Enhanced framework for COVID-19 prediction with computed tomography scan images using dense convolutional neural network and novel loss function☆

Anand Motwani [a], Piyush Kumar Shukla [b], Mahesh Pawar [c], Manoj Kumar [d], Uttam Ghosh [e], Waleed Alnumay [f], Soumya Ranjan Nayak [g],*

[a] *Faculty, School of Computing Science & Engineering, VIT Bhopal University, Sehore (MP), 466114, India*
[b] *Department of Computer Science & Engineering, University Institute of Technology, RGPV, Bhopal (MP), 462033, India*
[c] *Department of Information Technology, University Institute of Technology, RGPV, Bhopal (MP), 462033, India*
[d] *Faculty of Engineering and Information Sciences, University of Wollongong in Dubai, Dubai Knowledge Park, UAE*
[e] *Department of Computer Science & Data Science Meharry School of Computational Sciences Nashville, TN, USA*
[f] *Riyadh Community College, CS department, King Saud University, Riyadh, SA*
[g] *School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar 751024, Odisha, India*

ARTICLE INFO

ABSTRACT

Recent studies have shown that computed tomography (CT) scan images can characterize COVID-19 disease in patients. Several deep learning (DL) methods have been proposed for diagnosis in the literature, including convolutional neural networks (CNN). But, with inefficient patient classification models, the number of 'False Negatives' can put lives at risk. The primary objective is to improve the model so that it does not reveal 'Covid' as 'Non-Covid'. This study uses Dense-CNN to categorize patients efficiently. A novel loss function based on cross-entropy has also been used to improve the CNN algorithm's convergence. The proposed model is built and tested on a recently published large dataset. Extensive study and comparison with well-known models reveal the effectiveness of the proposed method over known methods. The proposed model achieved a prediction accuracy of 93.78%, while false-negative is only 6.5%. This approach's significant advantage is accelerating the diagnosis and treatment of COVID-19.

1. Introduction

"Severe Acute Respiratory Syndrome – CoronaVirus" (SARS-CoV) a deadly and new clinical syndrome first identified in first trimester of 2003 [1,2]. A new SARS, 2019-nCoV (coronavirus) recently out broke in Wuhan, Hubei Province of China in December 19 and named as SARS-CoV-2 by the "International Committee of Taxonomy of Viruses" (ICTV) [3]. The disease caused due to this virus is called as "coronavirus disease: COVID-19". COVID-19 rising more rapidly, as there are 16,27,73,940 confirmed cases, 33,75,573 confirmed deaths globally reported to WHO, as of 5:33 pm CEST, 17 May 2021 [4]. Figs. 1 and 2 show example CT-images of 3, COVID-19 positive and negative patients, respectively. The CT image is a reconstructed image created using an X-ray absorption profile [5]. It is preferable to reconstruct images with the lowest possible radiation dose and noise while maintaining image accuracy and
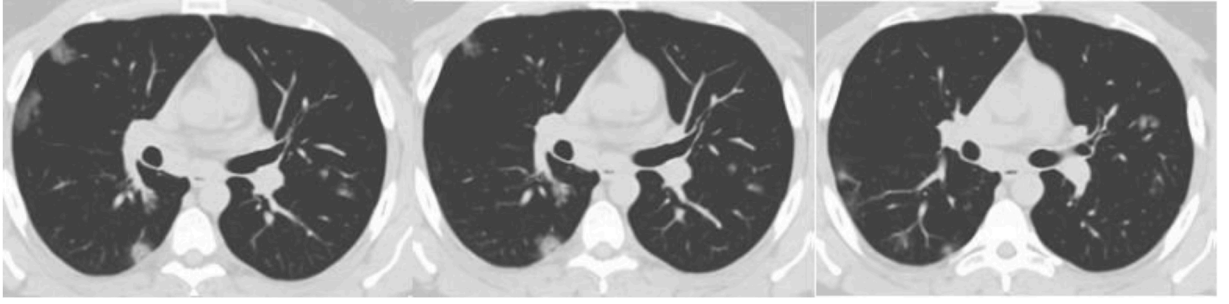

Fig. 1. CT-scan images of 3 COVID-19 patients.

Fig. 2. CT-scan images of 3 Non-COVID patients.

spatial resolution. In these images, the bilateral change is also depicted as a serious issue [6]. CNN is the most successful algorithm that has proven its ability to diagnose disease from medical images with high accuracy.

There are several deep learning optimization challenges including local minima, generalization error and proper convergence. To converge the algorithms in reasonable time is bit tricky which could take hours or days even on a modern GPU. The loss function based on a training dataset is usually an objective function for optimization algorithm. We must first define the loss function because updating our model requires computing the gradient of our loss function.

An ANN particularly DL network has a large number of weight parameters w_i (i=1, 2… N). There are M training samples each with input $\overrightarrow{X_k}$ and a corresponding correct output $\overrightarrow{Z_k}$ for (k=1, 2… N). For each input $\overrightarrow{X_k}$, the model predicts an output $\overrightarrow{Y_k} = G(\overrightarrow{X_k}, \overrightarrow{w})$, where the G (output function) is based on the DL architecture and weights $\overrightarrow{w}$. Thus, characterized by the loss function (Eq. (1)), the goal of learning is to minimize the loss (L) between the actual and predicted outcome.

$$L(\overrightarrow{w}) = M^{-1} \sum_{k=1}^{M} d\left(\overrightarrow{Y_k}, \overrightarrow{Z_k}\right) \tag{1}$$

Where $d\left(\overrightarrow{Y_k}, \overrightarrow{Z_k}\right)$ is distance between $\overrightarrow{Y_k}$ and $\overrightarrow{Z_k}$. In this study a novel loss for d has been used. This approach will lead to an optimal reduction in losses as shown in Fig. 5 below. As a result, selecting the appropriate loss function can have a significant impact on the model's output. In this work, we consider a DL model learned by minimizing loss (based on cross entropy), and show that the DNN model achieves efficient convergence. The authors [7] also obtained the fast convergence rate of a Deep Neural Network (DNN) classifier that is learned using a loss function.

The survey of related algorithms / models to be used for COVID-19 diagnosis is presented in Section 2.

1.1. Motivation

Diagnostic tests of suspected or confirmed cases of COVID-19 require specific medical kits, equipment and precautions in handling samples. Also, it is challenging to diagnose COVID-19 when other lung contagions such as pneumonia also affect the lungs. On the other hand, CT scans and X-ray images highlight typical manifestations of the disease. In fact, CT imaging is a beneficial tool in the detection of COVID-19 disease [8]. *However, image pixels have a spatial local correlation, so diagnosis and classification can only be done using advanced computational and deep learning (DL) techniques.*

COVID-19 is a serious illness that can cause a variety of symptoms from simple sneezing to severe respiratory illness and death [9]. The epidemic is highly contagious, and close contact with sick people within 2 weeks of symptom onset may be a potential cause of illness to others. Symptoms of COVID-19 disease appear after about 5.2 days of gestation [10], during which time the infected person

unknowingly spreads the disease between close contacts. *Therefore, the number of 'false negatives', when obtaining patient classification results, can put many lives at risk.*

Other challenging issues include limited availability of medical and research data, therefore, to date, there is no effective treatment to prevent the COVID-19 disease. Effective and timely diagnosis can prevent the tendency as well as spread of the disease. Furthermore, rapid diagnosis, quarantine and integrated interventions will have a major impact on the trends in disease outbreaks in the future. *Automatic and early diagnosis of disease is a quick option to prevent disease in people.*

Given the challenges, the primary motivations of the research study are:

- Automated detection and classification of COVID-19 disease using advanced Dense CNN from chest CT images.
- Effective diagnosis with fewer 'false-negatives' to prevent further spread of the disease.
- Prompt and effective diagnosis of COVID-19 disease to reduce the incidence of disease.
- Build and test models on novel and large datasets.
- Use a novel loss function for faster convergence and improved performance.

