Detection of COVID-19 from chest X-ray images: Boosting the performance with convolutional neural network and transfer learning

Sohaib Asif, Yi Wenhui, Kamran Amjad, Hou Jin, Yi Tao, Si Jinhai

Abstract
Coronavirus disease (COVID-19) is a pandemic that has caused thousands of casualties and impacts all over the world. Most countries are facing a shortage of COVID-19 test kits in hospitals due to the daily increase in the number of cases. Early detection of COVID-19 can protect people from severe infection. Unfortunately, COVID-19 can be misdiagnosed as pneumonia or other illness and can lead to patient death. Therefore, in order to avoid the spread of COVID-19 among the population, it is necessary to implement an automated early diagnostic system as a rapid alternative diagnostic system. Several researchers have done very well in detecting COVID-19; however, most of them have lower accuracy and overfitting issues that make early screening of COVID-19 difficult. Transfer learning is the most successful technique to solve this problem with higher accuracy. In this paper, we studied the feasibility of applying transfer learning and added our own classifier to automatically classify COVID-19 because transfer learning is very suitable for medical imaging due to the limited availability of data. In this work, we proposed a CNN model based on deep transfer learning technique using six different pre-trained architectures, including VGG16, DenseNet201, MobileNetV2, ResNet50, Xception, and EfficientNetB0. A total of 3886 chest X-rays (1200 cases of COVID-19, 1341 healthy and 1345 cases of viral pneumonia) were used to study the effectiveness of the proposed CNN model. A comparative analysis of the proposed CNN models using three classes of chest X-ray datasets was carried out in order to find the most suitable model. Experimental results show that the proposed CNN model based on VGG16 was able to accurately diagnose COVID-19 patients with 97.84% accuracy, 97.90% precision, 97.89% sensitivity, and 97.89% of F1-score. Evaluation of the test data shows that the proposed model produces the highest accuracy among CNNs and seems to be the most suitable choice for COVID-19 classification. We believe that in this pandemic situation, this model will support healthcare professionals in improving patient screening.

KEYWORDS
chest X-rays, COVID-19 detection, deep CNN, medical image analysis, transfer learning, VGG16
The novel coronavirus (COVID-19) is an acutely fatal disease that originated in Wuhan, China in December 2019 and spread globally. COVID-19 outbreak has always been a great concern of the health community because no cure has been discovered (Rothe et al., 2020). On 11th March 2020, World health Organization (WHO) declared this disease as a pandemic (World Health Organization, 2020). The structure of COVID-19 composed of single stranded RNA. Due to its mutating feature, this disease is very hard to cure. The new coronavirus disease is firstly a throat infection, and suddenly people have trouble in breathing. The other common symptoms of this virus are headache, fever, shortness of breath and cough. These symptoms do not indicate the COVID-19, but in some cases, it leads to pneumonia. This can lead us to diagnostic problem of the healthcare professional. To protect the healthy people from COVID-19, infected patients have to put in isolation, do proper examination and take other protective measures (Roosa et al., 2020). COVID-19 is a contagious disease, which can transfer from one person to another thorough breath, hand and mucus contacts (Yan et al., 2020). This virus mostly attacks the person with weak immune system but healthy people can also have this disease (Lancet, 2020; Razai et al., 2020). COVID-19 has different types and found in animals too. Despite the introduction of vaccines worldwide, some countries in the world still have new cases and new death records every day. Most of the new cases reported are due to the second and third wave, requiring immediate and long-term solutions for early detection of this disease. This is very important to find a solution for early detection to prevent further spread of COVID-19. Polymerize chain reaction (PCR) test is to diagnose the COVID-19 but the time required for this test is relatively long and it is expensive. In this type of test, a long swab is inserted into the patient's nostrils or back of the throat to get the sample and the results obtained within few days. In some cases, test results are negative, even if the patient shows progress on subsequent computed tomography (CT). Due to the limited availability of RT-PCR in some countries, it is recommended to use CT-scan and X-rays. By using CT-scans and X-rays, the detection of coronavirus symptoms has high accuracy in the lower parts of the lungs. In some cases, CT-scans and X-rays can also be used as a substitution of RT-PCR test. There are limited numbers of radiologists in the world and the people infected with this virus are quite a lot. There are also people who re-examine themselves to know the progress of this disease. To overcome the problem of delaying reports and assist radiologists, there is a need to develop non-invasive system for the early detection of Covid-19 using state of the art artificial intelligence (AI) tools.

Deep learning is a branch of machine learning with three or more layers called neural networks. These neural networks behave like a human brain and learn from a large amount of data. Artificial intelligence (AI) and deep learning is being used in different fields of biomedical research such as medical image analysis (AlZu'bi, Mughaid, et al., 2019), drug discovery, blood donation (AlZu'bi, Aqel, & Mughaid, 2021) and biomedicine. There are different types of models presented by researcher to diagnose, segment and predict the disease such as lungs and skin cancer detection, diabetes prediction and detection, heart disease detection, dengue and Parkinson's disease etc. AI is used in image classification, detection and medical image segmentation (AlZu'bi et al., 2020; Al-Zu'bi et al., 2017; Al-Zu'bi et al., 2021). Strong GPU support and the size and quality of the dataset are also factors that positively impact the network accuracy (AlZu'bi et al., 2018). Using computer-aided applications, Artificial intelligence (Li, Qin, et al., 2020) and deep learning (Li, Wang, et al., 2020) plays an important role in identify and classify COVID-19 cases, and have achieved excellent results (Liu et al., 2020). Deep learning models can also be used to forecast the nature of the virus and helps control the spread of this disease (Wiezorek et al., 2020). Chest X-rays and computed tomography (CT) are the most widely used methods for processing medical images and have been used by many researchers to develop models that can help radiologist predict disease. Deep learning methods produces state of the art results in medical image analysis. Recent works (LeCun et al., 2015; Litjens et al., 2017; Shin et al., 2016) using computer vision (CV) and artificial intelligence (AI) techniques, including the use of deep learning (DL) models, especially convolutional neural networks (CNNs) has proven to be a useful method for examining medical images. CNNs are widely used to automate the analysis of infections in chest radiographs to detect COVID-19 imaging patterns. However, one of the main limitations of current approaches is overfitting issues due to the limited amount of related COVID-19 data that is publicly available for use, making it much more difficult to create an efficient model with overall high performance. Some existing models reported lower accuracy scores and needed to be validated on a large dataset, and there was no evidence that standard performance measures such as Matthew's correlation coefficient were used in their approaches. Therefore, it is very important to develop a robust and effective system to automatically diagnose COVID-19. The aforementioned problems are addressed in this paper. The results of this study are promising and validate the efficiency of transfer learning in the automatic detection of COVID-19 disease.

