Article   Open Access

Deep Learning Convolutional Neural Network for SARS-CoV-2 Detection Using Chest X-Ray Images

Ali Mohammed Saleh Ahmed [1] 🆔**, Inteasar Yaseen Khudhair** [2] 🆔**,
Salam Abdulkhaleq Noaman** [1] 🆔

[1] College of Education for Pure Sciences, University of Diyala, Diyala, Iraq
[2] College of Basic Education, University of Diyala, Diyala, Iraq

Corresponding author: Ali Mohammed Saleh Ahmed (dr.alimahmed@uodiyala.edu.iq)

Abstract

The COVID-19 coronavirus illness is caused by a newly discovered species of coronavirus known as SARS-CoV-2. Since COVID-19 has now expanded across many nations, the World Health Organization (WHO) has designated it a pandemic. Reverse transcription-polymerase chain reaction (RT-PCR) is often used to screen samples of patients showing signs of COVID-19; however, this method is more expensive and takes at least 24 hours to get a positive or negative response. Thus, an immediate and precise method of diagnosis is needed. In this paper, chest X-rays will be utilized through a deep neural network (DNN), based on a convolutional neural network (CNN), to detect COVID-19 infection. Based on their X-rays, those with COVID-19 indications may be categorized as clean, infected with COVID-19 or suffering from pneumonia, according to the suggested CNN network. Sample pieces from every group are used in experiments, and categorization is performed by a CNN. While experimenting, the CNN-derived features were able to generate the maximum training accuracy of 94.82% and validation accuracy of 94.87%. The F1-scores were 97%, 90% and 96%, in clearly categorizing patients afflicted by COVID-19, normal and having pneumonia, respectively. Meanwhile, the recalls are 95%, 91% and 96% for COVID-19, normal and pneumonia, respectively.

Keywords

COVID-19; SARS-CoV-2; CNN; Deep neural network; RT-PCR; Computer model.

1   Introduction

In December 2019, the first case of COVID-19, a form of coronavirus illness that damages the pulmonary system of human beings, was discovered in Wuhan, which is located in China. As of 24 August 2022, the World Health Organization (WHO) states there have been a total of 602,772,663 patients diagnosed of COVID-19 all across the world, including 6,477,184 deaths. The water particles that are exhaled by an afflicted individual when they sneeze, cough or exhale transmit the virus from one individual to the next. The virus transmission is relatively straightforward. Nearly every nation on Earth is making a herculean effort to reduce the number of people who become infected with COVID-19. If the precise figure of COVID-19 instances in a particular area is recognized, then it will be ready to obtain preventative measures to slow down the overall disease incidence. However, this will only be viable if a sufficient number of people are tested for COVID-19.

It is important to consider COVID-19 test accuracy, speed and costs while choosing one. These are critical. Currently, a COVID-19 outbreak is diagnosed using the reverse transcription-polymerase chain reaction (RT-PCR) test (Giridhar & Sampathila, 2022) since it is the gold benchmark worldwide. The RT-PCR check for COVID-19 detects virus RNA. If you get a positive check, it means you have the coronavirus. To ensure accuracy, RT-PCR is done forty times. The genetic link of COVID-19 is verified by molecular data. This method has several false negatives. False-negative results are often caused by data collection errors, contaminated reagents, improper lab procedures, bad transportation or warehousing. A simple detection solution is needed because current techniques are lengthy. Variations in findings can be attributable to test type, viral inoculum amount, sample collection duration and test type. Poor viral concentration and sample errors can affect RT-PCR results. Thus, these procedures often yield deceptive findings and may require two or three attempts to prove their correctness. The intricate analysis can take hours or days to complete. In this case, the development of a trustworthy diagnostic tool for the diagnosis of COVID-19 is of utmost relevance (Corman et al., 2020; Kermany et al., 2018; Salehi et al., 2020).

It is critical to have reliable methods of estimating and predicting COVID-19 levels all through the present outbreak. It is possible to create a workable method with the help of big data analytics. It can lessen the likelihood of further viral transmission. Currently, there is no technique that can guarantee an outcome with a definite degree of certainty (Watson et al., 2020). The inability of even the RT-PCR assay to identify the infection at an extremely preliminary phase contributes to a spike in disease incidence and the delay of crucial preventative measures (Wikramaratna et al., 2020). A diagnosis and prognosis technology is necessary to protect the public and slow down the infection transmission. Therefore, a completely computerized framework is essential for virus identification. The amount of raw DICOM COVID-19 data that is currently available to the public is little. In the realm of medical imaging, the DICOM format is the standard for showing information. Loss of informative value and the introduction of a great deal of bias in our forecast occur during the process of converting data from DICOM to simple visual representations, often according to the JPEG and PNG standards. With the most recent neural network advancements, a simple and effective method based on a deep learning network may be created. However, different methodologies have been introduced in the literature to overcome the above issues (Abdulkareem et al., 2022; Hasoon et al., 2021; Mahmoudi et al., 2022; Mohammed et al., 2022; Nagi et al., 2022; Saeed et al., 2022; Shamim et al., 2022) starting from machine learning to deep learning, as will be seen in more detail in the next section.

Therefore, this paper presents a modification of the VGG-CNN-network. It consists of 16-layers plus input and output layers. The structure has an increasing density of kernels, starting from sixteen of three by three up to 512 of three by three filter size. We do not utilize a padding operation. The stride was one steep in all of the employed layers. The activation of the convolutional layers was ReLU. Furthermore, we remove redundant neighbours addressed by the pooling layers, where we use more than one pooling layer. Experiments found that the CNN-derived features achieved the highest training accuracy of 94.82%

and validation accuracy of 94.87%, with corresponding F1-scores of 97%, 90% and 96%, when it came to accurately classifying patients as either having COVID-19, being healthy or having pneumonia. The recall rates are 95% for COVID-19, 91% for normal and 96% for pneumonia. Our work can be summarized as follows:

- The problem statement is discussed to be more familiar for the reader than in previous works. This concerns issues such as the RT-PCR test accuracy and time required.
- A literature review for the state of the art is given, to show a deep insight into the suggested approaches with their weaknesses. Diverse methodologies are studied, such as custom deep neural networks and already suggested networks, which have been trained previously for a certain task.
- A modification to an existing deep neural network, called VGG, is suggested. It consists of 16 layers, input and output. Starting with sixteen three-by-three kernels, the structure has 512 three-by-three filters. Padding is not used. All layers have a steep stride. ReLU activates convolutional layers. We also use multiple pooling layers to remove redundant neighbours.

The rest of this paper is organized as follows: Section 1.1 reviews the literature, where the most significantly related works are presented, Section 2 presents the detailed CNN network with the layers used in this work, and Section 3 discusses the suggested methodology. Section 4 shows the simulation settings and a discussion of the results. Last but not least, Section 5 draws the concluding remarks.

1.1 Literature review

The use of chest X-ray photography has been useful in the identification of lung diseases, including COVID-19. Chest X-ray pictures may be used with deep learning techniques to identify cases with COVID-19 and those without it. Computed tomography (CT) is far more sensitive than RT-PCR diagnostics (Bai et al., 2020). The problem is that CT scanners are still rather expensive. Besides, the radiation exposure from an X-ray is far less than that from a CT scan, they are less costly, more easily obtainable and safer to use. Due to the availability of portable tools, X-ray scanning may be performed right at the patient's bedside, drastically lowering the possibility of illness. Researchers have employed artificial intelligence (AI) methods to create a computerized detection approach for COVID-19 infection (DeGrave et al., 2021), which might help alleviate the aforementioned problems. In the past decade, advances in AI and medical image processing have aided a wide range of industries, but none more so than healthcare. Consequently, deep neural networks (DNNs) (Wang et al., 2020) have seen widespread use in the medical field. CNNs (De La Iglesia Vayá et al., 2020) are the most popular type of DNN. To detect COVID-19 infection in chest X-rays using CNNs, a large amount of training data is required (Ng et al., 2020). However, there are no published X-ray datasets (Waheed et al., 2020) that have a representative number of both COVID-19 and healthy chest pictures. The imbalanced information (Zhao et al., 2018) may arise from a lack of guided information.

