

Article

Multi-Classification of Chest X-rays for COVID-19 Diagnosis Using Deep Learning Algorithms

Abeer A. Abdelhamid , Eman Abdelhalim , Mohamed A. Mohamed  and Fahmi Khalifa * 

Electronics and Communications Engineering Department, Faculty of Engineering, Mansoura University, Mansoura 35516, Egypt; abeer.abdelhamid@mans.edu.eg (A.A.A.); eman_haleim@mans.edu.eg (E.A.); mazim12@mans.edu.eg (M.A.M.)

* Correspondence: fahmikhalfifa@mans.edu.eg

Abstract: Accurate detection of COVID-19 is of immense importance to help physicians intervene with appropriate treatments. Although RT-PCR is routinely used for COVID-19 detection, it is expensive, takes a long time, and is prone to inaccurate results. Currently, medical imaging-based detection systems have been explored as an alternative for more accurate diagnosis. In this work, we propose a multi-level diagnostic framework for the accurate detection of COVID-19 using X-ray scans based on transfer learning. The developed framework consists of three stages, beginning with a pre-processing step to remove noise effects and image resizing followed by a deep learning architecture utilizing an Xception pre-trained model for feature extraction from the pre-processed image. Our design utilizes a global average pooling (GAP) layer for avoiding over-fitting, and an activation layer is added in order to reduce the losses. Final classification is achieved using a softmax layer. The system is evaluated using different activation functions and thresholds with different optimizers. We used a benchmark dataset from the kaggle website. The proposed model has been evaluated on 7395 images that consist of 3 classes (COVID-19, normal and pneumonia). Additionally, we compared our framework with the traditional pre-trained deep learning models and with other literature studies. Our evaluation using various metrics showed that our framework achieved a high test accuracy of 99.3% with a minimum loss of 0.02 using the LeakyReLU activation function at a threshold equal to 0.1 with the RMSprop optimizer. Additionally, we achieved a sensitivity and specificity of 99 and F1-Score of 99.3% with only 10 epochs and a 10^{-4} learning rate.

Keywords: COVID-19; transfer learning (TL); chest X-ray (CXR); deep learning (DL); artificial intelligence (AI)



Citation: Abdelhamid, A.A.; Abdelhalim, E.; Moahmed, M.A.; Khalifa, F. Multi-Classification of Chest X-rays for COVID-19 Diagnosis Using Deep Learning Algorithms. *Appl. Sci.* **2022**, *12*, 2080. <https://doi.org/10.3390/app12042080>

Academic Editor: Jan Egger

Received: 19 December 2021

Accepted: 8 February 2022

Published: 17 February 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The coronavirus (COVID-19) is possibly the biggest human threat of the twenty-first century. Since the disease outbreak in December 2019, the COVID-19 pandemic has become a worldwide health problem. According to the World Health Organization (WHO) statistics, in January 2022, there were more than 365 million confirmed COVID-19 cases, resulting in approximately 5.5 million deaths [1]. The COVID-19 pandemic disease is caused by severe acute respiratory syndrome coronavirus 2, or SARS-CoV-2. COVID-19 belongs to the family of viruses which cause cold-related diseases, such as Middle East respiratory syndrome (MERS-CoV) and extreme acute respiratory syndrome [2]. The emergence of disease transmission and the rise in mortality in a number of countries has necessitated the protection of health services and the community from COVID-19 spread. Thus, remote disease monitoring, including early diagnosis and quarantine, and follow-up is of immense importance. Everyday, thousands of cases have been reported in many countries around the world, which lead some governments to implement lockdown measures to contain the virus. Unfortunately, lockdowns due to the pandemic have not only had a global impact on health, but they have also harmed the economy. Thus, lockdown restrictions

were gradually released with strict obligations on wearing masks and maintaining physical distance. Health professionals from all over the world have since been working tirelessly to obtain drugs and vaccines for the disease [3].

Most COVID-19 infected patients have been diagnosed with pneumonia; thus, radiological examinations can be helpful for diagnosis and evaluation as well as for follow-up on disease progression. Initial screening using chest computed tomography (CT) revealed an over-sensitivity to reverse transcription-polymerase chain reaction (RT-PCR) and also reported COVID-19 infection in negative or low-positive RT-PCR cases. It can sometimes outperform the RT-PCR-based test, which has a low success rate of 70% and a sensitivity of 60–70% [4]. Recent surveys show a pooled sensitivity rate of 94% (95% confidence interval (CI) \in (91%, 96%)) but a low specificity rate of 37% (95% CI \in (26%, 50%)) for CT-dependent analysis [5]. As a result, CT-dependent analysis could be used to solve the insufficient PCR test sensitivity rate [6]. The most recent COVID-19 research has centred essentially on the observations of chest CTs. Nonetheless, a spike in the increase of COVID-19 prevalence limits the regular use of CT, as it puts a significant burden on patient health due to frequent radiation exposure and possible infections in CT suites.

Recent studies claim to obtain reliable results for the automated detection of the disease from chest X-ray (CXR) scans [7]. Since diagnosis has become a relatively rapid operation, the financial problems of diagnostic tests have affected patients worldwide, particularly in countries with private health systems or limited access to health systems due to prohibitive costs. There was an increase in the number of publicly available CXRs from stable cases as well as COVID-19 patients. This helps researchers around the globe analyze medical images and recognise potential trends that may contribute to an automatic diagnosis of the disease [8]. Moreover, CXR tests have the advantages of a fast operating speed, low cost, and ease of use for radiologists [9]. Thus, the need to identify COVID-19 features on CXR is growing [10].

Noninvasive approaches based on artificial intelligence (AI) for the analysis of patient data (e.g., CXRs and CTs) are extensively exploited for the successful diagnosis of COVID-19. In particular, deep learning (DL)-based techniques applied to radiomic features of thoracic imaging, CXR and CT as well as other clinical, pathological and genomics parameters have shown to provide valuable assistance in this direction. In the context of medical image analysis, DL automatically discovers hidden features required for detection and/or disease classification from raw data. Namely, images' pixel/voxels at the input are explicitly used instead of the representative (extracted or selected) features. This in turns reduces errors caused by incorrect segmentation and/or subsequent extraction of hand-crafted features. Literature research has demonstrated that machine learning (ML) and DL provide quick, automated, efficient strategies for detecting abnormalities and extracting key features of altered lung parenchyma; this may be connected to the unique signatures of the COVID-19 virus. However, the available datasets of COVID-19 are insufficient for the development of deep neural networks [11].

In short, early, accurate and rapid COVID-19 diagnosis plays a key role in timely quarantine and medical care. This is also of great significance for the prognosis of patients, the prevention of this epidemic and the protection of public health [12]. In this work, we propose a DL approach for the detection of COVID-19 from CXR images. The proposed pipeline is based on TL by employing the Xception pre-trained model for the feature extraction stage. Then, we used the global average pooling (GAP) layer to solve the vanishing gradient problem when using DL networks and to reduce overfitting probability; the activation layer was used for reducing the losses. We used a dataset which consisted of COVID-19, pneumonia and normal CXR images. We investigated different activation layers with different activation functions and optimization algorithms and compared between them to select the one that provides the best performance with minimum losses. We compared our approach with traditional pre-trained models and also with other literature studies. Our design achieves the highest performance compared to the previous studies and the pre-trained models.

The rest of the paper is partitioned into the following sections. Section 2 provides an overview of the related work for recent COVID-19 detection studies using different AI algorithms. Details of the methods and the processing stages of the developed framework are fully described in Section 3. Then, Section 4 describes the performance evaluation and validation methods/metrics for evaluating classification accuracy. The experimental results showing the potential of the proposed pipeline are given in Section 5, and Section 6 presents the results discussion. Finally, Section 7 contains the conclusions.