In our previous study in December 2020, we used the InceptionV3 model to diagnose COVID-19 using chest X-rays (Asif et al., 2020). We use the transfer learning method and remove the last layer of the InceptionV3 model, and then use the softmax activation function to retrain the last layer on the chest X-ray dataset. However, due to the lack of a COVID-19 dataset, we were only able to train our model on 864 COVID-19 images. After training the InceptionV3 model, we achieved a validation accuracy of 93% on the COVID-19 chest X-ray dataset. This work is an updated version of our previous work, but the method has been greatly expanded and improved. The shortcomings of the existing work identified by the current study led us to propose a deep learning-based system for efficient and cost-effective detection of COVID-19 with low misclassification rates. The diagnostic system presented in this paper utilizes a pre-trained model based on the concept of transfer learning to extract distinctive features present in chest radiographs of COVID-19 patients. Transfer learning is the best strategy to achieve accurate feature extraction and classification. It will save time and computing power and provide higher accuracy than competing models. Deep transfer learning was performed on six popular deep learning architectures (VGG16, DenseNet201, MobileNetV2, ResNet50, Xception, and EfficientNetB0) to
overcome the problem of insufficient training data for deep learning. An open dataset of chest X-ray images from the Kaggle repository was used to evaluate the performance of the proposed study. The dataset consists of 3886 chest X-ray images, of which 1200 are COVID-19, 1341 normal chest X-rays, and 1345 viral pneumonia images. Each model consists of two parts: a feature extractor and a classifier. The feature extractor consists of a convolutional base layers to extract features from the images, while the classifier serves to classify the extracted features. We retain the convolutional base layers and customize the final classification layer by adding new sets of layers such as multiple dense layers and batch normalization layers to alleviate the gradient disappearance problem, reduce the need for dropout, produce faster training, enhance learning, improve accuracy, and reduce overfitting. In this stage, only the additional layers are trained, while the other weights of the model remain frozen. The new set of layers consists of a flatten layer at the top, which transforms the data from the previous layer into a one-dimensional data vector. The classification part consists of two dense layers, followed by a batch normalization layer. We introduced the batch normalization layer between the two dense layers because it reduces the need for dropouts, resulting in faster training and higher accuracy. The first dense layer consists of 128 neurons with RELU activation function and the final output is produced by a dense layer with three neurons that uses SoftMax activation function for multi-class classification. This provides the respective probabilities of COVID-19, normal and viral pneumonia. Since the data in previous work was insufficient, this motivated us to further research and develop a more robust multi-classification deep CNN model based on the transfer learning method by increasing the amount of data to identify patients infected with COVID-19 to help healthcare professionals.

The main contributions of our proposed research are:

1. A novel and robust deep CNN model has been developed utilizing state-of-the-art deep learning architectures for multi-classification of COVID-19, viral pneumonia, and normal patients.
2. Six popular deep learning architectures, namely VGG16, DenseNet201, MobileNetV2, ResNet50, Xception, and EfficientNetB0 are trained by deploying transfer-learning technique on publicly available COVID-19 chest X-ray dataset.
3. Data augmentation is applied to reduce overfitting issues, increase the size of the data set, and improve the performance of the model for classifying COVID-19 cases.
4. We conduct a comprehensive analysis to assess the effectiveness of the proposed multi-class classification model using a variety of performance evaluation metrics, including accuracy, precision, sensitivity, F1-score, Matthew’s Correlation Coefficient (MCC), and confusion matrix.
5. Our experimental results showed that the proposed CNN model provides better classification accuracy compared to existing modern approaches. The proposed model is an end-to-end structure that eliminates the need for manual feature extraction and can reduce the workload of radiologists and identify COVID-19 more efficiently.

This article is structured as follows: Section 1 describes an introduction, the related work is described in Section 2. Section 3 details the datasets used, X-ray image pre-processing, the classification model used in this article, and model hyperparameters. Experimental results and discussion are presented in Section 4. Finally, we conclude our research, and areas for future research are discussed in detail in Section 5.

## 2 RELATED WORK

Many researchers are doing their best to use artificial intelligence technology to develop new systems that support the diagnosis of COVID-19. Most countries use chest X-rays and computed tomography (CT) images because they are relatively easy to obtain from patients and are inexpensive. In a variety of different image classification tasks, deep learning has been effectively applied in the medical field, and has significant performance compared with human-level performance. Several deep learning-based medical imaging systems have also been adopted to assist clinicians in classifying patients with COVID-19 infection. For example, Ozturk et al. in 2020 proposed an automatic COVID-19 detection system based on the DarkNet model to perform binary and multi-class classification tasks. This achieves a classification accuracy of 98.08%. Luz et al. in 2021 studied the deep learning architecture used to identify COVID-19. They used the COVID-19 chest X-ray dataset to prove the efficiency of the proposed model. The EfficientNet model achieved an accuracy of 93.9%. In Reference (Verma et al., 2020), the author used the InceptionV3 model to extract features from the dataset and then used an artificial neural network to classify the images. This model was able to classify the disease in a qualified way and achieved a very high accuracy of 99.01%. In Reference (Shelke et al., 2021), the author proposed a classification model for COVID-19 classification, which is analysed by VGG16 with an accuracy rate of 95.9%. In Reference (Apostolopoulos & Mpesiana, 2020), the author uses MobileNetV2 and VGG19 together for different data sets, and MobileNetV2 obtains 97.40% accuracy rate. Nayak et al. in 2021 proposes an automated approach supported by deep learning based on X-ray images. They evaluated the effectiveness of eight different pre-trained CNN models for COVID-19 classification from chest X-ray. The best performance was obtained with 98.33% accuracy by ResNet-34. Jain et al. in 2020 implemented the ResNet-101 model for COVID-19 classification and achieved 97.78% accuracy. Wang, Lin, & Wong in 2020 proposed a COVIDNet model for detecting COVID-19 using chest X-rays. The model was trained to classify chest X-rays into COVID-19, normal and pneumonia classes. They used 16,756 images to test the COVIDNET model and achieved a classification accuracy of 92.4%. Sethy and Behra in 2020 used the ResNet 50 model and SVM classifier to classify COVID-19 and Normal, achieving 95.38% accuracy. The author of Reference
(Hammoudi et al., 2021) leveraged the standard version of DenseNet-169 to achieve 95.72% accuracy. A pre-trained VGG-16 model using data augmentation technology to classify COVID-19 gives a 95% accuracy rate. Narin et al. in (2021) studied five pre-trained deep learning architectures to detect coronavirus infections in chest X-ray images. They achieved 99.5% accuracy with the ResNet50 model for COVID-19 chest X-ray classification. In another work, Ouchicha et al. in 2020 proposed a CVDNet model to classify COVID-19 cases by chest radiographs. They achieved 96.69% accuracy for multi-class chest X-ray classification. Makris et al. in 2020 used several pre-trained CNN models and compared their performance in a three-class classification of chest X-ray images. The VGG16 was found in their study to perform best with an overall accuracy of 95.88%. Farooq and Hafeez in 2020 proposed a COVID-ResNet model. They used a pretrained ResNet30 and achieved an overall accuracy rate of 96.23% for three-class classification of chest X-rays. Gupta et al. in 2021 proposed InstaCovNet-19 with an integrated stacked deep convolutional network, which utilizes five different pre-trained architectures. The proposed model achieves an accuracy of 99.53% in the binary classification and 99.08% in the three-category classification. Das et al. in 2020 proposed an Xception model for detecting COVID-19 patients. The model has been tested on the publicly available chest ray data set, with an accuracy rate of 97.40%. Xu et al. (2020) proposed ResNet+ Location Attention for three types of classification, using CT images to achieve an accuracy of 86.7%. Khan et al. in 2020 proposed the Xception-based Coronet model, which has been used for multi-class classification using X-ray images. They achieved an accuracy of 95% for the three-level classification. Jishi et al. in 2020 provides an in-depth analysis of COVID-19 transmission trends. The artificial intelligence-based tool can identify COVID-19 faster than traditional medical data reporting systems by processing epidemiological data. In Reference (Ahuja et al., 2021), the authors used 349 COVID-19 images and 397 normal images and applied data augmentation techniques to increase the size of the dataset. They used transfer-learning technique for binary classification and achieved 99.4% accuracy using the ResNet18 model. Singh et al. (2021) fine-tuned the VGG16 model to extract features from the images and performed automatic detection of COVID-19 using four different classifiers. By using Bagging Ensemble with SVM, a detection rate of 95.7% was achieved in 385 ms. AlZu’bi, Makki, et al. (2021) explains the economic impact of the Covid-19 pandemic in Jordan and proposes a mechanism to address the economic impact of the vaccine's spread across the country. AlZu’bi, Jararweh, et al. (2019) proposes a method based on multiresolution analysis to segment medical volumes under various conditions to facilitate the work of radiologists.