A number of studies have used a supervised learning approach to establish CNN models for X-ray detection of COVID-19. The COVIDX-Net model was introduced by Hemdan et al. (2020), which was the result of integrating seven convolutional networks. When compared to the separate models, COVIDX-Net performed more effectively. COVIDX-Net performance was only assessed on a small dataset. Wang et al. (2020) established a COVIDNet to improve pneumonia diagnosis and classification. Their classification accuracy was 92.4%. Apostolopoulos and Mpesiana (2020) added 224 COVID-19-positive individuals. For binary classification, this method has a 98.75% success rate, and for ternary classification, it is 93.48%. X-rays and ResNet-50 boosted diagnostic accuracy to 98% for Narin et al. (2021).

DeGrave et al. (2021) ran a few more tests to examine the stability and adaptability of CNN algorithms. With the COVID-x collection, they successfully replicated a number of supervised algorithms, including COVID-Net (Wang et al., 2020). Classification ability drops by 50% when tested on outside COVID and healthy samples (De La Iglesia Vayá et al., 2020). The detection of COVID-19 was accomplished with the

use of saliency maps and picture edges (Ng et al., 2020). Cardiovascular silhouette and diaphragm were two non-COVID measures that were proven to be unreliable predictors. Issues arise when attempting to extend highly accurate models from one dataset to another, which can be made difficult by the fact that pictures collected by various machines present unique challenges. The Decompose, Transfer and Compose (DeTraC) deep convolutional neural network was validated for COVID-19 X-ray categorization tasks by Abbas et al. (2021). There is a report on using a CNN model to classify chest X-ray scans (Kesim et al., 2019). The new network architecture as well as the scale of the source image are explored for the categorization of X-ray chest scans. To address this issue, the researchers devised a compact CNN design. It is seen that the suggested network successfully assigns chest images to 12 classes with a completion rate of around 86%, and training is accomplished rapidly because of the compact network topology.

Using freely available information, Minaee et al. (2020) compiled a dataset of 5,000 X-rays. Images of the COVID-19 infection were uncovered by a board-certified radiologist. Transfer learning was used to train four well-known CNNs to recognize COVID-19. For feature extraction from X-ray pictures, Nasiri & Alavi (2022) used a DenseNet169 network that had already been pretrained. In order to improve accuracy while decreasing computational complexity and training time, a feature selection method, namely analysis of variance (ANOVA), was used to select the features to be used. After that, eXtreme Gradient Boosting (XGBoost) was used to categorize chosen features. Two-class classification using the proposed method achieved a 98.72% success rate, whereas multi-class classification using COVID-19 no findings and pneumonia achieved a 92.8% success rate. Arias-Garzón et al. (2021) describe a novel method that builds on previous DNN research. Specifically, it seeks to improve the pretreatment phase in order to produce accurate and trustworthy outcomes in detecting COVID-19 from X-ray scans. In order to enhance the effectiveness of the classifiers, preprocessing includes a network to filter the images based on their projection (lateral or frontal), some common procedures such as normalization, quality control and resampling to reduce information variance, and a segmentation prototype to retrieve the lung area, which also includes the necessary details and rejects details of neighbours, which can generate inaccurate outcomes. A classifier (VGG16-19) is applied after the preprocessing phase, aided in effectiveness and time to completion by a transfer learning strategy that makes use of pre-trained weights from an even larger database, such as ImageNet.

Ahmed et al. (2022) shed light on the difficulty of generalizing deep learning-based models used in the disease detection process. They demonstrate how to use X-rays to diagnose COVID-19 with this scenario. They present and empirically validate strategies that prevent shortcut learning. With their recommended techniques, algorithms experience a statistically substantial decrease in effectiveness degradation when using previously unknown information. As a result, the efficiency decrease is reduced from 20% to 9%. Sae-Lim et al. (2022) proposed a different CNN architecture: a simpler CNN model for COVID-19 assessment, which can divide COVID-19 lesions into basically two look categories. The suggested prototype is superior since it requires fewer layers and makes use of a data preprocessing, augmentation strategy during the learning phase. They utilized a subset of a publicly available dataset, COVID-19 radiography collection, which is a set of 13,808 X-ray photographs, to assess the performance of the suggested methodology. To explain the COVID-19 forecasts, they used the Grad-CAM generative model to highlight the crucial areas in X-ray photographs.

In order to classify COVID-19 X-ray photographs, Kanjanasurat et al. (2022) evaluated two photograph-enhancing approaches and five CNN models. The X-ray source was examined against an improved version of the same photograph using a contrast-limited adaptive histogram (CLAHE) plus gamma adjustment. To identify COVID-19, the researchers employed five pre-trained CNN models that are all freely accessible to the public: MobileNet, MobileNetV2, DenseNet169, DenseNet201 and ResNet50V2. The publicly available COVID-19 radiology dataset served to verify their method. The researchers found that MobileNet using gamma adjustment is optimal for COVID-19 categorization, including an accuracy

of 87.53% during its initial epoch but 95.46% after 100 learning epochs using the least amount of computational effort. The purpose of the study by Singh & Sahu (2022) was to demonstrate a CNN architecture for automated identification of COVID-19 using X-ray photographs of the chest. They used the freely released Kaggle collection, which has a total of 42,330 photos divided into four groups. The research yielded an accuracy of 88.53% for learning and 86.19% for verification, which represents a superior performance for the largest amount of radiological photographs compared to previous research.

In the experiment described in Deo et al. (2022), a thick-layer CNN consisting of twenty-one layers was utilized to make a prediction about the kind of pneumonia based on X-ray chest photographs. The experiment findings demonstrate that the approach exceeds some alternative approaches that seem to be comparable in terms of both accuracy (97.38%) and specificity (93.6%). As a result, the risk of a false-positive result is reduced, which could also make the problem much simpler to diagnose.

However, there are different architectures of CNN networks. For instance, handwritten character recognition is handled quite effectively by the CNN network known as LeNet (LeCun et al., 1998). AlexNet seems to be the deep learning infrastructure responsible for making CNN commonplace. Its design was quite comparable to that of LeNet, although it was larger, more complex, and contained layered convolutional networks (Krizhevsky et al., 2017). The marriage of fully connected layers and CNNs is what makes ZFnet the CNN infrastructure of choice. Matthew Zeiler and Rob Fergus created ZFNet. It took first place in the 2013. Even though it has fewer variables than AlexNet, the model still managed to achieve outstanding accuracy with only 1000 photos of each class in the ILSVRC 2012 categorization assignment (Zeiler & Fergus, 2014). In 2013, researchers introduced the concept that would become VGG; by 2014, it had placed second in the ImageNet competition. When contrasted to AlexNet or even ZFNet, its simplicity makes it popular (Simonyan & Zisserman, 2015). While many strong networks were created in 2014, including VGG, GoogleNet emerged as the victor in the ImageNet competition. Inception modules, as suggested by GoogleNet, are mini-modules formed by skip interconnections that are then replicated inside the net at large (Szegedy et al., 2015). Furthermore, in 2015 ResNet, the scientists demonstrated experimentally that, contrary to "plain nets", wherein introducing a few layers led to increased training and testing defects, the percentage error could perhaps continue falling as more layers were added. Training it using an eight GPU system lasted about 2-3 weeks (He et al., 2016).

