets contain CXR images of COVID-19 positive patients, patients having pneumonia and other infections and healthy subjects, primarily collected from European countries. This data contain CXR images of different patients from which only the frontal images are considered and lateral images are discarded. This is because our region of interest is the lungs and lungs can be better examined with a frontal view than a lateral one. The original set of CXR images are labeled as COVID-19, Pneumonia, and Normal. However, for this work, we have segregated the images into two broad categories—COVID-19 POSITIVE (referred to as class 0) and COVID-19 NEGATIVE (referred to as class 1). For Class 0 there are 538 images similar to images shown in Fig. 7 whereas for Class 1 there are 468 images of COVID negative patients similar to images in Fig. 8.

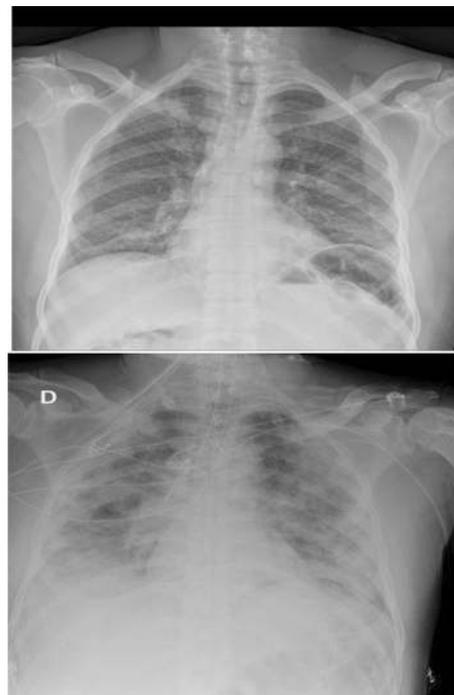
**Fig. 7** CXR images of COVID-19 positive subjects [45]



Fig. 8 CXR images of COVID-19 negative subjects [43]

4.2 Pre-processing

The consolidated images are first normalized and resized into 224×224 shaped images. Then the images are shuffled and split into training, test and validation data. The training data have 771 images where 438 images are for class 0 and 333 images are for class 1. The test data have 118 images where 43 images are for class 0 and 75 images for class 1. Similarly, the validation data have 117 images where 57 images are for class 0 and 60 images for class 1. To ensure effective training, we have adopted 5-fold cross validation while splitting the set of images. While a fixed set of images have been selected as validation data, the images for training and test datasets have been selected by splitting the combined set of images (i.e., training and test images combined) into 5 folds and picking up a different fold each time as the test set. Five iterations have been executed, one with each fold of test data and corresponding training data. However, if there are two or more images of the same patient, it is ensured that those images are either marked as training data or as test data—but not in both. In case the same patient’s images are kept both in training and test data, there is a possibility that the results will be overly promising because of patient overlap. With this, the image folders are ready for training the model followed by testing the efficacy of the trained model (see Table 1).

Table 1 Data distribution for training, test and validation

Dataset	COVID +ve (class 0)	COVID –ve (class 1)	Total
Training	438	333	771
Testing	43	75	118
Validation	57	60	117

4.3 Tools used

We have used Google Colab GPU (Tesla K80 12GB GDDR5 VRAM), Python 3.7 and TensorFlow 2.2.0. For the implementation of CNN, the deep learning library of TensorFlow 2.2.0 is used, and the training and the testing procedures are done in the Google Colab platform.

4.4 Performance metrics

To evaluate the performance of the proposed approach, the metrics adopted are classification accuracy, sensitivity and *F1*-score, measured as follows:

$$\text{Classification accuracy} = \frac{TP + TN}{TP + TN + FP + FN}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN}$$

