JCSSE 2022
International Joint Conference on Computer Science and Software Engineering

78th Anniversary Silpakorn University

Digital Transformation in the Post Pandemic Era

2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE 2022)

22 – 25 June, 2022 at Silpakorn University, Wang Tha Phra, Bangkok, Thailand

Organized by
Department of Computing
Faculty of Science,
Silpakorn University, Thailand

ISBN: 978-1-6654-8510-4
Online ISSN: 2642-6579

IEEE
IEEE THAILAND SECTION
IEEE THAILAND SECTION 40TH ANNIVERSARY SINCE 1981
IEEE COMPUTER SOCIETY Thailand Chapter
IEEE SMC Systems, Man, and Cybernetics Society THAILAND CHAPTER

2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE) | 978-1-6654-8510-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/JCSSE54890.2022.9836296

COVID-19 and Respiratory Diseases Classification using Deep Convolution Neuron Network

1st Orawan Chunhapran, 2nd Suteera Vongananusup, 3rd Tongjai Yampaka, 4th Rath Burirat
Department of Computer Science
Faculty of Business Administration and Information Technology
Rajamangala University of Technology Tawan-Ok
Vibhavadi Rangsit Road Dindeang Sub-District, Dindeang District Bangkok, Thailand, 10400
orawan_ch@rmutto.ac.th, suteera_vo@rmutto.ac.th, tongjai_ya@rmutto.ac.th, rath_bu@rmutto.ac.th

Abstract— **This study proposes COVID-19 and Respiratory Diseases Classification using Deep Convolution Neuron Network. ICBHI 2017 Respiratory Sound Database including COVID-19 from Coswara databased were used in our experiments. The potential results show that the left side model performances are 0.85 accuracy, 0.76 sensitivity, and 0.90 specificity. The right side model performances are 0.86 accuracy, 0.76 sensitivity, and 0.93 specificity. No side set model performances are 0.83 accuracy, 0.71 sensitivity, and 0.93 specificity. In addition, the lung characteristics and lung functions are different among left and right. Therefore, the breathing sound from left and right lung are difference. For this reason, the cross-model performances were evaluated to test this assumption. The cross-model performance results show that the left data is consistent with the left model. As same as the right data is consistent with the right model. Furthermore, the experiment found that mixing training data built the no side set model is the lowest performance. In addition, the proposed framework tends to achieve high performance when compared with a recent study.**

Keywords-component; COVID-19 classification; deep convolution neuron networks; breathing sound classification

I. INTRODUCTION

In the lung pandemic scenario, manual testing for medical diagnosis is not safe for the physician who can be infected with the virus spread from the patients [1]. Li F., et al. and Vinod D.N., et al. proposed self-diagnosis for telemedicine with preliminary symptom assessment as an alternative to manual testing [1], [2]. Ahmad Alhiyari M., et al. explain that breathing analysis utilizes respiratory abnormalities such as crackles (coarse and fine) and wheezes of patients were the basis for viral detection over other breathing sound abnormalities [3]. During the pandemic, people infected with COVID-19 required diagnostic solutions and new portable testing mechanisms [4], [5]. Therefore, many studies have urgently developed effective tools for respiratory diagnosis not only of general lung disease but also COVID-19. The sound-based medical diagnosis in a few studies [6-8] have utilized cough sounds for Covid-19 diagnosis. In the previous study, the respiratory sound has not separated the model for analysis from two sides of the lung. When the lung characteristics and lung functions are different among left and right, the digital stethoscopes should be separately considered. Chan H.P., et al. [9] explain that Deep Convolution Neuron Network models for medical diagnosis have been widely introduced in the context of the opportunities and challenges by Covid-19. The CNN is a higher classification performance [10], and utilizes a smaller dataset when using transfer learning such as VGGNet, InceptionNet, or ResNet [11]. Inspired by their success, this study proposes respiratory sound classification using Deep Convolution Neuron Network (DCNN) to distinguish spectrogram images from the respiratory sound in Bronchiectasis, Bronchiolitis, COPD, LRTI, Pneumonia, URTI, COVID-19, and Healthy.

