International Journal of Electrical and Computer Engineering (IJECE)
Vol. 12, No. 5, October 2022, pp. 5573~5579
ISSN: 2088-8708, DOI: 10.11591/ijece.v12i5.pp5573-5579

Automatic COVID-19 lung images classification system based on convolution neural network

Lamia Nabil Mahdy[1], Ahmed Ibrahim Bahgat El Seddawy[2], Kadry Ali Ezzat[1]
[1]Biomedical Engineering Department, Higher Technological Institute, 10th of Ramadan City, Egypt
[2]Computer Science and Information System, Arab Academy for Science and Technology and Maritime Transport, Giza, Egypt

Article Info

Article history:

Received Aug 27, 2021
Revised May 24, 2022
Accepted Jun 20, 2022

Keywords:

Classification
Computed tomography
Convolution neural network
COVID-19
X-Ray

ABSTRACT

Coronavirus disease (COVID-19) still has disastrous effects on human life around the world. For fight that disease. Examination on the patients who have been sucked in quick and cheap way is necessary. Radiography is most effective step closer to this target. Chest X-ray is readily obtainable and cheap option. Also, because COVID-19 is a virus, distinguish COVID-19 from common viral pneumonia from common viral pneumonia is difficult. In this study, X-ray images of 500, 500, 500, and 500 patients for healthy controls, typical viral pneumonia, bacterial pneumonia and COVID-19, were collected respectively. To our knowledge, this was the first quaternary classification study that also included classical viral pneumonia. To efficiently capture nuances in X-ray images, a new model was created by applying convolution neural network for accurate image classification. Our model outperformed to achieve an overall accuracy, sensitivity, specificity, F1-score, and area under curve (AUC) of 0.98, 0.97, 0.98, 0.97, and 0.99 respectively.

Corresponding Author:

Kadry Ali Ezzat
Department of Biomedical Engineering, Higher Technological Institute
Next to Small Industries Complex, Industrial Area, 10th of Ramadan City, Ash Sharqia Governorate
Email: kadry_ezat@hotmail.com

1. INTRODUCTION

Chest computed tomography (CT) performs essential part in patients evaluation with coronavirus disease (COVID-19). So far, deep learning (DL)-based chest CT analysis has mainly focused on diagnosing or classifying COVID-19 in order to distinguish between positive COVID-19 patients and patients with normal pneumonia. The widespread implementation of reverse transcription polymerase chain reaction (RTPCR) testing reduces the need to use screening radiographs. However, radiographic imaging is still important to detect lung intrusion in COVID-19 patients. In particular, it has been proposed to use CT scans of lesions as surrogate imaging of disease burden and to identify critically ill patients requiring hospitalization.

As assured by world health organization [1], coronavirus recorded the first formal cases in Wuhan, China. To date, millions of people have been killed, and there are millions of confirmed cases worldwide. Epidemics in the form of pandemics have a devastating impact on health of people all over world [2]. This infectious infection, expands quicker than other common flus. RTPCR used to distinguish COVID-19 [3]. This examination seems to be accurate. Even if the patient's sample contains only a small amount of the virus. However, polymerase chain reaction (PCR) examinations are complex, take long time and expensive. Therefore, not all health care institutions are able to do this. With these limitations in mind, a standalone approach to disease detection is to analyze chest radiographs to detect the presence or absence.

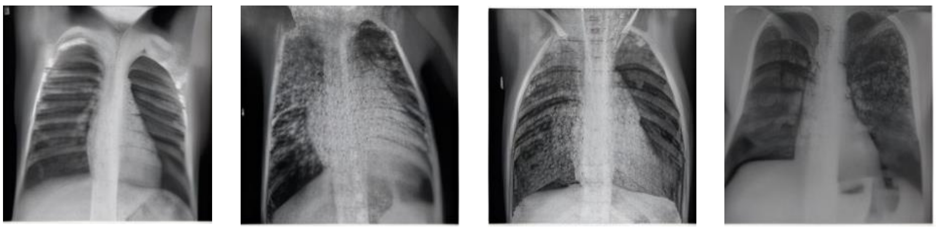
The main problem with radiographic-based diagnosis of COVID-19 patients is the lack of trained and well-available physicians. Furthermore, the radiation symptoms associated with COVID-19 are unknown, as many specialized have no expertise with COVID-19 infected patients. Recently, many methods have emerged that use artificial intelligence to distribute pneumonia and COVID-19 from CT or X-rays. There are basically two types: COVID-19 to non-COVID-19 and COVID-19 to other pneumonia. Xu *et al.* [4] introduced a distinguish system for classifying COVID-19, influenza, and healthy individuals with 86.7% accuracy. Li *et al.* [5] ResNet50 was used to distinguish COVID-19 from non-pneumonia with 90% sensitivity. Chen *et al.* [6], Ouyang. [7], Apostolopoulos *et al.* [8]. However, because COVID-19 is also a variant of the virus, all this work ignores the typical viral pneumonia that infects the most difficult and typical virus.