2. Related Work

Recent research work has proven that imaging tests (e.g., CXRs and CT) can provide rapid identification of COVID-19 and also help to monitor the spread of the disease. Various image-based diagnostic systems, ranging from hand-crafted features to feature learning, have been introduced to illustrate the possibility of the identification of COVID-19 [13]. Convolutional neural networks (CNN) have been found to be one of the most common and successful techniques for diagnosing COVID-19 from medical images. Yudong and Kulwa [14] used deep TL techniques for a multi-classification task on a small dataset by testing 15 different pre-trained models. Their dataset consisted of 860 images (260 for COVID-19, 300 normal and 300 for pneumonia). Their study revealed that VGG19 was the best algorithm, achieving a classification accuracy of 89.3% with average precision, recall, and F1-Score values of 0.90, 0.89, 0.90, respectively. Yama et al. [15] proposed an approach based on a convolution support estimation network (CSEN), which constructs a sparse support set of representation coefficients using a dictionary and a set of training samples. The dataset used included 6286 CXR images of COVID-19, as well as three other classifications: bacterial pneumonia, viral pneumonia, and normal. Using a 5-fold cross validation, their CSEN-based model achieved a sensitivity and specificity of more than 98% and 95%, respectively, for COVID-19 detection. A generative adversarial network (GAN)-based method was presented by Nour Eldeen et al. [16] to detect COVID-19 infection from CXR images. Their dataset consisted of 307 images divided into 4 classes (COVID-19, normal, bacterial pneumonia, and viral pneumonia). They used AlexNet, GoogleNet, and ResNet18 algorithms as TL. Their model achieved an accuracy of 80.6% on 4 classes (GoogleNet), 85.3% on 3 classes (AlexNet) and 100% on 2 classes (GoogleNet). A similar approach by Khan et al. [17] combined TL and the Xception pre-trained model. Their system attained a 89.6 % and 95% accuracy on classification for 4 and 3 classes, respectively. Another DCNN transfer learning-based pipeline by Asif et al. [18] utilized Inception V3 for the detection of COVID-19 in infected patients using chest X-ray scans. The test data contained 864 scans for COVID-19, 1345 for viral pneumonia and 1341 for normal scans. The model provided a classification accuracy of greater than 98% (training accuracy of 97% and validation accuracy of 93%). Suat and Alakus [19] produced a convolutional capsule network using chest scans for COVID-19 detection called CapsNet. Their system was evaluated using a total of 2331 images, of which 231 were COVID-19 and 1050 were normal and pneumonia. For binary and multi-class classification, the 11-layer architecture achieved 97.24% and 84.22% accuracy, respectively.

Kumari et al. [20] used the ResNet50 plus support vector machine (SVM) model as a feature extractor in a framework for identifying COVID-19 patients. They proposed a multi-classification task on 381 CXR images. Their model achieved an accuracy, sensitivity, FPR and an F1-Score of 95.33%, 95.33%, 2.33% and 95.34%, respectively. The limitation of their method is the leakage of the used dataset. In [21], the DenseNet121 model was used by Sarker et al. for both binary and multi-classification tasks of COVID-19 patients on a dataset consisting of 238 images for COVID-19, 6045 for pneumonia and 8851 for normal persons. Classification results of 2 and 3 classes obtained 96.49% and 93.71% accuracy, respectively. A study by Dilbag et al. [22] compared several types of TL with their proposed model for a multi-classification task of COVID-19 on a dataset that included three classes: COVID-19, pneumonia and other disease patients. Their model achieved a test accuracy of 97.4% between the used models. Barshooi et al. [23] proposed a model for screening for

COVID-19 infection using a GAN DL approach. Data augmentation were employed using different filter banks. A total of 4560 CXR images of patients with COVID (360 cases), as well as viral, bacterial, fungal, and other diseases, were used. Their model detection accuracy was compared to the performances of 10 existing COVID-19 identification techniques and achieved an accuracy of 98.5% in the 2-class classification task. The advantage of their method is that the utilization of different filters improved the performance of the classifiers. The limitations of their study is that the dataset used was small.

A 2-way classification framework for detecting 15 different types of chest diseases including COVID-19 using CXR images was proposed by Rehman et al. [24]. First, a CNN architecture with a softmax classifier was used. Then, TL was used with a fully connected layer of the proposed CNN to extract deep features. Deep features were then fed into traditional ML classifier. They performed 10-fold and 5-fold validation on the best-performing of the 7 ML classifiers. They collected 2800 CXR images belongs to 14 classes and 200 CXR images belonging to the COVID-19 class. Their method achieved an overall validation accuracy of 99.40% with the KNN-fine algorithm using 5-fold validation and 99.77% with the Bag-ensemble algorithm using 10-fold validation. Although their method has the advantages of being a fusion between the deep and machine learning models, the dataset used was small, and their method of training took a long time (500 epochs). In [25], Brima et al. designed a ResNet50 CNN-based architecture for detecting and classifying four types of classes using CXRs. Their dataset consisted of 21,165 CXR images divided into 6012 images for lung opacity, 3616 images for COVID-19, 1345 images for viral pneumonia and 10,192 images for the normal class. Their approach scored a test accuracy of 94% using the 5-fold cross validation technique. The advantage of this study is the large dataset used to compare different pre-trained models to detect the most suitable model. The limitation is the large number of epochs (100) used for training with the limited computing needed to perform hyper-parameter space searches, and, as a result, it took a long time. Additionally, the output accuracy was not high, although the dataset used was large and the principle of the TL was applied. To alleviate the burden placed on a single network, a multi-step classification by Albahli et al. was proposed in [26] for detecting COVID-19 and other chest diseases using X-ray images. They applied the TL with different pre-trained models using data augmentation and semantic segmentation in order to increase the model's accuracy in a 10-fold cross-validation. For the first level of classification (i.e., 3 classes), their technique achieved an average test accuracy of 92.52%. In the second level of classification (i.e., 14 classes), using a ResNet50 model, their technique achieved a maximum test accuracy of 66.63%. For all 16 classes, which were classified at once, the overall accuracy for COVID-19 detection decreased, achieving a rate of 71.91%. The advantage of this study is that the combination between the two classifiers provided a compatible accuracy on the dataset used, and the dataset contained more than two classes. The limitations of this study are that the accuracy was not high despite applying the TL and data augmentation techniques, and the dataset used was of a small size. Manokaran et al. proposed a modified DenseNet201 network [27] for COVID-19 detection that included a global averaging layer, a batch normalization layer, a dense layer with ReLU activation, and a final classification layer. Their model was trained using 8644 images (4000 normal and pneumonia cases and 644 COVID-19 cases) and tested on 1729 images (129 COVID-19, 800 normal, and 800 pneumonia) and yielded an overall accuracy of 92.19% compared to 7 pre-trained models. The advantage of this study is that the employment of the TL technique provided an acceptable accuracy. The limitation of this study is that the overall accuracy was not high although the dataset was large and the TL was employed. Additionally, the model was trained for 100 epochs; thus, training took a long time.

In summary, after the COVID-19 outbreak, a tremendous amount of research regarding the detection of COVID-19 has been conducted using different types of medical images (e.g., CXR and CT [28–32]). The existing techniques have their own advantages and disadvantages. Handcrafted vs. feature learning-based methods, time complexity, sample data size, binary and multi-level classification are the main criteria used for evaluation and

models' comparisons. In this paper, we propose and investigate a deep learning-based COVID-19 detection model that integrates TL with an Xception pre-trained module for feature extraction. Compared to [27], our system provides a multi-level approach for the accurate detection of COVID-19 using X-ray scans. In our model, we did not employ the pre-trained model as is. We applied the TL principle in the first stage of building the structure for the feature extraction step and benefited from its experience in order to achieve high performance. Additionally, we have tested our method on a larger dataset and investigated different activation layers, functions, and optimizers to attain the best performance with minimum losses.

3. Methodology

The proposed analysis pipeline is demonstrated in Figure 1, which consists of multiple analysis blocks: preprocessing, the proposed CNN model with TL, and model training and parameter setting. In the following subsections, details of those stages are fully illustrated.

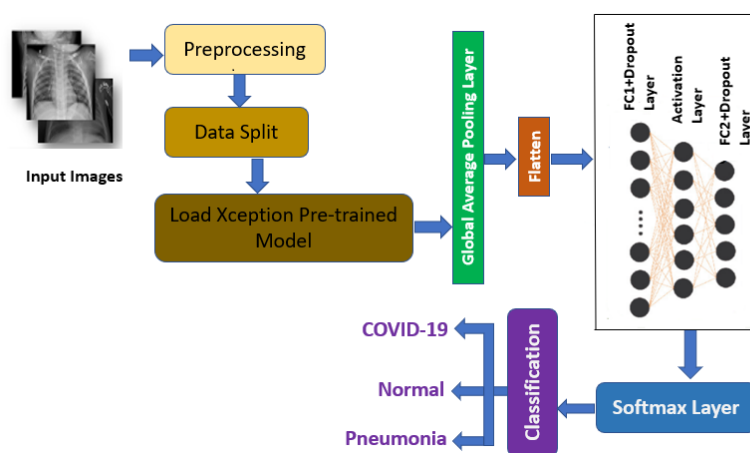


Figure 1. A block diagram of the structure of the proposed classification model.

3.1. Data Preprocessing

In the field of deep learning, data preprocessing represents the phase of data cleaning that improves the input data for tasks. This procedure is important for many reasons. First, it ensures the generalization of a given model, especially when tested on datasets outside its training cohort. Additionally, preprocessing reduces data noise and/or distortions and allows the network to operate more efficiently and faster. Because of the use of the transfer learning technique, the data must be rescaled to fit the input of the model. In our work, we initially employed data normalization where all pixel values were rescaled to $[-1, 1]$ using a pixel-wise multiplication factor of $1/255$, giving a set of grayscale images as follows:

$$R_i = \frac{I - I_{min}}{I_{max} - I_{min}} \quad (1)$$

where R_i is the normalized data, I is the original data, I_{max} and I_{min} are the maximum and minimum values in the input data, respectively. Then, histogram equalization [33] was applied to enhance the images' contrast. After that, images were resized to 200×200 before training began. Besides, some transformations were also applied to the images, including rotation by 25 degrees, a zoom range equal to 0.2 and setting the fill-mode to nearest. Data augmentation was used to improve our model's generalization ability because it mitigates overfitting by adding variations to the dataset.

