Detecting Respiratory Diseases from Recorded Lung Sounds by 2D CNN

Reetodeep Hazra
Electronics and Communication Engineering
Techno International New Town
Kolkata, India
reetodeep.hazra.2017@ece.tint.edu.in

Dr. Sudhan Majhi
Electrical Engineering
Indian Institute of Technology Patna
Patna, India
smajhi@iitp.ac.in

***Abstract*—Respiratory disease is among the leading causes of deaths around the world. A large amount of population is being affected regularly with some kinds of lung function disorders which eventually lead to respiratory diseases. Prevention and early detection are essential steps in managing respiratory diseases. To decrease the fatality, an efficient detection model is needed. In this paper, 2D convolutional neural network (CNN) is used to detect respiratory diseases from the recorded lung sounds at early stages. The proposed method can detect respiratory diseases like bronchiectasis, pneumonia, bronchiolitis, chronic obstructive pulmonary disease, upper respiratory tract infection, and healthy by using Mel-frequency cepstral co-efficients (MFCC). In the proposed scheme, a data frame is recorded and after extracting the statistical features from the audio clips, the data is loaded in the data frame where further classification is done using 2D CNN. The model is based on 2D CNN architecture where the number of layers is reduced to a certain extent to achieve more accuracy. The proposed model has only 13 CNN layers where each convolution layer is being associated with a pooling layer of max-pooling 2D type. The final convolution layer has a global-average pooling 2D layer. The proposed method obtained an accuracy of over 92.39%.**

***Index Terms*—Convolutional neural network (CNN), Mel-frequency cepstral co-efficient (MFCC), Respiratory diseases.**

I. Introduction

Lung sounds are important indicators of respiratory health and respiratory disorders. The sounds emitted when a person breathes are directly related to the air movements and changes in the lung tissue. Lung sounds can be classified into two main categories- normal lung sounds and abnormal lung sounds. Normal lung sounds consist of bronchial sounds and vesicular sounds whereas abnormal lung sounds consists of crackles and wheezes [1]. Based on lung sounds, respiratory sound diseases can be determined [2].

Here the work is centered on abnormal lung sounds such as crackles and wheezes which leads to respiratory sound disease classification. Bronchiectasis is a respiratory sound disorder where bronchial tubes of lungs get completely damaged. In chronic obstructive pulmonary diseases (COPD), breath sounds are diminished and are prolonged. Coarse crackles are heard at the beginning of inspiration and are common in COPD. Upper respiratory tract infection (URTI) is also a respiratory tract infection which directly affect the lungs. Other respiratory sounds classified here are Pneumonia and Bronchiolitis.

Respiratory illness which is becoming the most common mortality factors worldwide includes diseases and conditions such as Pneumonia, Bronchiolitis, lung cancer, Tuberculosis, Asthma, COPD and lower respiratory tract infection (LRTI). During lung auscultation (listening to the lung sounds produced), the detection of anomaly sounds such as crackles and wheezes is a very important aspect of a medical examination which further leads to the diagnosis of the common respiratory diseases. Both sounds are classified into a group of adventitious sounds, which may indicate pulmonary disorders. If some kind of automated methods can be developed which could detect these anomaly sounds, then early detection of respiratory diseases can be made in the near future. Automated analysis of respiratory sounds has a long history [1]. The research field attracted a little attention till machine learning techniques were developed. Most machine learning techniques use Mel-frequency cepstral co-efficients (MFCC) [2] to extract features and vectors. These features are then inserted to machine learning models such as convolutional neural networks (CNN) [3], support vector machines (SVM) [4], decision trees [5] etc. In [3], lung sounds are classified with the help of CNN. In [4], 5D vector classification is used to classify lung sounds using SVM whereas in [5], Chambres *et al.* used machine learning approach with a boosted decision tree model to classify lung sounds.