In this study, X-ray images of 500, 500, 500, and 500 patients for healthy controls, typical viral pneumonia, bacterial pneumonia and COVID-19, were collected respectively. According to our knowledge, this was the first quaternary classification study that also included classical viral pneumonia. To efficiently capture nuances in X-ray images, a new model was created by a recently developed convolution neural network for accurate image analysis. Our model outperformed the commonly used reference model, achieving an overall accuracy, sensitivity, specificity, F1-score, and AUC of 0.98, 0.97, 0.98, 0.97, and 0.99 respectively.

2. METHOD

2.1. Data set and model inputs

In current work, chest X-ray obtained from various open sources [9]–[15]. This data set contains 4,900 X-ray images, categorized into three sectors, namely natural, COVID-19 and pneumonia. The pneumonia category consists of bacterial pneumonia images and viruses pneumonia images. In Figure 1, the chest X-ray imaging dataset is rearranged into four types: normal, bacterial pneumonia, viral pneumonia, and COVID-19 as shown in Figures 1(a) to 1(d) respectively.



(a) (b) (c) (d)

Figure 1. Chest X-ray samples from the test datasets (a) normal, (b) bacteria, (c) virus and (d) COVID-19

2.2. Pre-processing operations

Since this data set consists of large and different lung images; and in order to handle reasonable computing time during the convolutional neural network (CNN) training experiment, all images are adjusted to a single dimension and rescaled to smaller images. 310×310 to adapt to the standard input of a custom structure. To standardize the input of CNN, the original image is preliminarily divided before the resizing step. Regarding the test, the trained model was tested on distinct dataset taken from IEEE DataPort [16].

2.3. Data augmentation

Data augmentation is method may drastically growth number of data times in dataset for training a model [17]. For the four categories image datasets (healthy controls, typical viral pneumonia, bacterial pneumonia and COVID-19, this method utilizes basal image procedures like flip, rotate, crop, and pad for enlargement. Next, the dataset will be expanded to include these converted images resulting from subsisting image set. This enlarges dataset size for training neural networks [18]–[24]. A data expansion method was used in this study to fix availability of few datasets problem that impacts proposed CNN performance. This technique has expanded the dataset. Moreover, extra learning countenances are available in knowledge model. In this study, image transformation procedures, mirroring, and rotation, were utilized for data expansion. In initial stage of data expansion, 100 X-ray images were mirrored to capture an additional 100 images. After using this operation, the resulting dataset was expanded to 180 images. In second stage, authentic 100 images are rotated another 90° to obtain another 100 images, after that, rotated 180 degrees to obtain another 100 images, and finally the original 100 images rotated the image of 270 degrees further and 100 more images were obtained. These operations created 500 datasets each for healthy controls, typical viral pneumonia, bacterial pneumonia, and COVID-19 X-rays. Table 1 introduces image transformation procedures applied on all image types and the related images generated number by operations. Figure 2 introduces impact of the extension technique performed to authentic image sample of dataset utilized in this work. Figure 2(a) is the ordinary image, Figure 2(b) is the rotated image with 90°, Figure 2(c) is the rotated image with 180°, Figure 2(d) is the rotated image with 270° and Figure 2(e) is the flipped image.

