Indonesian Journal of Electrical Engineering and Computer Science
Vol. 24, No. 1, October 2021, pp. 420~427
ISSN: 2502-4752, DOI: 10.11591/ijeecs.v24.i1.pp420-427

Convolution neural network and histogram equalization for COVID-19 diagnosis system

Bashra Kadhim Oleiwi Chabor Alwawi, Layla H. Abood
Department of Control and Systems Engineering, University of Technology, Baghdad, Iraq

Article Info

Article history:

Received Dec 27, 2020
Revised Jul 28, 2021
Accepted Aug 4, 2021

Keywords:

Chest X-ray images
Convolutional neural network
Coronavirus
Data augmentation
Dataset expansion
Histogram equalization

ABSTRACT

The coronavirus disease-2019 (COVID-19) is spreading quickly and globally as a pandemic and is the biggest problem facing humanity nowadays. The medical resources have become insufficient in many areas. The importance of the fast diagnosis of the positive cases is increasing to prevent further spread of this pandemic. In this study, the deep learning technology for COVID-19 dataset expansion and detection model is proposed. In the first stage of proposed model, COVID-19 dataset as chest X-ray images were collected and pre-processed, followed by expanding the data using data augmentation, enhancement by image processing and histogram equalization techniuque. While in the second stage of this model, a new convolution neural network (CNN) architecture was built and trained to diagnose the COVID-19 dataset as a COVID-19 (infected) or normal (uninfected) case. Whereas, a graphical user interface (GUI) using with Tkinter was designed for the proposed COVID-19 detection model. Training simulations are carried out online on using Google colaboratory based graphics prossesing unit (GPU). The proposed model has successfully classified COVID-19 with accuracy of the training model is 93.8% for training dataset and 92.1% for validating dataset and reached to the targeted point with minimum epoch's number to train this model with satisfying results.

This is an open access article under the CC BY-SA license.

Corresponding Author:

Bashra Kadhim Oleiwi Chabor Alwawi
Department of Control and Systems Engineering
University of Technology
Al-Wehda Neighborhood, Baghdad, Iraq
Email: 60010@uotechnology.edu.iq

1. INTRODUCTION

The first cases of the coronavirus disease-2019 (COVID-19) were appeared in December 2019 and reported in Wuhan city-China, and rapidly deployed around the world [1], [2]. Then the novel COVID-19 as a pandemic has been officially declared by the world health organization (WHO) in March 2020 [1]. This pandemic has negatively affected on daily life, public health and the global economy. Most of the COVID-19 infected cases of the patients has been diagnosed using typical Chest X-ray and computed tomography (CT) scan images. As a result of the the rapid increase in COVID-19 cases, few number of testing kits, sometimes lack of precision, the limited number of the limited number of medical specialists in remote locations and in order to save valuable time for medical professionals [3]. There is an urgent need to develop an automated assistant system for diagnosing COVID-19 cases for quickly treating affected patients. The emergence of deep learning as an artificial intelligence technique can be overcame these difficulties and helped with accurate disease detection. Deep learning introduced promising results in helping to analyze the CT scan images [4]-[7] has provided automatic model based DenseNet121-FPN for segmentation of the raw chest CT scan images

and for diagnosing COVID-19. In [8] has proposed online detection nd classification model using DarkNet based on 17 convolutional layers with various filters on each layer for diagnosing the COVID-19 on X-ray images. In [9] hs proposed deep learning model based convolution neural network (CNN) and the PyTorch frame using three binary decision-trees as classifiers the chest X-ray images as infected or uninfected. In [10] has proposed new fractional multichannel exponent moments utlizing framework based parallel multi-core calculation to reduce the computational process for extracting the feature from the chest X-ray images. Then, a modified manta ray foraging optimization approach utlizing differential evolution to choose the important features is used [11] has presented a five pre-trained transfer structures based deep CNN model called ResNet50, ResNet101, ResNet152, InceptionV3, and Inception-ResNetV2 for prediction COVID-19 and normal cases based on chest X-ray images. The results showed that the pre-trained ResNet50 model provided the highest prediction performance and accuracy among other four presented models. In [12] has used pre-trained transferring learning based VGG16 and InceptionV3 for detecting Covid-19 cases from a chest X-ray dataset and then applied support vector machines (SVM) for improving the classification performance. A hybrid model for detecting Covid-19 cases using chest X-ray images, based on machine learning and deep learning algorithms in terms of CNN and Softmax classifier is used in [13]. While [14] has proposed architecture for classifying COVID-19 based on chest X-ray images including seven classifiers called COVIDX-Net of the images with high accuracy performance of 90% for the VGG19 classifiers and the DenseNet201 [15]. Has proposed two 3D CNN models to classify the pulmonary CT images into COVID-19, Influenza-A viral pneumonia and healthy cases. The classical ResNet structure has been used and compared with the network using and without using the added location-attention mechanism. The experimental result of proposed model proved that the overall accuracy was 86.7 % of CT cases. In [16] performed a deep learning model on the gathered images of chest CT scan dataset for identifying the impacts of COVID-19.

