

# ASTHMA DETECTION FROM SPEECH SIGNALS

## Abstract

In recent years we find various categories of people suffering with asthma, which is a major cause of death and increased prominence of distress in minor categories of individuals. Numerous studies have also revealed that asthma patients generally have poor drug adherence, and that the intricacy of the regimen may be a contributing factor. The complexity of the regimen and its connection to adherence and asthma outcomes in Ethiopian asthma patients are not, however, known. As a result, this study evaluated how complex medication regimens affected asthmatic patients' medication adherence and asthma control.

Vocal cord vibration is a necessary component of speech production. However, asthmatic voice changes will happen because of the inflamed lung airways. Spirometry is a well-known method used to assess a patient's respiratory function and diagnose asthma. Speech data from the subjects included the words "She sells" and the vowel sounds /a:/, /e:/, /i:/, /o:/, /u:/, as well as the consonant /s:/. Praat software was used to analyse and create speech parameters from 33 samples.

This work specifies the prominence of detection of disease using speech signals. Machine Learning techniques plays a major role in analysing the Categorization of signals using classifiers in Social Spider Optimization in Genetic Algorithm (MSSO-GA) and other feature extraction techniques in Convolution Neural Networks (CNN)

**Keywords:** CNN, MSSO-GA, Spirometry

## Authors

### **Md. Asim Iqbal**

Department of ECE  
KUCE&T, Kakatiya University  
Warangal, Telangana, India.

### **Dr. K. Devarajan**

Department of ECE  
Annamalai University  
Tamil Nadu, India.

### **R Lakshman Naik**

Department of IT,  
KUCE&T, Kakatiya University  
Warangal, Telangana, India.

### **R. Sushmitha**

Department of CSE,  
KUCE&T, Kakatiya University  
Warangal, Telangana, India.

## I. INTRODUCTION

There are roughly 339 million individuals in the globe who suffer from asthma., leading to about a thousand deaths every day [1]. Due to inflammation of the airways, people who have asthma may suffer chest tightness, shortness of breath, wheezing, coughing, and other odd symptoms when breathing [2]. Spirometry, the widely accepted test for diagnosing and monitoring asthma symptoms, measures lung function by evaluating the speed and amount of exhaled air. During the spirometry test, patients are required to inhale deeply and then blow forcefully into a mouthpiece while wearing a nose clip to limit nasal passages. Spirometry data, such as forced expiratory volume in one second (FEV1) and forced expiratory capacity (FVC), as well as the ratio of FEV1/FVC, play a crucial role in determining the severity of asthma. The accuracy of spirometry results relies on patients' efforts, cooperation with the technician, and their willingness to engage in this demanding test [3-4], which can be particularly challenging for young children and the elderly.

## II. LITERATURE SURVEY

Using voice cues and the most effective classifier, Iqbal et al. [5-6] suggested a real-time detection and classification strategy for asthma disease. The suggested method makes use of a modified Social Spider Optimization with Genetic Algorithm (MSSO-GA)-based feature extraction in conjunction with a convolutional neural network (CNN) model for the purpose of asthma detection and classification. This allows for the categorization of normal and asthmatic people on the basis of the modifications in voice signals.

Iqbal et al. [7] proposed a method for the diagnosis and prognosis of asthma known as real-time detection and forecasting (RTDF). For the goal of diagnosing asthma and estimating the severity of the condition, the RTDF approach that is generally suggested makes use of the improved whale optimization (IWO) algorithm.

Iqbal et al. [8] proposed using a recurrent deep neural network (RDN-Net) for the purpose of advanced deep learning-based asthma prediction and classification. To begin, voice signals are pre-processed by using a method known as minimum mean-square-error short-time spectral amplitude (MMSE-STSA). This is done in an effort to eliminate the noises and enhance the linguistic aspects of the person's speech.

