Employing deep learning for lung sounds classification

Huda Dhari Satea[1], Amer Saleem Elameer[2], Ahmed Hussein Salman[1], Shahad Dhari Sateaa[1]
[1]Department of Computer Technologies Engineering, AL-Esraa University College, Baghdad, Iraq
[2]Biomedical Informatics College, University of Information Technology and Communications, Baghdad, Iraq

Article Info

Article history:

Received Jul 27, 2021
Revised Dec 29, 2021
Accepted Jan 21, 2022

Keywords:

Convolution neural network
Deep learning
Lung sounds
Respiratory diseases
Transfer learning

ABSTRACT

Respiratory diseases indicate severe medical problems. They cause death for more than three million people annually according to the World Health Organization (WHO). Recently, with coronavirus disease 19 (COVID-19) spreading the situation has become extremely serious. Thus, early detection of infected people is very vital in limiting the spread of respiratory diseases and COVID-19. In this paper, we have examined two different models using convolution neural networks. Firstly, we proposed and build a convolution neural network (CNN) model from scratch for classification the lung breath sounds. Secondly, we employed transfer learning using the pre-trained network AlexNet applying on the similar dataset. Our proposed model achieved an accuracy of 0.91 whereas the transfer learning model performing much better with an accuracy of 0.94.

Corresponding Author:

Huda Dhari Satea
Department of Computer Technologies Engineering, AL-Esraa University College
Al-Andalus Square, Baghdad, Iraq
Email: Huda@esraa.edu.iq

1. INTRODUCTION

Respiratory diseases indicate severe medical problems [1]. They cause death for more than three million people annually according to the world health organization (WHO) [2]. Recently, with coronavirus disease 19 (COVID-19) spreading the situation has become extremely serious. Thus, early detection of infected people is very vital in limiting the spread of respiratory diseases and COVID-19 [3]. The fundamental methods employed for diagnosing COVID-19 and respiratory diseases are computerized tomography (CT), chest X-rays, pulmonary function testing [1], [4]. However, these methods are high-priced and suffer from other issues such as radiation from the X-rays method. Alternately, auscultation pulmonary method, which is easy, fast, and much less expensive, was presented about 200 years ago by the French physician Laénnec [5]. He disclosed the relationship between respiratory disease detection and lung sounds.

Generally, lung sounds can be clustered as "normal" or "abnormal" whereas; normal lung sounds indicate that no disease exists. While abnormal lung sounds indicate that the disease is present [2]. However, the auscultation pulmonary method depends on the hearing ability and experience of the physician. It may lead to a misdiagnosis if performed by an untrained physician [6], [7]. Therefore, several computerized methods have been developed to support the auscultation procedure.

Li *et al.* [8] proposed a new classification method to distinguish between irregular and normal lung sounds, they have used active noise-canceling (ANC) for lung sounds enhancement, hidden Markov model (HMM) to characterize the lung sounds as irregular or normal, and then they have used deep neural networks (DNN) for the posterior probability estimation of HMM for every observation. Vaityshyn *et al.* [9] suggested a convolution neural network (CNN) model for classification diseases of bronchopulmonary by the lung sounds spectrogram. Serbes *et al.* [10] proposed a method for crackle and non-crackle classification using a dataset consisting of 6,000 audio files, timescale (TS) and time-frequency (TF) have been used as feature extraction methods whereas; k-nearest neighbors (KNN), support vector machines (SVM), and multi-layer sensor methods used in the classification stage. Kochetov *et al.* [11] proposed a CNN model for wheeze recognition in the lung sounds by using a dataset consisting of 817 spectrogram images with 232 normal and 585 sick lung sounds. Tariq *et al*. [12] proposed the lung disease classification (LDC) model which is based on deep learning by combined normalization and augmentation techniques as a preprocessing for effective classification of lung sounds. Güler *et al.* [13] proposed a two-stage classification model for respiratory sound patterns.