3.2. Base Model Stage

In order to reduce the learning time of DL techniques, specially in CNN-based classifications tasks, TL is employed to rapidly train a CNN model without initializing its weights

from scratch [34]. Additionally, TL is especially useful for tasks where adequate training samples are not available to train a model from scratch, such as medical image detection for rare or emerging diseases. Practically, TL imports the model parameters from a pre-trained model from other tasks. The transferred parameters have good initial values and only require a few minor adjustments to be better curated for the task at hand [35]. The benefit is that it takes less training time and is possible to achieve higher accuracy with small data samples. In our proposed model, we selected the Xception pre-trained model to be the base model, which was trained on the ImageNet dataset.

The Xception model is used for feature extraction with the CNN architecture. In 2017, Xception was developed by Francois Chollet, author of *Keras*. It is another enhancement of the Inception V3 proposed by Google [36]. It is a newer form than Inception and is called Xception because it was the “extreme version of Inception” [37]. It is a CNN architecture that consists of a linear stack of depth-wise separable convolution layers with residual connections. The feature extraction base is formed by 36 convolutional layers that are formulated into 14 modules with outlined linear residual connections [38]. Moreover, the Xception model works on one-by-one convolution first, followed by channel-wise spatial convolution. The separable convolution layer has fewer parameters and a lower computational cost than a traditional convolutional layer. There is no intermediate activation in Xception. As a result, we selected the Xception model in our design. Figure 2 demonstrates the structure of the Xception model, as described in [39].

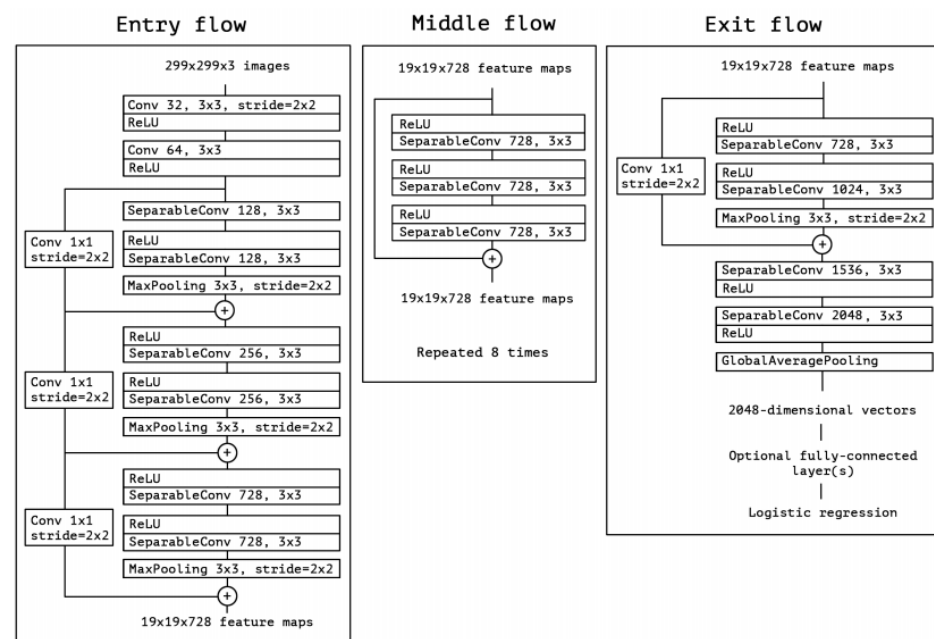


Figure 2. Detailed structure of the Xception model [39]. Reprinted with permission from [39]. Copyright 2017 IEEE.

3.3. Classification Stage

To make the model suitable for feature extraction, we do not include the classification layers at the top. After importing the base model from *Keras*, we have some additional layers in our structure. The global average pooling (GAP) layer is added immediately after the base model to reduce the total number of parameters that may lead the model to overfitting. In the next step, we applied the flatten layer, which converts the two-dimensional feature matrix into a vector. After the flatten process, two dense layers with 1024 and 512 neurons, respectively, were added which use the activation function with a threshold equal to alpha, α , followed by the dropout layer with a value of Γ . In the proposed model, we used two activation functions. The first is a LeakyReLU activation

function and the second activation is an exponential linear unit (ELU). Those functions are mathematically described by Equations (2) and (3), respectively:

$$\text{LeakyReLU}(R_i) = \begin{cases} \alpha R_i, & R_i < 0 \\ R_i, & R_i \geq 0 \end{cases} \quad (2)$$

$$\text{ELU}(R_i) = \begin{cases} R_i & R_i \geq 0 \\ \alpha(e^{R_i} - 1) & R_i < 0 \end{cases} \quad (3)$$

where R_i is the input data. The main goal of the dropout layer is to avoid the occurrence of overfitting. An activation layer with an activation function was added between the first dropout layer and the second fully connected layer with a threshold equal to α . This layer allows a small gradient during the training process, and it reduces the losses of the model. The main objective of our research is to study the effect of changing α values and applying different optimizers on the classification accuracy. We performed various experiments to examine the best performance of the model with minimum losses. The model has a softmax layer at the end to generate the output according to the softmax equation. The softmax activation function is defined by Equation (4):

$$Z(T_i) = \frac{e^{T_i}}{\sum_{k=1}^M e^{T_k}} \quad (4)$$

where T_i represents the input data to the softmax layer from the previous layer and M is the total number of classes. The softmax converts the scores (T_i) to a normalized probability, such that the input belongs to the i th class. The denominator in Equation (4) ensures that the output sums to unity (i.e., "1"). In the final step, the categorical cross entropy loss function was applied with the Adam optimizer (other optimization algorithms were also investigated). It uses the predicted class (or softmax output) and the ground truth probabilities of the class to calculate the loss, which is defined by Equation (5):

$$\text{CLF} = - \sum_{n=1}^S Y_n \cdot \log Y_n^* \quad (5)$$

where S is the total number of scalar values in the output, Y_n is the corresponding target value and Y_n^* is the n th scalar value in the output of the model. The proposed model has a total of 23,485,995 parameters, divided into 23,431,467 and 54,528 trainable and non-trainable parameters, respectively, as shown in Table 1.

Table 1. Summary of the parameter settings of the proposed structure.

Layer	Out Shape	No. of Parameters
Input	(200, 200, 3)	0
Xception (functional)	(None, 7, 7, 2048)	20,861,480
Global average pooling2D	(None, 2048)	0
Flatten	(None, 2048)	0
Dense	(None, 1024)	2,098,176
Dropout	(None, 1024)	0
Activation	(None, 1024)	0
Dense	(None, 512)	524,800
Dropout	(None, 512)	0
Dense	(None, 3)	1539
Total parameters		23,485,995
Trainable parameters		23,431,467
Non-trainable parameters		54,528
Gamma (dropout layer's threshold)		0.1
Batch size		32
No. of Epochs		10

4. Performance Evaluation and Validation

Generally, there are a variety of metrics that can be used to evaluate the performance of classification models. Those include classification accuracy, precision, and the F1-score. All those metrics depend on four expected outcomes: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). In our particular model, TP represents the COVID-19 patients identified correctly by our system, FP represents healthy subjects identified as patients, TN is the healthy subjects identified correctly, and FN are COVID-19 patients categorised as healthy subjects.

Based on those outcomes, various evaluation parameters can be measured. For example, the classification accuracy represents the number of labels correctly categorised divided by the total number of labels to be classified [40]. The recall (specificity) is another metric that calculates the fraction of positive (negative) patterns that are correctly categorised. The positive patterns that are correctly predicted from the total predicted patterns in the positive class are estimated by the precision metric. Finally, the harmonic mean between the REC and PER values is described by the F1-score. Those metrics are, respectively, defined mathematically by [41] as:

$$ACC = \frac{TN + TP}{TN + TP + FN + FP} \quad (6)$$

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

$$PER = \frac{TP}{TP + FP} \quad (8)$$

$$F1\text{-Score} = 2 * \frac{(PER \times REC)}{(PER + REC)} \quad (9)$$

In addition to accuracy metrics, the receiver operating characteristics (ROC) curve is also employed to confirm and support the robustness and accuracy of our deep learning model. ROCs depict a model's classification output based on the true positive and false positive rates at various classification thresholds. Quantitatively, an additional evaluation metric is usually employed in classification systems using the area under the curve (AUC) of an ROC. The AUC value illustrates how well the model performs by discriminating between classes [42]. AUC can be calculated according to Equation (10):

$$AUC = \frac{P_s - p_n(N_n + 1)/2}{p_n N_n} \quad (10)$$

where P_s is the sum of all the positive examples ranked, and p_n , N_n signifies the number of positive and negative examples.