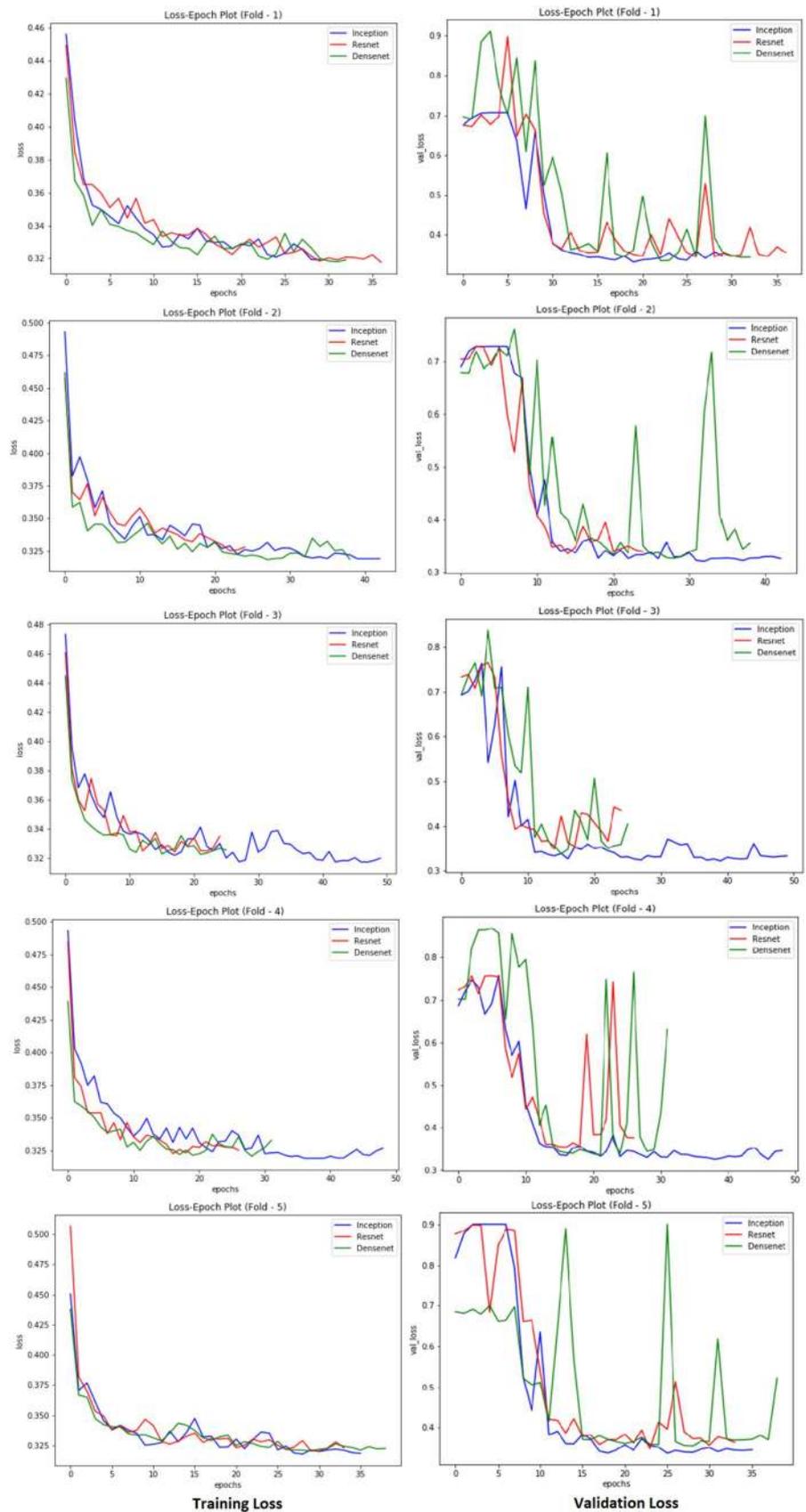
$$F1 \text{ Score} = \frac{2 \times \text{sensitivity} \times \text{precision}}{\text{sensitivity} + \text{precision}}$$

where TP stands for True Positive, FP for False Positive, FN for False Negative and TN for True Negative. In a confusion matrix, the COVID-19 +ve cases that are correctly classified by the model are termed as True Positive and incorrectly classified as COVID –ve are termed as False Positive. Similarly, COVID –ve subjects classified correctly are termed as True Negative and incorrectly classified as COVID +ve are termed as False Negative.

