



Learning effective embedding for automated COVID-19 prediction from chest X-ray images

Sree Ganesh T N¹ · Rishi Satish¹ · Rajeswari Sridhar¹

Received: 15 March 2022 / Accepted: 13 October 2022 / Published online: 26 October 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

The pandemic that the SARS-CoV-2 originated in 2019 is continuing to cause serious havoc on the global population's health, economy, and livelihood. A critical way to suppress and restrain this pandemic is the early detection of COVID-19, which will help to control the virus. Chest X-rays are one of the more straightforward ways to detect the COVID-19 virus compared to the standard methods like CT scans and RT-PCR diagnosis, which are very complex, expensive, and take much time. Our research on various papers shows that the currently researchers are actively working for an efficient Deep Learning model to produce an unbiased detection of COVID-19 through chest X-ray images. In this work, we propose a novel convolution neural network model based on supervised classification that simultaneously computes identification and verification loss. We adopt a transfer learning approach using pretrained models trained on imagenet dataset such as Alex Net and VGG16 as back-bone models and use data augmentation techniques to solve class imbalance and boost the classifier's performance. Finally, our proposed classifier architecture model ensures unbiased and high accuracy results, outperforming existing deep learning models for COVID-19 detection from chest X-ray images producing State of the Art performance. It shows strong and robust performance and proves to be easily deployable and scalable, therefore increasing the efficiency of analyzing chest X-ray images with high accuracy in detection of Coronavirus.

Keywords Siamese neural network · Medical image classification · VGG16 · COVID-19 prediction · Transfer learning · AlexNet · Multitask learning · Convolution neural network

1 Introduction

In 2019, The World Health Organisation (WHO) declared that COVID-19 is a pandemic. COVID-19 has been responsible for at least 3 million excess deaths in 2020. As of 31 December 2020, COVID-19 had infected over 82 million people and killed more than 1.8 million worldwide. Common Symptoms include cough, high fever, and long-term

breathing problems. Figure 1 represents current COVID-19 cases statistics across the globe. To prevent the spread of the virus and prevent complications, it is necessary to detect the virus as early as possible. Identifying and managing the pandemic is difficult because cases of COVID-19 increase daily, and current diagnostic tools have limitations. Research is currently going on around the world to find effective diagnostic measures for developing an effective vaccine to find a cure. Detection of COVID-19 includes blood tests, viral tests, and medical imaging. COVID-19 antibodies are detected in the blood via blood tests. However, the accuracy of this test is only 2% or 3% and cannot be used to detect COVID-19. Viral tests performed on patients help to determine the presence of COVID-19 antibodies in the respiratory tract. The tests for viral infection are usually performed using reverse transcription polymerize chain reaction. It is the standard test used in airports for flight travel. However, further research concluded that the test result sensitivity ranges from 55 to 67%. To ensure the test results of the diagnosis are correct, multiple RT-PCR tests are performed over

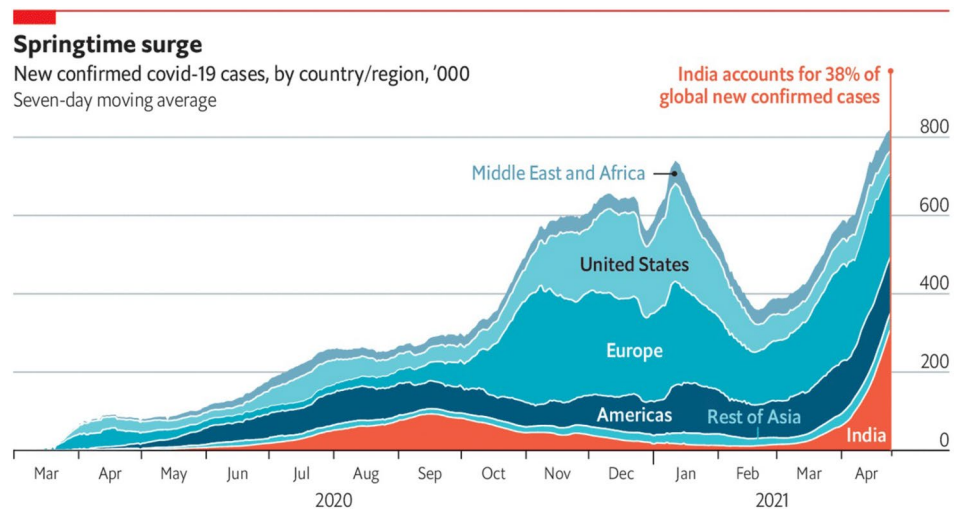
Communicated by X. Yang.

✉ Sree Ganesh T N
tnsreeganesh@gmail.com

Rishi Satish
rishi102001@gmail.com

Rajeswari Sridhar
srajeswari@nitt.edu

¹ Department of Computer Science and Engineering, National Institute of Technology, Tiruchirappalli, Tamil Nadu 620015, India

Fig. 1 Daily case statistics

14 days. A shortage of these test kits in several countries makes this stress-inducing for patients and expensive for healthcare institutions. Chest X-ray imaging is an effective method to diagnose the presence of COVID-19. Chest X-ray imaging is recommended in severely affected or limited-resource areas due to its widely available nature, cheap cost, and fast results. Methods of artificial intelligence, such as deep learning and machine learning has been proven promising for the fast diagnosis of the virus.

The literature review says that convolutional neural networks and other deep learning architectures have been used to review and examine COVID-19 images via chest X-ray. Several research papers have proposed deep learning architectural models for automatically detecting COVID-19 in chest X-ray images. A comprehensive and diverse public data set is urgently needed, even with encouraging results. For more vigorous, clear, and precise predictions, it is also necessary to conduct further investigations regarding explainable and justifiable decisions. Therefore, more error-free methods for the diagnosis are required. However, well-trained deep learning models can synthesize information that is not immediately apparent to the human eye, thereby reversing this perception. Our proposed architecture model reported accuracy of more than 98%. Detailed analysis of various research papers demonstrates that our proposed model is better than other existing models. We have conducted experiments on publicly available data sets, concluding that our model effectively gives accurate diagnostics of COVID-19.

In summary, this paper provides the following key contributions:

- A novel transfer learning based convolution neural network classifier that simultaneously computes the identification loss and verification loss using pre-trained Imagenet

models such as AlexNet and VGG16 as backbone model

- We also incorporate data augmentation techniques like rotation, translation for solving class imbalance.
- We achieve state-of-the-art COVID classification results on the COVID-19 chest X-ray images dataset in both Binary classification and Multiclass classification tasks.

The rest of the work is organised as follows: Sect. 2 consists of related works. Section 3 consists of detailed explanation of the proposed architecture of the proposed model. Section 4 consists of the results and discussions of the proposed model. Section 5 demonstrates the conclusion and the future works.

2 Related works

Ozturk et al. [1] used multi-class classification to develop the DarkCovidNet model for the detection of COVID-19. The model was developed inspired by Darknet, adding feature extraction techniques using an end-to-end architecture. The proposed model achieved a sensitivity of 85.35%.