Several research works are available that focus on lung sound analysis to detect the crackles and wheezes [4] - [6]. In [4], the authors used short time Fourier transform (STFT) spectrogram directly into the SVM. The final methodology used in the work is STFT as a pre-processing step to calculate the spectrum mean. In [6], Saraiva *et al.* classified lung sounds using CNN. They directly extracted the features from audio clips using MFCC without data normalization which eventually led to a less accuracy of 74%. In the previous works discussed, the classification is based on the lung sounds that are wheezes and crackles. But there is no work on classification of respiratory diseases such as bronchiectasis, pneumonia, bronchiolitis, COPD, URTI, and healthy. So in this field, only a few researches are there which are based on

978-1-7281-9180-5/20/$31.00 © 2020 IEEE

the detection of respiratory diseases [7]. In [7], Dubey *et al.* only gave a broad review of classification techniques based on neural networks for pulmonary obstructive diseases.

In the previous works, lung sounds are classified using different machine learning algorithms like CNN, SVM, etc. But respiratory diseases are not classified anywhere.

In this paper, we propose a deep learning approach for classifying respiratory diseases on the basis of different anomalies in lung sounds for early diagnosis and intervention of respiratory diseases. In this method, we introduce several preprocessing techniques such as data normalization and data augmentation for effective respiratory disease classifications. The features are extracted from the raw audio files using MFCC and the features are fed as inputs. We have used data augmentation techniques [8] which helps our 2D CNN model to get a better accuracy. The 2D CNN architecture used here reports an accuracy of over 92.39%. The dataset used in this work is obtained from the publicly available dataset [9] of ICBHI. We isolated the effects of various factors which are related to front-end feature extraction such as respiratory cycle length and time resolution to determine how they affect the detection performance.

Rest of the paper is organized as follows: Section II describes the related works which is followed by detailed explanation of the model in Section III. Section IV explains the results and discussion. Conclusion of our work is mentioned in Section V.

II. Related Works

One of the main challenges faced in this research was finding the publicly available data and consequently cleaning the data that are not recorded properly. Data corrupted by noises from heart, lungs, and other parts of the body cannot be used directly as input to the model. For overcoming this and improving the accuracy, data normalization is applied to the original audio files where features were extracted using MFCC to re-scale the audio samples. Due to the limited availability of the respiratory sounds, the research in this field is limited. We found several studies which focus on respiratory disease and respiratory sound classification with their own prediction models using machine learning algorithms. We are categorizing according to the algorithms used.

A. Predictions using CNN

Minami *et al.* [3] developed an automatic classifier of lung sounds using large-scale respiratory sound dataset based on CNN. They first transformed 1D signals into 2D time-frequency representation images using STFT and continuous wavelet transform. Then they classified the transferred images using CNN. However noise removal methods were not applied in the audio signals which would also result in the betterment of the accuracy. Saraiva *et al.* [6] classified respiratory sounds using CNN with the help of MFCC. The major drawback was that data normalization was not applied to the audio samples and the accuracy was reported to be less than 74%. Rupesh *et al.* [7] reviewed many classification and feature extraction techniques for diseases like COPD, asthma etc. In feature extraction techniques, they used fast Fourier transform (FFT) algorithm, spectrogram, wavelet transform and short time Fourier transform (STFT) algorithm. The best accuracy they reported was by CNN model. The paper is a complete overview of the works done previously. But no novel method is reported. Chen *et al.* [10] proposed a method for automatic early detection for heart and lung diseases using CNN. The accuracy found here is only on the basis of the CNN layers but no disease classification is present here. However, the dataset they used was too limited to have any consequences for results. Pham *et al.* [11] prepared a robust deep learning network for predicting respiratory anomalies and lung diseases. They provided an extensive analysis of how factors like respiratory cycle length, time resolution, and network architecture affects final prediction accuracy. Their work also considered deep learning networks like CNN-ensemble and CNN-mixture of expert algorithm (CNN-MoE) classifiers to predict anomalies in respiration. The sound files are divided into 6 seconds of cycle length and achieved an accuracy of 86% using CNN-MoE. Verma *et al.* [12] used X-ray images to classify pneumonia using CNN. They performed data normalization and data augmentation techniques for the betterment of the overall accuracy. But the paper only classifies pneumonia. The other lung diseases are not classified. Chatterjee *et al.* [13] developed an algorithm to detect respiration phase from audio data and then transformed it into 2D spectro-temporal image to develop a CNN model which can only detect wheezes. The paper also discussed about how the wheezed model may be used towards assessment and computation of patient severity. However, the classification is done only on wheezes. Crackles and other lung sounds were not classified. So the conclusion cannot be inferred. Vaityshyn *et al.* [14] developed a CNN model for classification of pulmonary system diseases with the use of CNN which achieved an accuracy of 80%. Here only pulmonary diseases are classified but the overall disease classification is not found in this work. The lung sounds are classified on the basis of X-ray images and can classify only a limited set of parameters. The CNN model consists of 18 layers resulting in a low overall accuracy of 78%.