4.5 Compared benchmark

In this paper, we have proposed a new weighted average based ensemble technique. However, alongside the results from the individual networks used in the ensemble, we have also compared the results from the proposed algorithm with another contemporary work on COVID-19 detection from X-ray images [7] which has used a concatenation of benchmark CNNs (henceforth referred to as ‘Concatenated network’). This gives a good

Fig. 9 Fold-wise training and testing loss of DenseNet201, ResNet50V2 and Inceptionv3



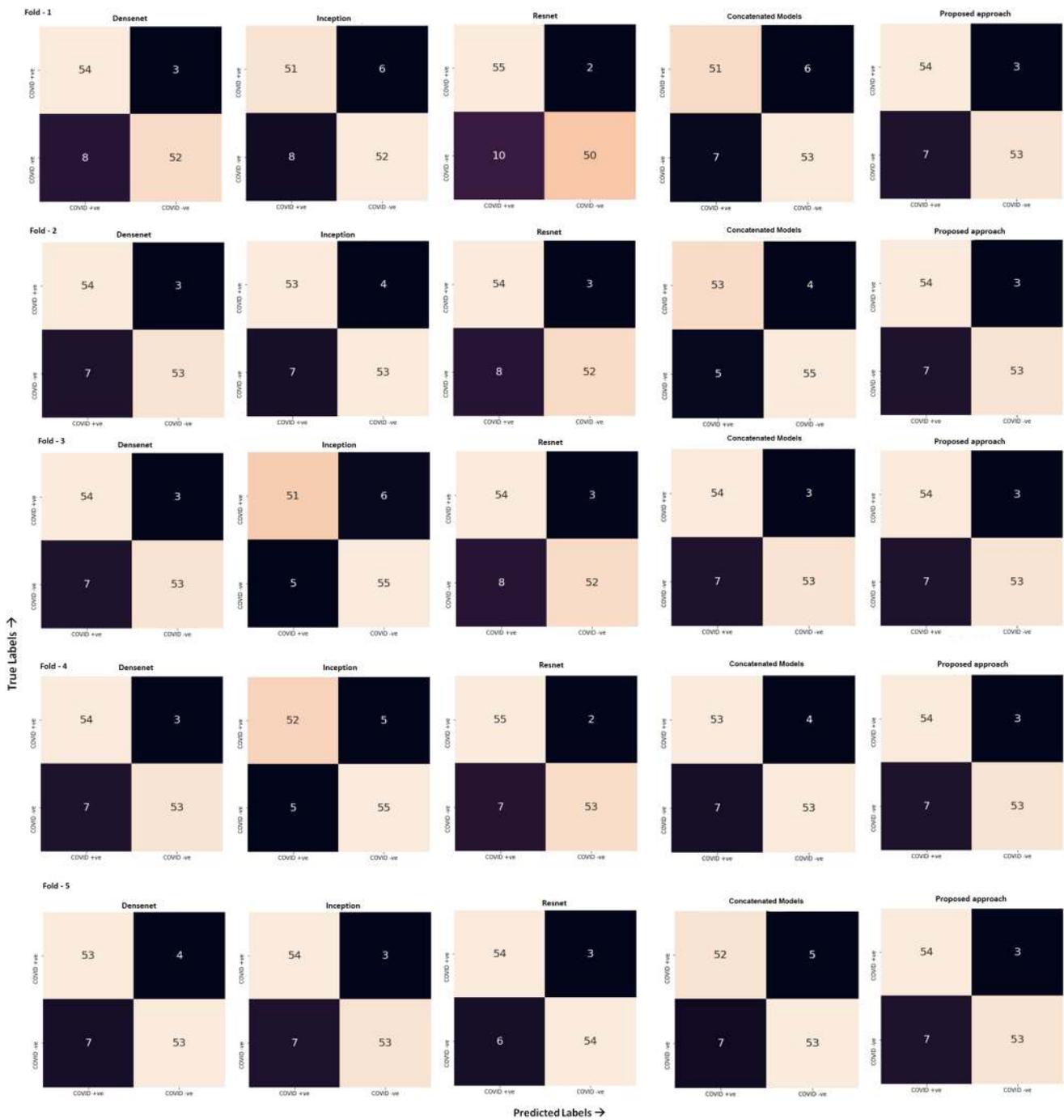


Fig. 10 Confusion matrix for the model outputs trained across different Folds

understanding of how the proposed algorithm performs compared to the benchmark CNNs which it has ensembled as well as compared to recent work in the same space.