It can be seen from the above research that the use of deep learning technology to identify the new coronavirus on radiological images may reduce the pressure on radiologists. However, it is still not clear which model will give the best results, as different researchers use different methods to classify COVID-19. From the explanation given above, it is obvious that most of the models proposed so far do not provide classification performance and indicate a limitation related to the number of samples to conduct experiments. Most of the methods have overfitting problems and there is no evidence that standard evaluation metrics, such as MCC, are used in their methods. In addition, most models use a small number of samples for training, and in most cases, the data is unbalanced, so they may lack robustness. This led us to develop a deep learning model that can be used to reduce the workload of clinicians and help provide them with a second opinion.

3 | MATERIALS AND METHODS

This section describes the proposed approach for classifying COVID-19 diseases based on the Deep CNN architecture. The purpose of our research is to use six different models to propose a deep learning model for COVID-19 classification. Figure 1 shows the workflow of our proposed method from beginning to end. The figure clearly shows that the model consists of three main stages: data preparation and preprocessing, deep learning models for classification, and hyperparameter tuning. The proposed model uses chest X-rays as input, and the final output is to classify the input image into one of three categories: COVID-19, Viral Pneumonia, and Normal.

3.1 | Transfer learning models for classification

Nowadays, AI research based on deep learning provides the most advanced solutions for computer vision. CNN is a deep learning technique that has proven to achieve excellent results in a wide range of image classification tasks. However, the availability of COVID-19 X-ray images is limited, it is difficult to train these models from scratch to predict COVID-19. In the field of medical imaging, there is a dearth of labelled data, and this is a major challenge when building a high-performance deep learning system. Therefore, these challenges were resolved through a pre-trained CNN network using the concept of transfer learning (TL). In TL, the weights of a particular model, pre-trained on some dataset, such as ImageNet, are used to solve a related problem with a relatively smaller dataset. Moreover, a pre-trained network has been found to perform better than a network trained from scratch. We have seen that CNN with TL plays an interesting role in classification, and it is faster and more efficient than developing classification solutions without TL. Thus, we have adopted six pre-trained CNN models based on the TL concept. The complete TL procedure is shown in Figure 2. The main motivation behind the development of our proposed CNN model is the automatic detection of COVID-19 patients with maximum efficiency and faster detection time. Here, in this work, VGG16, DenseNet201, MobileNetV2, ResNet50, Xception, EfficientNetB0 have been used as the base model, pre-trained on the ImageNet dataset for classification. These models are trained using a
transfer learning approach to develop CNN model. Each model consists of two parts: a feature extractor and a classifier. The feature extractor consists of a convolutional base layers to extract features from the images, while the classifier serves to classify the extracted features. We retain the convolutional base layers and customize the final classification layer by adding new sets of layers such as multiple dense layers and batch normalization layers to alleviate the gradient disappearance problem, reduce the need for dropout, produce faster training, enhance learning, improve accuracy, and reduce overfitting. In this stage, only the additional layers are trained, while the other weights of the models remain frozen. The new set of layers consists of a flatten layer at the top, which transforms the data from the previous layer into a one-dimensional data vector. The classification part consists of two dense layers, followed by a batch normalization layer. We introduced the batch normalization layer between the two dense layers because it reduces the need for dropouts, resulting in faster training and higher accuracy. The first dense layer consists of 128 neurons with RELU activation function and the final output is produced by a dense layer with three neurons that uses SoftMax activation function for multi-class classification. This provides the respective probabilities of COVID-19, normal and viral pneumonia. A detailed explanation of the models and how they are used for classification is mentioned in the next section.

3.1.1 | VGG16

VGG architecture was developed by a team of Oxford University. VGG16 (Simonyan & Zisserman, 2014) is trained on ImageNet database which contains more than 1 million images. Since the VGG-16 network has undergone extensive training, it can provide excellent accuracy even if the image data set is small. VGG16 has given the name because of the 16 layers in its architecture. Apart from having 16 layers in its architecture,
VGG16 has a 3x3 receptive field. It consists of five convolutional blocks and each block contain different convolutional layers combine with Relu activation layer and max pooling layer. In this work, the last fully connected layer is removed and replaced with the designed classifier using the flatten layer, dense layer and batch normalization layer as shown in Figure 3.

3.1.2 | DenseNet201

DenseNet (Huang et al., 2017) is known for its outstanding performance on four object recognition benchmark tasks such as CIFAR-100, CIFAR-10, ImageNet and SVHN. To maximize the flow of information between the various layers in the network, the architecture of DenseNet uses a feed forward connection to connect different layers. DenseNet has different benefits like alleviating the vanishing gradient problem, encouraging feature reuse, and dramatically decreasing the number of parameters. DenseNet201 has 201 layers in its architecture that is why it is called DenseNet201. This model can achieve high performance using less memory and low computational cost. In this work, the last fully connected layer is removed and replaced with the designed classifier using the flatten layer, dense layer and batch normalization layer as shown in Figure 4.

3.1.3 | MobileNetV2

MobileNetV2 (Sandler et al., 2018) architecture has been released in early 2018. MobileNetV2 is based on some ideas from MobileNetV1 and optimized with new ideas. Architecturally, MobileNetV2 adds two new modules to the architecture. First, is introduction of linear bottlenecks between layers and second is fast connections between bottlenecks. The core idea of MobileNetV2 is that bottlenecks encode intermediate inputs and outputs of the model, and inner layers are used to encapsulate the model from low-level concepts (e.g. pixels, etc.) to high-level descriptors (e.g. image categories etc.). Finally, as with traditional residual connections, shortcuts allow for faster training and higher accuracy. In this work, the last fully connected layer is removed and replaced with the designed classifier using the flatten layer, dense layer and batch normalization layer as shown in Figure 5.

3.1.4 | ResNet50

This network is very deep. We can use a standard network component called a residual module to form a more complex network (which can be called a network within a network) and train it using the standard stochastic gradient descent method. The ResNet (He et al., 2016) architecture has become a pioneering work that demonstrates that extremely deep networks can be trained with standard SGD by using the residual module. By using identity-mapping technique, accuracy can be achieved by updating the residual module. This architecture is deep because it uses global average pooling layer instead of fully connected layer and hence makes the size small of this architecture. It is called ResNet50 because it has

![VGG16 architecture](image3.png)  
**FIGURE 3** VGG16 architecture designed for multiclass classification.

![DenseNet201 architecture](image4.png)  
**FIGURE 4** DenseNet201 architecture designed for multiclass classification.
50 layers in its architecture. In this work, the last fully connected layer is removed and replaced with the designed classifier using the flatten layer, dense layer and batch normalization layer as shown in Figure 6.

3.1.5 | Xception

The Xception network is an extension of the Inception network. Xception (François, 2017) uses a new concept called depth-separable convolutional operations that can outperform Inception-v3 on a large image classification dataset with 350 million images and 17,000 classes. The Xception architecture has the same number of parameters as Inception-v3, and the performance gains are not due to increased capacity, but to more efficient use of model parameters. Xception has the smallest weight, which is only 91 MB. In this work, the last fully connected layer is removed and replaced with the designed classifier using the flatten layer, dense layer and batch normalization layer as shown in Figure 7.

