Evaluation of MFCC Based Cough Event Computing in an Acoustic Signal

Krishvadana S
School of Electronics Engineering
Vellore Institute of Technology
Vellore, TN
krishvadana06@gmail.com

M.K.Pragadeshwar Vishnu
School of Electronics Engineering
Vellore Institute of Technology
Vellore, TN
mkpvishnu@gmail.com

Keerthi Balaji
School of Electronics Engineering
Vellore Institute of Technology
Vellore, TN
keerthi0kiran@gmail.com

Rajesh Kumar M
School of Electronics Engineering
Vellore Institute of Technology
Vellore, TN
mrajeshkumar@vit.ac.in

Abstract — **Cough detection is very crucial in evaluating the condition of patients suffering from tuberculosis and other various diseases as it is a very common indicative of such infections and diseases. This paper aims at detecting cough events using the most widely implemented MFCC Feature Coefficients. We evaluate these features for cough detection in terms of variance between the instance of cough events and normal sounds. The results obtained show that classification based on this comparison among MFCC coefficients leads to better results compared to the complex architecture used in other related works. Moreover, this technique is simple to implement and robust in nature.**

Keywords— Cough Sounds, MFCC, Variance, FFT

1. INTRODUCTION

Cough is our body's natural way of expulsion of air that occurs when the airways are clogged up with irritants and thus, helps the body to clear the airways. Although Cough reflex which is occasional seems to be healthy , it can be a root cause for many viral infections and respiratory illness.

Chronic Cough is associated with many factors such as smoking, contact with respiratory irritants and also use of certain medications ie.Anti hypertensives like Beta Blockers and Angiotensin converting enzyme inhibitors which are mostly cough inclusive. Irritants and environmental pollutants will result in hyperactive airways which will result in cough and wheezing. Asbestos exposure will result in respiratory symptoms such as cough and breathing difficulty other factors which put one or at risk of acquiring cough are work related factors/work related exposure in case of traffic policemen, construction site workers, welders and electricians

Also coughing events in public or heavily crowded places can result in spread of the virus by the original carrier. So isolation of those cough events is immediately needed especially during pandemic events.

Recent research states that detection of pulmonary diseases are possible[1] by training the collected healthy and sick datasets using machine learning classifiers. The results argue that cough sounds alone have a significant utility as a screening tool in distinguishing pulmonary diseases.

When it comes to signal processing, most computations and analysis will be carried out in frequency domain representation of the signal. The paper [2] proposes time domain analysis of the cough signal by comparing the number of peaks observed in the signal which is obtained after passing through a series of filters.

So, an accurate and simple cough detection method is desperately needed for the diagnosis of various diseases which have symptoms related to coughs.

The methodology proposed in this paper involves MFCC and Variance analysis, where MFCC is the common method to analyze audio signals. MFCC gives a high accuracy rate when it comes to identification based on human vocal systems. Degenerative Variance has been introduced here as part of the analysis which is derived from the chi-square algorithm. This analysis is done using the first two coefficients of the MFCC that is., MFCC-1 and MFCC-2.

