

Article

A Deep-Learning-Based Framework for Automated Diagnosis of COVID-19 Using X-ray Images

Irfan Ullah Khan and Nida Aslam *

College of Computer Science and Information Technology, Imam Abdulrahman bin Faisal University, Dammam 1982, Saudi Arabia; iurab@iau.edu.sa

* Correspondence: naslam@iau.edu.sa

Received: 6 August 2020; Accepted: 25 August 2020; Published: 29 August 2020



Abstract: The emergence and outbreak of the novel coronavirus (COVID-19) had a devastating effect on global health, the economy, and individuals' daily lives. Timely diagnosis of COVID-19 is a crucial task, as it reduces the risk of pandemic spread, and early treatment will save patients' life. Due to the time-consuming, complex nature, and high false-negative rate of the gold-standard RT-PCR test used for the diagnosis of COVID-19, the need for an additional diagnosis method has increased. Studies have proved the significance of X-ray images for the diagnosis of COVID-19. The dissemination of deep-learning techniques on X-ray images can automate the diagnosis process and serve as an assistive tool for radiologists. In this study, we used four deep-learning models—DenseNet121, ResNet50, VGG16, and VGG19—using the transfer-learning concept for the diagnosis of X-ray images as COVID-19 or normal. In the proposed study, VGG16 and VGG19 outperformed the other two deep-learning models. The study achieved an overall classification accuracy of 99.3%.

Keywords: deep-learning; transfer-learning; COVID-19; coronavirus; pandemic

1. Introduction

A novel coronavirus (2019-nCoV or COVID-19 or SARS-CoV-2) case was reported in China in December 2019. The spread of the virus is increasing at an exponential rate globally and has become a pandemic. This COVID-19 outbreak has had an instigated a distressing impact on individuals' lives and community health and has led to an economic crisis. The COVID-19 pandemic has led to 7,981,067 cases, with 435,141 deaths, and is affecting 213 countries worldwide [1]. Some of the common symptoms associated with the virus are fever, cough, sore throat, shortness of breath, etc. The source of transmission is via human-to-human interaction and respiratory droplets [2]. Due to the highly contagious nature of the COVID-19 virus, early detection and control is crucial to control the outbreak. The long incubation period of COVID-19 has posed a substantial challenge regarding the control of this pandemic. Furthermore, the asymptomatic nature of COVID-19 in some patients is another reason for the outbreak. Owing to these reasons, early detection and control of the spread of COVID-19 is very hard.

Due to the above-mentioned challenges and the increase in the number of cases, efforts have been made to explore an effective method for accurate and easy diagnosis of COVID-19. Among the gold-standard laboratory methods used for the diagnosis of COVID-19 is real-time Reverse Transcription Polymerase Chain Reaction (RT-PCR) using an oral or nasopharyngeal swab [3]. The RT-PCR test requires a special kit to perform the diagnosis. It is a complex and time-consuming text. Furthermore, RT-PCR suffers from a high false-negative rate and results in poor sensitivity [4]. A high false-negative rate increases the prevalence of the virus, as some of the patients who carry the virus are diagnosed as negative, may interact with a substantial amount of people, and therefore, transmit the virus. These individuals will, in turn, be very precarious and can potentially increase the spread of the

pandemic. Consequently, there is a need for quintessential diagnosis methods with a reduced false-negative rate to control the magnitude of the risk associated with it.

However, radiologic imaging, such as computed tomography (CT) images and X-rays, is also widely used for the scanning and diagnosis of several diseases. Chest CT-scan and chest X-rays are also used for the diagnosis of COVID-19 [5]. PCR directly measures the amount of virus present and has specific markers for different viruses. PCR has very high specificity, it directly measures the viral load, and it can give an idea of the prognosis of the disease. The underlying pathology can sometimes hide viral disease, as in the case of a patient with pulmonary fibrosis. COVID-19 will be very hard to diagnose in these patients. Nonetheless, X-rays sometimes have false-negatives. They can only detect the complications of a disease, as in the case of pneumonia or acute respiratory distress syndrome (ARDS). However, those findings could be due to any other virus or bacteria.

Xie et al. [6] reported that in some patients who were diagnosed as COVID-19 negative in RT-PCR at the early stage, CT-scans contain abnormalities and identification of the virus. Subsequently, another study performed by Fang et al. [7] also compared the sensitivity of RT-PCR and CT-scan for COVID-19 patients. The study achieved a similar finding that the sensitivity of the chest CT is higher when compared with the RT-PCR. The study found that abnormalities in the CT-scan and X-rays can be seen even at a very early stage of COVID-19, before the presence of symptoms [8]. Furthermore, a subsequent discovery was made by Ai et al. [9], who showed that CT-scans have higher sensitivity when compared with the RT-PCR, and recommended a CT-scan for the diagnosis of COVID-19. To address this deficit, CT-scan and X-rays are used with the RT-PCR test for the diagnosis and monitoring of COVID-19 patients. Considerable effort and studies have been performed recently using a CT-scan for the diagnosis of COVID-19. Regardless of this advantage, CT-scans are not widely used, due to the high cost and unavailability of CT-scan devices in clinics and test centers. Due to the prevalence and accessibility of X-ray machines in developing countries widely used for the diagnosis of several other diseases, chest X-rays can be performed easily for the diagnosis of COVID-19. Studies have been performed to identify abnormalities or identification of COVID-19 using chest X-rays [2,10].

Due to the rapid advancement and integration of artificial intelligence (AI) in every domain of life, computer aided diagnosis can help medical professionals or radiologist in the diagnosis or analysis of radiologic scans. Several studies have been performed to combat the COVID-19 outbreak using computer-assisted technology [11–13]. Deep-learning has been widely used for the diagnosis, prognosis, and forecasting of various diseases including automated diagnosis of brain tumor, anesthesiology, ophthalmology, spine, pulmonary nodules, medical image segmentation, skin cancer, breast cancer, and many more [14–22] etc. Deep-Learning or Deep Convolutional Neural Network (DCNN) is a type of Convolutional Neural Network (CNN) that uses an automated method of feature extraction, image segmentation, and image classification. It can be used for supervised, unsupervised, and semisupervised learning [23].