3.1.6 | EfficientNetB0

Researchers at Google proposed a new model scaling method for EfficientNet (Tan & Le, 2019) base network model that uses simple and efficient composite coefficients to weigh network depth, width, and input image resolution. A series of EfficientNet models were obtained by scaling up the EfficientNet base model. This series of models beats all previous convolutional neural network models in terms of efficiency and accuracy. The core structure of EfficientNetB0 is the mobile inverted bottleneck convolution (MBConv) module, which also introduces the attention
mechanism of Squeeze-and-Excitation Network (SENet). In this work, the last fully connected layer is removed and replaced with the designed classifier using the flatten layer, dense layer and batch normalization layer as shown in Figure 8.

4 | EXPERIMENTAL SETUP AND RESULTS

This section describes the experimental setup, data sets used, data preprocessing and enhancement, details of model implementation, performance metrics, and then analyzes the performance of the proposed model. Finally, the performance of the proposed model is compared with the state-of-the-art methods.

4.1 | COVID-19 chest X-ray dataset

This study uses an open access database of chest X-ray images from the Kaggle repository (Chowdhury et al., 2020). A team of researchers from the University of Qatar and Dhaka University collaborated with doctors from Pakistan and Malaysia to create a chest X-ray image dataset and has been updated regularly. They have also added COVID-19 chest X-ray images from the Italian Society of Medical and Interventional Radiology (SIRM) database. The dataset is divided into three different classes: COVID, normal, and viral pneumonia. There are 1200 positive images of Covid-19, 1345 images of viral pneumonia, and 1341 normal images - a sample images belonging to all three classes depicted in Figure 9. Table 1 shows the number of X-ray images in each class.

4.2 | Image preprocessing

Image preprocessing is one of the important prerequisites for developing a more efficient detection system. Due to differences in the dimensions of the images in the dataset, all images in the dataset have been resized to 224 × 224 pixels for input to various architectures such as VGG16, DenseNet201, MobileNetV2, ResNet50, Xception, and EfficientNetB0. In addition, we split the data set into training images and test images. Seventy-five percent of the total images are reserved for training purposes, and 25% of the images are reserved for testing purposes.

4.3 | Data augmentation

The deep learning model we used in this research requires a lot of data to be effectively trained. However, the size of the input data is not large enough. Overfitting can happen on small datasets. Overfitting is a situation where the model is unable to identify unknown patterns. In this
situation, a data augmentation technique was applied to synthetically increase the size of the data over the original data, and also help reduce overfitting during CNN training. In this article, the following techniques are used to augment the image: (1) Horizontal flipping is applied with rotation range 20, (2) The rotation is done by rotating the image by 15° and (3) Scaling. It is worth noting that three augmentation strategies are used in order to generate a new training set. Initially, 2915 images were allocated to train the models for the multiclass classification task, resulting in 11,660 training images after the data augmentation. Which is four times larger than the original training images.

TABLE 1  Chest X-ray image dataset belonging to each class.

| Class              | Number of images |
|--------------------|------------------|
| COVID-19           | 1200             |
| Viral pneumonia    | 1345             |
| Normal             | 1341             |
| Total              | 3886             |

FIGURE 9  Sample of chest X-ray from the dataset. The first row represents COVID-19 images, the second row represents normal images, and the third row represents viral pneumonia images.

4.4 | Models implementation

In this study, all deep learning models-VGG16, DenseNet201, MobileNetV2, ResNet50, Xception, EfficientNetB0 are implemented using the TensorFlow package, which provides a Python application programming interface (API). We also used Keras API as the official front end of
TensorFlow for the development and implementation of CNN algorithms. In addition, all experiments were performed on Google Colab with 12 GB RAM. These models run on a 12 GB NVIDIA Tesla K80 GPU. We use the Adam optimizer to train and optimize the model, and the categorical cross-entropy loss function is used for 3-class classification. An epoch is a cycle of updating network weights using the entire training data. As the number of epochs increases, the performance of the model will improve over time. All models were run over 20 epochs with a learning rate of 0.001 and a batch size of 64. To overcome the problem of overfitting, we implemented an early stop technique based on validation performance. The parameters used to train the model are shown in Table 2. The achieved test accuracy is also shown in Table 2.

4.5 | Performance evaluation metrics

In this section, we mentioned different evaluation metrics used to evaluate the effectiveness of the proposed CNN model. The models were tested on 300 COVID-19 images, 336 normal images, and 337 viral pneumonia images. The performance of each model is evaluated based on different metrics, such as accuracy, sensitivity, precision, F1-score, and Matthews Correlation Coefficient (MCC). The confusion matrix was chosen to obtain the values of the performance metrics. The confusion matrix is shown in Figure 10. These metrics were computed using various parameters of the confusion matrix, such as TP, FP, TN, and FN, representing true positive, false positive, true negative, and false negative. TP indicates the number of positive patients correctly predicted, FN indicates infected patients incorrectly identified as normal or pneumonia cases, FP indicates normal or pneumonia patients predicted as COVID-19, and TN indicates healthy or pneumonia patients correctly predicted. Different performance evaluation metrics are defined as follows:

\[
\text{Accuracy} = \frac{(TP + TN)}{(TP + FP + TN + FN)},
\]

\[
\text{Sensitivity} = \frac{TP}{TP + FN},
\]

\[
\text{Precision} = \frac{TP}{TP + FP}.
\]

| Performance measures | VGG16 | DenseNet201 | MobileNetV2 | ResNet50 | Xception | EfficientNetB0 |
|-----------------------|-------|-------------|-------------|----------|-----------|----------------|
| Batch Size            | 64    | 64          | 64          | 64       | 64        | 64             |
| Layers                | 16    | 201         | 53          | 50       | 71        | 237            |
| Optimizer             | Adam  | Adam        | Adam        | Adam     | Adam      | Adam           |
| Parameters (M)        | 138   | 20.2        | 3.4         | 25.6     | 22.9      | 5.3            |
| Activation function   | Softmax | Softmax     | Softmax     | Softmax  | Softmax   | Softmax        |
| Learning rate         | 0.001 | 0.001       | 0.001       | 0.001    | 0.001     | 0.001          |
| Custom input size     | 224 x 224 | 224 x 224 | 224 x 224 | 224 x 224 | 224 x 224 | 224 x 224 |
| Testing accuracy      | 0.978 | 0.957       | 0.942       | 0.974    | 0.937     | 0.959          |

**Figure 10** Confusion matrix.
\[ F_1 – \text{score} = 2 \times \frac{\text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \]  
\[ \text{MCC} = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \]  

### 4.6 Performance analysis

This section presents detailed experiments and results of automatic detection and classification of COVID-19. Extensive experiments were conducted using several state-of-the-art models such as VGG16, DenseNet201, MobileNetV2, ResNet50, Xception and EfficientNetB0 to develop a robust classifier. The main goal of this study is to examine the success of the deep learning architectures and compare it with the performance of other CNN models in the literature. The six architectures used in the study were evaluated based on precision, sensitivity, and \( F_1 \)-score performance indicators. Table 3 shows the class-wise performance of the models. The different classes used in this study are COVID-19, viral pneumonia, and normal. Table 3 shows the performance metrics and macro average scores obtained for each model. From the results table, we can see