II. MATERIALS

A. Database

The respiratory sounds from the ICBHI 2017 Respiratory Sound Database were used in this present study [12]. The database consists of 920 annotated audio samples from 126 subjects. The diagnosis for each subject composes of Bronchiectasis, Bronchiolitis, COPD, LRTI, Pneumonia, URTI, and Healthy. In addition, the COVID-19-infected breath dataset from Coswara [13] were included in our experiment. When the lung characteristics and lung functions are different among left and right, the digital stethoscopes should be separately considered. Therefore, the database and model were divided into two sides. When considering the data is imbalanced, the class weight for each category per class was calculated to solve the skewed distribution of the classes by giving different weights to both the majority and minority classes. The difference in weights will influence the classification of the classes during the training phase. The majority class is penalized by reducing weight, while the minority class is setting a higher weight than the majority class. The detail of the database as shown in Table I.

978-1-6654-3831-5/21/$31.00 ©2021 IEEE

2022 19th International Joint Conference on Computer Science and Software Engineering (JCSSE) | 978-1-6654-8510-4/22/$31.00 ©2022 IEEE | DOI: 10.1109/JCSSE54890.2022.9836259

TABLE I. THE NUMBER OF IMAGE AND THE RATIO OF IMAGE

Class	Left side (#image)	Left class weight	Right side (#image)	Right class weight
Bronchiectasis	114	3.84	117	3.3
Bronchiolitis	105	4.1	126	3.0
COPD	205	2.1	186	2.0
LRTI	96	4.6	92	4.2
Pneumonia	95	4.6	72	5.4
URTI	127	3.4	90	4.2
COVID-19	438	1.0	386	1
Healthy	163	2.6	106	3.6

III. METHODOLOGY

This approach starts with breathing sound recording from digital stethoscope connected via Bluetooth on various regions of the chest called zones. When deep learning algorithms are employed to perform image processing, the data preparation using python package called librosa and soundfile [14] to covert the breathing sound to 2D-colour mapping of frequency in relation to time using digital signal processing. Then, the feature extraction using VGG19 pre-training and top layer fine tuning follow by classification layer. The proposed methodology was shown in fig 1. The detailed methodology was explained in the next section.

A. Data pre-processing

1) Breathing sound Breathing sounds were obtained using two digital stethoscopes with the Bluetooth module. When the lung characteristics and lung functions are different among left and right, the digital stethoscopes were separately collected the breathing sound. The chest zone is encoded as three ordered letters from the sets {A, P}, {L, R}, and {L, M, U} respectively. The letters have the following meanings: Anterior: A, Posterior: P, Left: L, Right: R, Lower: Lo, Upper: U, and Middle: M. Therefore, the database and model were divided into two sides. In addition, the no side set (mixed data both left and right) was included to compare with the single left side and the single right side.

2) The Fast Fourier transform (FFT) The recording sounds were transferred from the digital stethoscope to the FFT analyzer based on librosa and soundfile with python library. The breathing sound signal was transformed from the time-domain to the frequency domain using the Fourier transform. Using digital signal processing, the output is a 2D color matching of frequency in proportion to time.

3) Breathing sound spectrograms When deep learning algorithms are employed to perform image processing, the breathing sound spectrogram is transformed to a 2D vector image. All the spectrogram photos were resized in 224×224 standard pixels for transfer learning with VGG19 to extract the spectrogram image features.

B. Feature Extraction

1) Pre-training using Convolution Neuron Network Feature selection and extraction are involved selecting and extracting the relevant features from the input data. Especially, the high dimension data from image pixels requires reducing the number of a large set of data. Deep Convolutional Neural Network (DCNN) is the most widely used for extracting the feature from the image. However, training a deep CNN often requires computational resources. To address these challenges, transfer learning such as VGGNet, ResNet, or InceptionNet are introduced. Motivated by applicability of transfer learning of DCNNs, the VGG19 in the Keras library was used to extract the spectrogram images. Online Google Collaboratory with high performance GPU was used in this study. Fig 1 shows the VGG19 architecture. The architecture of VGG19 is shown in Fig 1; the default input image size of VGG19 is 224×224. It composes of 16 convolutional layers and 5 pooling layers for reducing a huge feature map followed by 3 fully connected layers.