B. Predictions using SVM

Grønnesby *et al.* [4] used 5D vector classification to classify lung sounds using SVM and the method reported a precision of 86% and recall of 84% for classifying a crackle in a window. The low-dimensional feature vector made their SVM work faster and the model was trained in 1.44 seconds and 39 crackles were classified in 1.08 seconds. However, only 100 cycles of cross validation were shuffled and classification of crackles is only present which resulted in a low accuracy. Fatih *et al.* [15] presented a novel solution for lung disease classification where they used visual geometry group-16 (VGG-16) for deep feature extraction model to better the classification accuracy. They further used SVM to predict all class labels and their classification accuracy was boosted by 7.62%. However, classification accuracy was boosted but the overall accuracy did not improve and it came 65% which is very less in respect to a prediction based model. Haider *et al.* [16] used several classifiers like SVM, k-nearest neighbor (KNN), logistic regression (LR), and decision tree to classify diseases into normal and COPD. They concluded that by combining lung based features with spirometry data, the accuracy of COPD diagnosis can be improved to a large extent. The proposed approach can be significantly used in a clinical scenario where it can assist clinicians for automated COPD diagnosis. They applied a few approaches like linear SVM, quadratic SVM, cubic SVM where there accuracy came as 100% which is not ideal. This suggests that the model is overfitted.

C. Other Algorithms

Zennat *et al.* [17] proposed a solution for lung disease classification using the publicly available dataset. They performed data normalization technique on the original data to rescale the audio samples in a better position and also applied it for a better accuracy. They developed their own lung disease classification (LDC) system for high-performance classification. All types of data normalization and data augmentation are performed in this model but as it is totally based on their own LDC, so the system becomes more complex. Hai *et al.* [18] proposed a triple classification of respiratory sounds using optimized ResNet architecture with optimized S-transform (OST). They implemented the ResNet method with STFT and OST where they gained maximum accuracy of 98% using the ResNet with OST method. Pramono *et al.* [19] evaluated features for classification of wheezes and normal respiratory sounds. A linear classifier was used to determine the best features for classification by evaluating several performance metrics. Totality index attained the highest accuracy of 87%. In the project, only wheezes and normal sounds are classified but the other abnormal sounds or respiratory diseases are not classified.

The related works in this field is limited and to achieve a higher accuracy rate, proper pre-processing of the dataset is needed for elimination of other sounds. We applied data normalization and data augmentation for pre-processing and achieved an overall accuracy of over 92% which is clearly higher than rest of the works in this field.

III. Proposed Model

This section describes the proposed methodology and the 2-dimensional CNN used for the task of classification. The block diagram in Fig. 1 presents the different stages involved in the proposed methodology for respiratory disease classification. It is composed of three components – data pre-processing, training and testing of CNN, and lastly performance and analysis.

A. Data Pre Processing

The data set used in this project is collected from ICBHI challenge. The recordings are collected from 126 patients where there are a total of 920 recorded samples. The recordings were recorded from the patients via Littman 3200 electronic stethoscope and Littman classic II SE stethoscope. The

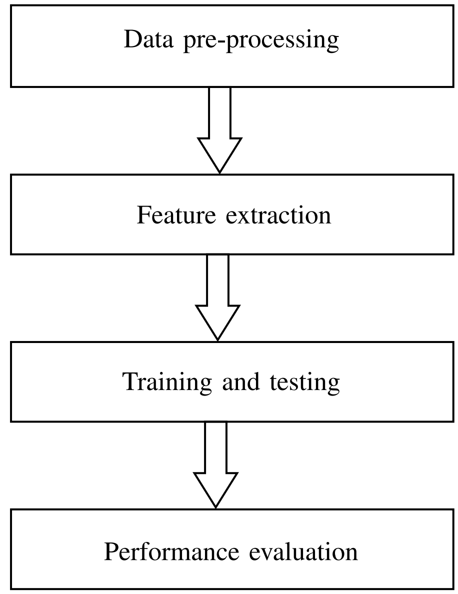

Fig. 1. Flowchart of main components of this literature

data set contains 6 types of data- bronchiectasis, pneumonia, bronchiolitis, COPD, URTI, and healthy. The dataset distribution is shown in Table I.