The purpose of this study is is two-fold. Firstly, COVID-19 chest X-ray image dataset using data augmentation and and histogram equalization technique will be expaneded. Secondly, COVID-19 detection model based new CNN layers and Keras library to predict COVID-19 in X-ray image dataset will be built and trained. The images will be fed into the new CNN architecture and trained to compute the network weights. The contributions of this research work can be defined as follows: A new CNN architecture is presented rather than using transfer learning techniques as pretrained network. The proposed CNN architecture is simpler than pre-trained networks, and it was evaluated and tested on augumented and contrast-enhanced images for classifying into two classes (COVID-19 and normal).

2. THE PROPOSED COVID-19 DETECTION MODEL

The feature extraction and dataset expansion of COVID-19 chest X-ray images utilizing deep learning on Google CoLab based graphics prossesing unit (GPU) are suggested in this study. The overall methodology for the proposed COVID-19 detection system is clearly comprised of four phases, starting with the phase of collection COVID-19 dataset, preparation phase in terms of dataset expansion using data augmentation, histogram equalization and data enhancement by preprocessing of the images, following by features extraction phase by designing the model of CNN layers based Keras library, and ending by the detection phase. The designed model will be trained until the desired accuracy achieved, then well-trained model will be saved and used in the detection stage. The full procedures of building the proposed model described in the following phases and in Figure 1 and Figure 2.

