



A deep learning-based framework for detecting COVID-19 patients using chest X-rays

Sohaib Asif¹ · Ming Zhao¹ · Fengxiao Tang¹ · Yusen Zhu²

Received: 30 October 2021 / Accepted: 9 March 2022 / Published online: 22 March 2022
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

Abstract

Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) has caused outbreaks of new coronavirus disease (COVID-19) around the world. Rapid and accurate detection of COVID-19 coronavirus is an important step in limiting the spread of the COVID-19 epidemic. To solve this problem, radiography techniques (such as chest X-rays and computed tomography (CT)) can play an important role in the early prediction of COVID-19 patients, which will help to treat patients in a timely manner. We aimed to quickly develop a highly efficient lightweight CNN architecture for detecting COVID-19-infected patients. The purpose of this paper is to propose a robust deep learning-based system for reliably detecting COVID-19 from chest X-ray images. First, we evaluate the performance of various pre-trained deep learning models (InceptionV3, Xception, MobileNetV2, NasNet and DenseNet201) recently proposed for medical image classification. Second, a lightweight shallow convolutional neural network (CNN) architecture is proposed for classifying X-ray images of a patient with a low false-negative rate. The data set used in this work contains 2,541 chest X-rays from two different public databases, which have confirmed COVID-19 positive and healthy cases. The performance of the proposed model is compared with the performance of pre-trained deep learning models. The results show that the proposed shallow CNN provides a maximum accuracy of 99.68% and more importantly sensitivity, specificity and AUC of 99.66%, 99.70% and 99.98%. The proposed model has fewer parameters and low complexity compared to other deep learning models. The experimental results of our proposed method show that it is superior to the existing state-of-the-art methods. We believe that this model can help healthcare professionals to treat COVID-19 patients through improved and faster patient screening.

Keyword COVID-19 detection · Convolutional neural network (CNN) · Deep learning · Chest X-ray · Transfer learning

1 Introduction

The outbreak of the new coronavirus shocked the world. The novel coronavirus or COVID-19 belongs to the “coronavirus” (CoV) virus family and was previously known as “Severe Acute Respiratory Syndrome Coronavirus 2” (SARS-CoV-2). It was named COVID-19 by the World Health Organization (WHO) in February 2020. In the initial stage, the COVID outbreak was declared a public health

emergency of international concern on January 30, 2020 [1], and was subsequently declared a pandemic by the WHO [2] on March 11, 2020. COVID-19 is spreading around the world at an unprecedented rate, and all countries are working hard to control the spread of the COVID-19 pandemic. To date, the number of confirmed positive COVID-19 cases worldwide is approximately 239,007,759 and 4,871,841 deaths. The virus has spread to more than 190 countries, among which the United States, China and India are seriously infected. This is the first recorded pandemic caused by the coronavirus. COVID-19 is a novel infectious disease that has created a global health crisis that profoundly affects our lives every day and is the greatest challenge we have faced since the era of the 1918 influenza pandemic. Due to its pathogenicity, it eventually became a major global health problem and spread around the world. Governments around the world have introduced border restrictions, social distancing and hygiene awareness. Most people infected

✉ Ming Zhao
meanzhao@csu.edu.cn

✉ Fengxiao Tang
tangfengxiao@csu.edu.cn

¹ School of Computer Science and Engineering, Central South University, Changsha, China

² School of Mathematics, Hunan University, Changsha, China

with the virus will experience mild to moderate respiratory illness, recover over time, and do not need any special treatment. However, some will become seriously ill and require medical attention. Older adults and people of any age with chronic conditions such as diabetes, chronic respiratory disease, cardiovascular disease, and chronic kidney disease are at increased risk for SARS-CoV-2 and the more serious COVID-19 infection [3, 4]. The main symptoms of COVID-19 are fever, cough, loss of smell, fatigue, loss of taste and shortness of breath [5]. In some cases, the infected person may have no symptoms. For this reason, detecting an infected person is a very difficult task. Medical facilities in many developed countries are exhausted due to the rapid increase in the number of people affected by the coronavirus. They are currently facing a shortage of testing kits and ventilator. Many countries have announced total lockdowns and require their residents to stay indoors.

Due to the above-mentioned challenges and the increase in the number of cases, people have tried to find an accurate and easy way to diagnose COVID-19. Reverse Transcription Polymerase chain reaction (RT-PCR) is the most commonly used clinical screening method and is considered the gold standard for confirming COVID-19 cases [6]. It also has some key limitations. Its testing capabilities are limited, it is too sensitive, and any contamination of the sample can lead to false results. Therefore, researchers are working to improve the detection capabilities of real-time RT-PCR [7]. RT-PCR kits are expensive and take 6–9 h to confirm infection in the patient, which are long compared to the rapid spread rate of COVID-19. As a result, many infected cases cannot be detected quickly and continue to infect others unconsciously. RT-PCR has a high false negative rate and low sensitivity. According to reports, the cost of the RT-PCR kit is about US\$120–130, and a dedicated biosafety laboratory is also required to accommodate the PCR machine. The cost per set may be between US\$15,000 and US\$90,000. These expensive screening tools with delayed test results can lead to disease epidemics and worsen scenarios. To solve this problem, recently, it has been discovered that medical imaging techniques such as X-rays and computer tomography (CT) play an important role in the examination of COVID-19 cases [8, 9]. Researchers have found that there are some visual markings on the lungs of patients with COVID-19 symptoms, such as ground glass opacities. These opacities can be used to distinguish between COVID-19-infected and non-COVID-19 patients. This method has several advantages over current RT-PCR tests in terms of early detection, but it requires an expert to understand the X-ray image. Chest X-rays are preferable to chest CT scans because of the ease of use of X-ray machines in hospitals and the low exposure to ionizing radiation on patients. However, the main problem with the use of chest radiographic imaging technology is that it takes a long time for radiologists to

read and interpret chest radiographic images. This is a time-consuming and error-prone task. Therefore, there is a strong need to automate chest X-ray analysis to achieve a faster and more accurate COVID-19 diagnosis.

Now a days, many biomedical health problems (such as brain tumor detection, skin cancer detection, etc.) use artificial intelligence (AI) solutions. The latest automated techniques use AI technologies (mostly deep learning (DL) algorithms) to improve the efficiency of X-ray imaging. In recent years, DL models, especially CNN, have become one of the most common methods in AI. CNN has proven to be more effective than traditional AI methods, and has been widely used to analyze various medical images, such as MRI [10], X-ray [11], and CT scan [12]. Recently, CNN has been successfully used to detect and classify pneumonia from chest X-ray images [13, 14]. With its ability to extract information from visual elements, CNN can be used to detect COVID-19 in patients based on chest X-rays.

1.1 Major contributions and objectives

The purpose of this article is to develop a method based on automatic deep learning to detect COVID-19 in chest X-rays. The purpose of this study is to devise a CNN based model that can use chest X-rays to classify COVID-19-infected patients. The main motivation for developing our shallow CNN architecture is to classify COVID-19 patients based on chest X-ray images while reducing the time required for detection and maintaining high accuracy. The main advantage of this network is its simple and lightweight architecture with fewer parameters compared to other models. This significantly reduces computational costs and avoids the possibility of overfitting. The proposed model requires much less execution time than existing models, so this model can run faster on low-performance computers.

The main contributions of this proposed study can be outlined as follows:

1. To overcome the low sensitivity of RT-PCR, this article uses chest X-ray images to detect and diagnose COVID-19.
2. Shallow CNN architecture with optimized parameters has been proposed that can accurately classify patients into COVID-19 and normal patients based on the patient's chest X-ray. To achieve our goal, we have optimized the model for excellent performance and an easy-to-understand design.
3. The proposed model is less complex and has fewer parameters than other state of art models, which can significantly reduce computational costs while maintaining a high level of accuracy. As a result, this model can be run accurately and fast on low-performance computers.

4. The proposed Shallow CNN and five deep learning models (including InceptionV3, Xception, MobileNetV2, NasNet and DenseNet201) were trained and comprehensively compared on the exact same experimental data set of chest X-ray images.
5. The proposed model outperforms the competitive models in terms of various performance measures, such as accuracy, precision (PPV), specificity, sensitivity or recall, F1-score, area under the ROC Curve (AUC), False Positive rate (FPR), False Negative rate (FNR), Negative predictive value (NPV) and Matthews correlation coefficient (MCC).

2 Related work

Recently, deep learning techniques such as CNNs and pre-trained frameworks have been widely used to classify patients infected with COVID-19. Researchers have done a great job in a short amount of time. Subsequently, to demonstrate the importance of X-ray images and deep learning in the diagnosis of COVID-19, this section will discuss some noteworthy studies. Wang et al. [15] proposed a deep convolutional neural network (DCNN) for identifying COVID-19 cases from chest radiographs. They also introduced a new benchmark data set called COVIDx, which contains 13,975 CXR images from 13,870 patient cases. The classification accuracy obtained using the model was 93.3%. I.D. Apostolopoulos and Mpesiana [16] proposed a set of existing CNN models were evaluated to classify COVID-19 cases, and the highest test accuracy of three-classification and two-classification was obtained, which were 93.48 and 98.75%, respectively. Hemdan et al. [17] proposed a model COVIDXnet for binary classification that achieves 90.0% accuracy. Due to the lack of a public COVID-19 data set, the proposed model was validated on 50 chest X-ray images, including 25 COVID-19 cases and 25 normal cases. Panwar et al. [18] used a neural network-based method on VGG-16 and proposed nCOVnet to achieve an overall accuracy of 88.10%. Afshar et al. [19] proposed a model COVID-CAPS, using a capsule network containing four CNNs and three capsule layers. They designed a capsule network to identify COVID-19 patients from chest X-ray images. The study used two open source data sets for binary classification and achieved 95.7% accuracy and 90% sensitivity. Mangal et al. [20] proposed a deep learning model called CovidAID to detect COVID-19 cases. They conducted experiments on CXR with 155 COVID-19 cases and reported an accuracy of 90.5%. Asif et al. [21] proposed a deep transfer learning-based approach using InceptionV3 model for the detection of COVID-19. They achieved 97% training accuracy and 93% validation accuracy on chest X-ray images. Oh et al. [22] proposed a patch-based CNN method, which has relatively

