ORIGINAL ARTICLE

Transfer learning–based ensemble support vector machine model for automated COVID-19 detection using lung computerized tomography scan data

Mukul Singh[1] · Shrey Bansal[1] · Sakshi Ahuja[2] · Rahul Kumar Dubey[3] ⑩ · Bijaya Ketan Panigrahi[2] · Nilanjan Dey[4]

Abstract

The novel discovered disease coronavirus popularly known as COVID-19 is caused due to severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and declared a pandemic by the World Health Organization (WHO). An early-stage detection of COVID-19 is crucial for the containment of the pandemic it has caused. In this study, a transfer learning–based COVID-19 screening technique is proposed. The motivation of this study is to design an automated system that can assist medical staff especially in areas where trained staff are outnumbered. The study investigates the potential of transfer learning–based models for automatically diagnosing diseases like COVID-19 to assist the medical force, especially in times of an outbreak. In the proposed work, a deep learning model, i.e., truncated VGG16 (Visual Geometry Group from Oxford) is implemented to screen COVID-19 CT scans. The VGG16 architecture is fine-tuned and used to extract features from CT scan images. Further principal component analysis (PCA) is used for feature selection. For the final classification, four different classifiers, namely deep convolutional neural network (DCNN), extreme learning machine (ELM), online sequential ELM, and bagging ensemble with support vector machine (SVM) are compared. The best performing classifier bagging ensemble with SVM within 385 ms achieved an accuracy of 95.7%, the precision of 95.8%, area under curve (AUC) of 0.958, and an F1 score of 95.3% on 208 test images. The results obtained on diverse datasets prove the superiority and robustness of the proposed work. A pre-processing technique has also been proposed for radiological data. The study further compares pre-trained CNN architectures and classification models against the proposed technique.

Keywords COVID-19 · CT scan data · Transfer learning · Ensemble SVM · VGG16

1 Introduction

A chest infection is a kind of infection that affects the proper functioning of the lungs (both larger and smaller airways) [1]. The severity of a lung infection depends on several factors like causes of lung infection (virus or bacteria) and the overall health of the infected person. The most common lung infections are pneumonia, chronic obstructive pulmonary disease (COPD), asthma, bronchitis, and lung cancer. Coronavirus disease popularly known as COVID-19 is a kind of lung infection disease. It is caused due to the novel discovered virus know as severe acute respiratory syndrome coronavirus 2 (SARS- CoV-2). Coronaviruses are

a family of viruses that are known to cause diseases like a common cold, severe acute respiratory syndrome (SARS), and Middle East respiratory syndrome (MERS)[2]. The coronavirus disease is first discovered in Wuhan, China, in December 2019. The unprecedented rise in COVID-19 cases is impacting the worldwide economy and declared a pandemic by the World Health Organization [3].

On 22 May 2020, a total 5,207,918 patients are infected with COVID-19, and 334,848 deaths are reported across 215 countries [4]. To control the spread of the COVID-19 virus, its accurate detection and treatment are required. Reverse transcriptase polymerase chain reaction (RT- PCR) is the standard diagnostic test for COVID-19 [5]. The high popularity of PCR is due to its high selectivity and sensitivity, i.e., over 90%. The limitations of the COVID-19 testing with PCR technique are (a) very time consuming, (b) expensive, and (c) shortage of kits due to long production time [6]. Considering the alarming rates of spread of COVID-19, a faster and cheaper testing mechanism is

✉ Rahul Kumar Dubey
  RahulKumar.Dubey@in.bosch.com

Extended author information available on the last page of the article.

required to tackle this outbreak. The need for a faster screening technique to control the COVID outbreak has also been studied by the authors in [7, 8]. Researchers have found that radiological analysis like X-rays and chest CT (computed tomography) scans have high accuracy in COVID-19 diagnosis and can be an effective tool for large scale screening. A high correlation between RT-PCR and radiological results for COVID-19 diagnosis is established in [9]. Also, COVID-19 infection is identified through ground-glass opacity patches (GGO) in radiographic scans of patients. This encouraged the development of a faster and cheaper COVID-19 screening mechanism using a radiological approach. Also, deep learning is playing a critical role in medical image analysis which motivates its use in screening of COVID-19. The growing role of deep learning has also been analyzed by the study of AI-based COVID classification techniques [10] where authors find that deep-learning-based techniques can provide very promising results for COVID classification. The details of the techniques available in the literature for COVID-19 diagnosis is put forth in Table 1.

From the detailed analysis of the state of the art of COVID-19 diagnosis field, it can be inferred that chest radiography (X-rays and CT scan) is the best alternative for COVID-19 detection in comparison to the RT-PCR test kits [31]. However, CT scan modality seems to be most efficient in comparison to chest X-ray due to the following reasons: (a) CT scan gives a detailed 3-dimensional view of the diagnosed organ whereas X-rays give a 2-D view, (b) the CT scan does not overlap the organ, whereas in X-rays ribs overlap the lungs and heart. Due to the high precision of a CT-scan based screening system, a deep learning-based 3 step model is proposed which consists of a transfer-learning-based feature extractor, a feature selector, and a feature classifier. In the proposed work, a truncated VGG16 architecture is proposed for extracting features. The last two blocks of the truncated architecture are fine-tuned with differential learning rates. PCA is applied to the features extracted by CNN. For the classification task, four different classifier models are compared.

The study addresses various issues with the current COVID-19 datasets and proposes various techniques to overcome these. Transfer learning capabilities of various models have been demonstrated and compared. Furthermore, various techniques like truncation and differential learning rates are proposed to increase robustness. The effect of various feature selection techniques has also been studied. Finally, bagging SVM is chosen for classification after a comparative study of popular classifiers. The proposed model within 385 ms achieved an accuracy of 95.7%, the precision of 95.8%, area under curve (AUC) of 0.958, and an F1 score of 95.3% on the 208 test images. The

results obtained on diverse datasets prove the superiority and robustness of the proposed work.

The rest of the paper is organized as follows: Section 2 illustrates the proposed methodology; Section 2.7 put-forth the details of different classifiers. Section 3 gives the details of results and discussion. Then Section 5 concludes the proposed work.

2 Proposed methodology

The chest CT scans of COVID-19 patients contain patches of ground glass opacity (GGO); thus, a multi-dimensional feature extractor is required for screening [32]. In the proposed work, the VGG16 architecture is fine-tuned and used to extract features from lung CT scan images. Since the size of the COVID-19 dataset is very small, a truncated version of the VGG16 architecture is used. PCA is used to reduce the dimensionality of the features obtained from truncated VGG-16. The final classification is performed using four different classifiers. The self-explanatory block diagram of the proposed methodology for COVID-19 classification is shown in Fig. 1.

