



# OPEN COVID-19 detection from chest X-ray images using transfer learning

Enas M. F. El Houby

COVID-19 is a kind of coronavirus that appeared in China in the Province of Wuhan in December 2019. The most significant influence of this virus is its very highly contagious characteristic which may lead to death. The standard diagnosis of COVID-19 is based on swabs from the throat and nose, their sensitivity is not high enough and so they are prone to errors. Early diagnosis of COVID-19 disease is important to provide the chance of quick isolation of the suspected cases and to decrease the opportunity of infection in healthy people. In this research, a framework for chest X-ray image classification tasks based on deep learning is proposed to help in early diagnosis of COVID-19. The proposed framework contains two phases which are the pre-processing phase and classification phase which uses pre-trained convolution neural network models based on transfer learning. In the pre-processing phase, different image enhancements have been applied to full and segmented X-ray images to improve the classification performance of the CNN models. Two CNN pre-trained models have been used for classification which are VGG19 and EfficientNetB0. From experimental results, the best model achieved a sensitivity of 0.96, specificity of 0.94, precision of 0.9412, F1 score of 0.9505 and accuracy of 0.95 using enhanced full X-ray images for binary classification of chest X-ray images into COVID-19 or normal with VGG19. The proposed framework is promising and achieved a classification accuracy of 0.935 for 4-class classification.

**Keywords** Classification, Convolution neural network, Coronavirus, COVID-19, Deep learning, Transfer learning

Since December 2019, coronavirus has been disseminated from China to many other countries. Coronavirus which is called SARS-CoV-2 causes COVID-19 as named by World Health Organization (WHO) on February 11, 2020. World Health Organization announced COVID-19 disease resulted from the coronavirus as a world pandemic in March 2020<sup>1</sup>. The disease has disseminated to nearly all countries, resulting in millions of people's deaths among confirmed cases based on the statistics of the WHO<sup>2</sup>. By July 2023, nearly 700 million confirmed cases, and almost 7 million confirmed deaths were recorded in the world<sup>3,4</sup>. Most patients with the virus experience mild to moderate respiratory illness and heal without needing special treatment. But, some suffer from complications and need medical attention. Older people and those with underlying medical conditions like cardiovascular disease, diabetes, chronic respiratory disease, or cancer are more likely to develop serious illnesses. Anyone can get sick with COVID-19 and become seriously ill or die at any age<sup>5</sup>.

Although the last diagnosis of COVID-19 depends on transcription-polymerase chain reaction (PCR) tests, in states of people with intensive respiratory symptoms the diagnosis protocol depends on medical imaging, which helps doctors to recognize the disease as the sensitivity of PCR is strongly variable<sup>6</sup>. As chest radiography imaging such as computed tomography (CT) imaging and X-ray have been used successfully for the diagnosis of pneumonia, they have a high sensitivity for the diagnosis of COVID-19<sup>2</sup>. The suspected case undergoes an X-Ray session and if more details are required, a computed tomography scan (CT-scan) session is taken. Therefore, X-ray<sup>7</sup> and CT scan images<sup>8</sup> are being used as diagnostic methods for COVID-19 and to detect the effects<sup>9</sup> of the virus<sup>6,10</sup>. The availability and accessibility of X-ray imaging in many imaging centers and clinics is more present even in rural regions as it is standard equipment in healthcare systems. Particularly, chest X-ray is more readily available than CT, because CT scanners require high equipment and maintenance costs. CT is not very suitable for COVID-19 screening as well because of its cost, imaging time, and radiation exposure whereas X-ray is more cost and time effective in dealing with such a common virus<sup>11</sup>.

Systems and Information Department, National Research Centre, Dokki 12311, Cairo, Egypt. email: enas\_mfahmy@yahoo.com; em.fahmy@nrc.sci.eg

The abnormalities can only be explained by expert radiologists. With the huge number of suspected cases and the limited number of available radiologists, automatic methodologies for the recognition of these precise abnormalities can aid in early diagnosis with high accuracy. The studies in Artificial Intelligence (AI) and machine learning, especially Deep Learning (DL), achieved high performance in the diagnosis of medical images. Therefore, DL techniques are robust tools for such issues.

Deep learning (DL) has been successfully used to predict COVID-19 from Chest images. Unlike the traditional machine learning techniques DL can be used to predict disease from raw images without feature extraction required. The role of deep learning is to learn the features using a trained model with a huge amount of data to improve the classification's accuracy which reduces the burden on physicians and decreases the effect of doctors' shortages of the struggle against the disease. Convolutional neural network (CNN) is the type of DL model intended for image analysis tasks and has already been utilized in many medical problems such as segmentation and classification<sup>12</sup>.

Many high-performing pre-trained CNN structures have been provided in the literature to be utilized in similar problems. These models were trained using ImageNet data which contains 1,000,000 images and 1000 classes to overcome the limitation of data and to reduce the training time<sup>13</sup>. These models can be used for image recognition based on transfer learning after fine-tuning these networks to the new problems. The learned weights of these pre-trained models are provided and used directly in the new problems<sup>14</sup>. The purpose of utilizing pre-trained models is to take advantage of learned features on a larger dataset, therefore the new model can converge faster and perform better with a smaller dataset. This gives us the advantage of DL independence of feature engineering over traditional methods without giving up the time, computational resources and cost efficiencies. Examples of these pre-trained CNN models are visual geometry group VGG (16, 19)<sup>15</sup>, EfficientNet (B0 to B7)<sup>16</sup>, MobileNet<sup>17</sup>, and residual neural network (ResNet)<sup>18</sup>, etc.

The contributions of this research can be summarized as follows:

- A framework has been developed to diagnose COVID-19 using chest X-ray images for both full and segmented images.
- A multiplication between each original image and the associated lung mask from the ground truth dataset provided by the database has been applied to get the segmented lung.
- Different image enhancement techniques have been applied to both full and segmented X-ray images to reach the best possible classification performance.
- CNN pre-trained models based on transfer learning have been used to classify both full and segmented chest X-ray images with all enhancement versions and achieved promising results.
- Since the purpose of utilizing pre-trained models is to take advantage of learned features on a larger dataset, therefore the smallest possible datasets that can achieve the best possible performance have been used for faster convergence.

Recently, many works have been developed to detect and diagnose COVID-19 and other lung diseases based on different medical image modalities using different machine learning techniques especially deep learning and transfer learning techniques. The purpose of all these works is to improve the performances of the methodologies used in the detection and classification of COVID-19 and other lung diseases. The focus in the research will be in X-ray images as the adopted medical image modality in this research.