5. Experimental Results

The test dataset contains 7395 images that have been prepared as a benchmark for the research community for the multi-classification task of detecting COVID-19 from CXR images. This dataset is collected from many links available on the "kaggle" website [43–46]. The COVID-19 class of images are labelled by a licensed radiologist, and only those with a specific sign are used for the purpose of the study. A total of 7395 images were used (1371 for COVID-19, 1751 for normal, and 4273 for pneumonia) and Figure 3 shows samples of every class in the dataset.

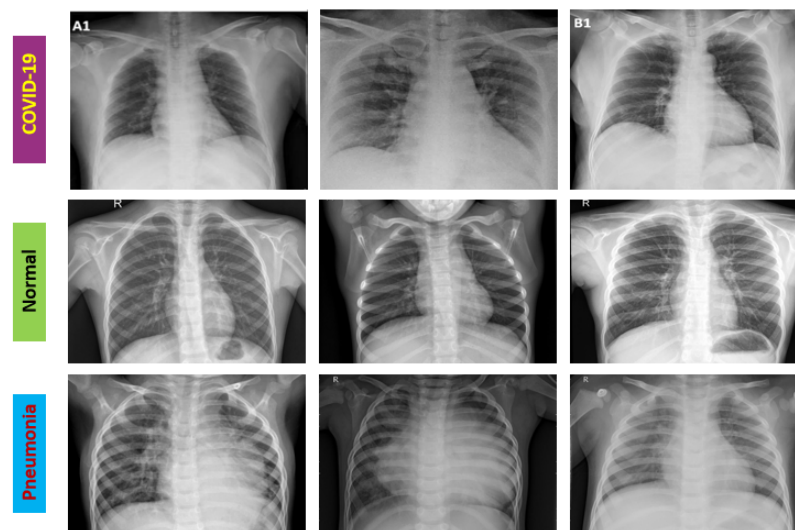


Figure 3. CXR examples for various classes from the test dataset.

In this research, to precisely identify COVID-19, we employed a DL algorithm that utilized TL for the multi-classification of chest X-rays into three classes: COVID-19, normal and pneumonia. As can be readily seen, the performance measures of the proposed model for COVID-19 classification depends on different optimizers and different activation functions. The model was developed using the python programming language. Particularly, we used the TensorFlow open-source model, which is commonly used for machine learning applications, such as neural networks and Keras, a high-level model of neural networks built on top of TensorFlow. The proposed model was trained with a learning rate η of 10^{-4} , a batch size of 32 and categorical cross-entropy with RMSprop was used as the loss and optimization function. The training used 80% of the dataset, and validation and testing used 10% each. During model training, the loss curve for the training vs validation showed that the validation loss became low after a few epochs, see Figure 4, thus we select a number of epochs equal to 10 in our experimnts. All images were resized to the target size of $200 \times 200 \times 3$ with scale, and the augmentation principle was applied before being fed into the neural network. In addition to quantitative accuracy, the system performance against the number of epochs was analyzed. First, we needed to define the parameters used in the DL criteria. Furthermore, the training and validation loss was calculated as the sum of the errors made for each example in the validation or training sets. Generally, DL-based classification systems, including CNNs, utilize gradient descent (GD)-based optimization techniques to lower the error (or loss score) during the training process to adjust the network parameters. In this work, we have examined multiple optimization algorithms like Adam, RMSprop, SGD, AdaGrad, AdaDelta and AdaMax [47] and compared their performance.

In the first set of experiments, we evaluated the performance using different optimizers to select the best network parameters for diagnosis with the LeakyReLU activation function at $\alpha = 0.1$ and η of 10^{-4} . The overall result accuracies for different optimizers are shown in Table 2. As can be readily seen, the performance measures of the proposed model for COVID-19 classification depends on the employed optimizer. Additionally, Table 3 summarizes the class-wise performance metrics of each optimizer in detail, and Figures 5 and 6 represent the confusion matrices and the ROC curves for different optimizers.

Similarly, the system's performance against different optimizers using the ELU function has been also studied. Table 4 illustrates the performance metrics of the proposed model using ELU at $\alpha = 0.2$, and the associated confusion matrices are shown in Figure 7. Furthermore, Figure 8 displays the ROC curves for the examined optimizers, and the detailed class-wise performance metrics are given in Table 5.

Table 2. Overall evaluation of the proposed model with different optimizers and the LeakyReLU activation function ($\alpha = 0.1$).

Optimizer	ACC	PER	REC	F1-Score	Losses	tp (sec)
AdaGrad	72.4%	84%	55%	53%	0.67	842
AdaDelta	83.9%	86.3%	79.6%	80.3%	0.44	841
SGD	91.8%	91%	92%	91%	0.25	890
Adam	95%	95%	98%	96%	0.08	845
AdaMax	97.4%	96.3%	98.3%	76%	0.07	847
RMSprop	99.3%	99%	99%	99.3%	0.02	862

Table 3. Comparative analysis of each class and overall accuracy using different optimizers with the LeakyReLU activation function set at $\alpha = 0.1$.

Optimizer	Metrics	Class		
		COVID-19	Normal	Pneumonia
AdaGrad	Precision	1.00	0.83	0.70
	REC	0.02	0.68	0.95
	F1-Score	0.04	0.74	0.81
	Accuracy	100%	82.5%	70.2%
AdaDelta	PER	1.00	0.73	0.86
	REC	0.57	0.93	0.89
	F1-Score	0.73	0.81	0.87
	ACC	100%	72.5%	86%
SGD	PER	0.99	0.79	0.95
	REC	0.90	0.95	0.91
	F1-Score	0.94	0.86	0.93
	ACC	99.1%	79%	95.3%
Adam	PER	1.00	0.81	1.00
	REC	0.99	1.00	0.92
	F1-Score	1.00	0.90	0.96
	ACC	100%	81.4%	100%
AdaMax	PER	1.00	0.89	1.00
	REC	1.00	0.99	0.96
	F1-Score	1.00	0.94	0.98
	ACC	100%	89.4%	99.7%
RMSprop	PER	1.00	0.97	1.00
	REC	0.99	0.99	0.99
	F1-Score	1.00	0.98	1.00
	ACC	100%	97.4%	99.7%

In addition to the evaluation of our system using different activations and optimizers, we evaluated the performance of our system against other DL methods. Particularly, we applied different pre-trained networks such as Inception V3, DenseNet121, ResNet50, MobilNet, and VGG16 on our dataset. The class identification was done by using the softmax layer of the pre-trained networks. Table 6 displays the performance of each model with all evaluation metrics and TP. Namely, the proposed model achieved a high validation accuracy of 99.3% compared to other traditional pre-trained models. Figure 9 presents the confusion matrices for different pre-trained models, and Figure 10 demonstrates their respective ROC curves. It is worth mentioning that our approach achieved high accuracy

with the advantage of being shallow (unlike the deep pre-trained models) with a low training burden (10 epochs). Our approach also investigated using different activation functions and optimizers. In addition, our approach results are promising and are on a comparable bar with other DL studies for COVID-19 detection. The results of other studies on DL with the details of their respective datasets are illustrated in Table 7.

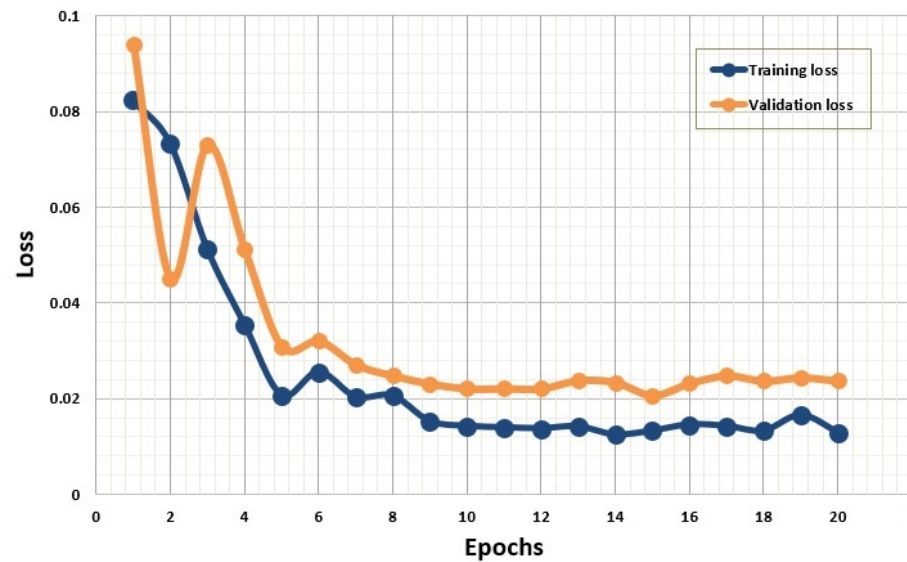


Figure 4. The loss curves of the proposed model.