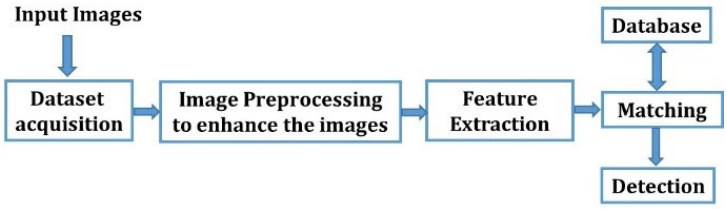


Figure 1. The Fundamental steps of the detection model

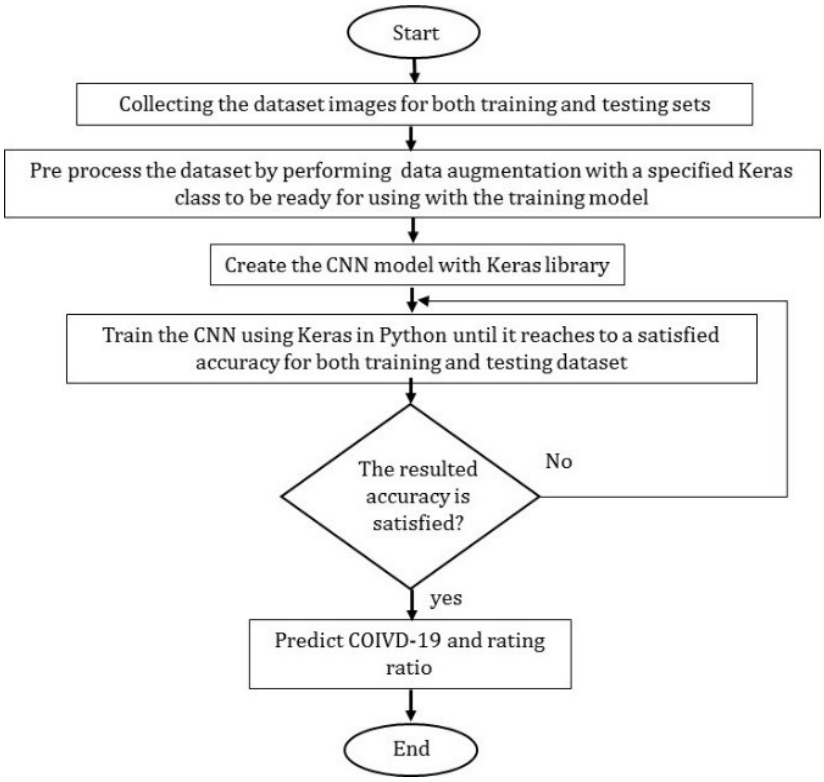


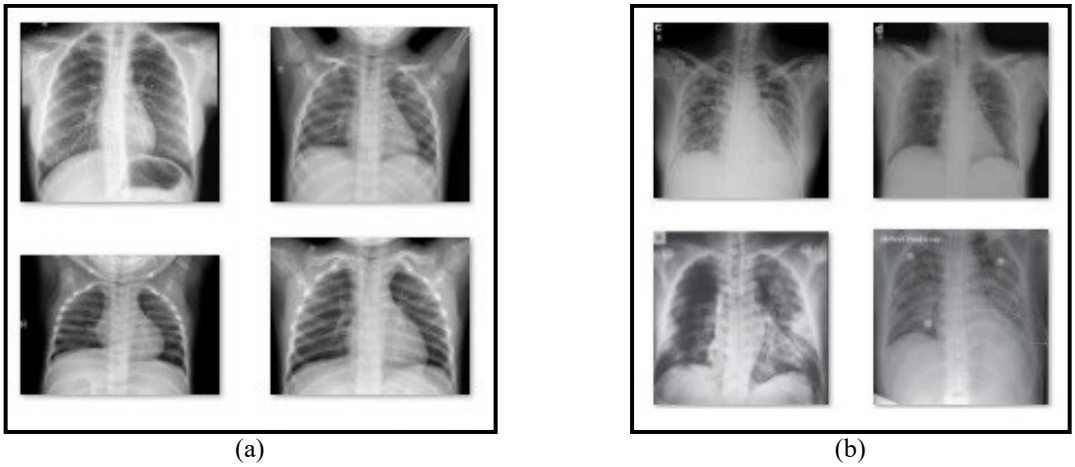
Figure 2. Flowchart describes the software part of the proposed model

2.1. Dataset preparing stage

In this part, the dataset acquisition and pre-processing operations will be presented.

2.1.1. COVID-19 dataset collection

The first step in the proposed system is the dataset acquisition of COVID-19 chest X-ray images from three different sources the open source GitHub repository introduced by Cohen [17], [18] and the Kaggle Chest X-rays [19]. The total number of collected chest X-rays images was 1102, the dataset consists of 500 COVID-19 cases from GitHub [17], 37 COVID-19 cases from GitHub [18] and 565 normal cases from Kaggle [19]. The COVID-19 raw dataset collected in JPG format with various sizes and 24-bit depth of X-ray images as a set of two categories, as illustrated in Figure 3.



(a) (b)

Figure 3. Samples of COVID-19 chest X-ray image dataset: (a) negative (uninfected) and (b) positive (infected) cases

2.1.2. Dataset augmentation and pre-processing

The objective of this step is to expand the dataset by increasing the number of images in each class using different data augmentation techniques and histogram equalization techniuque. The Keras deep learning library provides the library to use data augmentation and preprocessing automatically when training a model. This is achieved by using image data generator class [20], [21] based on necessary operations with a random parameters, such as the images shifting, rotation, cropping and resizing in order to make them in the same size, as demonstrated in Figure 4 the samples of the augmented dataset and in Table 1, parameters of the data augmentation techniques. While the contrast of the images is enhanced using histogram equalization techniuque by transforming the values in an intensity image inorder to match a specified histogram of the output image approximately matches [22]. The quality of the enhancement results has been increased as presente in Figure 5.