That is to say, subsequent to the aforementioned foundational architectures, a wide range of models have been proposed. The vast bulk of these ground-breaking innovations are either improvements upon or ensembles of previously influential concepts. While the enhanced performance of the new models is certainly appreciated, there has been no truly revolutionary thinking in the last few years.

2 Convolutional Neural Network Architecture

A deep learning method known as a convolutional neural network (CNN) is capable of accepting an image representation, attributing significance, trainable weights and biases known as parameters to distinct features in the image, and identifying one feature from another. When contrasted across alternative categorization techniques, a CNN network requires a far smaller amount of preparatory work than other techniques. When filters were hand-engineered, crude approaches were in use. CNN networks have the potential to learn these filters effectively provided that they receive sufficient training. The use of CNNs has made image categorization and object classification more accessible than ever before. These models use linear algebraic concepts, notably matrix multiplication, to detect patterns in input images. The arrangement of the visual cortex served as an inspiration for the design of a CNN, whose geometry mimics the connectivity network of neurons. A neuron's receptive field is the portion of the visual field in which it is sensitive to inputs. The whole field of view is covered by a mosaic of similar overlapping areas.

When compared to other kinds of neural networks, CNNs stand out by how well they handle visual or auditory data. Convolutional layers, pooling layers and a final fully-connected layer serve as the three primary types of layers in these architectures, listed in order from input to output. In other words, the very initial layer of a CNN network is a convolutional layer. The fully connected layer is the last layer in a neural network architecture, coming after the initial convolutional layer and any further convolutional or pooling layers. The complexity of the CNN grows with each additional layer, allowing it to recognize more and more information within the image. Primitive elements such as colours and borders are refined in the first few levels. Information from an image is fed into a CNN, where it is processed layer by layer until the goal item is recognized. Figure 1 shows a simple and general CNN network with the most important layers. For instance, as shown in Figure 1, an image is sent into the network as input, where it undergoes a series of operations including multiple convolutions, subsampling and a fully connected layer before producing any result.

Each neuron in the convolutional layer calculates the element-wise product of its weight-values and the small receptive area that it has linked in the input matrix in order to determine its response, and this layer is responsible for generating the network's output. A feature map is derived using the source image at the end of each calculation. Think of a picture as a 5x5 matrix of entries; then, using a 3x3 matrix as a sliding kernel window, crop the image to the desired proportions. The contents of the 3x3 kernel are multiplied by the picture elements that are presently being screened by the kernel at every place in the matrix. Thus, we obtain one score that stands for 'values of all the pictures within the given kernel window. Filtering occurs at this layer, with the kernel window scanning the picture for characteristics as it is moved around. The result of the convolution being multiplied by filters is what makes this method function.
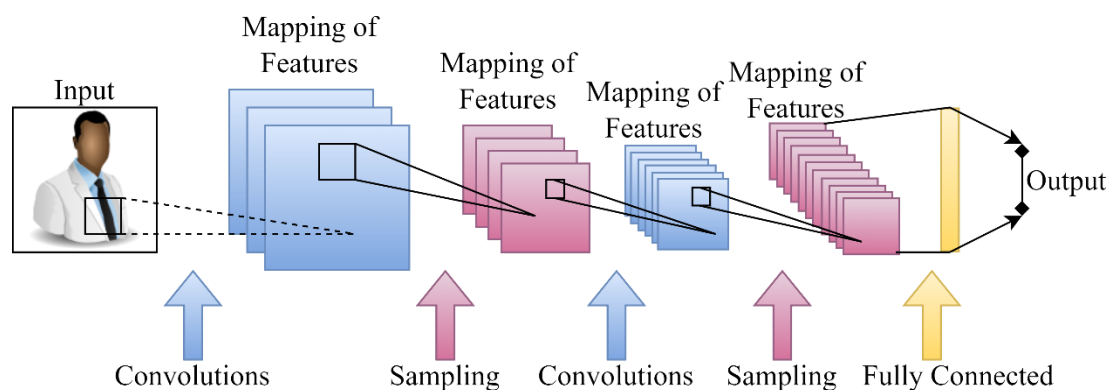
Figure 1. *Convolutional neural network, general architecture showing the most important and basic layers.*

The purpose of sampling is to provide a reduced-dimensional depiction of the input, which aids in the prevention of over-fitting. Max-pooling is a method of sampling. Using this method, we can choose the area with the greatest data point, taking into account the area size. To put it another way, max-pooling selects the greatest value in the region of the image that is now being processed by the kernel. Using a 2x2 max-pooling layer, for instance, one may choose the highest intensity amount among all the pixels in a 2x2 grid. After that, the pooling layer functions similarly to the convolutional layer, as previously suspected. Moving the window of the kernel across an image works the same way, with the exception that the procedure performed by the kernel or window is not linear. The whole pooling operation described above can be visualized in Figure 2.
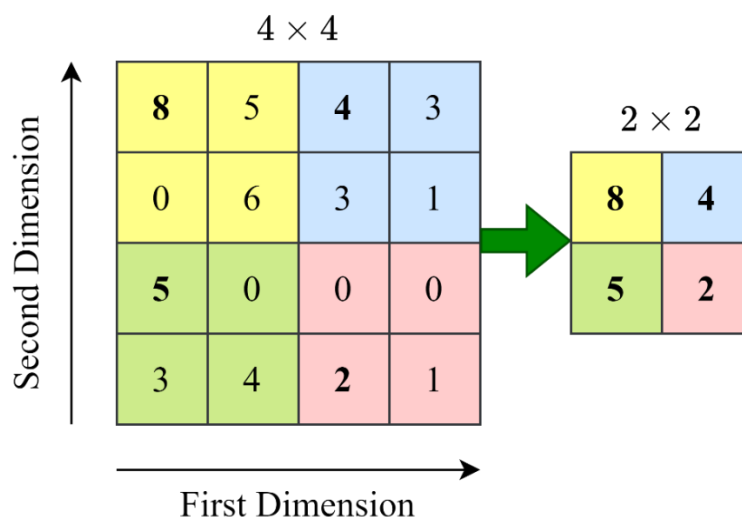
Figure 2. Sampling as represented by max-pooling layer, principle of operation.