TABLE I
RESPIRATORY CYCLE DATABASE

Sr. No.	Diseases	Numbers
1	Bronchiectasis	16
2	Pneumonia	37
3	Bronchiolitis	13
4	COPD	793
5	URTI	23
6	Healthy	35

B. Feature Extraction

Two types of techniques named data normalization and data augmentation were applied to the lung sound database to extract spectrogram features from the generated audio files. Here data augmentation was employed in the form of audio stretching, i.e., speeding up and speeding down. During data augmentation, iteration through each sound file and extraction of feature was done using MFCC. These spectrogram features were passed to the 2D CNN for further classification.

C. Training and Testing

After feature extraction, these spectrogram features are given to the classifier for further classification. The classifier consisted of CNN which is used to identify bronchiectasis, pneumonia, bronchiolitis, COPD, URTI, and healthy category of sounds from the MFCC features extracted from the audio samples. The classifier consists of different convolution layers and max-pooling layers which are then followed by activation and fully connected layers. A sequential model consisting of four 2D convolutional layers and the dense layer as output is used. The convolutional layers are designed for feature detection. They work by sliding a filter window over the input and performing a matrix multiplication and storing the result in a feature map. During the forward pass, the filters are convolved between the height and width of the inputs. It produces a 2D block which consists of the dot product of the height and widths. The pooling layer is another building block of CNN. It reduces the spatial size of the representation to reduce the number of parameters. Pooling is of two types- average pooling and max pooling. To prevent overfitting, dropout layer is used after each convolutional layer. Rectified linear unit (ReLu) is used as an activation function in each convolutional layer to introduce non-linearity from the input to the output. At the top, there are two fully connected layers with softmax as an activation function for the output layers.

The input layer is taken in the form of sample height, sample width and number of filters. The filter parameter specifies the number of nodes in each layer. The number of filters in each convolutional layer is chosen to be 16, 32, 64, and 128 respectively. The kernel size parameter specifies the size of the kernel window, which is 2, resulting in a 2x2 filter matrix.

Each convolutional layer has an associated pooling layer of max-pooling 2D type with the final convolutional layer having a global average pooling 2D type. The pooling layer is to reduce the dimensionality of the model (by reducing the parameters and subsequent computation requirements) which serves to shorten the training time and reduce over fitting. The max-pooling type takes the maximum size for each window and the global average pooling type takes the average for over whole frame. A dropout value of 20% is used in the CNN layers.

The final output layer consists of 6 neurons which corresponds to six different classes- bronchiectasis, pneumonia, bronchiolitis, COPD, URTI, and healthy. The CNN architecture for the proposed model is given in Table II.

TABLE II
ARCHITECTURE OF THE PROPOSED MODEL

Label	Type	Output shape	Parameters
1	Input	$40\times862\times1$	-
2	Convolution 2D	$39\times861\times16$	80
3	MaxPooling 2D	$19\times430\times16$	0
4	Dropout	$19\times430\times16$	0
5	Convolution $2D_1$	$18\times429\times32$	2080
6	MaxPooling $2D_1$	$9\times214\times32$	0
7	$Dropout_1$	$9\times214\times32$	0
8	Convolution $2D_2$	$8\times213\times64$	8256
9	MaxPooling $2D_2$	$4\times106\times64$	0
10	$Dropout_2$	$4\times106\times64$	0
11	Convolution $2D_3$	$3\times105\times128$	32896
12	MaxPooling $2D_3$	$1\times52\times128$	0
13	$Dropout_3$	$1\times52\times128$	0
14	Global Average Pooling 2D	128	0
15	Dense	6	774

D. *Performance Analysis*

For evaluation and analysis purpose we use confusion matrix. The matrix is useful in describing classification accuracy and characterizing errors. The accuracy can be calculated by dividing the sum of the main diagonal of the confusion matrix by the total number of samples collected which is given by-

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}.$$

In the above equation, TP = True Positive, TN = True Negative, FP = False Positive and FN = False Negative.