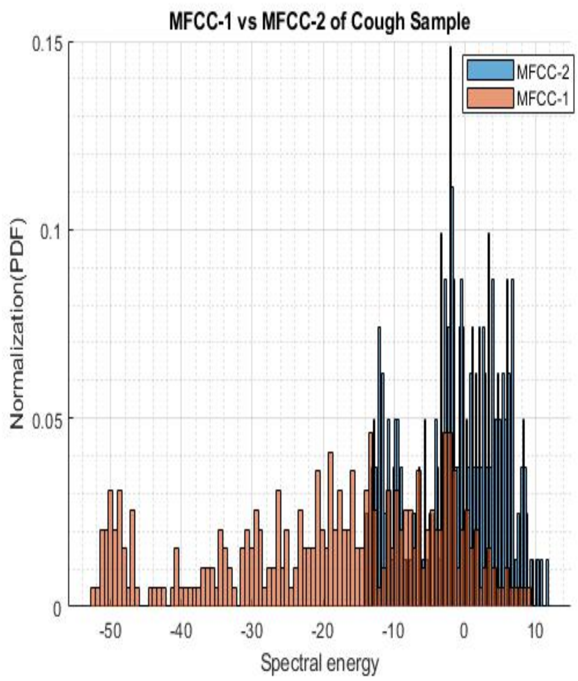

Fig 1

II. RELATED WORKS

During the course of research, various methods have been proposed for cough segmentation using various features and machine learning algorithms. Most of the time the common variable in cough segmentation seemed to be the classifier. From simple regression classifiers to a complex neural network has been used. For example a Convolutional Neural Network aids in detection of type of cough[3] by including spectrograms of coughing and comparing it with relevant respiratory sounds. By using Stochastic Gradient, ELU layer activation function the precision of this method in identifying cough comes to 0.75 to 0.80. Also another method which uses convolutional Neural Network is a K-band Continuous wave doppler sensor[4]. Even though the accuracy depends on the placement sensor it has an accuracy of 80% to 86.5% by including a fusion sensor system. Also two different classifiers are used in identifying cough sounds by addressing a sequence to sequence labelling problem [5] where the conclusion based network outperformed the recurrent network by 5% where the accuracy finally achieved is 92.7%.

A more complex Networks include a transformer networks which have constantly giving out accurate results in data classification even though they use a high amount of resources. One such application is where a self supervised Attention based transformer with three layered architecture, making use of Recurrent Neural Network[6] with the help of stacking up the predictions by pretraining, regularization and bootstrapping is argued to have the precision ranging between 0.74 and 0.8 which is better than the precision obtained without the use of transformer by 0.04%.

When it comes to simple regression classifiers, a three spectral feature model which uses a logistic regression model[7] gives out an accuracy of 90.31% by combining it with K-Nearest Neighbours classifier along local hu moments and 13 MFCC features. Again local hu moment has been used with KNN but also imposing effective tree structures[8] which can yield an accuracy of 93% and 18 times faster speed than the previous method.

A support Vector machine based classifier is also easy to implement with proper datasets. Such as exploiting a spectral structure using Non negative Matrix Factorization[9] Gives out an accuracy more than 90% by classifying the exploited structure with SVM.

Again a research work was made with SVM based classifier, including the gammatone based cepstral coefficient and evaluated by comparing the results with Mel frequency cepstral coefficients[10]. The GTCC outperforms MFCC by 0.02% with an accuracy of 94.28% thereby proving to be a better substitute for MFCC based feature Extraction techniques. A three classifier algorithm based detection, applying support vector machine, bayesian model and kNN is proposed[11] along with the implementation of hamming, rectangular and triangular windows with a resulting accuracy of 86.31%.

Other methods of identification also include wavelet transformation and few statistical parameter analysis of a cough signal by performing wavelet decomposition and by computing CWT coefficients. The type of disease is diagnosed by considering the threshold value of skewness and kurtosis[12].

Even though the above research yields good results they are difficult to implement and require good grasp in Machine learning. Hence we have proposed an efficient and simple system for cough event computing.

III.METHODOLOGY

The steps involved in this proposed work (As in Figure 2) :

a) Initially, cough samples used here have been compiled from various open sources. A total of 1200+ sample recordings were taken and compiled together with specific categories: WAV file format with 16 bit PCM encoding and driving through a single audio channel (1-MONO).

b) As the samples are susceptible to various background noises, it becomes necessary to perform the preliminary process of the raw samples. The various parameters that have been used to pre-process the signal are listed in Table1. First, the cough samples are resampled to 8000 Hz as most of the information about the cough is observed below 4000Hz. The removal of noise has been the most common initial process during any experimentation. The noise in the sample is removed using pre pre-emphasis filter by eliminating the unwanted samples.