The purpose of the fully connected layer is to aggregate all of the features, as well as to flatten the high-level characteristics that were already learned by the previous convolutional layers. It then sends the flattened result to the output layer, where it is used to predict the input class label using either a softmax classifier or a sigmoid. In our work, we employ the softmax classifier, as will be seen later. Learning nonlinear mixtures of the increased characteristics that are reflected by the production of the convolutional layers can be accomplished in a relatively inexpensive manner (in most cases) by introducing a fully connected layer. A potentially nonlinear function in that area is learned by the fully connected layer.
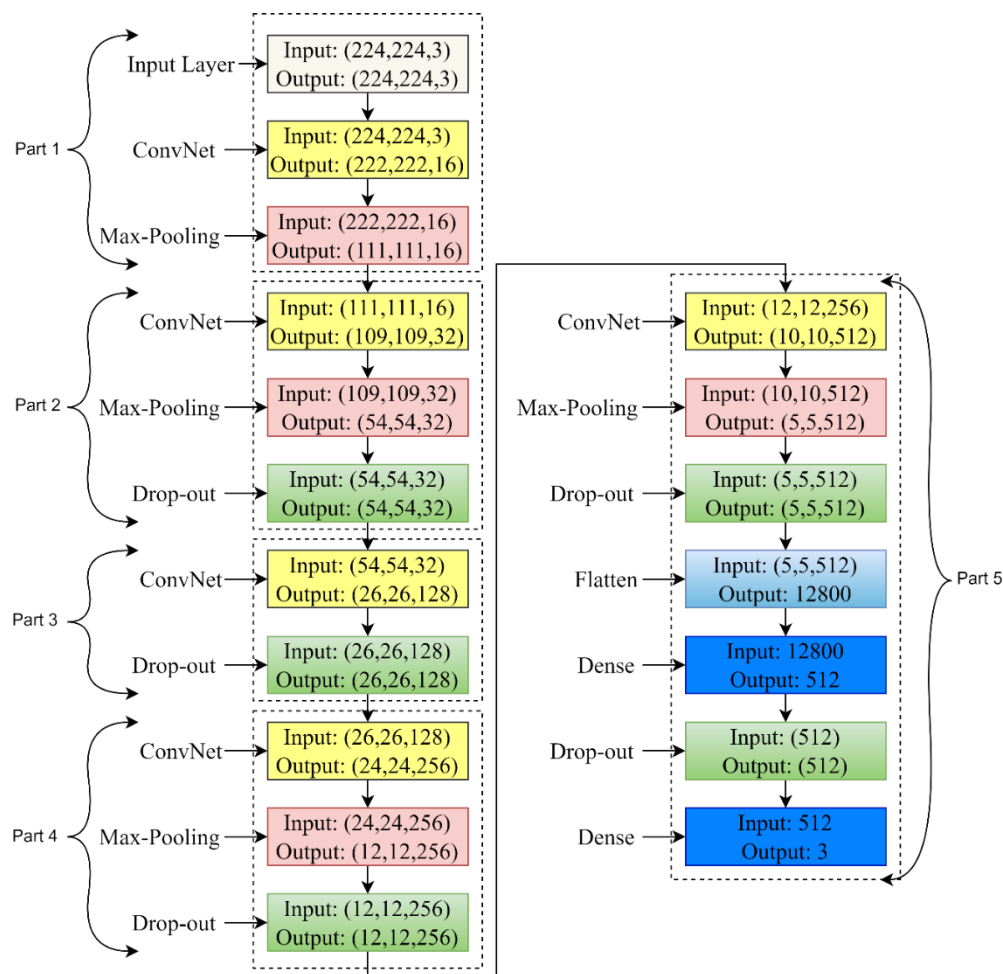
3   Proposed CNN Network

In this section, the suggested CNN network structure will be explained. According to the literature review in Section 1.1, most of the suggested methodologies are either a network that is already known or a mix of two different known networks, or maybe a new structure. That said, our proposed structure is different from previous works. The structure starts with input/output layers and between them, the convolutional and max-pooling layers will be located. Figure 3 depicts the general construction of the CNN network suggested in this paper. Namely, there are five convolutional layers, four max-pooling layers, five drop-out layers, one flattening layer, and two dense layers.
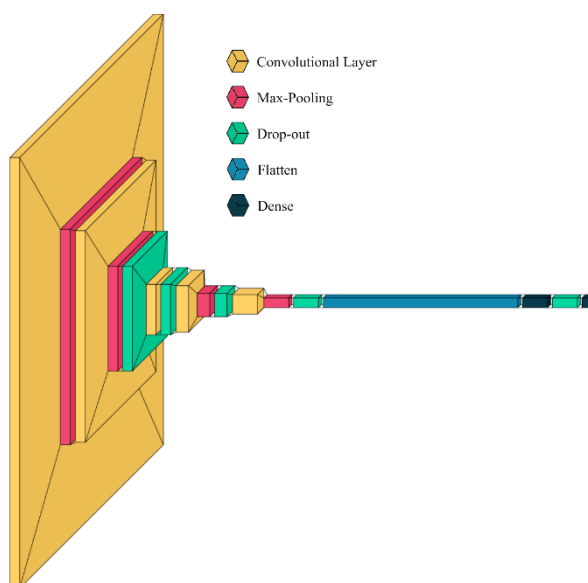
Figure 3. Structure demonstration of suggested CNN network.

Thus, the structure of the suggested network is mainly divided into five parts. Each part starts with a convolutional layer, except the first part which has an input layer. Thus, the first part consists of three layers: input, convolutional layer (ConvNet) and max-pooling. The first layer is the ConvNet. This ConvNet is configured as 16 filters of kernel window size 3 by 3. The activation function type is rectified linear unit (ReLU). Note that the input size of the image was (224, 224, 3); in other words, the image height equals 224, its width is 224, and it has 3 colours. Accordingly, there will be 448 learnable parameters and the output size is (222, 222, 16), where the last dimension, 16, stands for the number of filters of the layer, as shown in the flowchart in Figure 4. That is, as indicated in Figure 4, the output size of any layer is equivalent to the input size of the next layer. For instance, the first max-pooling layer input size is (222, 222, 16) which is the output size of the previous ConvNet layer. Nonetheless, the max-pooling layer accepts inputs and produces the input size divided by 2, according to our design in the suggested CNN network, but the number of filters remains unchanged. The number of learnable parameters in this max-pooling layer is zero.

Figure 4. Suggested CNN network, details of whole layers.

Next, the second part consists of three layers: ConvNet, max-pooling, and drop-out. The first layer in the second part is ConvNet of an input size (111, 111, 16) and the output becomes (109, 109, 32), to be fed to the next layer in the second part, which is the max-pooling layer, to give an output sized (54, 54, 32). The drop-out layer follows; hence, the number of learnable parameters is 4640. The third part of the suggested network is constructed using two layers, namely ConvNet and drop-out. The ConvNet layer accepts input equal to the output of the last layer, which is the drop-out layer of the second part; hence, the size is (54, 54, 32) and produces an output sized (26, 26, 128) to the next layer in this third part, which is a drop-out layer. Accordingly, the number of learnable parameters is 36,992. Next is the fourth part, which comprises

of three layers; it is similar to the second part but the input size equals (26, 26, 128) for its first layer, ConvNet, where its output size is (24, 24, 256). After this layer, a max-pooling layer sized (24, 24, 256) follows to introduce an output sized (12, 12, 256) to the next layer, which is a drop-out layer. This last layer in the fourth part produces an output of the same size as its input. Accordingly, the number of learnable parameters is 295,168.

Last but not least, the fifth part is the final stage in the structure of the suggested CNN network. This fifth part is an architecture from seven layers (including the output layer, a dense layer). As stated previously, each part of the proposed CNN network starts with a ConvNet layer. Consequently, the first layer in the fifth part is a ConvNet with an input size (12, 12, 256) and an output size (10, 10, 512), followed by a max-pooling layer with an input size (10, 10, 512) and output size (5, 5, 512) to be fed to the next layer. The next layer is a drop-out layer with an input size (5, 5, 512). For more convenience, Table 1 lists all the hyper-parameters of the suggested network.