Asim Iqbal and his colleagues [9] proposed an optimal method for the detection of asthmatic illness using speech signals and hybrid machine learning. They referred to this method as the OADD-HML methodology. The enhanced weed optimization (IWO) method was used by the inventors of the OADD-HML approach for the purpose of asthma identification and forecasting.

Shrivastava et al. [10] presented a method that makes effective use of a machine-learning paradigm as their solution. The approach that was devised was put to use in order to identify speech problems caused by respiratory illnesses. The PRAAT program, coupled with the MFCC and LPC coefficients, was used to extract a number of other parameters of speech, including F1, F2, F3, etc.

The Respiratory and Drug Actuation (RDA) Suite is a benchmarking and research tool that was proposed by Fakotakis et al. [11]. The RDA Suite provides users with a dataset that contains respiratory and drug actuation noises.

Haba, et al. [12] suggested Using classification algorithms to derive asthma severity levels from EEG data in a remote and individualized innovative way to monitoring asthma severity levels.

Nabi et al. [13] developed a telemedicine program that, by using techniques from the field of machine learning, is capable of automatically assessing the amount of asthma severity that a patient is experiencing. This was accomplished by the program's ability to remotely monitor patients' lung function. The MFCC-based characteristics were suggested by the authors as a method for classifying asthma illness based on speech signals.

In their study, Haider and colleagues [14] suggested using a lung sound (LS) based approach that could be implemented on computers to categorize asthma and COPD patients. In this work, utilizing the suggested method, researchers were able to differentiate between healthy participants, participants with asthma.

Shrivastava et al. [10] proposed the development of a method using machine learning to differentiate between normal speech and speech that is associated with a number of different respiratory disorders. The examination of linguistic patterns would be used to achieve this goal. The accuracy, specificity (Sp), and sensitivity (Se), in addition to the area under the receiver operating characteristic curve (AUC), were used in combination with the multi-fold cross-validation strategy and the holdout method, respectively, in order to evaluate the effectiveness of the various classifiers.

Nathan et al. [15] suggested a method that uses deep learning from beginning to finish in order to diagnose chronic obstructive pulmonary disease. After that, we used transfer learning by making use of a network that had been pre-trained for a different audio-based task, and after that, we applied our very own shallower network on top of that as a binary classifier in order to determine whether a specific voice clip belonged to a COPD patient or an asthma patient.

In the field of medicine, Singhal et al. [16] suggested a voice-signal-based method for illness diagnostics. The goal of artificial intelligence (AI) is to tackle difficult issues in a variety of different industries. As a result of the use of AI, the standard procedures for diagnosing and treating illnesses are undergoing transformation.

Warule and colleagues came up with the concepts of a successive harmonic peak ratio (SHPR), normalized harmonic peak with reference to the first harmonic peak (NHPF), and normalized harmonic peak with regard to the maximum value of harmonic peak (NHPM) [17].

Deep learning was suggested by Aptekarev et al. [18] as a method for diagnosing bronchial asthma by listening to recordings of respiratory sounds. The proposed method is distinct from other approaches in that the trained model makes it possible to diagnose bronchial asthma (including differential diagnoses) with a high level of accuracy, regardless

of factors such as the patient's gender and age, the stage of the illness, and the time at which the sound recording was made. This is in contrast to other methods, which are limited in their diagnostic capabilities due to the absence of a trained model.

In their study, Islam et al. [19] suggested an automatic and non-invasive vocal pathology diagnosis method that was founded on a convolutional neural network (CNN). The suggested system operates via a two-stage process. It starts by differentiating unhealthy sounds from healthy ones, and then it categorizes abnormal voices as belonging to one of the three separate disorders after it has discovered them. Two convolutional neural networks (CNNs) are utilized for these goals; one of them functions as a binary classifier to recognize sounds that are indicative of pathological conditions.

### III. PROPOSE METHOD

- 1. Speech Pre-processing:** The algorithm consists of two primary sections that are used to identify and detect the temporal start and offset of inhalations. These portions are utilized in conjunction with each other. These components might be separated into their own sections if necessary. The first stage is to identify and discriminate events in the recordings that are consistent with inhalation, and the second step is to get rid of false positives, which are also referred to as inaccurate inhalation classifications.