Deep-learning has been used for the thoracic radiologic images, such as for lung segmentation and reconstruction [24,25], diagnosis of pneumonia [26], and tuberculosis diagnosis [27]. Research has proven effective and accurate diagnosis of the respiratory diseases using deep-learning and chest X-rays [28]. COVID-19 is a form of pneumonia that causes inflammation in the lungs. Due to a huge number of COVID-19 patients worldwide, and the limited number of radiologists in hospitals, there is a need for automated artificial intelligence-based systems for accurate diagnosis. Although radiologists play a significant role due to their enormous knowledge of the analysis of CT-scans and X-rays, the implication of deep-learning technologies in radiology can be helpful to acquire precise diagnosis. These automated systems will provide timely assistance to radiologists in diagnosis.

To address the above-mentioned challenges and benefits, we endeavor to propose a deep-learning-based model for the diagnosis of COVID-19 using chest X-rays. The proposed model will effectively diagnose COVID-19 without using feature extraction. The model will serve as an assistive tool for radiologists in accurate diagnosis.

The study is organized as follows; Section 2 covers the literature review. Section 3 discusses the data set description; Section 4 covers the methodology along with the evaluation parameters used. Section 5 presents the experimental results, while Section 6 contains the comparison between the proposed study and the previous studies. Section 7 briefs a conclusion of the study.

2. Literature Review

Subsequently, manifesting the significance of X-ray images and deep-learning in COVID-19 diagnosis, several noticeable studies will be discussed in this section.

Recently, a study by Hamdan et al. [29] presented COVIDX-Net, containing a comparative analysis of seven deep-learning models for the diagnosis of COVID-19. The models used were VGG19, DenseNet201, ResNetV2, InceptionV3, InceptionResNetV2, Xception, and MobileNetV2 using a binary data set that consisted of 50 X-ray samples (25 healthy and 25 COVID-19). Experiments were performed using the X-ray images from two data sets: the COVID-19 X-ray image database [30], which consists of 123 frontal view X-rays, and the Adrain Rosebrock data set [31]. The study achieved the highest accuracy of 90% with VGG19 and DenseNet201. Nevertheless, the study suffers from the limitation of a small data set.

Likewise, VGG19 and ResNet50 models were used and compared with the proposed COVID-Net model in the study performed by Wang et al. [32] for COVID-19 diagnosis using a pretrained ImageNet model and Adam optimizer using a multiclass data set (normal, pneumonia, and COVID-19). They achieved a better accuracy rate of 93.3% when compared to the study mentioned earlier. There were 13,975 X-ray images that were taken from multiple open-source data sets [30,33–36]. However, to address the data imbalance issue a data augmentation technique was used. Similar to Wang et al., another study performed by Apostolopoulos et al. [37] also used VGG19. The data set used in the study was collected from four open-source data sets [30,34], Radiopaedia [38], and the Italian Society of Medical and Interventional Radiology (SIRM) [39]; the total number of X-ray images was 1427 (224 COVID-19, 700 pneumonia, and 504 normal). Experiments were performed for binary and multiclass categories. The highest accuracy achieved for the binary class was 98.75% and for multiclass, the highest accuracy was 93.48%. Like the two previously mentioned studies [29,32], VGG19 outperformed the other models in terms of accuracy.

Furthermore, Kumar et al. [40] used ResNet50 and Support Vector Machine (SVM) classification for the diagnosis of COVID-19 using X-ray images. The chest X-ray images used in the study were taken from two COVID-19 X-ray image databases: Cohen [30] and Kaggle Chest X-ray Images (Pneumonia) [41]; though, the experiments used 50 chest X-rays (25 COVID-19, 25 non-COVID). The study outperformed—with an accuracy of 95.38% for the binary class—and produced better results when compared to the study performed by Hamdan et al. [29]. Similarly, the study performed by Ali et al. [42] also found ResNet50 with the highest accuracy of 98% using some X-ray images from the data sets [30,41]. Nevertheless, the study achieved a very high outcome. However, the data set used in the study was small.

In addition, Ozturk et al. [43] proposed a DarkNet, a deep neural network model for COVID-19 diagnosis using X-ray images from two data sets, namely, the COVID-19 X-ray image database [30] and ChestX-ray8 database [44] for both binary (125 COVID, 500 no-findings) and multiclass (124 COVID, 500 pneumonia, and 500 no-findings) classification. The model contains a 17-layer convolutional network model, a leaky ReLu activation function, and a “You Only Look Once” (YOLO) object detection model. For multiclass, the accuracies of 82.02% and 98.08% for the binary class were achieved. DarkNet yielded a better classification performance when compared to the former study [40]. Despite all these benefits, the study suffers from the limitation of a small number of X-ray images of COVID-19 patients.

Consequently, due to the effectiveness of the ResNet and DenseNet models in the previous studies, Minaee et al. [45] proposed a Deep-COVID model using ResNet18, ResNet50, SqueezeNet, and DenseNet121. ResNet was trained using an ImageNet pretrained model. The study created a data

set of 5000 chest X-ray images (COVID X-ray 5k) using two open-source data sets [30,46]. The study achieved 97.5% sensitivity and 90% specificity for the binary classification. However, the number of samples for COVID-19 was 100, while for the non-COVID category was 5000 samples; this imitates a huge data imbalance.

To address the small size open-source data set for COVID-19, Afshar et al. [47] proposed a model COVID-CAPS, using a capsule network containing four CNN and three capsule layers. The study was performed using two open-source data sets [30,41], and achieved the highest accuracy of 95.7% and 95.8% specificity and a sensitivity of 90%. However, the study suffers from a huge data imbalance. Furthermore, Ucar et al. [48] proposed COVIDiagnosis-Net, a deep-learning-based model using SqueezeNet and Bayesian optimization techniques on a COVIDxNet data set [29]. The study produced an accuracy of 98.3% for multiclass classification. Likewise, the same data set was used in another model—COVID-ResNet by Farooq et al. [49]. They used pretrained ResNet30 with the aim to reduce training time and achieved an overall accuracy of 96.23% for multiclass. This study achieved lower accuracy when compared to the former study but covered the bacterial infection. Correspondingly, pretrained ResNet18 was used in another study performed by Oh et al. [50] and achieved an accuracy of 88.9%. Several data sets were used such as Japanese Society of Radiological Technology (JSTR) [51,52], U.S. National Library of Medicine (USNLM) collected Montgomery Country (MC) (NLM/MC) [53], CoronaHack [54], and the COVID-19 X-ray image database [30].