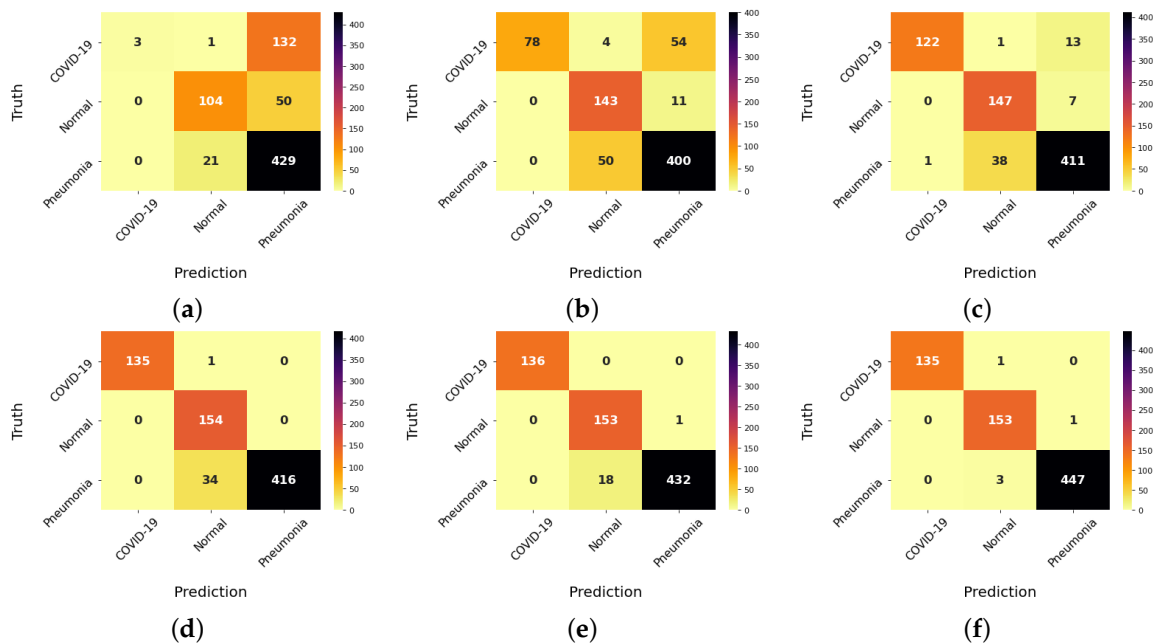


Figure 5. Confusion matrices of the proposed model with different optimizers using the LeakyReLU activation function set at $(\alpha = 0.1)$. (a) AdaGrad, (b) AdaDelta, (c) SGD, (d) Adam, (e) AdaMax, and (f) RMSprop.

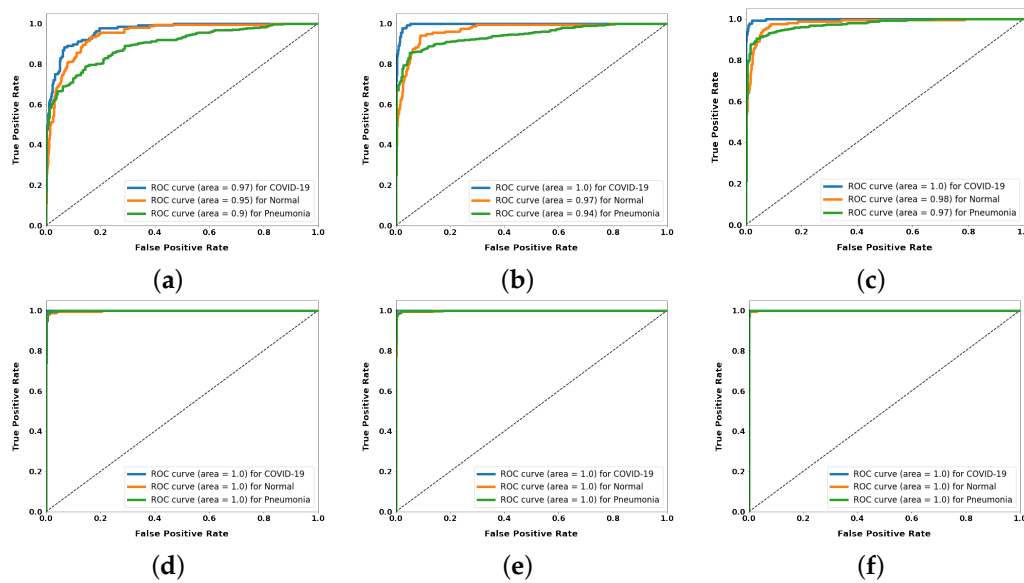


Figure 6. ROC curves of the proposed model with different optimizers using the LeakyReLU activation function set at ($\alpha = 0.1$). (a) AdaGrad, (b) AdaDelta, (c) SGD, (d) Adam, (e) AdaMax, and (f) RMSprop.

Table 4. Overall evaluation of the proposed model with different optimizers and ELU activation function ($\alpha = 0.2$)

Optimizer	ACC	PER	REC	F1-Score	Losses	tp (sec)
AdaGrad	73.5%	83.6%	57%	58%	0.66	859
AdaDelta	86%	87%	84%	84.3%	0.44	845
SGD	91.8%	90.6%	92.3%	91.3%	0.93	832
RMSprop	93.5%	92%	95.6%	93%	0.7	856
Adam	94%	92%	96%	94%	0.34	845
AdaMax	98.5%	98%	99%	98.6%	0.04	856

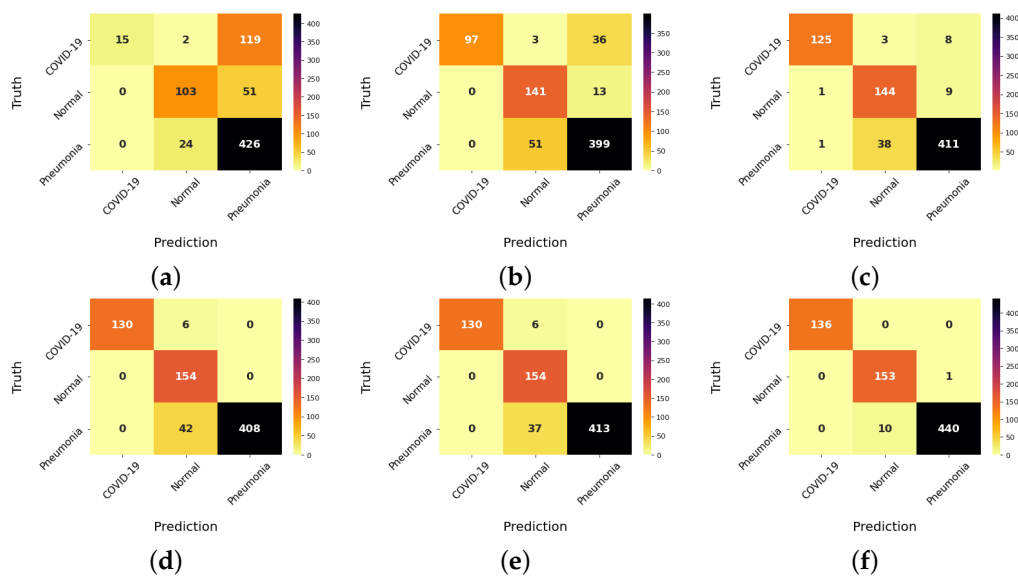


Figure 7. Confusion matrices of the proposed model using the ELU activation function ($\alpha = 0.2$) with different optimizers. (a) AdaGrad, (b) AdaDelta, (c) SGD, (d) RMSprop, (e) Adam, and (f) AdaMax.