Wang and Wong et al. [2] proposed a deep convolutional neural network model (DCNN), COVID-Net, developed by adopting a machine-driven design exploration strategy. A multi-class classification was performed using chest X-ray images to identify COVID-19 from normal and non-COVID diseases. The success of the model was determined by both qualitative and quantitative analysis. The proposed model achieved a sensitivity of 92% for detecting COVID-19.

Panwar et al. [3] introduced nCOVnet, an algorithm that was developed based on a transfer learning model for faster diagnosis of COVID-19 from chest X-ray images.

The authors use VGG16 as the base model. The model was able to achieve 96.72% in detecting COVID-19.

Sethy et al. [4], using the deep features extracted from the CNN layers and fed to the SVM classifier for classification, described a comparative study on Xception, ResNet18, ResNet50, Resnet 101, Inceptionv3, Inceptionresnetv2, GoogleNet, Densenet201, VGG16, VGG19, and AlexNet. Resnet, compared with other classification models on binary classification between COVID-19 and Pneumonia, performed better with a classification accuracy of 95.38%.

Jain et al. [5] utilized the ResNet50 model, which was used in the stage-I network model for distinguishing COVID-19 from Pneumonia and standard cases using chest X-ray images. ResNet101 was used in the stage-II network model to classify Pneumonia from COVID-19. The proposed model achieved an accuracy of 97% in COVID-19 detection.

Narin et al. [6] proposed three CNN models for detecting COVID-19 through chest X-ray images. The pre-trained models like ResNet50, InceptionV3, and Inception-ResNetV2 were chosen to conduct a comparative study. The authors proposed a binary classification between normal and COVID-19. The proposed ResNet50 model outperformed all the other models with the highest classification accuracy of 97.6

Ioannis et al. [7] compared different pre-trained models like VGG19, Mobile Net, Inception, Xception, and Inception-ResNetv2 for the automatic detection of COVID-19 using chest X-ray images. The authors implemented a multiclass classification among normal, COVID-19, and Pneumonia. MobileNetV2 achieved the highest sensitivity rate of 98.6%.

Asnaoui et al. [8] used the pre-trained DCNN models, namely VGG16, VGG19, Inception-ResNetV2, InceptionV3, ResNet50, DenseNet201, and MobileNetV2, for a comparative study on the multiclass classification of chest X-ray images into normal, bacteria, and coronavirus. Inception-ResNetV2 model provided a sensitivity rate of 93.11% in detecting coronavirus.

Ravi et al. [9] proposed large-scale learning with a stacked ensemble meta-classifier and a deep learning-based feature fusion approach for COVID-19 classification. This approach used a two-layer approach. In the first stage, random forest and support vector machine was applied and the second stage involved logistic regression classifier that classifies the data samples into COVID-19 and non-COVID-19 samples. One of the problems this paper faces is that its approach is very limited and will not handle a class imbalance in the Dataset.

Rahman et al. [10] have conducted a detailed survey of all the COVID-19 models to date. Their work has a comparison of more than 300 models. The results obtained from the paper show that the DenseNet201 model with quadratic

SVM classifier performs the best with an accuracy of 98.16%.

Misra et al. [11] presented a multi-channel transfer learning model based on ResNet architecture to facilitate the diagnosis of COVID-19 chest X-ray. Results show that the ensemble model is more accurate than the single ResNet model, which is also retrained using Dataset as it extracts more relevant semantic features for each class. The approach provides a precision of 94% and a recall of 100%.

In [12] Imran Ali et al. implemented an AI based screening solution to detect COVID, transferable through an intelligent mobile phone application was suggested, developed, and finally tested. The mobile App called AI4COVID-19 records and sends to AI-based clouds running in the cloud triple 3-s cough sounds and comeback reaction within two minutes. Generally, cough is a primary indication of over 30 medical conditions associated with non-COVID-19. This makes it a challenging multidisciplinary issue is cough alone to diagnose COVID disease. Investigating morphological direction changes with dissimilarities from cough respiratory achieves an accuracy of 88.76

Deep learning-based software system for automatic COVID-19 detection on chest CT was developed by Zheng et al. [13] using 3D CT volumes to detect COVID-19. A pre-trained UNet and a 3D deep neural network were used to predict the probability of COVID-19 infections on 630 CT scans. Of 1014 patients, 601 tested positive for COVID-19 based on RT-PCR, and the results were compared with the chest CT. CT scans, however, are very costly, and not enough CT scan images are present in order for the model to give high accuracy. However, the datasets are often not publicly available, which reduces their access to the wider research community and further development of classification techniques on standardized.

In [14], the authors used deep learning to tackle the problem of COVID-19 detection. They use the EfficientNet model on chest CT scan images. They used fixed, cyclic, and learning rates reduced on the plateau strategy to fine-tune their model, where the plateau strategy performed best on their testing model with an accuracy of 89.7% followed by the cyclic learning rate, which achieved an accuracy of 86%. In contrast, a constant learning rate yielded 83% accuracy on their test set. However, the datasets are often not publicly available, reducing their access to the broader research community and further developing classification techniques on standardized.

In [15], Mohamed Bader et al. proposed the significant model with the combination of Mel-Frequency Cepstral Coefficients (MFCCs) and SSP (Speech Signal Processing) for the extraction of samples from non-COVID and COVID, and it finds the person correlation from their relationship coefficients. These findings indicate high similarity between various breathing respiratory sounds and COVID

cough sounds in MFCCs, although MFCC speech is more robust between non-COVID-19 and COVID-19 samples. In addition, when recording their speech signals, the patients must sit with their heads straight in a comfortable way; three recordings for each speaker are acquired from smartphone devices in data collection, which can affect the quality of sound.

Hassan Abdelfatah et al. [16] implemented a system to diagnose COVID positivity using the RNN model; the authors illustrated the significant impact of RNN (Recurrent Neural Network) with the use of SSP (Speech Signal Processing) to detect the disease, and specifically, this LSTM (Long Short-Term Memory) used to evaluate the acoustic characteristics of patients' cough, breathing, and voice, in the process of early screening and diagnosing the COVID-19 virus. Compared to both coughing and breathing sound recordings, the model findings indicate poor precision in the speech test.

Fie et al. [17] presented an automatic algorithm with a Deep Learning (DL) approach for the detection of contagious points in the lungs. Xiawi et al. [18] created a primary screening model using CT imaging based on DL method to separate COVID-19 from Viral pneumonia. Brunese et al. [19] have utilized a developed VGG16-TL network to detect two COVID-19 and Healthy classes.

Stephen et al. [20] have proposed an effective DL model for pneumonia classification by chest X-ray imagery. Dansana et al. [21] used an X-ray and CT scan image dataset to train a CNN model-based VGG-19, Inception V2 and decision tree model for binary classification pneumonia. Oh et al. [22] presented a patch-based CNN approach with a relatively small number of trainable parameters for COVID-19 diagnosis. Nayak et al. [23] proposed a DL automated method for early detection of COVID-19 infection using X-ray images. Wang et al. [24] suggested a novel DTL

algorithm based on Pre-Trained Models (PTMs) to extract features.

