Available at www.sciencedirect.com

ScienceDirect

journal homepage: www.elsevier.com/locate/bbe

ELSEVIER

Original Research Article

Automated COVID-19 detection from X-ray and CT images with stacked ensemble convolutional neural network

Check for updates

Mahesh Gour [a,*], Sweta Jain [a]

[a] Maulana Azad National Institute of Technology, Bhopal, India

A B S T R A C T

Automatic and rapid screening of COVID-19 from the radiological (X-ray or CT scan) images has become an urgent need in the current pandemic situation of SARS-CoV-2 worldwide. However, accurate and reliable screening of patients is challenging due to the discrepancy between the radiological images of COVID-19 and other viral pneumonia. So, in this paper, we design a new stacked convolutional neural network model for the automatic diagnosis of COVID-19 disease from the chest X-ray and CT images. In the proposed approach, different sub-models have been obtained from the VGG19 and the Xception models during the training. Thereafter, obtained sub-models are stacked together using softmax classifier. The proposed stacked CNN model combines the discriminating power of the different CNN's sub-models and detects COVID-19 from the radiological images. In addition, we collect CT images to build a CT image dataset and also generate an X-ray images dataset by combining X-ray images from the three publicly available data repositories. The proposed stacked CNN model achieves a sensitivity of 97.62% for the multi-class classification of X-ray images into COVID-19, Normal and Pneumonia Classes and 98.31% sensitivity for binary classification of CT images into COVID-19 and no-Finding classes. Our proposed approach shows superiority over the existing methods for the detection of the COVID-19 cases from the X-ray radiological images.

© 2021 Nalecz Institute of Biocybernetics and Biomedical Engineering of the Polish Academy of Sciences. Published by Elsevier B.V. All rights reserved.

1.   Introduction

The novel coronavirus disease 2019 (COVID-19) pandemic has put the livelihoods and health of the massive population in a critical position. It has led to a disturbance in the public life of the world population. The severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) belongs to the family of coronavirus, which gets transmitted to the people based on the infection in the form of direct contact or fomites. The primary symptoms of coronavirus infection are fever, cough, and

fatigue. In several cases, coronavirus causes severe respiratory problems like Pneumonia, lung disorders, and kidney malfunction. The virus has serious consequences as its serial interval is 5 to 7.5 days, and the reproduction rate is 2 to 3 [1] people. The coronavirus infection can incite SARS (Severe Acute Respiratory Syndrome), which might unfold serious health impacts. This pandemic has brought new challenges to the medical world. People are not getting wards, ventilators, and even there is a shortage of doctors and nurses in hospitals. It also affected the diagnosis and treatment of non-communicable diseases [2]. A critical step to fight against the COVID-19 is to identify the infected people so that they get immediate treatment and isolate them to control the further spread of the infection.

The COVID-19 panic has increased due to the unavailability of fast and accurate diagnosis systems to test the infected people. According to the World Health Organization, the diagnosis of COVID-19 cases must be confirmed by the reverse transcription-polymerase chain reaction (RT-PCR) [3]. While RT-PCR has become a standard tool for confirmation of COVID-19, but it is a very time-consuming, laborious, and manual process, and there is a limitation of availability of diagnostic kits. The availability of COVID-19 testing kits is limited as compared to the increasing amount of infected people; hence there is a need to rely on different diagnosis methodologies.

The coronavirus targets the epithelial cells that affect patient's respiratory tract, which can be analyzed by the radiological images of a patient's lungs. Some early studies also show that patients present anomalies in chest X-ray and CT scan images, which are the typical characteristics of COVID-19 infected patients [4–6]. Hence, the development of the computer-aided diagnosis system for the automatic analysis of radiological images (CT scan or X-ray) can be very helpful in identifying infected patients at a faster rate [7].

Recently, deep learning-based computer aided diagnosis (CAD) systems have shown great success in the automated detection of COVID-19 disease using chest X-ray images. Wang et al. [8] have proposed a COVID-Net model based on the projection-expansion-projection design pattern for COVID-19 cases detection from the X-ray images. Ucar et al. [9] proposed a SqueezeNet convolutional neural network (CNN) model with Bayesian optimization for the classification of chest X-ray images into normal, pneumonia and COVID-19 classes. Jain et al.[10] have applied pre-trained deep networks in two stages. In the first stage, the ResNet50 model is used to classify the X-ray images into viral-induced pneumonia, bacterial-induced pneumonia, and normal cases. Further, in the second stage, they have detected COVID-19 cases from positive viral-induced pneumonia cases. Similarly, Apostolopoulos et al. [11] have evaluated the performances of state-of-the-art pre-trained CNNs on the chest X-ray images for COVID-19 detection, and they have achieved the best performance with VGG19 and MobileNet-v2 models. Joshi et al. [12] have used YOLO-v3 based-architecture to detect the COVID-19 from the X-ray images, in which they have used DarkNet-53 as a backbone network. They have developed their method for X-ray image classification in binary class as well as in multi-classes.

In this study, a new stacked convolutional neural network has been designed for the automatic diagnosis of COVID-19 disease from the chest X-ray and CT scan images. The contribution of this work are as following:

- The proposed stacked generalization approach hypothesises that different CNN sub-models learn different non-linear discriminative features and different levels of semantic image representation. Thus, a new powerful model can be developed by incorporating the best prediction from the different CNN sub-models. Therefore, in the proposed method, different sub-models are ensemble together by using Softmax classifier to build a reliable and accurate model for COVID-19 detection.

- We fine-tune VGG19 and Xception models on the X-ray and CT images. Thereafter, different sub-models have been obtained from the Xception and VGG19 models during the training to develop a stacked ensemble model.

- We investigate the performance of six pre-trained CNN models for the detection of COVID-19 from the chest X-ray images.

- We also investigate and compare the performances of various classifiers to build a stacked ensemble model with different classifiers.

- We collect CT images of COVID-19 patients to build a CT images dataset and also generate a dataset of chest X-ray images with the combination and modification of three publicly available datasets [13–15].

The organization of this paper as follows: Section 2 presents the related work. Section 3 describes the proposed stacked CNN model. Section 4 describes the COVID19CXr and COVID19CTs datasets, details the experimental results and performance comparison. Finally, the conclusion is drawn in Section 5.

2.   Related work

Over the past 40 years, many computer-aided systems have been developed for the diagnosis of lung diseases [16], and these systems have shown promising results for automatic detecting lung abnormality from the radiological images [17,18]. Recently, automatic CAD of COVID-19 using radiological images has drawn a lot of attention of researchers, and as a result, several approaches have been introduced in the literature. They have published a series of research articles [19,20] demonstrating the CAD systems for the detection of COVID-19 using radiological images. Xu et al. [21] have studied various CNN models technically and proposed a model with the combination of 2D and 3D CNN models for the classification of the CT images into COVID-19, Influenza viral pneumonia, or no-infection. Their approach achieved a sensitivity of 98.2% and specificity of 92.2%. Shah et al. [22] have developed a CTnet-10 model to classify CT scans images into

COVID-19 and non-COVID-19 classes. They have also tested the different pre-trained CNN networks, such as DenseNet-169, VGG-16, ResNet-50, Inception-V3, and VGG-19, for COVID-19 detection. They reported the best accuracy of 94.52% with the VGG19 model. Similarly, Kassania et al. [23] have used several CNN models for feature extraction from the CT and X-ray images, and various machine learning classifiers have been applied on the extracted feature for classifying them into COVID-19 and healthy classes. They reported the best classification accuracy of 99% with DenseNet121 features and Bagging tree classifier.

An attention-based deep 3D multiple instance learning-based approach has been proposed by Han et al. [24] for automatic screening of COVID-19 from the CT images. Their algorithm achieved an accuracy of 97.9%. Afshar et al. [25] developed a capsule network-based framework for the classification of the X-ray images into Normal, bacterial, Non-COVID, and COVID-19 cases. The authors reported accuracy of 95.7% and sensitivity of 90%. Ardakani et al. [26] have presented the application of deep learning in COVID-19 detection using CT images, in which they have tested the performances of ten pre-trained CNN models. Their experiment results showed that ResNet101 achieved the best AUC of 0.99 among all pre-trained networks. Benmalek et al. [7] have compared the performances of CT scans and X-ray images for COVID-19 disease detection using different pre-trained CNN models, namely Resnet-18, Inception-V3, and MobileNet-V2. They have reported the best sensitivity of 98.6% for CT scans with ResNet-18 and the best sensitivity of 92.3% for X-ray images with Inception-V3. Mishra et al. [27] have developed an algorithm based on VGG16 and ResNet50 architectures, using the transfer learning for COVID-19 detection from the CT scans. Their proposed approach achieved an accuracy of 99% for binary classification with both VGG16 and ResNet50 models and accuracy of 88.52% for multi-class classification with ResNet50.