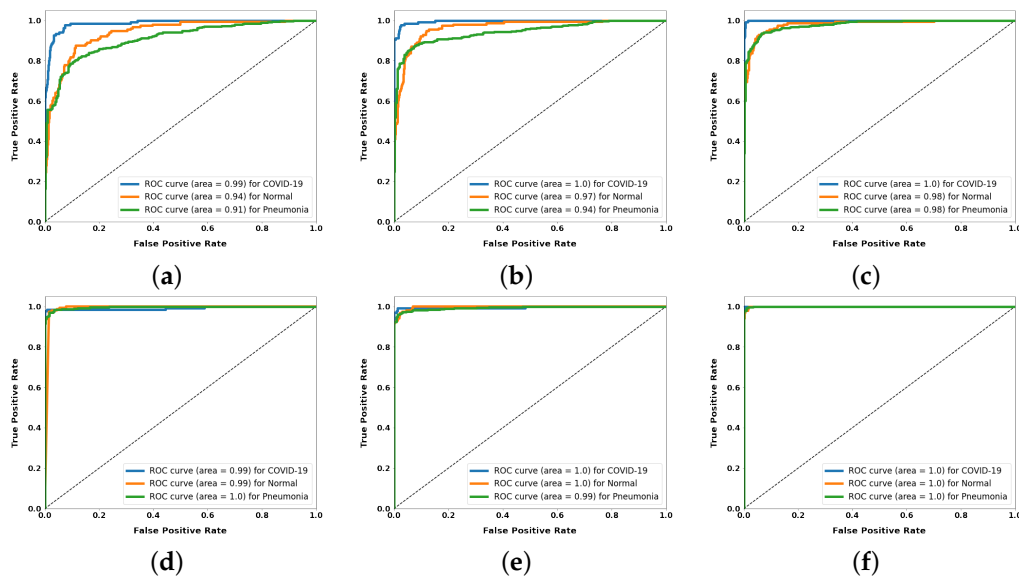


Figure 8. ROC curves of the proposed model using the ELU activation function ($\alpha = 0.2$) with different optimizers (a) AdaGrad, (b) AdaDelta, (c) SGD, (d) RMSprop, (e) Adam, and (f) AdaMax.

Table 5. Comparative analysis of each class and overall accuracy using different optimizers with the ELU activation function set at $\alpha = 0.2$.

Optimizer	Metrics	Class		
		COVID-19	Normal	Pneumonia
AdaGrad	PER	1.00	0.80	0.71
	REC	0.11	0.67	0.95
	F1-Score	0.20	0.73	0.81
	ACC	100%	79.8%	71.4%
AdaDelta	PER	1.00	0.72	0.89
	REC	0.71	0.92	0.89
	F1-Score	0.83	0.81	0.89
	ACC	100%	72.3%	89%
SGD	PER	0.98	0.78	0.96
	REC	0.92	0.94	0.91
	F1-Score	0.95	0.85	0.94
	ACC	98.4%	77.8%	96%
RMSprop	PER	1.00	0.76	1.00
	REC	0.96	1.00	0.91
	F1-Score	0.98	0.87	0.95
	ACC	100%	76.2%	100%
Adam	PER	1.00	0.78	1.00
	REC	0.96	1.00	0.92
	F1-Score	0.98	0.88	0.96
	ACC	100%	78.1%	100%
AdaMax	PER	1.00	0.94	1.00
	REC	1.00	0.99	0.98
	F1-Score	1.00	0.97	0.99
	ACC	100%	93.8%	99.7%

Table 6. Comparative accuracy of our model with other CNN models. Learning rate $\eta = 0.1$, ReLU activation function, and Adam optimizer for all pre-trained models.

Model	ACC	PER	REC	F1-Score	Losses	tp (sec)
VGG16	60.8%	61%	100%	76%	0.94	690
Xception	61%	61%	100%	76%	0.92	853
DenseNet121	72%	77%	79%	74%	0.58	693
InceptionV3	72.8%	80%	85%	81%	0.84	658
ResNet50	87%	84%	87%	85%	0.33	693
MobilNet	95.4%	77%	79%	74%	0.14	633
Proposed model	99.3%	99%	99%	99.3%	0.02	862

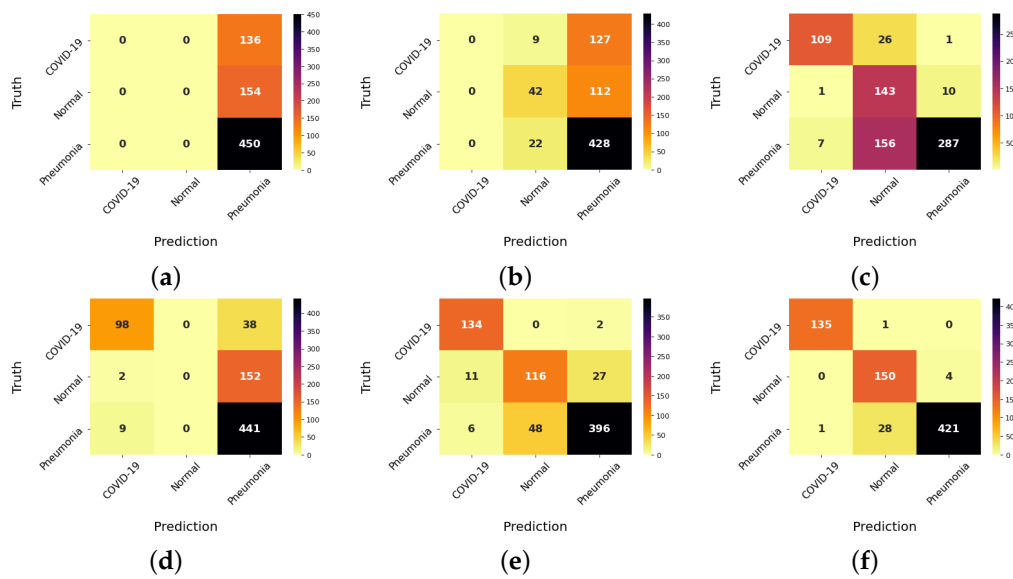


Figure 9. Confusion matrices of the pre-trained models with learning rate $\eta = 0.1$, ReLU activation function, and Adam optimizer. (a) VGG16, (b) Xception, (c) DenseNet121, (d) InceptionV3, (e) ResNet50, and (f) MobilNet.

Table 7. Previous works for COVID-19 classification. Here, MCSVM, CSEN, ML, TL, CV, and ACC stand for multi-class support vector machine, convolution support estimation network, machine learning, transfer learning, cross-validation, and accuracy, respectively.

Study	# Images	Data Split (Train:Test:Val)	# Classes	Model	ACC
Barshooi et al. [23]	4560	(60%:25%:15%)	2	DenseNet201	98.5%
Rehman et al. [24]	3000	5 and 10-fold CV	15	TL + ML classifiers	99.40%, 99.77%
Brima et al. [25]	21,165	(72%:10%:18%)	4	ResNet50	94%
Albahli et al. [26]	12,357	10-fold CV	16	ResNet50 + semantic segmentation	71.905%, 92.52%
Manokaran et al. [27]	10,373	(70%:20%:10%)	3	DenseNet201	92.19%
Yang et al. [32]	8461	(60%:20%:20%)	3	VGG16	99%
Mojab et al. [48]	2500	(64%:20%:16%)	2	EpistoNet	95%
Yama et al. [15]	6286	(80%:20%:0)	4	CSEN	95.4%
Proposed model	7395	(80%:10%:10%)	3	Xception + GAP& activation layers	99.3%

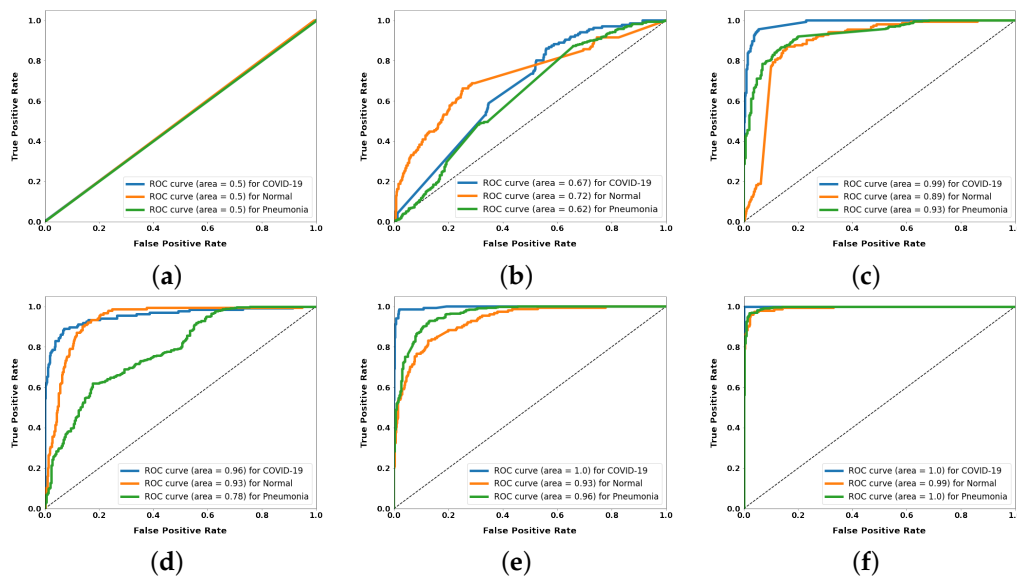


Figure 10. ROC curves of the pre-trained models with learning rate $\eta = 0.1$, ReLU activation function, and Adam optimizer. (a) VGG16, (b) Xception, (c) DenseNet121, (d) InceptionV3, (e) ResNet50, and (f) MobilNet.