All of the above research works modeled a paradigm that used the transfer learning approach using pre-trained deep Convolutional neural network by retraining them on the COVID X-ray image data set, but none of the existing work has considered the Siamese-based approach. All of the work has modeled the task as a classification task, but we propose to consider this task as a multitask learning by considering both a classification and verification task by incorporating both Siamese training and image classification training. With this approach, we propose that the model can effectively learn better representations of the images by maximizing the inter-class distance and minimize the intra-class difference in low dimensional semantic space, as well as classify the radiographic image to detect the presence of detection of COVID-19.

3 Methods

In this study, we introduce a novel end to end deep learning model architecture, a Siamese convolutional neural network model that learns by simultaneously computing the identification loss and verification loss. This proposed framework provides a simple, concise, and effective COVID-19 prediction from chest X-ray images supported by transfer learning of pretrained imagenet models as the backbone. Figures 2 and 3 shows the architecture and flowchart of our proposed model.

Unlike the previous methods, which classified the given chest X-ray images and predominantly used only classification loss as the primary objective function, It is indeed necessary that a Model must learn to distinguish between pairs of images along with classification by learning to

Fig. 2 Flow chart of the proposed model

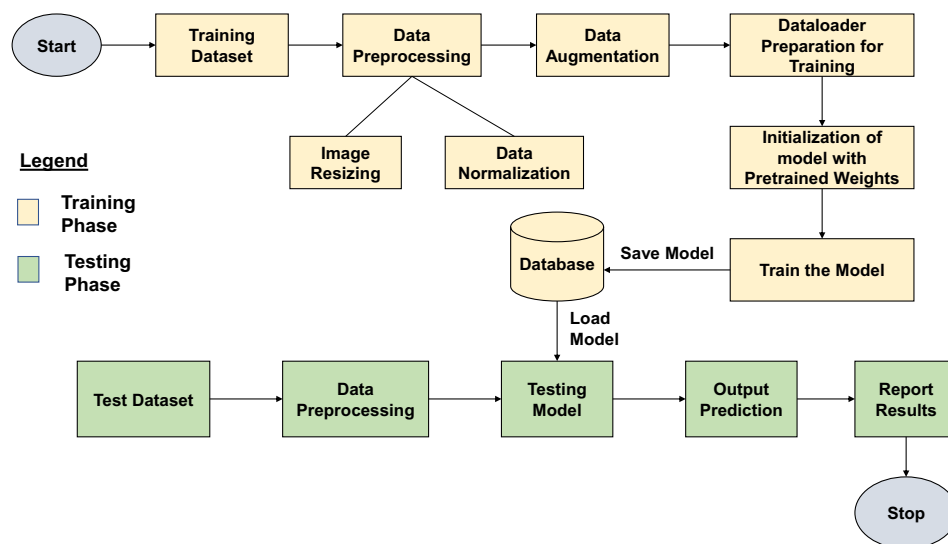
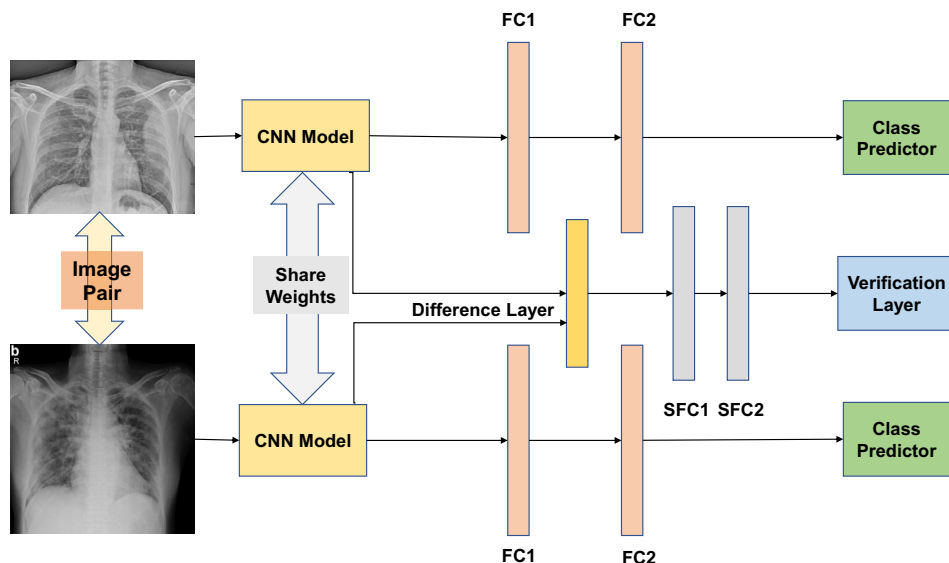


Fig. 3 Proposed model architecture



verify the classes as same or different. When these models are deployed for automated diagnostics in a real-time environment, the reliability factor of the model must be improved for model robustness and scalability. It is also quite intuitive that the majority of the chest X-ray images look similar, and it is essential to distinguish between chest X-ray images of different patients and validate between chest X-ray images. We infuse a verification loss into the model to enable the model to learn a better representation of the image feature. This verification loss additionally penalizes the model for learning more representative features for effective prediction. We approach the problem by a multi tasking Paradigm by optimizing the model that learns to perform accurate classification and identification of disease and validate and verify the chest X-ray images in a unified manner. We propose our network learns a discriminative embedding of the input images for reliable and accurate detection of COVID-19 from chest X-ray images and also validates the chest X-ray images by minimizing the intra-class distance and maximizing the inter-class distance. By this Paradigm, the model will be able to classify based on the classification score and rank the Inputs based on the similarity scores. It leverages the predictive power of the network not just to new data but to entirely new classes from unknown distributions. We optimized Two deep learning architectures and implemented development strategies toward achieving a more realistic development of AI-driven COVID-19.

We have used the pretrained imagenet models as the backbone CNN model. These models are then fine-tuned on our dataset to extract deep features and trained to make accurate predictions. This proposed architecture is scalable to other pretrained imagenet models as well. Due to the availability of limited data, data augmentation techniques were used to

enrich the existing dataset to solve the class imbalance problem for effective training and model evaluation.

In this proposed work, two models, namely, VGG16 [25] and AlexNet [26], pretrained on imagenet dataset as the backbone model as they are standard baseline models which are used in a wide variety of applications to learn high-level features in a simple deep structure and also shows good performance metrics. The models learn image patterns through a repeated stacking of convolutional filters with a small kernel to learn better about the Images and their texture. VGG16 and Alex Net models are standard CNN models used for Image classification trained on the Image Net database

VGG16 was developed by Simonyan et al. [25] in 2014 as part of the Imagenet Challenge. It has 14 convolution layers, five max-pooling layers, and three fully connected layers. It utilized a Deep layered network with small convolution filters to extract deep features and proved to be a baseline for deep architectures. AlexNet was developed by Krizhevsky et al. [26] in 2012, which is one of the earlier Deep convolutional neural network models for image classification. It has five convolution layers, three max-pooling layers, and two fully connected layers with fewer parameters. These models learn to classify the given image into respective classes by extracting deep features and learning from those embeddings while addressing the overfitting problem. Here we removed the final fully connected layers of the model and added additional fully connected layers.