5 Experiments and results

All the models have been trained for 60 epochs with Early Stopping callbacks (patience = 10 epochs). Adam optimizer, a combination of SGD with momentum and RMSProp, is used for faster convergence with the parameters as learning rate $\alpha = 0.0001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 1 \times 10^{-7}$. The same optimizer is used for all three models and then

Table 2 Number of true/false positives and true/false negatives for each class

Fold	Model	COVID+ Correctly detected (TP)	COVID+ Wrongly detected (FP)	COVID– Correctly Detected (FN)	COVID– Wrongly detected (TN)
1	InceptionV3	51	8	52	6
	Resnet50V2	55	10	50	2
	Densenet201	54	8	52	3
	Concatenated network	51	7	53	6
	Proposed approach	55	7	53	2
2	InceptionV3	53	7	53	4
	Resnet50V2	54	8	52	3
	Densenet201	54	7	53	3
	Concatenated network	53	5	55	4
	Proposed approach	54	7	53	3
3	InceptionV3	51	5	55	6
	Resnet50V2	54	8	52	3
	Densenet201	54	7	53	3
	Concatenated network	54	7	53	3
	Proposed approach	54	7	55	3
4	InceptionV3	52	5	55	5
	Resnet50V2	55	7	53	2
	Densenet201	54	7	53	3
	Concatenated network	53	7	53	4
	Proposed approach	54	7	53	3
5	InceptionV3	54	7	53	3
	Resnet50V2	54	6	54	3
	Densenet201	53	7	53	4
	Concatenated network	52	7	53	5
	Proposed approach	54	7	53	3

the models are saved as .h5 files. The time taken for model training is – 31 s/epoch for DenseNet201 and 17 s/epoch for each of the models ResNet50V2 and Inceptionv3. In Fig. 9 the gradual change in loss (both training as well as validation) through the epochs have been depicted for all the three models DenseNet201, ResNet50V2 and Inceptionv3.

Figure 10 presents the confusion matrix for the performance of different models trained across different folds. As can be observed, the proposed ensemble model gives the most consistent performance as well as one of the best performances with high true positive and true negative counts both for COVID +ve as well as for COVID –ve images.

Table 2 presents the count of correctly and incorrectly classified COVID +ve and COVID –ve images. It can be observed that the number of correctly classified COVID +ve images is consistently high for the proposed algorithm. Table 3 shows the summarization of accuracy, sensitivity and *F1*-score of the state-of-the-art CNN model, the competing benchmark algorithm and our proposed model. It is evident from the summary that the performance of our proposed solution is better than the other models in terms of accuracy. Also, it gives a consistently high value of sensitivity for both

COVID +ve and COVID –ve cases. This demonstrates the robustness of the model.

Also, to demonstrate the efficacy of the proposed ensemble based algorithm, a comparison of the proposed algorithm with a number of recent works on COVID detection using Chest X-Ray images have been presented in Table 4. It can be observed that the proposed method integrates the capabilities of state-of-the-art deep learning models to yield comparable or better results that the works where vanilla state-of-the-art deep learning models have been used [12, 27, 28]. Some of the new methods [29–31] show very promising results, though they may suffer from criticism due to the small size of the data used in the experimental set-up.

5.1 The prototype Tool

Based on the proposed solution, a simple desktop tool for the detection of COVID-19 positive and negative cases has been developed. This allows any medical personnel to browse a chest X-ray image and feeding it to the application. The application, in turn, will execute the model proposed in this work and provide the label for the given Chest X-Ray image.