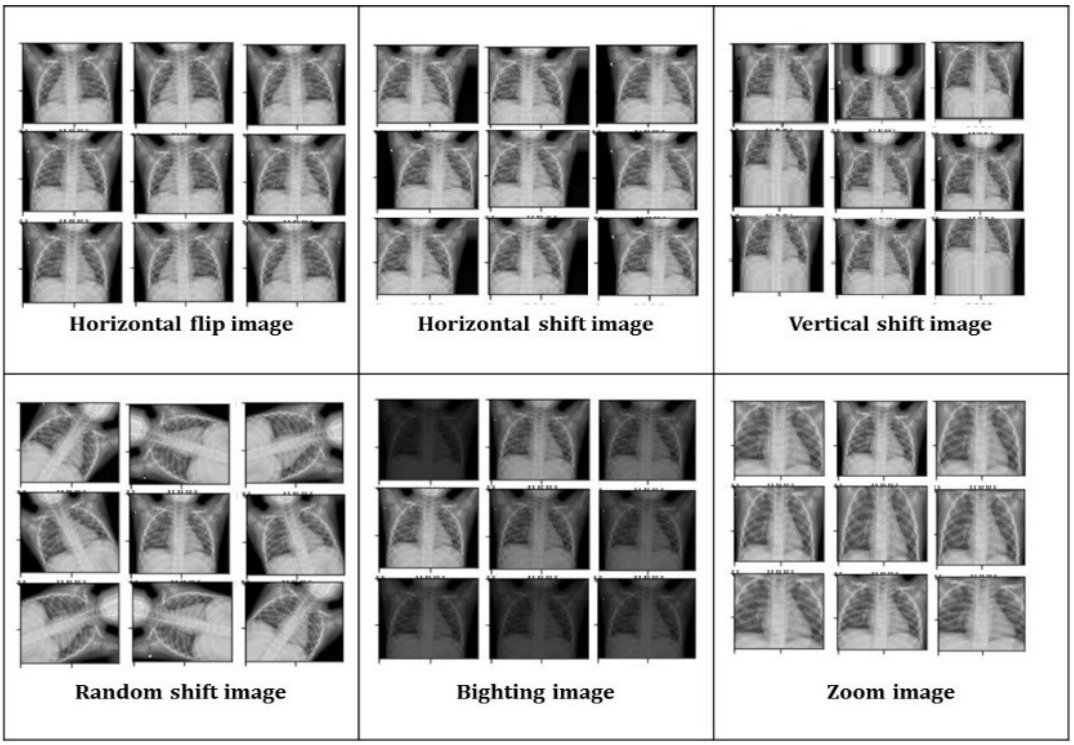


Figure 4. Samples of COVID-19 dataset expansion

Table 1. Overview of the used parameters of the data augmentation techniques

Name	Value
Scaling	1./255
Zooming	25%
Flipping	Vertical/Horizontal
Rotation	15
Shifting	Width/Height with range 0.05
Brightness	[0.1, 1.5]
Filing	nearest
Shearing	0.05
Crops	224*224

Figure 4 and Figure 5 presented the collected COVID-19 dataset after expanding using different data augmentation techniques, enhancement by image preprocessing and histogram equalization techniuque, respectively. After expanding the datase for the COVID-19 Chest X-ray images utlizing data augmentation and histogram equalization technique, the total number of the dataset became 5865 images.