Similarly, Narin et al. [28] have applied ResNet50, Inception-V3 and Inception-ResNet-V2 using transfer learning for classification of the X-ray images into Normal and COVID-19 classes. This method has achieved good performance with an accuracy of 98% for ResNet50. However, the number of X-ray images is only 100, which is very less for developing deep learning models. Oh et al. [29] have proposed a patch-based approach to train the ResNet18 model using image patches that have been extracted from the chest X-ray images. For decision making, they used the majority voting strategy, which resulted in an accuracy of 88.9%. An objected detection based DarkCovidNet model has been proposed by Ozturk et al. [30] for automatic detection of COVID-19 cases from the X-ray images. They have reported an accuracy of 98.08% for binary classification of X-ray images into COVID-19 and no-findings. For the multi-class classification of X-ray images into no-findings, COVID-19, and pneumonia, their approach achieved an accuracy of 87.02%. Pereira et al. [31] have proposed a hierarchical classification approach, in which they extracted deep features by Inception-V3 and tested texture descriptors. They investigated early and late fusion techniques for combining the strength of descriptors and classifiers. Their hierarchical classification approach achieved an F1-Score of 0.89 for the COVID-19 identification in the X-ray images. Sethy et al. [32]

extracted deep features of X-ray images from the pre-trained CNN, and support vector machine (SVM) has been applied to the extracted features to classify X-ray images. The authors achieved an accuracy of 95.38% using ResNet50 with the SVM classifier.

Similarly, Minaee et al. [33] also proposed a method based on transfer learning, in which they have fine-tuned four pre-trained networks on the COVID-19 chest X-ray images. They reported a sensitivity rate of 98%. Castiglioni et al. [34] have proposed ensemble-based model in which they have ensemble ten pre-trained CNNs. Their proposed approach achieved a sensitivity of 80%. Abraham et al. [35] have used multiple pre-trained CNNs with CFS technique to extract the features from the X-ray images, and authors have applied Bayesnet classifier on the extracted features to detect COVID-19. Their proposed approach achieved an accuracy of 91.16% for the classification of X-ray images into COVID-19 and non-COVID classes. Panwar et al. [36] have proposed a deep learning-based nCOVnet model to detect the COVID-19 cases from the X-ray images. They have reported a classification accuracy of 88.10%. Nigam et al. [37] have developed a Coronavirus diagnostic system using pre-trained CNN architectures, namely VGG16, DenseNet121, Xception, NASNet, and EfficientNet-B7. Authors reported an accuracy of 93.48% with EfficientNet-B7 for the chest X-ray image classification into COVID, normal, and other classes. Ashour et al. [38] have proposed a ensemble-based bag-of-features (BoF) model for classifying chest X-ray images into normal and COVID-19 classes. They have used the grid method for determining key points and speeded up robust features (SURF) for feature extraction from the image. Their proposed model achieved a classification accuracy of 98.6%.

3.   Methodology

The convolution neural network is the driving concept of deep learning algorithms in computer vision, which led to outstanding performance in most of the pattern recognition tasks such as image classification [39–42], object localization, segmentation, and detection [43–45]. It has also shown its superiority in the medical image analysis for image classification and segmentation problems [46–49], especially in lung-related diseases such as lung nodule detection [50], pneumonia detection [51], and pulmonary tuberculosis [52]. CNN automatically learns a low to the high level of useful feature representations and integrates feature extraction and classification stages in a single pipeline, which is trainable in an end-to-end manner without requiring any manual design and expert human intervention.

In this work, we have developed a deep learning-based stacked convolutional neural network for the rapid screening of COVID-19 patients using X-ray images. This study is a continuation and extension of the considerations presented in the preprint [53] publication from the Internet. The proposed COVID-19 detection method includes three modules, as shown in Fig. 1. In the first module, a pre-trained VGG19 [39] and Xception [54] models are fine-tuned on radiological images for the diagnosis of COVID-19 disease. In the second module, five CNN's sub-models are obtained during the training of Xception and VGG19 models. The outputs of CNN's
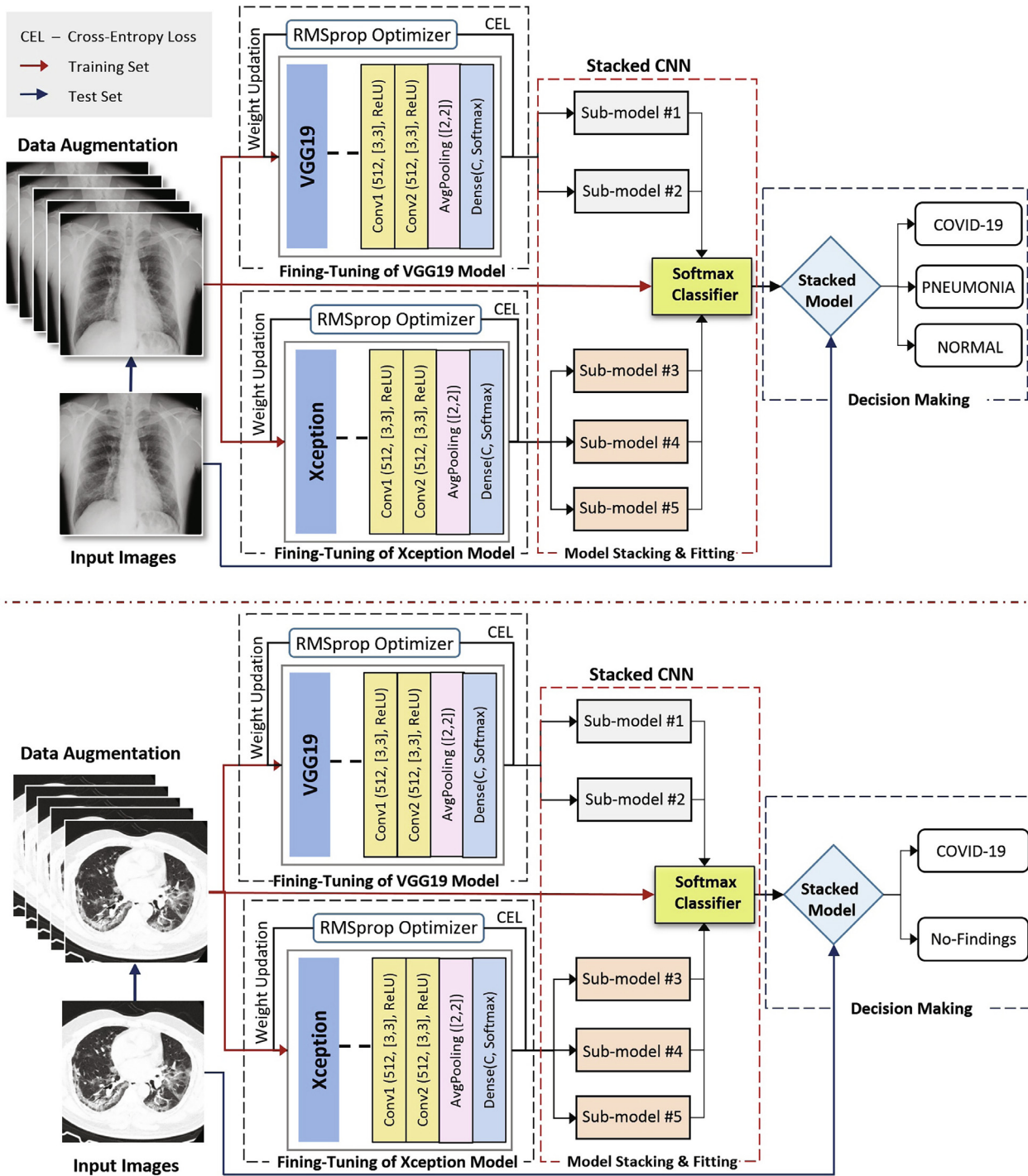
Fig. 1 – Block diagram of proposed Stacked CNN model.

sub-models are stacked together by applying softmax classifier for building a new final model for diagnosis of COVID-19 disease from CT and X-ray images. A detailed description of the proposed approach is given in the following section:

3.1.   Xception Model and Training

The Xception (Extreme Inception) model [54] is a state-of-the-art CNN architecture for image classification, developed by

Google. It achieved outstanding results for image classification, and it has outperformed the Inception-V3 on both ImageNet ILSVRC and JFT datasets. Xception model is the 1st runner-up in ILSVRC 2015 competition. The architecture of Xception CNN consists a linear stack of depthwise separable convolution layer (a depthwise convolution followed by a pointwise convolution) with residual connections.

Top layer (Fully-connected layer and Softmax layer) of pre-trained Xception model is replaced with the new top layer for

fine-tuning it on the radiological images. Xception model has been trained on the chest X-ray and CT images in a supervised manner. The cross-entropy loss function is used to calculate the training error and which is minimized using the RMSprop optimizer [55]. The cross-entropy loss function is mathematically represented in Eqn. (1).

$$J(T,P) = -\sum_{i=1}^{C} t_i \log(p_i) \tag{1}$$

where t_i and p_i are the target value and predicted probability respectively, for each class /in C.

In the experiment, the hyper-parameter values are set as follows: learning rate to 0.0001, the batch size to 16, and dropout probability to 0.15. We experimentally find that these are the best suitable values of hyper-parameters for network training.

3.2.   VGG19 Model and Training

The VGG19 is a pre-trained network that is trained on the ImageNet dataset, which achieved state-of-the-art performance on ILSVRC Challenge 2014. It also achieves outstanding performance on other image recognition datasets. Hence, we have also used VGG19 along with Xception model for generating sub-models.

To fine-tune the VGG19 on X-ray images, the top layers (Fully-connected layer and Softmax layer) of the VGG19 network are removed. We added new layers such as two Convolutional layers with ReLU activation, a Global Average Pooling layer, a Fully-connected layer, and a Softmax layer at the top of the VGG19 network. We have used same hyperparameter values as have used for in the training of Xception model.