Table 1. Total image number after augmentation

Image augmentation operation	Number of healthy controls images	Number of typical viral pneumonia images	Number of bacterial pneumonia images	Number of COVID-19 images
Ordinary	100	100	100	100
Rotation of ordinary images by 90˚	100	100	100	100
Rotation of ordinary images by 180˚	100	100	100	100
Rotation of ordinary images by 270˚	100	100	100	100
Flipping of ordinary images	100	100	100	100
Total	500	500	500	500

(a) (b) (c) (d) (e)

Figure 2. X-ray image augmentation methods on an (a) ordinary, (b) 90° rotation, (c) 180° rotation, (d) 270° rotation, and (e) flipping

2.4. Convolutional neural networks (CNNs)

CNNs are invigorated through the human brain’s visual framework. So, the background of CNNs is to allow computers to see the world people see. Therefore, CNNs can be utilized in image analysis and classification [25]. CNN is type of deep neural network with convolution, max pool, and non-linear activation level. The convolutional layer, which can be viewed as the CNN kernel layer, applies process known as "convolution," which CNN is named after. The base of convolutional layer is performed to input layer. Convolutional layers outputs are decreased to feature maps. In this work, rectified linear unit (ReLU) was utilized for activation task with a convolutional layer that helps to expand nonlinearity of input image because image is nonlinear. Therefore, CNNs with ReLU are easier and faster in the current situation. Since ReLU characterized as in (1).

$$Z = Max(0, i) \tag{1}$$

Here the function assumes that the output z is zero for all negative values, and positive values keep i constant.

2.4.1. The proposed system

As appeared in Figure 3, the study starts with the series of essential dataset consists of four image categories: healthy controls, typical viral pneumonia, bacterial pneumonia, and COVID-19. The X-ray image was preprocessed concurring to the preprocessing above steps. Dataset expanded utilizing standard expansion methods to expand its size. Resulting dataset was utilized to train model in next stage. Then, model's disease detection performance was tested. The proposed CNN model tests were done utilizing primary dataset test and independent validation dataset. Table 2 details the dataset, including training sets, test sets, and number of X-ray images for the four prediction classes. Finally, aggregation image-level predictions for image of each case were done to provide the reasonable result.

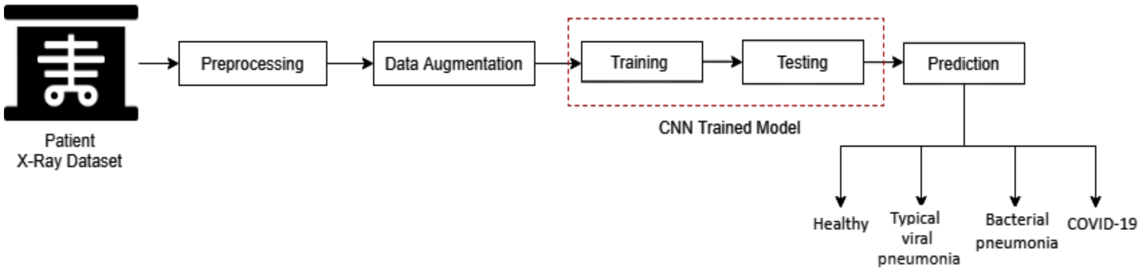


Figure 3. System workflow

Table 2. Images number for training and testing

Dataset	Healthy Controls	Typical Viral Pneumonia	Bacterial Pneumonia	COVID-19	Total
Images number	500	500	500	500	2,000
Training	400	400	400	400	1,600
Testing	100	100	100	100	400

3. RESULTS AND DISCUSSION

The proposed classification system for COVID-19 diagnostics was implemented on a processor with the programming language Python 3.8 and Intel R Core i58300H CPU @ 2.30GHz x 8 and 8 GB RAM, Windows 10, RADEON AMD 1050, and 4 GB graphics. Google Collab GPU, Python 3.7 and Tensor Flow 2.2.0. The Tensor Flow 2.2.0 deep learning library is used to implement CNN, and training and testing steps are performed on the Google Collab platform.

In this model, CNN used to effectively classify COVID-19. Proposed technique is trained on four public datasets and outperforms itself in all classes. CNN is used for the final classification with accuracy, sensitivity, specificity, F1 score and AUC giving the best results of 0.98, 0.97, 0.98, 0.97, and 0.99 respectively as in Table 3. So that, this approach using X-ray imaging and computational diagnostics can be used as a large, faster, and cheaper screening method. It also reduces the required testing time. Training with larger datasets and field testing with larger cohorts can be very helpful in making clinically effective predictions of COVID-19.