C. Classification

DCNN is not only feature extraction but also represents success in image recognition and classification. From the feature selection layer using VGG19, softmax function was used as a classifier to classify the input image based on the training data. (see Fig 1). The softmax function takes a vector with probabilities of each possible outcome.

D. Performance Evaluation

Confusion matrix was used to evaluate the class precision. This matrix composes of sensitivity (true positive rate), specificity (true negative rate) and accuracy of models. The predictive formulas were defined as:

$$Sensitivity = \frac{TP}{TP + FN} \quad (1)$$

$$Specificity = \frac{TN}{TN + FP} \quad (2)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (3)$$

IV. RESULTS AND DISCUSSION

A. Data pre-processing result

Breathing sound spectrograms were converted from signal to image to distinguish patterns across different Bronchiectasis, Bronchiolitis, COPD, LRTI, Pneumonia, URTI, COVID-19, and Healthy using FFT Analyzer based on librosa and soundfile with python library. Fig 2 shows the sample of spectrogram image.

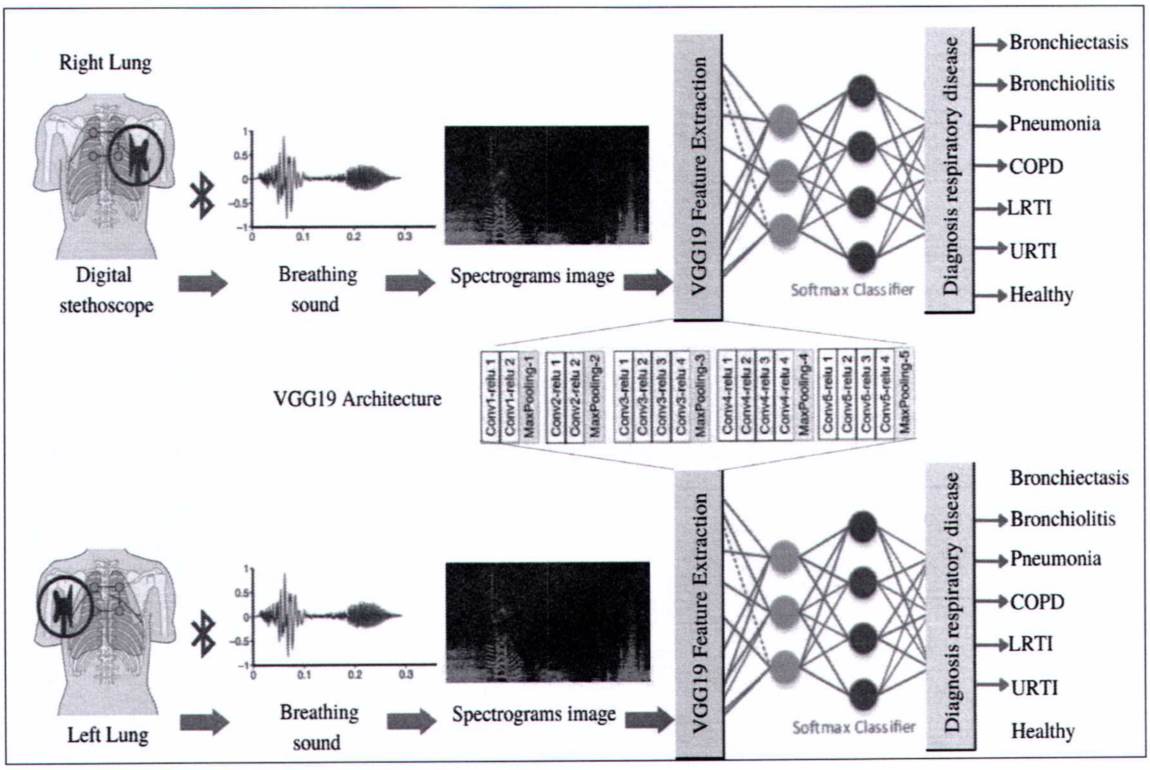


Fig 1 the proposed methodology consists of the following five phases: data acquisition, data preparation, feature extraction, training, and classification.