3.3.   Stacked Convolutional Neural Network

Stacked generalization [56] is an ensemble approach in which a new model learns how to incorporate the best predictions of multiple existing models. The proposed approach hypothesized that different CNN's sub-models learn non-linear discriminative features and semantic image representation from the images at different levels. Thus a stacked ensemble CNN model will be generalized and highly accurate. This section describes the proposed stacked convolutional neural network. The pseudo-code of the sub-models generation process is given in Algorithm 1. In this process, the COVID19CXr dataset is divided into a training set, validation set, and test set. The Xception and VGG19 are trained on chest X-ray images of the training set for the 3192 iterations. During training of VGG19, we have extracted the first *sub-model#1* after 1596 iterations and second *sub-model#2* after completion of the training. Similarly, during fine-tuning of Xception, we have extracted *sub-model#3* after 1064 iterations, *sub-model#4* after 2128 iterations, and *sub-model#5* at last. To deal with the class imbalance problem, we have assigned class weights while training of the networks. In this process, class weight in ratio

of 20:1:1 is assigned to COVID-19, Pneumonia, and Normal class, respectively.

Algorithm 1: Sub-model Generation process

Input: X-ray images of the chest
Output: sub-models
1: Divide the dataset into a training set, validation set, and test set.
2: Apply data augmentation on the training set.
3: Train VGG19 and Xception, and generate sub-models:
Initialisation: class_weight = [0:20, 1:1, 2:1]
4: **for** i = 1 to N **do**
5:    Train(VGG19, train_img, img_label, class_weight)
6:    **if** (i == l1) **then**
7:       sub-model#1 = save(VGG19)
8:    **end if**
9: **end for**
10: sub-model#2 = save(VGG19)
11: **for** i = 1 to M **do**
12:    Train(Xception, train_img, img_label, class_weight)
13:    **if** (i == l2) **then**
14:       sub-model#3 = save(Xception)
15:    **else**
16:       **if** (i == l3) **then**
17:          sub-model#4 = save(Xception)
18:       **end if**
19:    **end if**
20: **end for**
21: sub-model#5 = save(Xception)
22: **return** *sub-models*

Where N and M represent the total number of iterations for training VGG19 and Xception, respectively, and l1, l2 and l3 are the constant.

The performance of the sub-models varies across complex CAD systems, and it is reasonable to combine the strengths of sub-models, which might result in increased overall accuracy. Hence, we combined the sub-models predictions by applying softmax classifier [57] to build a highly accurate and reliable generalized model. The pseudo-code of the process of creating stacked CNN is presented in Algorithm 2.

Algorithm 2: Stacked Convolutional Neural network and X-ray image Classification

---

Input: Validation set, test set, and sub-models
Output: Classification results
1: Sub-models stacking:
2: **for** i = 1 to *length(validation set)* **do**
3:    **for** j = 1 to 5 **do**
4:       $[P1_{ji}, P2_{ji}, P3_{ji}]$ = sub-model#(j).predict(validation_img[i])
5:    **end for**
6:    P = concatenation($[P1_{ji}, P2_{ji}, P3_{ji}]$
7: **end for**
8: Training of softmax classifier on feature vector P
9: stacked_model = Train(P, validation_label)
10: Classification of the images
11: pred_label = classify(stacked_model, test_img)
12: **return** *pred_label*

The mathematical definition of the softmax classifier is as shown following:

$$\mathscr{H}(y_i|\hat{\theta}^j) = \frac{e^{\hat{\theta}^j}}{\sum_{i=0}^{c} e^{\hat{\theta}_c^j}} \tag{2}$$

$$where \quad \hat{\theta} = w_0 x_0 + w_1 x_1 + w_2 x_2 + \ldots + w_c x_c = \sum_{j=0}^{c} W_j X_j \tag{3}$$

where y_i denotes the associated label of image x_i, c denotes number of classes, W denotes weight matrix and $\mathscr{H}(y_i|\hat{\theta}^j)$ is estimated probability for each class i (where $i \in \{0 : COVID - 19, 1 : Normal, 2 : Pneumonia\}$) of an image x_i.

Next, to train the softmax classifier $H(y_i|\hat{\theta}^j)$ a dataset has been paprared. We have prepared dataset by providing X-ray images from the validation set to the each of the sub-models, and collected output class scores (predictions). In this case, each sub-model j is output three class scores $S0_{ji}$, $S1_{ji}$ and $S2_{ji}$ corresponding to each class (COVID-19, Normal, and Pneumonia classes) for the image x. We have concatenated the output class scores of these five sub-models that results the feature vector S of size 15 (5×3). Thus, for the M X-ray images of the validation set, we have created a dataset of the size of $(M \times 15)$ for the training of the softmax classifier. After training of the stacked model, on new input image x_i from the test set, to make a prediction, pick the class i that maximizes $\mathscr{H}(y_i|\hat{\theta}^j)$.

4. Experiments

This section presents the details of the dataset, evaluation metrics, experiment results, and performance comparison.

4.1. Datasets

In order to evaluate the performance of the proposed Stacked CNN model, we have build two datasets. The detailed description of the datasets are discussed in the following section:

4.1.1. COVID19CXr Dataset
We have generated first dataset of X-ray images, with the combination and modification of three publicly available datasets [13–15], which is referred to as COVID19CXr. The COVID19CXr dataset includes 3040 chest X-ray images of 1930 patients. Out of 3040 images, 546 images of 332 patients belong to a COVID-19 class, 1139 images of 1015 patients belong to a Normal class and 1355 images of 583 patients belong to a Pneumonia class. COVID-19 images are obtained from the two publicly available repositories: 1) *"Figure-1 COVID-19 Chest X-ray Dataset Initiative"* [13] and 2) *"COVID-19 Image Data Collection"* [14]. Pneumonia and Normal cases chest X-ray images are included from the *"Mendeley data"* [15]. Fig. 2 shows the sample chest X-ray images of COVID-19, Normal and Pneumonia classes from the COVID19CXr dataset.

For the performance assessment of the proposed method, we have used a 5-fold cross-validation strategy, where the dataset has divided into a training, a validation, and a test set, in the ratio of approximately 70:10:20, respectively, at the patient label. To make sure the proposed model generalizes to unseen patients, we guarantee that patients used to build the test set have not been used for the training and validation sets. The same strategy has been repeated five times to obtain the different folds. Table 1 gives details of the distribution of images in the training set, validation set, and test set corresponding to each fold. The training set and validation set are used while training the network, and a hold-out test set is used for the performance assessment of the proposed model.

4.1.2. COVID19CTs Dataset
We have built our second dataset by collecting Computed tomography (CT) images of the COVID-19 patients from the Noble Diagnostic Centre at Bhopal, India. We referred to it as the COVID19CTs dataset. The COVID19CTs dataset consists of 4645 chest CT scan images of 65 patients, including 30 females and 35 males. Patients included in the dataset are from the age group of 12 to 94, with an average age of 46 years. COVID19CTs dataset includes 2249 images of 36 patients of positive COVID-19 cases. 2396 images of 29 patients of healthy (no-Findings) cases. Out of 36 COVID-19 patients, 9 patients have COVID-19 related one lung (unilateral) acute respiratory distress syndrome (ARDS), and 27 patients have COVID-19 related bilateral ARDS. Sample CT images of COVID19CTs dataset are represented in Fig. 3. Obtained images size are varies from 515×512 to 630×512. Hence, we have resized all the images to 224×224.

For the performance assessment of the proposed model, we have divided the dataset at patient label into a training set, validation set, and test set in the ratio of 70:10:20. Class-wise images and patient distribution corresponding to the training set, validation set, and test set is given in Table 2.

4.2. Evaluation Metrics

To assess the performance of the proposed method, we have used sensitivity, specificity, accuracy, positive prediction value (PPV), F1-score, G-mean [58] and area under the ROC curve (AUC) as evaluation metrics. The mathematical definition for the evaluation metrics is given below (in Eqn. (4), Eqn. (5), Eqn. (6), Eqn. (7), Eqn. (8), and Eqn. (9) respectively):