Because MFCCs reflect the known shift in the essential bandwidth of the human ear with frequency, extracting mel frequency cepstral coefficients (MFCCs) is a popular technique to parameterization for vocalization. This is because MFCCs represent the change. This is because MFCCs model the relationship between frequency and the critical bandwidth of the human ear. It is well knowledge that the sounds produced by breathing have a distinct pattern that enables them to be differentiated from other noises [20]. This finding served as the basis for the development of an algorithm that could identify this pattern.

In the beginning, the system went through a training process using a collection of twenty different inhaler recordings that were chosen at random. Each signal was segmented into frames with a duration of 700 milliseconds that overlapped with one another every 20 milliseconds. After calculating 12 MFCCs for each frame included in the signal, a short-time cepstrogram of the signal was produced as a result. The cepstrogram was produced as a result. Applying the Singular Value Decomposition (SVD) method to the cepstrogram of the signal allowed for the generation of a normalized singular vector as a result of the analysis. In the process of analyzing breath sounds, MFCC calculations may be used to generate singular vectors, which can then be used to record the most important features of the sounds. An adaptive threshold is automatically established at a value that is 14% greater than the singular vector with the lowest value that was recorded by the inhaler. Those singular vectors that were below the adaptive threshold were not considered to be prospective inhaling events, while those that were above it were recognized as potential inhalation events. It was discovered via empirical research that using this adaptive threshold led to the most precise naming of occurrences and, as a consequence, inhalations in the training set.

In the second stage of the method, the zero-crossing rate (ZCR) (1) and the median amplitude were calculated in order to reduce the number of artifacts that were wrongly interpreted as being inhalations. These false positives were discovered by the algorithm. When compared to the parts of the training set in which there were no inhalations, it was discovered via empirical research that inhalations tend to have a ZCR that is much higher. In light of this information, a constant threshold value of 0.17 was chosen to serve as a fixed threshold. The inhalations always had a ZCR that was higher than the threshold value, and at the same time, false positives were effectively eliminated.

$$ZCR = \frac{1}{N} \sum_{n=N_0+1}^{N_0+N-1} \frac{1}{2} [\text{sign}(x[n]) - \text{sign}(x[n-1])] \quad (1)$$

A calculation was also done to determine the median amplitude of the planned inhaling event. A fixed threshold was implemented, in a manner similar to the ZCR threshold, with the goal of reducing the number of false positives using information acquired from the training set's empirical data. It was discovered that the median amplitude threshold value for inhalations was more than 0.012, and any artifact with a value that was lower than this threshold was disregarded. Through the use of empirical research, a specific combination of threshold values was determined in order to provide the most accurate identification of inhalations in the training set. As a direct consequence of this, it was applied to an entirely new validation set, which was made up of 255 separate files.

2. **Feature Extraction:** The wheezes feature extraction that was accomplished with the help of MFCCs is shown in Figure 2. First, a discrete FFT was applied to the respiratory sound segments in order to process them. In order to create the spectra, the next step was the computation of the log amplitudes of the squared absolute FFT. We made use of a Hanning window that was 256 points long (58 milliseconds) and had a 50% overlap in its points. As a consequence of this, In the spectrogram, the time scale interval was around 25 milliseconds, which is appropriate for wheezes.

An improved version of the Shabtai-Musih method [21] was implemented in order to maintain any wheezing peaks that may have been present. This was done in order to decrease the effect that spectrum noise has on MFCCs. In order to keep the potential of wheezing peaks, the mean value of each frequency band at 250 Hz was eliminated from a specific spectral segment. This was done in order to retain the possibility of wheezing peaks.

After that, the isolated peaks were normalized by dividing the standard deviation of the spectra. This was done in order to make the peaks comparable to one another. This was done to reduce the effect that variations in amplitude had on the process of obtaining respiratory sounds. In order to cut down on the number of peaks that were caused by the background noise, an experimentally determined peak threshold of 1.25 was used. A potential wheezing peak is indicated if there is an isolated peak that is greater than this peak threshold.