1.2. Research questions

Researchers have been inspired to find ways to identify diseases with the help of medical imaging. This study answers four key research questions (RQ):

- RQ1. What is the performance of different DL methods in predicting COVID-19?
- RQ2. Which DL method significantly reduces the incidence/exacerbation of the disease?
- RQ3. Which outcome parameters are important for disease characteristics?
- RQ4. Does the proposed method improve the performance of disease diagnosis (classification and prediction)?

1.3. Paper organization

The rest of the work is organized as follows: The related work and state of the art are presented in Section-2. The dataset description, proposed framework, model and approach are all covered in section-3. Section-4 contains findings and discussion. Section-5 brings the paper to a close with future directions.

2. Related work

CNN is the most successful algorithm that has proven its ability to diagnose disease from medical images with high accuracy. There are actually three major strategies that CNNs use to successfully classify medical images:

1. Training the new CNN from the start.
2. Utilizing the features from CNN, that have been pre-trained and undertaking unsupervised training and fine tuning.
3. Transfer Learning: In this strategy, medical image analysis tasks are used to tune the pre-trained CNN model.

Most of the literature uses techniques 2 or 3 above. Here, some well-known references are reviewed.

The authors [11] compiled a freely available dataset of CT scan images of real patients from hospitals of Sao Paulo, Brazil. The dataset encourages research and development of AI methods to identify a person infected with SARS-CoV-2. The authors used the eXplainable DL approach (xDNN) as a baseline to obtain results based on this dataset and achieved an F1 score of 97.31% which is quite encouraging.

In a multi-center case study, the authors [12] found that the overall accuracy of the DL model was 86.7 percent for three groups: 'Influenza-A viral pneumonia', COVID-19, and healthy subjects. In [12], a total of 618 CT images are first classified into three groups and then a ResNet-18 and location-attention classification model is applied and subsequently the Noisy or Bayesian function is applied to calculate the confidence score.

The authors [13] used 2D and 3D DL image analysis systems for the classification of COVID versus non-COVID disease patients. The system presents ambiguity measurements quantitatively and visually as an output. A total of 157 international (China and the US) patients were tested as part of the study. They achieved 95% accuracy.

The authors [8] proposed 3 CNN based models: Inception-ResNetV2, InceptionV3, and ResNet50 for the detection of coronavirus pneumonia infected patient using chest X-ray radiographs. For the experiment, 100 images of chest X-rays have been used, of which 50 are of normal cases and 50 are of COVID-19 patients. The results showed that the ResNet50 model provided the highest classification accuracy (98%) for very small test sets.

The authors [14] proposed the Inception migration-learning model and used 453 CT images of COVID-19 confirmed cases and patients diagnosed with typical viral pneumonia. They achieved a validation accuracy of 82.9% with a specificity of 80.5% and a sensitivity of 84%. Whereas, the accuracy achieved during testing on external test data is 73.1%.

The authors [6] used deep transfer learning and top-2 smooth loss function to classify COVID-19 infected patients and achieved a validation accuracy of 93%.

In work [15] a diagnostic method has been proposed in which a modified k-means algorithm is used to segment regions in medical
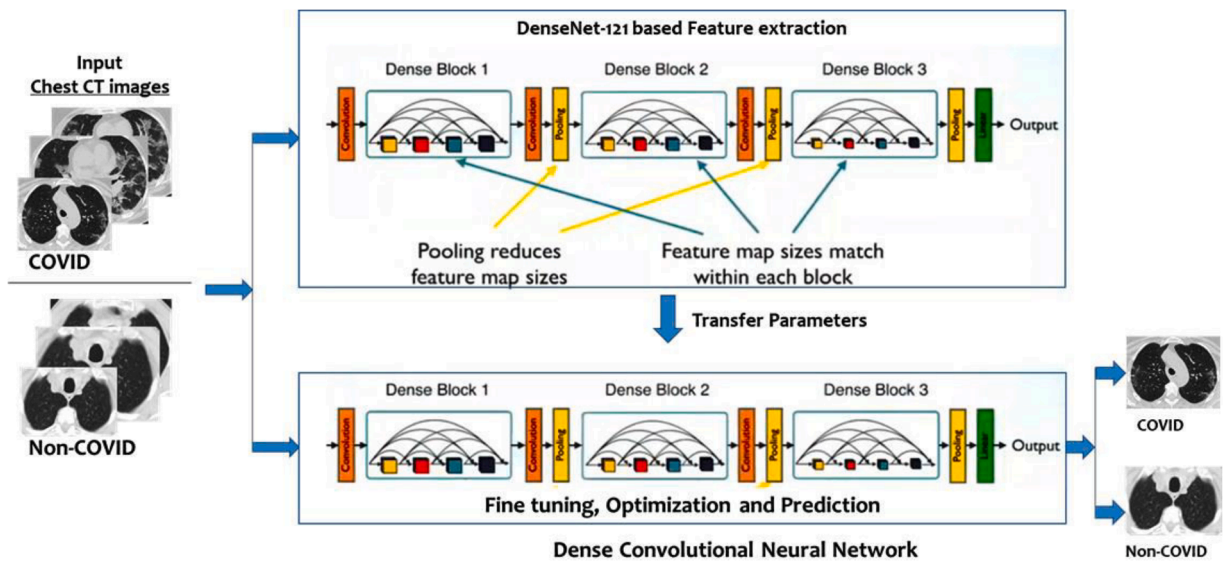
Fig. 3. Proposed Dense CNN and Transfer Learning based framework for COVID-19 prediction.

images. Support Vector Machine (SVM) and Radial Basis Function (RBF) have been used for the final classification.

In work [16] ResNet based on modified residual network based scheme was used for classification of COVID-19 patients with the help of chest X-ray images. The authors developed a multi-period cost function based on edge difference loss function (EDLF) and mean square error (MSE) to reduce the problem caused by over-mixing of noise and information in low-quality X-ray images.

The authors [17] used a non-supervised DL based Variable Auto-Encoder (UDL-VAE) model incorporating the Adaptive Wiener Filtering (AWF) based technique for pre-processing to enhance the image quality. UDL-VAE does not support any convergence time criterion.

The authors [18] propose a DL network where the pooling layer is a combination of pooling and Squeeze excitation block. To optimize convergence time and performance, batch normalization and Mish functions have been used.

The authors [19] proposed the CheXImageNet architecture for the detection of COVID-19 using digital images of chest X-rays trained with available images from open access datasets.

In another work, the authors [20] used six deep-learning techniques including ResNet and DenseNet to detect COVID-19 and other cardiovascular diseases (CVDs), where DenseNet201 outperformed other models.

From extensive studies, it has been found that CNN can provide a better clinical diagnosis before pathogen testing in case of COVID-19, thus saving significant diagnostic time which saves lives and spread of the disease. Therefore, the primary motivation of this study is to harness the concept of transfer learning with Dense CNN for COVID-19 patient classification and prediction using chest CT images.

2.1. State of the art

However, it has already been proven that CT-images of the chest can be used to diagnose the disease. Information extraction from images has to be done from the spatial local correlation that lies within its pixel/voxel. CNN is a class of DL models and has become a key element of many vision based systems. This is demonstrated using a range of popular and notable architectures: VGGNet, AlexNet, ResNet [21], Inception and DenseNet [22]. Although these CNNs were initially used for image classification, these designs also perform data mining tasks in other domains.

First, AlexNet, then VGGNet, ResNet and other representative CNNs demonstrated performance in the "ImageNet Large Scale Visual Recognition Challenge" (ILSVRC). The DenseNet model outperforms the state-of-the-art (SOTA), while requiring only fewer parameters and computation operations. For extracting features from parasitic and uninfected cells, DenseNet-121 received the award for best paper in CVPR 2017. In this paper, a data analysis framework is proposed that adopts enhanced DenseNet model, transfer learning and novel cross entropy optimization function. Unlike earlier DL systems, each network layer has direct access to the gradient, which boosts feature reuse. To promote the convergence of the algorithm, a new loss function has been implemented.