The rest of the paper is organized as follows. Related COVID-19 articles using deep learning are reviewed in the "[Related work](#)" section. Then, the proposed framework for COVID-19 classification is described in "[The proposed framework](#)" section. Next, the results of X-ray images obtained with the proposed framework are presented in "[Experimental results](#)" section. The discussion and comparison with literature are provided in "[Discussion](#)" section. Finally, the main "[Conclusions and future work](#)" are outlined.

## Related work

Recently, many works have been developed to detect and diagnose COVID-19 and other lung diseases based on different medical image modalities using different machine learning techniques especially deep learning and transfer learning. The purpose of all these works is to improve the performances of the methodologies used in the detection and classification of COVID-19 and other lung diseases. Where the proposed research will use X-ray images as a medical image modality, the focus in this section will be on the previous work based on X-rays.

Nishio, et al.<sup>19</sup> presented a system based on VGG16 to classify images of chest X-rays as healthy, COVID-19 pneumonia, and non-COVID-19 pneumonia. They applied the proposed system to 1248 X-ray images collected from 2 different public datasets. The collected X-ray images contain 500 healthy samples, 215 images for COVID-19 pneumonia patients and 533 images for non-COVID-19 pneumonia patients. The achieved accuracy was 83.6%, while the sensitivity was 90.9%.

Minaee et al.<sup>20</sup> applied deep learning to recognize COVID-19 cases using chest X-rays images. Transfer learning was used to train 4 CNN models which are DenseNet-121, SqueezeNet, ResNet50, and ResNet18 to binary classify images as COVID-19 or not. The training was applied to 84 (420 after augmentation) COVID-19 images and 2000 non-Covid images, while the test was applied to 100 COVID-19 images and 3000 non-COVID images. The best achieved sensitivity of these models was 98%, while the specificity was 92.9% for the SqueezeNet model.

Sahin<sup>21</sup> proposed a CNN model for binary classification of COVID-19 cases as COVID and Normal using chest X-ray images. Also, two pre-trained models which are ResNet50 and MobileNetv2 are applied to the used dataset of 13,824 X-ray images. The proposed CNN model achieved an accuracy of 96.71% and F1-score of

97%. MobileNetv2 achieved an accuracy of 95.73% and F1-score of 96%, while ResNet50 achieved an accuracy of 91.54% and F1-score of 91%.

Wang et al.<sup>22</sup> developed an open-source CNN called COVID-Net to detect COVID-19 cases using chest X-ray images. The proposed net can predict the case as one of three classes which are COVID-19 viral infection, non-COVID-19 infection, and normal. Also, an open access benchmark dataset COVIDx was introduced, it contains 13,975 X-ray images collected from 13,870 patients. The COVIDx dataset was generated using five different publically available datasets. The accuracy of COVID-Net reached 93.3%.

Panwar et al.<sup>23</sup> developed a deep learning model called nCOVnet for detecting COVID-19 based on X-rays. A dataset of 284 X-ray images was used of which 142 images are normal cases and 142 images are COVID-19 cases. The model achieved an accuracy of 88.1%.

Nigam et al.<sup>24</sup> used transfer learning to utilize 5 pre-trained models which are DenseNet121, NASNet, Xception, VGG16, and EfficientNet to classify Coronavirus suspected cases as normal, COVID-19 positive cases, and other classes. The used dataset contains 16,634 X-ray images, 6000 normal images, 5634 COVID images, and 5000 images for others. The achieved accuracies were 79.01%, 85.03%, 88.03%, 89.96%, and 93.48% for VGG16, NASNet, Xception, DenseNet121, and EfficientNet respectively.

Chow et al.<sup>25</sup> used transfer learning to utilize 18 CNN models including VGG-19, VGG-16, ShufNet, SqueezeNet. etc. to classify the cases as normal or COVID-19. The used dataset contains 700 X-ray images (350 normal cases and 350 COVID-19 cases) from both public and private institutes. The highest 4 models are VGG-19, VGG-16, ResNet-101, and SqueezeNet with accuracy ranging from 90.7 to 94.3% and F1-score from 90.8 to 94.3%. The VGG-16 is the highest with an accuracy of 94.3% and F1-score of 94.3%. The majority voting of the 18 models and the highest 4 models achieved an accuracy of 93.0% and 94.0%, respectively.

The proposed framework

In this section, the proposed framework has been explained. First, the used chest X-ray dataset has been described. Then, the developed framework, which includes “pre-processing” phase and the “Classification using CNN models based on transfer learning” phase, has been illustrated. Two different approaches have been used to train pre-trained CNN models using transfer learning. The first approach uses whole chest X-ray images, while the other approach uses lung-segmented images.

Used datasets

In this research, the data obtained from the “COVID-19 Radiography Database” has been used to apply the proposed framework. The database contains thousands of publicly available benchmark X-ray images and corresponding lung masks. The X-ray images are provided in Portable Network Graphics (PNG) format with a resolution of 299 × 299 pixels. The database includes 10,192 Normal cases, 3616 positive COVID-19 cases, 1345 Viral Pneumonia cases, and 6012 Lung Opacity images as shown in Table 1. This database was developed by a team from Qatar University, Dhaka University, Bangladesh with cooperators from Malaysia and Pakistan and cooperators of medical doctors<sup>26</sup>. Figure 1 illustrates samples from different classes in the COVID-19 Radiography Database.

Preprocessing

The purpose of the pre-processing phase is to prepare the X-ray images for classification using CNN pre-trained models. In this phase, different pre-processing steps are applied to improve the performance of the classification. The pre-processing steps can be summarized as follows:

Classes	Normal	Positive COVID-19	Viral pneumonia	Lung opacity
Number of cases	10,192	3616	1345	6012

Table 1. COVID-19 radiography database distribution.