In this paper, we initially i) developed a deep CNN model for classifying lung breath sounds. Moreover, ii) we employed transfer learning using a pre-trained network and the similar dataset. Then we iii) compared between both models. Lastly, iv) we compared the accuracy of our both experiments' results. The rest of the paper coordinates: "Dataset" presents all details of the dataset used, "methodology" presents the strategy of developing the proposed model with its architecture and transfer learning approach and all implementation details as well. "Evaluation" presents and discusses both the proposed model and transfer learning approach results and performance. "Conclusion" concludes the paper.

2. METHODOLOGY

2.1. Dataset

The dataset used in this paper is from COVID-19+ pulmonary abnormalities on Bhatia [14]. This dataset is a combination of generated and real sounds spectrogram images for human breathing as shown in Table 1. It contains 6 classes: crackle course, crackle fine, COVID-19, non-COVID-19, wheezes, and normal. Each class of them is with a various number of images and various sizes. However, to obtain a reasonable classification between these classes, we considered a subset of this dataset, thus each class was arranged within 322 images, then we resized all the images in this dataset into 227×227 pixels as a pre-processing step. Lastly, we randomly split the dataset into (70%) 225 images for training and (30%) 97 images for testing. Figure 1 shows Dataset splitting, samples, and class types.

Table 1. Dataset generated and real sounds of human breathing

Type	Class
Generated	Coarse crackles
Generated	Fine crackles
Generated	Normal
Generated	Wheezes
Real	COVID-19
Real	Non COVID-19

2.2. The proposed model

The proposed model architecture is illustrated in Figure 2. The essential aim of the proposed model is classifying the spectrogram images of the generated and real sounds of human breathing into 6 classes: crackle coarse, crackle fine, COVID-19, non-COVID-19, wheezes, and normal. It is a trainable multi-class classification model implemented by utilizing the deep learning algorithm CNN. It is a sequential model with four key blocks. Each block plays a role in classification tasks and extracting features from the input. The four blocks are trained as a single network. The first block consisted of a 2D convolutional layer with a kernel size of [7×7] and 4 filters; the convolutional layer is the fundamental layer of deep learning networks. Then, convolution layer is followed by batch normalization layer, which has the effect of soothing the learning process and vividly decreasing the amount of training epochs needed to train deep networks. Afterward, a rectified linear unit (ReLU) used as an activation nonlinear function to learn mapping between inputs and response classes. Moreover, to overcome the overfitting and reducing the convolved features size a max-pooling layer with a kernel size of [7×7] and a stride of 1 was used. This block is mainly designed to acquire better features from the input images. The three following blocks are similar to the first block except the receptive field size of convolution layers which are altered between the blocks as enlisted in Table 2. Finally, a fully connected layer and SoftMax classifier were used for classifying the features extracted from the previous blocks.

2.3. Transfer learning approach

The performance of deep CNN models depends on the amount of training data [15]. The larger dataset means more accurate results [16]. However, lack of training data is a common issue within deep learning. This typically arises as a result of the difficulties in gathering large datasets [17]. Currently, the transfer learning technique is employed to overcome the lack of dataset issues [18].