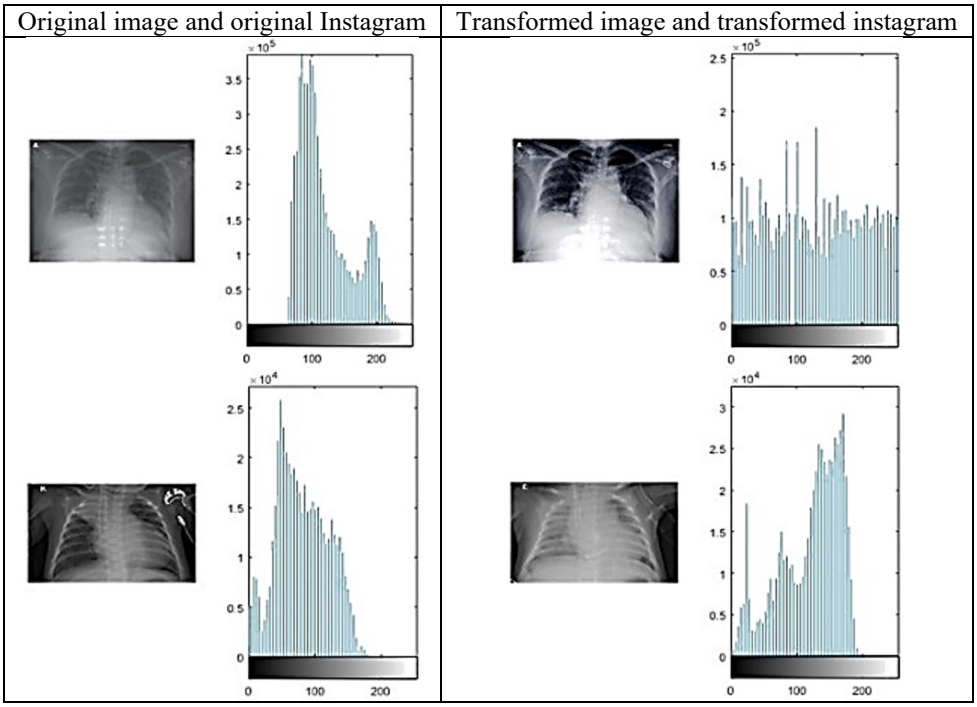


Figure 5. Histogram equalization applied to low contrast image

2.2. Training the COVID-19 detection system

The deep learning model based deep CNN model is proposed to predict automatically whether a patient has COVID-19 or not. The proposed model consists mainly of 10 layers including different types of layers in terms of convolutional layer, pooling layer and fully connected layer, which are combined together to introduce a well-defined CNN model and can be presented in Figure 6 and the used operators in this model as defined in Table 2.

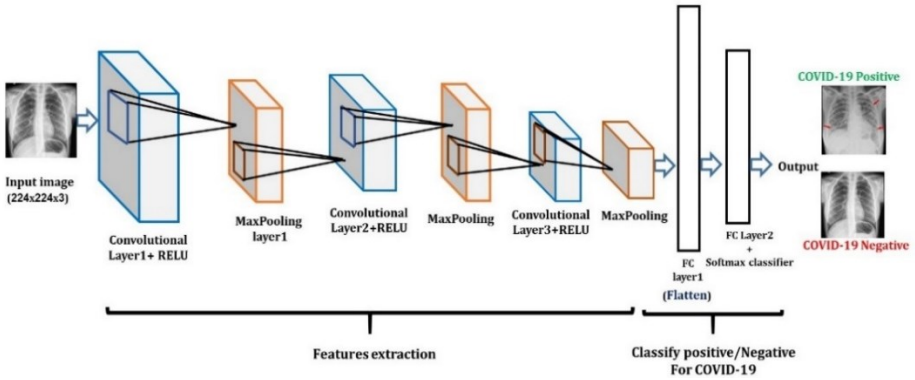


Figure 6. The used CNN architecture the proposed COVID-19 detection model

Table 2. The used operators in the proposed COVID-19 detection model

Parameters	Values
Patch size (PS)	32
Learning Rate (LR)	1e-4
Optimizer	Adam
Classification type	Categorical
Epochs	12

2.3. Implementation using Google golaboratory

All experimental processes of the proposed model are performed using Google because it represents a great cloud service that contains a complete library of Keras, high experimental specification based on Tesla P100

PCI-E 16GB GPU as free GPUs, 12.72GB RAM and cloud runtime with fully configuration for deep learning implementation. It allows users to interact with a sever through a Jupyter notebook environment [23], [24].

3. RESULTS AND DISCUSSION

The proposed model used deep learning based CNN and Keras library for detecting COVID-19 on a dataset of X-ray images. This research work focuses on using deep leraning technology based CNN to classify the COVID-19 chest X-ray image dataset. Dataset based on COVID 19 chest X-ray images will be collected and expanded utlizing data augmentation and then contrast-enhanced images using histogram equalization technique. Training simulations are implemented online on Google Colab. The overall methodology of the proposed model is designed and dealt with 80% and 20% of dataset for training and testing part, respectively. In the CNN model, the inputted image passes through different levels of convolution layers including a filters and pooling, fully connected layers and a Softmax filter. The training precess of the designed model will be carried out until the desired accuracy is reached, followed by saving a well-trained model and then prediction step to classify the image to its corresponding case. The full procedures of building the proposed model described in the following phases in Figure 7.