The framework consists of two identical backbone models that share their parameters and weights and learn to generate feature vectors and verify whether it belongs to the same class or is different, along with identifying which class the image belongs to. This Approach aims to regularize the training process, make the Model robust to class imbalance problems, and learn semantic similarities

between the images. This approach has been proven effective in different tasks like person re-identification [27], Face verification [28, 29], Signature verification [30] and even in brain tumor classification. Deepak and Ameer [31]

Given a pair of training images from the dataset, the network makes two predictions: the classes of the two input images, which contribute toward the identification loss and similarity score of the image pair. The chest X-ray images are of varying resolutions with variable aspect ratios. So all the training and testing images are resized to $224 \times 224 \times 3$ as a part of data pre-processing. Positive pair of images contributes to images of the same class, while images of different classes contribute to negative pair.

The network consists of 2 ImageNet Pretrained models, six additional Fully connected layers and a Square layer, and three losses. The proposed model takes 2 input images, and it is forward passed through the backbone CNN model to get the salient feature embedding. both feature embeddings are passed through fully connected layers FC1 and FC2 to predict the class and compute identification loss. We introduce a new layer called a difference layer for the Siamese training, which computes the squared difference between the feature embedding that learns the semantic similarity of the features in low dimensional space and the resultant vector is passed through 2 fully connected layers-SFC1 and SFC2 to verify both images belong to the same class or different and compute the verification loss and weighted sum of all the Loss is backpropagated and Model is trained in an end to end fashion. Rectified linear unit (ReLU) activation function has been used throughout the study. So the similarity between learned embeddings are eventually closer for images within the same class and far away for images in the different classes.

3.1 Loss

The loss function is the measure of performance of a deep learning model. Models learn weights and biases to minimize the loss function between original and predicted values. Identification loss is the loss associated with the class prediction of the model. Verification loss is associated with the similarity of the Images as to whether they belong to the same class or are different. We use cross-entropy loss as the loss function for all the losses. Consider p_1 and p_2 as the target class probabilities of the Training pair and p_s as the target verification probability of the pair, which is a K-dimensional vector with the index of the target class as 1 and the rest of the values as 0. Let \hat{p}_1 and \hat{p}_2 are predicted class probabilities. Let \hat{p}_s be the predicted verification probability and K be the number of classes. The classification Loss is given by:

$$\mathcal{L}_{\text{cls1}} = \sum_{i=1}^K -p_{1i} \log(\hat{p}_{1i}) \quad (1)$$

$$\mathcal{L}_{\text{cls2}} = \sum_{i=1}^K -p_{2i} \log(\hat{p}_{2i}) \quad (2)$$

Here p_{ji} and \hat{p}_{ji} denotes the true target probability and predicted probability of the i th class. The verification loss is given by:

$$\mathcal{L}_{\text{veri}} = \sum_{i=1}^2 -p_{si} \log(\hat{p}_{si}) \quad (3)$$

We have a defined a weighted objective function in which the model proposed in this work is optimized over these losses simultaneously under single overall objective. We propose The overall objective function with regularization parameter λ as given by

$$\mathcal{L} = L_{\text{veri}} + \lambda L_{\text{cls1}} + \lambda L_{\text{cls2}} \quad (4)$$

We empirically set the value of λ value as 0.5 throughout our experiments and use the same value of λ for both L_{cls1} and L_{cls2} . We have provided in the further implementation details Sect. 4.

3.2 Execution

According to the flowchart, first, the images are preprocessed, normalized, and augmented using data augmentation techniques. The dataset is split into train and test dataset and training, and the testing data loader is prepared based on the algorithm mentioned in the next section, and the model is then trained on the Augmented preprocessed images and saved in the database. During the testing phase, the test data loader is prepared by considering one image at a time, the model is loaded from the database, the output class is predicted using the trained model over the test dataset individually, and the performance metrics are analyzed. It is to note that we use only the class predictor and do not use the verification layers during the inference phase. In the next section, we cover the experimental analysis of this study.

4 Results and discussion

In this study, We have adopted two different pre-trained CNN models- VGG16 and AlexNet. Furthermore, We performed extensive experiments to detect and classify COVID-19 samples using X-ray images in two different scenarios. We experimentally verify the proposed algorithm using a publicly available dataset of chest X-ray images. The

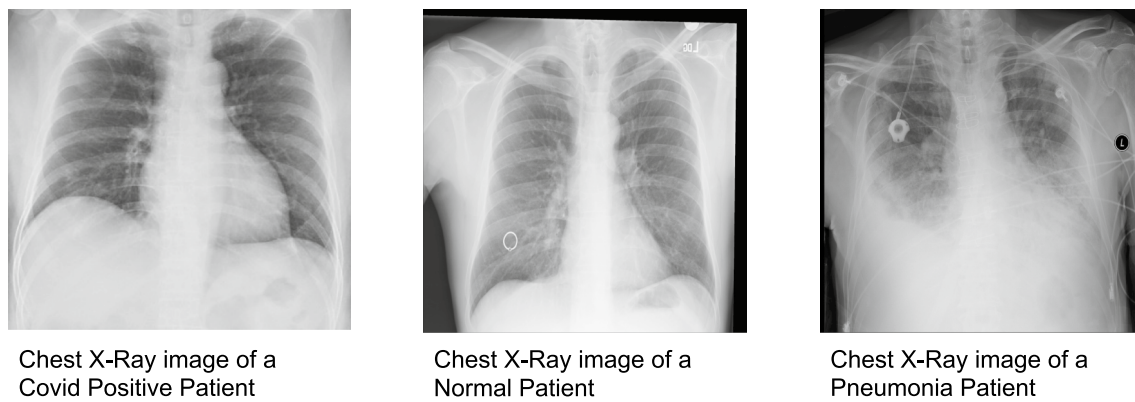


Fig. 4 Sample chest X-ray images from the dataset [10]

Table 1 Dataset distribution

Class	Train	Test
COVID-19	421	273
Normal	1884	1340
Pneumonia	2463	1719
Total	4768	3332

proposed framework is trained for two tasks—binary classification and multi-class classification. In binary classification, the model is trained to classify the given chest X-ray as COVID positive or COVID negative, while in the multi class classification, the model is trained to detect the chest X-ray as COVID positive, normal, and viral pneumonia. We have used the dataset prepared by [10] as our primary dataset.

4.1 Dataset

Rahman et al. [10] consists of COVID-19 X-rays collected from various sources. It consists of a total of 7879 distinct chest X-ray images classified into COVID-19, normal and viral pneumonia. Figure 4 shows some sample images from the dataset. The dataset in [10] consisted of 683 images of COVID 19 samples, 2924 of normal samples, and the rest are of viral pneumonia X-ray images. Rahman et al. [10] proposed a train-test split strategy where 4768, which is 60% of the total images, are used for training and the rest for testing. The Table 1 shows the dataset distribution of [10]

It is observed that data is very Imbalanced since the ratio of COVID samples to the rest is very low. This might lead to an overfitting problem and can lead to wrongful predictions. So, to tackle the imbalance, we are using data augmentation techniques—specifically-rotation, horizontal flipping, and translation as the techniques to enrich the dataset so that imbalance is reduced by a greater extent. The following Table 2 shows the training dataset before and after applying augmentation technique for dataset [10]. Figure 5 presents

Table 2 Train dataset after applying augmentation

Class	Before augmentation	After augmentation
COVID-19	421	1684
Normal	1884	1884
Pneumonia	2463	2463
Total	4768	6031

some of the augmented data samples from the training dataset.