| Model and class | Precision | Sensitivity | \( F_1 \)-score | Accuracy |
|-----------------|-----------|-------------|----------------|----------|
| **VGG16**       |           |             |                |          |
| COVID-19        | 0.99      | 0.99        | 0.99           |          |
| Viral pneumonia | 0.97      | 0.97        | 0.97           |          |
| Normal          | 0.98      | 0.97        | 0.97           |          |
| Macro average   | 0.98      | 0.98        | 0.98           | 0.978    |
| **DenseNet201** |           |             |                |          |
| COVID-19        | 0.99      | 0.95        | 0.97           |          |
| Viral pneumonia | 0.95      | 0.94        | 0.94           |          |
| Normal          | 0.94      | 0.98        | 0.96           |          |
| Macro average   | 0.96      | 0.96        | 0.96           | 0.957    |
| **MobileNetV2** |           |             |                |          |
| COVID-19        | 1.00      | 0.93        | 0.96           |          |
| Viral pneumonia | 0.89      | 0.96        | 0.92           |          |
| Normal          | 0.95      | 0.94        | 0.95           |          |
| Macro average   | 0.95      | 0.94        | 0.94           | 0.942    |
| **ResNet50**    |           |             |                |          |
| COVID-19        | 1.00      | 0.99        | 0.99           |          |
| Viral pneumonia | 0.95      | 0.97        | 0.96           |          |
| Normal          | 0.97      | 0.96        | 0.97           |          |
| Macro average   | 0.98      | 0.97        | 0.98           | 0.974    |
| **Xception**    |           |             |                |          |
| COVID-19        | 1.00      | 0.94        | 0.97           |          |
| Viral pneumonia | 0.89      | 0.94        | 0.92           |          |
| Normal          | 0.94      | 0.93        | 0.93           |          |
| Macro average   | 0.94      | 0.94        | 0.94           | 0.937    |
| **EfficientNetB0** |     |             |                |          |
| COVID-19        | 0.99      | 0.98        | 0.98           |          |
| Viral pneumonia | 0.98      | 0.91        | 0.94           |          |
| Normal          | 0.92      | 0.99        | 0.96           |          |
| Macro average   | 0.96      | 0.96        | 0.96           | 0.959    |
that the classifier works well for all classes. Based on these results, we consider the VGG16 to be our highest performing model. It exhibits the highest sensitivity, accuracy and F1-score for the COVID-19 class, while showing the highest F1-score for the viral pneumonia class. More specifically, our proposed VGG16 model achieves the highest sensitivity, precision, and F1-score of 0.99, 0.99 and 0.99 for COVID-19 classification, respectively. In addition, the model achieved the best overall F1-score of 0.97 in the normal and viral pneumonia class. On the other hand, Xception model provided relatively low sensitivity, precision, and F1-score values for all classes.

Table 4 shows the comparison of the detailed classification results of the proposed models in terms of different metrics. It can be seen that the proposed VGG16 model achieves the highest performance, with a precision of 0.979, a sensitivity of 0.978, an F1-score of 0.978, an accuracy of 0.978, and an MCC of 0.967. The ResNet50 network was found to be the second-best performing network in COVID-19 prediction, achieving 0.975 precision, 0.974 sensitivity, 0.975 F1-score, 0.974 accuracy, and 0.961 MCC. The proposed VGG16 models outperforms the other models in almost every performance metric, including accuracy, sensitivity and accuracy. After intensive experiments, considering the performance of different models, we chose VGG16 as our preferred model.

Figure 11 shows the comparison of the model in terms of different performance indicators such as accuracy, precision, sensitivity, F1-score, and Matthew Correlation Coefficient (MCC). It can be clearly seen from the figure that the proposed VGG16 model outperforms other models in terms of accuracy, sensitivity, F1-score and precision.

The confusion matrix is the main tool for evaluating errors in classification problems. It shows the number of images correctly and incorrectly recognized by the model. We constructed the confusion matrix of the six architectures proposed in the study to evaluate the performance, as shown in Figure 12. By observing the confusion matrix, the model performs well on the test set. The given model can be used to detect the presence of the virus in the human body in real time. From the matrix, we can conclude that out of 300 images, our proposed VGG16 model misclassifies only two COVID-19 images as viral pneumonia. The proposed model did not misclassify normal images as COVID-19, although nine normal

| Models     | Accuracy | Precision | Sensitivity | F1-score | MCC   |
|------------|----------|-----------|-------------|----------|-------|
| VGG16      | 0.978    | 0.979     | 0.978       | 0.978    | 0.967 |
| DenseNet201| 0.957    | 0.959     | 0.957       | 0.958    | 0.936 |
| MobileNetV2| 0.942    | 0.946     | 0.941       | 0.943    | 0.914 |
| ResNet50   | 0.974    | 0.975     | 0.974       | 0.975    | 0.961 |
| Xception   | 0.937    | 0.941     | 0.937       | 0.938    | 0.906 |
| EfficientNetB0 | 0.959 | 0.962 | 0.960 | 0.960 | 0.940 |

Figure 11. Accuracy, precision, F1-score, Matthew’s correlation coefficient (MCC) and sensitivity for the proposed models.
images were misclassified as viral pneumonia. Among the 337 images, only two images of viral pneumonia were misclassified as COVID-19, and eight images were misclassified as normal. It can be noticed that the model has no confusion between COVID-19 and the other two image classes. However, the high accuracy, sensitivity and $F_1$-score show that the proposed model does an excellent job of classifying most images. This shows that this deep learning technique is very reliable in distinguishing COVID-19 images from other X-ray images.

In addition to testing the classification performance of the models, comparisons in terms of computational cost were also performed. Table 5 shows the computational cost of the model in terms of total training and testing time, time per epoch (in seconds). The training phase of VGG16 takes 952 s and MobileNetV2 takes 808 s, which is the fastest models. The training phase of the Xception model takes 2389 s, which is the slowest model. It can be seen that the proposed model based on VGG16 is computationally less expensive than other CNN architectures. Furthermore, the results show that the proposed model can find the best accuracy in a reasonable time.

### 4.6.1 Performance comparison with and without data augmentation

Data augmentation is a popular technique for eliminating overfitting problems and improving the generalization of models. Therefore, to evaluate the strength and validity of the proposed model, we conducted experiments with and without data augmentation. Table 6 illustrates the
experiments related to the proposed CNN model based on the VGG-16 model, both with and without data enhancement. Our model achieved an accuracy of 0.978 with augmentation and 0.966 without augmentation. The proposed model records 0.978 sensitivity, 0.979 precision, 0.978 F1-score, and 0.967 MCC. The result of the experiment shows that incorporating data augmentation into the training of the proposed model performs significantly better and gives more consistent results.

4.6.2 | Result comparison with different optimization techniques

In this experimental setup, we chose various optimizers such as Adam, SGD, RMSProp and AdaDelta to evaluate the effectiveness of the proposed method. Initially, the Adam optimizer was used in the training phase, but we chose three other optimizers to further evaluate the strength of the proposed model. Table 7 shows the detailed results of the proposed model using different optimization algorithms. As shown in Table 7, the ADAM optimizer achieved the best classification performance than the other optimizers. Using the ADAM optimizer, the classification accuracy of the chest X-ray images was 0.978. Then, using SGD, RMSProp, and AdaDelta, the classification accuracy of the images was 0.951, 0.955 and 0.952, respectively. From these results, it can be seen that the proposed model adapts well to all four optimizers and exhibits fast learning.