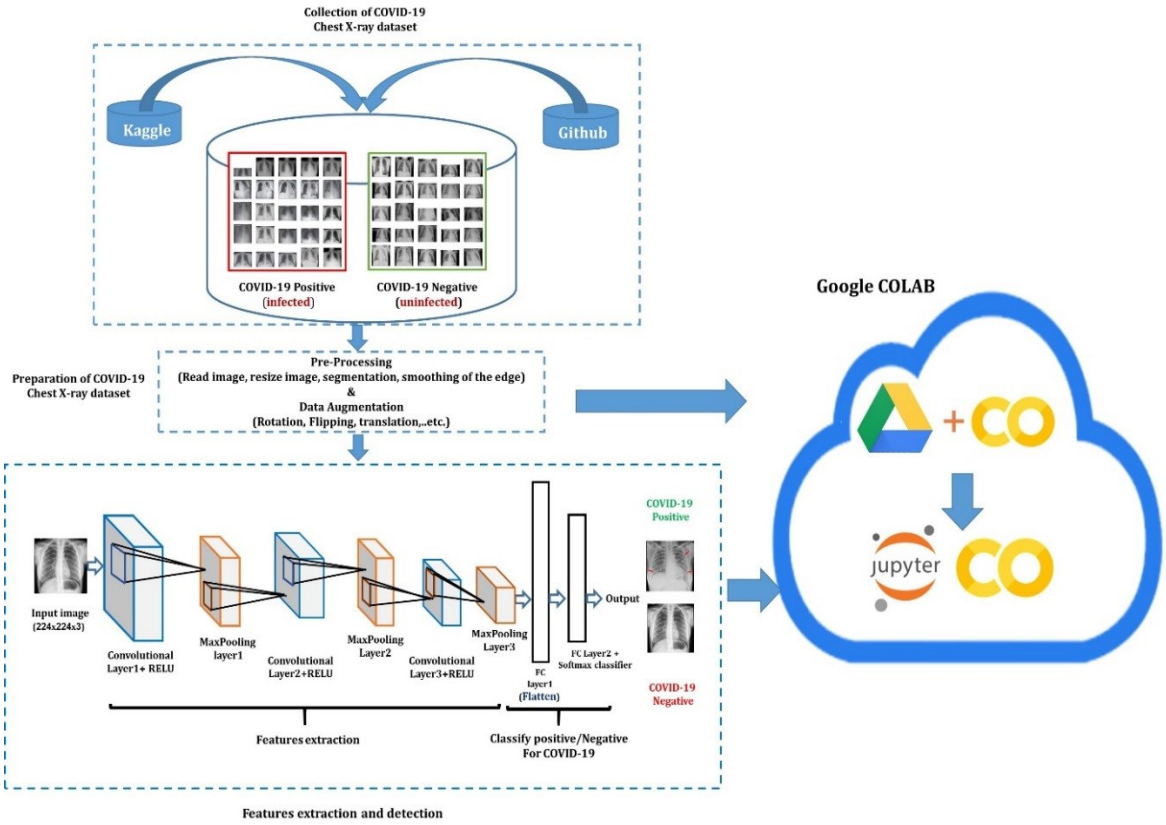


Figure 7. The overall methodology of the proposed COVID-19 detection model

The accuracy is calculated as follows:

$$\text{Accuracy} = \frac{\text{Number of the correctly classified samples}}{\text{Total number of samples}} * 100\%$$

The obtained training accuracy is 93.8% and 92.1% for training and testing dataset, respectively. It is worth mentioning that, the trained model achieved satisfying results with small number of epochs to produce the accuracy and loss metrics for training and validation dataset, as can be seen in the results in Figure 8 the training history plot presenting the accuracy and loss curves. A graphical user interface (GUI) for COVID-19 detection model with Tkinter is presented. Tkinter represents a GUI toolkit using the standard python library [25]. The GUI has been designed for uploading and classifying the COVID-19 image, as can be noticed in Figure 9.