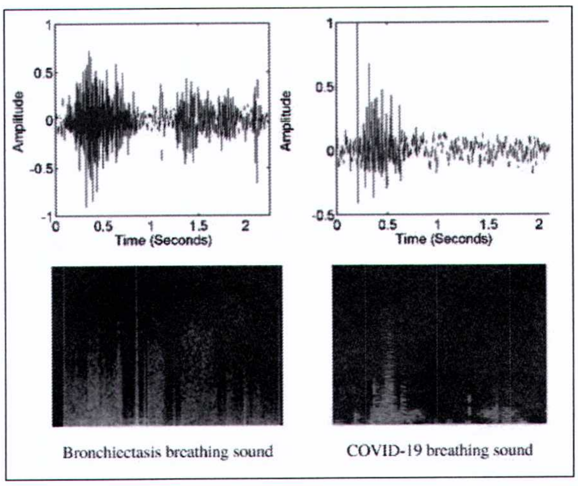


Fig 2 the sample of spectrogram image converting from sound wave to spectrogram image using librosa and soundfile with python library.

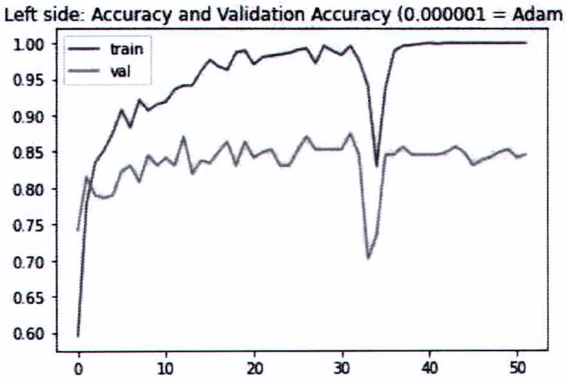
B. *Feature Extraction and Classification*

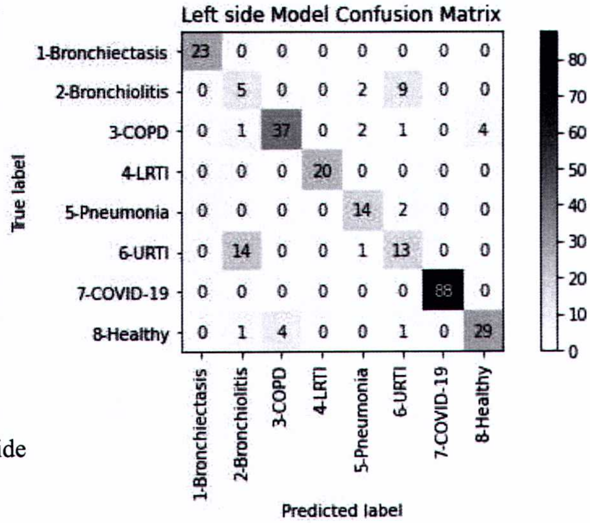
1) Pre-training using Convolution Neuron Network

The architecture of VGG19 was used to extract the features from spectrogram images followed by classification layer. There are two steps for extracting the image feature. First, the input image is fed to VGGNet backbone without fully training. Then, the output features from VGGNet layer are trained using CNN top-layer.