4.6.3 | Result comparison with different batch sizes

In this section, different batch size experiments are performed to show the effect of batch size on the proposed model, since batch size is considered as one of the most critical hyper-parameter. During training, different batch sizes such as 16, 32, 64 and 128 were applied. Table 8 shows

**TABLE 5** Comparison of computation time.

| Models        | Training time (s) | Testing time (s) | Time per epoch |
|--------------|-------------------|------------------|----------------|
| VGG16        | 952.20            | 17.31            | 47–48          |
| DenseNet201  | 1381.47           | 22.31            | 69–71          |
| MobileNetV2  | 808.99            | 16.36            | 40–41          |
| ResNet50     | 1010.64           | 17.80            | 50–51          |
| Xception     | 2389.72           | 26.97            | 119–120        |
| EfficientNetB0 | 1169.24         | 16.64            | 58–59          |

**TABLE 6** Performance comparison with and without data augmentation.

| Technique       | Accuracy | Precision | Sensitivity | F1-score | MCC  |
|-----------------|----------|-----------|-------------|----------|------|
| With augmentation | 0.978    | 0.979     | 0.978       | 0.978    | 0.967|
| Without augmentation | 0.966    | 0.968     | 0.966       | 0.967    | 0.949|

**TABLE 7** Comparison of performance among different optimizers.

| Optimizers | Accuracy | Precision | Sensitivity | F1-score | MCC  |
|------------|----------|-----------|-------------|----------|------|
| Adam       | 0.978    | 0.979     | 0.978       | 0.978    | 0.967|
| SGD        | 0.951    | 0.953     | 0.952       | 0.952    | 0.927|
| RMSProp    | 0.955    | 0.958     | 0.956       | 0.956    | 0.933|
| AdaDelta   | 0.952    | 0.955     | 0.952       | 0.953    | 0.929|

**TABLE 8** Performance comparison of proposed model with different batch sizes.

| Batch size | Accuracy | Precision | Sensitivity | F1-score | MCC  |
|------------|----------|-----------|-------------|----------|------|
| 16         | 0.954    | 0.958     | 0.955       | 0.956    | 0.932|
| 32         | 0.970    | 0.971     | 0.970       | 0.970    | 0.955|
| 64         | 0.978    | 0.979     | 0.978       | 0.978    | 0.967|
| 128        | 0.961    | 0.964     | 0.961       | 0.962    | 0.943|

Note: Bold indicates the best values.
the performance of the proposed model in terms of accuracy, precision, sensitivity, F1-score and MCC when trained using different batch sizes such as 16, 32, 64 and 128. Table 8 shows that a higher and more consistent test accuracy of 0.978 is achieved with batch size 64, while 0.979, 0.978, 0.978 and 0.967 are achieved in terms of precision, sensitivity, F1-score, and MCC, respectively.

4.7 | Discussion on results

Currently, the use of machine learning, especially deep neural networks, has attracted great attention in COVID-19 diagnosis, and several researchers have used different approaches to develop reliable diagnostic systems. However, most of them evaluate their methods on limited datasets. The concept of transfer learning has been widely used because it offers several beneficial features such as high accuracy, speed and ease of application. In this study, we propose a model based on deep transfer learning to detect COVID-19 from chest X-ray images. In this work, a total of 3886 chest radiographs (1200 COVID-19, 1341 normal, and 1345 pneumonia) obtained from publicly available dataset have been used to train the CNN models. Data augmentation was performed to increase the training samples in order to generalize the models and prevent overfitting. We comprehensively evaluate the effectiveness of six most effective CNN models including VGG16, ResNet50, DenseNet201, MobileNetV2, Xception and EfficientNetB0 for COVID-19 prediction. The architecture of all CNN models is enhanced by using multiple dense layers and batch normalization layer to alleviate the problem of gradient vanishing, reduce the need for dropout, produce faster training, higher accuracy, and reduce overfitting. Extensive experiments were performed on a large number of images for automated COVID-19 screening. The experimental results and overall comparison of the models are shown in Tables 3 and 4. The results demonstrate the superiority of our proposed CNN model based on VGG16, achieving the highest overall performance in multi-class classification of chest X-ray images. It can be seen that the proposed model achieved the highest performance with a precision of 0.979, a sensitivity of 0.978, an F1-score of 0.978, an accuracy of 0.978, and an MCC of 0.967. In addition, the proposed model achieved the highest sensitivity of 0.99 for the COVID-19 class. A higher sensitivity value will minimize missed cases of COVID-19, which is the main goal of this study. The ResNet50 network was found to be the second-best performing network in COVID-19 prediction, achieving 0.975 precision, 0.974 sensitivity, 0.975 F1-score, 0.974 accuracy, and 0.961 MCC. The Xception model show the lowest performance on the dataset. As described in Table 5, the training time of our proposed model is 952 s, which indicates that our model has a shorter training time than the other models. We can also note from Table 5 that the overall time taken by our proposed method to achieve an accuracy of 0.978 is much less. Thus, this indicates that our proposed method is not computationally complex and is a better solution that helps to reduce the spread of COVID-19 infections. From the results, we can say that this proposed model is more reliable and can be deployed in the medical field to assist radiologists in identifying COVID-19 patients. Although the current study has limitations in terms of data availability, the data used in this study came from one or two sources. However, in the future, we plan to use images from different sources to validate our proposed model.

4.8 | Comparison with the state-of-the-art methods

A comparative analysis was carried out to prove the effectiveness of the proposed CNN model for COVID-19 classification. We compared the results of our proposed system with the recently proposed DL methods for the automatic diagnosis of COVID-19 using chest X-rays in Table 9. As shown in Table 9, there are some high-quality studies that have developed DL models to detect COVID-19. However, the main issue that should be noted is that most of them use a limited number of COVID-19 case data. Keeping all these in mind, CNN models have been developed for COVID-19 detection based on various pre-trained architectures. In this work, a total of 3886 chest X-rays obtained from a publicly available database were used to train and test the CNN models. Our proposed CNN model based on the pre-trained VGG16 architecture achieves the highest accuracy rate of 0.978 in the multi-class classification of COVID-19, normal and viral pneumonia. It can be seen that the proposed CNN model provides better performance than other existing methods.

5 | CONCLUSION AND FUTURE WORK

The COVID-19 pandemic clearly poses a threat to human survival. Early diagnosis of COVID-19 infection is essential to prevent the disease from spreading to others. Several researchers from all over the world are working hard to deal with the COVID-19 pandemic. During this emergency, it is important that no positive case goes unrecognized. Deep learning techniques have proven to be an indispensable tool for more accurate analysis of big data. This study demonstrates the usefulness of modern deep learning models using a transfer learning approach to automatically detect COVID-19 from chest X-ray images. This study utilizes open source chest X-ray images of normal, viral pneumonia and COVID-19 cases. The curated chest X-ray dataset consists of 1200 COVID-19 images, 1341 normal images, and 1345 viral pneumonia images. In this work, six popular and best-performing CNN-based deep learning models were tested to detect COVID-19, viral pneumonia, and normal patients from chest X-rays. Through extensive experiments and results performed on the collected datasets, the results showed that the proposed CNN model based on the
The VGG16 model outperformed all other models. The proposed CNN model built on VGG16 achieved 97.84% accuracy, 97.89% sensitivity, 97.90% precision, 97.89% F1-score, and 96.75% of MCC. Chest radiographs of COVID-19 and viral pneumonia contain similar features that are difficult for radiologists, so defining a COVID-19 detection problem can provide better control over model development. According to our research results, it is believed that due to its higher performance, this study has the potential to reduce the pressure on radiologists caused by the increasing number of COVID-19 patients to make decisions in clinical practice. In the future, we will use other deep learning techniques, such as GAN, and large number of X-ray images.