6. Discussion

The recent statistical data from WHO and CDC has shown a significant increase in COVID-19 cases, hospitalizations, and deaths in the first week of January 2022. In the US, 98% of cases are caused by the spread of the Omicron variant of the Coronavirus [49]. Therefore, the early prediction of COVID-19 is of immense importance to help in avoiding the disease spread and thus protecting immune-vulnerable people. As in many other diseases, early detection can be very helpful for diagnosis and evaluation, as well as for follow-up on disease progression. Thus, an automated system that utilizes patient data could be a valuable decision support tool for COVID-19 detection and provide significant downstream treatment/follow-up implications.

For that purpose, we have developed a deep learning system (DLS) to predict COVID-19. The proposed DLS is validated experimentally on a benchmark dataset of chest X-ray images, which contains three classes: COVID-19, normal control, and pneumonia. The proposed pipeline utilized a pre-trained Xception model and modified the network structure by including the global average pooling layer. Then, we used the layer to solve the vanishing gradient problem and reduce the probability of overfitting; the activation layer was used for reducing losses. Furthermore, we have explored different activation functions at different thresholds to improve performance and reduce the losses of the proposed pipeline. Additionally, extensive evaluation has been conducted using different optimizers.

Developing an analysis pipeline with high accuracy is our ultimate goal, for which various experiments have been conducted to evaluate the performance using various evaluation metrics, such as precision, recall, accuracy, F1-score, and area under the ROC. Before performing any experiments, data augmentation (i.e., rotation by 25 degrees, a zoom range equal to 0.2 and fill mode set to nearest) has been applied before feeding the data into the DLS to overcome the dataset imbalance.

Experimental results have revealed that the proposed DLS performance can be improved by carefully employing a nonlinear activation function and an optimization method. Particularly, the proposed model achieved an accuracy of 99.3% and a minimum loss of 0.02 using the combination of the RMSprop optimizer and LeakyReLU with $\alpha = 0.1$. Other optimization techniques demonstrated closer results over shorter times (e.g., AdaMax). This has been documented using overall and class-wise accuracies, ROC curves and confusion matrices. Similar experiments were conducted using the ELU activation at $\alpha = 0.2$. As

Figure 7 showed, the COVID-19 class samples (i.e., 130) were correctly classified, while six normal cases resulted in negative and positive prediction errors. Overall, the LeakyReLU activation function at a threshold of α equal 0.1 combined with the RMSprop optimizer is the best case (see Tables 2 and 3 and Figures 5 and 6). It is worth mentioning that, additionally, the use of TL helped our model to start from good initial parameters, and few training epochs were needed to achieve better performance. Particularly, during model training, the loss/accuracy curves for the training vs validation showed that the validation accuracy became high with small losses after a few epochs (=10).

Additional comparisons between the developed DLS and other pre-trained deep learning-based methods, including Inception V3, DenseNet121, ResNet50, MobilNet, and VGG16, have been conducted. We unified the parameters for all systems by using the ReLU activation function, $\eta = 0.1$, and the Adam optimizer. The results shown in Table 6 and Figures 9 and 10 emphasize the benefits of the proposed DLS, especially its advantage of being shallow (unlike the deep pre-trained models) with a low training burden (10 epochs). The comparative results with other deep learning systems, tested on different COVID datasets, highlighted that although the proposed model is shallow, it provides high accuracy in a short time when tested on CXR images. The above results document that the proposed DLS system can be helpful in healthcare systems in several ways, such as decreasing consultation-associated costs, enhancing detection consistency, and thus lowering the risk of disease spread.

Despite the promising results, our analysis pipeline has some limitations. Firstly, we trained and tested our approach from a single benchmark dataset. Thus, using external test datasets from other different centers should further enhance the robustness of our automated DLS. Secondly, our DLS only utilized a single input data type: chest X-ray images. In clinical practice, multiple data sources are usually used, such as clinical biomarkers and, sometimes, chest CT scans, which can provide more localized views of the affected lungs. Recent studies, e.g., [28–32], explored the various DL methods available to diagnose COVID-19 based on chest CT images. The studies have documented high accuracy rates for both binary and multi-class classification tasks. Therefore, an update to the proposed model could integrate multiple inputs and give a more robust prediction at the patient level. Finally, our system concentrated on two activation functions without further investigation of the effect of α or other activation functions.

7. Conclusions and Future Work

This paper has introduced a multi-level X-ray-based diagnostic system for the accurate detection of COVID-19. The proposed analysis pipeline is a deep learning system (DLS) that utilizes transfer learning. The potential of the proposed framework is documented by the presented results to accurately identify COVID-19 cases in a cohort of 7395 CXR images. Extensive evaluation of the model parameters' optimization is conducted to improve performance and reduce the losses of the proposed model. The accuracy of our system is comparable to other deep learning-based methods for COVID-19 classification on various X-ray datasets. The proposed algorithm has great potential to be used in clinical applications and can benefit front-line medical personnel for accurate and rapid COVID diagnosis. In future work, we will extend our system's ability to fuse multi-input data, including other imaging modalities (i.e., chest CT) and clinical markers, to provide more comprehensive markers (images as well as clinical) that can help physicians in providing more appropriate personalized medicine. In addition, we will try to test our model with different and more diverse datasets.

Author Contributions: Conceptualization: E.A., M.A.M. and F.K.; methodology, formal analysis, and visualization: A.A.A., E.A. and F.K.; project administration; E.A., M.A. and F.K.; supervision: E.A., M.A.M.; writing—original draft: A.A.A., E.A. and F.K.; writing—review & editing: E.A., M.A.M. and F.K.; software and data curation: A.A.A. All authors have read and agreed to the submitted version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not Applicable: This study used publicly available datasets.

Informed Consent Statement: Not Applicable: This study used publicly available datasets.

Data Availability Statement: The datasets used in this work are publicly available.

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

References

1. WHO Coronavirus (COVID-19) Dashboard. Available online: <https://covid19.who.int/> (accessed on 28 January 2022).
2. Narin, A.; Kaya, C.; Pamuk, Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *Pattern Anal. Appl.* **2021**, *24*, 1–14. [[CrossRef](#)]
3. Majhi, B.; Thangeda, R.; Majhi, R. A Review on Detection of COVID-19 Patients Using Deep Learning Techniques. In *Assessing COVID-19 and Other Pandemics and Epidemics using Computational Modelling and Data Analysis*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 59–74.
4. Kanne, J.P.; Little, B.P.; Chung, J.H.; Elicker, B.M.; Ketaj, L.H. Essentials for Radiologists on COVID-19: An update—Radiology Scientific Expert Panel. *Radiology*. **2020**, *296*, E113–E114. [[CrossRef](#)] [[PubMed](#)]
5. 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](#)]
6. Van Kasteren, P.B.; van Der Veer, B.; van den Brink, S.; Wijsman, L.; de Jonge, J.; van den Brandt, A.; Molenkamp, R.; Reusken, C.B.; Meijer, A. Comparison of seven commercial RT-PCR diagnostic kits for COVID-19. *J. Clin. Virol.* **2020**, *128*, 104412. [[CrossRef](#)] [[PubMed](#)]
7. Apostolopoulos, I.D.; Aznaouridis, S.I.; Tzani, M.A. Extracting possibly representative COVID-19 Biomarkers from X-Ray images with Deep Learning approach and image data related to Pulmonary Diseases. *J. Med Biol. Eng.* **2020**, *40*, 462–469. [[CrossRef](#)] [[PubMed](#)]
8. 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*, 1. [[CrossRef](#)]
9. Akter, S.; Shamrat, F.; Chakraborty, S.; Karim, A.; Azam, S. COVID-19 detection using deep learning algorithm on chest X-ray images. *Biology* **2021**, *10*, 1174. [[CrossRef](#)] [[PubMed](#)]
10. 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](#)]
11. Basu, S.; Mitra, S.; Saha, N. Deep learning for screening covid-19 using chest X-ray images. In Proceedings of the 2020 IEEE Symposium Series on Computational Intelligence (SSCI), Canberra, ACT, Australia, 1–4 December 2020; pp. 2521–2527.
12. Zheng, C.; Deng, X.; Fu, Q.; Zhou, Q.; Feng, J.; Ma, H.; Liu, W.; Wang, X. Deep learning-based detection for COVID-19 from chest CT using weak label. *IEEE Trans. Med Imaging* **2020**, *39*, 2615–2625.
13. Abbas, A.; Abdelsamea, M.M.; Gaber, M.M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Appl. Intell.* **2021**, *51*, 854–864. [[CrossRef](#)]
14. Rahaman, M.M.; Li, C.; Yao, Y.; Kulwa, F.; Rahman, M.A.; Wang, Q.; Qi, S.; Kong, F.; Zhu, X.; Zhao, X. Identification of COVID-19 samples from chest X-ray images using deep learning: A comparison of transfer learning approaches. *J. X-ray Sci. Technol.* **2020**, *28*, 821–839. [[CrossRef](#)] [[PubMed](#)]
15. Yamaç, M.; Ahishali, M.; Degerli, A.; Kiranyaz, S.; Chowdhury, M.E.; Gabbouj, M. Convolutional Sparse Support Estimator-Based COVID-19 Recognition From X-Ray Images. *IEEE Trans. Neural Networks Learn. Syst.* **2021**, *32*, 1810–1820. [[CrossRef](#)] [[PubMed](#)]
16. Loey, M.; Smarandache, F.; Khalifa, N.E.M. Within the lack of COVID-19 benchmark dataset: A novel gan with deep transfer learning for corona-virus detection in chest X-ray images. *Symmetry* **2020**, *12*, 651. [[CrossRef](#)]
17. Khan, A.I.; Shah, J.L.; Bhat, M.M. CoroNet: A deep neural network for detection and diagnosis of COVID-19 from chest x-ray images. *Comput. Methods Programs Biomed.* **2020**, *196*, 105581. [[CrossRef](#)] [[PubMed](#)]
18. Asif, S.; Wenhui, Y.; Jin, H.; Jinhai, S. Classification of COVID-19 from Chest X-ray images using Deep Convolutional Neural Network. In Proceedings of the 2020 IEEE 6th International Conference on Computer and Communications (ICCC), Chengdu, China, 11–14 December 2020; pp. 426–433.
19. Toraman, S.; Alakus, T.B.; Turkoglu, I. Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Solitons Fractals* **2020**, *140*, 110122. [[CrossRef](#)]
20. Sethy, P.K.; Behera, S.K. Detection of Coronavirus Disease (COVID-19) Based on Deep Features. 2020. Available online: <https://pdfs.semanticscholar.org/9da0/35f1d7372cfe52167ff301bc12d5f415caf1.pdf> (accessed on 28 January 2022)
21. Sarker, L.; Islam, M.M.; Hannan, T.; Ahmed, Z. COVID-DenseNet: A Deep Learning Architecture to Detect COVID-19 from Chest Radiology Images. 2020. Available online: <https://pdfs.semanticscholar.org/c6f7/a57a37e87b52ac92402987c9b7a3df41f2db.pdf> (accessed on 28 January 2022)
22. Das, N.N.; Kumar, N.; Kaur, M.; Kumar, V.; Singh, D. Automated deep transfer learning-based approach for detection of COVID-19 infection in chest X-rays. *IRBM* **2020**, *in press*. [[CrossRef](#)]