few trainable parameters and can be used for COVID-19 diagnosis. Several data sets were used and achieved an accuracy of 88.9%. Ozturk et al. [23] proposed a model for detecting COVID-19 by chest X-ray radiography based on deep learning. They use the 17-layer Darknet CNN classifier with the YOLO target detection system. Their accuracy of binary classification reached 98.08%, and the accuracy of multi-class classification reached 87.02%. Sethy et al. [24] uses pre-trained ResNet50 for feature extraction and SVM for classification. The accuracy of binary classification reaches 95.38%. Waheed et al. [25] used a pre-trained VGG-16 deep learning model using synthetic data augmentation techniques to classify COVID-19 from chest X-ray images, resulting in 95% validation accuracy. Khan et al. [26] proposed a CNN-based Coronet model that was used to classify 2, 3 and 4 classes using chest X-ray images. They achieved 99, 95 and 89.6% accuracy, respectively. Ucar and Korkmaz [27] proposed a COVIDiagnosis-Net model based on SqueezeNet using Bayesian optimizer to achieve 98.3% test accuracy in the three-class classification. They performed offline data enhancement to address data imbalance and fine-tune the parameters to make the proposed model more robust. Asnaoui and Chawki [28] published a comparative study of seven different deep learning architectures for detecting COVID-19 symptoms on chest radiographs. Inception-ResNet V2 provided 92.18% classification accuracy. Shelke et al. [29] proposed a classification model for detecting COVID-19 using chest X-rays images. The accuracy obtained with DenseNet-161 is 98.9%, while ResNet-18 achieved 76% accuracy in severity classification. However, they tested a very small data set of only 22 X-ray images. Narin et al. [30] proposed the Deep CNN ResNet-50 model for binary class classification and achieved 98% accuracy. Das et al. [31] used various applied deep transfer learning methods for COVID-19 detection. They achieved a training accuracy of 99.52% and a test accuracy of 97.40%. Song et al. [32] proposed a DRE-Net model based on a modified version of the ResNet-50 model. They used the Feature Pyramid Network (FPN) to extract details from each image and classified COVID and non-COVID patients with 86% accuracy and 93% recall rate. Wang et al. [33] modified the inception transfer-learning model to establish the algorithm and achieved internal and external accuracy of 89.5 and 79.3%. Sedik et al. [34] performed an interesting analysis of the DA method that was used to perform a binary classification ("Normal" and COVID-19). However, they used a small data set and did not mention the data source used. Maghdid et al. [35] classified CXR as normal or COVID-19 pneumonia with 94.1 and 98% accuracy, respectively, using a customized CNN and a pre-trained AlexNet model. Shui-Hua et al. [36] used rank-based average pooling to develop the NRAPM module and presented an improved data augmentation technique. They proposed DRAPNet for

the COVID-19 classification and achieved a micro-average F1-score of 95.49%. Zhang et al. [37] used different data augmentation techniques and proposed a multi-input DCNN model using the convolution block attention module. They achieved a very good accuracy of 98.02% for the diagnosis of COVID-19.

Through extensive review, it was found that deep learning models can achieve significant results from chest X-ray images to classify COVID-19 diseases. However, deep neural networks are computationally intensive and consume a lot of memory. This article proposes a shallow CNN architecture for detecting positive COVID-19 cases from non-COVID-19 cases using CXR images. Five most successful pre-trained CNN models, namely, Inception-V3, Xception, MobileNetV2, NasNet and DenseNet201 have been taken into consideration based on the concept of Transfer Learning (TL). A comparative analysis was performed between the shallow CNN and the five pre-trained models, and finally the best performing model was obtained and compared with the state-of-the-art method.

3 Preliminaries

This section describes basic concepts, such as CNN and transfer learning techniques. Understanding and implementing the proposed CNN model requires a basic knowledge of these concepts.

3.1 Convolutional neural network (CNN)

CNN is a deep learning method that can be used for different tasks, such as object detection, image classification, and other computer vision related tasks. It is attracting attention in various fields including radiology. CNN uses a system that closely resembles a multi-layer perceptron designed to reduce processing requirements. CNN is made up of multiple building blocks: (1) The input layer represents the

input image into the CNN. (2) The convolution layer is the core of CNN's architecture, which is used to extract features from the input image. The convolutional layer uses a set of parameters, including padding, kernel size, filter, and stride. (3) The pooling layer reduces network parameters and overall computation by down-sampling the image. Pooling is a down-sampling operation that reduces the size of each feature map to minimize the possibility of overfitting. There are basically two types of pooling operations: max pooling and average pooling. (4) Fully connected layer (FC) and Softmax function can classify images with probability values between 0 and 1. Figure 1 shows an overview of the CNN architecture.

3.2 Transfer learning

Transfer learning is a machine learning technique for training networks on small data sets. Where models are already pre-trained on very large data sets, such as ImageNet. Transfer learning strategies have been successfully applied to solve many deep learning problems caused by insufficient labeled training data. Several studies have demonstrated the benefits of transfer learning in improving DNN performance and solving many difficult problems in computer vision. Figure 2 shows the concept of transfer learning. In image classification, transfer learning [38] is a technique in which a model trained on one task is reused for another related task, requiring minimal re-training or fine-tuning. For example, task-specific learning for COVID-19 detection on smaller data sets can be initiated using a CNN model trained on the ImageNet data set. Since deep learning requires large amounts of training data to learn specific patterns, the need for large amounts of labeled data is a big problem, especially in the field of medical imaging. The effectiveness of transfer learning can be very limited when switching data content from one type to another. Therefore, transfer learning in this case is no better than training from scratch. Transfer learning can be done in two ways to customize a pre-trained model.

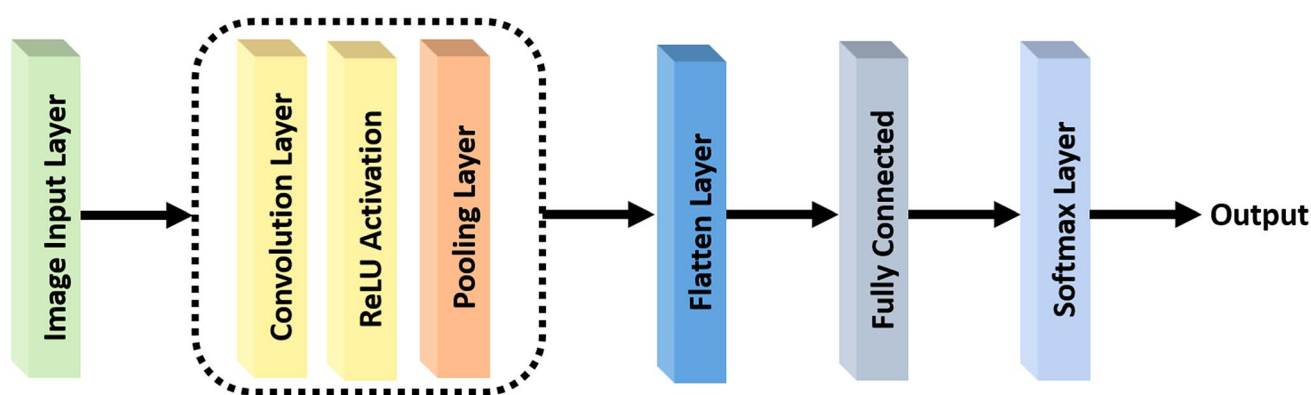
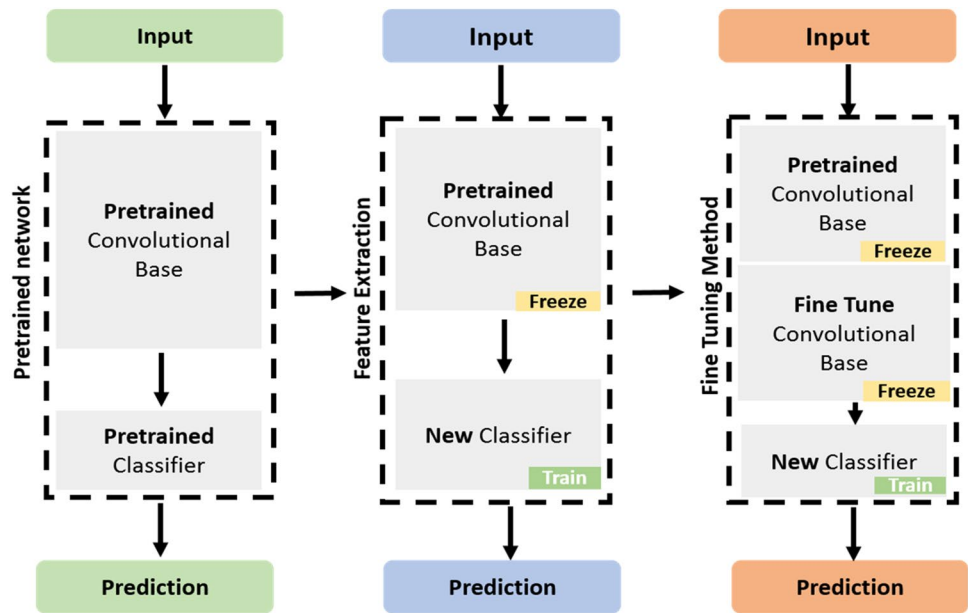


Fig. 1 Convolutional neural network (CNN) architecture

Fig. 2 Concept of fine tuning and feature extraction



Feature extraction is the process of training a network on an ImageNet data set. The fully connected layers of the model are then removed. The remaining network is then processed as a feature extractor. Fine-tuning is a technique that can not only replace the fully connected layer of the model with a new fully connected layer to retrain on the required data set, but also fine-tune all or part of the kernel.

following steps: data set collection, image preprocessing, training of transfer learning models using InceptionV3, Xception, MobileNetV2, NasNet, and DenseNet201 neural networks, and the architecture of the proposed CNN model. The steps are then explained in more detail in the following subsections.

4 Materials and methods

The main motivation for developing our proposed model is to automatically distinguish people with COVID-19. Figure 3 shows the workflow of our proposed method for diagnosing the COVID-19 virus in chest X-ray images. We present a detailed description of the proposed methodology designed to detect COVID-19. It includes the

4.1 Data set

The proposed model was trained and tested on a combined data set, which consists of chest X-ray images taken from two different sources for the diagnosis of COVID-19. The combined data set has images of three different classes, COVID-19, Viral Pneumonia and Normal. This combination makes our model more reliable and less prone to biases. The source of the data set is as follows:

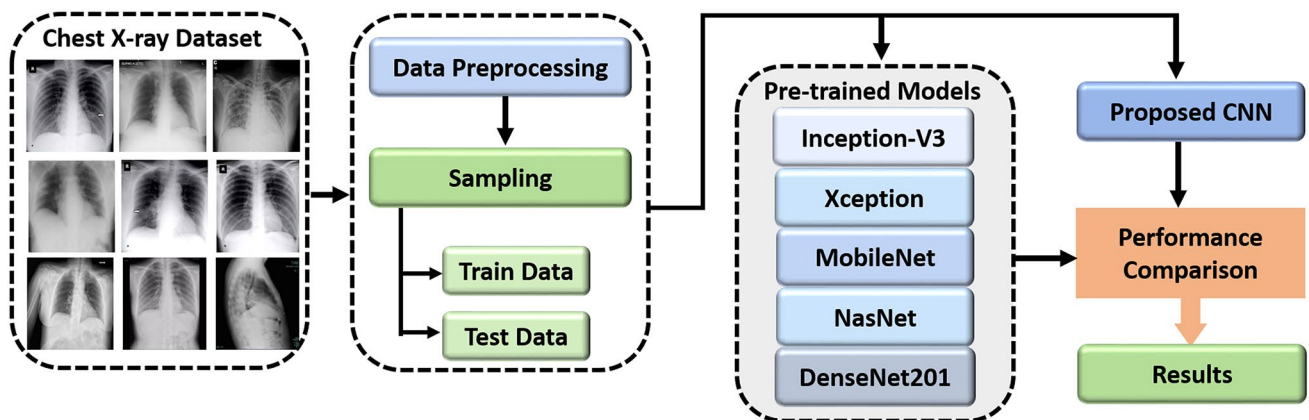


Fig. 3 Workflow of the proposed COVID-19 automated diagnostic method using chest radiographs

1. Chest X-ray images of patients infected with COVID-19 have been obtained from the Kaggle COVID-19 radiography database [39]. It was created by researchers from Dhaka University and Qatar University, as well as doctors and collaborators. This database is constantly updated, so the number of images available in these repositories may change in the future. This data set contains 1200 chest X-ray images of COVID-19 class.
2. The second source of chest X-rays was obtained from the Kaggle repository [40]. It contains 1341 normal chest X-rays and 1345 images of viral pneumonia.