The main contributions of the research study are:

- An effective Dense CNN model (DenseNet) and transfer learning has been proposed for automated categorization of COVID-19 disease, utilizing chest CT images.
- Novel Cross Entropy Loss function for faster convergence.
- Fewer 'false-negatives' achieved to prevent disease progression and spread.
- Better validation accuracy achieved.
- The proposed Model is built and tested on a recently published large dataset.

Algorithm 1

Dense CNN based COVID-19 classification model.

Step 1. Input CT-scan images of COVID-19 (+) and COVID-19 (-) cases from dataset

Step 2. Randomly Split the train and test dataset.
Step 3. Pre-process Images
Step 3.1 Read image
Step 3.2 Resize image and reshape as per CNN
Step 4. Pass the image to Input Convolutional layer
Step 5. Input to Dense Block / Layer of DenseNet network
Step 5.1 Extract feature-map of potential features
Step 6. Apply Transition Layer (Conv. + Pooling) to reduce feature-map
Step 7. Input extracted features to Next Dense Block / Layer
Step 8. Repeat steps 5-7 till classification layer reached.
Step 9. Apply Novel Loss Function, Optimize, Tune Hyper-parameters
Step 10. Transfer the learned parameters
Step 11. Apply Prediction Model to unseen data of COVID-19 CT images
Step 12. Return Model score

Layers	Output Size	DenseNet-121		DenseNet-169		DenseNet-201		DenseNet-264	
Convolution	112 × 112	7 × 7 conv, stride 2							
Pooling	56 × 56	3 × 3 max pool, stride 2							
Dense Block (1)	56 × 56	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 6	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 6	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 6	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 6
Transition Layer (1)	56 × 56	1 × 1 conv							
	28 × 28	2 × 2 average pool, stride 2							
Dense Block (2)	28 × 28	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 12	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 12	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 12	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 12
Transition Layer (2)	28 × 28	1 × 1 conv							
	14 × 14	2 × 2 average pool, stride 2							
Dense Block (3)	14 × 14	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 24	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 32	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 48	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 64
Transition Layer (3)	14 × 14	1 × 1 conv							
	7 × 7	2 × 2 average pool, stride 2							
Dense Block (4)	7 × 7	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 16	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 32	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 32	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix}$	× 48
Classification	1 × 1	7 × 7 global average pool							
Layer		1000D fully-connected, softmax							

Fig. 4. DenseNet implementation details, feature maps (k) = 32, for DenseNet-121.

- Comparison of proposed Model with SOTA models.

3. Materials and methods

3.1. Dataset

The proposed framework employs 2482 CT scan images of the SARS-CoV-2 CT dataset collected from real patients in hospitals from Sao Paulo, Brazil [11]. The dataset includes 1252 CT scans of patients diagnosed with SARS-CoV-2 infection (COVID-19) and 1230 CT scans of non-COVID patients, but these patients were suffering from other pulmonary diseases. Figs. 1 and 2 depict some CT scans of SARS-CoV-2 infected and Non-SARS-CoV-2 patients. For ethical reasons, the detailed characteristics of the patients have been omitted.

3.2. Proposed framework: Dense CNN with transfer learning

DenseNet achieves significant improvements in various benchmark tasks, with less memory and computation required to achieve higher performance [22]. In this paper, DenseNet is considered with Transfer Learning. Fig. 3 shows the proposed framework in an illustrative manner. The algorithmic flow of the model appears in Algorithm 1.

Table 1
Model parameters.

Model parameters	Values
DenseNet weights	'imagenet'
DropOut	0.5
Activation at Dense layer	ReLU
Activation at Output	Softmax
Loss Function	Novel Cross Entropy (Proposed)
Optimizer	Adam
Learning Rate (l)	0.002
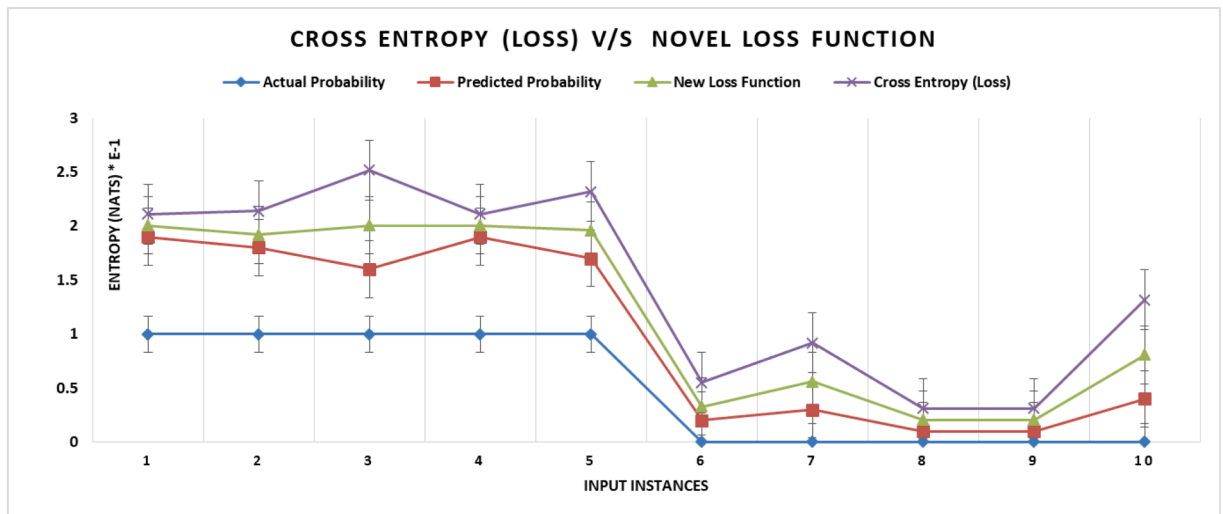
Fig. 5. Comparison of Losses with Cross Entropy Loss and Novel Loss function.

Prediction Accuracy = (True results) / (Total Cases)
= (TP+TN) / (TP+TN+FP+FN)

Precision = TP / (TP+FP)

Recall or Sensitivity = TP / (TP+FN)

F-score = 2* (Precision*Recall) / (Precision+Recall)

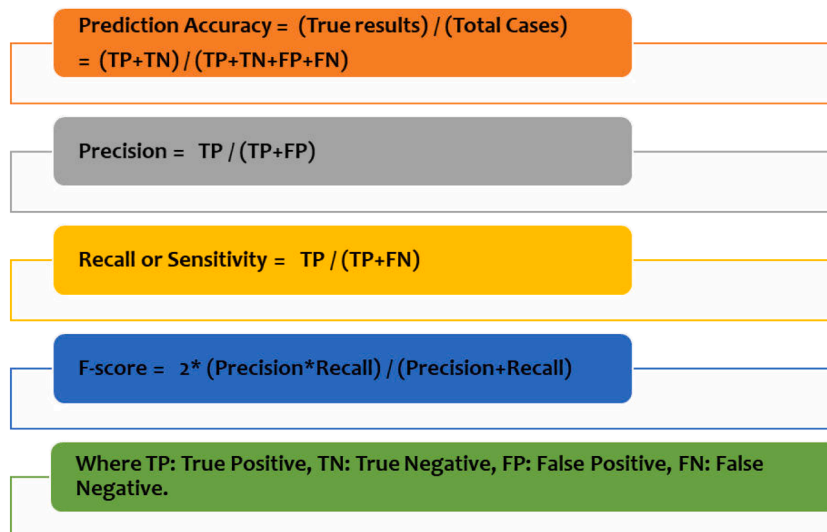Where TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative.

Fig. 6. Performance evaluation parameters.