2.1 Training data

In the proposed work, the dataset is collected from three different sources to ensure the robustness of the model. The brief details of datasets used are:

a) **Dataset 1 (D1)-** A CT scan dataset of 617 COVID and non-COVID images, compiled by Zhao et. al [14]
b) **Dataset 2 (D2)-** COVID-19 image data collection (53 COVID CT scans): Joseph Paul Cohen, Paul Morrison, and Lan Dao [24]
c) **Dataset 3 (D3)-** Italian society of medical and interventional research (60 COVID-19 CT scans) [15]

The final split of the data is summarized below:

– **Training** : 432 images (204 COVID and 228 non-COVID before augmentation
– **Validation** : 62 images (29 COVID and 33 non-COVID)
– **Test** : 208 images (111 COVID and 97 non-COVID)

Some images in these datasets had markings and other non-removable artifacts and hence had to be dropped. Table 2 puts forth the details of CT scan images available in D1, D2, and D3 along-with the details of training, validation, and test set used. The minimum, average, and maximum height are 153, 491, and 1853. The minimum, average, and maximum width are 124, 383, and 1485. These images are from 216 patient cases. For patients labeled with

Table 1 Summary of techniques available in literature for COVID19 screening

Ref.	Technique	Key findings	Dataset
[11]	Transfer learning on ResNet-50 (CNN model)	Achieved 10-fold cross-validation accuracy of 93.01% on 109 test images.	413 COVID-19 (+) images and 439 images of normal or pneumonia infected patients. [12]
[13]	Resnet50 and VGG16 (deep learning)	COVID-19 positive cases and pneumonia cases of X-ray modalities are classified with an accuracy of 89.2%.	There were 135 COVID-19 cases obtained from JP Cohen [14], radiopedia and SIRM [15].
[16]	Two-step transfer learning model	Two-step transfer learning pipeline based on the deep neural network framework COVID19XrayNet. The approach achieved a maximum accuracy of 91.4%.	The study uses 189 COVID-19 Chest X-Ray images (131 train and 41 test). [14]
[17]	EfficientNet	An accuracy of 93.9%, a sensitivity of 96.8%, and positivity prediction of 100% are obtained on 231 test X-ray images (COVID19-positive cases—31, pneumonia—100, and normal cases—100).	The model is trained on 13569 X-ray images (COVID19 positive cases-152, pneumonia-5421, and normal cases-7966) [18]
[19]	Pre-trained CheXNet and DenseNet	An accuracy of 90.5% and a sensitivity of 100% are achieved on 654 test X-ray images (COVID19-positive cases—30, pneumonia—390, and normal cases—234).	The model is trained on 5323 chest X-ray images (COVID19 positive cases-115, pneumonia-3867, and normal cases - 1341) [14].
[20]	Domain extension transfer learning (DETL) with gradient class activation map (Grad-CAM)	Fivefold cross-validation accuracy of 90.13% and test set accuracy of 95.3% are obtained on the proposed X-ray dataset.	A total of 305 COVID-19 X-Ray images were used in the study. [14]
[21]	ResNet, Inception, and GoogleNet	The classification of COVID-19 positive cases based on X-ray modality is done. The approach achieved 98% of accuracy with VGG19, 95% with Resnet50, and 96% with InceptionV3.	A dataset of nearly 100 subject, among them 50 x-ray images subjects were tested positive with COVID-19 and 50 x-ray images of normal subjects. [14, 15]
[22]	ResNet18, ResNet50, SqueezeNet, and DenseNet-121	The model achieved a specificity of 90% and sensitivity of 96.5% on testing data of 3000 chest X-rays (COVID and non-COVID patients).	The model is trained on 5000 chest x-rays dataset [14]
[23]	Joint classification and segmentation (JCS)	Classification is done with a specificity of 93% and a sensitivity of 95%. A dice score of 78.3% is obtained for the segmentation task.	JCS system is implemented on 400 COVID-19 patients (144,167 images) and 350 Non-COVID patients. [14, 24, 25]

Table 1 (continued)

Ref.	Technique	Key findings	Dataset
[26]	Pruned efficient net-based model on chest CT scans and X-rays	Classification into two binary classes, i.e., COVID and non-COVID. The highest accuracy of 85.22% is achieved with the ResNet50 pre-trained CNN model.	The CNN based pre-trained models are trained on 596 chest CT scans. [14, 15]
[27]	Detail-oriented capsule networks (DECAPS) +Peekaboo (patch crop and drop strategy)	An accuracy of 87.6%, recall of 91.5%, precision of 84.3%, and AUC of 96.1 are achieved for binary classification (COVID-19 and non-COVID) of chest CT scan.	Uses a total of 746 chest CT images - COVID-19 (349 images) and non-COVID-19 (347 images). [15]
[28]	Transfer learning on Xception net	For binary classification of chest CT scan of COVID and non-COVID dataset, the model achieved a sensitivity of 96.1%, the specificity of 93.4%, and AUC of 0.92.	It contains three classes as COVID-19 (+), pneumonia (+) but COVID-19 (-) with 504 images. [25]
[29]	Multi-objective differential evolution (MODE) deep learning	In comparison to authentic CNN models, the performance parameters of MODE outperforms by 2.09% of F-measure, 1.82% of sensitivity, 1.68% of specificity, and 1.927% of Kappa statistics.	A study of 73 patients with 205 COVID positive images. [30]

positive, 169 of them have age information and 137 of them have gender information. The images are from multiple sources including hospital donations, pre-prints, and reports released by the ISMIR [15]. The exact patient and image details can be found at [14, 15, 24].

2.2 Pre-processing module

As the input images are of different sizes, thus all the input images are resized to $112 \times 112 \times 3$ to maintain the uniformity.

CT scans have artifacts like beam hardening, noise, and scatter, which reduce the accuracy of the model. To overcome this, first, a median filter has been applied. Median filtering is a widely used nonlinear method used to remove noise from images while preserving edges. The median filter operates by replacing each pixel value with the median value of neighboring pixels. 5×5 is a popular median filter kernel choice for biomedical images and has hence been chosen [33]. Finally, morphological close transformation is applied to the image. A morphological close operation is a dilation operation followed by an erosion operation. It removes holes and any remaining salt and pepper noise from the images. It has been shown to be highly effective on binary and gray-scale images [34].

Since the images are of different scales and have labelings and markings around corners, an adaptive ROI selector has been applied to the images. The filter first centers and straightens the image, then applies an elliptical mask on the images such that the non-lung parts get cropped out. The ellipse is fitted to the image using the abrupt pixel-value changes as we move close to the rib-cage bones (black to white color change). Furthermore, all the masked images are manually checked to ensure that no image is over or under-cropped. Such images are manually fixed. Figure 2 shows the pictorial representation of the various pre-processing module used in the study.

2.3 Image augmentation

Data augmentation allows the model to learn a more diverse set of features and also increases the size of the dataset thereby preventing the model from overfitting. Each training image is augmented by a random affine transformation, random flip, and random changes in hue, brightness, and saturation of the image. The random affine transformation consists of shearing and rotation. The details of image augmentation parameters include (a) rotation—within range of 0 to 30°, (b) shearing—0.2, (c) zooming—0.2, and (d) changing the brightness level—within range of 0.75 to 1.5. The augmentation parameters were chosen based on the study of the effectiveness of image augmentation techniques on deep networks [35]. The parameters chosen have also

Fig. 1 Self-explanatory block diagram of the proposed methodology of COVID-19 screening
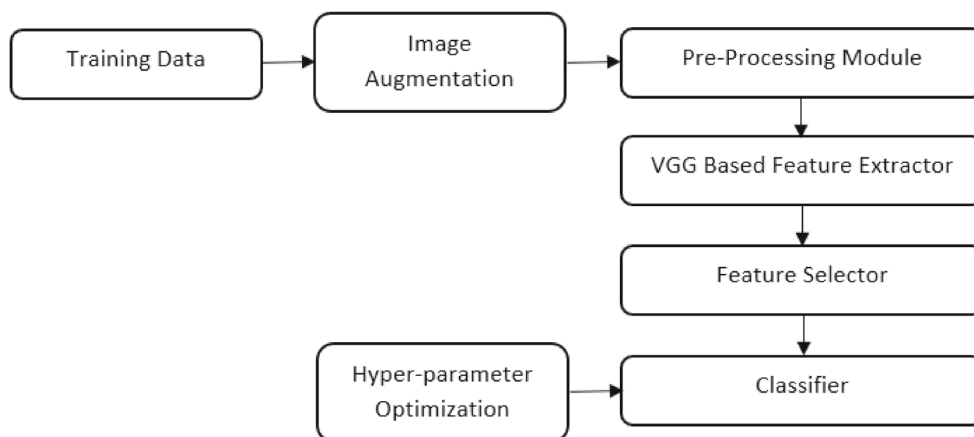
been used in [36] where authors have achieved great results with these settings on a CT scan classification problem similar to this.