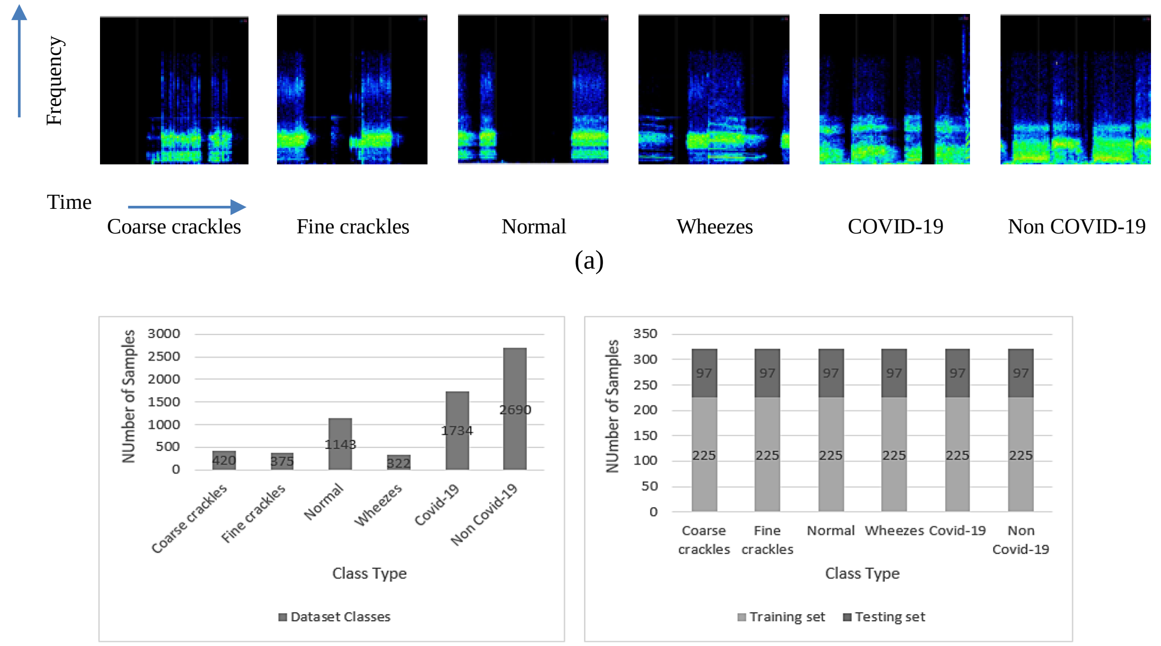


Figure 1. Splitting of dataset samples and class types (a) samples of COVID-19+pulmonary abnormalities dataset and (b) dataset splitting into training and testing

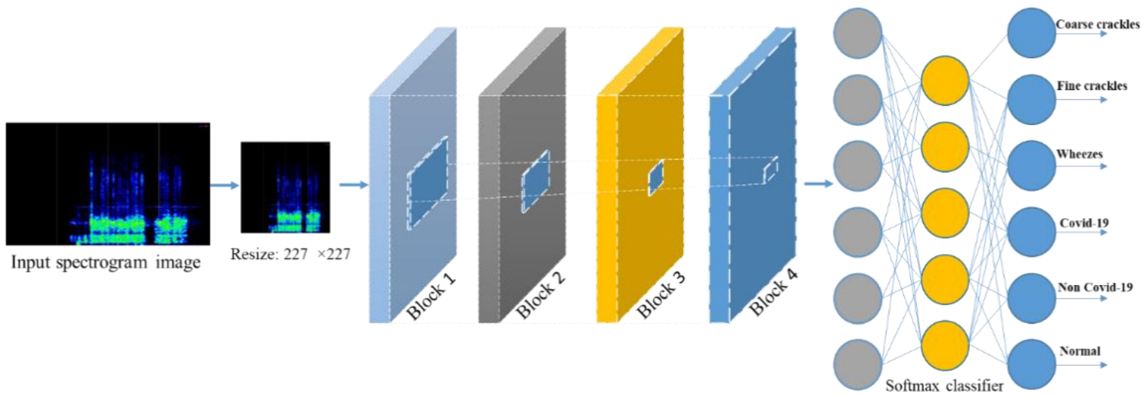


Figure 2. The proposed model architecture

Table 2. Dataset generated and real sounds of human breathing

Block	Layers
Block 1	Convolution layer(4,7×7) Batch Normalization Layer RELU Max pooling layer(7×7)
Block 2	Convolution layer (8,5×5) Batch Normalization Layer RELU Max pooling (7×7)
Block 3	Convolution layer (16,5×5) Batch Normalization Layer RELU Max pooling layer (7×7)
Block 4	Convolution layer (32,7×7) Batch Normalization Layer RELU Max pooling layer (7×7)
Classification block	Fully Connected layer SoftMax classifier

Transfer learning is a mechanism commonly used in DL. Which attempts to improve the traditional deep learning models by transferring knowledge from one source domain to another domain (target domain) [19]. CNNs are typically trained on larger datasets, and then fine-tuned for use on a smaller dataset [20].