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{4}$$

$$PPV = \frac{TP}{(TP + FP)} \tag{5}$$

$$Sensitivity = \frac{TP}{(TP + FN)} \tag{6}$$

$$Specificity = \frac{TN}{(TN + FP)} \tag{7}$$

$$F1 - Score = \frac{2TP}{(2TP + FP + FN)} \tag{8}$$

$$G - mean = \\ \sqrt[3]{Sensitivity_{COVID19} \times Sensitivity_{Normal} \times Sensitivity_{Pneumonia}} \tag{9}$$
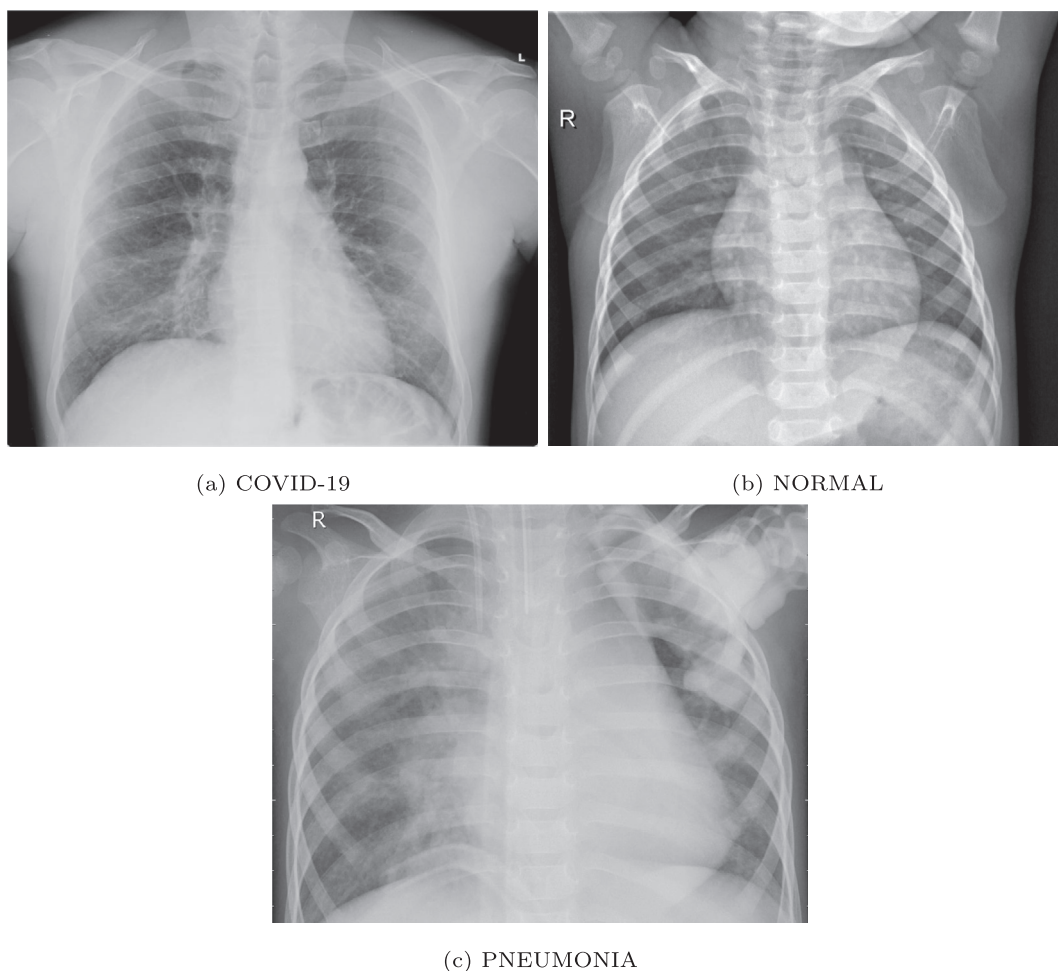
(a) COVID-19   (b) NORMAL

(c) PNEUMONIA

Fig. 2 – Sample X-ray images of Chest from the COVID19CXr dataset; where image in (a) COVID-19, (b) NORMAL, (c) PNEUMONIA.

Table 1 – Images distribution in the training set, validation set and test set, corresponding to folds of COVID19CXr dataset.

Fold (s)	Data set (s)	COVID-19	Normal	Pneumonia	Total
	Train Set	382	798	948	**2128**
	Validation set	55	113	136	**304**
Fold1	Test Set	109	228	271	**608**
	Train Set	381	799	951	2131
	Validation set	54	113	135	302
Fold2	Test Set	111	227	269	607
	Train Set	382	798	950	2130
	Validation set	55	113	134	302
Fold3	Test Set	109	228	271	608
	Train Set	380	797	945	2122
	Validation set	56	114	136	306
Fold4	Test Set	110	228	274	612
	Train Set	382	798	952	2132
	Validation set	56	113	136	305
Fold5	Test Set	108	228	267	603

Where true positive (TP), true negative (TN), false positive (FP), and false negative (FN) are the parameters of the confusion matrix. The present study deals with a multi-class problem; therefore, to get the overall metric score of the method, the mean of each metric is calculated.
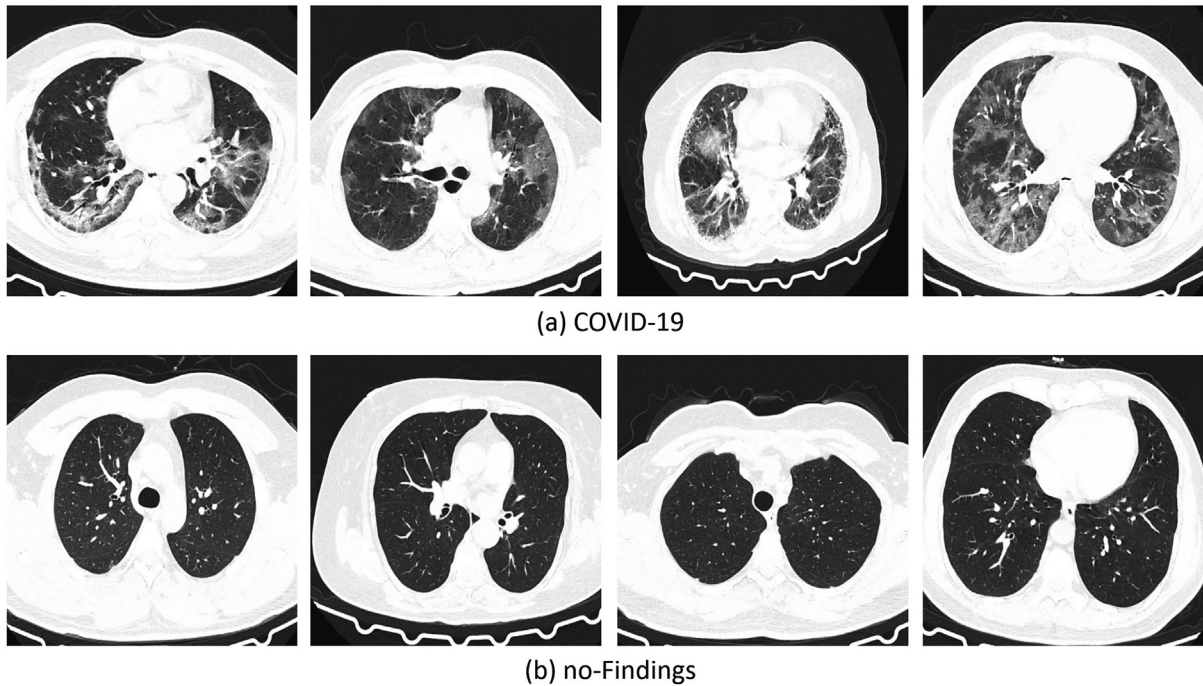
(a) COVID-19

(b) no-Findings

Fig. 3 – Sample CT images of different patients from the COVID19CTs dataset.; where images in (a) COVID-19 and in (b) no-Findings.

Table 2 – Images distribution in the training set, validation set and test set of the COVID19CTs dataset.

Set (s)	Class (s)	Number of Patients	Number of images
Training	COVID-19	24	1478
	no-Findings	19	1606
Validation	COVID-19	4	281
	no-Findings	4	280
Test	COVID-19	7	490
	no-Findings	7	510
Total		65	4645

4.3.   Results and Discussion

In order to evaluate the performance of the proposed stacked CNN, a set of experiments have been conducted. In the first experiment, data augmentation techniques such as flip, rotation, shear, zoom, and shift have been applied on a training set of the both datasets. Thereafter, the augmented training set is utilized for the training of the Xception model and VGG19 model. In the second experiment, the stacked CNN model is trained on the validation set. Finally, evaluation results are produced on the test set. The same set of experiments has been repeated five number of times for each fold. The following sections represent the experimental results and performance comparison.

4.3.1.   Discrimination power of stacked CNN model on COVID19CXr Dataset

Table 3 presents the diagnostic performance of stacked CNN. The proposed model shows good discrimination abil-ity for the diagnosis of the COVID-19 from the chest X-ray images. The proposed model achieved mean sensitivity of 97.62%, the specificity of 98.52%, PPV of 97.36%, accuracy of 97.27%, and G-mean of 97.59% to classify the COVID-19, Normal and Pneumonia X-ray images. The proposed method achieved good sensitivity; therefore, the chances of miss classification of the COVID-19 positive cases are very small.

For a deeper exploration of the performances of the proposed method, the confusion matrix and receiver operating characteristic (ROC) curve corresponding to each fold are evaluated and are shown in Fig. 4 and Fig. 5, respectively. It can be observed from the confusion matrix that the proposed model produces very little false negatives and false positives, specifically for the COVID-19 cases compared to other cases of the COVID19CXr dataset. For COVID-19 cases, it is essential to minimize the wrong diagnosis. On the other hand, the ROC curve shows the stability of the proposed stacked CNN model, and the present model achieved a mean AUC of 0.998 for COVID-19 class and a mean AUC of 0.996 for all categories.

Furthermore, we have investigated the computation time (or prediction time) of the proposed model. The proposed stacked CNN model requires, on average, 0.029 s of computation time to detect the disease from the image on the system with 16 GB GPU (NVIDIA Tesla P100). On the other hand, it requires, on average, 3.66 s to detect the disease from the image on the system with 16 GB CPU (Intel(R) Core(TM) i5-3470). The average computation time has been computed based on the total time required to make the prediction of 612 images of the test set. Based on the prediction time, the developed stacked CNN model is observed to be computationally efficient.

Table 3 – Diagnosis performance of stacked CNN model on COVID19CXr dataset.