Following that, the pre-processed spectra were mapped by the utilization of the triangular overlapping windows, and their frequency bands were positioned logarithmically in accordance with the mel scale. The frequencies that constitute the mel

scale come quite close to approaching an approximation of the response of the human auditory system. After that, MFCCs were established by using the cosine transform on the spectrum outputs that were mapped via the use of mel scale filter banks.

Let us represent an N-point spectrum that has been pre-processed at iteration t by the notation  $s(t)$ , and let FB equal  $FB = [F_1, F_2, \dots, F_{N_f}]^T$  denotes a series consisting of  $N_f$  mel scale filter banks, with  $F_j$  denoting the  $j$ -th mel scale filter bank and  $N_s \times 1$  serving as the vector. Calculations may be made using to determine the outputs of spectra that have been mapped using mel scale filter banks.

$$\left[ O_1(t), O_2(t), \dots, O_{N_f}(t) \right]^T = FB \times S(t) \quad (2)$$

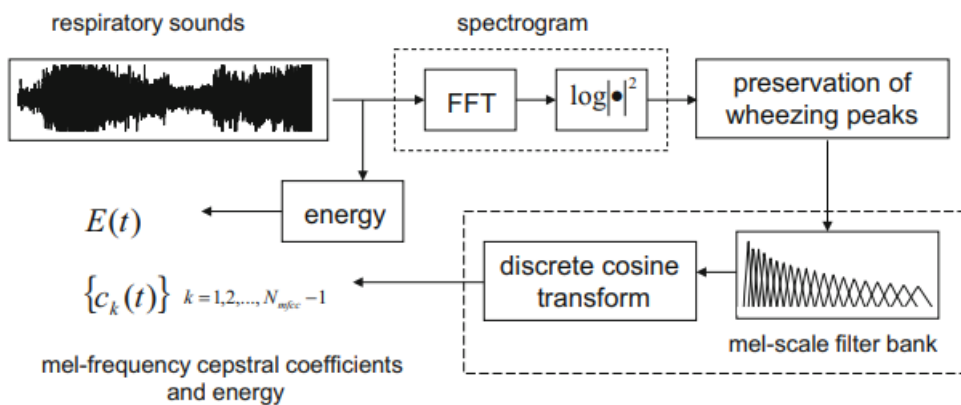
where  $O_j(t)$  indicates the result of applying the  $j$ th mel scale filter bank on the spectrum and mapping the results. It is possible to compute the  $k$ -th MFCC by using:

$$C_k(t) = \sum_{j=1}^{N_f} O_j(t) \cos \left( k(i + 0.5) \frac{\pi}{N_f} \right) \quad (3)$$

where  $N_{mfcc}$  is the total number of MFCCs in the system. The MFCCs and the energy of the respiratory sound segments were both used as feature parameters in the process of identifying wheezing. Let's assume the amplitude of the respiratory sounds at iteration t is represented by the sign  $w_t$ ; this would mean that it was measured in watts. Then, the value of the energy parameter  $E_t$  may be expressed as:

$$E(t) = \frac{\sum_{i=0}^{255} w(t-i)^2}{256} \quad (4)$$

The lower and higher coefficients of the MFCCs are used to characterize the envelopes and harmonic components of the spectra, respectively.



**Figure 1:** Feature extraction of respiratory sounds

In actual calculations, the zeroth cepstral coefficient, denoted by the symbol  $c_0(t)$ , which stands for the average value of the spectrum, is often disregarded. As a result, the computed formula for the removed feature vector  $v(t)$  at repetition  $t$  is as follows:

$$v(t) = \{c_1(t), c_2(t), \dots, CN_{mfcc}, E(t)\} \quad (5)$$

- 3. Acoustic GMM Models for Respiratory Sounds:** When developing a recognition system, it is necessary to collect a significant number of signals for each possible scenario; the data collected is then split into two halves. The first partition is going to be helpful for the construction of the acoustic model (the training set). The remaining partition is going to be utilized for assessment reasons, so that we can get an accurate indication of how accurate the system is. Because we do not have a sufficient number of signals, we conducted our analysis using the cross-validation approach, which is often used in situations in which there is an inadequate amount of data.