Undoubtedly, the significance of the chest X-ray for the diagnosis of COVID-19 and the implication of a deep convolutional model for the automated analysis of X-ray [28] motivated the need for further exploration. Despite these advantages, there is a conundrum to find an open-source data set containing a large number of COVID-19 X-ray images. Most of the previous studies suffer from a small data set or data imbalance. To avoid these drawbacks, we used the data set which is the combination of a number of open-source data sets. Several studies have already been performed, but indeed, there is still a need for further exploration.

3. Data Set Description

The X-ray images used were taken from four open-source chest X-ray image data sets, with a total number of X-rays of 1683. The details of the images taken from each data set are mentioned below.

- a. COVID-19 X-ray image database collected by Cohen et al. [30] consists of a total of 660 images; some of the images in the data set were CT-scan, and some were nonfrontal chest X-rays. CT-scan and nonfrontal X-rays of non-COVID-19 patients X-rays were removed. Moreover, the images tagged with pneumonia were also removed from the data set. The selected frontal chest X-ray of positive COVID-19 patients from Cohen's data set was 390 X-rays.
- b. Furthermore, 25 X-ray images of COVID-19 patients were selected from the COVID-19 chest X-ray data initiative [33] data set. The original data set consisted of 55 X-rays. Some of the images in the data set were not clear and were not considered in our experiments.
- c. Additionally, 180 X-ray images of COVID-19 were also selected from the Actualmed COVID-19 chest X-ray data initiative [35]. Originally, the data set consisted of 237 scans.
- d. Finally, the X-ray images of both the normal and COVID-19 categories were selected from the COVID-19 radiography database [36]. The data set contained 1057 X-ray images (219 COVID-19, 1341 normal, and 1345 viral pneumonia). In our study, we selected 195 X-rays for COVID-19 and 862 images for the normal category.

The total number of images per category used for training, testing, and validation is shown in the Table 1, while Figure 1 indicates the number of images per category (normal and COVID-19). The distribution of the data set used in the study was stratified in order to alleviate the data imbalance issue. Figure 2 contains the sample of COVID-19 and normal X-ray in the data set.

Table 1. Number of images per category used in training, validation, and testing phases without data augmentation.

Split	COVID-19	Normal
Training	630	642
Validation	60	60
Testing	100	100

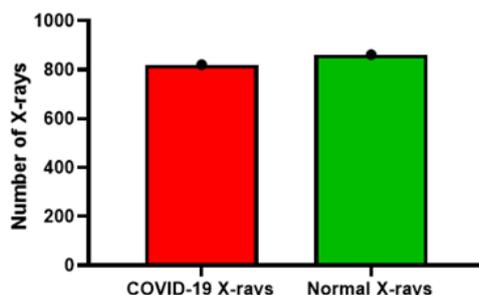


Figure 1. Number of X-ray images per category in the data set.

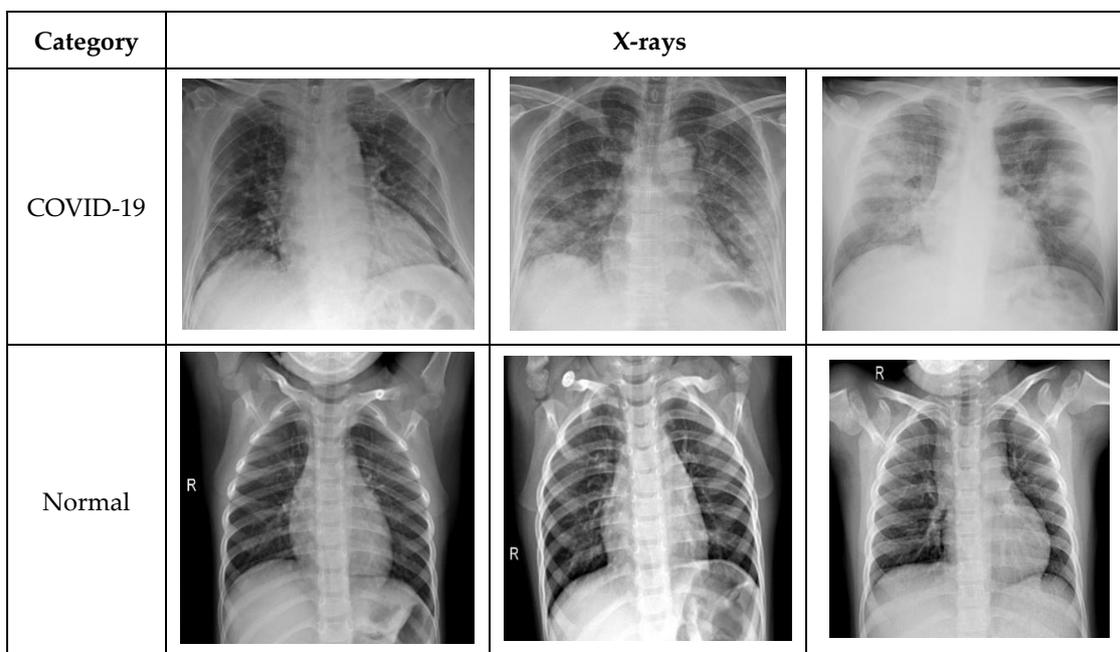


Figure 2. Chest X-ray sample images from the data set used in the study.

4. Methodology

The model consists of two steps: preprocessing and data augmentation and transfer-learning using pretrained deep-learning models: ResNet50, VGG16, VGG19, and DenseNet121. This study classified the chest X-ray images into binary classes as normal and COVID-19. The description of the stages is discussed below.