Figure 4 shows some sample images from the data set. A normal chest X-ray shows a clear lung and there are no irregular "opaque" areas. Like other pneumonia, COVID-19 pneumonia increases lung density. This can be seen as lung whiteness on radiographs. The most common changes in the lungs include nodular shadowing, consolidation and ground glass opacities (GGO). They mainly affect the periphery and lower part of the lungs. Peripheral distribution of GGO, where lower lobe predilection is the most common finding on chest X-rays. After getting a balanced data set. The data set was shuffled and split into a training set and a test set. Of these, 80% images are used for training purposes and 20%

images are used for testing purposes. Table 1 summarizes the partitions of this data set.

4.2 Data preprocessing and augmentation

The next step is to preprocess the images, because the chest X-ray images in the data set are of different sizes. One of the important steps in data preprocessing is to adjust the size of the X-ray image, because different CNNs have different input requirements. For example, the Xception and InceptionV3 architecture expects an image size of $299 \times 299 \times 3$, and the DenseNet201 architecture expects an input size of $224 \times 224 \times 3$. To accommodate the common dimensions of all the architectures used in this study, we first resized all chest X-ray images to a $(224 \times 224 \times 3)$ shape. All images are normalized according to the pre-trained model standard.

Table 1 Details of the images used in the training and test set

Types	Total images	Training	Testing
COVID-19	1200	900	300
Normal	1341	1005	336
Viral pneumonia	1345	1008	337

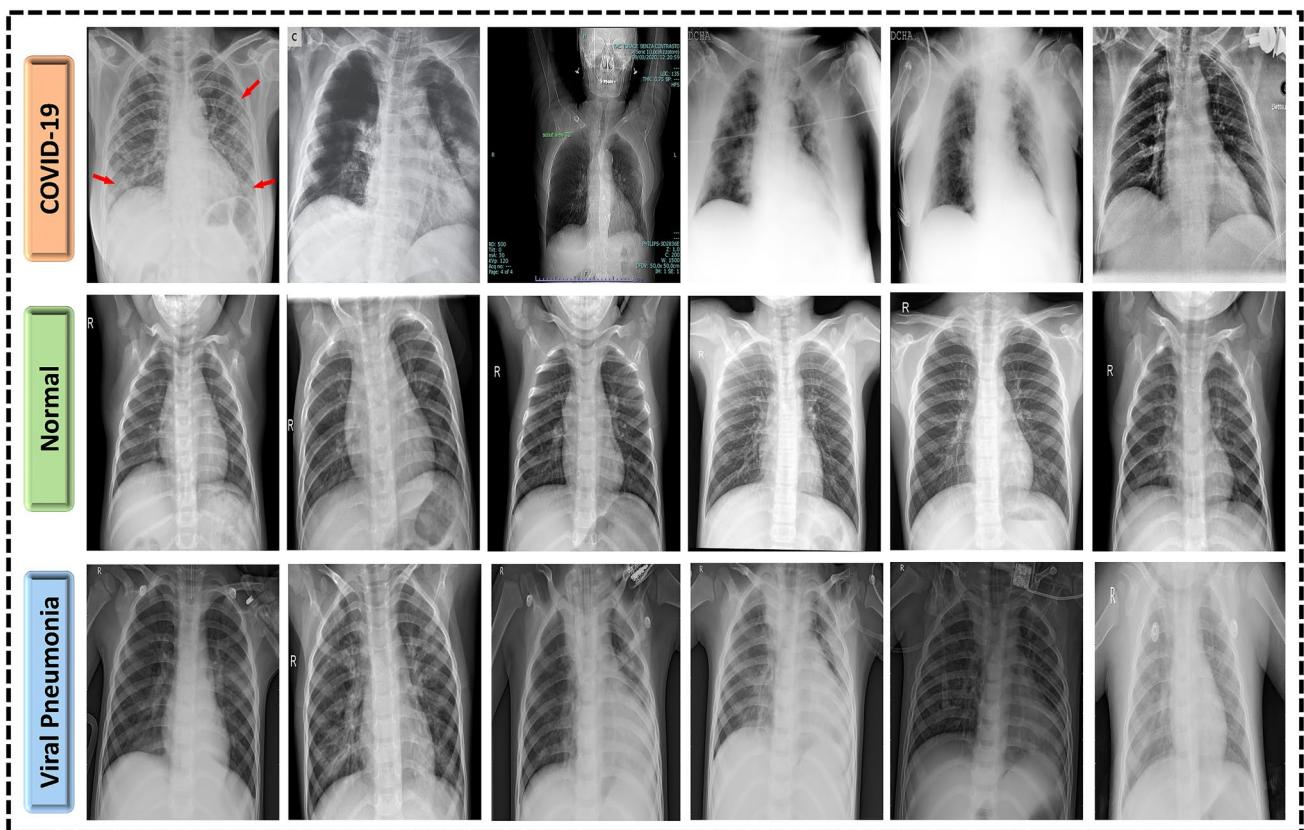


Fig. 4 COVID-19, viral pneumonia and normal samples of chest X-ray images from the prepared data set

Data normalization was performed for better training and was used to normalize the pixel values of the input images in the range from 0 to 1. One of the reasons for doing this is to help with the problem of gradient propagation. Using normalization speeds up the training of the CNN model and is more likely to stabilize the gradient propagation problem. The data set used for this task contains a grayscale image that has been rescaled by multiplying 1/255 by a pixel value. Image augmentation techniques are applied to images that increase the size of the data set to alleviate the problem of model overfitting. In this study, we used three augmentation strategies to create new training sets (horizontal flip, rotation, and translation).

4.3 COVID-19 prediction using deep transfer learning models

The CNN model has proven to give excellent results in a wide range of medical imaging applications. However, due to the limited availability of X-ray samples, it is difficult to train these models from scratch to predict COVID-19 cases. Applying a pre-trained model using the concept of transfer learning can be useful in such situations. In this work, the transfer learning strategy has been used for a number of reasons, including: (1) To avoid the problem of overfitting due to insufficient chest X-ray images available,

(2) To improve the prediction accuracy, and (3) Reduces computational complexity during the training process. In this study, five well-known pre-trained deep learning CNN InceptionV3 [41] Xception [42] MobileNetV2 [43] NasNet [44] and DenseNet201 [45] were trained and tested to classify COVID-19. These models were originally trained on the ImageNet data set and then retrained on the prepared data set. Fine tuning was done by removing the last layer in these models, flatten layer is added followed by dropout rate of 0.5 and finally a new fully connected layer is added with softmax activation function with an output size of two that represents two different classes (COVID-19 and normal). Data shuffling was enabled during model training to shuffle the data before each epoch. Five pre-trained CNN models are trained for 20 epochs using Adam optimizer with a learning rate of 0.001, and batch size set to 64. In all cases, the cross-entropy was used as the loss function. Hyperparameters play a key role in training these DL models and are constant to obtain fair comparisons. Figure 5 shows the architectures of InceptionV3, Xception, MobileNetV2, NasNet and DenseNet201.

4.3.1 InceptionV3

InceptionV3 [41] is a CNN-based network for classification. The Inception module's network has additional layers

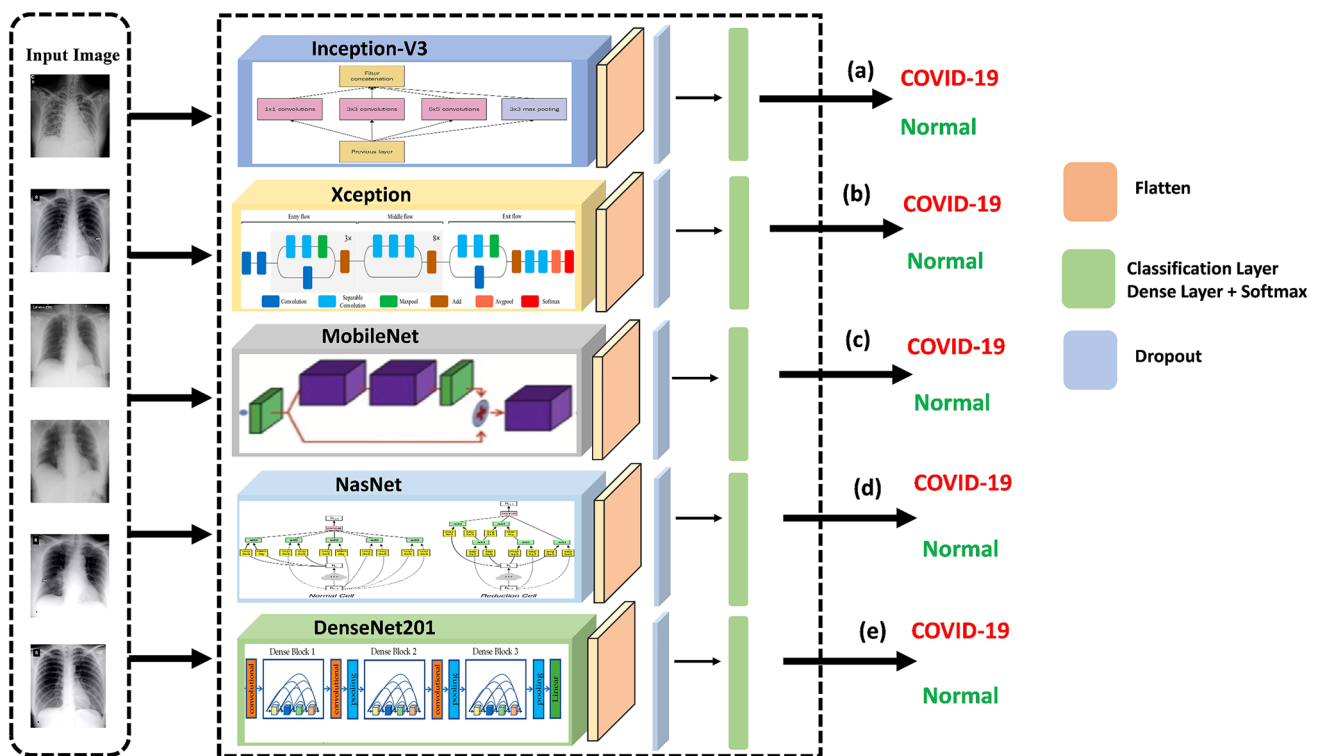


Fig. 5 Architectures of deep transfer learning models. a InceptionV3, b Xception, c MobileNetV2, d NasNet, e DenseNet201

consisting of 1×1 , 3×3 , and 5×5 convolutions. This allows the model to use all types of kernels in the image and get results from all of them. The InceptionV3 network has multiple building blocks, where each block has multiple branches of convolution, pooling, concatenation, loss, and fully connected layers. The architecture used is shown in Fig. 5a.