4.2 Dataloader preparation

For each sample in the training dataset Img_1 , we randomly choose an image from the training dataset Img_2 . The class Label of both the images is counted into $Label_1$ and $Label_2$. A pair label is considered where it is set to 1 if both the images belong to the same class or 0 if it is different. A 5 member tuple is prepared for each training sample consisting of the original image Img_1 , selected pair image Img_2 , $Label_1$, $Label_2$, and pair label, and this tuple is processed in each iteration. In testing phase, only the testing images in the test dataset and their respective class labels are only considered for preparing the test data loader, and the model is evaluated on it.

4.3 Implementation details

We have used Pytorch as the deep learning framework to implement the model architecture in Python. We have adopted Stochastic Gradient Descent as our Optimizer with a fixed learning rate set to 0.001. The batch size has been set to 32 based on the computational resource available, and models were trained for 30 epochs. ReLU activation function was used

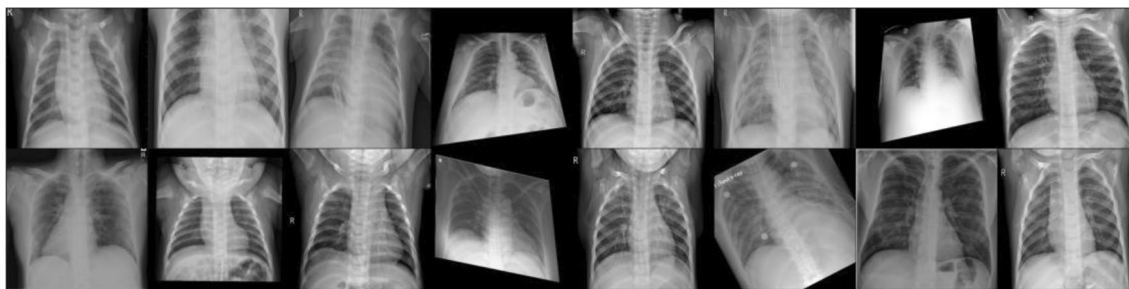


Fig. 5 The sample images from training dataset after applying data augmentation

throughout the network, and Softmax activation was used in the last classification layer of the model. The model is initialized with pre-trained weights trained over the Imagenet dataset, and weights are updated iteratively to allow the model to fine-tune over the X-ray dataset to learn effective representations. In our experiments, we have set the value of $\lambda = 0.5$ as the regularization parameter. We used a tensor board for visualizing. Training and testing loss and accuracy values. Google Colab provides free NVIDIA GPU, and we have used the Google Colab for training and testing of the models.

4.4 Experimental results

We analyze the performance of the proposed work in various test cases and compare them with the established deep learning-based automatic COVID-19 prediction models based on chest X-ray images. We use well-known metrics such as Accuracy, Precision, Recall, F1 score, and Specificity as the performance measure for this classification task. The four significant outcomes—True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN) are required to compute these metrics.

- Accuracy is the fraction of correctly classified data samples to total data samples.
- Precision is the calculation of the model's correct predictions over all predictions.
- Recall, also known as sensitivity or true positive rate, measures the proportion of actual positives that was identified and classified correctly.
- F1 score is the harmonic mean of Precision and Recall
- Specificity is the measure of True Negative rate, which measures the proportions of True Negatives classified correctly.

The Equations are given by

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{FN} + \text{TN}} \quad (5)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (6)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (7)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{FP} + \text{TN}} \quad (8)$$

The model needs to have a better metric performance in all of these metrics for a Better Model performance. The model's misclassification rate should be as low as possible by reducing the false positives and false negatives, as these mispredictions might cause wreak havoc when the model is deployed.

Table 3 shows the result metrics of the proposed system in the Multiclass classification and Binary class classification task. We can observe that the proposed systems show promising results over Binary Class predictions and Multiclass Predictions. It is to point out that VGG16 based system performs better than the AlexNet model. VGG16 shows better performance over all the six criteria over the multiclass classification, while both models have a meager difference in binary classification.

To demonstrate the positive effect of the proposed architectural solution, we have conducted ablation studies to analyze the performance gains obtained through our proposed system. We have conducted 4 sets of experiments for each of the Imagenet models- VGG16 and Alexnet. We consider VGG16 and Alexnet Model without pre-trained weights as our baseline model. Our Ablation Studies include Baseline model, Baseline model initialized with Pretrained Imagenet weights, Baseline model initialized with Pretrained Weights, and Data augmentation and Baseline model initialized with Pretrained weights trained in a Multitask paradigm with combined Classification and verification loss. Table 4 presents the Ablation Study results achieved, proving that the proposed system has significant performance gain in Prediction and Diagnostics. We can observe that the Baseline model has

Table 3 Performance results of the model in datasets

Proposed system	Accuracy (%)	Specificity (%)	Precision (%)	Recall (%)	F1 score (%)
VGG16 ^a [10]	98.79	97.8	98.8	98.79	98.8
AlexNet ^a [10]	96.64	96.5	96.72	96.64	96.64
VGG16 ^b [10]	99.93	100	99.93	99.93	99.93
AlexNet ^b [10]	99.69	99.93	99.68	99.69	99.68

^aThe model is trained for multiclass classification to predict COVID, normal and pneumonia

^bThe model is trained for binary classification to predict COVID and normal

Table 4 Ablation study of the proposed approach in various experiment scenarios

Experiment scenario	Accuracy (%)	Precision (%)	Recall (%)	F1 score (%)
VGG16 baseline ^a	71.34	61.58	62.57	61.83
AlexNet baseline ^a	69.26	58.93	61.58	60.23
VGG16 pretrained baseline ^a	91.45	84.79	86.31	85.23
AlexNet pretrained baseline ^a	89.79	82.37	85.17	83.75
VGG16 pretrained-augmentation ^a	93.94	88.47	90.45	88.39
AlexNet pretrained-augmentation ^a	91.84	86.18	89.30	87.71
VGG16 pretrained baseline ^b	95.43	91.78	93.45	92.47
AlexNet pretrained baseline ^b	93.63	90.25	92.68	91.45

^aThe model is trained with classification loss as the objective function

^bThe model is trained with both classification and verification loss as the objective function

Table 5 Class specific accuracy of the proposed system

Proposed system	COVID positive (%)	Normal (%)	Viral pneumonia (%)
VGG16 [10]	99.94	98.83	98.83
AlexNet [10]	99.7	96.43	96.49

the least performance since the model has been initialized with random weights, and when the model has been initialized with pre-trained imagenet weights, a significant performance can be observed. The class imbalance harnesses the performance by increasing false positives and negatives, and adding Data augmentation improves the precision and recall of the proposed approach. We observe significant performance gain when the Pretrained model is trained with a combined loss of Classification loss and Verification loss rather than just Classification loss. This proves when a model trained with combined loss can learn better representation for effective prediction than just Classification

Table 5 shows class specific accuracy of proposed model over the predicted classes. Results proves that the model performs a accurate classification of COVID-19 as well as the other classes.