To evaluate the model performance, precision and recall are given as-

$$Recall = \frac{TP}{TP+FN},$$

$$Precision = \frac{TP}{TP+FP}.$$

F1 score is a metric that takes precision and recall into account which is given as-

$$F1score = 2 \times \frac{Precision \times Recall}{Precision + Recall}.$$

IV. RESULTS AND DISCUSSION

We divided the dataset in the ratio of 80:20; 80% for the training set and 20% for the test set because the training set is large enough to yield statistically meaningful results and also it is a representative of the whole data set. The annotations comprised of six classes of respiratory diseases- bronchiectasis, pneumonia, bronchiolitis, COPD, URTI, and healthy. The disease count in the sound files is shown in Fig. 2.

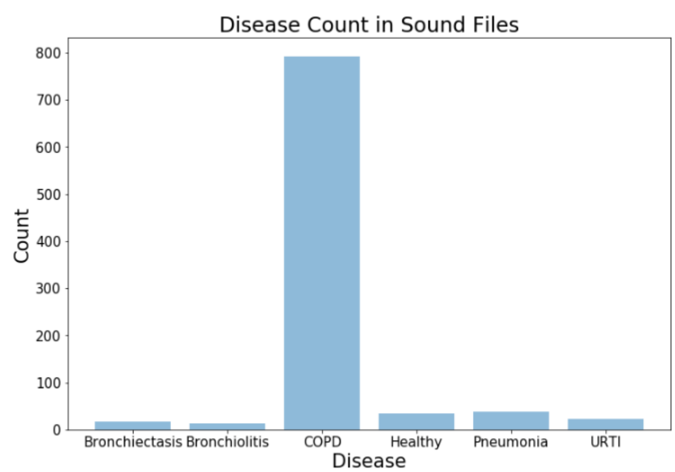

Fig. 2. Disease counts

So based on the division of training and testing set, different respiratory diseases are identified. For statistical analysis, the confusion matrix is developed. Based on the confusion matrix, the percentage of accuracy, recall, precision, and F1 score are achieved. The values are given in Table III.

It is observed from [8] that using 2D CNN, they achieved an overall accuracy of 82% but further they applied different types of normalization techniques like normal peak value augmentation, normal RMS value augmentation. After simulating

TABLE III
RESULT CLASSIFICATION OF RESPIRATORY DISEASES

Diseases	Precision	Recall	F-1 Score
Pneumonia	0.44	0.57	0.50
Bronchiolitis	1.00	0.67	0.80
Bronchiectasis	0.25	0.33	0.29
COPD	0.95	0.99	0.97
Healthy	0.67	0.29	0.40
URTI	1.00	0.20	0.33

the model with these, they gained the highest accuracy of 97%. We only applied 2D CNN where our accuracy came 92.39% which outperformed their accuracy.

From [15] it can be seen that using the same dataset, they extracted the features using the VGG16 classifier and classified the model using CNN and SVM. They gained an accuracy of 65.5% where they only classified the lung sounds such as crackles and wheezes. We used the recorded lung sounds and added featured classification which eventually helped us to classify respiratory diseases from the recorded lung sounds and achieved a better accuracy over them.

Minami *et al.* in his project [3] proposed two methods for classification. Firstly, they used Short-time fourier transform and wavelet transform to transform 1D into 2D time frequent representation images. Secondly, they classified the transferred images using CNN. The reported individual accuracies is harmonic score of 81%, sensitivity of 54% and specificity of 42% whereas our overall accuracy came much better than them.

The batch size was set to 128 in this project and the epochs was fixed at 500 to prevent any over or under fitting. Based on the epochs and batch size, the overall accuracy of the model was determined. The training parameters of the CNN model and the accuracy is given in Table IV.