4.1. Data PreProcessing and Augmentation

During this stage, a data augmentation technique was applied, with the aim to alleviate the problem of model overfitting. Due to the in-depth nature of the pretrained model, there is a high risk of overfitting if the size of the data set is small. To circumvent this drawback, additional images were generated using data augmentation. The data augmentation method increases the generalization of the data, specifically for the X-ray data sets [55,56]. Augmentation was applied using three steps—resizing,

flipping, and rotation. For resizing the images, a dimension of $224 \times 224 \times 3$ was used. Moreover, random horizontal flip was used to increase the generalization of the model for all possible locations of the COVID-19 symptoms in the X-rays. Finally, some images were generated by applying the rotation of 15 degrees. The applied augmentation techniques endeavored to boost the proposed model generalization. The data augmentations were only applied on the X-ray training data set.

4.2. Deep Neural Networks and Transfer-Learning

The emergence of the Convolution Neural Networks (CNN) or deep-learning has enhanced the image classification task. Training the deep neural network model requires a large training data set. The performance of a deep-learning model highly depends on the number of images used to train the model because the model has the innate ability to extract the features (temporal and spatial) using filters. However, deep-learning can also be employed in the domain where the size of the data set is not huge by using the concept of transfer-learning. In transfer-learning, the feature extracted from the specified data using a CNN model is transferred to solve related tasks, including new data (small data set), where building the CNN from scratch is unsuitable [57]. Among the widely used methods for transfer-learning in a medical domain is training a model with a huge data set i.e., ImageNet [58], a pretrained model for object detection and classification. The selection of the deep-learning model for the transfer-learning depends on the ability of the model to extract the features related to the domain.

Transfer-Learning is implemented using two steps such as feature extraction and parameter tuning (optimization strategy). During the feature extraction, the pretrained model holds the new features from the data set using the training data. Secondly, to optimize the performance of a model in the current applied domain, the model architecture needs to be reconstructed and updated along with the parameter tuning. Using this pretrained model alleviates the drawback of a small data set and reduces the computational cost.

The pretrained models used in this study are discussed below.

Pretrained models such as DenseNet121 [59], ResNet50 [60], VGG16, and VGG19 [61] were used. These models were pretrained using the ImageNet data set and were further trained using the X-ray data set.

- a. **DenseNet121:** The dense convolutional neural network (DenseNet) is a feed forward fully connected neural network. Each layer in DenseNet consists of a feature map. The feature map of each layer serves as an input to the next layer. Among the advantages of the DenseNet is that it requires less parameters. The number of filters or feature maps used in DenseNet is 12. Traditional convolutional neural networks consisting of L layers contain L number of connections, while in DenseNet, the number of direct connections is $\frac{L(L+1)}{2}$ [60]. The dense connectivity of the model circumvents the need for redundant learning. In addition to this, DenseNet decreases the chance of model overfitting due to the small size training data set by applying regularization.
- b. **ResNet50:** ResNet, also known as the deep residual network, was initially proposed in 2015 with the motivation of a “identity shortcut connection”. It is also among the pretrained models using ImageNet. ResNet skips one or more layers and handles the gradient vanishing issue. Among the key advantages of ResNet is easier optimization. Moreover, the accuracy of the model can be enhanced with the increase in the depth of the model [60]. ResNet model skips one, two, or more layers and is directly connected to any layer, not necessarily the adjacent layer, using a ReLu nonlinear activation function. ResNet uses the forward and backward propagation method.
- c. **VGG:** VGG, also known as a very deep convolutional network, was first introduced in 2014. VGG is an advanced version of AlexNet with an increased number of layers. The increase in the number of layers increases the generalization of the model [61]. The benefit of VGG is the use of only 3×3 convolutional filters. The only difference between VGG16 and VGG19 is the number

of layers. However, the convolutional neural network is used for analyzing the object of the image. We used both models for COVID-19 X-ray.

Finally, for training the network, the input image size $224 \times 224 \times 3$ was used, and the initial learning rates (1×10^{-3}) for all models were kept fixed. The number of Epochs was kept at 30 for each model, and the ReLU activation function was used in order to make the feature extraction range of neurons more extensive.

4.3. Model Evaluation

For evaluating the performance of the proposed module, several standard evaluation parameters were used. In the current study, we compared the performance of all the developed models in terms of Accuracy (ACC), Sensitivity (SEN), Specificity (SPE), Positive Predicted Value (PPV), F1 Score (F1), False Negative Rate (FNR), and False-Positive Rate (FPR) [62]. Accuracy (ACC) represents the total number of X-ray samples classified correctly as COVID-19 or healthy divided by the total number of X-rays of the data set. Sensitivity is among the widely used measures in the health domain. Sensitivity indicates the true positive rate of the proposed technique. As mentioned in the first section, the sensitivity of the RT-PCR test was low; so, in the current study, sensitivity is among the key evaluation measures. It represents the ratio of the number of patients predicted COVID-19 using the model and the number of overall COVID-19 cases in the data set. Specificity, also known as the true negative rate, represents the ratio of the number of patients that were predicted normal and all patients that were normal in the data set.

Moreover, the likelihood ratio of the proposed models was evaluated in terms of positive predicted value (PPV), also known as the precision. It indicates if the X-ray image is predicted as COVID-19 and how well it represents the actual presence of COVID-19 disease.

In addition, other measures used were F1, FNR, and FPR. F1 is a combination of predicted positive value (precision) and sensitivity (true positive rate), also known as recall. However, false negative and positive rate measure the number of X-ray samples predicted mistakenly as COVID-19 positive or healthy. The smaller the value of false negative and false positive, the higher the significance of the proposed model.

5. Experimental Results

Experiments were conducted using the four trained models—DenseNet121, ResNet50, VGG16, and VGG19. The below figures (Figures 3–6) show the training accuracy, training loss, testing accuracy, and testing loss of the implemented models with the number of epochs. Figure 3 represents the number of epochs and accuracy for all the implemented deep-learning techniques for training the model. However, Figure 4 represents the training loss. These Figures represent the significance of VGG19 and VGG16 in the training phase. VGG19 had the highest accuracy and reduced loss. Similarly, testing the accuracy curve (see Figure 5) and loss curve (see Figure 6) showed the effectiveness of both VGG16 and VGG19 deep-learning models. However, the performance of DenseNet121 had a much lower stability and produced the worst outcome when compared with the other applied models.