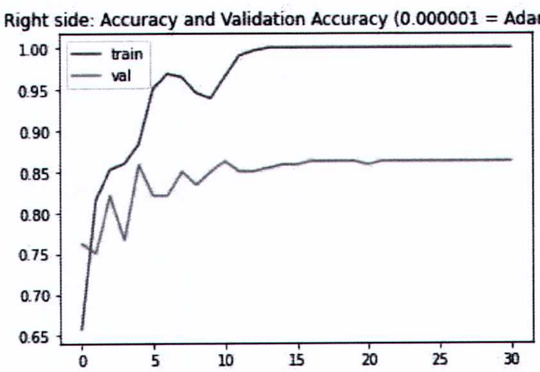
2) Classification

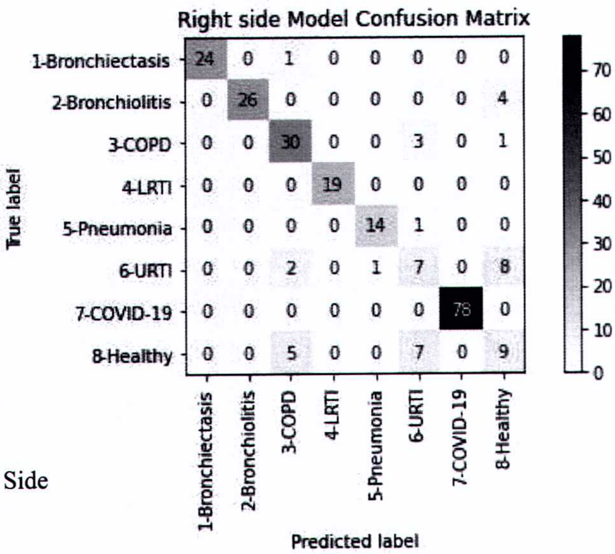
The automated classification framework and implementation using the DCNN model were held in a Python 3.7.13 environment using Keras, and TensorFlow on Google Colab Pro with Nvidia K80/T4. The model was performed using two databases from ICBHI 2017 and Coswara. These datasets were divided into 80% for training and 20% for testing. In all three splitting sides i.e., Left side, Right side, and no side set are presented efficient experiment results in the training report in Fig 3 model performance is presented in Fig 3. From the experiments, the left side model performances are 0.85 accuracy, 0.76 sensitivity, and 0.90 specificity. The right side model performances are 0.86 accuracy, 0.76 sensitivity, and 0.93 specificity. No side set model performances are 0.83 accuracy, 0.71 sensitivity, and 0.93 specificity. Figure 3(c) shows unstable in long run. It seems to train using the large network with small batch side. However, the model performances were tested with testing data. Although these results show that all models are no significant differences in classification performance. Nevertheless, the main difference between right lung and left lung is that the anterior border of the left lung consists of a deep cardiac notch whereas right lung is straight. It means that the exchange of respiratory gas between left and right is difference. Therefore, the breathing sound from left and right lung are difference. For this reason, the cross-model performances were evaluated to test this assumption. The cross-model performances are shown in Table II.



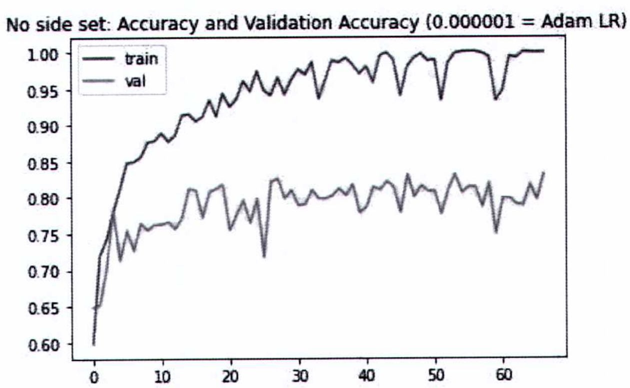


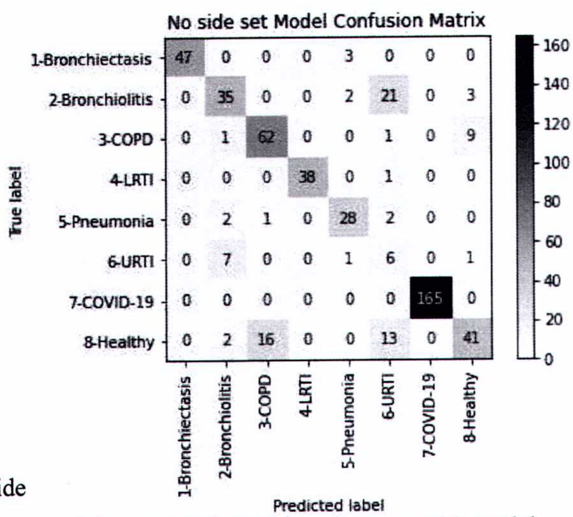
(a) Left Side





(b) Right Side





(c) No Set Side

Fig 3 the training and testing report and overall confusion matrix of the proposed model performance (a) model performance from the left side (b) model performance from the right side (c) model performance from the no side set