Figure 6 shows the training and validation curves of VGG16 backed model for multiclass classification and Binary classification in dataset [10].

The Fig. 6 shows that VGG16 model is converging well over both training and validation datasets. We could observe that the training accuracies and validation accuracy are high as 99% and loss as low as 0.01 as well as the validation loss. The Highs and Lows in validation set curves initially is attributed to the limited COVID samples in Validation set and the Model was able to learn better representation and got better in each epoch. When the Model was experimented on Alex Net based model, a Similar Pattern of results were observed which proves that both the Model perform substantially as the accuracies were more than 96% in the validation set Figure. Figure 7 shows the training and validation curves of AlexNet system for multiclass classification and Binary classification in Dataset [10]. This proves us the model has learnt a better feature representation by utilizing the power of both identification and verification.

The confusion matrix gives an understandable interpretation of the proposed framework and its effectiveness through a detailed classification report. The confusion matrix of the models over the test set is shown in Fig. 8 for multi-class classification and in Fig. 9 for binary classification. In both multiclass and binary classification, the VGG16 model correctly classified 272 COVID samples out of 273 samples. In binary classification, The number of mispredictions is less than 5 out of 1673 samples of COVID and normal chest X-ray images for both VGG16 and AlexNet model. This proves the potential Impact and robustness of the model classification. In multiclass classification, the VGG16 model correctly classified 3292 samples from a total of 3332 X-ray

Fig. 6 Training and Validation curves of VGG16 based system. Red curves indicate Training and Blue curve indicate validation. **a** Represents the plot for multiclass classification and **b** represents the plot for binary classification (colour figure online)

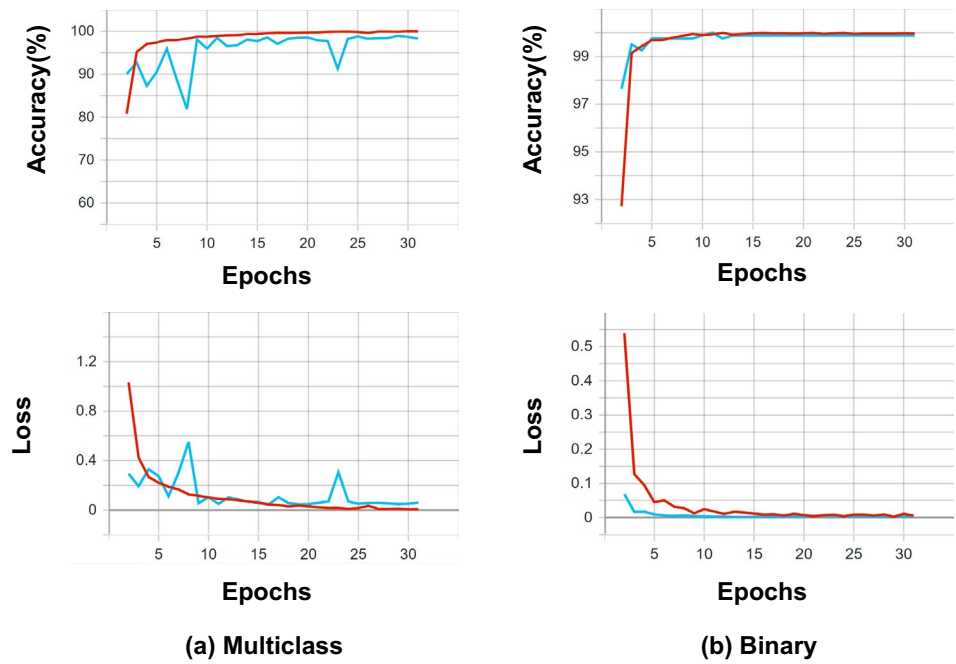
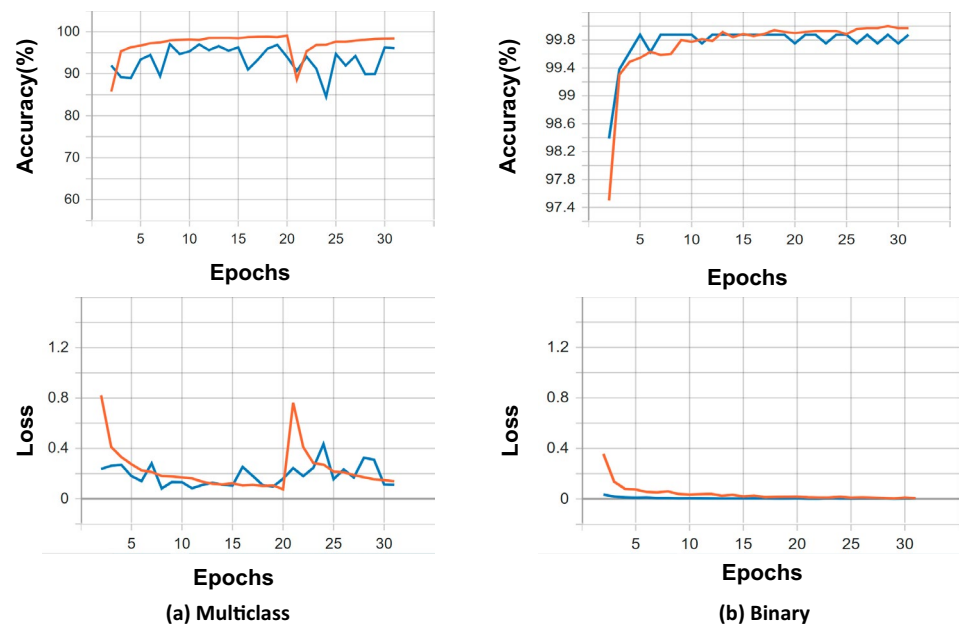


Fig. 7 Training and validation curves of AlexNet based system. Orange curves indicate Training and Blue curve indicate validation. **a** Represents the plot for multiclass classification and **b** represents the plot for Binary classification (colour figure online)



samples leaving just 40 mispredictions, while the AlexNet-based model competitively predicted 3209 samples correctly from the total. The Proposed models have correctly identified the majority of the samples with high accuracy and high precision with a minimal number of False Positives and False Negatives. It is worth noting that the models show higher accuracy and higher recall, more than 97%, for the COVID Positive, as the model has correctly classified the samples by reducing the false negatives. False Negative predictions are dangerous since they might mispredict a Positive

sample as negative and might lead to serious problems for the person and society, causing concern to public health. These models must have high recall and precision while optimizing for higher accuracy.

To prove the evaluate the proposed method, The proposed method is compared against some of the established Automatic COVID-19 diagnosis and transfer learning based methods based upon chest X-rays. Table 6 summarizes the comparisons of the performance efficiency of the proposed method with other related works.