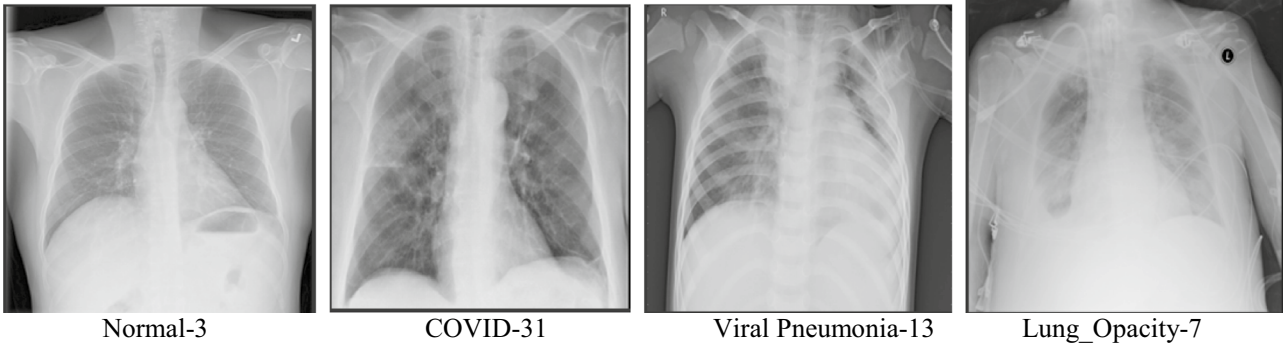


Figure 1. Samples from COVID-19 radiography chest database representing different classes.

### Image enhancement

Enhancing images is a significant step for the correct classification. It increases image contrast in order to improve classification performance. Different techniques can be applied to enhance the images. In this research, some of these techniques have been applied to the original X-ray images before introducing them to the classification models, they are as follows:

1. **Histogram Equalization (HE):** The purpose of histogram equalization (HE) is to spread the gray levels inside the image. It modifies the brightness and contrast of the images to improve the image quality<sup>27</sup>. The original X-ray images' intensity has been enhanced using histogram equalization (HE).
2. **Contrast Limited Adaptive Histogram Equalization (CLAHE):** It originated from Global Histogram Equalization (GHE), it is based on dividing the image into non-overlapping blocks, and after that, the histogram of each block is gotten using a pre-specified value<sup>28</sup>. In this research, CLAHE has been used to enhance the contrast of original X-ray images.
3. **Image Complement:** The complement or inverse of X-ray images transforms the dark positions to lighter and the light positions to darker. As this is a standard process, which is similar to that used by radiologists, it may aid a deep learning model for improving classification performance. The complement of the binary image can be obtained by changing the zeros to ones and ones to zeros. Whereas for a grayscale image, each pixel is subtracted from 255.

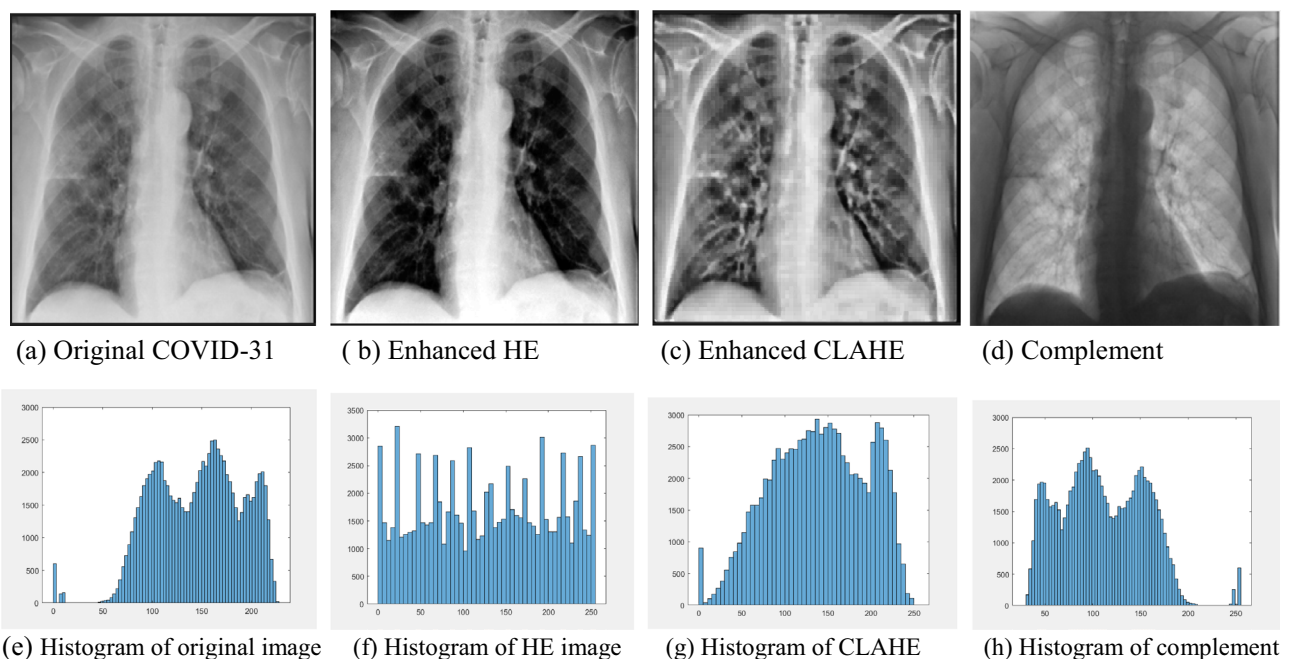
Figure 2 shows an original X-ray image and its enhanced versions after applying HE, CLAHE and image complement on the original image with the corresponding histogram plots for each version.

### Segmentation

In the segmentation step, the regions of interest (ROI), which are the lungs region in our case, are cropped from the associated image. In this research, the ground truth lungs' masks which are provided by the database have been used. A modified U-Net model was applied by the authors of the database on the X-ray images to get the lung masks associated with the full X-ray images. In this research, multiplication between each original image and the associated lung mask has been applied to get the segmented lungs. The same process of multiplication between different enhanced image versions and the associated masks has been applied to get different versions of segmented datasets with different enhancements. All these versions are introduced to CNN models as segmented versions of data. Figure 3 shows the segmented images of the original image and of the different enhanced images for one of the COVID samples.