As shown by the following equation, the density of a Gaussian mixture is equal to the weighted sum of the densities of all M of its components:

$$p(\vec{x}|\lambda) = \sum_{i=1}^M m_i b_i(\vec{x}) \quad (6)$$

Where  $\vec{x}$  is a D-dimensional random vector ( $\vec{x}$  is an MFCC front-end),  $b_i(\vec{x})$   $i = 1, \dots, M$  are the module concentrations, and  $m_i$ ,  $i = 1, \dots, M$  are the combination masses.  $x$  is an MFCC front-end. Each component density may be represented by a Gaussian function with D-variables.

$$b_i(\vec{x}) = \frac{1}{(2\pi)^{D/2} |\Sigma_i|^{1/2}} \exp \left\{ -\frac{1}{2} (\vec{x} - \vec{\mu}_i) \Sigma_i^{-1} (\vec{x} - \vec{\mu}_i) \right\} \quad (7)$$

All of the mixture weights, as well as the mean vectors and the covariance matrices, are obtained from the Gaussian densities and are used to parameterize the whole GMM model. The following sentence is representative of the model that does the collecting of the parameters.

$$\lambda = \{m_i, \vec{\mu}_i, \Sigma_i\} i = 1, \dots, M \quad (8)$$

In order to compute the models  $\lambda = \{m_i, \vec{\mu}_i, \Sigma_i\}$ , the GMM approach employs the EM algorithm; in this scenario, M refers to the number of densities produced by each model, and  $m_i$  stands for the weight that is assigned to each Gaussian density that is included in the mixture. In order to optimize the model, the GMM computation was performed with a varied number of densities (ranging from 1 to 20), with the goal of finding the optimal balance between producing the best possible results and maintaining the lowest possible density numbers throughout the mixture. Using the Bayes decision rule and excluding  $p(x)$  from consideration since it does not change throughout the process of maximizing, the essential formula that gives the best hypothesis in automated signal identification is as follows:

$$C = \arg_{1 \leq i \leq j} \max p(X|\lambda_i) \quad (9)$$

since of the total number of recordings that represent both disease and individually normal breathing, a significant number of characteristic vectors are acquired. In addition, since it is assumed that the MFCC vectors are statistically independent of one another, the method to pick the best hypothesis leads to the following:

$$\prod_{t=1}^T p(\vec{x}_t|\lambda_i) > \prod_{t=1}^T p(\vec{x}_t|\lambda_r) \quad (10)$$

In the field of scientific speech processing and statistical pattern recognition [17-20], the term "similarity function" refers to the formula " $\prod_{t=1}^T p(\vec{x}_t|\lambda_i)$ ." It is vital, however, to simplify the computations by avoiding computational overrun in order to accommodate the fact that each recording comprises a large number of acoustic vectors. In view of the fact that the log function is monotonous, it is permissible to use it in the formulation of equation (10).

In the end, the product operation is converted into an addition operation, and the resulting expression is given the name of the rule of maximal similarity decision:

$$L(\lambda_l) = \sum_{t=1}^T \log p(\vec{x}_t|\lambda_i) \quad (11)$$

This last equation is employed in the comparing procedure, which is really an expression that is used to determine which hypothesis is the most likely to be correct. To put it another way, the input signal is connected to the acoustic model in our codebook that corresponds to the one that is the most probable.