Fold (s)	Sensitivity (in %)	Specificity (in %)	G-mean (in %)	Accuracy (in %)	PPV (in %)	F1-score	AUC
Fold1	98.96	99.37	98.95	98.85	99.02	0.989	1.00
Fold2	98.13	98.74	98.12	97.69	97.91	0.980	0.996
Fold3	98.28	99.07	98.27	98.19	98.13	0.982	0.995
Fold4	95.30	96.93	95.21	94.44	94.75	0.950	0.994
Fold5	97.42	98.51	97.42	97.18	97.01	0.972	0.996
Mean	97.62±1.26	98.52±0.85	97.59±1.29	97.27±1.52	97.36±1.46	0.975±0.01	0.996±0.002

(a) Fold1

(b) Fold2

(c) Fold3

(d) Fold4

(e) Fold5
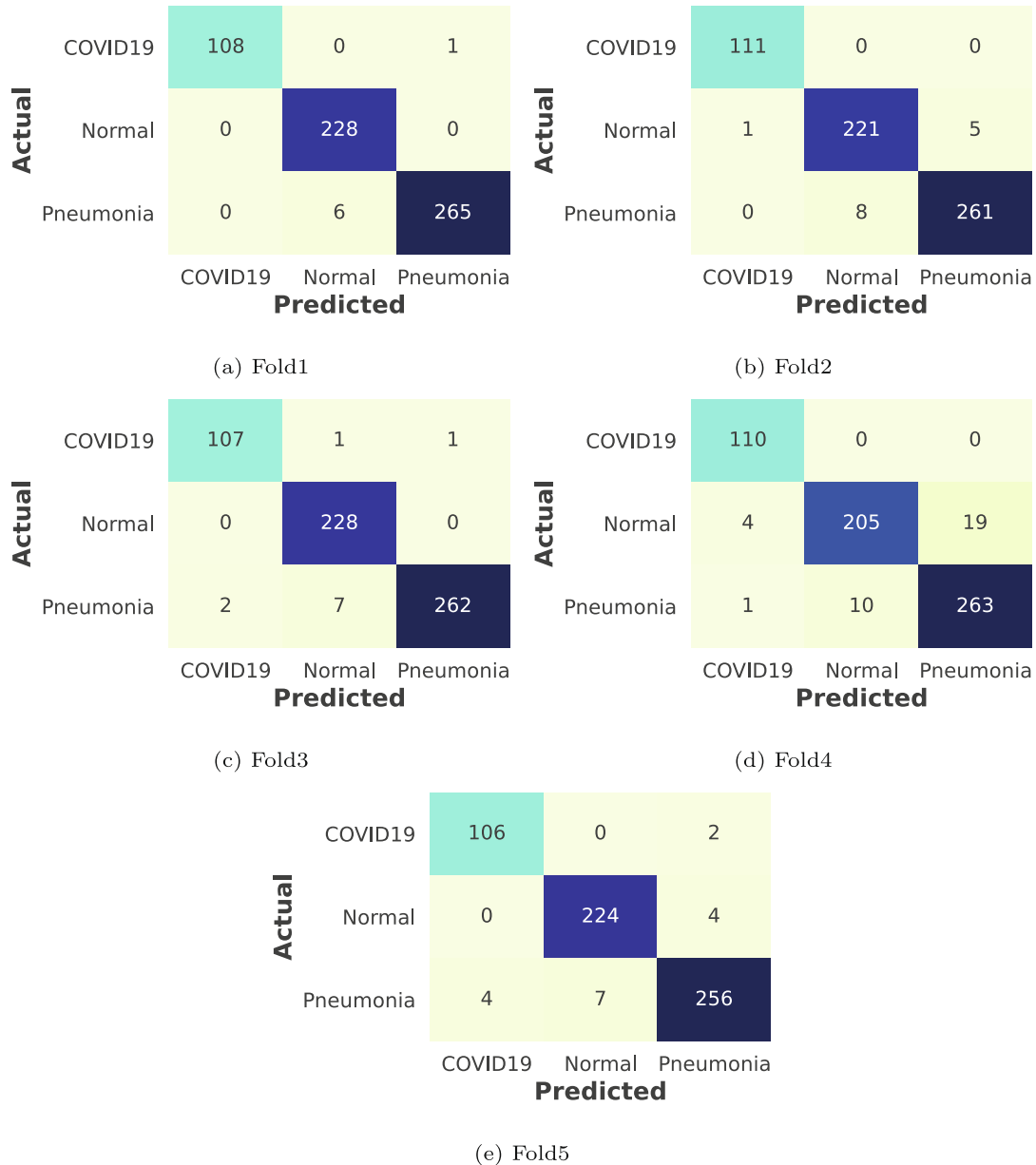
Fig. 4 – Confusion matrix for the stacked CNN model on the different folds.

4.3.2. Performance of stacked CNN model on COVID19CTs Dataset

The performance of the Stacked CNN model on the COVID19CTs dataset is represented in Table 4. It can be observed from Table 4 that the proposed model achieved a sensitivity of 98.31% and classification accuracy of 98.30% for the CT images classification into COVID-19 and no-Findings classes. Fig. 6 shows performance of proposed method in-terms of

(a) Fold1

(b) Fold2

(c) Fold3

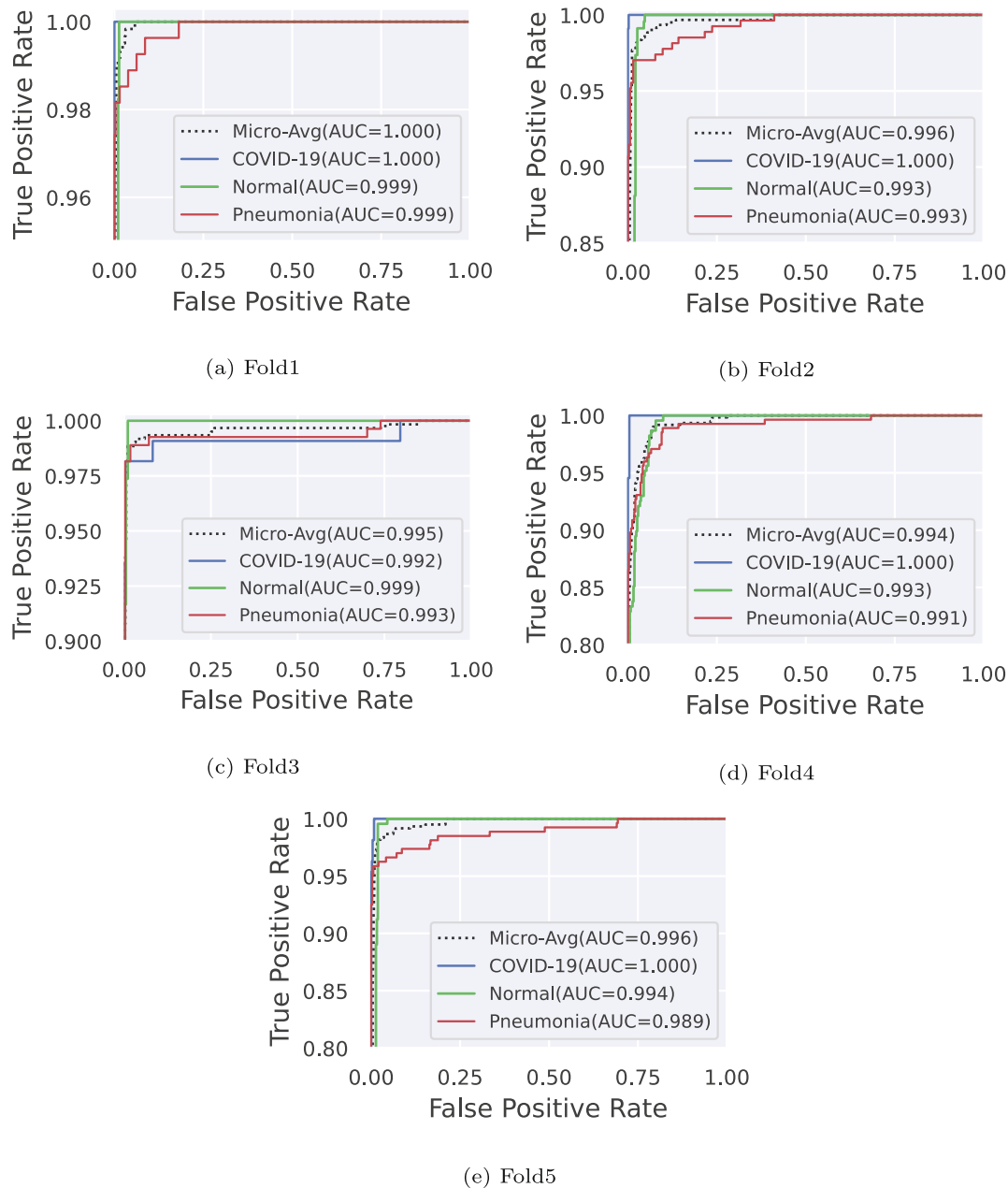

(d) Fold4

(e) Fold5

Fig. 5 – ROC curve for stacked CNN model on the different folds.

the confusion matrix and ROC curve. The proposed method produces very few false positives and false negatives and it achieves AUC of 0.999.