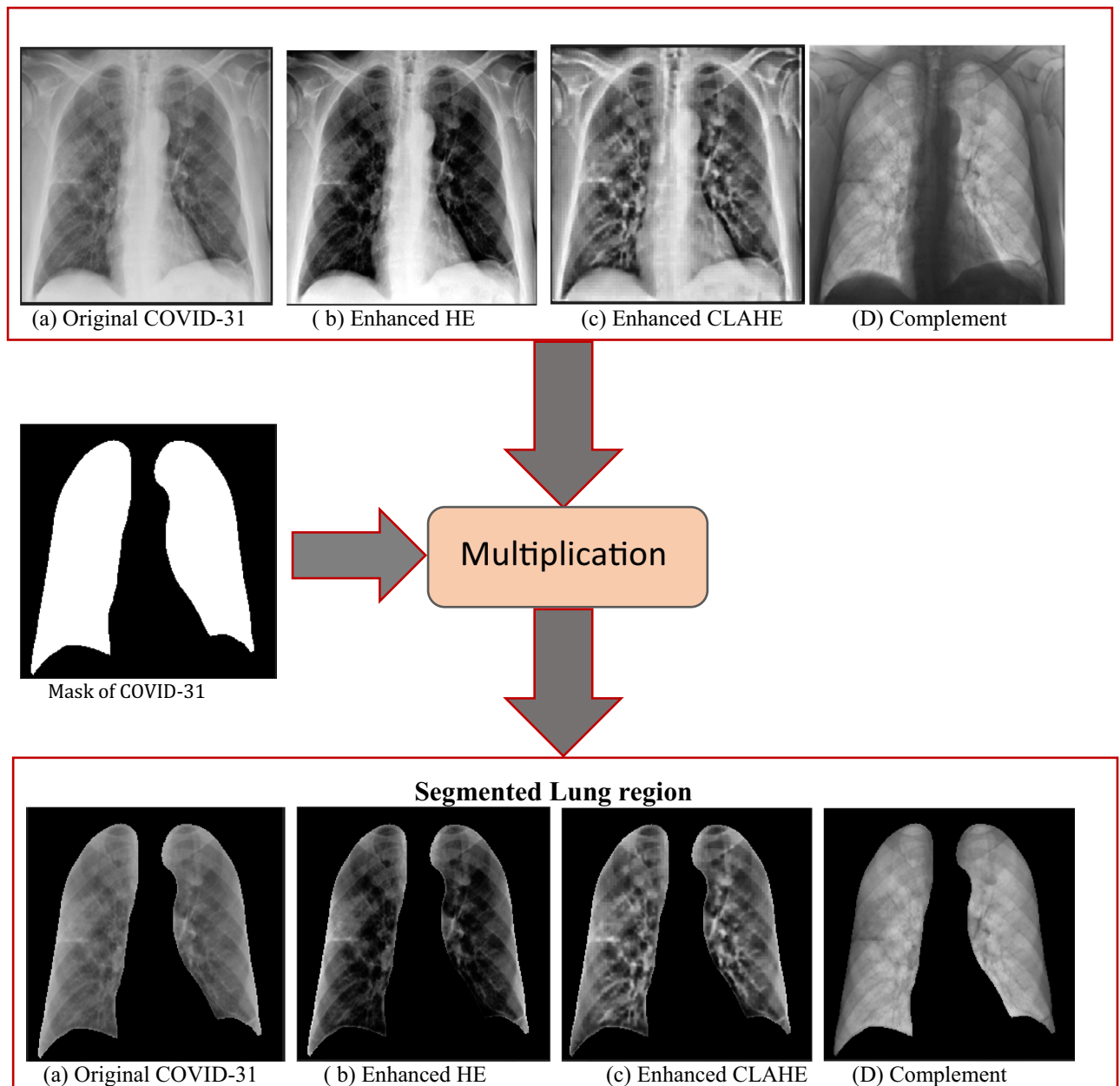
### Resizing images phase

Resizing the images is an essential process to satisfy the requirement of CNN of equally sized input images. In this research, the process of resizing X-ray images has been done to fit all X-ray images to the input size of the used pre-trained CNN models which are VGG19 and EfficientNetB0. Therefore, all images' versions either full or segmented versions were resized to fit the CNNs input image size which is  $224 \times 224$  pixels. To expedite



**Figure 2.** An X-ray image and its enhanced versions after applying HE, CLAHE and complement to the original image and the corresponding histogram plots.





**Figure 3.** X-ray original image and its enhanced versions and the segmented lung region of each version.

the training process, it was found that the size of  $112 \times 112$  pixels expedited the training without affecting the performance metrics.

#### Classification using pre-trained convolution neural network model

In this research, different versions of either full or segmented chest X-ray images have been introduced to CNN models to train the classifiers. Different experiments have been carried out on the original and segmented lung X-ray images both with their different enhanced versions. The classification has been done using VGG19<sup>14</sup> and EfficientNetB0<sup>16</sup> pre-trained CNN models. After the calculation of different performance metrics, the best model has been selected as the adopted model. The next subsections give a brief description of the used pre-trained models.

##### VGG19 model

VGG19 is a variant of the VGG CNN model which was created by Visual Geometry Group (VGG) at Oxford University. VGG19 was one of the winners of the Image Net Large Scale Visual Recognition Challenge (ILSVRC) in 2014. The size of the input image to VGG19 is  $(224 \times 224)$ . VGG19 contains 16 convolution layers, 5 max-pooling layers and 3 fully connected layers. The convolution layers are with  $(3 \times 3)$  filters' size, stride of 1 pixel and padding of 1 pixel. The max-pooling layers are with a size of  $2 \times 2$  and a stride of 2. The rectification (ReLU)

activation function is utilized for all hidden layers. Then, the first 2 fully connected layers with 4096 channels each are utilized followed by the last layer of 1000 channels to represent the different 1000 classes of the ImageNet with soft-max activation function<sup>15</sup>.

#### *EfficientNetB0 model*

Google research group designed a family of models, called EfficientNets using a scaling method and achieved better efficiency and accuracy than previous ConvNets. EfficientNet is based on scaling CNNs and reaching better performance by balancing network width, depth, and resolution. Therefore, the focus is to present a scaling method to uniformly scale the 3 dimensions with a simple highly effective compound coefficient. Thus, it can be considered as an optimization problem to find the best coefficients for depth, width, and resolution that maximizes the accuracy of the network given the constraints of the available resources. The primary building block of the EfficientNet models is MBConv. The network's dimension equation was used to get the family of neural networks EfficientNet-B0 to B7<sup>16</sup>. In this research, EfficientNetB0 was used for the classification of the chest X-ray images. Figure 4 sums up the framework of the adopted methodology in this research.

#### Ethical approval

This article does not contain any studies with human participants or animals performed by the author.