### TABLE 9  Accuracy comparison between the proposed CNN model and other existing state-of-the-art methods.

| Author                        | Method                  | Dataset                                              | Class | Accuracy  |
|-------------------------------|-------------------------|------------------------------------------------------|-------|-----------|
| Wang et al. (2020)            | COVID-Net               | 53 COVID-19                                          | 3     | 0.933     |
|                               |                         | 5526 Non-COVID-19 and 8066 Normal                    |       |           |
| Apostolopoulos & Mpesiana (2020) | VGG19                  | 224 COVID-19                                         | 3     | 0.934     |
| National Health Commission of People’s Republic of China (2020) (He, 2020) | ResNet50 + SVM         | 25 COVID-19 and 25 Normal                           | 3     | 0.953     |
| Loey et al. (2020)            | AlexNet                 | 79 COVID-19                                         | 3     | 0.851     |
| Ozturk et al. (2020)          | DarkNet                 | 127 COVID-19                                        | 3     | 0.870     |
| Keles et al. (2021)           | COV19-ResNet            | 210 COVID-19                                        | 3     | 0.976     |
| Oh et al. (2020)              | Patch-Based CNN         | 8851 Normal                                          | 3     | 0.889     |
| Elzeki et al. (2021)          | CXRVN                   | 221 COVID-19                                        | 3     | 0.930     |
| El Asnaoui & Chawki (2020)    | VGG-16                  | 2780 bacterial pneumonia                             | 3     | 0.921     |
|                               | Inception_Resnet_V2    | 231 Covid19 and 1583 normal                          | 3     | 0.748     |
| Khan et al. (2020)            | CoroNetXception         | 310 Normal and 327 Viral pneumonia and 284 COVID-19  | 3     | 0.896     |
| Jain et al. (2020)            | ResNet101               | 440 COVID-19 and 480 Viral pneumonia                 | 2     | 0.977     |
| Nigam, et al. (2021)          | VGG16, Xception, EfficientNet | 795 COVID-19 and 795 Normal                          | 3     | 0.790, 0.880, 0.934 |
| Alhudhaif et al. (2021)       | DenseNet201             | 268 COVID-19 and 850 Other Pneumonia                 | 2     | 0.949     |
| Gilanie et al. (2021)         | CNN                     | 7021 Pneumonia Images and 7021 Normal Images         | 3     | 0.966     |
| Fayemiwo et al. (2021)        | VGG16                   | 1300 Pneumonia                                      | 3     | 0.938     |
| Gaur et al. (2021)            | VGG16                   | 1345 Viral Pneumonia and 420 COVID-19 and 1341 Normal | 3     | 0.878     |
| Proposed method               | VGG16, ResNet50         | 1200 COVID-19 and 1341 Normal                        | 3     | 0.978     |

VGG16 model outperformed all other models. The proposed CNN model built on VGG16 achieved 97.84% accuracy, 97.89% sensitivity, 97.90% precision, 97.89% F1-score, and 96.75% of MCC. Chest radiographs of COVID-19 and viral pneumonia contain similar features that are difficult for radiologists, so defining a COVID-19 detection problem can provide better control over model development. According to our research results, it is believed that due to its higher performance, this study has the potential to reduce the pressure on radiologists caused by the increasing number of COVID-19 patients to make decisions in clinical practice. In the future, we will use other deep learning techniques, such as GAN, and large number of X-ray images.
ACKNOWLEDGEMENTS

This work was supported by the National Key R&D Program of China (2019YFA0706400, 2019YFA0706402), the Pre-Research Funds for Equipment of China (61409220115), the National Natural Science Foundation of China (61275179, 60970815), the Natural Science Basic Research Plan in Shaanxi Province of China (program no. 2016JM8047), the Fundamental Funds for the Central Universities (no. xjj2015112), Shaanxi Province Administration of Traditional Chinese Medicine (15-ZY036), and Xi’an Medical University State Fund Cultivation Project (2016GJF01).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request.

ORCID

Sohaib Asif https://orcid.org/0000-0003-0707-470X
Yi Wenhui https://orcid.org/0000-0002-4044-0444

REFERENCES

Ahuja, S., Panigrahi, B. K., Dey, N., Rajnikanth, V., & Gandhi, T. K. (2021). Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices. Applied Intelligence, 51(1), 571–585.
Alhudaif, A., Polat, K., & Karaman, O. (2021). Determination of COVID-19 pneumonia based on generalized convolutional neural network model from chest X-ray images. Expert Systems with Applications, 180, 115141.
Al-Zubi, S., Al-Ayyoub, M., Jararweh, Y., & Shehab, M. A. (2017). Enhanced 3D segmentation techniques for reconstructed 3D medical volumes: Robust and accurate intelligent system. Procedia Computer Science, 113, 531–538.
Al-Zubi, S., Al-Qatawneh, S., & Alsmirat, M. (2018). Transferable HMM trained matrices for accelerating statistical segmentation time. Paper presented at the 2018 fifth international conference on social networks analysis, management and security (SNAMS).
Al-Zubi, S., Aqel, D., & Mughaid, A. (2021). Recent intelligent approaches for managing and optimizing smart blood donation process. Paper presented at the 2021 international conference on information technology (ICIT).
Al-Zubi, S., Hawashin, B., Mughaid, A., & Baker, T. (2021). Efficient 3D medical image segmentation algorithm over a secured multimedia network. Multimedia Tools and Applications, 80(11), 16887–16905.
Al-Zubi, S., Jararweh, Y., Al-Zoubi, H., Elbes, M., Kanan, T., & Gupta, B. (2019). Multi-orientation geometric medical volumes segmentation using 3D multi-resolution analysis. Multimedia Tools and Applications, 78(17), 24223–24248.
Al-Zubi, S., Makki, Q. H., Ghani, Y. A., & Ali, H. (2021). Intelligent distribution for COVID-19 vaccine based on economical impacts. Paper presented at the 2021 international conference on information technology (ICIT).
Al-Zubi, S., Mughaid, A., Hawashin, B., Elbes, M., Kanan, T., Alrawashdeh, T., & Aqel, D. (2019). Reconstructing big data acquired from radioisotope distribution in medical scanner detectors. Paper presented at the 2019 IEEE Jordan international joint conference on electrical engineering and information technology (JEET).
Al-Zubi, S., Shehab, M., Al-Ayyoub, M., Jararweh, Y., & Gupta, B. (2020). Parallel implementation for 3d medical volume fuzzy segmentation. Pattern Recognition Letters, 130, 312–318.
Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. Physical and Engineering Sciences in Medicine, 43(2), 635–640.
Asif, S., Wenhui, Y., Jin, H., & Jinhai, S. (2020). Classification of COVID-19 from chest X-ray images using deep convolutional neural network. Paper presented at the 2020 IEEE 6th international conference on computer and communications (ICCC).
Chowdhury, M. E., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M. A., Mahbub, Z. B., Islam, K. R., Khan, M. S., Iqbal, A., Al Emadi, N., Reaz, M. B. I., & Islam, M. T. (2020). Can AI help in screening viral and COVID-19 pneumonia? IEEE Access, 8, 123655–123676.
Das, N. N., Kumar, N., Kaur, M., Kumar, V., & Singh, D. (2020). Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays. IBM, 43, 114–119.
El Asnaoui, K., & Chawki, Y. (2020). Using X-ray images and deep learning for automated detection of coronavirus disease. Journal of Biomolecular Structure and Dynamics, 39(10), 3615–3626. https://doi.org/10.1080/07391102.2020.1767212
Elzaki, O. M., Shams, M., Sarhan, S., Abd Elfattah, M., & Hassanien, A. E. (2021). COVID-19: A new deep learning computer-aided model for classification. PeerJ Computer Science, 7, e358.
Farooq, M., & Hafeez, A. (2020). Covid-resnet: A deep learning framework for screening of Covid-19 from radiographs. arXiv preprint, arXiv:2003.14395.
Fayemiwo, M. A., Olowookere, T. A., Arekete, S. A., Ogunde, A. O., Odim, M. O., Oguntunde, B. O., Olaniyi, O. O., Ojewumi, T. O., Oyetade, I. S., Aremu, A. A., & Kayode, A. A. (2021). Modeling a deep transfer learning framework for the classification of COVID-19 radiology dataset. PeerJ Computer Science, 7, e614.
François, C. (2017). Xception: Deep learning with depthwise separable convolutions. Paper presented at the 2017 IEEE conference on computer vision and pattern recognition (CVPR).
Gaur, L., Bhatia, U., Jhanji, N., Muhammad, G., & Masud, M. (2021). Medical image-based detection of COVID-19 using deep convolution neural networks. Multimedia Systems, 11, 1–10.
Gilanie, G., Bajwa, U. I., Warih, M. M., Ashgar, M., Kousar, R., Kashif, A., Aslam, R. S., Qasim, M. M., & Rafique, H. (2021). Coronavirus (COVID-19) detection from chest radiology images using convolutional neural networks. Biomedical Signal Processing and Control, 66, 102490.
Gupta, A., Gupta, S., & Katarya, R. (2021). InstaCovNet-19: A deep learning classification model for the detection of COVID-19 patients using chest X-ray. Applied Soft Computing, 99, 106859.
Hammoudi, K., Benhabiles, H., Melkemi, M., Dornaika, F., Arganda-Carreras, I., Collard, D., & Scherpereel, A. (2021). Deep learning on chest X-ray images to detect and evaluate pneumonia cases at the era of covid-19. *Journal of Medical Systems*, 45(7), 1–10.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Paper presented at the proceedings of the IEEE conference on computer vision and pattern recognition.