Table 3. Model overall performance primarily based totally on unbiased validation data

Evaluation parameters	Degree
Accuracy	0.98
Sensitivity	0.97
Specificity	0.98
F-1 score	0.97
ROC AUC	0.99

Receiver operating characteristic area under curve (ROC AUC)

To examine the overall performance, classification accuracy, sensitivity, specificity and F1-score, measured:

$$\text{Accuracy}=\frac{\text{TPR + TNR}}{\text{TPR + FPR + TNR + FNR}} \tag{2}$$

$$\text{Sensitivity}=\frac{\text{TP}}{\text{TP+FN}} \tag{3}$$

$$\text{Specificity}=\frac{\text{TN}}{\text{TN+FP}} \tag{4}$$

$$\text{F1-score}=\frac{2\text{*TP}}{2\text{*TP+FP+FN}} \tag{5}$$

where TP stands for true positive, TN for true negative, FP for false positive, FN for false negative and FP for false positive TN for true negative. Figure 4 shows train and test losses of the proposed CNN model. As can be seen in Figure 4, proposed CNN does not take much time to converge. Figure 5 shows a graph of training and test accuracy done by the proposed CNN model. Figures 6 and 7 shows the characteristic curve of the receiver operating characteristics at AUC=0.99 and the sensitivity-specificity curves for the various case types respectively.