The training data after augmentation is (a) 612 of COVID-19 images and (b) 684 of non-COVID-19 images.

2.4 VGG-based feature extractor

Table 3 shows the result of the comparative study considering the popular CNN architectures. The accuracy reported is on the test set by adding a classification layer to the model, and it can be seen that VGG16 outperforms other deeper architectures. This is an interesting observation that deeper models are performing poorer on the current COVID-19 datasets. This is probably due to the size and quality of the currently available datasets. Due to the better performance of VGG16, it is the model of choice in this study. Similar observations have been made in [37] where authors use VGG16 to classify COVID-19 from a multi-modal input.

Initially, the VGG model is trained on the ImageNet database with over 14 million images [38]. Instead of using large receptive fields, VGG16 uses very small receptive fields (3×3 with a stride of 1). VGG16 incorporates 1×1 convolution layers to make the decision function more non-linear without changing the receptive fields. Since the COVID-19 dataset is much smaller with only 591 training images (before augmentation), the high complexity of the feature set will be difficult to generalize. To prevent this, a truncated VGG16 architecture is proposed which limits the complexity of the features. The first four convolution blocks of the VGG16 architecture are used for the proposed truncated architecture as shown in Fig. 3. The truncation of the architecture reduces model complexity and number of trainable parameters, which eventually helps in reducing overfitting. This technique has also been used in Inception-net based COVID classification in [39] where the authors truncated the architecture to reduce overfitting. The truncation layer is determined by evaluating performance on the validation set with different points of truncation as detailed out in Table 4.

2.5 Transfer learning

Training a neural network from scratch requires huge amounts of data. As the COVID-19 dataset available is significantly smaller, transfer learning is applied to extract an accurate and concise feature set from the training data. This is a popular technique and has also achieved great results in [11, 13, 17, 19].

In the proposed methodology, a representation learning-based approach is used. A pre-trained VGG-16 model is fine-tuned and its intermediate outputs act as a representation of raw data. This representation serves as features for the classifier module. The first four blocks of the VGG16 architecture pre-trained on ImageNet weights are used for this purpose [13, 17]. Since the Image net set is non-overlapping to the problem, the last 8 layers, i.e.,

Table 2 The brief details of the dataset for the proposed model

Dataset	COVID	Non-COVID	Total
D1	233 images (training—204, and validation—29)	358 images (training—228, validation—33, and test—97)	591
D2	53 images (test—53)	0 images	53
D3	58 images (test—58)	0 images	58
Total	344 images	358 images	702
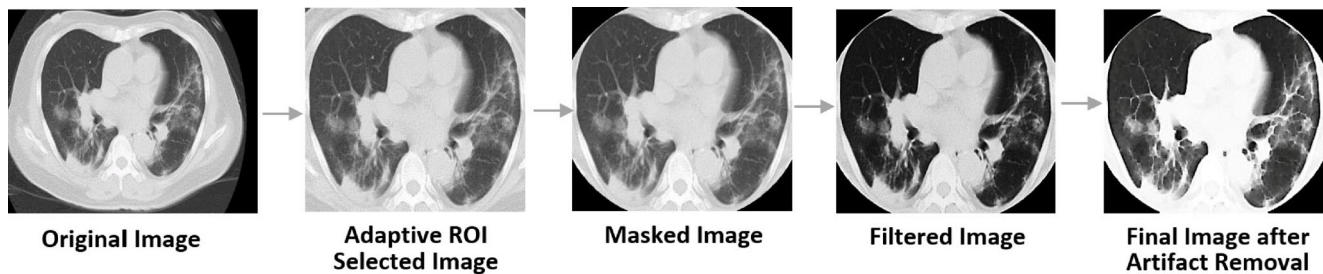
Fig. 2 Pictorial representation of various stages of the pre-processing module

the third and fourth convolution blocks are fine-tuned on the augmented CT scan training data [40]. While training these, it is desired that the fourth block adapts more to the data compared to the third block. The third block carries relatively fewer complex features that do not need to change much. Hence, a higher learning rate has been used for the fourth convolutional block compared to the third convolutional block while fine-tuning [11]. The extracted features are displayed as a color map as shown in Fig. 4. Figure 5 shows the confusion matrices of the proposed model with and without fine-tuning of the VGG16 based feature extractor.

The feature extractor module reduces the dimension of the data to 25,000 features per image for an image size of $112 \times 112 \times 3$ pixels. However, with only 591 training examples (before augmentation), the model would still overfit the features. To prevent this, feature selection and dimensionality reduction of data are performed.

2.6 Feature selector

Principal component analysis (PCA), autoencoders, and variance-based selectors are the most popular feature selectors for image data. [41] finds PCA to perform significantly better as a feature selector on biomedical data. PCA finds the eigenvectors of a covariance matrix with the highest eigenvalues and then uses those to project the data into a new subspace of equal or fewer dimensions. Autoencoders compress the input to a lower dimension. Variance-based methods select the features which have the highest variance over the data. PCA, autoencoder, and variance-based selector have been used to reduce the

Table 3 Comparative study of various popular CNN architectures

Sr. No.	CNN architecture	Accuracy on validation set (%)
1	VGG16	79.1
2	VGG19	77.2
3	Resnet50	70.8
4	InceptionV3	72.2
5	DenseNet21	68.5

dimensionality of the feature set, and then their accuracies on the validation set are compared after classification with an SVM. Applying PCA with 95% variance representation yields 359 components. Since 95% variance is a standard value for variance-based reductions, the autoencoder and variance-based selectors were also configured to retain 95% variance of the original feature set. The results of the analysis are tabulated in Table 5. For the proposed model, PCA gives the highest accuracy because it represents the low-dimensional sample and synchronized variables. Furthermore, the extracted features from the training set are used to train the classification module to screen COVID-19 CT scans. The better performance of PCA as a feature selector has also been emphasized in [42].

2.7 Classification

For the classification task, the required features are extracted using the truncated VGG16 model and selected using PCA. In machine-learning, no single algorithm is suitable for all problems. Thus, for achieving the highest performance, 4 different classification models are evaluated. Various classification techniques used in the proposed work are as follows: (a) deep CNN, (b) bagging ensemble with SVM, (c) extreme learning machine (ELM), and (d) online sequential ELM (OS-ELM).
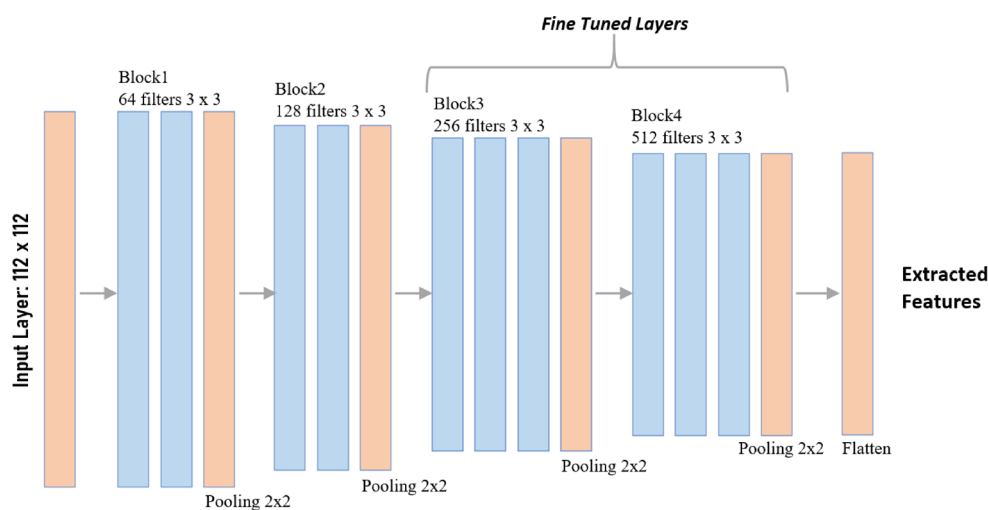
2.7.1 Deep CNN