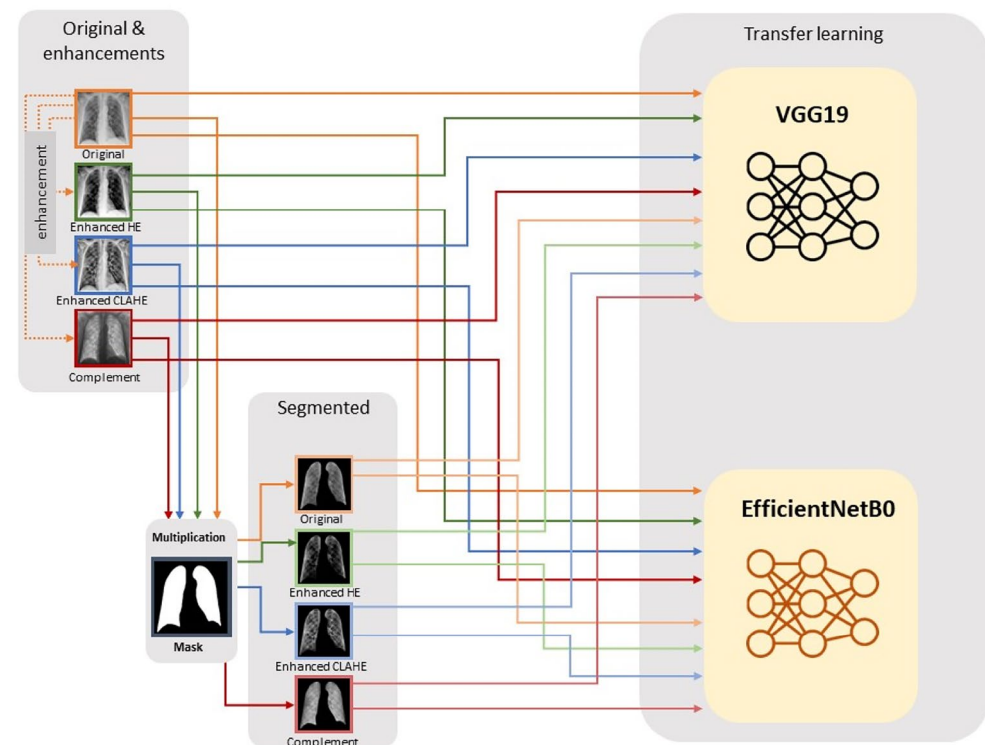
#### Experimental results

This section presents the experimental results of the proposed framework to evaluate its performance and study the effect of different enhancements applied to the X-ray images on the performance of the CNN classification model.

#### Experimental setup

Keras Python deep learning library on top of TensorFlow was utilized for implementing CNN models on a machine with the following specification; an Intel® Core™ i7 CPU@ 3.6 GHz with 32 GB RAM and a Titan × Pascal Graphics Processing Unit (GPU). Extensive experiments were carried out to obtain the best settings of the CNN models that achieve the best possible results. It is worth noting that the pre-processing for enhancing the images has been carried out using MATLAB® 18 software.

Both the full and segmented datasets with their enhanced versions have been used to train VGG19 and EfficientNetB0 CNN pre-trained models. The training was carried on with Adam optimizer, learning rate of 0.001, batch size of 32, and the number of epochs (10–30) epochs, SoftMax classifier. The fine-tuned pre-trained models were used for feature extraction; therefore, the weights of the pre-trained models were frozen, and they were not updated during the training to maintain ImageNet's initial weights. The top layers were fine-tuned to adjust the



**Figure 4.** The framework of the used methodology for Chest X-ray images classification.

network according to the used chest X-ray data and to the current problem output which is (2/4) rather than 1000 in the ImageNet data. To avoid overfitting, a dropout of 0.3 was applied in the fully connected layers. Figures 5 and 6 illustrate the fine-tuned top layers based on VGG19 and EfficientNetB0 respectively for binary classification.

### Performance evaluation

A benchmark dataset was employed to validate the performance of the proposed framework. For binary classification, a set of 1600 X-ray images (800 of each class) have been used to train CNN models using transfer learning. The dataset has been divided into three subsets training, validation, and test sets. The training set is used for learning the model and adjusting the parameters. The validation set is to test the model during the training phase and fine-tune the parameters. The test set is to evaluate the trained model. The division was done as 400 samples (25%) of the used X-ray images were selected randomly for testing (200 images for each class), and the remaining 75% samples were split again into training and validation splits (80–20%). For 4-classes classification, a set of 3200 X-ray images (800 of each class) have been used. Before training CNN models, different pre-processing steps were implemented to enhance the images of both full and segmented lungs chest X-ray images to investigate the classification performance of the CNN models using the different versions.

The following metrics were used for the evaluation of the different CNN models trained using various dataset versions:

$$\text{Sensitivity/Recall(\%)} = \frac{TP}{TP + FN} \quad (1)$$

$$\text{Precision(\%)} = \frac{Tp}{TP + FP} \quad (2)$$

$$\text{Specificity(\%)} = \frac{TN}{TN + FP} \quad (3)$$

$$\text{Accuracy(\%)} = \frac{TP + TN}{TP + FP + TN + FN} \quad (4)$$

dense (Dense)	(None, 1024)	25691136
dense_1 (Dense)	(None, 512)	524800
dense_2 (Dense)	(None, 256)	131328
dropout (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 2)	514
Total params: 39,292,738		
Trainable params: 26,347,778		
Non-trainable params: 12,944,960		

**Figure 5.** Fine-tuned top layers based on VGG19 pre-training model.

dense (Dense)	(None, 1024)	5243904
dense_1 (Dense)	(None, 512)	524800
dense_2 (Dense)	(None, 256)	131328
dropout (Dropout)	(None, 256)	0
dense_3 (Dense)	(None, 2)	514
Total params: 9,534,117		
Trainable params: 5,900,546		
Non-trainable params: 3,633,571		

**Figure 6.** Fine-tuned top layers based on EfficientNetB0 pre-training model.

$$F1score = \frac{2TP}{2TP + FP + FN} \quad (5)$$

where TP is a true-positive value, FP is a false-positive value, TN is a true-negative value and FN is a false-negative value.

## Results

Tables 2 and 3 show the results of training VGG19 using the original and different enhanced versions of full X-ray images and of segmented versions respectively. As it is shown in Table 2, applying different image enhancement techniques has improved the performance of the classification model. The accuracy of classification for the model trained using the original images' version was 0.913, however, it has been improved for the enhanced versions to reach 0.94, 0.95, and 0.9475 for histeq, CLAHE, and complement respectively. The detailed results including TP, TN, FP, FN, sensitivity, specificity, precision, F1 score, test accuracy, and AUC of the proposed full X-ray images using VGG19 are shown in Table 2. The performance of the model trained using the CLAHE version is the best.