Table 3 Summary of overall performance

Fold	Model	Accuracy (%)	COVID+ sensitivity (%)	COVID- sensitivity (%)	COVID+ F1-Score (%)	COVID- F1-Score (%)	AUC-Score (%)
1	InceptionV3	88.03	89.47	86.67	87.93	88.14	88.07
	Resnet50V2	89.74	96.49	83.33	90.16	89.29	89.91
	Densenet201	90.60	94.74	86.67	90.76	90.43	90.70
	Concatenated network	88.89	89.47	88.33	88.70	89.08	88.90
	Proposed approach	92.31	96.49	88.33	92.44	92.17	92.41
2	InceptionV3	90.60	92.98	88.33	90.60	90.60	90.66
	Resnet50V2	90.60	94.74	86.67	90.76	90.43	90.70
	Densenet201	91.45	94.74	88.33	91.53	91.38	91.54
	Concatenated network	92.31	92.98	91.67	92.17	92.44	92.32
	Proposed approach	91.45	94.74	88.33	91.53	91.38	91.54
3	InceptionV3	90.60	89.47	91.67	90.27	90.91	90.57
	Resnet50V2	90.60	94.74	86.67	90.76	90.43	90.70
	Densenet201	91.45	94.74	88.33	91.53	91.38	91.54
	Concatenated network	91.45	94.74	88.33	91.53	91.38	91.54
	Proposed approach	91.45	94.74	88.33	91.53	91.38	91.54
4	InceptionV3	91.45	91.23	91.67	91.23	91.67	91.45
	Resnet50V2	92.31	96.49	88.33	92.44	92.17	92.41
	Densenet201	91.45	94.74	88.33	91.53	91.38	91.54
	Concatenated network	90.60	92.98	88.33	90.60	90.60	90.66
	Proposed approach	91.45	94.74	88.33	91.53	91.38	91.54
5	InceptionV3	91.45	94.74	88.33	91.53	91.38	91.54
	Resnet50V2	92.31	94.74	90.00	92.31	92.31	92.37
	Densenet201	90.60	92.98	88.33	90.60	90.60	90.66
	Concatenated network	89.74	91.23	88.33	89.66	89.83	89.78
	Proposed approach	91.45	94.74	88.33	91.53	91.38	91.54
Average	InceptionV3	90.43	91.58	89.33	90.31	90.54	90.46
	Resnet50V2	91.11	95.44	87.00	91.28	90.93	91.22
	Densenet201	91.11	94.39	88.00	91.19	91.03	91.19
	Concatenated network	90.60	92.28	89.00	90.53	90.66	90.64
	Proposed approach	91.62	95.09	88.33	91.71	91.54	91.71

As a result, this will detect the COVID +ve and COVID -ve cases along with their probabilities as shown in Fig. 11. This can be used on platforms like Windows, Mac, and Linux. This interface can be used in any COVID-19 testing centers or other health facilities for the fast detection of the disease. This ready to use tool along with the underlying code for data preparation and model training is available publicly at <https://github.com/CUIEMCOVIDProject/COVID-19-Detection-Using-Ensemble-Learning>.

6 Conclusion

Fast and timely detection of COVID +ve patients is necessary to avoid spreading the disease and keeping it in control. This research work has been done to detect the COVID +ve patients from Chest X-Ray images in a simple and inexpensive way. In the work proposed in this paper, three state-of-the-art deep learning models have been adopted and ensembled. The proposed model has achieved a classification accuracy of 91.62%. Even more important fact is it yields a sensitivity of around 95% for COVID +ve cases i.e., out of 100 COVID +ve patients, more than 95 can be correctly diagnosed by our proposed model. It is believed that this research work along with the GUI interface will

Table 4 Comparative study of previous works in COVID detection using Chest X-ray images