In this paper, we have used the pre-trained network AlexNet [21]. The transfer learning model architecture is illustrated in Figure 3. AlexNet network has been trained on over a million of different images. We then transferred its knowledge to classify COVID-19+pulmonary abnormalities dataset into six classes by replacing the latest three layers as illustrated in algorithm 1.

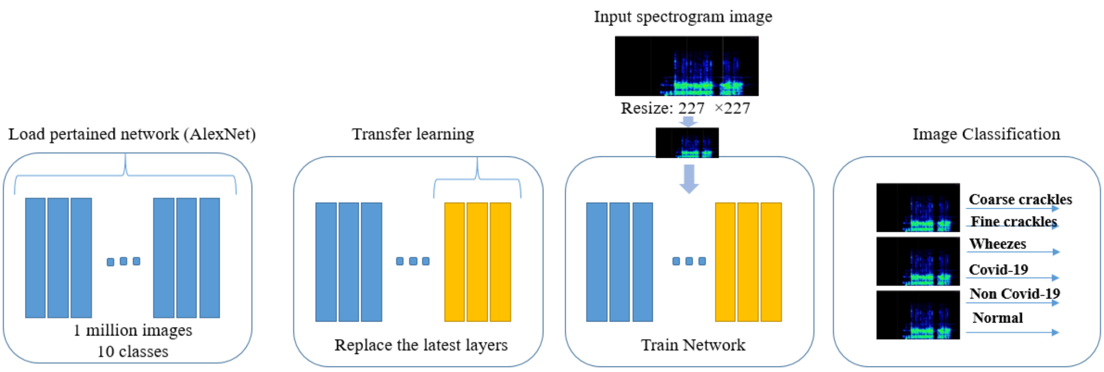


Figure 3. Transfer learning model architecture

Algorithm 1: Transfer learning approach
Input: COVID-19+Pulmonary abnormalities dataset+AlexNet network
Output: Re-fined trained network.
Begin
Step1: TR \gets Load training dataset.
L \gets Length (TR).
Step2: Resize (TR) //[227×227] pixels.
Step3: Replace the latest layers of AlexNet network
Step4: Train the network.
For i\gets L do
Classify (TR)//classifying training dataset
Step5: Evaluate the tainted network.
End.

2.4. Training

Both experiments were achieved by using similar training options. Whereas, stochastic gradient descent (SGD) has been employed as an optimizer with a momentum of 0.9, the initial learning rate was set to 0.001 and it remains constant during the training process, and the mini batch was 100. The most classical cost function used was the loss function which decreases the error between actual and output labels. In the first experiment, the training process was stabilized in 250 epochs while the number of epochs increased to more than 1,000 in the second experiment. Finally, all coding was implemented on MATLAB (R) 2020a with a toolbox of deep learning installed on a personal computer (PC) running with Intel Core (i7) inside, 16 GB RAM, CPU of 1.99 GHz, and Nvidia GPU GeForce MX130.

3. EVALUATION

In this section, we explain the results achieved by our proposed model and the model used as fine-tuning of transfer-learning for classification of the spectrogram images into 6 classes: crackle coarse, crackle fine, COVID-19, non-COVID-19, wheezes, and normal. As mentioned before, we divided the dataset into two parts 70% for training and 30% for testing. Evaluation of the performance of the models is accomplished against testing set using some metrics such as accuracy, sensitivity, specificity, precision, and F-score. Formulas of these metrics are given in equations (1-5) respectively [22], [23]. Table 3 exhibits the results achieved by both the proposed model and the transfer learning approach.

$$\text{Accuracy} = (\text{TP} + \text{TN})/(\text{TP} + \text{FP} + \text{TN} + \text{FN}) \quad (1)$$

$$\text{Sensitivity} = \text{TP}/(\text{TP} + \text{FN}) \quad (2)$$

$$\text{Specificity} = \text{TN}/(\text{TP} + \text{FN}) \quad (3)$$

$$\text{Precision} = \text{TP}/(\text{TP} + \text{FP}) \quad (4)$$

$$\text{F1 score} = 2 \times (\text{Precision} \times \text{Recall})/(\text{Precision} + \text{Recall}) \quad (5)$$

Where true positive (TP) is the total number of spectrogram images correctly classified, true negative (TN) is the total number of spectrogram images that do not fit another class and also have not been classified as that class. False positive (FP) is the total number of spectrogram images that were wrongly classified, and false negative (FN) is the total number of spectrogram images that have been perceived as an incorrect class [24], Additionally, we used confusion matrices shown in Figures 4(a) and 4(b) indicating all details about correct or wrong classification for both experiments.