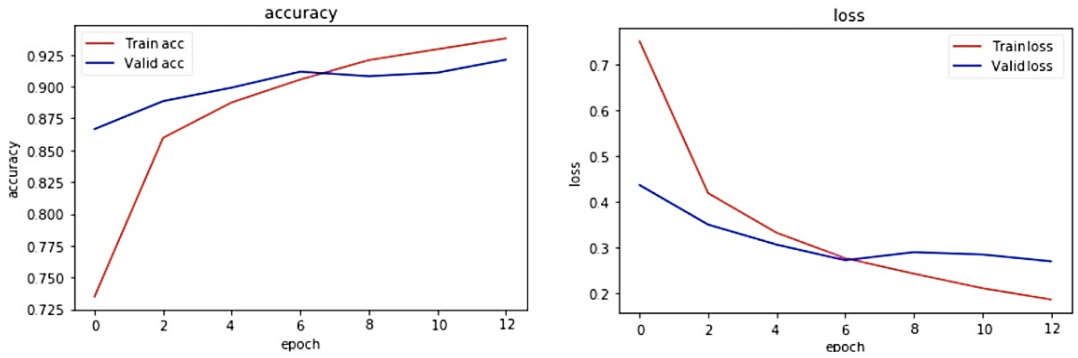


Figure 8. Accuracies of the training and validation according to each epoch

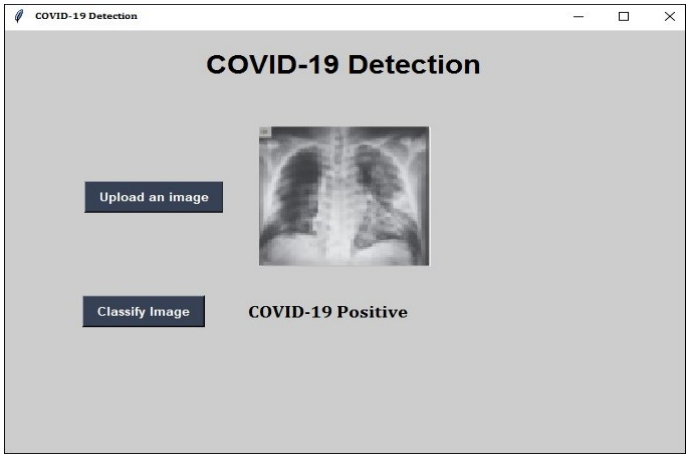


Figure 9. Designed GUI for COVID-19 detection system

4. CONCLUSION

In this study, deep learning model based new CNN architecture using Keras library is presented for automatic COVID-19 detection in raw and augmented chest X-ray images. The COVID-19 dataset as chest X-ray images was collected and pre-processed, followed by expanding the dataset using data augmentation and histogram equalization. Then, CNN architecture was built and trained to recognize cell images as Infected/Uninfected. The overall accuracy of the proposed model is 93.8% and 92.1% for training dataset and for validating dataset, respectively. And finally, for the future work the proposed model can be tested with a larger database and can be implemented based multi-modal system.

REFERENCES

[1] World Health Organization (WHO), "Pneumonia of unknown cause - China," *Emergencies Preparedness, Response, Disease Outbreak News, WHO*, Jan. 2020. Available: https://www.who.int/emergencies/disease-outbreak-news/item/2020-DON229

[2] European Centre for Disease Prevention and Control, "An agency of the European Union," 2019. Available: https://www.ecdc.europa.eu/en/geographical-distribution-2019-ncov-cases

[3] T. Ozturk, M. Talo, E. A. Yildirim, U. B. Baloglu, O. Yildirim, and U. R. Acharya, "Automated detection of COVID-19 cases using deep neural networks with X-ray images," *Computers in Biology and Medicine*, vol. 121, no. 103792, pp. 1-11, June 2020, doi: 10.1016/j.compbiomed.2020.103792.

[4] S. Wang, *et al.,* "Predicting EGFR mutation status in lung adenocarcinoma on computed tomography image using deep learning," *European Respiratory Journal*, vol. 53, no. 3, pp. 1-11, 2019, doi: 10.1183/13993003.00986-2018.

[5] S. L. F. Walsh, S. M. Humphries, A. U Wells, and K. K Brown, "Imaging research in fibrotic lung disease; applying deep learning to unsolved problems," *The Lancet Respiratory Medicine*, vol. 8, no. 11, pp. 1144-1153, Feb. 2020, doi: 10.1016/S2213-2600(20)30003-5.

[6] E. Angelini, S. Dahan, and A. Shah, "Unravelling machine learning: insights in respiratory medicine," *European Respiratory Journal*, vol. 54. no. 6, 2019, doi: 10.1183/13993003.01216-2019.

[7] S. Wang, *et al.*, "A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis," *European Respiratory Journal*, vol. 56, no. 2, pp. 1-44, 2020, doi: 10.1183/13993003.00775-2020.