Table 1. *Hyper-parameters of proposed CNN network.*

| Part	Layer no.	Type	Input size	Output size	Learnable parameters
1	1	Input	(224, 224, 3)	(224, 224, 3)	None
	2	ConvNet	(224, 224, 3)	(222, 222, 16)	448
	3	Max-pooling	(222, 222, 16)	(111, 111, 16)	0
2	4	ConvNet	(111, 111, 16)	(109, 109, 32)	4640
	5	Max-Pooling	(109, 109, 32)	(54, 54, 32)	0
	6	Drop-out	(54, 54, 32)	(54, 54, 32)	0
3	7	ConvNet	(54, 54, 32)	(26, 26, 128)	36,992
	8	Drop-out	(26, 26, 128)	(26, 26, 128)	
4	9	ConvNet	(26, 26, 128)	(24, 24, 256)	295,168
	10	Max-pooling	(24, 24, 256)	(12, 12, 256)	0
	11	Drop-out	(12, 12, 256)	(12, 12, 256)	0
5	12	ConvNet	(12, 12, 256)	(10, 10, 512)	1,180,160
	13	Max-pooling	(10, 10, 512)	(5, 5, 512)	0
	14	Drop-out	(5, 5, 512)	(5, 5, 512)	0
	15	Flatten	(5, 5, 512)	12,800	0
	16	Dense	12,800	512	6,554,112
	17	Drop-out	512	512	0
	18	Dense	512	3	1539

Further, there are first and second dense layers (see Figure 3 and Figure 4) with 6,554,112 and 1539 learnable parameters, respectively. Accordingly, there are a total of 8,073,059 learnable parameters. Note that the third ConvNet layer has a 2-step stride, being the only employed stride in the suggested CNN network. It is worth noting that the output from the last dense layer is only three; this is because the suggested network will be employed for classifying three classes: pneumonia, COVID-19 and normal, as will be seen in the next section, where the detailed pre-processing and dataset source with the simulation results are discussed comprehensively.

4   Results and Discussion

Before starting to use the proposed CNN network, it is essential to dive into the dataset that will be used in this work. Thus, the dataset utilized in our work is that found on the Kaggle website (Prashant, 2020). This dataset consists of 6432 sample of X-rays of the chest. There are three classes in this dataset, 'PNEUMONIA', 'NORMAL', 'COVID19', which is why we have three outputs in our CNN design (Figure 4). There are two folders in this dataset, first for training, which contains 5144 images in total and 1288 images for testing purposes. The image size was (224, 224, 3). Figure 5 presents a sample of an infected-type X-ray image, while Figure 6 shows an X-ray image classified normal, and Figure 7 indicates an X-ray for a pneumonia-infected person. Alternatively speaking, the samples shown in Figures 5-7 are samples drawn from the dataset, which is utilized in our paper.

Figure 5. *Sample of X-ray image classified as infected with COVID-19 drawn from the dataset. Source: (Prashant, 2020).*
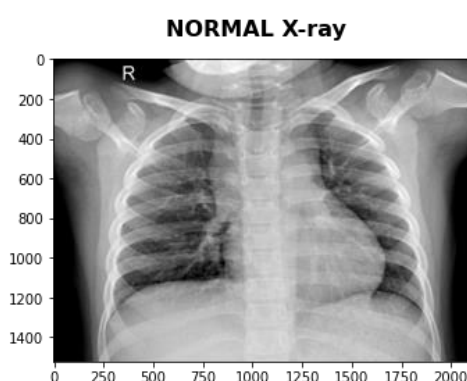
Figure 6. *Sample of X-ray image classified Normal drawn from the dataset. Source: (Prashant, 2020).*
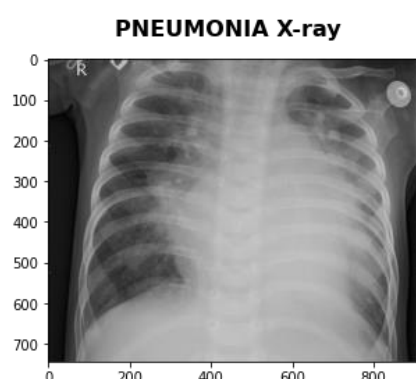
Figure 7. *Sample of X-ray image classified as pneumonia drawn from the dataset. Source: (Prashant, 2020).*

As a first step in the simulations, images are normalized, then the images with sizes other than (224, 224, 3) are corrected to the required size. After that, images are arranged in patches. Each patch consists of 16 images, each image sized (224, 224, 3), as stated previously. Regarding the training phase of the structured CNN network, the part of the images that is dedicated for the training is fed to the network. There are 200 iterations/epochs. In each epoch, there are 64 steps for the iteration, where it is the cumulative number of steps (sample batches) that the generator must produce before the current epoch is declared over and a new one begins. In fact, the simulation code is implemented using Keras TensorFlow package. Keras, however, is a Python-based deep learning application programming interface that operates on top of TensorFlow, a ML framework. Rapid prototyping was a primary design goal during the development. The key to successful research is a rapid time between concept and conclusion.

Consequently, Figure 8 shows the epoch history. As indicated above, 200 iterations are conducted in the simulation. The training losses and validation losses in Figure 8 indicate that the suggested system was trained efficiently during the 200 iterations. A maximum training loss reached at the end of the training phase was 0.1375, while the validation loss was 0.1526, as shown in Figure 8.
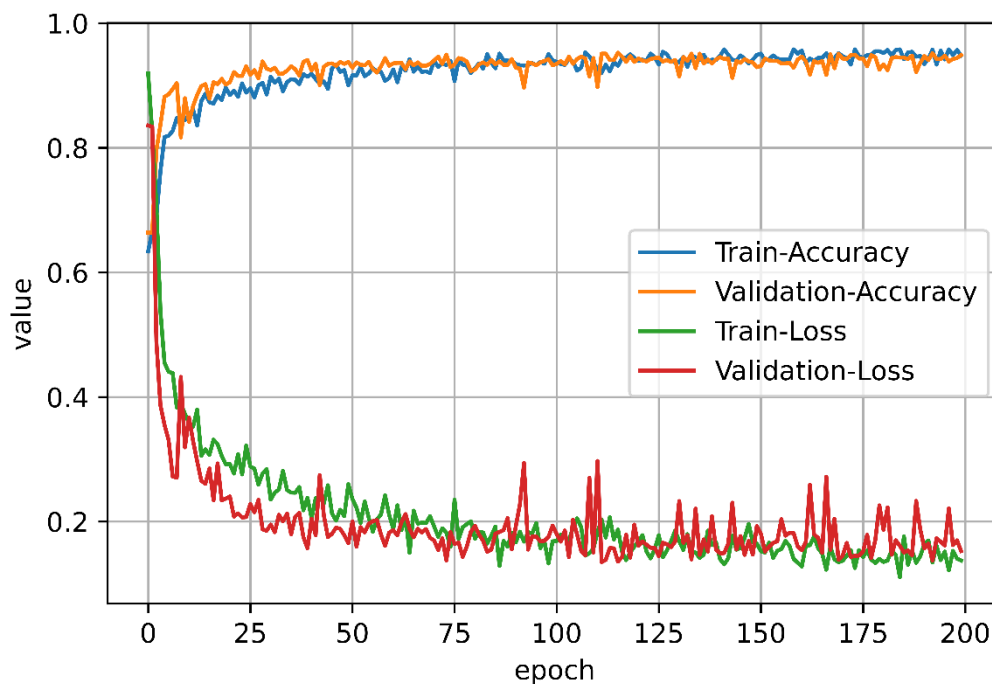