4.3.2 Xception

Xception [42] CNN was developed by Google Inc. and is an "advanced" version of the Inception model. The Xception model is 36 layers deep, except for the final fully connected layer. It replaces regular inception modules with distinguishable depth convolutions. In the classification of the ImageNet data set, Xception surpassed InceptionV3. Its parameter size is similar to the Inception Net, but it performs better on the ImageNet data set than the Inception Net. The architecture we use is shown in Fig. 5b.

4.3.3 MobileNetV2

MobileNetV2 [43] builds on the ideas of MobileNet-V1, which uses depth-separable convolution as effective building blocks. MobileNetV2 is an architecture designed to run on mobile and embedded systems or devices that lack compute power. It is based on an inverted residual structure with linear bottleneck. There are two types of blocks in the network. One is a residual block with a stride of 1, and the other is a block with a stride of 2, which is used to reduce the input from the previous layer. The architecture we use is shown in Fig. 5c.

4.3.4 NasNet

NasNet [44] refers to the neural architecture search network, which is a machine learning model. This idea was realized through the NAS concept from the Google ML group. Several changes have been made based on the number of layers, weights, regularization methods, etc. to improve network efficiency. NasNet can efficiently work with a small data set (CIFAR-10) as well as a large data set (ImageNet). The architecture we use is shown in Fig. 5d.

4.3.5 DenseNet201

DenseNet [45] is an architecture that uses dense connections between layers through Dense Blocks, so it is called Densely Connected Convolutional Network. DenseNet has four different variants: DenseNet169, DenseNet121, DenseNet201, and DenseNet264. In this article, we used DenseNet201. Its depth is 201 layers. For each layer in the network, the feature maps of all previous layers are

used as input, and the output feature maps are used as the input of each subsequent layer. The architecture we use is shown in Fig. 5e.

4.4 Architecture of proposed shallow convolutional neural network

Methods based on deep learning have shown better performance than traditional machine learning methods. Recently, many researchers have begun to use deep learning models to diagnose suspected COVID-19 infections. These models show significantly better results. However, the main drawback is that these models still suffer from overfitting issues, large number of parameters, usually require considerable training time, and are costly in real-world applications. The main motivation for developing our shallow CNN architecture is to classify COVID-19 patients based on chest X-ray images while reducing the time required for detection and maintaining high accuracy. The main advantage of this network is its simple and lightweight architecture with fewer parameters compared to other models. This significantly reduces computational costs and avoids the possibility of overfitting. As a result, this model can run fast on low-performance computer.

In this section, we propose a shallow CNN structure consisting of six layers for COVID-19 classification. Figure 6 shows the architecture of the proposed shallow CNN. Our proposed model is based on five components, namely, convolutional layer, pooling layer, dense layer, flatten layer and activation function. The details of each component of our proposed network and their respective parameters are described in next section.

4.4.1 Input layer

The input layer basically depends on the size of the image. The size of the chest X-ray image has different sizes in the data set. In our network, all images must have the same size and fixed at 224×224 .

4.4.2 Convolution layers

The convolutional layer is the main component used in convolutional neural networks. This layer performs most of the calculations of the CNN model. It has some parameters and hyperparameters, namely, number of filters, kernel size and K. The network consists of two convolutional layers connected to the first input layer. The first convolution layer used a 3×3 kernel size with 32 filters, and the second convolution layer used a 3×3 kernel size with 64 filters.

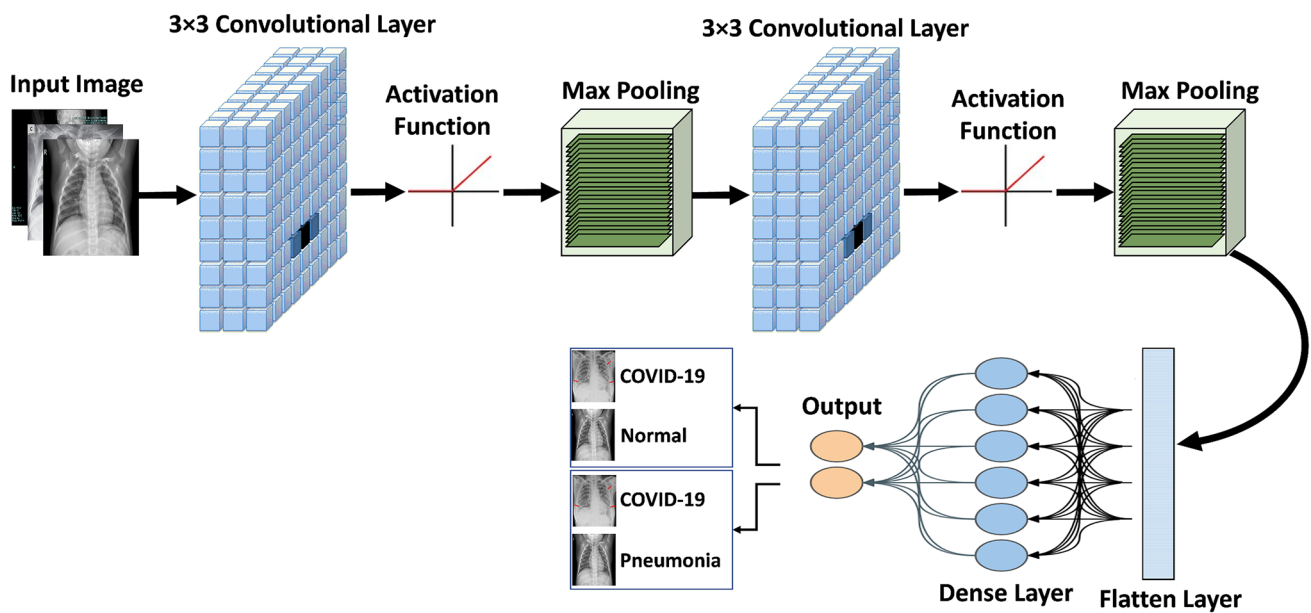


Fig. 6 Proposed CNN architecture for the classification of COVID-19 patients

4.4.3 Pooling layer

A pooling layer is a down sampling operation, usually applied after a convolutional layer, that provides some spatial invariance. This layer can speed up the calculation and prevent overfitting. There are different types of pooling layers, including maximum pooling, sum pooling, and average pooling. A maximum pooling layer is used in this work. It uses the maximum value of the window as output and reduces the size of the feature map. Note that we have used one max pooling layer for each convolution layer. Therefore, in our proposed model, we use two max pooling layers, and the window pooling size is 2×2 . In addition, the stride of the max pooling operation is set to 2.

4.4.4 Flatten layer

Flattening transforms the input into a single column connecting the convolutional layer and the dense layer. In our model, we used a flattened layer after the pooling layer to flat the entire network.

4.4.5 Fully connected layer

After the convolution and pooling process, we use a dense layer in the CNN for final classification. This layer flattens inputs from past layers from matrices to vectors. After flattening, the volume of the previous layer is input to a fully connected layer, such as a neural network. This layer determines the features that closely match a particular class by looking at the output of the previous layer. The dense layer

is also called a fully connected layer. The output of a fully connected layer is provided by an activation function that predicts the class score. CNNs often use dense layers to map spatial features to image labels.

4.4.6 Activation functions

The activation function adds non-linearity to the neural network and helps the neural network to learn powerful operations. The functions used are as follows:

1. We have used Rectified Linear Unit (ReLU) function with convolution layers. It provides decent calculations and avoids the vanishing gradient problem. Mathematically given as: $f(x) = \max(0, x)$.
2. We use the softmax activation function with the dense layer to predict the class of the input X-ray image. Softmax activation function can be computed as

$$\text{softmax}(z)_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K. \quad (1)$$

During the training process, we used the Adam Optimizer with a learning rate of 0.001 to train the proposed model on the prepared data set. Cross entropy loss function is used to minimize the loss. Data shuffling is used to shuffle the data during training before each epoch. Batch size is one of the most important hyperparameters for training a model. During training, different batch sizes were applied, such as 16, 20, 25, 32, 50, and 64. We aim to find out which batch size is the best possible result for the model. For example, in the

beginning we started with a batch size of 16, which gives an accuracy of 97.95%.

5 Results and discussion

In this section, various experiments were performed on the data set to test the effectiveness of the proposed deep learning models. We conducted a comprehensive experimental analysis on X-ray images to predict COVID-19. All models were prepared and trained individually as described above. The proposed work is implemented using the Keras package with Python and TensorFlow. The experiment was conducted on a Google Colaboratory with a Tesla P100–PCIe 16 GB graphics card, an Intel Core i5 processor with a 2.4 GHz CPU, 16 GB of RAM, and a 128 GB hard disk.

5.1 Performance evaluation metrics

Performance evaluation metrics are very important for evaluating the performance of deep learning classification models. After the training phase was completed, the performance of each model on the test data set was evaluated and compared based on ten performance metrics: accuracy, precision (PPV), specificity, sensitivity or recall, F1-score, area under the ROC Curve (AUC), False Positive rate (FPR), False Negative rate (FNR), Negative predictive value (NPV) and Matthews correlation coefficient (MCC):

$$\text{Accuracy} = \frac{\text{TN} + \text{TP}}{\text{TN} + \text{TP} + \text{FN} + \text{FP}}, \quad (2)$$

$$\text{Precision(PPV)} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \quad (3)$$

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}, \quad (4)$$

$$\text{Sensitivity(Recall)} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad (5)$$

$$F1 - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}, \quad (6)$$

$$\text{NegativePredictiveValue(NPV)} = \frac{\text{TN}}{\text{TN} + \text{FN}}, \quad (7)$$

$$\text{MCC} = \frac{(\text{TP} \times \text{TN} - \text{FP} \times \text{FN})}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}}. \quad (8)$$

These metrics are calculated from different parameters of the confusion matrix, including "True Positive" (TP), "True

Negative" (TN), "False Positive" (FP) and "False Negative" (FN). TP represent the number of COVID-19 images identified as COVID-19, TN represent the number of normal images identified as normal, FP represent the number of normal images that are incorrectly identified as COVID-19 images, FN represent the number of COVID-19 images that are incorrectly identified as normal.