[8] M. Chung, *et al.,* "CT imaging features of 2019 novel coronavirus (2019-nCoV)," *Radiology*, vol. 295, no 1, pp. 202-207, 2020, doi: 10.1148/radiol.2020200230.
[9] S. H. Yoo *et al.,* "Deep learning-based decision-tree classierfor COVID-19 diagnosis from chest X-rayImaging," *Front Med (Lausanne)*, vol. 7, no. 427, pp. 1-8, 2020, doi: 10.3389/fmed.2020.00427.
[10] M. A. Elaziz, K. M. Hosny, A. Salah, M. M. Darwish, S. Lu, and A. T. Sahlol, "New machine learning method for image-based diagnosis of COVID-19," *PLoS ONE*, vol. 15, no. 6, p. e0235187, 2020, doi: 10.1371/journal.pone.0235187.
[11] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using X-ray images and deep convolutional neural networks," *Pattern Analysis and Applications*, pp. 1-14, 2021, doi: 10.1007/s10044-021-00984-y.
[12] S. S. Yadav and S.M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *Journal Big Data*, vol. 6, no. 1, p. 113, 2019, doi: 10.1186/s40537-019-0276-2.
[13] W. J. Choi, L. Ki-Nam, K. Eun-Ju Kang, and L. Hyuck, "Middle East respiratory syndrome-coronavirus infection: a case report of serial computed tomographic findings in a young male patient," *Korean Journal Radiol,* vol. 17, no. 1, pp. 166-170, 2016, doi: 10.3348/kjr.2016.17.1.166.
[14] E. El-Din Hemdan, M. A. Shouman, and M. E. Karar, "Framework of deep learning classifiers to diagnose covid-19 in X-ray images," *arXiv preprint arXiv:2003.11055*, pp. 1-14, 2020.
[15] C. Butt, C. J. Gill, D. Chun, and B. A. B. "Deep learning system to screen coronavirus disease 2019 pneumonia," *Appl Intell* 2020, doi: 10.1007/s10489-020-01714-3.
[16] A. M. V. Dadário, *et al.,* "Regarding artificial intelligence distinguishes COVID-19 from community acquired pneumonia on chest CT," Apr. 2020, doi: 10.1148/radiol.2020201178.
[17] J. P. Cohen, P. Morrison, and L. Dao, "COVID-19 Image Data Collection," Mar. 2020. Available: https://ui.adsabs.harvard.edu/abs/2020arXiv200311597C
[18] A. Chung, "COVID-chestxray-dataset," 2020. Available: https://github.com/agchung/Figure1-COVID-chestxray-dataset
[19] P. Mooney Kaggle, "Kaggle's Chest X-Ray Images (Pneumonia) dataset," 2020. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia
[20] F. Chollet, "Deep Learning with Python", book of *Simon and Schuster*, 2017.
[21] R. C. Gonzalez and R. E. Woods, "Digital Image Processing," *Third Edition, Pearson International Edition prepared by Pearson Education*, 2008.
[22] M. Thoma, "Analysis and optimization of convolutional neural network architectures," *arXiv:1707.09725*, 2017.
[23] T. Carneiro, R. V. Medeiros Da NóBrega, T. Nepomuceno, G. Bian, V. H. C. De Albuquerque, and P. P. R. Filho, "Performance Analysis of Google Colaboratory as a Tool for Accelerating Deep Learning Applications," in *IEEE Access*, vol. 6, pp. 61677-61685, 2018, doi: 10.1109/ACCESS.2018.2874767.
[24] Google colabirity. 2018. Available: https://www.kdnuggets.com/2018/02/google-colab-free-gpu-tutorial tensorflow-keras-pytorch.html
[25] "Tkinter - Python interface to Tcl/Tk - Python v2.6.1 documentation," 2021. Available: https://docs.python.org/3/library/tkinter.html

BIOGRAPHIES OF AUTHOR

Bashra Kadhim Oleiwi Chabor Alwawi, born in Baghdad-Iraq, completed a Master degree in Mechatronics Engineering/Control and Systems Engineering Department at University of Technology (UOT) Bagdad-Iraq. She finished her PhD degree at Control Engineering Department (RST), Siegen University, Germany. She is currently working as a faculty member in control and systems engineering department at UOT. Her research interests are Deep learning, Machine learning, Artificial intelligence systems, IOT, Optimization, Image processing, Robotics, Path planning, Automation, control, Microcontrollers and Microprocessor (Arduino and Raspberry), Sensors.

Layla H. Abood received her B.Eng. and M.Sc. and PHD degrees in Electronic and Communication Engineering from the University of Technology-Baghdad. She is currently an Academic staff member in the Department of Control and System Engineering, University of Technology, Baghdad, Iraq. Her research interest is Artificial Intelligent, Optimization Technique, Modeling & Simulation, and FPGA and Embedded systems.