He, Y. (2020). Translation: Diagnosis and treatment protocol for novel coronavirus pneumonia (trial version 7). *Infectious Microbes & Diseases*, 2(2), 48–54.

Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. Paper presented at the proceedings of the IEEE conference on computer vision and pattern recognition.

Jain, G., Mittal, D., Thakur, D., & Mittal, M. K. (2020). A deep learning approach to detect Covid-19 coronavirus with X-ray images. *Biocybernetics and Biomedical Engineering*, 40(4), 1391–1405.

Joshi, A., Dey, N., & Santosh, K. (2020). Intelligent systems and methods to combat covid-19. Springer.

Keles, A., Keles, M. B., & Keles, A. (2021). COVID19-CNNet and COVID19-ResNet: Diagnostic inference engines for early detection of COVID-19. *Cognitive Computation*, 13, 1–11.

Khan, A. I., Shah, J. L., & Bhat, M. M. (2020). CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. *Computer Methods and Programs in Biomedicine*, 196, 105581.

Lancet, T. (2020). COVID-19: Too little, too late? Lancet, 395(10226), 755.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436–444.

Li, D., Wang, D., Dong, J., Wang, N., Huang, H., Xu, H., & Xia, C. (2020). False-negative results of real-time reverse-transcriptase polymerase chain reaction for severe acute respiratory syndrome coronavirus 2: Role of deep-learning-based CT diagnosis and insights from two cases. *Korean Journal of Radiology*, 21(4), 505–508.

Li, L., Qin, L., Xu, Z., Yin, Y., Wang, X., Kong, B., Bai, J., Lu, Y., Fang, Z., Song, Q., Cao, K., Liu, D., Wang, G., Xu, Q., Fang, X., Zhang, S., Xia, J., & Xia, J. (2020). Artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT. *Radiology*, 296, E65–E71.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.

Liu, B., Yan, B., Zhou, Y., Yang, Y., & Zhang, Y. (2020). Experiments of federated learning for covid-19 chest X-ray images. arXiv preprint, arXiv:2007.05592.

Loey, M., Smarandache, F., & Khalifa, N. E. M. (2020). Within the lack of chest COVID-19 X-ray dataset: A novel detection model based on GAN and deep transfer learning. *Symmetry*, 12(4), 651.

Luz, E., Silva, P., Silva, R., Silva, L., Guimarães, J., Miozzo, G., Moreira, G., & Menotti, D. (2021). Towards an effective and efficient deep learning model for COVID-19 patterns detection in X-ray images. *Research on Biomedical Engineering*, 38(1), 149–162.

Makis, A., Kontopoulos, I., & Tsperes, K. (2020). COVID-19 detection from chest X-ray images using deep learning and convolutional neural networks. Paper presented at the 11th Hellenic conference on artificial intelligence.

Narin, A., Kaya, C., & Pamuk, Z. (2021). Automatic detection of coronavirus disease (Covid-19) using X-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, 24, 1–14.

Nayak, S. R., Nayak, D. R., Sinha, U., Arora, V., & Pachori, R. B. (2021). Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: A comprehensive study. *Biomedical Signal Processing and Control*, 64, 102365.

Nigam, B., Nigam, A., Jain, R., Dodia, S., Arora, N., & Annappa, B. (2021). COVID-19: Automatic detection from X-ray images by utilizing deep learning methods. *Expert Systems with Applications*, 176, 114883.

Oh, Y., Park, S., & Ye, J. C. (2020). Deep learning covid-19 features on cxr using limited training data sets. *IEEE Transactions on Medical Imaging*, 39(8), 2688–2700.

Ouchicha, A., Ammor, O., & Meknassi, M. (2020). CVDNet: A novel deep learning architecture for detection of coronavirus (Covid-19) from chest X-ray images. *Chaos, Solitons & Fractals*, 140, 110245.

Ozturk, T., Talo, M., Yildirim, E. A., Baloglu, U. B., Yildirim, O., & Acarya, A. R. (2020). Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in Biology and Medicine*, 121, 103792.

Razai, M. S., Doerholt, K., Ladhani, S., & Oakeshott, P. (2020). Coronavirus disease 2019 (covid-19): A guide for UKGPs. Lancet, 393(10126), 196–205.

Razai, M. S., Doerholt, K., Ladhani, S., & Oakeshott, P. (2020). Coronavirus disease 2019 (covid-19): A guide for UKGPs. *Lancet*, 393(10126), 755.

Rothe, C., Schunk, M., Sothmann, P., Bretzel, G., Wallrauch, C., Zimmer, T., Thiel, V., Janke, C., Guggemos, W., Selmaier, M., Drosten, C., Vollmar, P., Zvirginmaier, K., Zange, S., Wölfel, R., & Hoelscher, M. (2020). Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. *New England Journal of Medicine*, 382(10), 970–971.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Sethy, P. K., & Bhera, S. K. (2020). Detection of coronavirus disease (covid-19) based on deep features. Preprints. https://doi.org/10.20944/preprints202003.0300.v1

Shelke, A., Inanmard, M., Shah, V., Tiwari, A., Hussain, A., Chaferak, T., & Mehendale, N. (2021). Chest X-ray classification using deep learning for automated COVID-19 screening. *SN Computer Science*, 2(4), 1–9.

Shin, H.-C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., Yao, J., Mollura, D., & Summers, R. M. (2016). Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Transactions on Medical Imaging*, 35(5), 1285–1298.

Simonoyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint, arXiv:1409.1556.

Singh, M., Bansal, S., Ahuja, S., Dubey, R. K., Panigrahi, B. K., & Dey, N. (2021). Transfer learning-based ensemble support vector machine model for automated COVID-19 detection using lung computed tomography scan data. *Medical and Biological Engineering and Computing*, 59(4), 825–839.

Tan, M., & Le, Q. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. Paper presented at the international conference on machine learning.

Verma, D., Bose, C., Tufchi, N., Pant, K., Tripathi, V., & Thapliyal, A. (2020). An efficient framework for identification of tuberculosis and pneumonia in chest X-ray images using neural network. *Procedia Computer Science*, 171, 217–224.

Wang, L., Lin, Z. Q., & Wong, A. (2020). Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest X-ray images. *Scientific Reports*, 10(1), 1–12.
How to cite this article: Asif, S., Wenhui, Y., Amjad, K., Jin, H., Tao, Y., & Jinhai, S. (2023). Detection of COVID-19 from chest X-ray images: Boosting the performance with convolutional neural network and transfer learning. Expert Systems, 40(1), e13099. https://doi.org/10.1111/exsy.13099