TABLE II. COMPARISON BETWEEN THE PERFORMANCE OF THE CROSS-MODEL PERFORMANCES

Testing data/Model	Left model	Right model	No side set model
Left	**0.85**	0.76	0.68
Right	0.78	**0.86**	0.65
No side set	0.68	0.65	**0.83**

TABLE III. COMPARISON THE PROPOSED FRAMEWORK AND OTHER METHODS

Study	Sensitivity	Specificity	Accuracy
Liu et al. (2019) [15]	-	-	0.81
Minami et al. (2019) [16]	0.28	0.81	-
Saraiva et al. (2020) [17]	-	-	0.74
Gairola et al. (2021) [18]	0.40	0.72	-
Pham et al. (2021) [19]	0.26	0.68	-
Petmezas et al. (2021) [20]	0.52	0.84	0.76
Proposed method	**0.74**	**0.91**	**0.85**

Finally, a comparison in average performance between the performance of the proposed framework and other methods in recent relevant studies of lung sounds is presented in Table III.

The proposed framework tends to achieve high performance when compared with a recent study. A highly sensitive test means that there are few false negative results, and thus fewer cases of the disease are missed. A highly specific test means that there are few false positive results. It may not be feasible to use a test with low specificity for screening, since many people without the disease will screen positive, and potentially receive unnecessary diagnostic procedures.

V. CONCLUSIONS

This study proposes respiratory sound classification using Deep Convolution Neuron Network (DCNN) to distinguish spectrogram images from the respiratory sound in Bronchiectasis, Bronchiolitis, COPD, LRTI, Pneumonia, URTI, COVID-19, and Healthy. ICBHI 2017 Respiratory Sound Database including COVID-19 from Coswara databased were used in our experiments. The potential results show that single left model, single right model, and no side set model are no significant differences in classification performance. However, the lung characteristics and lung functions are different among left and right. It means that the exchange of respiratory gas between left and right is difference. Therefore, the classification model should be separately considered. The cross-model performance results show that the left data is consistent with the left model. As same as the right data is consistent with the right model. Furthermore, the experiment found that mixing training data built the no side set model is the lowest performance. In addition, the proposed framework tends to achieve high performance when compared with a recent study.

On the other hand, the limitations of the present study could be summarized in three parts. First, the other location such as the Upper, and Lower lungs should be included to complete the lung location. Second, other pre-training model such as Inception, or ResNet should be performed to compare the model performance. Finally, more datasets should be considered for more reliable.

REFERENCES

[1] Li F., Valero M., Shahriar H., Khan R.A., Ahamed S.I. Wi-COVID: A COVID-19 symptom detection and patient monitoring framework using WiFi. Smart Health. 2021;19 doi: 10.1016/j.smhl.2020.100147.

[2] Ahmad Alhiyari M., Ata F., Islam Alghizzawi M., Bint I Bilal A., Salih Abdulhadi A., Yousaf Z. Post COVID-19 fibrosis, an emerging complicationof SARS-CoV-2 infection. IDCases. 2021;23 doi: 10.1016/j.idcr.2020.e01041.

[3] Quirino M.G., Colli C.M., Macedo L.C., Sell A.M., Visentainer J.E.L. Methods for blood group antigens detection: cost-effectiveness analysis of phenotyping and genotyping. Hematol. Transfus. Cell Ther. 2019;41:44–49. doi: 10.1016/j.htct.2018.06.006.

[4] Kilic T., Weissleder R., Lee H. Molecular and immunological diagnostic tests of COVID-19: Current status and challenges. IScience. 2020;23 doi: 10.1016/j.isci.2020.101406.