Regarding the segmented versions, as it is shown in Table 3 the accuracy of classification of the model trained using the original segmented dataset version was 0.887. However, the accuracy has been improved for the enhanced segmented dataset versions to reach 0.91 using Histeq techniques, 0.9049 for CLAHE and 0.9075 for the complement version. It is clear that the accuracies using different enhanced versions are close to each other, and they are better than that of the original segmented version. The detailed results using the different metrics are shown in Table 3.

Table 4 shows the results of training EfficientNetB0 using the original and different enhanced versions of full X-ray images. The accuracy of classification using the original full images' version was 0.915, it reached 0.94, 0.938, and 0.94 for histeq, CLAHE, and complement versions respectively. The accuracies for the models trained using different enhanced versions are better than that of the original version. The detailed results using the different metrics are shown in Table 4.

Regarding the segmented versions, as shown in Table 5 the accuracy of the classification for the EfficientNetB0 model trained using the original segmented lung dataset version was 0.885. However, the accuracy has been improved to 0.905, 0.905, and 0.9075 for Histeq, CLAHE and complement versions respectively. As with VGG19, the accuracies of training EfficientNetB0 using different enhanced segmented versions are close to each other, but they are all better than that of the original segmented version. The detailed results using the different metrics are shown in Table 5.

It is clear that the performance of all enhanced versions is better than that of their associated original version using either VGG19 or EfficientNetB0 models and for both full and segmented versions. By comparing the results

Dataset	Tn	Fp	Tp	Fn	Sensitivity	Specificity	Precision	F1 score	Test acc	AUC
Original	170	30	195	5	0.975	0.85	0.8667	0.9176	0.913	0.9125
Histeq	192	8	184	16	0.92	0.96	0.9583	0.9388	0.94	0.94
CLAHE	188	12	192	8	0.96	0.94	0.9412	0.9505	0.95	0.95
Complement	187	13	192	8	0.96	0.935	0.93659	0.9482	0.9475	0.9475

**Table 2.** The results of applying VGG19 to original and different enhanced versions of the used dataset.

Dataset	Tn	Fp	Tp	Fn	Sensitivity	Specificity	Precision	F1 score	Test acc	AUC
Segmented original	157	43	198	2	0.99	0.785	0.8216	0.898	0.887	0.8875
Segmented Histeq	173	27	191	9	0.955	0.865	0.876	0.9139	0.91	0.91
Segmented CLAHE	173	27	189	11	0.945	0.865	0.875	0.9086	0.905	0.9049
Segmented complement	169	31	194	6	0.97	0.845	0.862	0.9129	0.908	0.9075

**Table 3.** The results of applying VGG19 to original and different enhanced versions of the used dataset after segmentation.

Dataset	Tn	Fp	Tp	Fn	Sensitivity	Specificity	Precision	F1 score	Test acc	AUC
Original	173	27	193	7	0.965	0.865	0.877	0.91905	0.915	0.915
Histeq	192	8	184	16	0.92	0.96	0.9583	0.9387	0.94	0.94
CLAHE	194	6	181	19	90.5	0.97	0.9679	0.9354	0.938	0.9375
Complement	193	7	183	17	0.915	0.965	0.963	0.9385	0.94	0.94

**Table 4.** The results of applying EfficientNetB0 to original and different enhanced versions of the used dataset.



Dataset	Tn	Fp	Tp	Fn	Sensitivity	Specificity	Precision	F1 score	Test acc	AUC
Segmented original	160	40	194	6	0.97	0.8	0.829	0.894	0.885	0.885
Segmented Histeq	181	19	181	19	0.905	0.905	0.905	0.905	0.905	0.905
Segmented CLAHE	183	17	179	21	0.895	0.915	0.9133	0.904	0.905	0.905
Segmented complement	174	26	189	11	0.945	0.87	0.879	0.9108	0.9075	0.9075

**Table 5.** The results of applying EfficientNetB0 to original and different enhanced versions of the used dataset after segmentation.

of the full X-ray images and segmented images using the same CNN model, reductions in the performance of the CNN models trained using the segmented datasets rather than those trained using full-image datasets were observed. The models built using full images achieved better performance in general. The reason for that might be that the full images may have more details outside the lung region in the surroundings region that contribute to the classification and help to improve the performance.

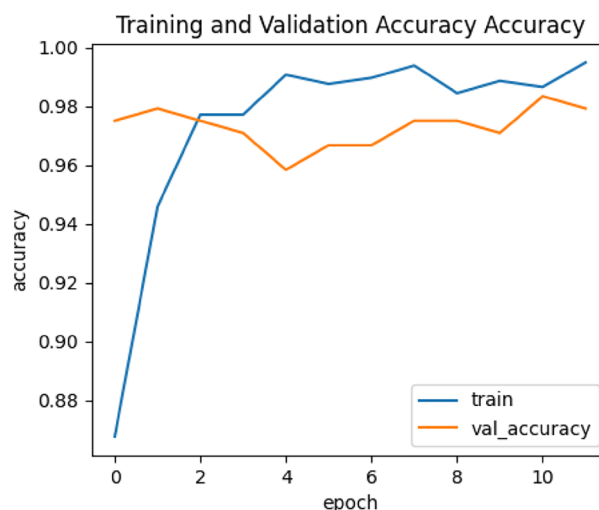
Regarding the comparison of the used 2 CNNs models, the results of the two models are close to each other, however, the results of VGG19 are a bit better than those of EfficientNetB0. The best achieved performance is of the VGG19 model trained using the CLAHE version of full X-ray images for binary classification of chest X-ray images into COVID-19 or normal. It achieved a sensitivity of 0.96, specificity of 0.94, precision of 0.9412, F1 score of 0.9505 and accuracy of 0.95. Figures 7, 8, 9 and 10 show the training and validation accuracies, the training and validation losses, the training and test ROC curve and the confusion matrix of that best achieved model respectively.