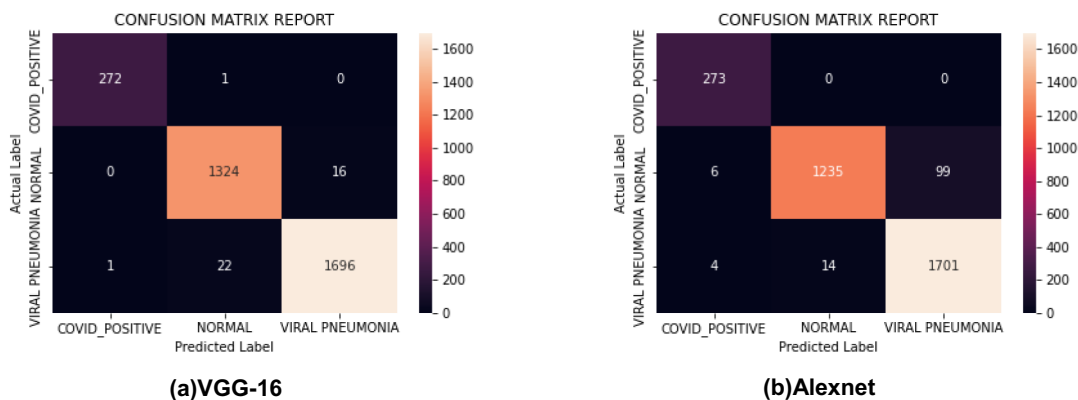


Fig. 8 Confusion Matrix of the proposed Model in Multi class classification. **a** Represents confusion matrix for VGG16 based system and **b** represents confusion matrix for AlexNet based system

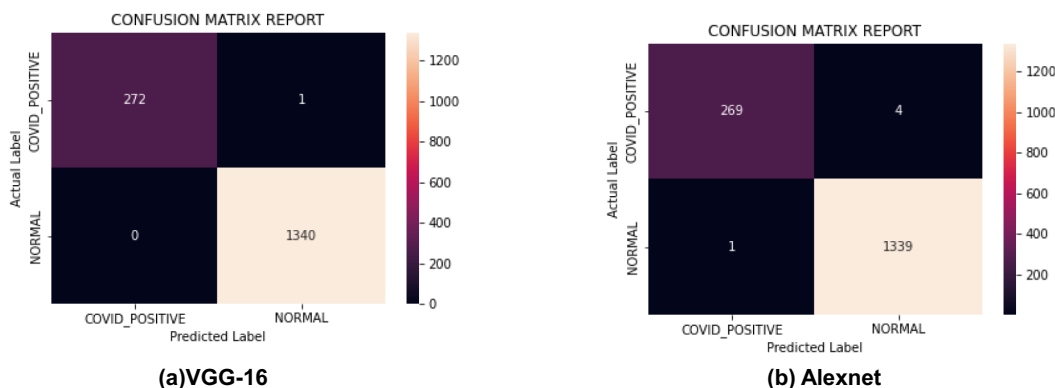


Fig. 9 Confusion matrix of the proposed model in binary classification. **a** Represents confusion matrix for VGG16 based system and **b** represents confusion matrix for AlexNet based system

Our model outperformed the existing deep learning state of art methods by a significant percentage due to practical usage of both identification and verification loss for training as the model learned a discriminative representation for the classes, and also due to the enriching the dataset by data augmentation technique. As it is evident from the above result is that our proposed model has outperformed by a greater extent over all the metrics in multi-class and binary class classification due to data augmentation and Siamese training model. The encouraging results of the automated and intelligent deep learning model for detecting COVID-19 in radiographic image detection show that computer vision and deep learning models will boost the automatic diagnosis, aiding the health workers in fighting the pandemic. The proposed work is scalable to different pre-trained models, and their performance analysis is a subject of future work.

5 Conclusion and future work

The emergence of Coronavirus Disease 2019 (COVID-19) has caused immense damage to people’s health around the world. RT-PCR is a method used to detect COVID, but many hospitals worldwide face a shortage of these kits; moreover, it is a complex and time-consuming process. In this project, we have proposed a novel convolutional neural network model that uses verification and identification losses to detect COVID-19 from chest X-ray images. The model is trained in a Siamese approach, making the model learn effective feature representation to make compelling predictions. Transfer Learning models such as VGG16, and Alex Net, which are pre-trained on Image Net data set, are used as Backbone models, and data augmentation techniques are used to solve the data imbalance since

Table 6 Comparison of the proposed work with existing automatic COVID-19 diagnosis model that are operated upon chest X-Ray

Work	Model	Task	Accuracy (%)
Ozturk et al. [1]	DarkNet	Binary	98.08
Ozturk et al. [1]	DarkNet	Multiclass	87.02
Sethy et al. [4]	ResNet50	Binary	95.38
Panwar et al. [3]	VGG16	Binary	96.72
Ioannis et al. [7]	VGG19	Multiclass	93.48
Guefrechi et al. [32]	VGG16	Binary	98.3
Wang et al. [2]	COVID-Net	Multiclass	92.4
Narin et al. [6]	InceptionV3	Binary	97
Rahman et al. [10]	DenseNet201	Multiclass	98.23
Hemdan et al. [33]	COVIDX-Net	Binary	90
Minaee et al. [34]	SqueezeNet	Binary	92.3
Misra et al. [11]	ResNet18	Multiclass	93.3
Fie et al. [17]	VBNet	Binary	78.06
Xiawi et al. [18]	VNet	Multiclass	81.06
Brunese et al. [19]	VGG16	Binary	98
Stephen et al. [20]	VGG16	Binary	93.7
Dansana et al. [21]	VGG-19	Binary	91
Oh et al. [22]	GoogLeNet	Multiclass	89
Nayak et al. [23]	ResNet34	Multiclass	92.4
Wang et al. [24]	CCSHNet	Multiclass	96.7
Proposed method	VGG16	Multiclass	98.79
Proposed method	AlexNet	Multiclass	96.64
Proposed method	VGG16	Binary	99.93
Proposed method	AlexNet	Binary	99.69

The Bold signifies the Maximum Accuracy among all other models in the literature. It validates the proposed methodology provides state of the art results

the Number of COVID Positive samples are limited. The experiments on publicly available datasets show that our model is performing well in providing accurate diagnostics for binary classification (COVID vs. normal) and multi-class classification (COVID vs. normal vs. pneumonia).

However, in terms of available data, we still have significantly fewer data to make our model robust. As the number of COVID-19 patients is increasing worldwide and the symptoms, formation, and variants of the virus are evolving daily, we will be able to extend our work, and the usability of our model as the data available is increasing.

For future works, we want to scale our convolutional neural network model to various other automated diagnosis applications like cancer predictions, where it could be used as an architectural paradigm, and generative adversarial networks are currently being used for generating realistic data sets as part of data augmentation techniques. So different GAN models could be leveraged to generate realistic data to enrich the dataset. These early diagnostic Models can be helpful for Hospitals and Doctors, which can be integrated

into a web or an application and deployed for real-time use. Using different loss functions like focal loss and circle loss for class imbalance could be explored further.

Funding The authors declare that they have received no funding from any external organization.