Table 3. The measurement results were achieved by both the proposed model and the transfer learning approach

		Our proposed model		
Class	Sensitivity	Specificity	Precision	F-score
Crackle Coarse	1	1	1	1
Crackle Fine	1	1	1	1
Wheezes	1	1	1	1
COVID-19	1	1	1	1
Non-COVID-19	0.82	0.93	0.64	0.72
Normal	0.71	0.97	0.86	0.78
		Transfer learning approach		
Crackle Coarse	1	1	1	1
Crackle Fine	1	1	1	1
Wheezes	1	1	1	1
COVID-19	1	1	1	1
Non-COVID-19	0.81	0.96	0.82	0.82
Normal	0.82	0.96	0.81	0.81

	Crackle Coarse	Crackle Fine	Wheezes	Normal	COVID-19	Non-COVID-19
Crackle Coarse	97	0	0	0	0	0
Crackle Fine	0	97	0	0	0	0
Wheezes	0	0	97	0	0	0
Normal	0	0	0	79	0	17
COVID-19	0	0	0	0	97	0
Non-COVID-19	0	0	0	18	0	80

(a)

	Crackle Coarse	Crackle Fine	Wheezes	Normal	COVID-19	Non-COVID-19
Crackle Coarse	97	0	0	0	0	0
Crackle Fine	0	97	0	0	0	0
Wheezes	0	0	97	0	0	0
Normal	0	0	0	84	0	34
COVID-19	0	0	0	0	97	0
Non-COVID-19	0	0	0	13	0	63

(b)

Figure 4. Confusion matrices (a) the proposed model and (b) the transfer learning approach

As cited in Table 4, the comparisons are drawn among our both models, the proposed model and the transfer learning model, with method of [25]. Comparisons examined the performance according to accuracy. This comparison showing that the transfer learning model performance is effective and more accurate on lung breath sounds classification process.

Table 4. Comparative results of the accuracy

Method	Accuracy
Unais Sait *et al.* [25]	80%
Our proposed model	91%
Our transfer learning approach	**94%**

4. CONCLUSION

In this paper, we have examined two different models using convolution neural networks. Firstly, we proposed and build a CNN model from scratch for classifying lung breath sounds into six classes: crackle course, crackle fine, COVID-19, non-COVID-19, wheezes, and normal utilizing COVID-19+pulmonary abnormalities dataset. Secondly, we employed transfer learning using the pre-trained network (AlexNet) applying on the similar dataset, which in turn divided into two parts 70% for training and 30% for testing. Next, we have used several measurement criteria for evaluating the performance of both models such as accuracy, sensitivity, specificity, precision, and F-score. After that, we have compared between both models' results. Our proposed model achieved an accuracy of 0.91, whereas the transfer learning model performing much better with an accuracy of 0.94. Which means that the transfer learning model is effective and more accurate on lung breath sounds classification. Finally, we plan to improve the accuracy performance by using different pre-trained networks. Also, we plan to use two or more different datasets in one experiment to increase the challenge of classification task.

REFERENCES

[1] İ. Güler, H. Polat, and U. Ergün, "Combining neural network and genetic algorithm for prediction of lung sounds," *Journal of Medical Systems*, vol. 29, no. 3, pp. 217–231, Jun. 2005, doi: 10.1007/s10916-005-5182-9.

[2] F. Demir, A. M. Ismael, and A. Sengur, "Classification of lung sounds with CNN model using parallel pooling structure," *IEEE Access*, vol. 8, pp. 105376–105383, 2020, doi: 10.1109/ACCESS.2020.3000111.