3.3. Model architecture

We use the insight in this paper that DenseNet links each layer to every other layer in a feed-forward fashion. These networks have $L*(L+1)/2$ direct connections compared to L connections in traditional CNNs; where L represents the count of layers. Each layer in the DenseNet network adopts its own feature-maps as well as the feature-maps of all preceding layers. The following features differentiate
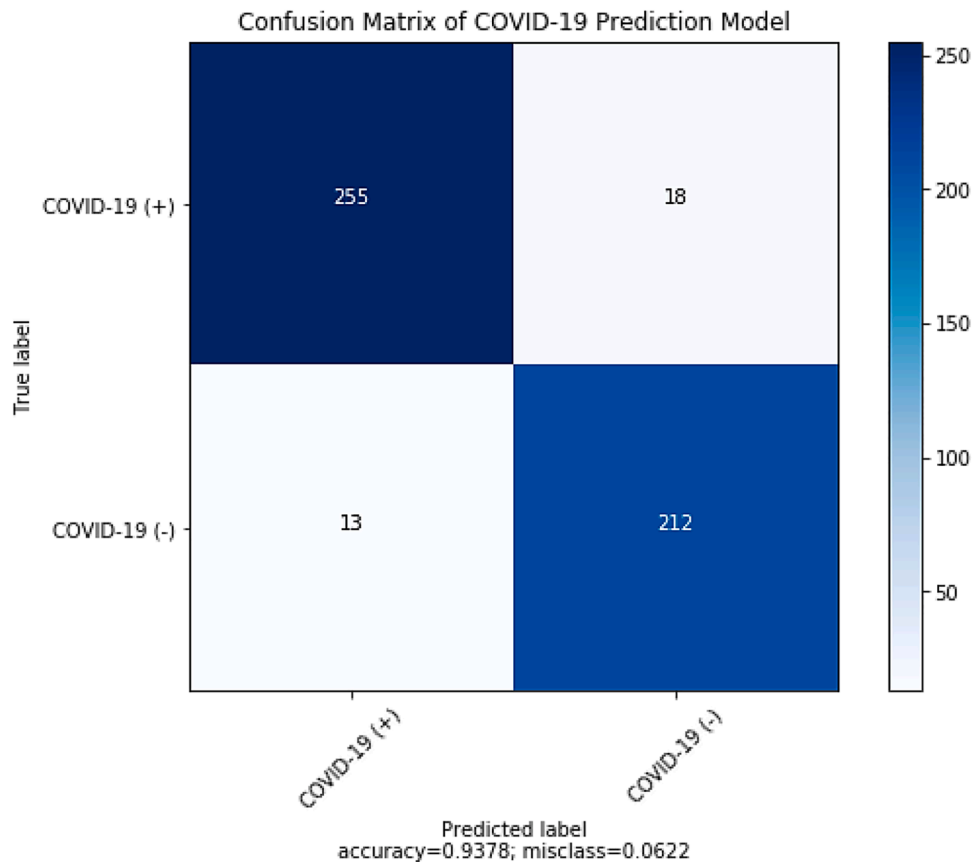
Fig. 7. Confusion matrix for performance evaluation of proposed prediction model.

Table 2
Performance analysis of covid-19 prediction models using key metrics.

S. N.	Models	TP	FP	TN	FN	Precision	Sensitivity	Specificity	F-Score	Accuracy (%)
1	ResNet-50 [21]	199	99	126	74	0.667	0.728	0.56	0.696	65.26
2	ResNet-18 with Noisy-or Bayesian Function [12]	236	17	208	37	0.813	0.867	0.924	0.84	89.15
3	ResNet-50 and Cross validation [28]	213	16	209	60	0.93	0.78	0.928	0.85	84.73
4	Inception Migration [8]	202	75	150	71	0.6	0.74	0.667	0.63	70.68
5	MobileNet [29]	247	43	182	26	0.85	0.904	0.80	0.876	86.14
6	D-CNN with binary cross entropy	241	15	210	32	0.941	0.88	0.933	0.91	90.54
7	**This study**	**255**	**13**	**212**	**18**	**0.951**	**0.934**	**0.942**	**0.942**	**93.78**

DenseNet from other variants of CNN:

- encourage feature reuse,
- lessen the 'vanishing-gradient' problem,
- strengthen feature proliferation, and
- Substantially lessen the dimension.

The architectural configuration of DenseNet is presented in Fig. 4. The values of the major model parameters are listed in Table 1. To implement the framework model, we have used Keras, built on top of Google TensorFlow. Keras is a high level API that provides an easy way to build DNNs.

We have used DenseNet which was previously trained on ImageNet [23] dataset which is a huge annotated image data set. Various ImageNet-based pre-trained CNN models extract high-level features. CNNs provide effective solutions to the generalization error that might occur due to the difference between minority and majority classes via transfer learning where tens of million attributes (of CNN architectures) need to be trained.
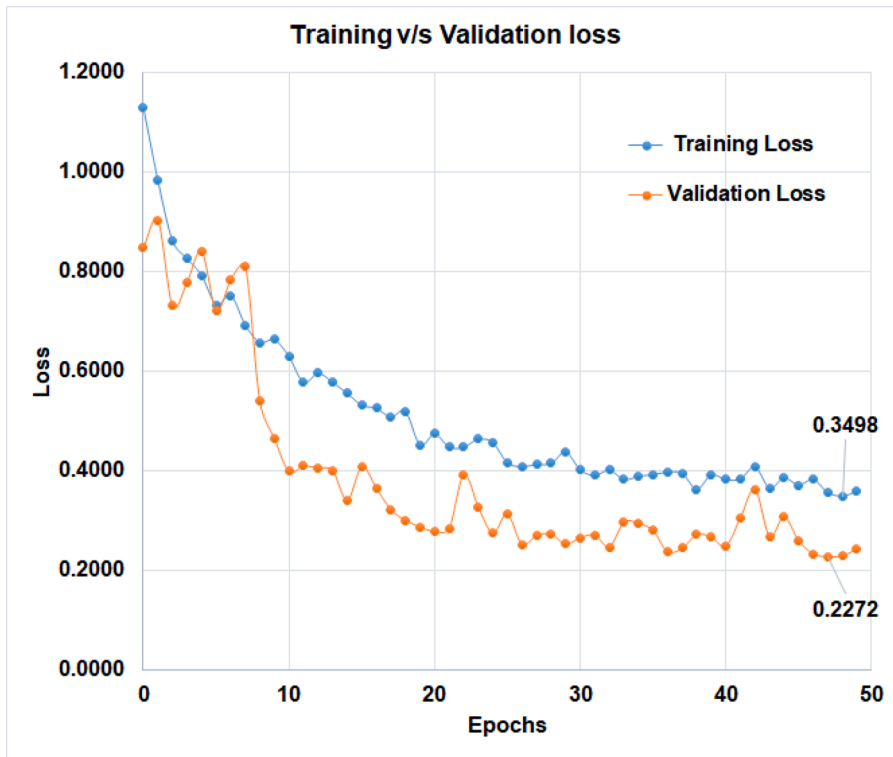
Fig. 8. Comparison of training and validation loss of D-CNN with binary cross entropy.

Dropout performs an adaptive regularization and was introduced by Hinton et al. [24]. It randomly drops a subset of features at each iteration of a training process and provides a way to control overfitting. The dropout value varies in the range of 0.1 - 0.9 according to the classification problem and the data set. It is generally set to 0.5.

The 'Non-Linear Activation Function (NLAF)-ReLU' is preferred for its effectiveness in classifying pathological images compared to conventional NLAF. So, for the last stage, usually NLAF - ReLU is selected.

Softmax is a math function that transforms a vector of numerical values into a vector of probabilities $\in [0, 1]$, where the probabilities are proportional to the values in vector on a relative scale.

We have chosen Adaptive learning rate optimization algorithm (Adam) which is a stochastic optimization method. Adam is well suitable for the problems mentioned below: (a) high noise and/or sparse gradients; (b) huge data and/or parameters; (c) non-stationary objectives. Additionally, Adam is computationally efficient, simple to implement and requires less memory.