CNN can successfully capture the spatial and temporal dependencies in an image through the application of relevant filters. The architecture performs a better fitting to the image dataset due to the reduction in the number of parameters involved and the re-usability of weights [43]. Since VGG is itself a CNN architecture, for the deep CNN model, a fully connected layer of size 1024 is added to the truncated VGG architecture followed by a softmax layer for classification. This gives us the most direct classification model where the feature extraction and classification are in the same CNN architecture. The deep CNN utilizes the fine-tuned weights and uses it to directly predict the output. Similar model has been used by authors in [11] where

Fig. 3 Architecture of truncated VGG16 model

they have used fine-tuned Resnet-50 for chest CT scan classification and achieved an accuracy of 93%.

2.7.2 Extreme learning machine

ELMs are single-hidden layer feedforward neural networks (SLFNs) that randomly choose hidden nodes and analytically determines the output weights of SLFNs through the generalized inverse operation of the hidden layer output matrices. The implementation of ELM is as described in [44]. The number of hidden nodes in the model is experimentally determined with the best-suited gamma (width multiplier for RBF distance). L2-normalized RBF activation function has been used. Experimentation has also been conducted with varying numbers of neurons in the hidden layer. Using validation set accuracy, it was observed that the highest accuracy was observed at 1000 hidden nodes. The performance of ELM is comparable to [45] where authors have used ELM to classify COVID-19 chest C-rays and have achieved an F1 score of 0.95.

2.7.3 Online sequential ELM

OS-ELM can learn data chunk by chunk with varying chunk size and provides faster sequential learning. The implementation of the model is the same as that described in [46]. It uses the idea of ELMs with a sequential

Table 4 Summary of various VGG16 truncation point accuracy evaluated on the validation set with SVM as classifier

Sr. No.	Truncation point	Accuracy on validation set (%)
1	3 blocks	73.6
2	4 blocks	84.2
3	Un-truncated	79.1

determination of the output weights through the recursive least-squares (RLS) algorithm. OS-ELM consists of two phases, namely an initialization phase and a sequential learning phase. In the initialization phase, a base extreme learning machine model is trained using a small chunk of initial training data. For classification using OS-ELM, SLFN is implemented with a sigmoid activation function with 2500 hidden layers. As the model has very few hyperparameters, they have been optimized using a grid-search optimization.

2.7.4 Bagging ensemble with SVM

To improve the limited performance of the SVM (accuracy of 93.4%) due to the high complexity of time and space, the SVM ensemble with bagging is used. A single classifier may have a high test-error, but many small classifiers can produce a low test error and increase robustness because diversity compensates for error. For classification using the bagging SVM, the dataset is randomly divided into 10 parts. The individual classifiers are trained independently with the bootstrap technique and aggregated to make a joint decision by the deterministic averaging process. The proposed classifier model with "RBF" kernel and tuned hyperparameters is used as the base estimator. Bagging ensemble with SVM achieves the highest accuracy of 95.7% on the testing data. Due to the high accuracy of the bagging with the SVM model, it is the proposed classification method for COVID-19 screening. SVM also achieved exceptional results in biomedical studies such as [47, 48].
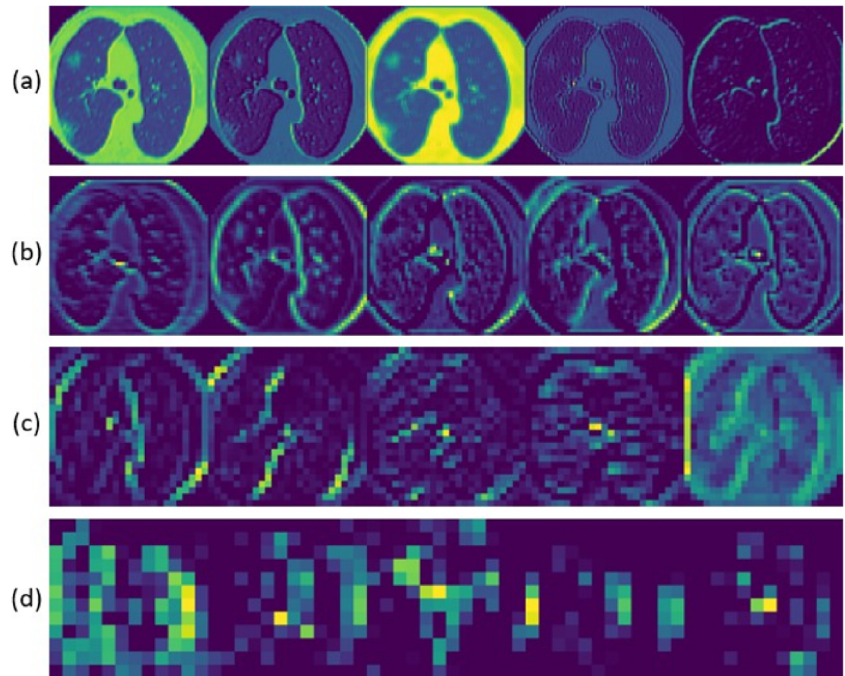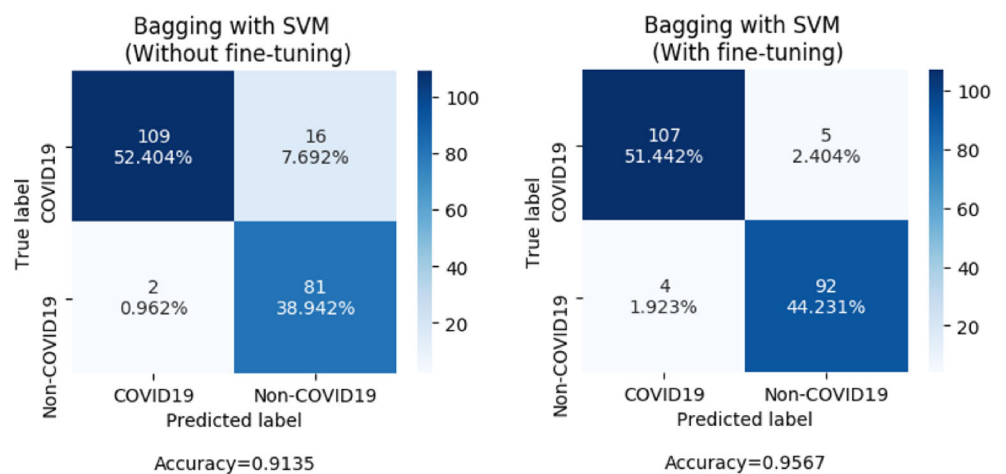
2.8 Evaluation metrics

Confusion matrices for different classifiers are shown in Fig. 8. The classifiers are evaluated on the test set with 111

Fig. 4 Intermediate color-mapped outputs. **a** Layer 1. **b** Layer 4. **c** Layer 8. **d** Layer 14