[5] Adebisi Y.A., Lucero Prisno III D.E. Towards an inclusive health agenda: People who inject drugs and the COVID-19 response in Africa. Public Health. 2020 doi: 10.1016/j.puhe.2020.10.017.

[6] Pal A., Sankarasubbu M. 2020. Pay attention to the cough: Early diagnosis of COVID-19 using interpretable symptoms embeddings with cough sound signal processing; pp. 1–15.

[7] Aygün H., Apolskis A. The quality and reliability of the mechanical stethoscopes and Laser Doppler Vibrometer (LDV) to record tracheal sounds. Appl. Acoust. 2020;161 doi: 10.1016/j.apacoust.2019.107159.

[8] Hui Huang Y., jun Meng S., Zhang Y., et al. 2020. The respiratory sound features of COVID-19 patients fill gaps between clinical data and screening methods.

[9] Chan H.P., Samala R.K., Hadjiiski L.M., Zhou C. Deep learning in medical image analysis. Adv. Exp. Med. Biol. 2020;1213:3–21. doi: 10.1007/978-3-030-33128-3_1.

[10] Dong N., Zhao L., Wu C.H., Chang J.F. Inception v3 based cervical cell classification combined with artificially extracted features. Appl. Soft Comput. 2020;93 doi: 10.1016/j.asoc.2020.106311.

[11] Chhikara P., Singh P., Gupta P., Bhatia T. Deep convolutional neural network with transfer learning for detecting pneumonia on chest x-rays. Adv. Intell. Syst. Comput. 2020;1064:155–168. doi: 10.1007/978-981-15-0339-9_13.

[12] ICBHI 2017 Challenge, Bhichallenge.med.auth.gr, 2022. [Online].Available: https://bhichallenge.med.auth.gr/ICBHI_2017_Challenge. [Accessed: 10- May- 2022].

[13] Sharma, N.; Krishnan, P.; Kumar, R.; Ramoji, S.; Chetupalli, S.R.; Ghosh, P.K.; Ganapathy, S. Coswara–A Database of Breathing, Cough, and Voice Sounds for COVID-19 Diagnosis. arXiv 2020, arXiv:2005.10548.

[14] librosa 0.9.1 documentation", Librosa.org, 2022. [Online]. Available: https://librosa.org/doc/main/index.html. [Accessed: 10- May- 2022].

[15] Liu, R.; Cai, S.; Zhang, K.; Hu, N. Detection of Adventitious Respiratory Sounds based on Convolutional Neural Network. In Proceedings of the International Conference on Intelligent Informatics and Biomedical Sciences (ICIIBMS), Shanghai, China, 21–24 November 2019.

[16] Minami, K.; Lu, H.; Kim, H.; Mabu, S.; Hirano, Y.; Kido, S. Automatic Classification of Large-Scale Respiratory Sound Dataset Based on Convolutional Neural Network. In Proceedings of the 2019 19th International Conference on Control, Automation and Systems (ICCAS), Jeju, Korea, 15–18 October 2019.

[17] Saraiva, A.; Santos, D.; Francisco, A.; Sousa, J.; Ferreira, N.; Soares, S.; Valente, A. Classification of Respiratory Sounds with Convolutional Neural Network. In Proceedings of the 13th International Joint Conference on Biomedical Engineering Systems and Technologies, Valletta, Malta, 24–26 February 2020.

[18] Gairola, S.; Tom, F.; Kwatra, N.; Jain, M. Respirenet: A deep neural network for accurately detecting abnormal lung sounds in limited data setting. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Mexico City, Mexico, 1–5 November 2021; pp. 527–530.

[19] Pham, L.D.; Phan, H.; Palaniappan, R.; Mertins, A.; McLoughlin, I. CNN-MoE Based Framework for Classification of Respiratory Anomalies and Lung Disease Detection. IEEE J. Biomed. Health Inform. 2021, 25, 2938–2947.

[20] G. Petmezas et al., "Automated Lung Sound Classification Using a Hybrid CNN-LSTM Network and Focal Loss Function", Sensors, vol. 22, no. 3, p. 1232, 2022. Available: 10.3390/s22031232.