[3] S. Gairola, F. Tom, N. Kwatra, and M. Jain, "RespireNet: a deep neural network for accurately detecting abnormal lung sounds in limited data setting," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Nov. 2021, pp. 527–530, doi: 10.1109/EMBC46164.2021.9630091.

[4] 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.

[5] A. Hashemi, H. Arabalibeik, and K. Agin, *Classification of wheeze sounds using cepstral analysis and neural networks*. IOS Press, 2012.

[6] A. H. Falah and J. Jondri, "Lung sounds classification using stacked autoencoder and support vector machine," in *2019 7th International Conference on Information and Communication Technology (ICoICT)*, Jul. 2019, pp. 1–5, doi: 10.1109/ICoICT.2019.8835278.

[7] D. Bardou, K. Zhang, and S. M. Ahmad, "Lung sounds classification using convolutional neural networks," *Artificial Intelligence in Medicine*, vol. 88, pp. 58–69, Jun. 2018, doi: 10.1016/j.artmed.2018.04.008.

[8] L. Li, W. Xu, Q. Hong, F. Tong, and J. Wu, "Classification between normal and adventitious lung sounds using deep neural network," in *2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP)*, Oct. 2016, pp. 1–5, doi: 10.1109/ISCSLP.2016.7918407.

[9] V. Vaityshyn, M. Chekhovych, and A. Poreva, "Convolutional neural networks for the classification of bronchopulmonary system diseases with the use of lung sounds," in *2018 IEEE 38th International Conference on Electronics and Nanotechnology (ELNANO)*, Apr. 2018, pp. 383–386, doi: 10.1109/ELNANO.2018.8477483.

[10] G. Serbes, C. O. Sakar, Y. P. Kahya, and N. Aydin, "Pulmonary crackle detection using time-frequency and time-scale analysis," *Digital Signal Processing*, vol. 23, no. 3, pp. 1012–1021, May 2013, doi: 10.1016/j.dsp.2012.12.009.

[11] K. Kochetov, E. Putin, S. Azizov, I. Skorobogatov, and A. Filchenkov, "Wheeze detection using convolutional neural networks," in *Progress in Artificial Intelligence*, Springer International Publishing, 2017, pp. 162–173.

[12] Z. Tariq, S. K. Shah, and Y. Lee, "Lung disease classification using deep convolutional neural network," in *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, Nov. 2019, pp. 732–735, doi: 10.1109/BIBM47256.2019.8983071.

[13] E. Ç. Güler, B. Sankur, Y. P. Kahya, and S. Raudys, "Two-stage classification of respiratory sound patterns," *Computers in Biology and Medicine*, vol. 35, no. 1, pp. 67–83, Jan. 2005, doi: 10.1016/j.compbiomed.2003.11.001.

[14] R. Bhatia, "Spectrogram images of breathing sounds for COVID-19 and other pulmonary abnormalities," *Mendeley Data, V1*. 2021, doi: 10.17632/pr7bgzxpgv.1.

[15] T. Tian, Institute of Electrical and Electronics Engineers, IEEE Computer Society, National Science Foundation (U.S.), and Haerbin gong ye da xue, *Proceedings, 2016 IEEE International Conference on Bioinformatics and Biomedicine : Dec 15-18, 2016, Shenzhen, China*. IEEE.

[16] S. D. Veloz, "Spatially autocorrelated sampling falsely inflates measures of accuracy for presence-only niche models," *Journal of Biogeography*, vol. 36, no. 12, pp. 2290–2299, Dec. 2009, doi: 10.1111/j.1365-2699.2009.02174.x.

[17] L. Alzubaidi, O. Al-Shamma, M. A. Fadhel, L. Farhan, J. Zhang, and Y. Duan, "Optimizing the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model," *Electronics*, vol. 9, no. 3, Mar. 2020, doi: 10.3390/electronics9030445.

[18] L. Perez and J. Wang, "The effectiveness of data augmentation in image classification using deep learning," *Computer Vision and Pattern Recognition*, Dec. 2017.