Fig. 5 Comparision of confusion matrices before and after fine-tuning of VGG16 by evaluation on the test set with bagging SVM as the classifier

Table 5 Performances analysis of feature selection techniques on validation set using SVM as classifier

Sr. No.	Feature selection technique	Validation accuracy (%)
1	PCA	93.4
2	Autoencoder	89.6
3	Variance-based selector	87.3

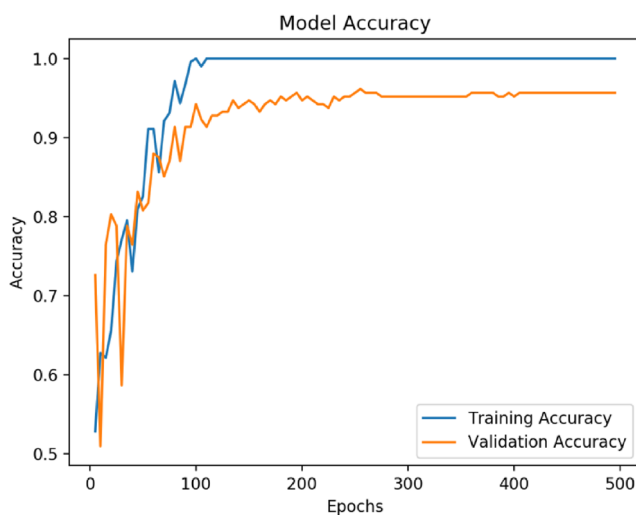**Table 6** Performance parameters of different classifiers on testing data

Classifier	TP	TN	FP	FN	AUC	PRE	NPV	S1	S2	F1	ACC
Bagging with SVM	107	92	5	4	95.8	0.955	0.958	0.963	0.948	0.959	0.957
ELM	107	88	9	4	93.8	0.922	0.956	0.963	0.907	0.942	0.937
OS-ELM	107	90	7	4	94.9	0.938	0.957	0.963	0.927	0.951	0.947
Deep CNN	103	82	15	8	89.5	0.872	0.911	0.927	0.845	0.899	0.889

COVID-19 images and 97 non-COVID images. The features for the model are extracted using the truncated VGG16 model and selected using PCA. The screening performance of the model is assessed using generalized performance parameters derived from the confusion matrix. Table 6 put forth the generalized performance parameters, namely, true positive (TP), false positive (FP), true negative (TN), false negative (FN), AUC, accuracy (ACC), precision (PRE), sensitivity (S1), specificity (S2), negative predictive value (NPV), and F1 score (F1).

3 Experimentation

3.1 Testing environment

The proposed methodology is implemented on python software, run on a CPU. The system requirements are an Intel Core i7 processor with a 4 GB graphic card, a 64-bit operating system at 1.80 GHz, and 16 GB RAM.
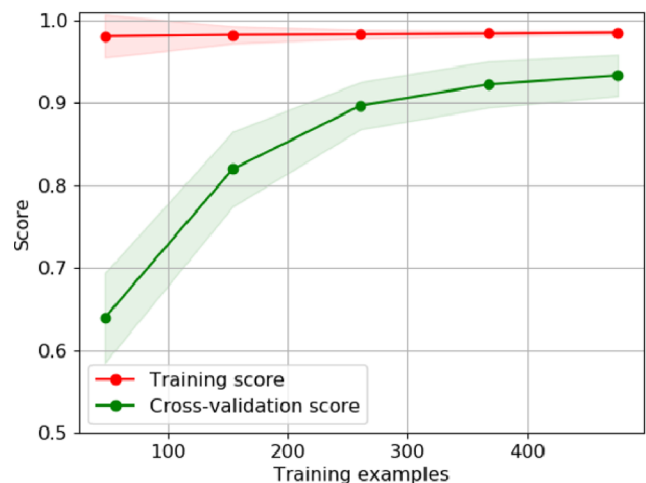
3.2 Screening of COVID-19 based on different classifier

Figure 6 shows the convergence graph of training and validation accuracy of the transfer learning–based CNN model for the bagging ensemble classifier with SVM. Figure 5 shows the confusion matrices of the proposed architecture with and without fine-tuning of the VGG16 based feature extractor. The confusion matrices are obtained by evaluating the models on the test set with bagging SVM as the classifier.

3.3 *K*-fold cross-validation

K-fold cross-validation divides the training set into k-subsets and trains k-models such that one of the subsets is left out while training the model. The accuracy of these k models is then averaged. The benefit of k-fold is that the entire dataset is used for both training and validation, and each sample is used for validation exactly once. The 10-fold cross-validation curve obtained for the bagging SVM is shown in Fig. 7. This is a 10-fold cross-validation curve

Fig. 6 Convergence graph of accuracy vs epoch for proposed methodology (VGG16+PCA+bagging ensemble with SVM)

Fig. 7 Learning curve for proposed method using 10-fold cross-validation

plotted against the number of training examples. It compares the robustness of the model with experience and shows that the proposed model is well generalized. The average scores of 10-folds have been shown in the plot.
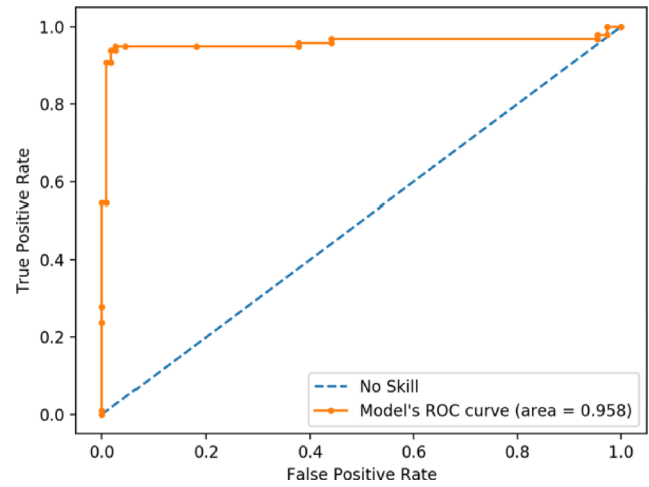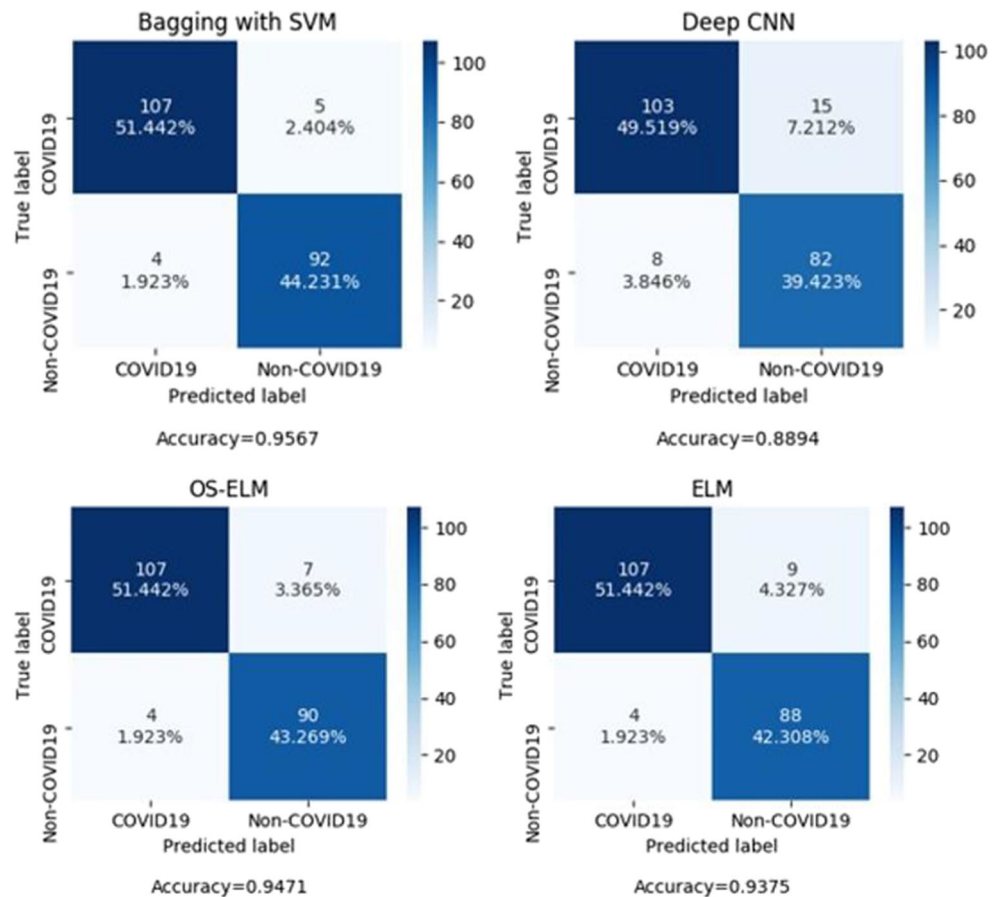
3.4 Adversial defense