23. Barshooi, A.H.; Amirkhani, A. A novel data augmentation based on Gabor filter and convolutional deep learning for improving the classification of COVID-19 chest X-Ray images. *Biomed. Signal Process. Control.* **2022**, *72*, 103326. [CrossRef]
24. Rehman, N.u.; Zia, M.S.; Meraj, T.; Rauf, H.T.; Damaševičius, R.; El-Sherbeeney, A.M.; El-Meligy, M.A. A self-activated cnn approach for multi-class chest-related COVID-19 detection. *Appl. Sci.* **2021**, *11*, 9023. [CrossRef]
25. Brima, Y.; Atemkeng, M.; Tankio Djikap, S.; Ebiele, J.; Tchakounté, F. Transfer Learning for the Detection and Diagnosis of Types of Pneumonia including Pneumonia Induced by COVID-19 from Chest X-ray Images. *Diagnostics* **2021**, *11*, 1480. [CrossRef]
26. Albahli, S.; Yar, G.N.A.H. Fast and Accurate Detection of COVID-19 Along With 14 Other Chest Pathologies Using a Multi-Level Classification: Algorithm Development and Validation Study. *J. Med Internet Res.* **2021**, *23*, e23693. [CrossRef]
27. Manokaran, J.; Zabihollahy, F.; Hamilton-Wright, A.; Ukwatta, E. Detection of COVID-19 from chest x-ray images using transfer learning. *J. Med Imaging* **2021**, *8*, 017503. [CrossRef] [PubMed]
28. Khan, M.A.; Alhaisoni, M.; Tariq, U.; Hussain, N.; Majid, A.; Damaševičius, R.; Maskeliūnas, R. COVID-19 case recognition from chest CT images by deep learning, entropy-controlled firefly optimization, and parallel feature fusion. *Sensors* **2021**, *21*, 7286. [CrossRef]
29. Wang, S.H.; Zhang, X.; Zhang, Y.D. DSSAE: Deep stacked sparse autoencoder analytical model for COVID-19 diagnosis by fractional Fourier entropy. *ACM Trans. Manag. Inf. Syst. (TMIS)* **2021**, *13*, 1–20. [CrossRef]
30. Scarpiniti, M.; Ahrabi, S.S.; Baccarelli, E.; Piazzi, L.; Momenzadeh, A. A novel unsupervised approach based on the hidden features of Deep Denoising Autoencoders for COVID-19 disease detection. *Expert Syst. Appl.* **2021**, *192*, 116366. [CrossRef] [PubMed]
31. Song, Y.; Zheng, S.; Li, L.; Zhang, X.; Zhang, X.; Huang, Z.; Chen, J.; Wang, R.; Zhao, H.; Zha, Y.; et al. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **2021**, *18*, 2775–2780. [CrossRef]
32. Yang, D.; Martinez, C.; Visuña, L.; Khandhar, H.; Bhatt, C.; Carretero, J. Detection and analysis of COVID-19 in medical images using deep learning techniques. *Sci. Rep.* **2021**, *11*, 1–13. [CrossRef]
33. Cheng, H.D.; Shi, X. A simple and effective histogram equalization approach to image enhancement. *Digit. Signal Process.* **2004**, *14*, 158–170. [CrossRef]
34. Kandel, I.; Castelli, M. Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review. *Appl. Sci.* **2020**, *10*, 2021. [CrossRef]
35. 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]
36. Liu, S.; Ou, X.; Che, J.; Zhou, X.; Ding, H. An Xception-GRU Model for Visual Question Answering in the Medical Domain. CLEF (Working Notes). 2019. Available online: http://www.dei.unipd.it/~ferro/CLEF-WN-Drafts/CLEF2019/paper_127.pdf (accessed on 28 January 2022).
37. Mulligan, K.; Rivas, P. Dog breed identification with a neural network over learned representations from the xception cnn architecture. In Proceedings of the 21st International Conference on Artificial Intelligence (ICAI'19), Luxor, Las Vegas, NV, USA, 29 July–1 August 2019; pp. 246–249.
38. Patil, S.; Golellu, A. Classification of COVID-19 CT Images using Transfer Learning Models. In Proceedings of the 2021 International Conference on Emerging Smart Computing and Informatics (ESCI), Pune, India, 5–7 March 2021; pp. 116–119.
39. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, 17–19 June 2017; pp. 1251–1258.
40. Higashinaka, R.; Funakoshi, K.; Kobayashi, Y.; Inaba, M. The dialogue breakdown detection challenge: Task description, datasets, and evaluation metrics. In Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16), Portorož, Slovenia, 23–28 May 2016; pp. 3146–3150.
41. Hossin, M.; Sulaiman, M.N. A review on evaluation metrics for data classification evaluations. *Int. J. Data Min. Knowl. Manag. Process.* **2015**, *5*, 1.
42. Ahmed, M.; Al-qaysi, Z.; Shuwandy, M.L.; Salih, M.M.; Ali, M.H. Automatic COVID-19 pneumonia diagnosis from x-ray lung image: A Deep Feature and Machine Learning Solution. *J. Phys. Conf. Ser.* **2021**, *1963*, 012099. [CrossRef]
43. COVID-19 Radiography Database. Available online: <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database/> (accessed on 6 January 2021).
44. Chest X-ray (Covid-19 & Pneumonia). Available online: <https://www.kaggle.com/prashant268/chest-xray-covid19-pneumonia/> (accessed on 6 January 2021).
45. COVID-19 & Normal Posteroanterior(PA) X-rays. Available online: <https://www.kaggle.com/tarandeep97/covid19-normal-posteroanteriorpa-xrays/> (accessed on 6 January 2021).
46. COVID-19 Patients Lungs X Ray Images 10000. Available online: <https://www.kaggle.com/nabeelsajid917/covid-19-x-ray-10000-images/> (accessed on 6 January 2021).
47. Bera, S.; Shrivastava, V.K. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. *Int. J. Remote. Sens.* **2020**, *41*, 2664–2683. [CrossRef]
48. Mousavi Mojab, S.Z.; Shams, S.; Fotouhi, F.; Soltanian-Zadeh, H. EpistoNet: An ensemble of Epistocracy-optimized mixture of experts for detecting COVID-19 on chest X-ray images. *Sci. Rep.* **2021**, *11*, 1–13. [CrossRef]
49. Centers for Disease Control and Prevention: Covid Data Tracker Weekly Review. Available online: <https://www.cdc.gov/coronavirus/2019-ncov/covid-data/covidview/index.html/> (accessed on 28 January 2022).