Related work	Number of CXR images included	Evaluation method	Accuracy	Sensitivity	Specificity	Classification method used	Overall summary
Apostolopoulos et al. [12]	1,428 (COVID = 224, Pneumonia = 700, Normal = 504)	10-fold cross-validation	90.5% ($\pm 6.97\%$)	41% ($\pm 50.5\%$)	99% ($\pm 1.2\%$)	VGG 19, Mobile Net v2, Inception, Xception, Inception Resnet v2	Comparison of existing state-of-the-art CNN models; High on accuracy and very high specificity. There seems to be some issues mentioned in the reported sensitivity data
Duran-Lopez et al. [32]	6,926 (COVID = 2,589, Normal = 4,337)	5-fold cross validation	94.43%	92.53%	96.33%	New Method - COVID-Xnet	New method - COVID-Xnet proposed. Demonstrates high accuracy, sensitivity and specificity. However, number of images used is quite low
Jain, R et al. [27]	6,432 (COVID = 576, Pneumonia = 4,273, Normal = 1,583)	Train: Test = 5467 : 965	95.3% (+ / - 2.1%)	Normal - 92.7% ($\pm 5\%$), Pneumonia - 99.3%, COVID - 92.3% ($\pm 2.5\%$)	98.2% ($\pm 1.2\%$)	Inception V3, Xception, ResNeXt	Comparison of existing state-of-the-art CNN models; Demonstrates high accuracy, sensitivity and very high specificity
Khan et al. [29]	594 (COVID = 284, Normal = 310)	4-fold cross validation	95.3% ($\pm 4.9\%$)	97.5% ($\pm 1.1\%$)	98.60%	New method - CoroNet (Xception)	New method - CoroNet proposed. Demonstrates high accuracy, sensitivity and specificity
Mahmud et al. [31]	610 (COVID = 305, Normal = 305)	5-fold cross validation	97.40%	97.80%	94.70%	Stacked Multi-Resolution CovXNet	New method - CovXNet proposed. Demonstrates very high accuracy, sensitivity and specificity. However, number of images used is quite low
Minaei et al. [28]	5,184 (COVID = 184, Normal = 5,000)	Train: Test = 2084 : 3100	90.89% ($\pm 1.9\%$)	98% ($\pm 2.7\%$)	87.1% ($\pm 1.7\%$)	ResNet18, ResNet50, SqueezeNet, and DenseNet-121	Comparison of existing state-of-the-art CNN models; Demonstrates very high on sensitivity

Table 4 (continued)

Related work	Number of CXR images included	Evaluation method	Accuracy	Sensitivity	Specificity	Classification method used	Overall summary
Ozturk et al. [30]	625 (COVID- 125, Normal - 500)	5-fold cross validation	98.08%	95.13%	95.30%	New Method - DarkNet	New method - DarkNet proposed. Demonstrates very high accuracy, sensitivity and specificity. However, number of images used is quite low
Wang, L et al. [5]	13,975 (COVID = 5,538, Normal = 8,066)	Train: Test = 50 : 1 (approx.)	93.30%	Normal - 95%, Non-COVID- 94%, COVID- 91%	95%	New method - COVIDNet	New method - COVIDNet proposed. Demonstrates high accuracy, sensitivity and specificity
Proposed Method	1,006 (COVID = 538, non-COVID = 468)	5-fold cross validation	91.62%	95.09%	88.33%	New ensemble method combining InceptionV3, Resnet50V2 and Densenet201	Unique ensemble based technique proposed. Demonstrates high accuracy, sensitivity and specificity

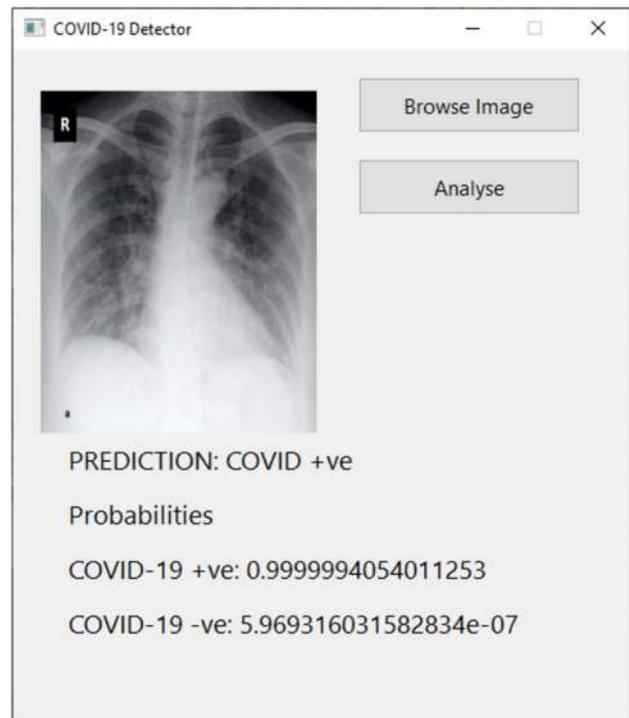


Fig. 11 GUI-based tool for COVID-19 detection

help the doctors to detect the affected patients with the help of computer-aided analysis, that too within a few seconds. We believe that this will significantly add value to the medical field.