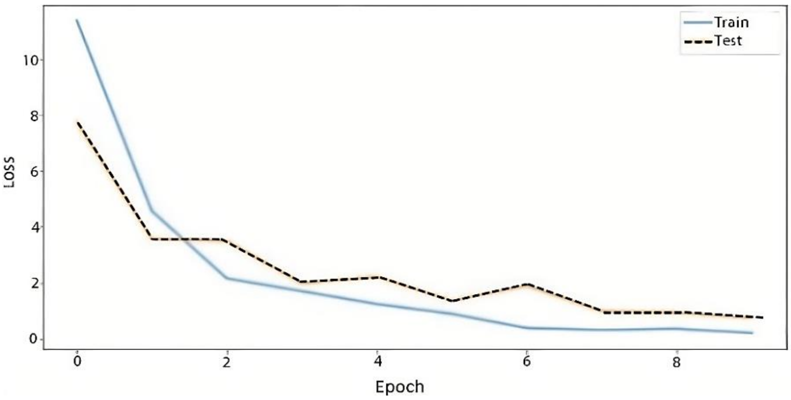


Figure 4. Train and test loss plot by using CNN

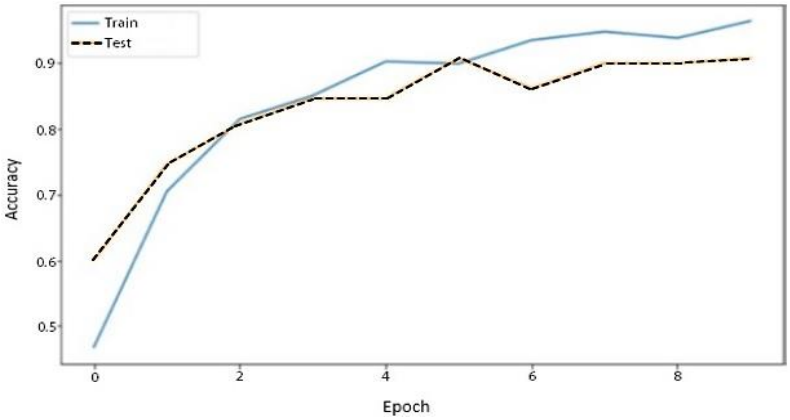


Figure 5. Train and test accuracy plot by using CNN

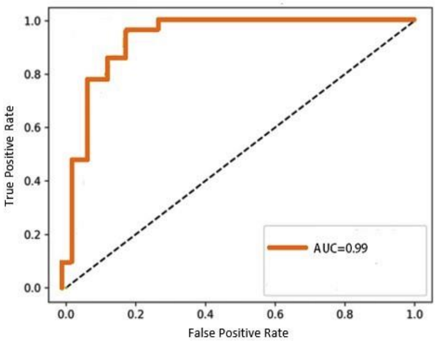


Figure 6. ROC curve of classification

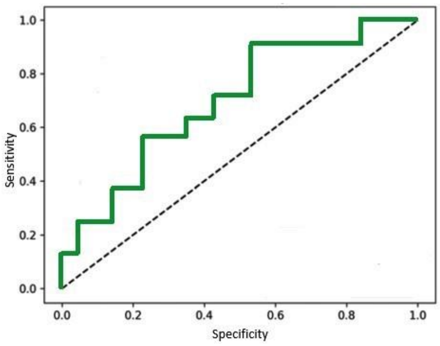
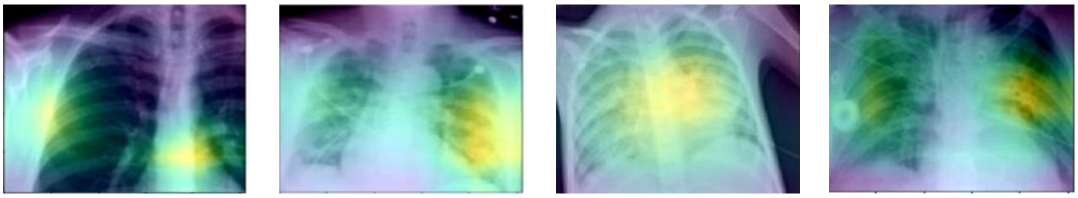


Figure 7. Sensitivity and specificity curve of classification

As evidence of introduced CNN system in classification importance for detecting COVID-19 from X-ray images, trained model was examined on distinct dataset from IEEE DataPort. The distinct test dataset consists of 100 COVID-19 X-rays. 100 regular images were added to dataset for testing. In addition, a feature map of four different X-RAY images types excerption by the classification model and examined overall learning feature then renders possibility of network. As in Figure 8, area that the lesion is existed shows a superior level of response, indicating that the model was eligible to learn underlying properties from radiography of three different patients' cases. Figure 8(a) shows healthy X-Ray image, Figure 8(b) shows COVID-19 X-Ray image, Figure 8(c) shows viral pneumonia image and Figure 8(d) shows bacterial pneumonia image.



(a) (b) (c) (d)

Figure 8. The X-Ray images feature maps four types of people extracted by proposed classification model (a) healthy X-Ray image, (b) COVID-19 X-Ray image, (c) viral pneumonia image, and (d) bacterial pneumonia image

4. CONCLUSION

By applying CNN, X-ray deep learning diagnostic system for rapid COVID-19 diagnosis was created. This model can individually detect COVID-19 X-rays from healthy individuals, X-rays from patients with bacterial pneumonia, and X-rays from patients with typical viral pneumonia. This model able to differentiate so various different kinds of pneumonia at once. Results have shown that introduced model fulfill high accuracy, specificity, sensitivity, AUC ROC and F1 score that shows the reliability of the model.

REFERENCES

[1] D. Cucinotta and M. Vanelli, "WHO declares COVID-19 a pandemic," *Acta bio-medica : Atenei Parmensis*, vol. 91, no. 1, pp. 157–160, 2020, doi: 10.23750/abm.v91i1.9397.

[2] F. Rustam *et al.*, "COVID-19 future forecasting using supervised machine learning models," *IEEE Access*, vol. 8, pp. 101489–101499, 2020, doi: 10.1109/ACCESS.2020.2997311.

[3] R. Lu *et al.*, "Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding," *The Lancet*, vol. 395, no. 10224, pp. 565–574, Feb. 2020, doi: 10.1016/S0140-6736(20)30251-8.

[4] X. Xu *et al.*, "A deep learning system to screen Novel coronavirus disease 2019 pneumonia," *Engineering*, vol. 6, no. 10, pp. 1122–1129, Oct. 2020, doi: 10.1016/j.eng.2020.04.010.

[5] L. Li *et al.*, "Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT," *Radiology*, vol. 296, no. 2, pp. E65--E71, Aug. 2020, doi: 10.1148/radiol.2020200905.