In short, the learning rate is a regulated hyper-parameter used to train a NN and has a positive range of 0.0 and 1.0. Multiple learning rates (l) can be simulated during training to get an effective result. When there is a progressive development in the DL training phase, it is appropriate to set a restricted (l). Here, we have simulated for different values and narrowed the value to 0.002 for optimal training results.

3.4. Novel cross entropy loss

Large-batch "stochastic gradient descent" (SGD) is widely used for training in DL because of its training-time efficiency. SGD has poor generalization and it easily converges to sharp minima rather than good minima. To address the problem, we propose a novel loss function based on cross entropy loss to effectively smooth out sharp minima created by other loss functions such as cross entropy loss. The proposed loss function gives a smoothing effect. Then, we have used Adam [25] optimization for hyper-parameter tuning. In CNN it optimizes the loss function.

The cross entropy loss for binary classification task is given in Eq. (2). Whereas $y_i and (\widehat{y})$ are actual and predicted probabilities.

$$E(W) = -\sum\nolimits_{k}^{i=1} y_i \log(\widehat{y}) \tag{2}$$

For proposed model, if the cross entropy $E(W)_{i+1}$ of subsequent inputs is greater than $E(W)_{i=1}$ i.e. of first input instance, then cross entropy is updated using the Eq. (3).

$$z_i = \left[y_i \log(\widehat{y}) - E(W)_1\right] \text{ if } y_i \log(\widehat{y}) > E(W); \text{ where } i = 2...N \tag{3}$$

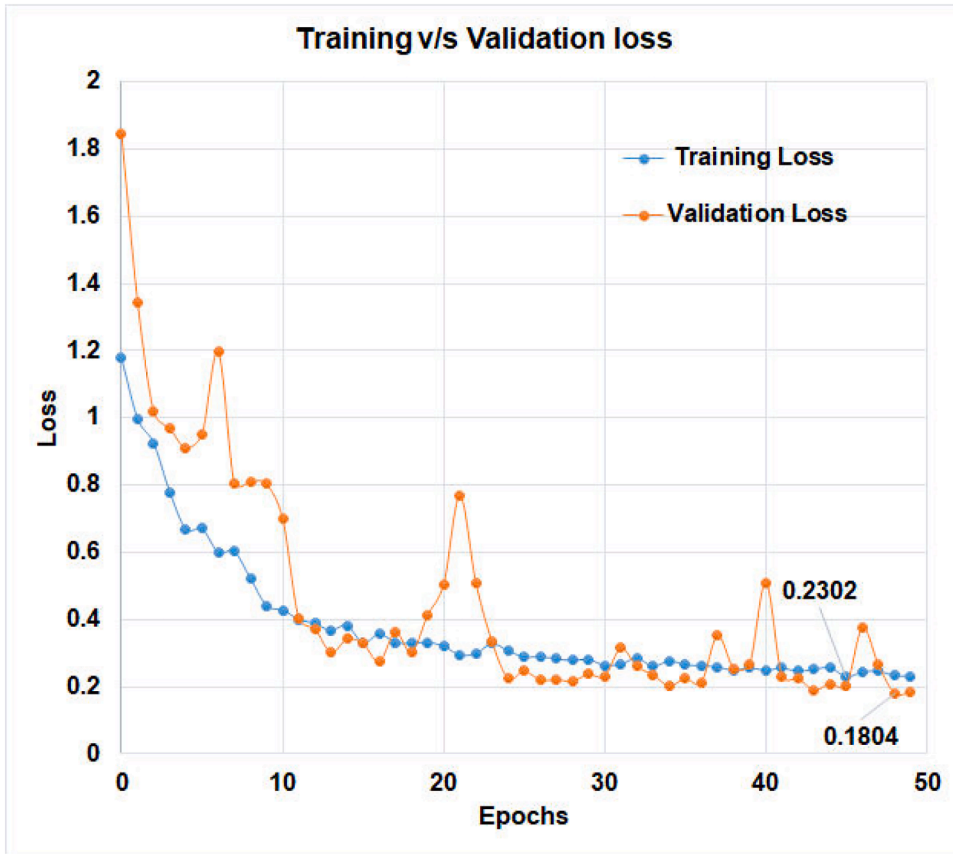Then, using the newer entropy vector z, the novel loss function is represented as:

Fig. 9. Comparison of training and validation loss of D-CNN with novel cross entropy.

$$E(W) = - \sum_{i=1}^{k} z_i \tag{4}$$

Then computing the loss function using (4) will convert the sharp minima to the good minima. The novel loss / cost function gradually improves the performance of optimization algorithms which further improves convergence of DL algorithm.

Illustrative example. Given two vectors with actual and predicted probabilities:

Actual = [1, 1, 1, 1, 1, 0, 0, 0, 0]

Predicted = [0.9, 0.8, 0.6, 0.9, 0.7, 0.2, 0.3, 0.1, 0.1, 0.4]

Cross entropy vector obtained with Eq. (2) is

[0.105, 0.223, 0.511, 0.105, 0.357, 0.223, 0.357, 0.105, 0.105, 0.511]

Now calculating the average cross entropy loss which is (refer Eq. (5))

$$E(W) = -\frac{1}{N} \sum_{k}^{i-1} y_i \log(\widehat{y}) \tag{5}$$

$$E(W) = -\frac{1}{10} \sum_{i=1}^{k} \begin{pmatrix} 0.105, & 0.223, & 0.511, & 0.105, & 0.357, \\ 0.223, & 0.357, & 0.105, & 0.105, & 0.511 \end{pmatrix} = 0.260 \ nats$$

Now, applying Eq. (3), for instances where $y_i \log(\widehat{y}) > E(W)_{i=1}$ to derive new cross entropy vector

The new vector obtained with values of cross entropy z_i is

[0.105, 0.118, 0.406, 0.105, 0.261, 0.127, 0.261, 0.105, 0.105, 0.406].

Then again calculating the average of $E(W)$ with vector of z_i values

$$E(W) = -\frac{1}{10} \sum_{i=1}^{k} \begin{pmatrix} 0.105, & 0.118, & 0.406, & 0.105, & 0.261, \\ 0.127, & 0.261, & 0.105, & 0.105, & 0.406 \end{pmatrix} = 0.199 \ nats$$

Here we can see that the average loss has come down significantly. Fig. 5 represent the Cross Entropy Loss v/s Loss with novel loss function. Also the loss is shown in relation to the actual and predicted probabilities. Our proposed model uses the newly proposed loss
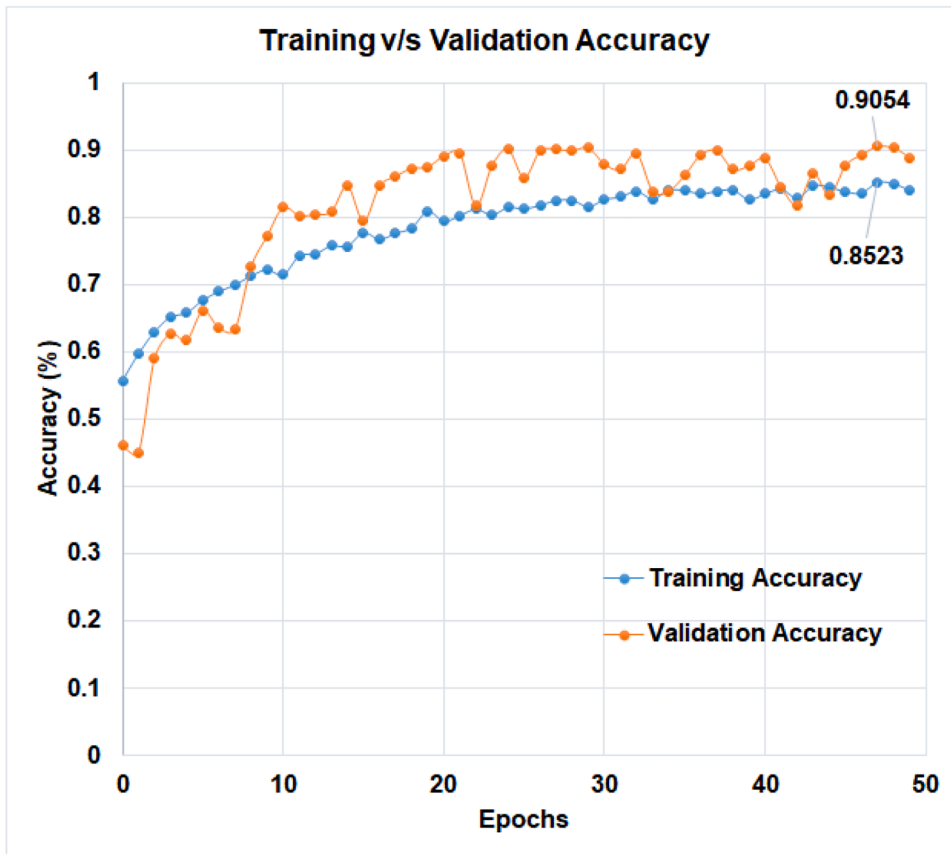
Fig. 10. Comparison of training and validation accuracy of D-CNN with binary cross entropy.