5.2 Hyperparameter optimization

The main goal of this task is to design a robust model with fewer parameters to classify COVID-19 patients while reducing the time required for detection and maintaining high accuracy. This can only be achieved by finding the best hyperparameter configuration so that the best accuracy can be obtained. Table 2 presents the hyperparameters used to train the proposed CNN model and the pretrained model.

5.3 Classification performance of deep transfer learning models

Since the release of the COVID-19 data set, the research on chest X-ray images used to accurately predict COVID-19 infection has attracted much attention. Different techniques have been used to develop robust diagnostic models using deep learning methods. The concept of transfer learning has been widely used in CNN-based models. However, most of the earlier methods were estimated using unbalanced data. In addition, in some cases, data are limited for evaluating methods. In this study, we comprehensively evaluated the effectiveness of the five most effective CNN models. A large number of experiments were performed on the chest X-ray data set to determine the best model for predicting COVID-19 infection in X-ray images. Independently, five versions of the pre-trained CNN model: InceptionV3, Xception, NasNet, MobileNetV2, and DenseNet201 on the ImageNet data set were fine-tuned on the chest X-ray data set. Table 3 compares

Table 2 Model's parameter configuration

Metrics	Proposed CNN	Pretrained models
Input shape	224×224	224×224
Shuffling	Each epoch	Each epoch
Optimizer	Adam	Adam
Batch size	16, 20, 25, 32, 50, 64	64
Learning rate	0.001	0.001
Epoch	20	20
Re-scaling	1/255	1/255
Flip	Horizontal Vertical	Horizontal Vertical
Activation	Softmax (Final Classification Layer)	Softmax (Final Classification Layer)

Table 3 Comparison of classification results of five different deep transfer learning models on the test data set

Models	ACC	Prec	Spec	Sens	F1-score	NPV	MCC	AUC
InceptionV3	96.22	98.24	98.51	93.66	96.53	94.57	92.50	97.17
Xception	98.11	99.65	99.70	96.33	97.96	96.82	96.25	99.06
NasNet	96.85	97.29	97.61	96.00	96.77	96.47	93.69	99.12
MobileNetV2	99.37	99.33	99.40	99.33	99.32	99.40	98.73	99.82
DenseNet201	99.21	99.66	99.70	98.66	99.15	98.82	98.42	99.49

Acc. Accuracy, Prec. Precision, Spec. Specificity, FPR False Positive Rate, FNR False Negative Rate

the detailed classification results of all networks in terms of various metrics. The same amount of data was selected for both groups to handle the problem of data imbalance. From Table 3, it is clear that all pre-trained models performed very well in the classification of COVID-19 and normal images. MobileNetV2 delivers outstanding test accuracy of 99.37% while achieving 99.33, 99.40, 99.33 and 99.32% in precision, specificity, recall and F1-score, respectively. The accuracy of DenseNet201 on the test set reaches 99.21%. It also achieved precision, specificity, recall and F1-score of 99.66, 99.77, 98.66 and 99.15%. The Xception Network was the third best performer for COVID-19 prediction, with 99.65% precision, 96.33% sensitivity, 99.70% specificity, 97.96% F1 score, 98.11% accuracy, and 99.06 AUC. %. The lowest performance was achieved with the InceptionV3 model with 96.22% test accuracy, 93.66% recall, 98.24% precision, and 96.53% F1-score. DenseNet201 and MobileNetV2 outperform other pre-trained models in almost all performance metric, including accuracy, precision, specificity, recall, f1-score, NPV and AUC.

Table 4 shows the confusion matrix for the studied pre-trained models on a test set containing COVID-19 and normal samples. True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN) values are calculated for COVID-19 cases. We can conclude that our MobileNetV2 model only made 4 misclassifications in the test data set. These misclassifications have occurred in the COVID-19 class and the normal class. We see that out of

636 tests, our model is correct in 632 tests. By testing the DenseNet201 model on the test set, 4 images infected with COVID-19 were incorrectly identified as normal images, and 1 normal image was incorrectly classified as a COVID-19 positive case.

5.4 Classification performance of proposed shallow CNN

This study constructs binary classification experiments to evaluate the classification performance of the proposed model and demonstrate its effectiveness. We conducted two different experiments on two different data sets. These experiments can be divided into (1) Binary image classification of COVID-19 positive vs normal patients (2) Binary image classification of COVID-19 positive patients vs viral pneumonia patients.

5.4.1 Experiment 1: COVID-19 vs normal

To analyze the performance of our proposed shallow CNN model, we calculated the accuracy, precision, specificity, sensitivity, F1 score, NPV, AUC and confusion matrix for different batch sizes. Batch size is one of the most important hyperparameters for model training. This experiment investigates the effect of various batch sizes on CNN. During the training period, various batch sizes such as 16, 20, 25, 32, 50, 64 were applied. Table 5 shows the performance of our

Table 4 Confusion matrices for the deep transfer learning models on the testing data set

Models	TP	FP	TN	FN	Labels (actual values)	Predicted values	
						COVID-19	Normal
InceptionV3	281	5	331	19	COVID-19	281	19
					Normal	5	331
Xception	289	1	335	11	COVID-19	289	11
					Normal	1	335
NasNet	288	8	328	12	COVID-19	288	12
					Normal	8	328
MobileNetV2	298	2	334	2	COVID-19	298	2
					Normal	2	334
DenseNet201	296	1	335	4	COVID-19	296	4
					Normal	1	335

Table 5 Comparative analysis of CNN model with different batch sizes

Batch size	Acc	Prec	Spec	Sens	F1-score	FPR	FNR	NPV	MCC	AUC
16	97.95	97.67	97.91	98.00	97.83	2.127	2.040	98.20	95.90	99.85
20	99.21	99.00	99.10	99.33	99.16	0.900	0.671	99.40	98.42	99.80
25	98.27	97.68	97.91	98.66	98.16	2.127	1.351	98.79	96.53	99.92
32	98.42	98.33	98.51	98.33	98.33	1.510	1.694	98.51	96.84	99.76
50	98.27	98.32	98.51	98.00	98.15	1.510	2.040	98.21	96.52	99.72
64	99.68	99.66	99.70	99.66	99.65	0.298	0.334	99.70	99.36	99.97

Acc. Accuracy, Prec. Precision, Spec. Specificity, FPR False Positive Rate, FNR False Negative Rate

proposed CNN model for different batch sizes in terms of accuracy, precision, specificity, sensitivity, F1-score, NPV, MCC and AUC. It can be seen from Table 5 that a higher accuracy of 99.68% is achieved with a batch size of 32, while 99.66, 99.70, 99.66 and 99.65% are achieved in terms of precision, specificity, sensitivity and F1-score, respectively. The proposed CNN model achieved 99.21% accuracy from a batch size of 20, while the accuracy, sensitivity, and F1-scores were 99, 99.33, and 99.16%, respectively. The value of the false positive rate is also one of the key indicators, it means that the patient is mistakenly considered positive. In addition, the rapid increase in COVID-19 pandemics makes patient management difficult in most countries due to the large number of patients. The proposed CNN model achieved low false positive and false negative rates of 0.298 and 0.334%. High FPR increases problems in the healthcare management system. Which is sometimes not suitable for real positive patients.

Table 6 shows the confusion matrix of the proposed CNN using a data set containing COVID-19 and normal samples. The confusion matrix clearly visualizes that shallow CNNs perform relatively well over other algorithms. First, we can see that there are very few false negative (FN) counts with batch sizes of 20 and 64 (i.e., 0 and 2), which helps to increase sensitivity. FN indicates that the model classifies

COVID-19 patients as normal, whereas the patients are infected. Second, the model with batch sizes of 20 and 64 also showed very few false positives (FP) cases (i.e., 0–3), which were mistaken for COVID-19 infection, resulting in higher specificity and precision. From Table 6, it can be seen that the proposed CNN with a batch size of 64 incorrectly classified only one COVID-19 image out of 300 images as normal, and only one image out of 336 images was incorrectly classified as COVID-19. However, for a batch size of 20, 333 non-COVID-19 images were correctly identified as normal and 298 COVID-19 images were correctly identified as COVID-19. This shows that the model is very powerful in distinguishing COVID-19 images from X-ray images.

5.4.2 Experiment 2: COVID-19 vs viral pneumonia

The proposed model mainly focuses on binary classification of chest X-ray images for rapid and accurate detection of COVID-19. In this experiment, the proposed model was applied to classify COVID-19-infected patients from viral pneumonia. The goal of this experiment is to study the relationship between COVID-19 and pneumonia. We used 1341 chest X-ray images of pneumonia-infected patients, whereas for COVID-19 positive patients we considered the same data set used in Experiment 1. The confusion matrix used

Table 6 Confusion matrices obtained for different batch sizes

Batch size	TP	FP	TN	FN	Labels (actual values)	Predicted values	
						COVID-19	Normal
16	294	7	329	6	COVID-19	294	6
					Normal	7	329
20	298	3	333	2	COVID-19	298	2
					Normal	3	333
25	296	7	329	4	COVID-19	296	4
					Normal	7	329
32	295	5	331	5	COVID-19	295	5
					Normal	5	331
50	294	5	331	6	COVID-19	294	6
					Normal	5	331
64	299	1	335	1	COVID-19	299	1
					Normal	1	335

Table 7 Confusion matrix for COVID-19 vs viral pneumonia

Actual class	Predicted class	
	COVID-19	Pneumonia
COVID-19	TP = 289	FN = 11
Pneumonia	FP = 10	TN = 327

Table 8 Performance metrics of the proposed model on the COVID-19 vs viral pneumonia test set

Metrics	Proposed CNN
Accuracy	96.70
Precision	96.66
Specificity	97.03
Sensitivity	96.33
F1-score	96.49
FPR	0.029
FNR	0.036
NPV	96.75
MCC	93.38

to evaluate the performance of the proposed model on the test set is provided in Table 7. Table 7 shows that only eleven images of COVID-19 out of 300 images were misclassi-

$$\text{Imbalance Ratio (COVID - 19 vs Normal)} = \frac{1341(\text{Majority class})}{640(\text{Minority class})} = 2.09 \tag{9}$$

$$\text{Imbalance Ratio (COVID - 19 vs Pneumonia)} = \frac{1345(\text{Majority class})}{654(\text{Minority class})} = 2.06 \tag{10}$$

fied as pneumonia, and only ten images out of 337 images were misclassified as COVID-19. Based on this confusion matrix, the results were further analyzed against various performance evaluation measures, such as accuracy precision, specificity, sensitivity, F1-score, and MCC, as detailed in Table 8. From the results obtained, we can conclude that the proposed model can identify patients with COVID-19 with

an accuracy of 96.70%. In addition, accuracy, sensitivity and F-1 reached 96.66%, 96.33 and 96.49%, respectively.