[6] J. Chen *et al.*, "Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography," *Scientific Reports*, vol. 10, no. 1, Dec. 2020, doi: 10.1038/s41598-020-76282-0.

[7] X. Ouyang *et al.*, "Dual-sampling attention network for diagnosis of COVID-19 from community acquired pneumonia," *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2595–2605, Aug. 2020, doi: 10.1109/TMI.2020.2995508.

[8] I. D. Apostolopoulos and T. A. Mpesiana, "COVID-19: automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Physical and Engineering Sciences in Medicine*, vol. 43, no. 2, pp. 635–640, Jun. 2020, doi: 10.1007/s13246-020-00865-4.

[9] J. Rasheed, A. A. Hameed, C. Djeddi, A. Jamil, and F. Al-Turjman, "A machine learning-based framework for diagnosis of COVID-19 from chest X-ray images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 13, no. 1, pp. 103–117, Mar. 2021, doi: 10.1007/s12539-020-00403-6.

[10] Y. Pathak, P. K. Shukla, and K. V Arya, "Deep bidirectional classification model for COVID-19 disease infected patients," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 18, no. 4, pp. 1234–1241, Jul. 2021, doi: 10.1109/TCBB.2020.3009859.

[11] J. P. Kanne, "Chest CT findings in 2019 novel coronavirus (2019-NCoV) infections from Wuhan, China: key points for the radiologist," *Radiology*, vol. 295, no. 1, pp. 16–17, Apr. 2020, doi: 10.1148/radiol.2020200241.

[12] P. Mooney, "Chest X-Ray images (Pneumonia)," *Kaggle*, 2019. Accessed: Sep 1, 2019. [Online]. Available: https://www.kaggle.com/paultimothymooney/chest-xray-pneumonia.

[13] D. Singh, V. Kumar, Vaishali, and M. Kaur, "Classification of COVID-19 patients from chest CT images using multi-objective differential evolution-based convolutional neural networks," *European Journal of Clinical Microbiology and Infectious Diseases*, vol. 39, no. 7, pp. 1379–1389, Jul. 2020, doi: 10.1007/s10096-020-03901-z.

[14] M. Chung *et al.*, "CT imaging features of 2019 novel Coronavirus (2019-nCoV)," *Radiology*, vol. 295, no. 1, pp. 202–207, Apr. 2020, doi: 10.1148/radiol.2020200230.

[15] W. Kong and P. P. Agarwal, "Chest imaging appearance of COVID-19 infection," *Radiology: Cardiothoracic Imaging*, vol. 2, no. 1, Feb. 2020, doi: 10.1148/ryct.2020200028.

[16] Z.-H. Chen, "Mask-RCNN detection of COVID-19 pneumonia symptoms by employing stacked autoencoders in deep unsupervised learning on low-dose high resolution CT," *IEEE DataPort*, 2020, doi: 10.21227/4kcm-m312.

[17] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1,

Dec. 2019, doi: 10.1186/s40537-019-0197-0.
[18] D. Ho, E. Liang, and R. Liaw, “1000x faster data augmentation – the berkeley artificial intelligence research blog,” *Berkeley Artificial Intelligence Research (BAIR)*, 2019.
[19] A. S. Al-Waisy *et al.*, “COVID-CheXNet: hybrid deep learning framework for identifying COVID-19 virus in chest X-rays images,” *Soft Computing*, Nov. 2020, doi: 10.1007/s00500-020-05424-3.
[20] Z. Han *et al.*, “Accurate screening of COVID-19 using attention-based deep 3D multiple instance learning,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2584–2594, Aug. 2020, doi: 10.1109/TMI.2020.2996256.
[21] H. Kang *et al.*, “Diagnosis of coronavirus disease 2019 (COVID-19) with structured latent multi-view representation learning,” *IEEE Transactions on Medical Imaging*, vol. 39, no. 8, pp. 2606–2614, Aug. 2020, doi: 10.1109/TMI.2020.2992546.
[22] S. Albahli, “A deep neural network to distinguish COVID-19 from other chest diseases using X-ray images,” *Current Medical Imaging Formerly Current Medical Imaging Reviews)*, vol. 17, no. 1, pp. 109–119, Mar. 2021, doi: 10.2174/1573405616666200604163954.
[23] A. Jaiswal, N. Gianchandani, D. Singh, V. Kumar, and M. Kaur, “Classification of the COVID-19 infected patients using DenseNet201 based deep transfer learning,” *Journal of Biomolecular Structure and Dynamics*, vol. 39, no. 15, pp. 5682–5689, Oct. 2021, doi: 10.1080/07391102.2020.1788642.
[24] R. Zhang *et al.*, “COVID19XrayNet: a two-step transfer learning model for the COVID-19 detecting problem based on a limited number of chest X-Ray images,” *Interdisciplinary Sciences: Computational Life Sciences*, vol. 12, no. 4, pp. 555–565, Dec. 2020, doi: 10.1007/s12539-020-00393-5.
[25] D. T. Mane and U. V Kulkarni, “A survey on supervised convolutional neural network and its major applications,” *International Journal of Rough Sets and Data Analysis*, vol. 4, no. 3, pp. 71–82, Jul. 2017, doi: 10.4018/IJRSDA.2017070105.