4.3.3. Performance comparison

Table 5 shows the performance comparison of the proposed model, and pre-trained CNN models, namely ResNet50 [40], Inception-V3 [59], Xception [54], DenseNet-121 [60], MobileNet [61], and VGG19 on the COVID19CXr dataset. It is observed from Table 5 that the proposed stacked CNN model achieves the best G-mean of 97.59% compared to pre-trained CNNs. The proposed model improves the G-mean by $0.72\% \sim 17.26\%$. In terms of sensitivity, specificity, and PPV the proposed model also shows performance improvement by $0.70\% \sim 16.65\%, 0.40\% \sim 8.12\%$, and $0.65\% \sim 14.59\%$, respec-

tively. While comparing performances of the VGG19 and Xception with pre-trained networks, it is observed that VGG19 and Xception have shows superiority over the pre-trained networks in terms of G-mean and sensitivity. Table 6 represents the performance of the individual sub-models and the proposed stacked CNN model for each fold. Our stacked ensemble CNN model has incorporated the best prediction of sub-models and achieved better performance than all the sub-models for each fold. Furthermore, we have also investigated and compared the performance of various classifiers to build a stacked ensemble model with different classifiers. Therefore, four classifiers, such as support vector machine (SVM) [62], decision tree (DT)[63], neural network (NN), and K-Nearest Neighbor (KNN) [64], have been applied to the output of sub-models and obtained different stacked model.

Table 4 – Diagnosis performance of stacked CNN model on COVID19CTs dataset.						
Sensitivity (in %)	Specificity (in %)	G-mean (in %)	Accuracy (in %)	PPV (in %)	F1-score	AUC
98.31	98.31	98.30	98.30	98.30	0.98	0.999

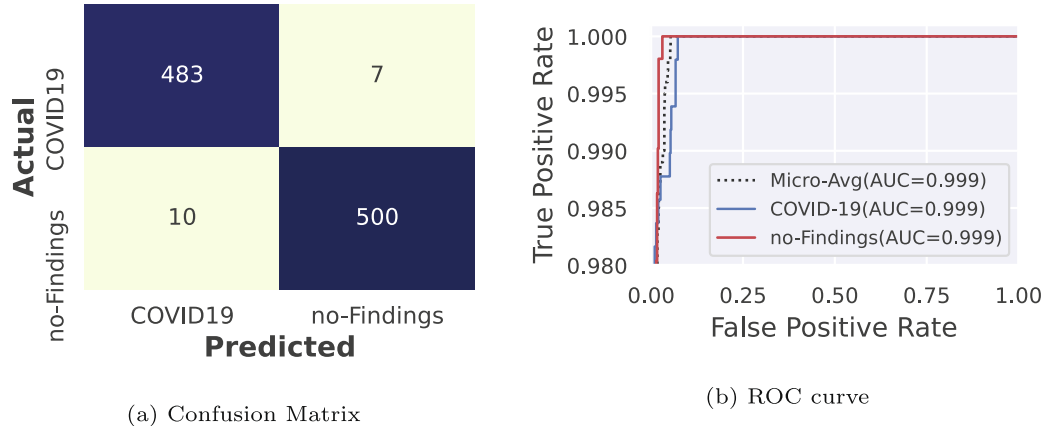(a) Confusion Matrix

(b) ROC curve

Fig. 6 – Confusion matrix and ROC curve for the stacked CNN model on the COVID19CTs dataset.

Table 5 – Performance comparison of different methods.					
Method(s)	Sensitivity (%)	Specificity (%)	G-mean (%)	PPV (%)	F1-score	AUC
ResNet50	80.97±2.47	90.40±0.91	80.33±2.97	82.77±2.18	0.81±0.02	0.93±0.008
Inception-V3	88.98±2.49	94.08±1.64	88.72±2.74	90.49±2.82	0.89±0.03	0.98±0.010
DenseNet-121	88.52±1.67	93.97±0.96	88.18±1.94	90.64±2.15	0.89±0.01	0.97±0.005
MobileNet	88.67±3.20	94.43±1.75	88.36±3.47	91.75±2.44	0.90±0.03	0.98±0.008
VGG19	94.83±0.93	97.10±0.54	94.80±1.12	95.02±1.34	0.95±0.01	0.99±0.005
Xception	96.92±1.48	98.12±0.94	96.87±1.53	96.71±1.57	0.97±0.02	0.99±0.005
Proposed	**97.62 ±1.26**	**98.52 ± 0.85**	**97.59±1.29**	**97.36±1.46**	**0.975±.01**	**0.996±0.002**

Table 6 – Performance evaluation of the sub-models (in terms of G-mean (%)).						
Model (s)		Fold1	Fold2	Fold3	Fold4	Fold5
VGG19	Sub-model1	90.30	93.80	94.22	91.83	94.87
	Sub-model2	95.84	96.41	96.82	93.87	97.44
Xception	Sub-model3	97.28	97.70	95.16	93.40	95.58
	Sub-model4	97.31	97.74	96.36	93.65	96.31
	Sub-model5	98.70	97.76	97.05	94.12	96.70
Proposed Model		**98.95**	**98.12**	**98.27**	**95.21**	**97.42**

Table 7 – Performance comparison of different classifiers (in terms of G-mean (%)).					
Fold (s)	SVM	NN	DT	KNN	Proposed
Fold1	98.33	98.93	99.3	98.66	**98.95**
Fold2	97.97	96.81	96.89	97.52	98.12
Fold3	96.97	97.1	95.53	97.22	**98.27**
Fold4	93.99	92.76	**93.28**	94.15	95.21
Fold5	96.68	95.29	94.34	96.69	97.42
Mean	96.79±1.53	96.18±2.06	95.87±2.10	96.85±1.50	97.59±1.29
The best result among classifiers for each fold is shown in bold.					

Table 8 – Performance comparison with existing methods

Author (s)	Method	Dataset/Subjects	Classification Task	Results
Jain et al. [10]	Image pre-processing, Data augmentation, ResNet50, ResNet101	250 COVID-19, 350 viral pneumonia	Binary: COVID-19, viral pneumonia	97.77% (Acc) 97.14% (Rec) 97.14% (Prec)
Kassania et al. [23]	Deep features, Various classifiers	137 COVID-19, 137 healthy	Binary: COVID-19, Healthy	99% (Acc)
Narin et al. [28]	Pre-Train CNNs	50 COVID-19, 50 Normal	Binary: COVID-19, Normal	97% (Acc)
Sethy et al. [32]	Deep feature, SVM	25 COVID-19+, 25 COVID-19-	Binary: COVID-19+, COVID-19-	95.38% (Acc)
Minaee et al. [33]	Pre-trained CNN	184 COVID-19, 5000 Non-COVID	Binary: COVID-19, Non-COVID	98% (Sens) 90% (Sp)
Castiglioni et al. [34]	Ensemble model Pre-trained CNNs	250 COVID-19, 250 non-COVID-19	Binary: COVID-19, Non-COVID-19	80% (Sens) 81% (Sp)
Abraham et al. [35]	Multi-CNN Features, CFS, Bayesnet	453 COVID-19, 497 non-COVID	Binary: COVID-19, non-COVID	91.16% (Acc) 0.963 (AUC)
Panwar et al. [36]	nCOVnet, VGG16	142 COVID-19, 142 Normal	Binary: COVID-19, Normal	88.10% (Acc) 0.881 (AUC)
Ashour et al. [38]	Ensemble-based BoF, Grid method, SURF	200 COVID-19, 200 Normal	Binary: COVID-19, Normal	98.6% (Acc)
Oh et al. [29]	ResNet18	191 Normal, 54 Bacterial, 57 Tuberculosis, 20 Viral, 180 COVID-19	Multiclass: Normal, Bacterial, Tuberculosis, Viral, COVID-19	88.9% (Acc)
Pereira et al. [31]	Deep features, Texture features, Fusion techniques	200 Normal, 22 SARS, 20 MERS, 180 COVID-19, 20 Varicella, 22 Pneumocystis, 24 Streptococcus	Multiclass: Normal, SARS, COVID-19, MERS, Varicella, Pneumocystis, Streptococcus	0.89 (F1-score)
Apostolopoulos et al. [11]	Pre-trained CNN	224 COVID-19, 504 Normal, 714 Pneumonia	Binary: COVID-19, Pneumonia; Multiclass: COVID-19, Normal, Pneumonia	96.78% (Acc); 94.72% (Acc)
Joshi et al. [12]	DarkNet-53	194 COVID-19, 583 Normal, 2265 Pneumonia	Binary: COVID, Non-COVID; Multiclass: COVID, normal, pneumonia	99.81% (Acc); 97.11% (Acc) 0.951 (F1-score)
Ozturk et al. [30]	DarkCovidNet	127 COVID-19, 500 no-findings, 500 Pneumonia	Binary: COVID-19, No-findings; Multiclass: COVID-19, No-findings, Pneumonia	98.08% (Acc); 87.02% (Acc)
Benmalak et al. [7]	InceptionV3, ResNet-18, MobileNetV3	1530 Normal, 1778 COVID-19, 1718 Viral pneumonia	Multiclass: Normal, COVID-19, Viral pneumonia	93.4% (Prec), 92.3% (Sens), 92.8% (F1-Score)
Wang et al. [8]	COVID-Net	183 COVID-19, 8066 Normal, 5538 non-COVID19	Multiclass: COVID-19, Normal, Non-COVID19	92.6% (Acc)
Ucar et al. [9]	SqueezeNet CNN	1583 Normal, 4290 Pneumonia, 76 COVID-19	Multiclass: Normal, Pneumonia, COVID-19	95.7% (Acc), 90% (Sens)
Nigam et al. [37]	Pre-trained CNNs	6000 Normal, 5634 COVID-19, 5000 others	Multiclass: Normal, COVID, other	93.48% (Acc)
Proposed Method	Stacked CNN: VGG19, Xception, Softmax classifier	**COVID19CXr:** 546 COVID-19, 1139 Normal, 1355 Pneumonia; **COVID19CTs:** 2249 COVID-19, 2396 no-Findings	Multiclass: COVID-19, Normal, Pneumonia; Binary: COVID-19, no-Findings	97.27% (Acc), 97.62% (Sens), 0.975 (F1-Score); 98.30% (Acc), 98.31% (Sens), 0.98 (F1-Score)

"Acc": Accuracy, "Sens": Sensitivity, "Sp": Specificity,"Rec": Recall, "Prec": Precision,"AUC": Area under the ROC curve.