Deep learning models are often fooled with noise perturbations in the image. Such perturbations or attacks lead to miss classification of images. To defend the model against such noise attacks, a defense module has been designed. To remove noise from an image before prediction, three image denoiser have been applied namely total variation, Gaussian filter, and wavelet denoising. The prediction of all three denoised images is passed to an ensemble which finally classifies the image as shown in Fig. 8. On evaluating this module with the test set after adding random noise, the model gave an accuracy of 82.34%.

4 Results

In the proposed work, the best performing model achieves an accuracy of 95.67% along with a precision of 96.83%.

Fig. 9 ROC characteristics curve for the proposed methodology (VGG16+PCA+bagging ensemble with SVM)

The area under the ROC curve (AUC) obtained is 95.8, as shown in Fig. 9. The proposed method aims to reduce the false-negative rate as much as possible since false-positive cases can potentially be identified in subsequent tests, but false-negative cases might not have that chance. The proposed model has a false negative rate of 4.33%,

Fig. 8 Confusion matrices of the proposed methodology with different classifiers

Table 7 Comparative analysis of COVID-19 detection proposed methodology with techniques available in the literature on the used dataset

Sr. No.	Techniques	Dataset	Performance evaluation
1.	DECAPS + Peekaboo [27]	Binary classification of total 746 chest CT images COVID-19 and non-COVID-19 [15]	Accuracy - 87.6%, AUC- 0.961, and precision - 84.3%.
2.	Resnet50 and VGG16 [13]	Total 102 X-ray images of COVID-19-positive and -pneumonia patients [24], [49], [50], [14]	Overall accuracy achieved is 89.2%.
3.	AI methods (JCS and DenseNet169) [15]	Binary classification of CT scan data into COVID (349 images) and non-COVID (463 CT images) [51]	Accuracy - 0.83, AUC - 0.95, and F1 - 0.85.
4.	Proposed methodology VGG16+PCA+Bagging Ensemble with SVM	Binary classification (COVID-19 and non-COVID-19) using 702 CT scan images (344 COVID-19 images and 358 non-COVID images) [14], [24], and [15].	Prediction time is 385ms, Accuracy - 95.7%, Precision - 95.8%, AUC - 0.958, and F1 score - 95.3%.

which is significantly lower than other COVID-19 CT scan screening models. The model proposed in this study achieves a very high accuracy of 95.67% on the testing data with a very low prediction time of 358 ms. This proves that deep learning-based approaches could be used to effectively and accurately screen COVID-19 at very large scales. Table 7 puts forth the comparative analysis of the proposed methodology with other existing techniques.

5 Conclusion

A deep learning-based truncated VGG16 model is proposed in this study to screen COVID-19 patients using chest CT scans. The VGG16 architecture is fine-tuned and used to extract features from CT scan images. An interesting observation has been that pre-trained models are able to learn features very effectively with fine-tuning. The study demonstrates that VGG16 outperforms other models on biomedical image feature extraction. Another important observation has been that truncation improves the model's performance on the limited dataset. The performance also improves on applying dimensionality reduction techniques, indicating a high correlation among features that need to be removed to boost the performance of the classifier. It was experimentally found that PCA performs much better than auto-encoders for biomedical image features, as has been verified by several other studies. Finally, a comparison of various popular classification techniques shows us that a bagging ensemble of SVM gives the best results and outperforms direct CNN classification, ELM, and OS-ELM.

Our study addresses the challenges associated with the limited and poor quality of COVID-19 radiology data. The study has proposed techniques like transfer learning, fine-tuning, model-truncation, image augmentation, and feature-reduction to overcome these. This should be helpful for practitioners aiming to use these datasets for their research and development. Furthermore, the importance of various pre-processing techniques has also been highlighted. While comparing different pre-trained models, it was found that both VGG16 and VGG19 gave great results within the constraints of the small dataset. While deeper networks were found to struggle, they will perform better when larger datasets are available. For currently available datasets, we conclude that VGG16 with appropriate truncation and fine-tuning gives the best feature maps. These features give good results when classified directly but the results are found to improve significantly when a feature selector like PCA is applied before classification.

With the quality of available COVID-19 data increasing, clinically verified, and trusted deep learning models may be developed for fast diagnosis of COVID-19. The superior performance of the deep models may result in AI-based

diagnosis for various diseases especially in times of outbreak where rapid screening and early identification are crucial for effective containment. As the quality of available data increases, deep models can be very effective for diagnosis. Future works on the study would focus on making the model interpretable and subjective in classification with a severity score which may be useful in the case of screening. Furthermore, a region of infection marking may assist the medical staff to treat the infection.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

References

1. WebMD Lung Diseases Overview (2020) https://www.webmd.com/lung/lung-diseases-overview#1, Last accessed date 22nd May

2. Novel Coronavirus World Health Organization (WHO) (2020) https://www.who.int/emergencies/diseases/novel-coronavirus-2019, Last accessed date 22nd May

3. Hamid RA, Alwan JK, Albahri AS (2020) Role of biological data mining and machine learning techniques in detecting and diagnosing the novel coronavirus (covid-19): a systematic review. J Med Syst 44:122. https://doi.org/10.1007/s10916-020-01582-x

4. Worldometer Global COVID-19 Case Counter (2020) https://www.worldometers.info/coronavirus, Last accessed date 22nd May

5. Tahamtana A, Ardebili A (2020) Real-time rt-pcr in covid-19 detection: issues affecting the results. Expert review of molecular diagnostics 20:453–454. https://dx.doi.org/10.1080%2F14737159.2020.1757437

6. Tao A, Yang Z, Hou H, Zhan C, Chen C, Lv W, Tao Q, Sun Z, Xia L (2020) Correlation of chest ct and rt-pcr testing for coronavirus disease 2019 (covid-19) in china: A report of 1014 cases. Radiology 296(2):E32–E40. https://doi.org/10.1148/radiol.2020200642. https://doi.org/10.1148/radiol.2020200642. PMID: 32101510

7. Albahri OS, Al-Obaidi JR, Zaidan AA, et al. (2020) Helping doctors hasten covid-19 treatment: towards a rescue framework for the transfusion of best convalescent plasma to the most critical patients based on biological requirements via ml and novel mcdm methods. Computer Methods and Programs in Biomedicine 196:105617. https://doi.org/10.1016/j.cmpb.2020.105617. ISSN 0169-2607.