BIOGRAPHIES OF AUTHORS

Lamia Nabil Mahdy received B.Sc. degree in biomedical engineering from Helwan University, Cairo, Egypt in 2004 and the M.Sc. and Ph.D. degree in biomedical engineering from Cairo University, Egypt, in 2009 and 2016 respectively. In 2004 she joined the biomedical engineering department in the higher technological institute as researcher assistant then she promoted as assistant Lecturer in 2009 and then she promoted to be Lecturer in 2016. Her current research interests are: diagnostic imaging, robotics, hospital design, artificial intelligence, reproductive surgery, image processing, expert systems, biomechanics, computer programming, data transmission, data structures, biomedical instrumentation and electronics, pattern recognition, microcontrollers, modeling and simulation, internet of things (IoT). She is now Lecturer (Ph.D.) in Biomedical Engineering Department at Higher Technological Institute in 10th of Ramadan city since 2016, Member in Scientific Research Group in Egypt (SRGE), Head of Bright Minds students (BMS) Research Group in Higher Technological Institute; she supervised on more than 13 under graduate bachelor projects. She can be contacted at email: englamia_82@yahoo.com.

Ahmed Ibrahim Bahgat El Seddawy is Associated Professor, Vice Dean of Educational Affairs, and Head of Business Information System Department. He received his bachelor from Business Administration, MIS department and M.Sc. degree in college of Computer Science and Information Systems, I.S Department 2008 from, both from AASTMT. Cairo, Egypt. On December 2014, he received his doctoral degree from the Department of Information Systems, College of Computer Science, and Information Systems, I.S department, Helwan University. Working as visitor professor in Ludwigsburg university in Germany and Proto University Italy. He has authored/co-authored several research publications in Selected fields such as, knowledge discovery, data science, data analytics decision support systems, knowledge base systems, project management, and leadership management using ERP. He can be contacted at email: ahmed.bahgat@aast.edu.

Kadry Ali Ezzat received B.SC degree in biomedical engineering from the higher technological institute, Cairo, Egypt in 2004 and the M.Sc. and Ph.D. degree in biomedical engineering from Cairo University, Egypt, in 2011 and 2017 respectively. In 2004 he joined the biomedical engineering department in the higher technological institute as researcher assistant then he promoted as assistant Lecturer in 2011 and then he promoted to be Lecturer in 2017. His current research interests are: diagnostic imaging, robotics, satellite communications, artificial intelligence, image processing, expert systems, biomechanics, data transmission, data structures, biomedical instrumentation and electronics, pattern recognition, microcontrollers, modeling and simulation. He is now Lecturer in Biomedical Engineering Department at Higher Technological Institute in 10th of Ramadan city since 2017, Member in Scientific Research Group in Egypt (SRGE), Consultant member of the Egyptian Remote Sensing and Space Science Authority (Satellite Alliance Project), Instructor in High Technology Center, Faculty of Engineering, Cairo University. He is reviewer for papers in many international conferences: Cairo International Biomedical Engineering Conference (CIBEC), International Undergraduate Research Conference in Military Technical College (IUGRC). He is reviewer in Australasian Physical and Engineering Sciences in Medicine (APES) journal. He can be contacted at email: kadry_ezat@hotmail.com.