5.4.3 Evaluation on imbalanced data set

When we estimate the incidence of COVID-19, it becomes clear that the data we encounter in real life will be imbalanced; as a result, models for the minority class tend to have low predictive accuracy. Medical data sets are often imbalanced due to lack of data. In this section, we performed another set of experimental tests to further demonstrate the effectiveness of our proposed model on the data imbalance problem, classifying COVID-19 vs normal cases and COVID-19 vs viral pneumonia cases. We reduced the sample for COVID-19 to get better results with a small number of COVID-19 images. In this experiment, 640 COVID-19 and 1341 normal X-ray images were considered for COVID-19 vs normal cases. 654 COVID-19 and 1345 pneumonia X-ray images were considered for COVID-19 vs pneumonia cases. We can see that the number of images in each class is different. Using the equation below, we calculated the imbalance ratio of our data set. If the imbalance rate is greater than 1.5, the data set can be considered imbalanced [46]:

The overall results are shown in Tables 9 and 10, where Table 9 shows the classwise results along with the confusion matrix, and the results are further analyzed using different performance evaluation metrics detailed in Table 10. It is noted that when the sample size of the target class (COVID-19) decreases, the results do not deviate much. We achieved 97.74% accuracy for COVID-19 vs normal

Table 9 Classwise results and confusion matrix for imbalanced distribution of COVID-19 data

Class	Precision	Sensitivity	F1-score	Labels (actual values)	Predicted values	
					COVID-19	Normal
COVID-19	0.96	0.97	0.96	COVID-19	145	5
Normal	0.99	0.98	0.98	Normal	6	330
COVID-19	0.94	0.91	0.92	COVID-19	140	14
Pneumonia	0.96	0.97	0.97	Pneumonia	9	328

Table 10 Performance evaluation metrics for proposed CNN model on imbalanced data set

Metrics	COVID-19 vs Normal	COVID-19 vs Pneumonia
Accuracy	97.74	95.32
Precision	96.03	93.96
Specificity	98.21	97.33
Sensitivity	96.67	90.91
F1-score	96.35	92.41
FPR	0.017	0.026
FNR	0.033	0.090
NPV	98.51	95.91
MCC	94.71	89.05

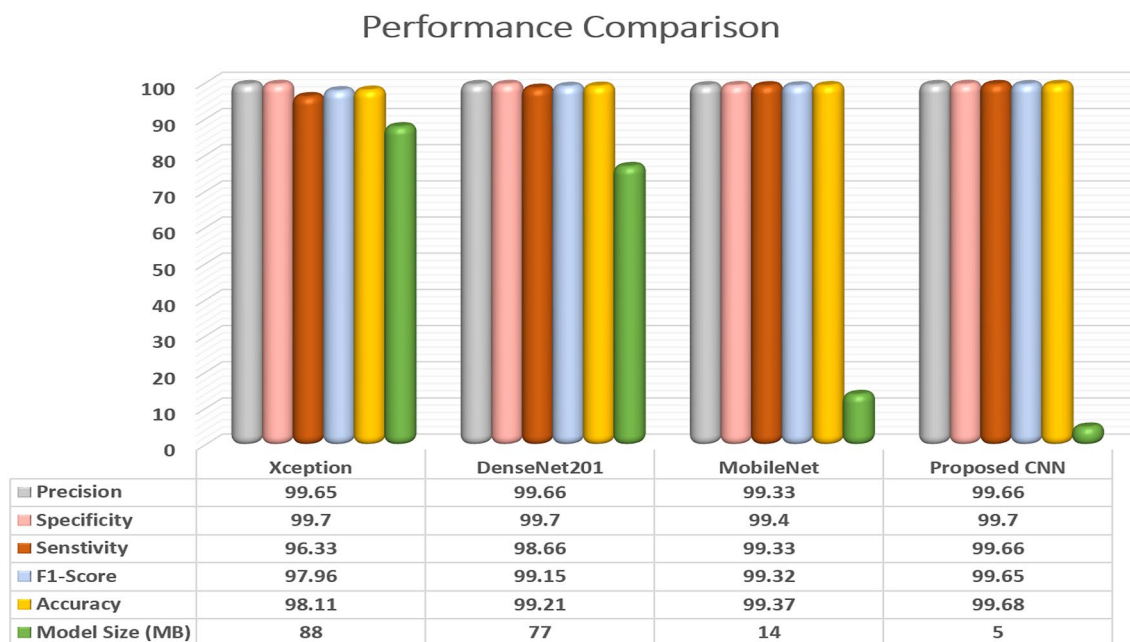
cases and 95.32% for COVID-19 vs pneumonia on such an imbalanced data set. It is clear that the proposed model is able to counter the class imbalance, as there are fewer cases of COVID-19 than cases of normal and pneumonia.

5.5 Performance comparison of the proposed CNN with pretrained models

The three best classification models based on deep transfer learning previously evaluated in Sect. 5.2 were selected and compared with the performance of the proposed shallow CNN. Figure 7 presents the overall performance analysis of the proposed CNN and the other pretrained models along with the model sizes. MobileNetV2 performed best, followed

by DenseNet201 and Xception. MobileNetV2 achieved an overall accuracy of 99.37%. The proposed shallow CNN produces maximum accuracy (99.68%) due to its shallow configuration. The DenseNet201 and Xception architectures are capable of providing the same precision and specificity as the proposed model. The proposed model competes with other pre-trained models in terms of sensitivity, specificity, precision and F1-score. The proposed model consists of 420,712 parameters and has less storage space than other models. MobileNetV2 consists of 3,329,218 parameters and takes up about 14 mb of disk space. Xception consists of 21,062,186 parameters and occupies 88 mb of storage space. DenseNet201 requires 18,510,146 parameters and occupies approximately 77 Mb of storage space. Proposed model parameters are 98% less than Xception, 97.72% less than DenseNet201, 87.36% less than MobileNetV2. The proposed shallow architecture continues to perform well compared to other pre-trained deep CNN architectures due to its light-weight computation. The confusion matrix (Fig. 8) clearly shows that the performance of the proposed CNN is better than other models.

In terms of validation accuracy and validation loss at each epoch, the performance of the proposed model is compared with the performance of other deep CNN architectures, as shown in Fig. 9. It is clear that the proposed algorithm meets the maximum accuracy and minimum loss within 20 epoch counts. The excellent results achieved by our model, along with its simplicity and low computational cost, show that the model is capable of detecting COVID-19 very efficiently from chest X-ray images.

**Fig. 7** Performance comparison of the proposed CNN with other deep learning models

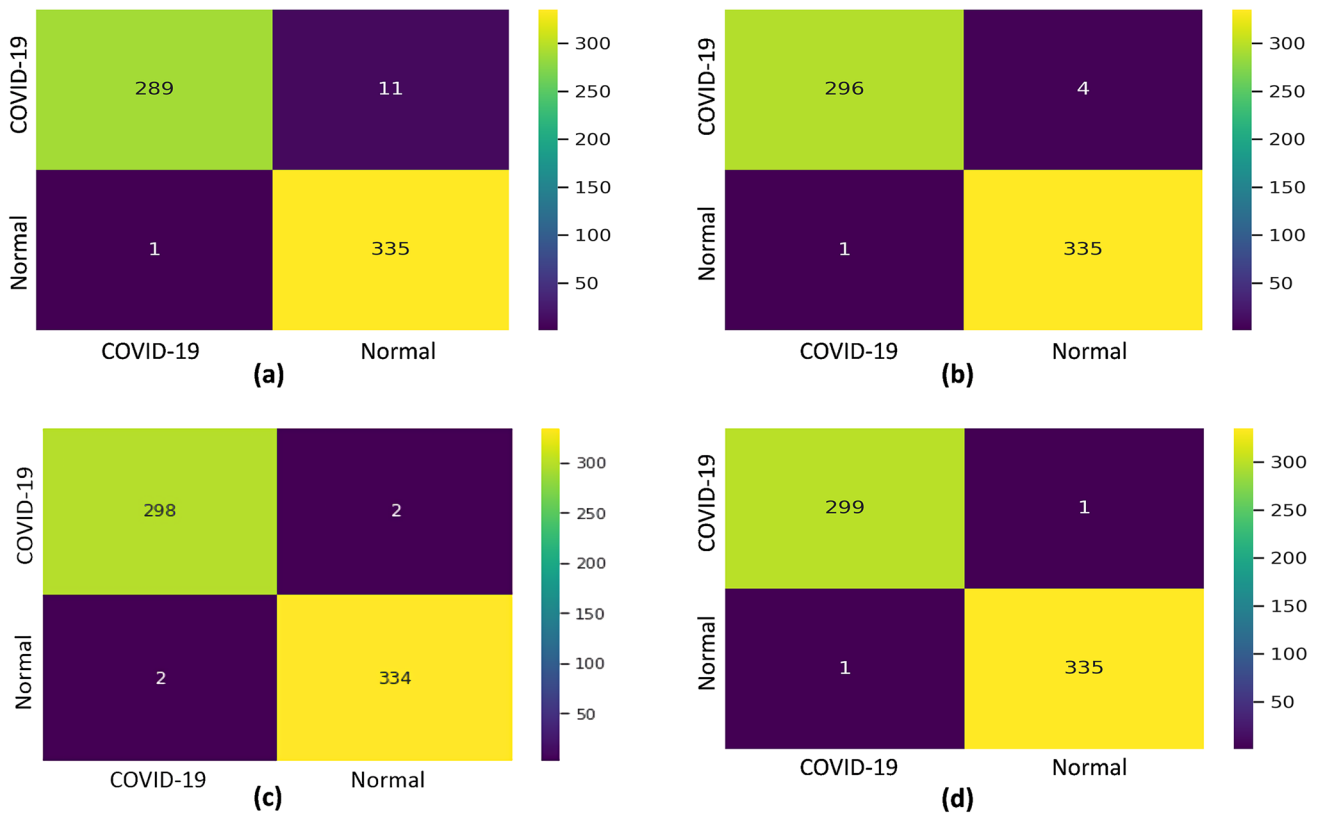


Fig. 8 Confusion matrices for **a** Xception, **b** DenseNet201, **c** MobileNetV2, **d** proposed Shallow CNN

In Fig. 10, the shallow CNN trained with the proposed method outperforms the entire set of pretrained CNN models with the least number of parameters, exactly 420,712. This is because the model holds the minimum depth and the total number of parameters involved in execution is very small compared to other models. As the depth increases, the total number of parameters increases exponentially, resulting in more storage space. The overall performance of the model is based on optimizing the hyperparameters contained in the network. Compared with the largest model Xception, with an accuracy of 98.11% and 21,062,186 million parameters, the Shallow CNN model achieves the best performance and parameter size ratio. With MobileNetV2 and Shallow CNN, even with a slight difference in performance, the proposed method still had a huge advantage with fewer parameters. Suppose we need COVID-19 testing in remote areas without adequate medical facilities. Ideally, the device needs a lightweight model that is computationally inexpensive and provides the highest accuracy in predicting COVID-19 cases. With minimal performance trade-offs, the proposed model is a lightweight architecture with fewer parameters compared to deep models. This significantly reduces the computational cost and can be run efficiently on IoT and edge devices. The low computation time required to train and test the proposed CNN makes it more suitable for large-scale

population screening, especially for diagnosing COVID-19 infection cases in resource-limited areas.