In our study, a total of four models were developed, and the performance of each model was evaluated in terms of the measures discussed in the previous section. The comparative analysis of all four models is shown in the Table 2. Our results demonstrate that VGG16 and VGG19 outperformed the other two models. The highest accuracy achieved in the study was 99.33%. Both VGG16 and VGG19 achieved similar accuracy. However, the sensitivity of VGG19 was higher than VGG16. The main motivation of using X-ray images for the early diagnosis of COVID-19 was alleviating the impediment of RT-PCR i.e., lower sensitivity. RT-PCR laboratory tests have a higher false positive rate. The main challenge in controlling the outbreak of the pandemic is that the gold-standard test used for the diagnosis of the COVID-19 has a lower sensitivity. Higher sensitivity indicates the only a few COVID-19 patients will be undetected which will ultimately reduce the spread of the disease.

VGG-19 achieved the highest sensitivity of 100%. Similarly, the specificity indicates the true negative rate. VGG-16 achieved the highest specificity of 99.38%.

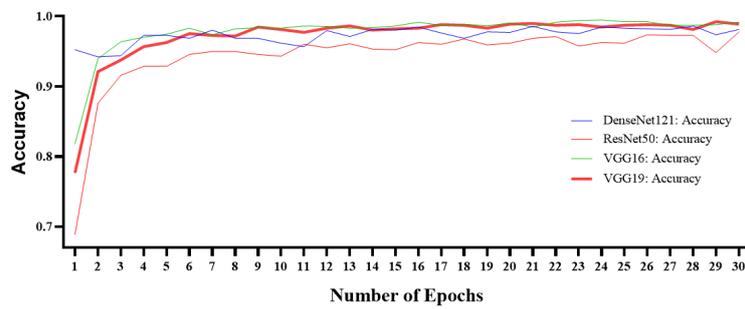


Figure 3. Training accuracy curve for DenseNet121, ResNet50, VGG16, and VGG19.

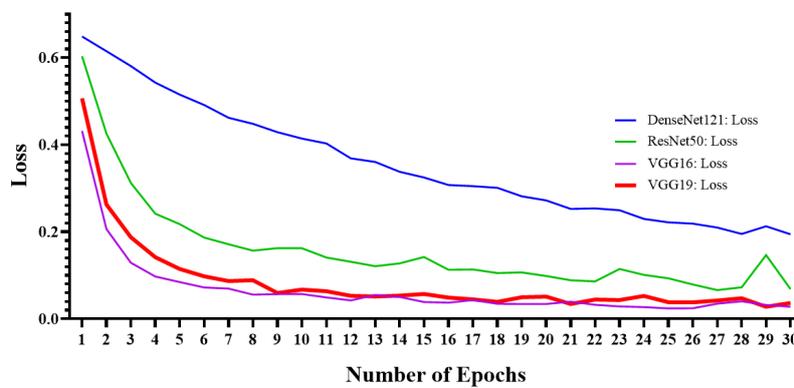


Figure 4. Training loss curve for DenseNet121, ResNet50, VGG16, and VGG19.

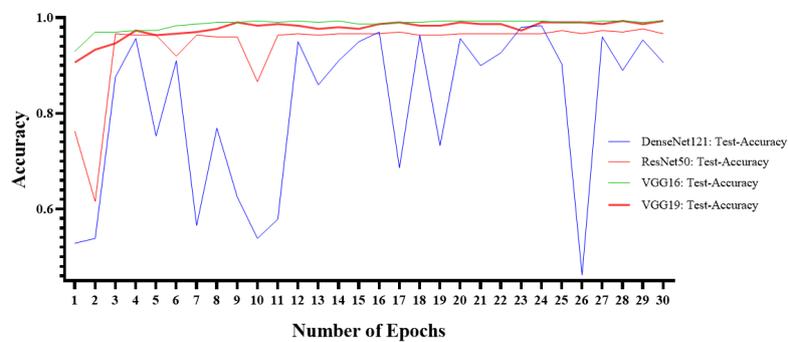


Figure 5. Testing accuracy curve for DenseNet121, ResNet50, VGG16, and VGG19.

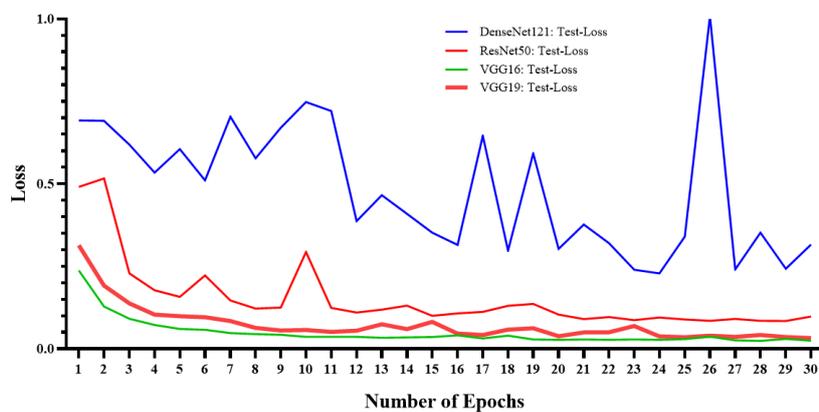


Figure 6. Testing loss curve for DenseNet121, ResNet50, VGG16, and VGG19.

Table 2. Comparative performance of ResNet50, VGG16, VGG19, and DenseNet121 for COVID-19 diagnosis using X-ray images.

Measures	ResNet50	VGG16	VGG19	DenseNet121
Accuracy (ACC)	97%	99.33%	99.33%	96.66%
Sensitivity (SEN)	98.48%	99.28%	100.00%	99.23%
Specificity (SPE)	95.21%	99.38%	98.77%	94.67%
False Negative Rate (FNR)	1.52%	0.72%	0.00%	0.77%
False Positive Rate (FPR)	4.79%	0.62%	1.23%	5.33%
Positive Predicted Value (PPV)	94.20%	99.28%	98.55%	93.48%
F1 Score (F1)	96.30%	99.28%	99.27%	96.27%

Consequently, the false positive rate value is also among the key measures because if a patient is wrongly predicted as positive and treated by the same health professional, that might expose the patient to the virus. Moreover, due to the exponential increase in the COVID-19 pandemic, most countries are facing difficulties in managing the patients as the number of patients is very high, and the resources are not enough to treat the patients. A high FPR increases the burden on the health care system, due to the increased number of required resources (RT-PCR tool kits, other medical resources), which will sometimes fail to accommodate the actual positive patients.