Declarations

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

References

- Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, Ö., Acharya, U.R.: Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* **121**, 103792–103792 (2020)
- Wang, L., Lin, Z.Q., Wong, A.: COVID-net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Sci. Rep.* **10**, 1–12 (2020)
- Panwar, H., Gupta, P.K., Siddiqui, M.K., Morales-Menéndez, R., Singh, V.: Application of deep learning for fast detection of COVID-19 in X-rays using ncovnet. *Chaos Solitons Fractals* **138**, 109944–109944 (2020)
- Sethy, P.K., Behera, S.K.: Detection of coronavirus disease (COVID-19) based on deep features. Preprints 2020, 2020030300 (2020). <https://doi.org/10.20944/preprints202003.0300.v1>
- Jain, G., Mittal, D., Thakur, D., Mittal, M.K.: A deep learning approach to detect COVID-19 coronavirus with X-ray images. *Biocybern. Biomed. Eng.* **40**, 1391–1405 (2020)
- Narin, A., Kaya, C., Pamuk, Z.: Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Anal. Appl.* **24**, 1207–1220 (2021)
- Apostolopoulos, I.D., Bessiana, T.: COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **43**, 635–640 (2020)
- Asnaoui, K.E., Chawki, Y.: Using X-ray images and deep learning for automated detection of coronavirus disease. *J. Biomol. Struct. Dyn.* **39**, 3615–3626 (2020)
- Ravi, V., Narasimhan, H., Chakraborty, C., Pham, T.D.: Deep learning-based meta-classifier approach for COVID-19 classification using ct scan and chest X-ray images. *Multimed Syst* **28**, 1401–1415 (2021)
- Rahman, S., Sarker, S., Miraj, M.A.A., Nihal, R.A., Haque, A.K.M.N., Noman, A.A.: Deep learning-driven automated detection of COVID-19 from radiography images: a comparative analysis. *Cogn Comput* 1–30 (2021). <https://doi.org/10.1007/s12559-020-09779-5>
- Misra, S., Jeon, S., Lee, S., Managuli, R., Kim, C.: Multi-channel transfer learning of chest X-ray images for screening of COVID-19. *arXiv:2005.05576* (2020)
- Imran, A.S., Posokhova, I., Qureshi, H.N., Masood, U., Riaz, S., Ali, K., John, C.N., Nabeel, M.: Ai4covid-19: Ai enabled preliminary diagnosis for COVID-19 from cough samples via an app. *Inform. Med. Unlocked* **20**, 100378–100378 (2020)

13. Zheng, C., Deng, X., Fu, Q., Zhou, Q.-F., Feng, J., Ma, H., Liu, W., Wang, X.: Deep learning-based detection for COVID-19 from chest ct using weak label. *medRxiv* (2020)
14. Anwar, T., Zakir, S.: Deep learning based diagnosis of COVID-19 using chest ct-scan images. 2020 IEEE 23rd International Multi-topic Conference (INMIC), 1–5 (2020)
15. Alsabek, M.B., Shahin, I., Hassan, A.: Studying the similarity of COVID-19 sounds based on correlation analysis of mfcc. 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), pp. 1–5 (2020)
16. Hassan, A., Shahin, I., Alsabek, M.B.: COVID-19 detection system using recurrent neural networks. 2020 International Conference on Communications, Computing, Cybersecurity, and Informatics (CCCI), pp. 1–5 (2020)
17. Shan, F., Gao, Y., Wang, J., Shi, W., Shi, N., Han, M., Xue, Z., Shen, D., Shi, Y.: Lung infection quantification of COVID-19 in ct images with deep learning. [arXiv:2003.04655](https://arxiv.org/abs/2003.04655) (2020)
18. Xu, X., Jiang, X.-G., Ma, C., Du, P., Li, X., Lv, S., Yu, L., Chen, Y., Su, J.-W., Lang, G.-J., Li, Y.-T., Zhao, H., Xu, K., Ruan, L., Wu, W.: A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering (Beijing, China)* **6**, 1122–1129 (2020)
19. Brunese, L., Mercaldo, F., Reginelli, A., Santone, A.: Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. *Comput. Methods Prog. Biomed.* **196**, 105608–105608 (2020)
20. Stephen, O., Sain, M., Maduh, U.J., Jeong, D.-U.: An efficient deep learning approach to pneumonia classification in healthcare. *J. Healthc. Eng.* **2019**, 1–7 (2019)
21. Dansana, D., Kumar, R., Bhattacharjee, A., Hemanth, D.J., Gupta, D., Khanna, A., Castillo, O.: Early diagnosis of COVID-19-affected patients based on X-ray and computed tomography images using deep learning algorithm. *Soft Comput* 1–9 (2020). <https://doi.org/10.1007/s00500-020-05275-y>
22. Oh, Y., Park, S., Ye, J.C.: Deep learning COVID-19 features on cxr using limited training data sets. *IEEE Trans. Med. Imaging* **39**, 2688–2700 (2020)
23. Nayak, S.R., Nayak, D.R., Sinha, U., Arora, V., Pachori, R.B.: Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study. *Biomed. Signal Process. Control* **64**, 102365–102365 (2020)
24. Wang, S., Nayak, D.R., Guttery, D.S., Zhang, X., Zhang, Y.-D.: COVID-19 classification by ccshnet with deep fusion using transfer learning and discriminant correlation analysis. *Int. J. Inf. Fusion* **68**, 131–148 (2021)
25. Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. *CoRR* [arXiv:1409.1556](https://arxiv.org/abs/1409.1556) (2015)
26. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. *Commun. ACM* **60**, 84–90 (2012)
27. Chung, D., Tahboub, K., Delp, E.J.: A two stream siamese convolutional neural network for person re-identification. 2017 IEEE International Conference on Computer Vision (ICCV), pp. 1992–2000 (2017)
28. Schroff, F., Kalenichenko, D., Philbin, J.: Facenet: a unified embedding for face recognition and clustering. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 815–823 (2015)
29. Lu, T., Zhou, Q.-F., Fang, W., Zhang, Y.: Discriminative metric learning for face verification using enhanced siamese neural network. *Multimed. Tools Appl.* **80**, 8563–8580 (2021)
30. Dey, S., Dutta, A., Toledo, J.I., Ghosh, S.K., Lladós, J., Pal, U.: Signet: Convolutional siamese network for writer independent offline signature verification. [arXiv:1707.02131](https://arxiv.org/abs/1707.02131) (2017)
31. Deepak, S., Ameer, P.M.: Retrieval of brain mri with tumor using contrastive loss based similarity on googlenet encodings. *Comput. Biol. Med.* **125**, 103993103993 (2020)
32. Guefrechi, S., Jabra, M.B., Ammar, A., Koubaa, A., Hamam, H.: Deep learning based detection of COVID-19 from chest X-ray images. *Multimed. Tools Appl.* **80**, 31803–31820 (2021)
33. Hemdan, E.E.-D., Shouman, M.A., Karar, M.E.: Covidx-net: a framework of deep learning classifiers to diagnose COVID-19 in X-ray images. [arXiv:2003.11055](https://arxiv.org/abs/2003.11055) (2020)
34. Minaee, S., Kafieh, R., Sonka, M., Yazdani, S., Soufi, G.J.: Deep-COVID: predicting COVID-19 from chest X-ray images using deep transfer learning. *Med. Image Anal.* **65**, 101794–101794 (2020)

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.