Their performances are represented in Table 7. As observed from Table 7, performances of the stacked model with different classifiers being quite similar, and among them, our proposed model (stacked with Softmax classifier) giving a little better results.

A variety of deep learning-based studies have already been proposed recently for the diagnosis of COVID-19 disease from the chest X-ray images. The performance comparison of the proposed method with some of the related studies is shown in Table 8. Since COVID-19 is a new pandemic and a limited number of COVID-19 X-ray images are available publicly for developing CAD systems for COVID-19 detection.

Studies in [32,28] have just developed their deep learning models on the dataset of a very small size, which consists of 50 and 100 images, respectively. Other studies in Table 8 have used less than 250 COVID-19 images for developing their methods, except the studies in [7,35,37]. In this study, a total of 3040 X-ray images have been used to develop the stacked CNN model, including 546 COVID-19 images, which is the relatively larger number of COVID-19 images among most of the studies presented in Table 8, except the studies in [7,37].

We can see in Table 8 that the studies in [10,23,28,32–36,38] have evaluated for binary classification task, studies in [7–9,29,31,37] have evaluated for multi-class classification task and studies in [11,12,30] have evaluated for binary as well as multi-class classification tasks. For the binary classification of X-ray images, the method proposed by Joshi et al. [12] has outperformed the existing methods. On the other hand, for the multi-class classification task, the proposed stacked CNN model shows superiority over the existing methods.

Some of the salient features of stacked CNN can be summarized as:

- The proposed method is based on the stacked generalization of CNN's sub-models, which minimizes the variance of predictions and reduces generalization error. As a result, stacked CNN yields higher diagnosis accuracy in the both CT and X-ray images.

- The proposed stacked CNN model produces very little false positives (type 1) and false negatives (type 2) error, which confirms that the stacked CNN is reliable for clinical use.

- The proposed model is developed based on less complex networks, which is computationally efficient, and shows its stability on a small dataset.

- The stacked CNN model requires, on average, 0.029 s of computation time to detect the disease from an image. Therefore, this model could be utilized for rapid screening of the COVID-19 disease.

5.   Conclusion

In this paper, we introduced a new stacked convolutional neural network for the automatic diagnosis of the COVID19 from the chest X-ray and CT images. In the proposed method, CNN's

sub-models have been obtained from the pre-trained Xception and the VGG19 models. The proposed stacked CNN model ensemble the sub-models using Softmax classifier, for deriving a powerful model for image classification than individual sub-models. The stacked CNN model is able to learn the image discriminative features and retrieved the diverse information present in the radiological images of the chest. It achieves a classification accuracy of 97.27% on the chest X-ray images of the COVID19CXr dataset and 98.30% on the CT scan images of the COVID19CTs dataset. Our proposed approach shows its superiority over the existing methods for the diagnosis of the COVID-19 from the X-ray images.

Our experiments results show the effectiveness of the stacked CNN for the classification of COVID-19, Normal, and Pneumonia X-ray images. More importantly, the proposed model outperforms the pre-trained CNNs, including ResNet50, Inception-V3, Xception, DenseNet, and MobileNet, for the classification of chest X-ray images. In the future, we would like to explore the stacked CNN model for the further classification of X-ray images into bacterial pneumonia, non-COVID-19-viral pneumonia, COVID-19-viral pneumonia, and normal lung classes.

CRediT authorship contribution statement

Mahesh Gour: Conceptualization, Methodology, Software, Writing - original draft, Writing - review & editing. **Sweta Jain:** Supervision, Visualization, Project administration, Resources, Formal analysis, Investigation.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

R E F E R E N C E S

[1] Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (covid-19) infections. Int J Infectious Diseases 2020;93:284–6. https://doi.org/10.1016/j.ijid.2020.02.060.

[2] Einstein AJ, Shaw LJ, Hirschfeld C, Williams MC, Villines TC, Better N, Vitola JV, Cerci R, Dorbala S, Raggi P, et al. International impact of covid-19 on the diagnosis of heart disease. J Am Coll Cardiol 2021;77(2):173–85. https://doi.org/10.1016/j.jacc.2020.10.054.

[3] World Health Organization, Laboratory testing for 2019 novel coronavirus (2019-nCoV) in suspected human cases, WHO - Interim guidance 2019 (January) (2020) 1–7, [Accessed 9 May 2020]. URL:https://www.who.int/publications/i/item/laboratory-testing-of-2019-novel-coronavirus-(-2019-ncov)-in-suspected-human-cases-interim-guidance-17-january-2020.

[4] M.-Y. Ng, E.Y.P. Lee, J. Yang, F. Yang, X. Li, H. Wang, M.M.-s. Lui, C.S.-Y. Lo, B. Leung, P.-L. Khong, et al., Imaging profile of the covid-19 infection: Radiologic findings and literature review, Radiology: Cardiothoracic Imaging 2 (1) (2020) e200034. doi:10.1148/ryct.2020200034.

[5] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, et al. Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. Lancet 2020;395(10223):497–506. https://doi.org/10.1016/S0140-6736(20)30183-5.

[6] Tang Z, Zhao W, Xie X, Zhong Z, Shi F, Ma T, Liu J, Shen D. Severity assessment of COVID-19 using CT image features and laboratory indices. Phys Med Biol 2021;66(3) . https://doi.org/10.1088/1361-6560/abbf9e 035015.

[7] Benmalek E, Elmhamdi J, Jilbab A. Comparing ct scan and chest x-ray imaging for covid-19 diagnosis. Biomed Eng Adv 2021;1 . https://doi.org/10.1016/j.bea.2021.100003 100003.

[8] L. Wang, Z.Q. Lin, A. Wong, Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest radiography images, Scientific Reports 10 (1). doi:10.1038/s41598-020-76550-z.

[9] Ucar F, Korkmaz D. Covidiagnosis-net: Deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (covid-19) from x-ray images. Med Hypotheses 2020;140 . https://doi.org/10.1016/j.mehy.2020.109761 109761.

[10] Jain G, Mittal D, Thakur D, Mittal MK. A deep learning approach to detect covid-19 coronavirus with x-ray images. Biocybernetics Biomed Eng 2020;40(4):1391–405. https://doi.org/10.1016/j.bbe.2020.08.008.

[11] Apostolopoulos ID, Mpesiana TA. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys Eng Sci Med 2020;43(2):635–40. https://doi.org/10.1007/s13246-020-00865-4.

[12] Joshi RC, Yadav S, Pathak VK, Malhotra HS, Khokhar HVS, Parihar A, Kohli N, Himanshu D, Garg RK, Bhatt MLB, et al. A deep learning-based covid-19 automatic diagnostic framework using chest x-ray images. Biocybernetics Biomed Eng 2021;41(1):239–54. https://doi.org/10.1016/j.bbe.2021.01.002.

[13] L. Wang, A. Wong, Z.Q. Lin, P. McInnis, A. Chung, H. Gunraj, Figure 1 covid-19 chest x-ray data initiative, 2020, [Accessed 15 June 2020]. URL:https://github.com/agchung/Figure1-COVID-chestxray-dataset.

[14] J.P. Cohen, P. Morrison, L. Dao, K. Roth, T.Q. Duong, M. Ghassemi, Covid-19 image data collection, arXiv 2006.11988. URL:https://github.com/ieee8023/covid-chestxray-dataset.

[15] D. Kermany, K. Zhang, M. Goldbaum, Labeled Optical Coherence Tomography (OCT) and chest X-Ray images for classification, Mendeley Data, v2. URL: https://data.mendeley.com/datasets/rscbjbr9sj/2.

[16] Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Computerized Med Imag ggraphics 2007;31(4–5):198–211. https://doi.org/10.1016/j.compmedimag.2007.02.002.

[17] B. Van Ginneken, S. Katsuragawa, B. Ter Haar Romeny, K. Doi, M. Viergever, Automatic detection of abnormalities in chest radiographs using local texture analysis, IEEE Transactions on Medical Imaging 21 (2) (2002) 139–149. doi:10.1109/42.993132.

[18] Jaeger S, Karargyris A, Candemir S, Folio L, Siegelman J, Callaghan F, Xue Z, Palaniappan K, Singh RK, Antani S, et al. Automatic tuberculosis screening using chest radiographs. IEEE Trans Med Imaging 2014;33(2):233–45. https://doi.org/10.1109/TMI.2013.2284099.

[19] Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, He K, Shi Y, Shen D. Review of artificial intelligence techniques in imaging data acquisition, segmentation, and diagnosis for covid-19. IEEE Rev Biomed Eng 2021;14:4–15. https://doi.org/10.1109/RBME.2020.2987975.