## Discussion

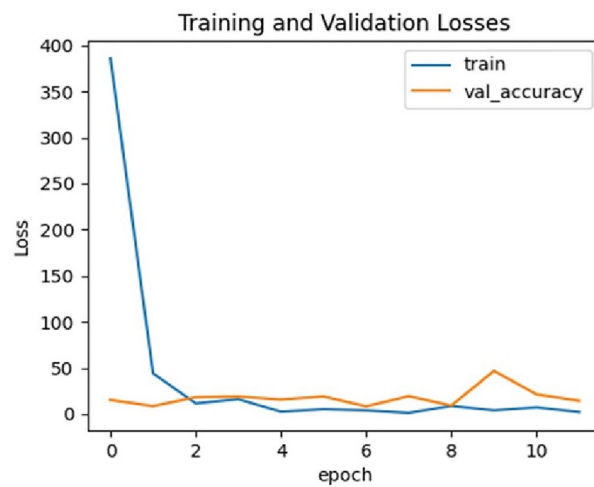
Distinguishing COVID-19 from normal and other classes is one of the important issues since the pandemic in 2019. The contribution of this research is to develop a framework to classify Coronavirus suspected cases as normal or COVID-19 positive cases. Different pre-processing steps have been applied to improve the performance of the classification. After that, multiplication between the original images and the associated lung masks has been applied to get the segmented lungs. The same process of multiplication has been applied between different enhanced image versions and the associated masks to get different enhanced versions of segmented datasets. All these versions are introduced to CNN models which are VGG19 and EfficientNetB0. Therefore, two different approaches have been used to train pre-trained CNN models using transfer learning. The first approach uses full chest X-ray images, while the other approach uses lung segmented images.

From the results of conducted experiments, it has been observed that the proposed framework has achieved a good performance using either full or segmented images, however the performance using full images is better than that using segmented. Moreover, it has been observed that the performance of the classification models has been improved after applying enhancement techniques.

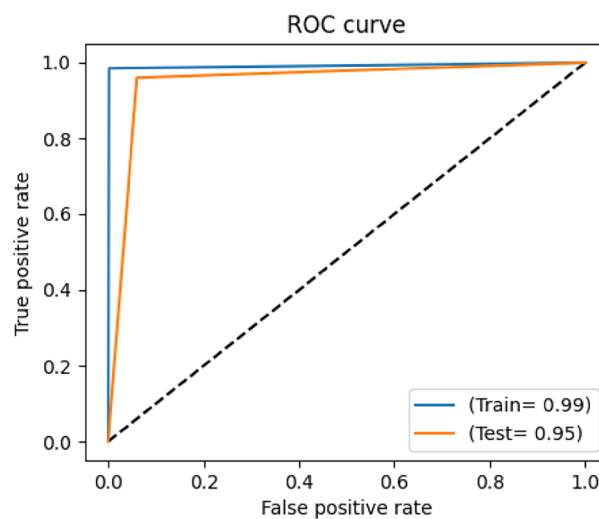
To evaluate the proposed framework with respect to the state-of-the-art works in COVID-19, it has been compared with the related works reviewed in this research as described in Table 6. It is worth mentioning that, the comparison is not an easy task in COVID research as the pandemic broke out in the world suddenly, all Covid research used different sources of data either local, public, or combined from different databases. Some of the public datasets are collected from different other databases. Even the research that used the same public



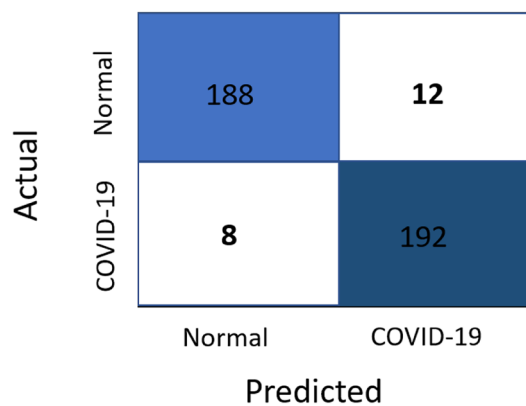
**Figure 7.** Training and validation accuracy of the best model.



**Figure 8.** Training and validation losses of the best model.



**Figure 9.** Training and test ROC curve of the best achieved model.



**Figure 10.** Confusion matrix of the best model.

Ref	Techniques	Modalities	Task	Accuracy
<sup>19</sup>	CNN (VGG16)	1248 chest X-ray	3-classes classification: (Healthy/COVID-19/pneumonia/non-COVID-19 pneumonia)	83.6% accuracy, 90.9% sensitivity
<sup>20</sup>	SqueezeNet	5184 Chest X-rays	Binary classification of images as COVID-19 or not	98% sensitivity, 92.9 specificity
<sup>21</sup>	CNN model	13,824 X-ray images	Binary classification of images as COVID-19 or normal	96.71% accuracy, 97% F1-score
<sup>22</sup>	CNN called COVID-Ne	13,975X-ray	Prediction of cases as (normal/pneumonia/COVID-19)	93.3% accuracy
<sup>23</sup>	CNN called nCOVnet	284 X-ray images	(Normal/COVID-19)	88.1% accuracy
<sup>24</sup>	EfficientNet	16,634 X-ray images	Classify cases as: (Normal/COVID-19/other)	93.48% accuracy
<sup>25</sup>	VGG-16	700 X-ray images	binary classification (Normal/COVID-19)	94.3% accuracy, 94.3% F1-score
Proposed model	VGG19	1600 X-ray images	Binary classification (Normal/COVID-19)	95% accuracy, 0.9505% F1-score, 0.96% sensitivity, 0.94% specificity

**Table 6.** Comparison of proposed method with the relevant researches.

Dataset	Model	F1 score	Test acc
Histeq	EfficientNetB0	0.93495	0.935
CLAHE	VGG19	0.9339	0.93375
Complement	VGG19	0.9337865	0.93375

**Table 7.** The results of best achieved models for 4-classes classification.

datasets used different number of samples. Some research performed binary class classification, where others performed multi-class classification. Thus, the proposed work has been compared with others that used the same modality which is X-ray with mentioning the number of used samples and the task.

By comparing the results of the proposed framework with recent literature, it was found that the proposed framework outperforms most of the state-of-the-art works. However, <sup>21</sup> slightly outperforms the proposed framework. Where the accuracy and F1-score of <sup>21</sup> are 96.71% and 97% respectively, the corresponding values of the proposed framework are 95% and 0.9505 respectively. Taking in consideration that in <sup>21</sup> un-balanced data has been used; the number of COVID images used is 3626 where the number of normal images is 10,198 as mentioned in the manuscript.