8. Albahri AS, Al-Obaidi JR, Zaidan AA, et al. (2020) Multi-biological laboratory examination framework for the prioritization of patients with covid-19 based on integrated ahp and group vikor methods. International Journal of Information Technology & Decision Making 19(05):1247–1269. https://doi.org/10.1142/S0219622020500285. https://doi.org/10.1142/S0219622020500285

9. Ye Z, Zhang Y, Wang Y, Huang Z, Song B (2020) Chest ct manifestations of new coronavirus disease 2019(covid-19): a pictorial review. European Radiology 30:4381–4389. https://doi.org/10.1007/s00330-020-06801-0

10. Albahri OS, Zaidan AA, et al. (2020) Systematic review of artificial intelligence techniques in the detection and classification of covid-19 medical images in terms of evaluation and benchmarking: taxonomy analysis, challenges, future solutions and methodological aspects. Journal of Infection and Public Health 13(10):1381–1396. https://doi.org/https://doi.org/10.1016/j.jiph.2020.06.028. ISSN 1876-0341

11. Pathak Y, Shukla PK, Tiwari A, Stalin S, Singh S, Shukla PK Deep transfer learning based classification model for covid-19 disease. IRBM, 2020. ISSN 1959-0318. https://doi.org/https://doi.org/10.1016/j.irbm.2020.05.003

12. Chowdhury MEH, Rahman T, Khandakar A, Mazhar R, Kadir MA, Mahbub ZB, Islam KR, Khan MS, Iqbal A, Al-Emadi N, et al. (2020) Can ai help in screening viral and covid-19 pneumonia?. arXiv:2003.13145

13. Hall LO, Paul R, Goldgof DB, Goldgof GM (2020) Finding covid-19 from chest x-rays using deep learning on a small dataset. arXiv:2004.02060

14. Cohen JP, Morrison P, Dao L (2020) Covid-19 image data collection. arXiv:2003.11597

15. Zhao J, Zhang Y, He X, Xie P (2020) Covid-ct-dataset: a ct scan dataset about covid-19. arXiv:2003.13865

16. Zhang R, Guo Z, et al. Covid19xraynet: a two-step transfer learning model for the covid-19 detecting problem based on a limited number of chest x-ray images. 12(4):555–565, Dec 2020. ISSN 1867-1462. https://doi.org/10.1007/s12539-020-00393-5

17. Luz E, Silva PL, Silva R, Silva L, Moreira G, Menotti D (2020) Towards an effective and efficient deep learning model for covid-19 patterns detection in x-ray images. arXiv:2004.05717

18. Wang L, Lin ZQ, Wong A (2020) A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images. Scientific Reports 10(1):1–12

19. Mangal A, Kalia S, Rajgopal H, Rangarajan K, Namboodiri V, Banerjee S, Arora C (2020) Covidaid: Covid-19 detection using chest x-ray. arXiv preprint arXiv:2004.09803

20. Basu S, Mitra S, Saha N (2020) Deep learning for screening covid-19 using chest x-ray images. arXiv:2004.10507

21. Ilyas M, Rehman H, Nait-ali A (2020) Detection of covid-19 from chest x-ray images using artificial intelligence: an early review. arXiv:2004.05436

22. Minaee S, Kafieh R, Sonka M, Yazdani S, Soufi GJ (2020) Deep-covid: predicting covid-19 from chest x-ray images using deep transfer learning. arXiv:2004.09363

23. Yu-Huan W, Gao S-H, Mei J, Jun X, Fan D-P, Zhao C-W, Cheng M-M (2020) Jcs: an explainable covid-19 diagnosis system by joint classification and segmentation. arXiv:2004.07054

24. Italian Society of Medical and Interventional Radiology (2020) https://www.sirm.org/en/category/articles/covid-19-database/, Last accessed date 7th May

25. Sajid N (2020) Covid-19 patients lungs x ray images 10000. https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images

26. Jaiswal A, Tiwari P, Rathi V, Qian J, Pandey H, Albuquerque VHC (2020) Covidpen: a novel covid-19 detection model using chest x-rays and ct scans. 07. https://doi.org/10.1101/2020.07.08.20149161

27. Mobiny A, Cicalese PA, Zare S, Yuan P, Abavisani M, Wu CC, Ahuja J, de Groot PM, Van Nguyen H (2020) Radiologist-level covid-19 detection using ct scans with detail-oriented capsule networks. arXiv:2004.07407

28. Narayan Das N, Kumar N, Kaur M, Kumar V, Singh D (2020) Automated deep transfer learning-based approach for detection of covid-19 infection in chest x-rays. IRBM. ISSN 1959-0318. https://doi/https://doi.org/10.1016/j.irbm.2020.07.001. http://www.sciencedirect.com/science/article/pii/S1959031820301172

29. Singh D, Kumar V, Vaishali, Kaur M (2020) Classification of covid-19 patients from chest ct images using multi-objective differential evolution-based convolutional neural networks. European journal of clinical microbiology & infectious diseases : official publication of the European Society of Clinical Microbiology 39(7):1379–1389. ISSN 1435-4373. https://doi.org/10.1007/s10096-020-03901-z

30. Liu K-C, Xu P, Lv W-F, Qiu X-H, Yao J-L, Gu J-F, Wei W (2020) Ct manifestations of coronavirus disease-2019: a retrospective analysis of 73 cases by disease severity. European Journal of Radiology 126:108941. ISSN 0720-048X. https://doi/https://doi.org/10.1016/j.ejrad.2020.108941. http://www.sciencedirect.com/science/article/pii/S0720048X20301303

31. Gozes O, Frid-Adar M, Sagie N, Zhang H, Ji W, Greenspan H (2020) Coronavirus detection and analysis on chest ct with deep learning. arXiv:2004.02640

32. Kanne J, et al. (2020) Essentials for radiologists on covid-19: an update—radiology scientific expert panel. Radiology 296(2):E113–E114. https://doi.org/10.1148/radiol.2020200527. https://doi.org/10.1148/radiol.2020200527

33. Zubair AR, Busari H (2018) Robustness of median filter for suppression of salt and pepper noise (spn) and random valued impulse noise (rvin). IJIP 12:12–27

34. Jamil N, Sembok TMT, Bakar ZA Noise removal and enhancement of binary images using morphological operations. 4:1–6, 2008. https://doi.org/10.1109/ITSIM.2008.4631954

35. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. Journal of Big Data 6(1):60. ISSN 2196-1115. https://doi.org/10.1186/s40537-019-0197-0. https://doi.org/10.1186/s40537-019-0197-0

36. Silva P, Luz E, Silva G, Moreira G, Silva R, Lucio D, Menotti D (2020) Covid-19 detection in ct images with deep learning: a voting-based scheme and cross-datasets analysis. Informatics in Medicine Unlocked 20:100427. ISSN 2352-9148. https://doi/https://doi.org/10.1016/j.imu.2020.100427. http://www.sciencedirect.com/science/article/pii/S2352914820305773

37. Horry MJ, Chakraborty S, Paul M, Ulhaq A, Pradhan B, Saha M, Shukla N (2020) Covid-19 detection through transfer learning using multimodal imaging data. IEEE Access 8:149808–149824. https://doi.org/10.1109/ACCESS.2020.3016780

38. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition

39. Das D, Santosh KC, Pal U (2020) Truncated inception net: Covid-19 outbreak screening using chest x-rays. Physical and Engineering Sciences in Medicine 43(3):915–925. ISSN 2662-4737. https://doi.org/10.1007/s13246-020-00888-x. https://doi.org/10.1007/s13246-020-00888-x