References

1. WHO—Emergencies preparedness, response. World Health Organization (2019) Pneumonia of unknown cause China. https://www.who.int/csr/don/05-january-2020-pneumonia-of-unknown-cause-china/en/?mod=article_inline. Accessed 29 Mar 2020
2. The SARS-CoV-2 outbreak: what we know 94:44–48
3. Bustin SA (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* 25(2):169–193
4. Makris A, Kontopoulos I, Tserpes K (2020) COVID-19 detection from chest X-ray images using deep learning and convolutional neural networks. medRxiv. Accessed 29 Mar 2020
5. Wang L, Lin ZQ, Wong A (2020) COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports*, Vol. 10
6. Mangal A, Kalia S, Rajgopal H, Rangarajan K, Namboodiri V, Banerjee S, Arora C (2020) COVIDAID: COVID-19 detection using chest X-ray. [arXiv:2004.09803](https://arxiv.org/abs/2004.09803)
7. Rahimzadeh M, Attar A (2020) 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. *Inf Med Unlock* 19:100360
8. Rajpurkar P, Irvin J, Zhu K, Yang B, Mehta H, Duan T, Ding D, Bagul A, Langlotz C, Shpanskaya K, et al (2017) Chexnet: radiologist-level pneumonia detection on chest x-rays with deep learning. [arXiv preprint arXiv:1711.05225](https://arxiv.org/abs/1711.05225)

9. Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM (2017) Chestx-ray8: hospital-scale chest x-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In: Proceedings of the IEEE Conference on computer vision and pattern recognition, pp 3462–3471
10. Wang L, Wong A (2020) COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest radiography images
11. Yan Q, Wang B, Gong D, Luo C, Zhao W, Shen J, Shi Q, Jin S, Zhang L, You Z (2020) COVID-19 Chest CT image segmentation—a deep convolutional neural network solution. [arXiv:2004.10987](https://arxiv.org/abs/2004.10987)
12. Apostolopoulos ID, Bessiana T (2020) COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys Eng Sci Med*, pp 1–6
13. Ilyas M, Rehman H, Nat-Ali A (2020) Detection of COVID-19 from chest X-ray images using artificial intelligence: an early review. [arXiv:2004.05436](https://arxiv.org/abs/2004.05436)
14. Smolyakov V (2017) Ensemble learning to improve machine learning results. <https://blog.statsbot.co/ensemble-learning-d1dcd548e936>. Accessed 29 Mar 2020
15. Chouhan V, Singh S, Khamparia A, Gupta D, Tiwari P, Moreira C, Damaevius R, Albuquerque V (2020) A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Appl Sci* 10:559
16. Jaiswal A, Tiwari P, Kumar S, Gupta D, Khanna A, Rodrigues JJ (2019) Identifying pneumonia in chest X-rays: a deep learning approach. *Measurement* 145:511–518
17. Ranjan E, Paul S, Kapoor S, Kar A, Sethuraman R, Sheet D (2018) Jointly learning convolutional representations to compress radiological images and classify thoracic diseases in the compressed domain. Proceedings of the 11th Indian Conference on computer vision, graphics and image processing
18. Andre E, Brett K, Roberto A et al (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542(7639):115–118
19. Grewal M, Srivastava MM, Kumar P, Varadarajan S (2017) Radiologist level accuracy using deep learning for hemorrhage detection in CT scans. [arXiv:1710.04934](https://arxiv.org/abs/1710.04934)
20. Singh D, Kumar V, Vaishali, Kaur M (2020) Classification of COVID-19 patients from chest CT images using multi-objective differential evolution based convolutional neural networks. *Eur J Clin Microbiol Infect Dis* 39:1379–1389
21. Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shen D, Shi Y (2020) Lung infection quantification of COVID-19 in CT images with deep learning. [arXiv:2003.04655](https://arxiv.org/abs/2003.04655)
22. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J (2020) Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology*
23. Dansana D, Kumar R, Bhattacharjee A, Hemanth DJ, Gupta D, Khanna A, Castillo O (2020) Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm. *Soft Comput*, pp 1–9
24. Ng M-Y, Lee EY, Yang J, Yang F, Li X, Wang H, Lui M-M-S, Lo CS-Y, Leung B, Khong P-L, Hui CK-M, Yuen K-Y, Kuo MD (2020) Imaging prole of the COVID-19 infection radiologic endings and literature review. *Radiol Cardiothorac Image* 2(1):e200034. Accessed 29 Mar 2020
25. Huang C et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 395:497–506. Accessed 29 Mar 2020
26. Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Y L, Chen Y, Su J, Lang G, Li Y, Zhao H, Xu K, Ruan L, Wu W (2020) Deep learning system to screen Coronavirus disease 2019 pneumonia. [arXiv:2002.09334](https://arxiv.org/abs/2002.09334)
27. Jain R, Gupta M, Taneja S, Hemanth DJ (2020) Deep learning based detection and analysis of COVID-19 on chest X-ray images. *Appl Intell*, pp 1–11. Accessed 29 Mar 2020
28. Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ (2020) Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning. *Med Image Anal* 65:101794
29. Khan AI, Shah J, Bhat M (2020) CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Comput Methods Programs Biomed* 196:105581. Accessed 29 Mar 2020
30. Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Acharya UR (2020) Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput Biol Med* 121:103792
31. Mahmud T, Rahman MA, Fattah SA (2020) CovXNet: a multidilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Comput Biol Med* 122:103869
32. Duran-Lopez L, Dominguez-Morales JP, Corral-Jaime J, Vicente-Diaz S, Linares-Barranco A (2020) COVID-XNet: a custom deep learning system to diagnose and locate COVID-19 in chest X-ray images. *Appl Sci* 10(16):5683
33. Waheed A, Goyal M, Gupta D, Khanna A, Al-Turjman F, Pinheiro PR (2020) COVIDGAN: data augmentation using auxiliary classifier GAN for improved COVID-19 detection. *IEEE Access* 8:91916–91923
34. Tao S (2019) Deep neural network ensembles. [arXiv:1904.05488](https://arxiv.org/abs/1904.05488)
35. Frazao X, Alexandre LA (2014) Weighted convolutional neural network ensemble. In: Iberoamerican Congress on pattern recognition, Springer, Cham
36. Ju C, Bibaut A, van der Laan M (2018) The relative performance of ensemble methods with deep convolutional neural networks for image classification. *J Appl Stat* 45(15):2800–2818
37. Huang G, Liu Z, Weinberger Kilian Q (2017) Densely connected convolutional networks. In: 2017 IEEE Conference on computer vision and pattern recognition (CVPR) (2017), pp 2261–2269
38. He K, Zhang X, Ren S, Sun J (2016) Identity mappings in deep residual networks. [arXiv:1603.05027](https://arxiv.org/abs/1603.05027)
39. Szegedy Ch, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: 2015 IEEE Conference on computer vision and pattern recognition (CVPR), pp 1–9
40. <https://twitter.com/ChestImaging/status/1243928581983670272>
41. <https://www.sirm.org/category/senza-categoria/COVID-19/>
42. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Illcus S, Chute C, Marklund H, Haghgoo B, Ball RL, Shpanskaya KS, Seekins J, Mong DA, Halabi SS, Sandberg JK, Jones R, Larson DB, Langlotz CP, Patel BN, Lungren MP, Ng AY (2019) CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. In: AAAI
43. <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia>
44. Cohen JP, Morrison P, Dao L (2020) COVID-19 image data collection. [arXiv:2003.11597](https://arxiv.org/abs/2003.11597)
45. Wang L, Wong A, Lin ZQ, Lee J, McInnis P, Chung A, Ross M, VanBerlo B, Ebadi A (2020) Figure 1 COVID-19 chest X-ray dataset initiative. <https://github.com/agchung/Figure1-COVID-chestxray-dataset>
46. Kong W, Agarwal PP (2020) Chest imaging appearance of COVID-19 infection. *Radiol Cardiothorac Imaging* 2(1):e200028. <https://doi.org/10.1148/ryct.2020200028>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.