VGG16 achieved the lowest false positive rate of 0.62%, while the false negative rate of VGG19 was the best. Similarly, VGG16 achieved the highest outcomes in terms of positive predicted value and F1 score.

6. Comparison with Existing Studies

In the proposed study, four deep-learning models—DenseNet121, ResNet50, VGG16, and VGG19 were used. The data set consisted of a total of 1272 X-ray images (642 normal, 630 COVID19). To compare the performance of the proposed techniques, the outcome of the study was compared with the benchmark studies. The criterion for the benchmark was the studies using X-ray radiology images for the diagnosis of COVID-19. Table 3 contains the comparison of the proposed technique with the benchmark studies in the literature using X-ray images for the diagnosis of COVID-19.

Therefore, based on Table 3, the proposed study outperformed the studies in the benchmark. Most of the previous studies had a very limited number of COVID-19 X-ray radiology images. Novel coronavirus (COVID-19) is a new pandemic; limited open-source X-ray radiology images are available for developing a deep-learning based automated diagnosis model. Nevertheless, a huge number of X-ray images is present for other respiratory diseases. Most of the previous studies suffered from data imbalance issues. Widely used deep-learning models in the literature for COVID-19 diagnosed using X-ray radiology images were VGG19, ResNet, Inception, DenseNet, and SqueezeNet. However, in our study Xception and SqueezeNet did not provide good outcomes.

The main contributions of the current study are:

1. The study does not suffer from data imbalance.
2. The model was trained using a large number of COVID-19 X-ray radiology images when compared to the previous studies.
3. The proposed model is a fully automated diagnosis method and does not require any separate feature extraction or annotation prior to the diagnosis.
4. Data augmentation was applied to increase the generalization of the proposed model.
5. The model outperforms the benchmark studies.

Despite the above-mentioned advantages, the study also suffers from some limitations:

1. The proposed system needs to be trained for other respiratory diseases. The current model only diagnoses COVID-19 and healthy individuals and is unable to diagnose other kinds of pneumonia and respiratory infections.

2. The number of COVID-19 X-ray radiology images needs to be increased for better model training. The deep-learning model performance can be further enhanced with the increase in the size of the data set.
3. The current study was based on the data set curated using several open-source chest X-ray images. These samples were collected from various research publications or uploaded by volunteers. Therefore, these X-ray images were not collected in rigorous manner.

To alleviate the above mention limitations, there is a need for developing a model using X-ray data samples from the hospital.

Table 3. Comparison of the proposed method with the benchmark studies.

Study	Number of Samples	Technique	ACC	SEN	SPE
Hamdan et al. [29]	50 (25 healthy and 25 COVID-19)	COVIDX-Net	90%	-	-
Wang et al. [32]	13,975 (normal, pneumonia, and COVID-19)	Tailored CNN (COVID-Net)	93.3%	-	-
Apostolopoulos et al. [37]	1427 (224 COVID-19, 700 pneumonia, 504 normal)	VGG19, MobileNet, Inception, Xception, Inception ResNetV2	98.75%	99.1%	-
Kumar et al. [40]	50 (25 healthy and 25 COVID-19)	ResNet50 + SVM	95.38%	-	-
Ali et al. [42]	100 (50 normal and 50 COVID-19)	ResNet50	98%	-	-
Ozturk et al. [43]	625 (125 COVID-19, 500 no-findings)	DarkNet	98.08%	-	-
Minaee et al. [45]	(100 COVID-19, 5000 Non-COVID-19)	Deep-COVID	-	97.5%	90%
Ucar et al. [48]	2839 (1203 normal, 1591 pneumonia and 45 COVID-19)	COVIDiagnosis-Net	98.30%	-	-
Farooq et al. [49]	2813 (1203 normal, 931 bacterial pneumonia, 660 viral pneumonia, 19 COVID-19 cases)	COVID-ResNet	96.23%	-	-
Oh et al. [50]	502 (191 normal, 54 bacterial, 57 tuberculosis, 20 viral, 180 COVID-19)	ResNet18	88.90%	-	96.4%
	Proposed Study	VGG16, VGG19	99.38%	100%	99.33%

7. Conclusions

In this study, we used transfer-learning for an automated COVID-19 diagnosis using X-ray images. The motivation of using X-ray images for the diagnosis of COVID-19 is the lower sensitivity of the gold-standard RT-PCR diagnosis test. The proposed system achieved the highest sensitivity of 100% and specificity of 99.3% when compared to the studies in the benchmark. The system can assist radiologists in the early diagnosis of COVID-19. Generalization in the model was achieved by generating the data using augmentation. Moreover, the study attempted to use a large number of COVID-19 X-ray images by combining several open-source data sets. Despite combining multiple open-source data sets, there is still a need for an increased number of COVID-19 positive X-ray sample images. An increased number of COVID-19 X-ray samples will enhance the model's performance.

Author Contributions: Conceptualization, investigation and software, I.U.K.; data curation, formal analysis, writing—original draft and writing—review & editing, N.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