[20] Dong D, Tang Z, Wang S, Hui H, Gong L, Lu Y, Xue Z, Liao H, Chen F, Yang F, et al. The role of imaging in the detection and management of covid-19: A review. IEEE Rev Biomed Eng 2021;14:16–29. https://doi.org/10.1109/RBME.2020.2990959.

[21] Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yu L, Ni Q, Chen Y, Su J, et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. Engineering 2020;6(10):1122–9. https://doi.org/10.1016/j.eng.2020.04.010.

[22] Shah V, Keniya R, Shridharani A, Punjabi M, Shah J, Mehendale N. Diagnosis of covid-19 using ct scan images and deep learning techniques. Emergency Radiol 2021;28:497–505. https://doi.org/10.1007/s10140-020-01886-y.

[23] Kassania SH, Kassanib PH, Wesolowskic MJ, Schneidera KA, Detersa R. Automatic detection of coronavirus disease (covid-19) in x-ray and ct images: A machine learning based approach. Biocybernetics Biomed Eng 2021;41(3):867–79. https://doi.org/10.1016/j.bbe.2021.05.013.

[24] Wang J, Bao Y, Wen Y, Lu H, Luo H, Xiang Y, Li X, Liu C, Qian D. Prior-attention residual learning for more discriminative covid-19 screening in ct images. IEEE Trans Med Imaging 2020;39(8):2572–83. https://doi.org/10.1109/TMI.2020.2994908.

[25] Afshar P, Heidarian S, Naderkhani F, Oikonomou A, Plataniotis KN, Mohammadi A. Covid-caps: A capsule network-based framework for identification of covid-19 cases from x-ray images. Pattern Recogn Lett 2020;138:638–43. https://doi.org/10.1016/j.patrec.2020.09.010.

[26] Ardakani AA, Kanafi AR, Acharya UR, Khadem N, Mohammadi A. Application of deep learning technique to manage covid-19 in routine clinical practice using ct images: Results of 10 convolutional neural networks. Comput Biol Med 2020;121 . https://doi.org/10.1016/j.compbiomed.2020.103795 103795.

[27] Mishra NK, Singh P, Joshi SD. Automated detection of covid-19 from ct scan using convolutional neural network. Biocybernetics Biomed Eng 2021;41(2):572–88. https://doi.org/10.1016/j.bbe.2021.04.006.

[28] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, Pattern Analysis and Applications doi: 10.1007/s10044-021-00984-y.

[29] Oh Y, Park S, Ye JC. Deep learning covid-19 features on cxr using limited training data sets. IEEE Trans Med Imaging 2020;39(8):2688–700. https://doi.org/10.1109/TMI.2020.2993291.

[30] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of covid-19 cases using deep neural networks with x-ray images. Computers Biol Med 2020;121 . https://doi.org/10.1016/j.compbiomed.2020.103792 103792.

[31] Pereira RM, Bertolini D, Teixeira LO, Silla CN, Costa YM. Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios. Comput Methods Programs Biomed 2020;194 . https://doi.org/10.1016/j.cmpb.2020.105532 105532.

[32] P.K. Sethy, S.K. Behera, P.K. Ratha, P. Biswas, Detection of coronavirus disease (covid-19) based on deep features and support vector machine, Preprints 2020030300.

[33] Minaee S, Kafieh R, Sonka M, Yazdani S, Jamalipour Soufi G. Deep-covid: Predicting covid-19 from chest x-ray images using deep transfer learning. Med Image Anal 2020;65 . https://doi.org/10.1016/j.media.2020.101794 101794.

[34] I. Castiglioni, D. Ippolito, M. Interlenghi, C.B. Monti, C. Salvatore, S. Schiaffino, A. Polidori, D. Gandola, C. Messa, F. Sardanelli, Artificial intelligence applied on chest x-ray can aid in the diagnosis of covid-19 infection: a first experience from lombardy, italy, medRxiv doi: 10.1101/2020.04.08.20040907.

[35] Abraham B, Nair MS. Computer-aided detection of covid-19 from x-ray images using multi-cnn and bayesnet classifier. Biocybernetics Biomed Eng 2020;40(4):1436–45. https://doi.org/10.1016/j.bbe.2020.08.005.

[36] Panwar H, Gupta P, Siddiqui MK, Morales-Menendez R, Singh V. Application of deep learning for fast detection of covid-19 in x-rays using ncovnet. Chaos, Solitons & Fractals 2020;138 (4) . https://doi.org/10.1016/j.chaos.2020.109944 109944.

[37] Nigam B, Nigam A, Jain R, Dodia S, Arora N, Annappa B. Covid-19: Automatic detection from x-ray images by utilizing deep learning methods. Expert Syst Appl 2021;176 . https://doi.org/10.1016/j.eswa.2021.114883 114883.

[38] Ashour AS, Eissa MM, Wahba MA, Elsawy RA, Elgnainy HF, Tolba MS, Mohamed WS. Ensemble-based bag of features for automated classification of normal and covid-19 cxr images. Biomed Signal Process Control 2021;68 . https://doi.org/10.1016/j.bspc.2021.102656 102656.

[39] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Adv Neural Inform Processing Syst 2012;25:1097–105.

[40] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). p. 770–8. https://doi.org/10.1109/CVPR.2016.90.

[41] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 arXiv:1409.1556.

[42] Abbas A, Jain S, Gour M, Vankudothu S. Tomato plant disease detection using transfer learning with c-gan synthetic images. Computers Electron Agric 2021;187 . https://doi.org/10.1016/j.compag.2021.106279 106279.

[43] Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation, in. In: 2014 IEEE Conference on Computer Vision and Pattern Recognition. p. 580–7. https://doi.org/10.1109/CVPR.2014.81.

[44] Girshick R. Fast r-cnn. In: 2015 IEEE International Conference on Computer Vision (ICCV). p. 1440–8. https://doi.org/10.1109/ICCV.2015.169.

[45] Gour M, Jain S, Agrawal R. Deeprnnetseg: Deep residual neural network for nuclei segmentation on breast cancer histopathological images. In: International Conference on Computer Vision and Image Processing. Springer; 2019. p. 243–53. https://doi.org/10.1007/978-981-15-4018-9_23.

[46] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical image computing and computer-assisted intervention, Springer, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.

[47] M. Gour, S. Jain, T. Sunil Kumar, Residual learning based cnn for breast cancer histopathological image classification, International Journal of Imaging Systems and Technology 30 (3) (2020) 621–635. doi: 10.1002/ima.22403.

[48] Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK. Medical image analysis using convolutional neural networks: a review. J Medical Syst 2018;42(11):226. https://doi.org/10.1007/s10916-018-1088-1.

[49] Shen D, Wu G, Suk H-I. Deep learning in medical image analysis. Ann Rev Biomed Eng 2017;19:221–48. https://doi.org/10.1146/annurev-bioeng-071516-044442.

[50] Huang X, Shan J, Vaidya V. Lung nodule detection in ct using 3d convolutional neural networks. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). p. 379–83. https://doi.org/10.1109/ISBI.2017.7950542.

[51] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, K. Shpanskaya, et al., Chexnet: Radiologist-level pneumonia detection on chest x-rays with deep learning, arXiv preprint arXiv:1711.05225 arXiv:1711.05225.

[52] Liu C, Cao Y, Alcantara M, Liu B, Brunette M, Peinado J, Curioso W. Tx-cnn: Detecting tuberculosis in chest x-ray images using convolutional neural network. In: 2017 IEEE International Conference on Image Processing (ICIP). p. 2314–8. https://doi.org/10.1109/ICIP.2017.8296695.

[53] M. Gour, S. Jain, Stacked convolutional neural network for diagnosis of covid-19 disease from x-ray images, arXiv preprint arXiv:2006.13817.

[54] Chollet F. Xception: Deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). p. 1800–7. https://doi.org/10.1109/CVPR.2017.195.

[55] Tieleman T, Hinton G. Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude. COURSERA: Neural networks for machine learning 2012;4 (2):26–31.

[56] Wolpert DH. Stacked generalization. Neural Networks 1992;5 (2):241–59. https://doi.org/10.1016/S0893-6080(05)80023-1.

[57] Breiman L. Stacked regressions. Mach Learn 1996;24(1):49–64. https://doi.org/10.1007/BF00117832.

[58] M. Kubat, S. Matwin, et al., Addressing the curse of imbalanced training sets: one-sided selection, in: Icml, Vol. 97, Citeseer, 1997, pp. 179–186.

[59] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). p. 2818–26. https://doi.org/10.1109/CVPR.2016.308.

[60] Huang G, Liu Z, Van Der Maaten L, Weinberger KQ. Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). p. 2261–9. https://doi.org/10.1109/CVPR.2017.243.

[61] A.G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, H. Adam, Mobilenets: Efficient convolutional neural networks for mobile vision applications, arXiv preprint arXiv:1704.04861 arXiv:1704.04861.

[62] Cortes C, Vapnik V. Support-vector networks. Mach Learn 1995;20(3):273–97. https://doi.org/10.1007/BF00994018.

[63] Safavian S, Landgrebe D. A survey of decision tree classifier methodology. IEEE Trans Systems, Man, Cybern 1991;21 (3):660–74. https://doi.org/10.1109/21.97458.

[64] Peterson LE. K-nearest neighbor. Scholarpedia 2009;4(2):1883. https://doi.org/10.4249/scholarpedia.1883.