Figure 8. *Loss history over 200 epochs for suggested CNN network.*

Consequently, the accuracy of the model was monitored over the same 200 epochs. It was found that the accuracy followed the same scenario of the loss model of Figure 8. That is, the accuracy of the training phase starts at 0.63, which corresponds to a validation accuracy of 0.66. Thus, at the end of the training phase, the accuracies are 0.9482 and 0.9488, corresponding to training and validation, respectively, as shown in Figure 9. Nonetheless, note that the accuracy reached 0.9 after the 25th epoch as shown in Figure 9. However, Figure 9 also shows that the accuracy was fluctuating above and below 0.95 starting from epoch 175 till the end of the epochs. As a comparison, findings indicated in Figure 8 and those found in Figure 9 are displayed together in one comparison area, as shown in Figure 10. It can be observed that the performance of our suggested CNN network in terms of losses and accuracy is sufficient. Furthermore, the loss fluctuation was significant between epoch 88 and the last epoch. However, during this history of iterations, the accuracy performance was adequate.
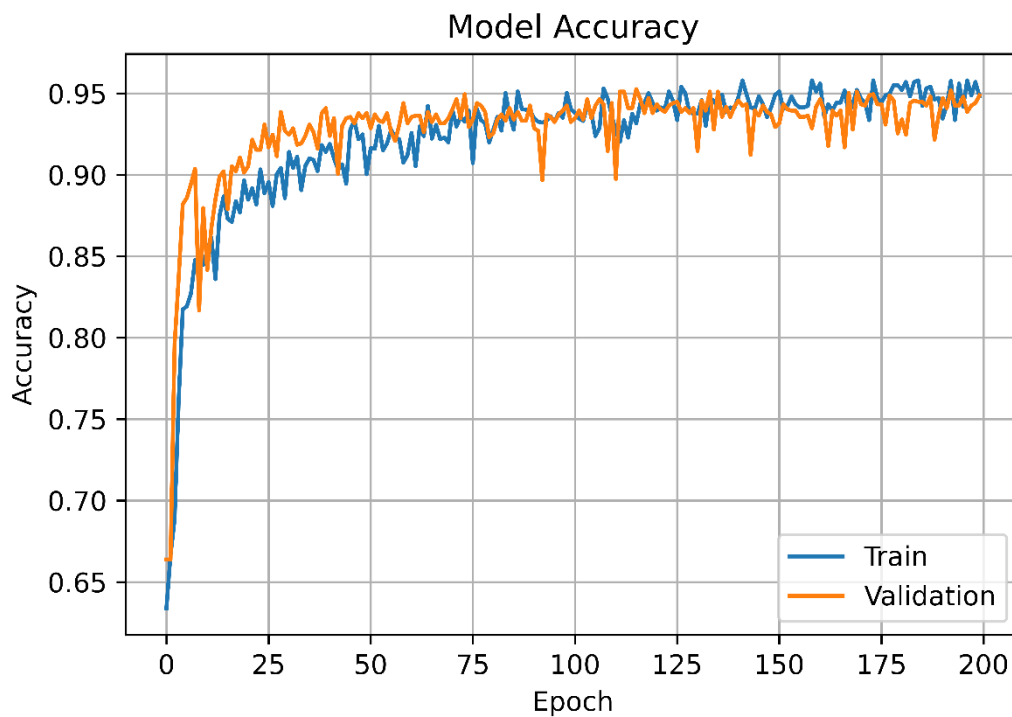
Figure 9. *Accuracy history over 200 epochs for suggested CNN network.*


Figure 10. *Accuracy/loss history over 200 epochs for suggested CNN network.*

Figure 11, on the other hand, shows prediction results for three X-ray images drawn randomly from the three classes of the dataset images (COVID19, NORMAL and PNEUMONIA). According to the trained CNN network, the randomly-selected images are classified correctly, with a boundary highlighted in yellow line for images of COVID19 or PNEUMONIA. Thus, the first sample was drawn from the COVID19 dataset class, the second was taken from the NORMAL class, while the last image is from the PNEUMONIA class.
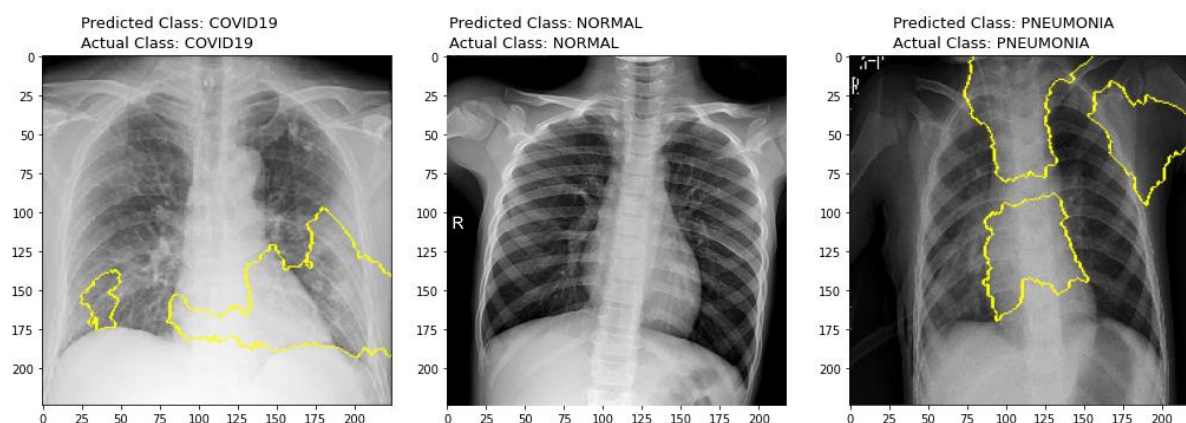
Figure 11. Predicted images from CNN network after it was trained properly.

The confusion matrix in Figure 12 indicates that there is 96.55% COVID19 with 0.86% of NORMAL predicted as COVID19 as well as 0.12% of PNEUMONIA also predicted as COVID19. The NORMAL class prediction was 84.48% with 0% and 2.79% of COVID19 and PNEUMONIA, respectively, predicted as NORMAL. Last but not least, 97.09% of PNEUMONIA class with 3.45% and 14.66% of COVID19 and NORMAL classes, respectively, are also classified as PNEUMONIA.
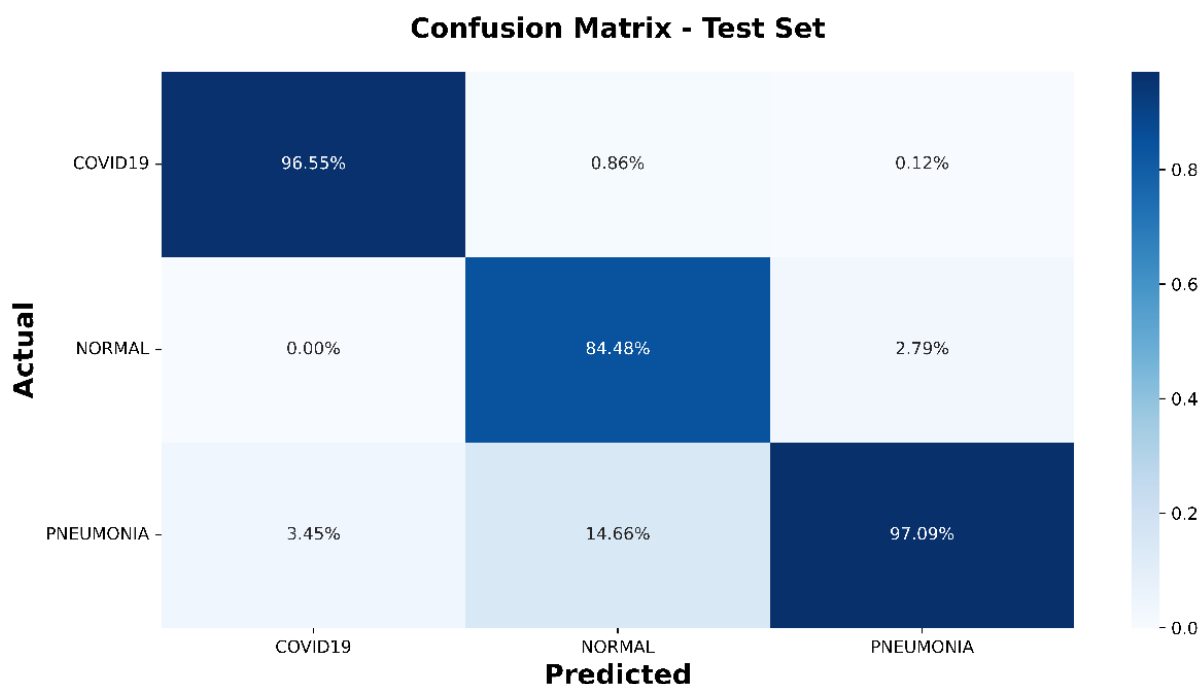
Figure 12. Confusion matrix results of CNN network.