- COVID-19 Worldwide Statistics. Available online: <https://www.worldometers.info/coronavirus/> (accessed on 15 June 2020).
- Huang, C.; Wang, Y.; Li, X.; Ren, L.; Zhao, J.; Hu, Y.; Zhang, L.; Fan, G.; Xu, J.; Gu, X.; et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **2020**, *395*, 497–506. [[CrossRef](#)]
- West, C.P.; Montori, V.M.; Sampathkumar, P. COVID-19 Testing: The Threat of False-Negative Results. *Mayo Clin. Proc.* **2020**, *95*, 1127–1129. [[CrossRef](#)]
- Guyatt, G.; Rennie, D.; Maureen, O.M.; Cook, D.J. *Users' Guides to the Medical Literature: A Manual for Evidence-Based Clinical Practice*, 3rd ed.; McGraw-Hill Medical: New York, NY, USA, 2015.
- Yoon, S.H.; Lee, K.H.; Kim, J.Y.; Lee, Y.K.; Ko, H.; Kim, K.H.; Park, C.M.; Kim, Y.-H. Chest Radiographic and CT Findings of the 2019 Novel Coronavirus Disease (COVID-19): Analysis of Nine Patients Treated in Korea. *Korean J. Radiol.* **2020**, *21*, 494–500. [[CrossRef](#)]
- Xie, X. Chest CT for Typical Covid-19 pneumonia. *Radiology* **2020**, *21*, 494–500.
- Fang, Y.; Zhang, H.; Xie, J.; Lin, M.; Ying, L.; Pang, P.; Ji, W. Sensitivity of Chest CT for COVID-19: Comparison to RT-PCR. *Radiology* **2020**, *296*, E115–E117. [[CrossRef](#)]
- Chan, J.F.-W.; Yuan, S.; Kok, K.-H.; To, K.K.-W.; Chu, H.; Yang, J.; Xing, F.; Liu, J.; Yip, C.C.-Y.; Poon, R.W.-S.; et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: A study of a family cluster. *Lancet* **2020**, *395*, 514–523. [[CrossRef](#)]
- Ai, T.; Yang, Z.; Hou, H.; Zhan, C.; Chen, C.; Lv, W.; Tao, Q.; Sun, Z.; Xia, L. Correlation of Chest CT and RT-PCR Testing in Coronavirus Disease 2019 (COVID-19) in China: A Report of 1014 Cases. *Radiology* **2020**, *296*, E32–E40.
- Guan, W.-J.; Ni, Z.-Y.; Hu, Y.; Liang, W.-H.; Ou, C.-Q.; He, J.-X.; Liu, L.; Shan, H.; Lei, C.-L.; Hui, D.S.; et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708–1720. [[CrossRef](#)]
- Latif, S.; Usman, M.; Manzoor, S.; Iqbal, W.; Qadir, J.; Tyson, G.; Castro, I.; Razi, A.; Boulos, M.N.K.; Weller, A.; et al. Leveraging Data Science to Combat COVID-19: A Comprehensive Review. *TechRxiv* **2020**, 1–19. [[CrossRef](#)]
- Wynants, L.; Van Calster, B.; Collins, G.S.; Riley, R.D.; Heinze, G.; Schuit, E.; Bonten, M.M.J.; Dahly, D.L.; Damen, J.A.A.; Debray, T.P.; et al. Prediction models for diagnosis and prognosis of covid-19: Systematic review and critical appraisal. *BMJ* **2020**, *369*, m1328. [[CrossRef](#)]
- Swapnarekha, H.; Behera, H.S.; Nayak, J.; Naik, B.; Hanumanthu, S.R. Role of intelligent computing in COVID-19 prognosis: A state-of-the-art review. *Chaos Solitons Fractals* **2020**, *138*, 109947. [[CrossRef](#)] [[PubMed](#)]
- Hesamian, M.H.; Jia, W.; He, X.; Kennedy, P. Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges. *J. Digit. Imaging* **2019**, *32*, 582–596. [[CrossRef](#)] [[PubMed](#)]
- Balyen, L.; Peto, T. Promising Artificial Intelligence-Machine Learning-Deep Learning Algorithms in Ophthalmology. *Asia Pacific J. Ophthalmol.* **2019**, *8*, 264–272. [[CrossRef](#)]
- Connor, C.W. Artificial Intelligence and Machine Learning in Anesthesiology. *Anesthesiology* **2019**, *131*, 1346–1359. [[CrossRef](#)]
- Galbusera, F.; Casaroli, G.; Bassani, T. Artificial intelligence and machine learning in spine research. *JOR Spine* **2019**, *2*, e1044. [[CrossRef](#)]
- Gjoreski, M.; Gradišek, A.; Budna, B.; Gams, M.Z.; Poglajen, G. Machine Learning and End-to-End Deep Learning for the Detection of Chronic Heart Failure from Heart Sounds. *IEEE Access* **2020**, *8*, 20313–20324. [[CrossRef](#)]
- Kumar, A.; Fulham, M.; Feng, D.; Kim, J. Co-Learning Feature Fusion Maps From PET-CT Images of Lung Cancer. *IEEE Trans. Med. Imaging* **2020**, *39*, 204–217. [[CrossRef](#)]