For more validation, different classes of the datasets have been used for training the CNN models. To check the capability of the proposed framework for 4-classes classification. The dataset versions that achieved the highest binary classification have been utilized for multi-classes classification. Since the performance of all models that trained using enhanced full image versions is close to each other, therefore, these versions have been utilized for 4-classes classification. A set of 3200 X-ray images (800 of each class) have been used to train CNN models. The newly added classes are Viral Pneumonia and Lung Opacity in addition to COVID-19 and normal classes.

It was found that the best-achieved accuracy of 4-classes classification using the full image versions reached 0.935 for histeq version by EfficientNetB0. While it reached 0.93375 for both CLAHE, and complement versions using VGG19. Table 7 shows the results of the best-achieved models for 4-classes classification.

## Conclusion and future work

In this research, a framework has been developed for automatically classifying chest X-ray images as COVID-19 positive cases or normal cases. Different techniques such as histeq, CLAHE, and complement have been applied to enhance the original X-ray images and therefore, both the original and enhanced versions have been introduced to the selected CNN pre-trained models. Two pre-trained CNN models which are VGG19 and EfficientNetB0 have been used to train different versions with the last dense layer set to (2/4) according to the number of classification classes.

Two approaches have been utilized to train pre-trained CNN models which are using whole chest X-ray images and using lung segmented images with their enhanced versions. The best binary classification accuracy reached 95% for the model trained using CLAHE full images version utilizing VGG19. The best achieved accuracy for a model trained using a segmented dataset is 91% for the model trained using Histeq version utilizing VGG19. By testing the framework for 4-classes classification, it achieved promising results which reached 0.935 accuracy.

It is obvious from the results that, the proposed framework can be employed in the future to support physicians and decrease the effect of doctors' shortages in the struggle against the disease. However, extra validations are required before applying any system, as more accuracy and more careful experiments are needed when things are related to human life. In the future, the authors are willing to try the proposed model on local data.

## Data availability

The used data has been obtained from an available online database and it has been referenced in the manuscript. The link to the database used in the study: <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database>

Received: 7 January 2024; Accepted: 8 May 2024

Published online: 21 May 2024

## References

1. <https://www.who.int/europe/emergencies/situations/covid-19>.
2. Kanne, J. P. *et al.* Essentials for radiologists on COVID-19: An update—Radiology scientific expert panel. *Radiology* **296**, E113–E114 (2020).
3. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019>.
4. <https://www.worldometers.info/coronavirus/>.
5. [https://www.who.int/health-topics/coronavirus#tab=tab\\_1](https://www.who.int/health-topics/coronavirus#tab=tab_1).
6. Yang, D. *et al.* Detection and analysis of COVID-19 in medical images using deep learning techniques. *Sci. Rep.* **11**(1), 1–13 (2021).
7. Alsattar, H. A. *et al.* Developing deep transfer and machine learning models of chest X-ray for diagnosing COVID-19 cases using probabilistic single-valued neutrosophic hesitant fuzzy. *Expert Syst. Appl.* **236**, 121300 (2023).
8. Wang, X. *et al.* Broad learning solution for rapid diagnosis of COVID-19. *Biomed. Signal Process. Control* **83**, 104724 (2023).
9. Mezina, A. & Burget, R. Detection of post-COVID-19-related pulmonary diseases in X-ray images using Vision Transformer-based neural network. *Biomed. Signal Process. Control* **87**, 105380 (2024).
10. Gaur, P. *et al.* COVID-19 disease identification from chest CT images using empirical wavelet transformation and transfer learning. *Biomed. Signal Process. Control* **71**, 103076 (2022).
11. Wong, H. Y. F. *et al.* Frequency and distribution of chest radiographic findings in patients positive for COVID-19. *Radiology* **296**(2), E72–E78 (2020).
12. Badrinarayanan, V., Kendall, A. & Cipolla, R. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(12), 2481–2495 (2017).
13. Russakovsky, O. *et al.* Imagenet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
14. Kermany, D. S. *et al.* Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **172**(5), 1122–11319 (2018).
15. Simonyan, K., Zisserman, A. *Very deep convolutional networks for large-scale image recognition*. arXiv preprint arXiv:1409.\*\*\*1556 (2014).
16. Tan, M., Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In *International Conference on Machine Learning*. (PMLR, 2019).
17. Howard, A. G., *et al.* Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:\*\*\*.1704.04861 (2017).
18. He, K., *et al.* Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. (2016).
19. Nishio, M. *et al.* Automatic classification between COVID-19 pneumonia, non-COVID-19 pneumonia, and the healthy on chest X-ray image: Combination of data augmentation methods. *Sci. Rep.* **10**(1), 1–6 (2020).
20. Minaee, S. *et al.* Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. *Med. Image Anal.* **65**, 101794 (2020).
21. Sahin, M. E. Deep learning-based approach for detecting COVID-19 in chest X-rays. *Biomed. Signal Process. Control* **78**, 103977 (2022).
22. Wang, L., Lin, Z. Q. & Wong, A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest X-ray images. *Sci. Rep.* **10**(1), 1–12 (2020).
23. Panwar, H. *et al.* Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. *Chaos Solitons Fractals* **138**, 109944 (2020).
24. Nigam, B. *et al.* COVID-19: Automatic detection from X-ray images by utilizing deep learning methods. *Expert Syst. Appl.* **176**, 114883 (2021).
25. Chow, L. S. *et al.* Quantitative and qualitative analysis of 18 deep convolutional neural network (CNN) models with transfer learning to diagnose COVID-19 on chest X-ray (CXR) images. *SN Comput. Sci.* **4**(2), 141 (2023).
26. Rahman, T., COVID-19 radiography database. <https://www.kaggle.com/tawsifurrahman/covid19-radiography-database> (2021).
27. Veluchamy, M. & Subramani, B. Image contrast and color enhancement using adaptive gamma correction and histogram equalization. *Optik* **183**, 329–337 (2019).
28. Zimmerman, J. B. *et al.* An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. *IEEE Trans. Med. Imaging* **7**(4), 304–312 (1988).

## Author contributions

E.M.F.E.H. is the only author of this manuscript and hence is the corresponding author and the only contributor for this manuscript.

## Funding

Open access funding provided by The Science, Technology & Innovation Funding Authority (STDF) in cooperation with The Egyptian Knowledge Bank (EKB).

## Competing interests

The author declares no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to E.M.F.E.H.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

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



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024