Furthermore, the precision of the classifier for the three classes, COVID19, NORMAL and PNEUMONIA, are 100%, 89%, 96%, respectively. Moreover, and as expected, the recall of the model reached 95%, 91% and 96% corresponding to COVID19, NORMAL and PNEUMONIA, respectively. Furthermore, the F1-scores of COVID19, NORMAL and PNEUMONIA classification classes are 97%, 90% and 96%, respectively, as indicated in the confusion matrix in Figure 12 and Table 2. Finally, one more thing has to be stated, which is the execution time. The whole training time was about 208 minutes, where each iteration needed between 60 and 63 seconds.

Table 2. Classification measure parameters derived from confusion matrix.

Layer mo.	Precision (%)	Recall (%)	F1-score (%)
COVID19	100	95	97
NORMAL	89	91	90
PNEUMONIA	96	96	96

As is seen from the confusion matrix, the precision, recall and the F1-score should all be as high as possible, which is the goal of this work. In other words, measuring one factor, such as precision, is not enough to evaluate the suggested model; hence, we determined the other two factors, recall and F1-score, as listed in Table 2. The results of our methodology are considered improved when compared to the work of Nasiri and Alavi (2022), since they used DenseNet169 that had already been pretrained and consisted of 169 layers, while our network is more compact. However, Nasiri and Alavi (2022) achieved an accuracy of 98.72%, whereas multi-class classification using COVID-19, no findings and pneumonia achieved 92.8%. Furthermore, Wang et al. (2020) achieved 92.4% accuracy in their classifications, while Apostolopoulos & Mpesiana (2020) achieved 93.48%.

5 Conclusion

Regarding the COVID-19 pandemic, which affected all the globe, there must be a fast and cost-effective tool to predict the infection to prevent the outbreak. A convolutional neural network was designed specifically in this paper to distinguish the infection from normal and pneumonia chest X-ray pictures of individuals. A CNN can classify picture features using significance, trainable weights and biases. CNN categorization algorithms require less preparation. Handcrafted filters were basic. CNNs can learn these filters. CNNs simplify image and object classification. Linear algebra, especially matrix multiplication, helps these models discover patterns in input photographs. CNN form mimics the brain's visual neuronal connectivity network and a neuron's optical receptive field. The designed architecture consisted of more than 15 layers. An efficient dataset, extracted from the Kaggle website, was employed in the training of our suggested network. There were 200 epochs to train the model. Adequate results have been reached; therefore, we recommend to continue developing our network for future investigations.

The findings of our suggested method are comparable to other works, but the proposed CNN network is more compact than other works. Furthermore, other works have used either pre-trained networks or new ones, but more complicated than our network. Therefore, our proposed work has outperformed others. It is intended to expand the suggested CNN network in this paper to be capable of detecting more than three classes. This may require addition of more layers with different settings.

Additional Information and Declarations

Conflict of Interests: The authors declare no conflict of interest.

Author Contributions: A.M.S.A.: Conceptualization, Methodology, Software. I.Y.K.: Writing – original draft, Visualization, Investigation. S.A.N.: Supervision, Software, Validation, Writing – review & editing.

Institutional Review Board Statement: Not applicable because the authors used external datasets that already exist.

Data Availability: The data that support the findings of this study are openly available in www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia.

References

Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2021). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Applied Intelligence*, 51(2), 854-864. https://doi.org/10.1007/s10489-020-01829-7

Abdulkareem, K. H., Mostafa, S. A., Al-Qudsy, Z. N., Mohammed, M. A., Al-Waisy, A. S., Kadry, S., Lee, J., & Nam, Y. (2022). Automated System for Identifying COVID-19 Infections in Computed Tomography Images Using Deep Learning Models. *Journal of Healthcare Engineering*, 2022, Article no. 5329014. https://doi.org/10.1155/2022/5329014

Ahmed, K. B., Hall, L. J., Goldgof, D. B., & Fogarty, R. B. (2022). Achieving Multisite Generalization for CNN-Based Disease Diagnosis Models by Mitigating Shortcut Learning. *IEEE Access*, 10, 78726–78738. https://doi.org/10.1109/access.2022.3193700

Apostolopoulos, I. D., & Mpesiana, T. A. (2020). Covid-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, 43(2), 635–640. https://doi.org/10.1007/s13246-020-00865-4

Arias-Garzón, D., Alzate-Grisales, J. A., Orozco-Arias, S., Arteaga-Arteaga, H. B., Bravo-Ortiz, M. F., … Tabares-Soto, R. (2021). COVID-19 detection in X-ray images using convolutional neural networks. *Machine Learning With Applications*, 6, 100138. https://doi.org/10.1016/j.mlwa.2021.100138

Bai, H. X., Hsieh, B., Xiong, Z., Halsey, K., Choi, J. Y., Tran, T. H., Pan, I., Shi, L., Tang, B. Z., Mei, J., Jiang, X., Zeng, Q., Egglin, T. K., … Liao, W. (2020). Performance of Radiologists in Differentiating COVID-19 from Non-COVID-19 Viral Pneumonia at Chest CT. *Radiology*, 296(2), E46–E54. https://doi.org/10.1148/radiol.2020200823

Corman, V. M., Landt, O., Kaiser, M., Molenkamp, R., Meijer, A., Chu, D. I., Bleicker, T., Brünink, S., Schneider, J., Schmidt, M., Mulders, D. G., Haagmans, B. L., … Drosten, C. (2020). Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance*, 25(3). https://doi.org/10.2807/1560-7917.es.2020.25.3.2000045

De La Iglesia Vayá, M., Saborit, J., Montell, J., Pertusa, A., Bustos, A., Cazorla, M., … García-García, F. (2020). BIMCV COVID-19+: A large annotated dataset of RX and CT images from COVID-19 patients. *arXiv preprint arXiv:2006.01174*. https://doi.org/10.48550/arXiv.2006.01174

DeGrave, A. J., Janizek, J. D., & Lee, S. (2021). AI for radiographic COVID-19 detection selects shortcuts over signal. *Nature Machine Intelligence*, 3(7), 610–619. https://doi.org/10.1038/s42256-021-00338-7

Giridhar, A., & Sampathila, N. (2022). Application of Artificial Intelligence to Predict the Degradation of Potential mRNA Vaccines Developed to Treat SARS-CoV-2. In *Proceedings of International Conference on Machine Learning and Big Data Analytics (ICMLBDA) 2021* (pp. 85–94). Springer. https://doi.org/10.1007/978-3-030-82469-3_8