40. Horry MJ, Chakraborty S, Paul M, Ulhaq A, Pradhan B, Saha M, Shukla N (2020) Covid-19 detection through transfer learning using multimodal imaging data. IEEE Access 8:149808–149824. https://doi.org/10.1109/ACCESS.2020.3016780

41. Shaltout N, Moustafa M, Rafea A, Moustafa A, ElHefnawi M (2015) Comparing pca to information gain as a feature selection method for influenza-a classification. In: 2015 International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), pages 279–283. https://doi.org/10.1109/ICIIBMS.2015.7439550

42. Ceci M, Corizzo R, Japkowicz N, Mignone P, Pio G (2020) Echad: embedding-based change detection from multivariate time series in smart grids. IEEE Access 8:156053–156066. https://doi.org/10.1109/ACCESS.2020.3019095

43. He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. arXiv:1512.03385

44. Huang G-B, Zhu Q-Y, Siew C-K (2006) Extreme learning machine: theory and applications. Neurocomputing 70(1):489–501. ISSN 0925-2312. https://doi.org/10.1016/j.neucom.2005.12.126. Neural Networks

45. Rajpal S, Rajpal A, Lakhyani N, Kumar N (2020) Cov-elm classifier: an extreme learning machine based identification of covid-19 using chest x-ray images. arXiv:2007.08637

46. Liang N, Huang G, Saratchandran P, Sundararajan N (2006) A fast and accurate online sequential learning algorithm for feedforward networks. IEEE Transactions on Neural Networks 17(6):1411–1423. https://doi.org/10.1109/TNN.2006.880583

47. Mignone P, Pio G, D'Elia D, Ceci M (2019) Exploiting transfer learning for the reconstruction of the human gene regulatory network. Bioinformatics 36(5):1553–1561. ISSN 1367-4803. https://doi.org/10.1093/bioinformatics/btz781. https://doi.org/10.1093/bioinformatics/btz781

48. Krzysztof A, Cyran JK, Kawulok M, Stawarz M, Michalak M, Pietrowska M, Widłak P, Polańska J Support Vector Machines in Biomedical and Biometrical Applications, pages 379–417. Springer Berlin Heidelberg, Berlin, Heidelberg, 2013. ISBN 978-3-642-28699-5. https://doi.org/10.1007/978-3-642-28699-5_15.

49. Covid-19 chest ct scans Radiopedia (2020) https://radiopaedia.org/search?utf8=%E2%9C%93&q=covid&scope=all&lang=us.. Last accessed date 7th May 2020

50. NIH (2020) Nih chest x-ray dataset. https://www.kaggle.com/nih-chest-xrays/data. Last accessed date 7th May 2020

51. He X, Yang X, Zhang S, Zhao J, Zhang Y, Xing E, Xie P (2020) Sample-efficient deep learning for covid-19 diagnosis based on ct scans. medRxiv. https://doi.org/10.1101/2020.04.13.20063941. https://www.medrxiv.org/content/early/2020/04/17/2020.04.13.20063941

Mukul Singh is an undergraduate student in the Department of Computer Science and Engineering at Indian Institute of Technology Delhi, New Delhi, India. His research interests include deep learning, computer vision, image processing, data analytics and cybersecurity.

Shrey Bansal is currently doing his B.Tech with the Department of Computer Science Engineering, Indian Institute of Technology Delhi, New Delhi, India. His interests include neural networks, computer vision and image processing, and data analytics.

Sakshi Ahuja received a Masters degree in Electrical Engineering from Deenbandhu Chhotu Ram University of Science and Technology Haryana, India, in 2016. She is currently working toward a Ph.D. degree with the Department of Electrical Engineering, Indian Institute of Technology Delhi, and New Delhi, India. Her interest includes neural networks, computer vision and image processing, and data analytics.

Rahul Kumar Dubey received his Ph.D. degree in power system engineering from the Department of Electrical Engineering, IIT, New Delhi, India, in 2016. He was a Research Engineer (2014–2016) with the Robert Bosch Corporate Research, India, where he was mainly involved in the area of power-system and quality. He was a Research and Development Lead (2016–2017) with the ABB Ability Innovation Centre, PGGA Unit, and India, where he was mainly involved in the areas of power system protection and monitoring. He was a Postdoctoral Member (2017–2018) with IEPG group TU Delft, The Netherlands. He is currently working as an Senior Architect in Robert Bosch Corporate Research, India, where he was mainly involved in the areas of research and innovation in AI/ML. His research interests include artificial intelligence, deep learning,vehicle telemetries data analytic,battery management(RUL and SOC), connected mobility, natural language processing, power system protection, digital signal processing, soft computing, flexible ac transmission systems, and wide-area monitoring systems. He is the recipient of the 2014 prestigious Prime Minister's Fellows Award by the Government of India, the 2015 Clayton Griffin award at Georgia Tech USA and, the 2017 POSOCO award for best research work by Power Grid India. He serves as an Associate Editor for the Cambridge Scholar Press Book, International Journal of Electrical Engineering Education(University of Manchester UK, SAGE Publishing) and a Senior Member of the IEEE.

Bijaya Ketan Panigrahi received the Ph.D. degree in power systems from Sambalpur University, Sambalpur, India, in 2004. Since 2005, he has been an Associate Professor with the Department of Electrical Engineering, Indian Institute of Technology (IIT)Delhi, New Delhi, India, where he became a Professor in 2017. Prior to joining IIT, he was a Lecturer with the University College of Engineering, Sambalpur, India, for 13 years. His research interests include intelligent control of flexible ac transmission system devices, digital signal processing, power quality assessment, and application of soft computing techniques to power system planning, operation, and control. He is an associate editor for IEEE Systems Journal and a Senior Member of the IEEE.

Nilanjan Dey was born in Kolkata, India, in 1984. He received his B.Tech. and M.Tech. degrees in information technology from the West Bengal University of Technology, in 2005 and 2011, respectively, and the Ph.D. degree in digital image processing from Jadavpur University, India, in 2015. In 2011, he was appointed as an Assistant Professor with the Department of Information Technology, JIS College of Engineering, Kalyani, India, followed by the Bengal College of Engineering College, Durgapur, India, in 2014. He is currently an Assistant Professor with the Department of Information Technology, JIS University, Kolkata, India. His research interests include signal processing, machine learning, and information security. He is also an Associate Editor of IEEE Access, the Editor-in-Chief of the International Journal of Ambient Computing and Intelligence, and a Series co-editor of Springer Tracts of Nature-Inspired Computing (STNIC).

Affiliations

Mukul Singh[1] · Shrey Bansal[1] · Sakshi Ahuja[2] · Rahul Kumar Dubey[3] (iD) · Bijaya Ketan Panigrahi[2] · Nilanjan Dey[4]

Mukul Singh
Mukul.Singh.cs118@cse.iitd.ac.in

Shrey Bansal
Shrey.Bansal.cs118@cse.iitd.ac.in

Sakshi Ahuja
Sakshi.Ahuja@ee.iitd.ac.in

Bijaya Ketan Panigrahi
bkpanigrahi@ee.iitd.ac.in

Nilanjan Dey
nilanjan.dey@jisuniversity.an

[1] Computer Science and Engineering Department,
Indian Institute of Technology Delhi, New Delhi,
110016, India

[2] Electrical Engineering Department, Indian Institute
of Technology Delhi, New Delhi, 110016, India

[3] Robert Bosch Engineering and Business Solutions Private
Limited Head Office, 123, Hosur Rd, 7th Block, Koramangala,
Bengaluru, Karnataka 560095, India

[4] JIS University, Kolkata, India