function above [26,27].

The proposed DL method with the novel loss function in this work can be further extended to other image classification tasks. Furthermore, the method can be adapted for automated medical image analysis on point-of-care devices because of its potential benefits. The method will benefit in terms of minimizing losses to a small number and thus helps in converging the optimization to a good minimum.

4. Experimental results and discussion

We have chosen the metrics that best express the performance of the predictive model. The best evaluation parameters F-score, accuracy, precision and sensitivity are depicted in Fig. 6. The F-score and accuracy are important to compare and prove the model efficacy of the COVID-19 prediction model. All parameters have been calculated using the resulting confusion matrix (see Fig. 7).

4.1. Performance evaluation and analysis

A detailed performance analysis of the recently published models and the proposed model has been performed, and is presented in Table 2. Since the dataset is sufficiently large and the disease is severe, the training-test split is 80:20. The number of epochs taken is 50. In the random division of the training and test sets, we got 273 CT images of COVID and 225 CT images of non-COVID in the test set. In this study, experiments have been performed 5 times with a 20% hold-out set and each time a different hold-out set is randomly selected.

The learning curve, which is calculated using the training dataset, reflects how effectively the model is learning. The learning curve obtained from the hold-out validation dataset can be used to estimate how well the model generalizes. Optimization learning curves are being optimized on metrics such as loss. Learning curves have been evaluated and selected based on metrics such as accuracy. The training v/s validation loss of 'D-CNN with binary cross entropy loss', and the training v/s validation loss of 'D-CNN with novel cross entropy loss' has been shown in Figs. 8 and 9 respectively. Also, the training v/s validation accuracy of 'D-CNN with binary cross entropy loss', and the training v/s validation accuracy of 'D-CNN with novel cross entropy loss' has been shown in Figs. 10 and 11respectively. Results graphs (Figs. 8-11) show that the model is a well-fitted model.

The comparison of False-Negatives is shown in Fig. 12. Figs. 13 and 14 represent the F-score and accuracy, respectively. It is clear
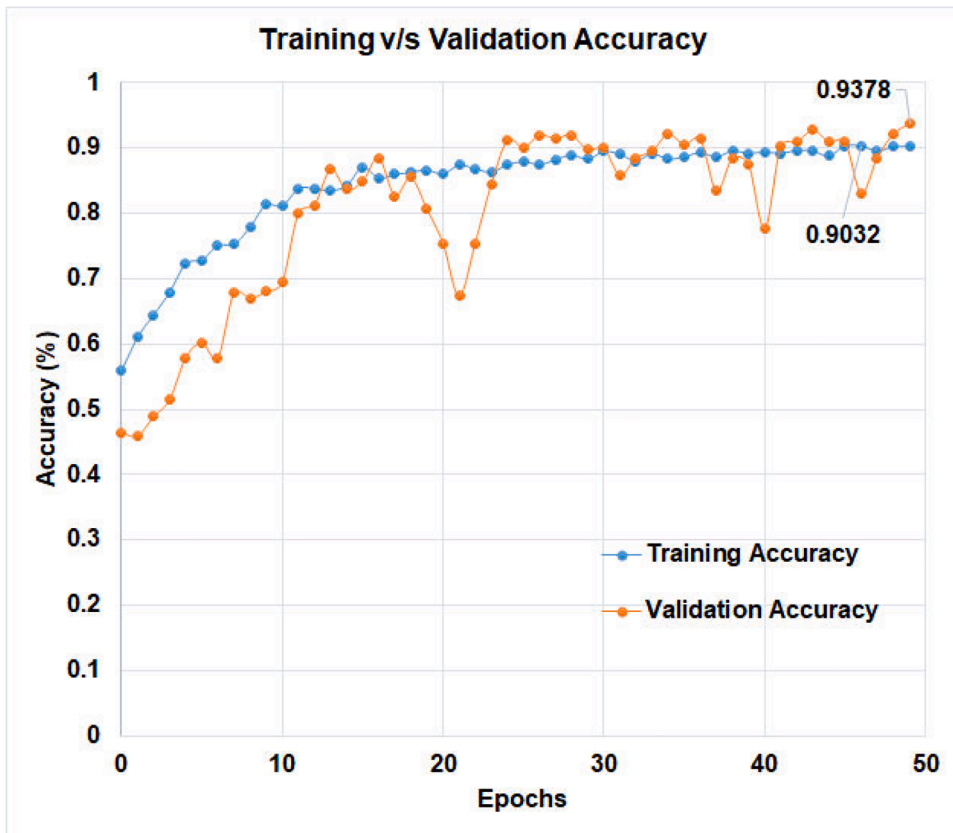
Fig. 11. Comparison of training and validation accuracy of D-CNN with novel cross entropy.
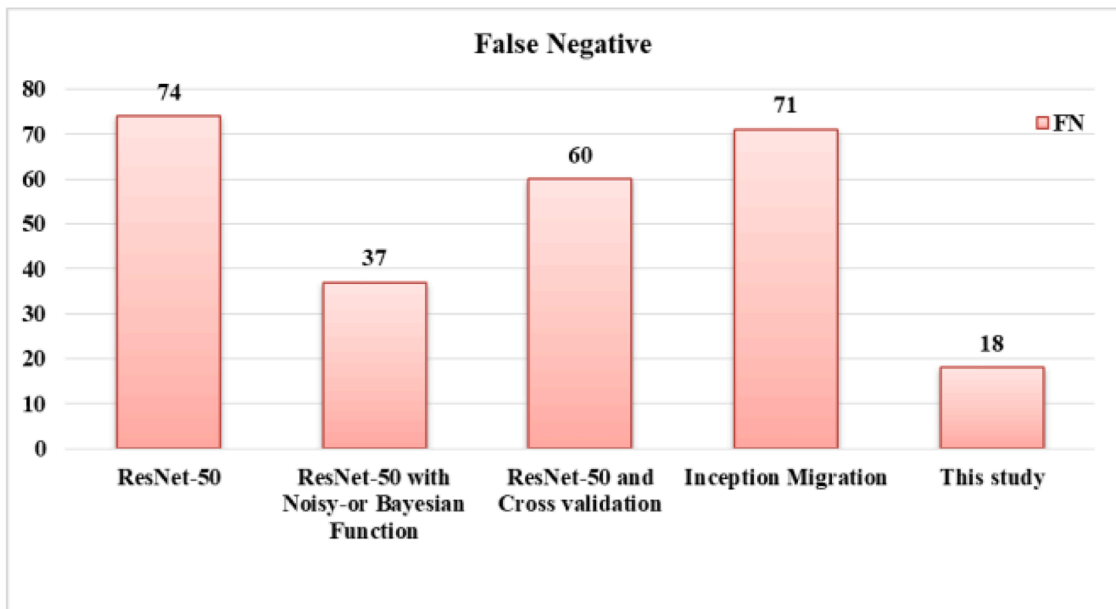