Table 11 shows the computation cost for total training time and test time (in seconds). It can be seen from Table 11 that the average training time of the proposed CNN is about 556 s. All other deep learning models have relatively long training times, ranging from 1062 to 4903 s. The computation time of the proposed CNN for test data is 22.44 s, which is the shortest run time among other test models. The table shows that the proposed CNN has lower computational costs due to a small number of parameters and minimal depth, while deep models have more layers and trainable parameters, which leads to longer training times.

5.6 Comparison with state-of-the-art methods

We also compared the results of our proposed method with existing methods that used a similar set of data sets. Table 12 presents previous studies conducted in the automated diagnosis of COVID-19 and their comparison with our proposed model. It is obvious from the table below that the proposed model outperforms the state of art models. In addition, the proposed model was trained on a balanced data set with the same number of images, making it more robust and easier to combat this pandemic in the near future.

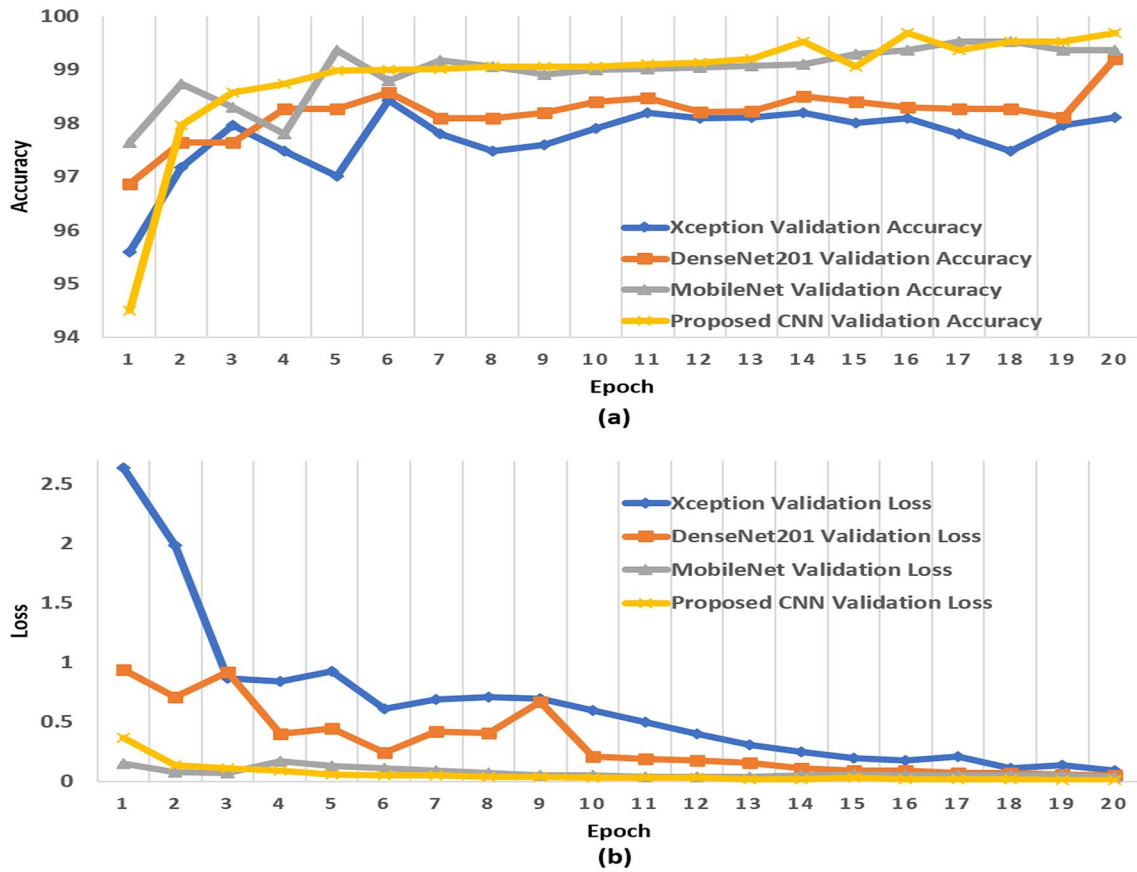
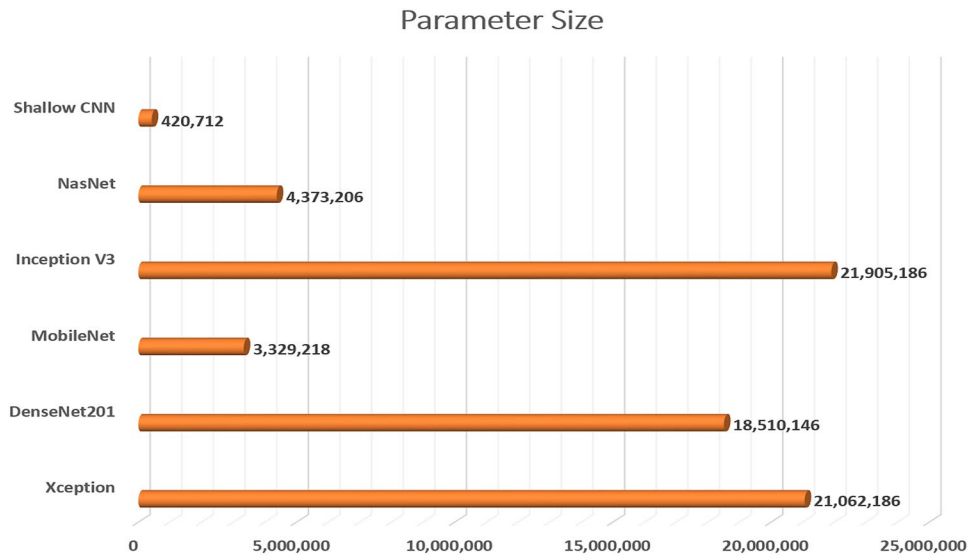


Fig. 9 Accuracy and loss comparison with Epoch

Fig. 10 Comparison of parameter sizes of proposed CNN with other state-of-the-art pre-trained CNNs



6 Conclusions

The COVID-19 pandemic is growing every day. As the number of cases continues to increase, it may be necessary to

quickly test cases in batches. The rapid outbreak of COVID-19 presents a global challenge to the scientific community, requiring it to develop new diagnostic tests to solve this public health problem. Rapid diagnosis plays a key role in combating pandemic situations, such as the current COVID-19.

Table 11 Comparison of computational time for training and testing in seconds

Model	Training time	Testing time
NasNet	3552.85	130.73
InceptionV3	1653.02	72.87
MobileNet	1062.76	59.79
DenseNet201	4903.24	245.24
Xception	1123.92	62.71
Proposed CNN	556.72	22.44

Researchers are very concerned about building a deep learning model with a lightweight architecture so that the symptoms of COVID-19 can be quickly and effectively identified. We aimed to propose a robust deep learning approach for detecting COVID-19 cases from chest X-rays to offer early and automatic detection with high accuracy and low false-negative rate. The proposed method is a lightweight shallow CNN with optimized parameters designed to identify COVID-19 cases using chest X-ray images. The study uses

open source chest X-rays from healthy people and patients with COVID-19. Our model achieves 99.68% accuracy in the classification of healthy and COVID-19-infected patients. The values for precision, sensitivity and specificity are also very good. The proposed approach has a precision, sensitivity, and specificity of 99.66, 99.66 and 99.70%, respectively. We have compared the performance of the proposed shallow CNN with Xception, DenseNet201, MobileNetV2 models and the existing techniques. The proposed method demonstrates the superiority of this method over existing state-of-the-art methods. Contrary to previous research work, the proposed model is computationally cheaper, requires fewer parameters, and achieves promising results on the data set. We believe that this study can be used as an initial screening that can help radiologists and healthcare professionals in screening COVID-19 positive cases. In the future, we will increase the size of the data set when it becomes available to make the model more robust.

Table 12 Comparison of proposed method with other existing deep learning methods

Authors	Method	No. of images	Accuracy (%)
Horry et al. [47]	VGG-19	140 COVID-19 400 Normal	87.87
Hemdan et al. [17]	COVIDX-Net	25 COVID-19 25 Normal	90
Khan et al. [26]	CoroNet	284 COVID-19 310 Normal	99
Rahimzadeh et al. [48]	Xception and ResNet50V2	180 COVID-19 851 Normal	91.4
Apostolopoulos et al. [16]	VGG-19	224 COVID-19 504 Normal	93.4
Toğaçar et al. [49]	SqueezeNet and MobileNetV2, SMO and SVM	295 COVID-19 65 Normal	98.25
Farooq and Hafeez [50]	Covid-ResNet	68 COVID-19 660 Normal	96.23
Ozturk et al. [23]	DarkCovidNet	125 COVID-19 500 Normal	98.08
Toraman et al. [51]	CapsNet	231 COVID-19 500 Normal	89.1
Das et al. [31]	Xception	–	97.4
Mukherjee et al. [52]	Lightweight CNN	168 COVID-19 168 Normal	96.28
Ucar et al. [27]	COVIDiagnosis-Net	45 COVID-19 1203 Normal	98.30
Wang et al. [53]	Xception + SVM	537 COVID-19 565 Normal	99.33
Gilanie et al. [54]	CNN	1066 COVID-19 1341 Normal	96.68
Das et al. [55]	VGG16	219 COVID-19 1341 Normal	97.67
Our proposed method	Shallow CNN MobileNetV2	1200 COVID-19 1341 Normal	99.68 99.37

Funding This work was supported by the Natural Science Foundation of Hunan Province, China (Grant No. 2020JJ4757).