20. Podnar, S.; Kukar, M.; Gunčar, G.; Notar, M.; Gošnjak, N.; Notar, M. Diagnosing brain tumours by routine blood tests using machine learning. *Sci. Rep.* **2019**, *9*, 14481–14487. [CrossRef]
21. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [CrossRef]
22. Celik, Y.; Talo, M.; Yildirim, O.; Karabatak, M.; Acharya, U.R. Automated invasive ductal carcinoma detection based using deep transfer learning with whole-slide images. *Pattern Recognit. Lett.* **2020**, *133*, 232–239. [CrossRef]
23. Deng, L. Deep Learning: Methods and Applications. *Deep Learn. Methods Appl.* **2014**, *7*, 197–387. [CrossRef]
24. Gaál, G.; Maga, B.; Lukács, A. Attention U-Net Based Adversarial Architectures for Chest X-ray Lung Segmentation. *arXiv* **2020**, arXiv:2003.10304.
25. Souza, J.C.; Diniz, J.O.B.; Ferreira, J.L.; Da Silva, G.L.F.; Silva, A.C.; De Paiva, A.C. An automatic method for lung segmentation and reconstruction in chest X-ray using deep neural networks. *Comput. Methods Programs Biomed.* **2019**, *177*, 285–296. [CrossRef]
26. Rajpurkar, P.; Irvin, J.; Zhu, K.; Yang, B.; Mehta, H.; Duan, T.; Ding, D.; Bagul, A.; Langlotz, C.; Shpanskaya, K.; et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-rays with Deep Learning. *arXiv* **2017**, arXiv:1711.05225.
27. Lakhani, P.; Sundaram, B. Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* **2017**, *284*, 574–582. [CrossRef] [PubMed]
28. Chen, D.; Liu, F.; Li, Z. A Review of Automatically Diagnosing COVID-19 based on Scanning Image. *arXiv* **2020**, arXiv:2006.05245.
29. Hemdan, E.E.D.; Shouman, M.A.; Karar, M.E. COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-ray Images. *arXiv* **2020**, arXiv:2003.11055.
30. Cohen, J.P.; Morrison, P.; Dao, L. COVID-19 Image Data Collection. *arXiv* **2020**, arXiv:2003.11597v1.
31. Detecting COVID-19 in X-ray Images with Keras, TensorFlow, and Deep Learning. Available online: <https://www.pyimagesearch.com/2020/03/16/detecting-covid-19-in-x-ray-images-with-keras-tensorflow-and-deep-learning/> (accessed on 16 June 2020).
32. Wang, L.; Wong, A. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-ray Images. *arXiv* **2020**, arXiv:2003.09871.
33. COVID-19 Chest X-ray Dataset Initiative. Available online: <https://github.com/agchung/Figure1-COVID-chestxray-dataset> (accessed on 16 June 2020).
34. RSNA Pneumonia Detection Challenge. Available online: <https://www.kaggle.com/c/rsna-pneumonia-detection-challenge/data> (accessed on 16 June 2020).
35. Actualmed-COVID-chestxray-dataset. Available online: <https://github.com/agchung/Actualmed-COVID-chestxray-dataset> (accessed on 16 June 2020).
36. COVID-19 Radiography Database. Available online: <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database> (accessed on 16 June 2020).
37. Apostolopoulos, I.D.; Mpesiana, T.A. Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **2020**, *43*, 635–640. [CrossRef]
38. Radiopaedia. Available online: <https://radiopaedia.org/> (accessed on 16 June 2020).
39. Italian Society of Medical and Interventional Radiology (SIRM). Available online: <https://www.sirm.org/en/italian-society-of-medical-and-interventional-radiology/> (accessed on 16 June 2020).
40. Kumar, P.; Kumari, S. Detection of coronavirus Disease (COVID-19) based on Deep Features. *Preprints* **2020**. [CrossRef]
41. Chest X-ray Images (Pneumonia). Available online: <https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia> (accessed on 16 June 2020).
42. Narin, A.; Kaya, C.; Pamuk, Z. Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. *arXiv* **2020**, arXiv:2003.10849.
43. 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. Boil. Med.* **2020**, *121*, 103792. [CrossRef] [PubMed]
44. Wang, X.; Peng, Y.; Lu, L.; Lu, Z.; Bagheri, M.; Summers, R.M. ChestX-ray8: Hospital-scale chest X-ray database and benchmarks on weakly-supervised classification and localization of common thorax diseases. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 3462–3471.

45. Minaee, S.; Kafieh, R.; Sonka, M.; Yazdani, S.; Soufi, G.J. Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. *Med. Image Anal.* **2020**, *65*, 101794. [[CrossRef](#)] [[PubMed](#)]
46. Irvin, J.; Rajpurkar, P.; Ko, M.; Yu, Y.; Ciurea-Ilcus, S.; Chute, C.; Marklund, H.; Haghgoo, B.; Ball, R.; Shpanskaya, K.; et al. CheXpert: A Large Chest Radiograph Dataset with Uncertainty Labels and Expert Comparison. *Proc. AAAI Conf. Artif. Intell.* **2019**, *33*, 590–597. [[CrossRef](#)]
47. 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. *arXiv* **2020**, arXiv:2004.02696.
48. Ucar, F.; Korkmaz, D. COVIDiagnosis-Net: Deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Med. Hypotheses* **2020**, *140*, 109761. [[CrossRef](#)] [[PubMed](#)]
49. Farooq, M.; Hafeez, A. COVID-ResNet: A Deep Learning Framework for Screening of COVID19 from Radiographs. *arXiv* **2020**, arXiv:2003.14395.
50. Oh, Y.; Park, S.; Ye, J.C. Deep Learning COVID-19 Features on CXR using Limited Training Data Sets. *IEEE Trans. Med. Imaging* **2020**, *39*, 2688–2700. [[CrossRef](#)]
51. Shiraishi, J.; Katsuragawa, S.; Ikezoe, J.; Matsumoto, T.; Kobayashi, T.; Komatsu, K.-I.; Matsui, M.; Fujita, H.; Kodera, Y.; Doi, K. Development of a Digital Image Database for Chest Radiographs with and without a Lung Nodule. *Am. J. Roentgenol.* **2000**, *174*, 71–74. [[CrossRef](#)]
52. Van Ginneken, B.; Stegmann, M.B.; Loog, M. Segmentation of anatomical structures in chest radiographs using supervised methods: A comparative study on a public database. *Med. Image Anal.* **2006**, *10*, 19–40. [[CrossRef](#)] [[PubMed](#)]
53. Jaeger, S.; Candemir, S.; Antani, S.; Wang, Y.-X.J.; Lu, P.-X.; Thoma, G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant. Imaging Med. Surg.* **2014**, *4*, 475–477. [[PubMed](#)]
54. CoronaHack—Chest X-ray-Dataset. Available online: <https://www.kaggle.com/praveengovi/coronahack-chest-xraydataset> (accessed on 17 June 2020).
55. Chouhan, V.; Singh, S.K.; Khamparia, A.; Gupta, N.; Tiwari, P.; Moreira, C.; Damasevicius, R.; De Albuquerque, V.H.C. A Novel Transfer Learning Based Approach for Pneumonia Detection in Chest X-ray Images. *Appl. Sci.* **2020**, *10*, 559. [[CrossRef](#)]
56. Rahman, T.; Chowdhury, M.E.H.; Khandakar, A.; Islam, K.R.; Islam, K.F.; Mahbub, Z.B.; Kadir, M.A.; Kashem, S.; Rahman, T. Transfer Learning with Deep Convolutional Neural Network (CNN) for Pneumonia Detection Using Chest X-ray. *Appl. Sci.* **2020**, *10*, 3233. [[CrossRef](#)]
57. Pan, S.J.; Yang, Q. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2009**, *22*, 1345–1359. [[CrossRef](#)]
58. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. ImageNet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[CrossRef](#)]
59. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 2261–2269.
60. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; pp. 770–778.
61. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, 7–9 May 2015; pp. 1–14.
62. Hossin, M.; Sulaiman, M.N. A Review on Evaluation Metrics for Data Classification Evaluations. *Int. J. Data Min. Knowl. Manag. Process.* **2015**, *5*, 1–11. [[CrossRef](#)]