#### IV. RESULTS AND DISCUSSIONS

##### Performance Evaluation

**Table 1: Speech signals-based asthma detection systems performance**

Metric	MSSO-GA [11]	RTDF [12]	MMSE-STSA [13]	Proposed CNN
<b>Accuracy (%)</b>	97.37	93.49	94.12	98.08
<b>Sensitivity (%)</b>	93.59	97.33	97.09	99.07
<b>Specificity (%)</b>	94.27	95.04	95.80	99.92
<b>F-measure (%)</b>	94.34	97.07	95.31	98.42
<b>Precision (%)</b>	93.49	93.13	93.99	99.80
<b>MCC (%)</b>	97.34	97.77	96.75	98.22
<b>Dice (%)</b>	97.83	97.49	95.03	99.36
<b>Jaccard (%)</b>	94.60	95.65	97.92	99.00

**Table 2: cough-based asthma detection systems performance**

Metric	MSSO-GA [11]	RTDF [12]	MMSE-STSA [13]	Proposed CNN
<b>Accuracy (%)</b>	97.37	93.49	94.12	98.08
<b>Sensitivity (%)</b>	93.59	97.33	97.09	99.07
<b>Specificity (%)</b>	94.27	95.04	95.80	99.92
<b>F-measure (%)</b>	94.34	97.07	95.31	98.42
<b>Precision (%)</b>	93.49	93.13	93.99	99.80
<b>MCC (%)</b>	97.34	97.77	96.75	98.22
<b>Dice (%)</b>	97.83	97.49	95.03	99.36
<b>Jaccard (%)</b>	94.60	95.65	97.92	99.00



**Table 3: lung sound-based asthma detection systems performance**

<b>Metric</b>	<b>MSSO-GA [11]</b>	<b>RTDF [12]</b>	<b>MMSE-STSA [13]</b>	<b>Proposed CNN</b>
<b>Accuracy (%)</b>	97.37	93.49	94.12	98.08
<b>Sensitivity (%)</b>	93.59	97.33	97.09	99.07
<b>Specificity (%)</b>	94.27	95.04	95.80	99.92
<b>F-measure (%)</b>	94.34	97.07	95.31	98.42
<b>Precision (%)</b>	93.49	93.13	93.99	99.80
<b>MCC (%)</b>	97.34	97.77	96.75	98.22
<b>Dice (%)</b>	97.83	97.49	95.03	99.36
<b>Jaccard (%)</b>	94.60	95.65	97.92	99.00

## V. CONCLUSION

In recent years, many of the people suffering with asthma, which is a major cause of death and increased prominence of distress in minor categories of individuals. Numerous studies have also revealed that asthma patients generally have poor drug adherence, and that the intricacy of the regimen may be a contributing factor. This work specifies the prominence of detection of disease using speech signals. Machine learning techniques used to analysing the Categorization of signals using classifiers in Social Spider Optimization in Genetic Algorithm (MSSO-GA) and other feature extraction techniques in Convolution Neural Networks (CNN) to detect the asthma. The implementation of proposed system shows the 98% accuracy in predicting the asthma disease. And also shown that proposed model is best suitable for the prediction of asthma disease as compared with other models.