TABLE IV
ACCURACY AND OTHER PARAMETERS

Optimiser	Batch Size	Epochs	Training Time(minutes)	Accuracy
Adam	128	500	95	92.39%

After 500 epochs the total accuracy of the model came 92.39%. Another similar project of Saraiva *et al.* [6] was studied where they also extracted features from the audio files using MFCC and classified the respiratory sounds using CNN. They divided the batch size on the basis of the learning rate of the model. For 0.0001 learning rate and 0.001 learning rate, the batch size was divided into 128 and 200 sets respectively, and the accuracies reported were 74.3% and 72% respectively. Moreover the training time was over 5hrs which is much more in respect to a training of a deep learning model. We divided the batch size to 128 and achieved an accuracy of 92.39% which is much better in comparison with the work proposed by Saraiva *et al.* [6]. Also our training time is only 95 minutes where we had run 500 epochs. The related works are shown in Table V.

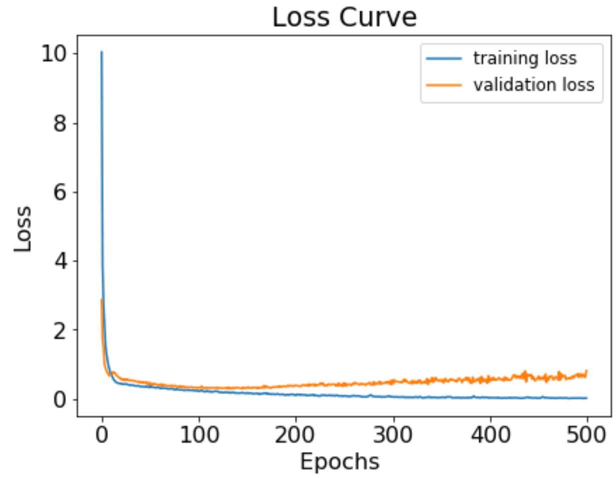

Fig. 3. Loss Curve

TABLE V
RELATED WORKS AND THEIR ACCURACY

Model	Reference	Accuracy
Proposed Method	-	0.92
KNN	[4]	0.84
Linear SVM	[4]	0.71
CNN (RMS Normalization)	[8]	0.87
CNN (Peak Value Normalization)	[8]	0.86
CNN (Ensemble Architecture)	[11]	0.80
CNN (SoftMax Architecture)	[15]	0.63
SVM	[15]	0.65
CNN (ResNet Architecture)	[18]	0.97
CNN (VGG-16 Architecture)	[20]	0.65
CpNN	[20]	0.84
BPNN	[20]	0.88
STFT	[21]	0.84

In Table V, our work is placed at the top with an accuracy of 92%. Most of the related works which are done by CNN and SVM have achieved an accuracy which is lesser than the present work. The work of Hai *et al.* [18] has only achieved an accuracy better than our work. In [12], respiratory sounds have been classified but in our work we have classified respiratory diseases on the basis of respiratory sounds. For the less availability of data of the diseases like bronchiectasis and pneumonia, our model cannot generate higher precision, recall and F1 scores of these diseases. We have introduced novel methods like data augmentation and data normalization for data pre-processing and cancelling unnecessary noises coming from other parts of the body whereas this is not present in any of the related works in this field. The related works where respiratory diseases have been classified, the proposed work outperforms their work by a large amount.

V. CONCLUSION

CNN model is implemented to train the dataset of respiratory diseases. Respiratory sound disorders are classified into 6 categories- bronchiectasis, bronchiolitis, pneumonia, COPD, URTI, and healthy. The overall accuracy of the model stands

92.39%. The overall article is divided into 3 parts. First the pre-processing of the data set is described which leads to the extraction of features from the audio clips. Then training and implementation of the 2D CNN model is done and finally results are achieved. From the results we can conclude that CNN are a feasible tool for detection of respiratory sounds from a lung sound database. This paper is expected to inspire other researchers to enable further research for the analysis of respiratory diseases.

REFERENCES

[1] S. Reichert, R. Gass, and E. Andrès, "Analysis of respiratory sounds: state of the art," *Clinical medicine. Circulatory, respiratory and pulmonary medicine*, vol. 2, pp. 45-58, May 2008.

[2] H. Yamamoto, S. Matsunaga, M. Yamashita, K. Yamauchi, and S. Miyahara , "Classification between normal and abnormal respiratory sounds based on stochastic approach," *20th International Congress on Acoustics*, vol.5, 2010, pp. 4144-4148.