Fig. 12. Comparison of FN of proposed COVID-19 prediction model with known models.
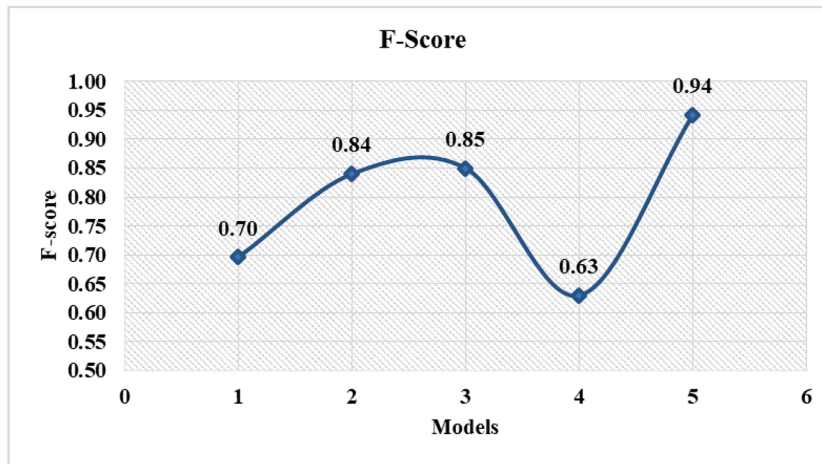
Fig. 13. Comparison of F-score of proposed COVID-19 prediction model with known models.
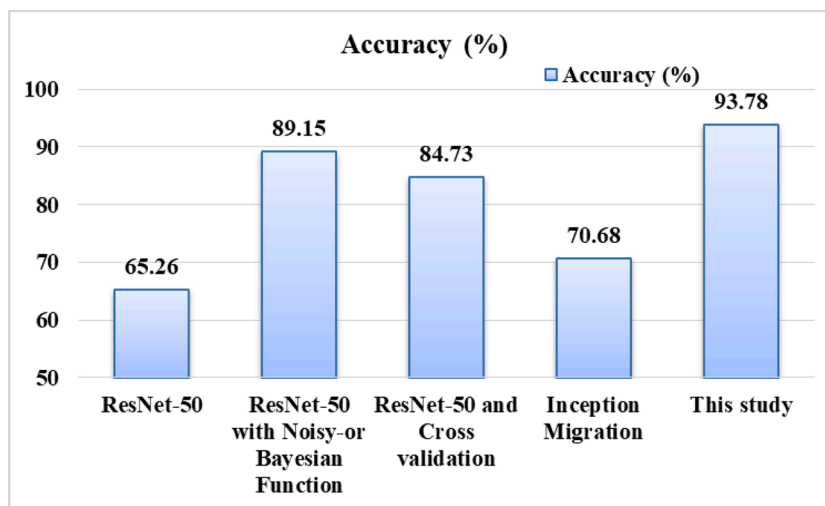

Fig. 14. Comparison of Accuracy of proposed COVID-19 prediction model with known models.

from result representation (refer Table 2) that the number of False-Negatives is significantly reduced with proposed model. In addition, important metrics: F-score and accuracy have been enhanced. Moreover, the model achieved good Precision and Sensitivity values. With the proposed model, the improvement in accuracy is 5 – 28%, while F-score is improved by 9 – 24%.

Sensitivity (recall) shows what proportion of actual positive cases have been correctly identified. Sensitivity is more important when the cost of FN is high. Thus sensitivity (recall) evaluates the proposed model best out of all the parameters. For example, in the detection of COVID-19, if someone has the disease but his test is negative, the patient will miss the best time for treatment and this will lead to further spread of the disease. This study achieved the highest recall (93.4%).

Precision is one of the more valuable metrics in COVID-19 testing. This ensures that healthy people do not have to be diagnosed as positive. Ensuring high precision will ensure careful use of health care resources and prevent exposure of healthy people to the potentially harmful environment of hospitals and healthcare centres. This study achieved the highest precision (95.1%).

Specificity has been assessed to determine the effectiveness of the proposed method in the determination of a non-COVID patient. It is the ratio of the number of TNs to the total number of persons without the disease (non-COVID). Significant results have been obtained in terms of specificity with the proposed method.

4.2. Comparative analysis

With the referenced dataset [11], we achieved considerable improvement in accuracy and F-score values using the proposed model. Although we have presented a comparative analysis of recent research with the proposed work, Table 3 represents the authors, the methods employed, the experiment information and their obtained results.

Table 3

Comparative analysis of reference works on covid-19.

Authors	Methods	Experiment details	Findings
Xu X. et al. [12]	3D DL model and Noisy-or Bayesian function.	CT image set of 618 people. 219 COVID-19, 214 Influenza-A and 175 healthy people.	Overall Accuracy = 86.7%
Gozes O. et al. [13]	Pre-trained Deep CNN (ResNet-50) model for 2D and 3D volume display.	Thoracic CT images of 157 international patients (China and U.S)	AUC = 0.996, sensitivity = 98.2% and specificity = 92.2%
Hall LO. et al. [28]	Pre-trained Deep CNN (ResNet-50) 10-fold cross validation	Tuned on Chest X-ray of 102 COVID-19 cases and 102 other pneumonia cases. Test set of 33 new COVID-19 and 218 pneumonia cases	Overall Accuracy = 91.4% TP rate = 0.78 TN rate = 0.93
Wang S. et al. [14]	Inception migration-learning model	453 CT images of COVID-19 confirmed cases and of viral pneumonia patients	Testing Accuracy = 73.1% Specificity of 67% and Sensitivity of 74%.
Narin A. et al. [8]	3 pre-trained models: InceptionV3, Inception-ResNetV2 and ResNet50 with k-fold cross validation	50 CT images of COVID-19 and 50 of healthy subjects; testing set is 20% i.e. 10 images only.	Accuracy of Inception-ResNetV2, ResNet-50, and InceptionV3 is 87%, 98% and 97% respectively.
This Study	Dense CNN (DenseNet-121) model with Novel Cross Entropy Loss	Big Data of 2482 CT images of real patients. 1252 COVID-19 infected and 1230 patients infected from other pulmonary diseases	**Prediction Accuracy = 93.78%** **Precision = 0.951** **Sensitivity = 0.934** **Specificity = 0.942** **F-score = 0.942**

4.3. Threat to validity

The SARS-CoV-2 virus has the ability to change over time, resulting in genetic heterogeneity in a viral population. For clinical laboratory personnel and health care providers who perform multiple diagnostic tests to diagnose COVID-19 disease, the "U.S. Food and Drug Administration (FDA)" recommends addressing potential FN results. Currently, the proposed model is trained and tested only on known dataset. In the future the proposed method can be tested on CT images of patients affected by different virus forms.

5. Conclusion and future directions

It is highly necessary that an affordable, quick and effective diagnostic technique should be developed for symptomatic and asymptomatic COVID-19 patients, which can detect the affected patient before the condition worsens. It has already been proven that DL methods can detect radiographic changes in CT images of patients infected with the SARS-CoV-2 virus. Thus, using clinical diagnostic methods prior to pathogenic tests saves time which may be important for disease control. In this research study, an enhanced framework for automatic identification and classification of COVID-19 patients using CT images based on Dense-CNN model, transfer learning and novel cross entropy loss function is proposed. The model has been developed and subsequently evaluated on a large dataset of patients. Our work mainly focuses on reducing False-Negatives which will further reduce the chances of spreading the disease to our loved ones. The prediction accuracy of proposed model is **93.78%** while False-Negative is 6.5% only, which means it effectively diagnose the COVID-19 positive cases. We have also presented the comprehensive study and comparison with well-known models. An overall improvement in accuracy, specificity and F-score values compared to recent models has also been achieved. Another important contribution of the paper is that we propose to use a loss function that effectively smooths the sharp minima so that the optimization can obtain a better minima.