## REFERENCES

- [1] Bor-Shing Lin, Bor-Shyh Lin. "Automatic Wheezing Detection Using Speech Recognition Technique", Journal of Medical and Biological Engineering, 2016
- [2] Holmes, Martin S., Marine Le Menn, Shona D'Arcy, Viliam Rapcan, Elaine MacHale, Richard W. Costello, and Richard B. Reilly. "Automatic identification and accurate temporal detection of inhalations in asthma inhaler recordings", 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2012.
- [3] "Proceedings of Data Analytics and Management", Springer Science and Business Media LLC, 2023
- [4] Pankaj Warule, Siba Prasad Mishra, Suman Deb, Jarek Krajewski. "Sinusoidal model-based diagnosis of the common cold from the speech signal", Biomedical Signal Processing and Control, 2023
- [5] MD. Asim Iqbal, Krishnamoorthy Devarajan, Syed Musthak Ahmed. "Real time detection and forecasting technique for asthma disease using speech signal and DENN classifier", Biomedical Signal Processing and Control, 2022
- [6] Iqbal, Md Asim, K. Devarajan, and Syed Musthak Ahmed. "Optimal convolutional neural network classifier for asthma disease detection using speech signals." International Journal of Healthcare Management (2023): 1-13.
- [7] Iqbal, MD Asim, Krishnamoorthy Devarajan, and Syed Musthak Ahmed. "Real time detection and forecasting technique for asthma disease using speech signal and DENN classifier." Biomedical Signal Processing and Control 76 (2022): 103637.
- [8] Md. Asim Iqbal, K. Devarajan, Syed Musthak Ahmed. "RDN-NET: A Deep Learning Framework for Asthma Prediction and Classification Using Recurrent Deep Neural Network", International Journal of Image and Graphics, 2023
- [9] Asim Iqbal, Mohammed, Krishnamoorthy Devarajan, and Syed Musthak Ahmed. "An optimal asthma disease detection technique for voice signal using hybrid machine learning technique." Concurrency and Computation: Practice and Experience 34.11 (2022): e6856.

- [10] Shrivastava, Poonam, and Neeta Tripathi. "Comparison of different Classification techniques for the detection of Speech affected due to Respiratory disorders." *Journal of Physics: Conference Series*. Vol. 2273. No. 1. IOP Publishing, 2022.
- [11] Nikos Dimitris Fakotakis, Stavros Nousias, Gerasimos Arvanitis, Evangelia I. Zacharaki, Konstantinos Moustakas. "AI Sound Recognition on Asthma Medication Adherence: Evaluation with the RDA Benchmark Suite", *IEEE Access*, 2023
- [12] Haba, Rotem, et al. "A remote and personalised novel approach for monitoring asthma severity levels from EEG signals utilizing classification algorithms." *Expert Systems with Applications* 223 (2023): 119799.
- [13] Nabi, Fizza Ghulam, et al. "A telemedicine software application for asthma severity levels identification using wheeze sounds classification." *Biocybernetics and Biomedical Engineering* 42.4 (2022): 1236-1247.
- [14] Haider, Nishi Shahnaj, and A. K. Behera. "Computerized lung sound-based classification of asthma and chronic obstructive pulmonary disease (COPD)." *Bio cybernetics and Biomedical Engineering* 42.1 (2022): 42-59.
- [15] Nathan, Viswam, et al. "Utilizing Deep Learning on Limited Mobile Speech Recordings for Detection of Obstructive Pulmonary Disease." 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE, 2022.
- [16] Singhal, Abhishek, and Devendra Kumar Sharma. "Voice signal-based disease diagnosis using IoT and learning algorithms for healthcare." *Implementation of Smart Healthcare Systems using AI, IoT, and Blockchain*. Academic Press, 2023. 59-81.
- [17] Warule, Pankaj, et al. "Sinusoidal model-based diagnosis of the common cold from the speech signal." *Biomedical Signal Processing and Control* 83 (2023): 104653.
- [18] Aptekarev, Theodore, et al. "Application of deep learning for bronchial asthma diagnostics using respiratory sound recordings." *PeerJ Computer Science* 9 (2023): e1173.
- [19] Islam, Rumana, Esam Abdel-Raheem, and Mohammed Tarique. "Voice pathology detection using convolutional neural networks with electroglottographic (EGG) and speech signals." *Computer Methods and Programs in Biomedicine Update* 2 (2022): 100074.
- [20] Junting Yang, Zulu Yang, Jing Zhou, Yong Zhao, Qifei Dai, Fuchi Li. "Parallel Attention with Weighted Efficient Network for VideoBased Person Re-Identification", 2021 the 5th International Conference on Innovation in Artificial Intelligence, 2021
- [21] Viswam Nathan, Korosh Vatanparvar, Keum San Chun, Jilong Kuang. "Utilizing Deep Learning on Limited Mobile Speech Recordings for Detection of Obstructive Pulmonary Disease", 2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2022