[3] K.Minami *et al.*, " Automaic classification of large-scale respiratory sound dataset based on convolutional neural network," *International Conference on Control, Automation and Systems (ICCAS)*, 2019, pp. 804-807.

[4] M. Grønnesby, J. C. A. Solis, E. Holsbø, H. Melbye, and L. A. Bongo, " Feature extraction for machine learning based crackle detection in lung sounds from a health survey," *arXiv*, 2017, pp. 1-15.

[5] G. Chambres, P. Hanna, and M. Desainte-Catherine," Automatic detection of patient with respiratory diseases using lung sound analysis," *2018 International Conference on Content-Based Multimedia Indexing (CBMI)*, pp. 1–6, 2018.

[6] A.A.Saraiva *et al.*, "Classification of respiratory sounds using convolutional neural network," *11th International Conference on Bioinformatics Models, Methods and Algorithms*, January 2020, pp. 138-144.

[7] R.Dubey and R.M.Bodade, "A review Of classification techniques based on neural networks for pulmonary obstructive diseases," *Proceedings of Recent Advances in Interdisciplinary Trends in Engineering and Applications (RAITEA)*, 2019.

[8] I.Rebai, Y.BenAyed, W.Mahdi and JP.Lorre,"Improving speech recognition using data augmentation and acoustic model fusion," *Procedia Computer Science*, vol. 112, pp.316-322, 2017.

[9] Rocha *et al.*, " A respiratory sound database for the development of automated classification," *Springer Precision Medicine powered by pHealth and connectedHealth*, pp.33-37, 2017.

[10] Q. Chen*et al.*, "Automatic heart and lung sounds classification using convolutional neural networks," *IEEE Xplore*, pp. 1-4, 2016.

[11] Lam Pham *et al.*, "Robust deep learning framework for predicting respiratory anomalies and diseases", January 2020, submitted in *arXiv*.

[12] Garima Verma, Shiva Prakash, "Pneumonia classification using deep learning in healthcare," *International Journal of Innovative Technology and Exploring Engineering (IJITEE)*, vol. 9, pp. 1715-1723, February 2020.

[13] Soujanya Chatterjee, Md Mahbubur Rahman, Ebrahim Nemati, Jilong Kuang, "Respiration phase based wheeze detection using acoustic data from pulmonary patients under attack,"*13th EAI International Conference on Pervasive Computing Technologies for Healthcare - Demos and Posters*, 2018.

[14] V. Vaityshyn, H. Porieva and A. Makarenkova, "Pre-trained convolutional neural networks for the lung sounds classification," *IEEE 39th International Conference on Electronics and Nanotechnology (ELNANO)*, 2019, pp. 383-386.

[15] F.Demir, A.Sengur, V.Bajaj, " Convolutional neural networks based efficient approach for classification of lung diseases,"*Health information science and systems*, pp. 1-8, 2020.

[16] Haider NS, Singh BK, Periyasamy R, Behera AK, "Respiratory sound based classification of chronic obstructive pulmonary disease: a Risk Stratification Approach," *Machine Learning Paradigm. J Med Syst*, June 2019, pp. 254-266.

[17] Zennat Tariq, Sayed Khushal Shah, Yugyung Lee, "Classification using deep convolutional neural network," *International Conference on Bioinformatics and Biomedicine*, 2019, pp. 732-735.

[18] H.Chen, X.Yuan, Z.Pei, M.Li and J.Li, " Triple-classification of respiratory sounds using optimized S-transform and deep residual networks," *IEEE Access*, pp. 32845-32852, 2019.

[19] Renard Pramono, Anas Imtiaz, Esther Rodriguez-Villegas, "Evaluation of features for classification of wheezes and normal respiratory sounds," *PLOS ONE*, pp. 1-21, 2019.

[20] Rahib H. Abiyev, Mohammad Khaleel Sallam, "Deep convolutional neural networks for chest diseases detection," *Hindawi, Journal of Healthcare Engineering*, pp. 1-11, 2018.

[21] Lukui Shi, Kang Du, Chaozong Zhang, Hongqi Ma, Wenjie Yan, " Lung sound recognition algorithm based on VGGish-BiGRU," *IEEE Access*, pp. 1-1, 2019.