The proposed model can be used as a clinical tool to diagnose patients before pathogenic tests and can save lives and spread. Further, various nature-inspired optimization methods can be applied to further improve the model accuracy and faster convergence of DL algorithms. The framework can also be integrated into health care tools to accurately classify other pulmonary and chronic diseases. Framework can be tested for Quality of Service, energy and other performance parameters in Cloud and Social Networking Service (SNS).

Research involving Human Participants and/or Animals

This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent

Informed consent was obtained from all individual participants included in the study.

Author Contributions

All authors have equally contributed to this work, read and agreed to the published version of the manuscript.

Funding

Declaration of Competing Interest

The authors declare that this manuscript has no conflict of interest with any other published source and has not been published previously (partly or in full). No data have been fabricated or manipulated to support our conclusions.

Data availability

Data will be made available on request.

References

[1] Holmes KV. SARS coronavirus: a new challenge for prevention and therapy. J Clin Invest 2003;111:1605–9.
[2] W.H. Organization. SARS (Severe Acute Respiratory Syndrome).
[3] Gorbalenya AE. Severe acute respiratory syndrome-related coronavirus–The species and its viruses, a statement of the Coronavirus Study Group. BioRxiv 2020.
[4] W.H. Organization. Coronavirus disease (COVID-19) pandemic.
[5] Sharma N, Aggarwal LM. Automated medical image segmentation techniques. J Med Phys/Ass Med Phys India 2010;35:3.
[6] Pathak Y, Shukla PK, Tiwari A, Stalin S, Singh S, Shukla PK. Deep transfer learning based classification model for COVID-19 disease. IRBM 2020.
[7] Kim Y, Ohn I, Kim D. Fast convergence rates of deep neural networks for classification. Neural Netw 2021;138:179–97.
[8] Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. arXiv preprint arXiv 2020:10849.
[9] Rasmussen SA, Smulian JC, Lednicky JA, Wen TS, Jamieson DJ. Coronavirus Disease 2019 (COVID-19) and pregnancy: what obstetricians need to know. Am J Obstet Gynecol 2020.
[10] Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. N Engl J Med 2020.
[11] E. Soares, P. Angelov, S. Biaso, M. Higa Froes, D. Kanda Abe. SARS-CoV-2 CT-scan dataset: A large dataset of real patients CT scans for SARS-CoV-2 identification. 2020. p. 2020.04.24.20078584.
[12] X. Xu, X. Jiang, C. Ma, P. Du, X. Li, S. Lv, et al. Deep learning system to screen coronavirus disease 2019 pneumonia. arXiv preprint arXiv:09334. (2020).
[13] O. Gozes, M. Frid-Adar, H. Greenspan, P.D. Browning, H. Zhang, W. Ji, et al. Rapid ai development cycle for the coronavirus (covid-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis. arXiv preprint arXiv:05037. (2020).
[14] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, et al. A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). MedRxiv. (2020).
[15] Shahin OR, Alshammari HH, Taloba AI, Abd El-Aziz RM. Machine learning approach for autonomous detection and classification of COVID-19 Virus. Comput Electr Eng 2022:108055.
[16] Ghosh SK, Ghosh A. ENResNet: A novel residual neural network for chest X-ray enhancement based COVID-19 detection. Biomed Signal Process Control 2022; 72:103286.
[17] Mansour RF, Escorcia-Gutierrez J, Gamarra M, Gupta D, Castillo O, Kumar S. Unsupervised deep learning based variational autoencoder model for COVID-19 diagnosis and classification. Pattern Recognit Lett 2021;151:267–74.
[18] JavadiMoghaddam S, Gholamalinejad H. A novel deep learning based method for COVID-19 detection from CT image. Biomed Signal Process Control 2021;70: 102987.
[19] Shastri S, Kansal I, Kumar S, Singh K, Popli R, Mansotra V. CheXImageNet: a novel architecture for accurate classification of Covid-19 with chest x-ray digital images using deep convolutional neural networks. Health Technol 2022:1–12.
[20] Rahman T, Akinbi A, Chowdhury ME, Rashid TA, Şengür A, Khandakar A, et al. COV-ECGNET: COVID-19 detection using ECG trace images with deep convolutional neural network. Health Inf Sci Syst 2022;10:1–16.
[21] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–8.
[22] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2017. p. 4700–8.
[23] Deng J, Dong W, Socher R, Li L-J, Li K, Fei-Fei L. Imagenet: a large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. IEEE; 2009. p. 248–55.
[24] G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R.R. Salakhutdinov. Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:. (2012).
[25] D.P. Kingma, J. Ba. Adam: A method for stochastic optimization. arXiv preprint arXiv:. (2014).
[26] Motwani A, Shukla PK, Pawar M. Novel framework based on deep learning and cloud analytics for smart patient monitoring and recommendation (SPMR). J Ambient Intell Humanized Comput 2021:1–16.
[27] Motwani A, Shukla PK, Pawar M. In: Smart Predictive Healthcare Framework for Remote Patient Monitoring and Recommendation Using Deep Learning with Novel Cost Optimization. International Conference on Information and Communication Technology for Intelligent Systems. Springer; 2020. p. 671–82.
[28] L.O. Hall, R. Paul, D.B. Goldgof, G.M. Goldgof. Finding covid-19 from chest x-rays using deep learning on a small dataset. arXiv preprint arXiv:02060. (2020).
[29] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:04861. (2017).

Anand Motwani is a prolific academician and researcher with 20+ years of experience in training, teaching, and corporate employees. He has worked with renowned educational institutions in senior positions and is still associated with other organizations as a consultant, trainer, and examiner for academic and research work. He is an expert in cloud computing, machine and deep learning applications.

Piyush Kumar Shukla has completed Post Doctorate Fellowship (PDF) under "Information Security Education and Awareness Project Phase II" funded by the Ministry of Electronics and IT (MeitY). He is the editor and reviewer of various prestigious SCI, SCIE, and WOS-indexed journals. He has over 300 publications in highly indexed journals and prestigious conferences, including many books.

Mahesh Pawar has the rich industrial experience, which he has used to teach various professional and programming skills. He has empowered university students with employable skills. His research interests include Cloud and Ubiquitous computing, Machine learning, Healthcare Internet of Things, Image Processing, Blockchain.

Manoj Kumar has completed his Ph.D. from The Northcap University and M.Sc. (Information Security and Digital Forensics) from Technological University Dublin (Formerly ITB Blanchardstown) Ireland in 2013. He is currently working on the post of Associate Professor in the University of Wollongong in Dubai, UAE. He has published over 115 articles in International refereed journals and conferences.

Uttam Ghosh is currently an Associate Professor with the Computer Science and Data Science, School of Applied Computational Sciences, Meharry Medical College, Nashville, TN, USA. He received the Ph.D. degree in electronics and electrical communication engineering from the Indian Institute of Technology Kharagpur, Kharagpur, India.

Waleed Alnumay received the bachelor's degree and master's degree in computer science domain, from King Saud University, and University of Atlanta, USA, and the Ph.D. degree in computer science from Oklahoma University, USA, in 2004. He is currently working as an Associate Professor at King Saud University. His main research interests include computer networks and distributed computing information-centric networking etc.

Soumya Ranjan Nayak is currently working as Assistant Professor at KiiT University, India. He received his Ph.D. degree under MHRD Govt. of India fellowship; with a preceded degree of M. Tech and B. Tech in CSE. He has published over 100 articles in peer-reviewed journals and conferences of international repute which includes pattern recognition, fractal graphics and computer vision.