[19] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, and G. Zhang, "Transfer learning using computational intelligence: A survey," *Knowledge-Based Systems*, vol. 80, no. January, pp. 14–23, 2015, doi: 10.1016/j.knosys.2015.01.010.

[20] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.

[21] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.

[22] C. Li, H. Du, and B. Zhu, "Classification of lung sounds using CNN-Attention." EasyChair Preprint no. 4356, 2020.

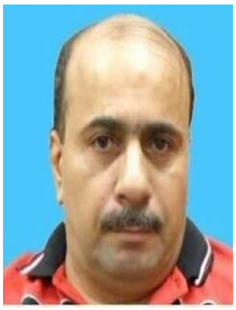
[23] A. Saber, M. Sakr, O. M. Abo-Seida, A. Keshk, and H. Chen, "A novel deep-learning model for automatic detection and classification of breast cancer using the transfer-learning technique," *IEEE Access*, vol. 9, pp. 71194–71209, 2021, doi: 10.1109/ACCESS.2021.3079204.

[24] V. Maeda-Gutiérrez *et al.*, "Comparison of convolutional neural network architectures for classification of tomato plant diseases,"

Applied Sciences, vol. 10, no. 4, Feb. 2020, doi: 10.3390/app10041245.
[25] U. Sait *et al.*, "A deep-learning based multimodal system for Covid-19 diagnosis using breathing sounds and chest X-ray images," *Applied Soft Computing*, vol. 109, Sep. 2021, doi: 10.1016/j.asoc.2021.107522.

BIOGRAPHIES OF AUTHORS

Huda Dhari Satea was born in Baghdad, Iraq in 1992. She received B.Sc. degree in Network engineering from Al-Nahrieen University, Baghdad, in 2014. And Higher Diploma in Information Technology/Website Technology from informatics Institute for Postgraduate Studies (IIPS), Baghdad, Iraq 2016. And M.Sc. degree in Computer sciences from informatics Institute for Postgraduate Studies (IIPS), Baghdad, Iraq in 2019. Her research interests in image processing and deep learning. She is working as Assistant Lecturer at AL-Esraa University, Baghdad since 2014. She can be contacted at email: huda@esraa.edu.iq.



Amer Saleem Elameer is a Consultant Engineer, from the Iraqi Commission for Computers and Informatics (ICCI), in the Informatics Institute for Postgraduate Studies (IIPS) in Baghdad, Iraq. He obtained his PhD degree in Internet Portals and specialized in e-learning portals from the Universiti Sains Malaysia (USM), Penang, Malaysia, and introduced the Elameer-Idrus Orbital IT e-learning ramework as a major outcome from his postgraduate research; the framework now has taken the shape of an e-education and MOOCs framework for Iraqi education and higher education. Also, he is the designer of the Numerical Iraqi Social Safety network (ISSN). The founder of the first Iraqi MOOC website (MOOC-IIPS) and the designer of the first Iraqi e-learning framework for the Blended learning (BL) and lifelong learning. The designer and the programmer of the first Iraqi learning management system (LMS). He can be contacted at email: dr.amerelameer@gmail.com

Ahmed Hussein Salman was born in 1989 in Baghdad. He was graduated from Informatics Institute for Postgraduate Studies (IIPS), M.Sc. degree in Computer Sciences, Baghdad, Iraq 2019, and B.Sc. degree in Computer sciences, in 2014 from Dijlah Uni. His interested in theoretical fields and analysis of; image processing, data compression, IoT, pattern recognition, data mining, security, and multimedia systems. He is working as Assistant Lecturer at AL-Esraa University, Baghdad since 2014. He can be contacted at email: ahmed@esraa.edu.iq.

Shahad Dhari Sateea was born in Baghdad, Iraq in April of 1988. Received her B.Sc. degree in Electrical and Electronics Engineering from the University of Technology, Iraq in 2010, and the M.Sc. degree in Communication Engineering from the Department of Electrical Engineering, University of Technology, Iraq in 2013. Currently, she is Instructor at the computer Techniques Engineering department, Al-Esraa University College, Baghdad, Iraq. She can be contacted at email: shahad@esraa.edu.iq.