Declarations

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

References

- Wu, Z., McGoogan, J.M.: Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* **323**(13), 1239–1242 (2020). <https://doi.org/10.1001/jama.2020.2648>
- WHO. WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020. <https://www.who.int/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020> (2020)
- Botros, N., Iyer, P., Ojcius, D.M.: Is there an association between oral health and severity of COVID-19 complications? *Biomed J.* **43**(4), 325–7 (2020). <https://doi.org/10.1016/j.bj.2020.05.016>
- Halpin, D.M., Faner, R., Sibila, O., Badia, J.R., Agusti, A.: Do chronic respiratory diseases or their treatment affect the risk of SARS-CoV-2 infection? *Lancet Respir. Med.* **8**(5), 436–438 (2020). [https://doi.org/10.1016/S2213-2600\(20\)30167-3](https://doi.org/10.1016/S2213-2600(20)30167-3)
- Chen, N., et al.: Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**(10223), 507–513 (2020). [https://doi.org/10.1016/S0140-6736\(20\)30211-7](https://doi.org/10.1016/S0140-6736(20)30211-7)
- Corman, V.M., et al.: Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* **25**(3), 2000045 (2020)
- Xie, X., Zhong, Z., Zhao, W., Zheng, C., Wang, F., Liu, J.: Chest CT for typical coronavirus disease 2019 (COVID-19) pneumonia: relationship to negative RT-PCR testing. *Radiology* **296**(2), E41–E45 (2020). <https://doi.org/10.1148/radiol.2020200343>
- Dong, D., et al.: The role of imaging in the detection and management of COVID-19: a review. *IEEE Rev. Biomed. Eng.* **14**, 16–29 (2020). <https://doi.org/10.1109/RBME.2020.2990959>
- Kanne, J.P., Little, B.P., Chung, J.H., Elicker, B.M., Ketaj, L.H.: Essentials for radiologists on COVID-19: an update—radiology scientific expert panel. *RSNA* (2020). <https://doi.org/10.1148/radiol.2020200527>
- Zou, L., Zheng, J., Miao, C., Mckeown, M.J., Wang, Z.J.: 3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI. *IEEE Access* **5**, 23626–23636 (2017). <https://doi.org/10.1109/ACCESS.2017.2762703>
- Liu, C. et al.: TX-CNN: Detecting tuberculosis in chest X-ray images using convolutional neural network. In 2017 IEEE international conference on image processing (ICIP), pp. 2314–2318 (2017)
- Zhao, X., Liu, L., Qi, S., Teng, Y., Li, J., Qian, W.: Agile convolutional neural network for pulmonary nodule classification using CT images. *Int. J. Comput. Assist. Radiol. Surg.* **13**(4), 585–595 (2018). <https://doi.org/10.1007/s11548-017-1696-0>
- Chouhan, V., et al.: A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Appl. Sci.* **10**(2), 559 (2020). <https://doi.org/10.3390/app10020559>
- Lakhani, P., Sundaram, B.: Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* **284**(2), 574–582 (2017). <https://doi.org/10.1148/radiol.2017162326>
- 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), 1–12 (2020). <https://doi.org/10.1038/s41598-020-76550-z>
- Apostolopoulos, I.D., Mpesiana, T.A.: Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci.* **43**(2), 635–6402 (2020). <https://doi.org/10.1007/s13246-020-00865-4>
- 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. Preprint at arXiv.2003.11055 (2020)
- Panwar, H., Gupta, P.K., Siddiqui, M.K., Morales-Menendez, R., Singh, V.: Application of deep learning for fast detection of COVID-19 in X-RAYS using nCOVnet. *Chaos Solit. Fractals* **138**, 109944 (2020). <https://doi.org/10.1016/j.chaos.2020.109944>
- Afshar, P., Heidarian, S., Naderkhani, F., Oikonomou, A., Plataniotis, K.N., Mohammadi, A.: Covid-caps: a capsule network-based framework for identification of covid-19 cases from x-ray images. *Pattern Recognit. Lett.* **138**, 638–643 (2020). <https://doi.org/10.1016/j.patrec.2020.09.010>
- Mangal, A. et al.: CovidAID: COVID-19 detection using chest X-ray. Preprint arXiv.2004.09803 (2020)
- Asif, S., Wenhui, Y., Jin, H., Jinhai, S.: Classification of COVID-19 from chest X-ray images using deep convolutional neural network. In 2020 IEEE 6th International Conference on Computer and Communications (ICCC), 426–433 (2020). <https://doi.org/10.1109/ICCC51575.2020.9344870>
- Oh, Y., Park, S., Ye, J.C.: Deep learning covid-19 features on cxr using limited training data sets. *IEEE Trans. Med. Imag.* **39**(8), 2688–2700 (2020). <https://doi.org/10.1109/TMI.2020.2993291>
- Ozturk, T., Talo, M., Yildirim, E.A., Baloglu, U.B., Yildirim, O., Acharya, U.R.: Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* **121**, 103792 (2020). <https://doi.org/10.1016/j.combiomed.2020.103792>
- Sethy, P.K.; Behera, S.K.; Ratha, P.K.: Detection of coronavirus disease (COVID-19) based on deep features and support vector machine. Preprints (2020)
- Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., Pinheiro, P.R.: Covidgan: data augmentation using auxiliary classifier gan for improved covid-19 detection. *IEEE Access.* **8**, 91916–91923 (2020). <https://doi.org/10.1109/ACCESS.2020.2994762>
- Khan, A.I., Shah, J.L., Bhat, M.M.: CoroNet: a deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Comput. Methods Programs Biomed.* **196**, 105581 (2020). <https://doi.org/10.1016/j.cmpb.2020.105581>
- Ucar, F., Korkmaz, D.: COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Med. Hypotheses* **140**, 109761 (2020). <https://doi.org/10.1016/j.mehy.2020.109761>
- El Asnaoui, K., Chawki, Y.: Using X-ray images and deep learning for automated detection of coronavirus disease. *J. Biomol. Struct. Dyn.* **39**(10), 3615–3626 (2020). <https://doi.org/10.1080/0739102.2020.1767212>
- Shelke, A., Inamdar, M., Shah, V., Tiwari, A., Hussain, A., Chafekar, T., Mehendale, N.: Chest X-ray classification using deep learning for automated COVID-19 screening. *SN Comput. Sci.* **2**(4), 1–9 (2021). <https://doi.org/10.1007/s42979-021-00695-5>
- 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). <https://doi.org/10.1007/s10044-021-00984-y>

31. Das, N.N., Kumar, N., Kaur, M., Kumar, V., Singh, D.: Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays. *IRBM* (2020). <https://doi.org/10.1016/j.irbm.2020.07.001>
32. Song, Y., et al.: Deep learning enables accurate diagnosis of novel corona. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **31**(8), 6096–6104 (2021)
33. Wang, S., et al.: A deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19). *Eur. Radiol.* **31**, 1–9 (2021). <https://doi.org/10.1007/s00330-021-07715-1>
34. Sedik, A., et al.: Deploying machine and deep learning models for efficient data-augmented detection of COVID-19 infections. *Viruses* **12**(7), 769 (2020). <https://doi.org/10.3390/v12070769>
35. Maghdid, H.S., Asaad, A.T., Ghafoor, K.Z., Sadiq, A.S., Mirjalili, S., Khan, M.K.: Diagnosing COVID-19 pneumonia from X-ray and CT images using deep learning and transfer learning algorithms. Preprint arXiv.2004.00038 (2020)
36. Shui-Hua, W., Khan, M.A., Govindaraj, V., Fernandes, S.L., Zhu, Z., Yu-Dong, Z.: Deep rank-based average pooling network for COVID-19 recognition. *Comput. Mater. Continua* **70**(2), 2797–2813 (2022)
37. Zhang, Y.-D., Zhang, Z., Zhang, X., Wang, S.-H.: MIDCAN: a multiple input deep convolutional attention network for Covid-19 diagnosis based on chest CT and chest X-ray. *Pattern Recognit. Lett.* **50**, 8–16 (2021). <https://doi.org/10.1016/j.patrec.2021.06.021>
38. Pan, S.J., Yang, Q.: A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22**(10), 1345–1359 (2009). <https://doi.org/10.1109/TKDE.2009.191>
39. Chowdhury, M.E., et al.: Can AI help in screening viral and COVID-19 pneumonia? *IEEE Access.* **8**, 132665–132676 (2020). <https://doi.org/10.1109/ACCESS.2020.3010287>
40. Mooney, P.: Chest X-Ray Images (Pneumonia). <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia> (2018)
41. Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., Wojna, Z.: Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 2818–2826 (2016)
42. François, C.: Xception: Deep learning with depthwise separable convolutions. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 1800–1807 (2017)
43. Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., Chen, L.C.: Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4510–4520 (2018)
44. Zhang, Q., Wu, Y.N., Zhu, S.C.: Interpretable convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. pp. 8827–8836 (2018)
45. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 4700–4708 (2017)
46. Fernández, A., García, S., del Jesus, M.J., Herrera, F.: A study of the behaviour of linguistic fuzzy rule based classification systems in the framework of imbalanced data-sets. *Fuzzy Sets Syst.* **159**(18), 2378–2398 (2008). <https://doi.org/10.1016/j.fss.2007.12.023>
47. Horry, M.J., et al.: COVID-19 detection through transfer learning using multimodal imaging data. *IEEE Access.* **8**, 149808–149824 (2020). <https://doi.org/10.1109/ACCESS.2020.3016780>
48. Rahimzadeh, M., Attar, A.: A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. *Inform. Med. Unlocked* **19**, 100360 (2020). <https://doi.org/10.1016/j.imu.2020.100360>
49. Toğaçar, M., Ergen, B., Cömert, Z.: COVID-19 detection using deep learning models to exploit social mimic optimization and structured chest X-ray images using fuzzy color and stacking approaches. *Comput. Biol. Med.* **121**, 103805 (2020). <https://doi.org/10.1016/j.combiomed.2020.103805>
50. Farooq, M., Hafeez, A.: Covid-resnet: A deep learning framework for screening of covid19 from radiographs. Preprint arXiv.2003.14395 (2020)
51. Toraman, S., Alakus, T.B., Turkoglu, I.: Convolutional capsnet: a novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Solit. Fractals.* **140**, 110122 (2020). <https://doi.org/10.1016/j.chaos.2020.110122>
52. Mukherjee, H., Ghosh, S., Dhar, A., Obaidullah, S.M., Santosh, K.C., Roy, K.: Deep neural network to detect COVID-19: one architecture for both CT Scans and Chest X-rays. *Appl. Intell.* (2020). <https://doi.org/10.1007/s10489-020-01943-6>
53. Wang, D., Mo, J., Zhou, G., Xu, L., Liu, Y.: An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images. *PLoS One* **15**(11), e0242535 (2020). <https://doi.org/10.1371/journal.pone.0242535>
54. Gilanie, G.: Coronavirus (COVID-19) detection from chest radiology images using convolutional neural networks. *Biomed. Signal Process Control.* **66**, 102490 (2021). <https://doi.org/10.1016/j.bspc.2021.102490>
55. Das, A.K., Kalam, S., Kumar, C., Sinha, D.: TLCoV-An automated Covid-19 screening model using transfer learning from chest X-ray images. *Chaos Solit. Fractals.* **144**, 110713 (2021). <https://doi.org/10.1016/j.chaos.2021.110713>

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