Gouda, W., Almurafeh, M., Humayun, M., & Zaman, N. (2022). Detection of COVID-19 Based on Chest X-rays Using Deep Learning. *Healthcare*, 10(2), no. 343. https://doi.org/10.3390/healthcare10020343

Hasoon, J. N., Fadel, A., Hameed, R. T., Mostafa, S. A., Khalaf, B. A., Mohammed, M. A., & Nedoma, J. (2021). COVID-19 anomaly detection and classification method based on supervised machine learning of chest X-ray images. *Results in Physics*, 31, 105045. https://doi.org/10.1016/j.rinp.2021.105045

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. https://doi.org/10.1109/cvpr.2016.90

Hemdan, E. E.-D., Shouman, M. A., & Karar, M. E. (2020). COVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images. *ArXiv, abs/2003.11055*. https://doi.org/10.48550/arXiv.2003.11055

Kanjanasurat, I., Domepananakorn, N., Archevapanich, T., & Purahong, B. (2022). Comparison of image enhancement techniques and CNN models for COVID-19 classification using chest x-rays images. In *2022 8th International Conference on Engineering, Applied Sciences, and Technology (ICEAST)*. IEEE. https://doi.org/10.1109/ICEAST55249.2022.9826319

Kermany, D., Goldbaum, M. H., Cai, W., Valentim, C. C. S., Liang, H., Baxter, S. L., McKeown, A., Yang, G., Wu, X., Yan, F., Dong, J., Prasadha, M. K., Pei, J., Ting, M. Y. L., Zhu, J., Li, C. W., Hewett, S., Dong, J., Ziyar, I., … Zhang, K. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell*, 172(5), 1122-1131.e9. https://doi.org/10.1016/j.cell.2018.02.010

Kesim, E., Dokur, Z., & Olmez, T. (2019). X-Ray Chest Image Classification by A Small-Sized Convolutional Neural Network. In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)*. IEEE. https://doi.org/10.1109/EBBT.2019.8742050

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6), 84–90. https://doi.org/10.1145/3065386

LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278–2324. https://doi.org/10.1109/5.726791

Mahmoudi, R., Benameur, N., Mabrouk, R., Mohammed, M. A., Garcia-Zapirain, B., & Bedoui, M. H. (2022). A Deep Learning-Based Diagnosis System for COVID-19 Detection and Pneumonia Screening Using CT Imaging. *Applied Sciences*, 12(10), no. 4825. https://doi.org/10.3390/app12104825

Minaee, S., Kafieh, R., Sonka, M., Yazdani, S., & Soufi, G. J. (2020). Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. *Medical Image Analysis*, 65, 101794. https://doi.org/10.1016/j.media.2020.101794

Mohammed, M. A., Al-Khateeb, B., Yousif, M., Mostafa, S. A., Kadry, S., Abdulkareem, K. H., & Garcia-Zapirain, B. (2022). Novel Crow Swarm Optimization Algorithm and Selection Approach for Optimal Deep Learning COVID-19 Diagnostic Model. *Computational Intelligence and Neuroscience*, 2022. https://doi.org/10.1155/2022/1307944

Nagi, A., Awan, M. J., Mohammed, M. A., Mahmoud, A., Majumdar, A., & Thinnukool, O. (2022). Performance Analysis for COVID-19 Diagnosis Using Custom and State-of-the-Art Deep Learning Models. *Applied Sciences*, 12(13), 6364. https://doi.org/10.3390/app12136364

Narin, A., Kaya, C., & Pamuk, Z. (2021). Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, 24(3), 1207–1220. https://doi.org/10.1007/s10044-021-00984-y

Nasiri, H. R., & Alavi, S. M. (2022). A Novel Framework Based on Deep Learning and ANOVA Feature Selection Method for Diagnosis of COVID-19 Cases from Chest X-Ray Images. *Computational Intelligence and Neuroscience*, 2022, Article ID 4694567. https://doi.org/10.1155/2022/4694567

Ng, M., Lee, E. C., Yang, J., Yang, F., Li, X., Wang, H., Lui, M. M. S., Lo, C., Leung, B., Khong, P., Hui, C. K., Yuen, K., & Kuo, M. D. (2020). Imaging Profile of the COVID-19 Infection: Radiologic Findings and Literature Review. *Radiology*, 2(1), e200034. https://doi.org/10.1148/ryct.2020200034

Prashant, P. (2020). Dataset contains chest x-ray images of Covid-19, Pneumonia and normal patients. *Kaggle*. https://www.kaggle.com/datasets/prashant268/chest-xray-covid19-pneumonia

Sae-Lim, W., Suwannanon, R., & Aiyarak, P. (2022). A Simplified Convolutional Neural Network Design for COVID-19 Classification on Chest X-ray Images. In *2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE)*. IEEE. https://doi.org/10.1109/jcsse54890.2022.9836299

Saeed, M., Ahsan, M., Saeed, M., Rahman, A. U., Mehmood, A., Mohammed, M. A., Jaber, M. M., & Damaševičius, R. (2022). An Optimized Decision Support Model for COVID-19 Diagnostics Based on Complex Fuzzy Hypersoft Mapping. *Mathematics*, 10(14), 2472. https://doi.org/10.3390/math10142472

Salehi, S., Abedi, A., Balakrishnan, S., & Assadi, M. (2020). Coronavirus Disease 2019 (COVID-19): A Systematic Review of Imaging Findings in 919 Patients. *American Journal of Roentgenology*, 215(1), 87–93. https://doi.org/10.2214/ajr.20.23034

Shamim, S., Awan, M. J., Zain, A. M., Naseem, U., Mohammed, M. A., & Garcia-Zapirain, B. (2022). Automatic COVID-19 Lung Infection Segmentation through Modified Unet Model. *Journal of Healthcare Engineering*, 2022, Article ID 6566982. https://doi.org/10.1155/2022/6566982

Simonyan, K., & Zisserman, A. (2015). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*. https://doi.org/10.48550/arXiv.1409.1556

Singh, P., & Sahu, S. (2022). Detection of COVID-19 using CNN from Chest X-ray Images. In *2022 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES)*. IEEE. https://doi.org/10.1109/CISES54857.2022.9844330

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S. M., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE. https://doi.org/10.1109/CVPR.2015.7298594

Waheed, A., Goyal, M., Gupta, D., Khanna, A., Al-Turjman, F., & Pinheiro, P. R. (2020). CovidGAN: Data Augmentation Using Auxiliary Classifier GAN for Improved Covid-19 Detection. *IEEE Access*, 8, 91916–91923. https://doi.org/10.1109/access.2020.2994762

Wang, L., Lin, Z., & Wong, A. (2020). COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-76550-z

Watson, J., Whiting, P., & Brush, J. E. (2020). Interpreting a covid-19 test result. *BMJ*, m1808. https://doi.org/10.1136/bmj.m1808

Wikramaratna, P. S., Paton, R. S., Ghafari, M., & Lourenço, J. (2020). Estimating the false-negative test probability of SARS-CoV-2 by RT-PCR. *Eurosurveillance*, 25(50). https://doi.org/10.2807/1560-7917.es.2020.25.50.2000568

Zeiler, M. D., & Fergus, R. (2014). Visualizing and Understanding Convolutional Networks. In *Computer Vision – ECCV 2014* (pp. 818–833). Springer. https://doi.org/10.1007/978-3-319-10590-1_53

Zhao, D., Zhu, D., Lu, J., Luo, Y., & Zhang, G. (2018). Synthetic Medical Images Using F&BGAN for Improved Lung Nodules Classification by Multi-Scale VGG16. *Symmetry*, 10(10), 519. https://doi.org/10.3390/sym10100519