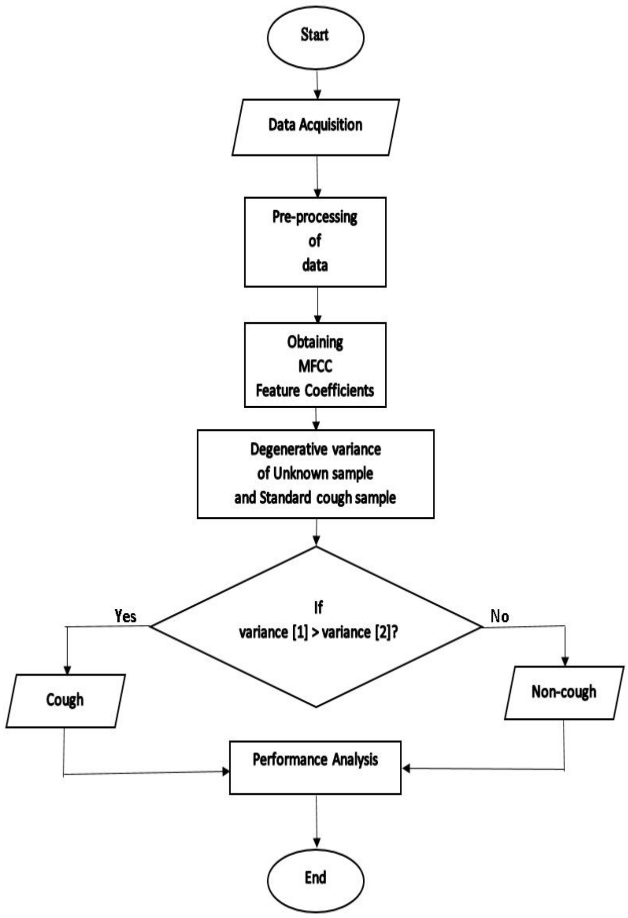

Fig 2. Process Chart of the proposed work

Table1. Pre-processing parameters used in this work:

Pre-processing parameters	
Sampling Rate	8000Hz
Filter used	Pre-emphasis filter, [-0.95, 1]
Window function	Blackman Window
Size of the Window	256ms
Overlapping (%)	50% i.e., 128ms

c) MFCC Coefficients of the FFT of these audio samples are evaluated by multiplying them with 20-Mel Filter Banks. These computed values are taken log and converted to the cepstral domain by,

$$X_k = \sum_{n=0}^{N-1} x_n \cos\left[\frac{\pi}{N}\left(n+\frac{1}{2}\right)k\right] \quad k = 0,\ldots, N-1$$

This gives out Mel-Frequency coefficients for each of the windows in the audio sample.

d) Event though 20 Coefficients are calculated only the first two coefficients are taken into account while calculating Degenerative Variance(D_v).The Coefficients of both the given sample and Identified cough sample is calculated from their respective Mfcc coefficients.It is given by,

$$D_v(j) = \frac{1}{i}\sum_i \frac{(m_{ji}-n_{ji})^2}{|m_{ji}+n_{ji}|}$$

That is MFCC-1 of the given audio sample and MFCC-1 of the Identified Cough sample is used in the above equation for finding the variance. MFCC-2 is also calculated in the same procedure. The result will be two values $D_v(1)$ and $D_v(2)$ for each given sample.

e) In the Decision tree, if the Degenerative variance of MFCC-1($D_v(1)$) is greater than the Degenerative variance of MFCC-2($D_v(2)$) the given sample is identified as a Cough sample, otherwise as Non-cough Sample.

IV.Experimental Results And Discussion

In this paper we experimented with 1200+ samples, containing different types of cough events and Non-cough Events. Fig 2.a gives the output of Cough signal with respect to an amplitude. The output of the MFCC feature set taken for 20 Filter banks from 0Hz to 8000Hz is shown in Fig 2.b. Here, MFCC-1 of the Coug signal varies from -35dB to +10db whereas MFCC-2 varies from -22dB to +8dB . This mapping of Energy to the frequency bands is called Mel Binning. As Discrete Cosine Transform is an Orthogonal Transformation, the output produced are mostly unrelated features. So it will easier to calculate each feature Coefficient separately without worrying about the overall deviation from the actual feature set.

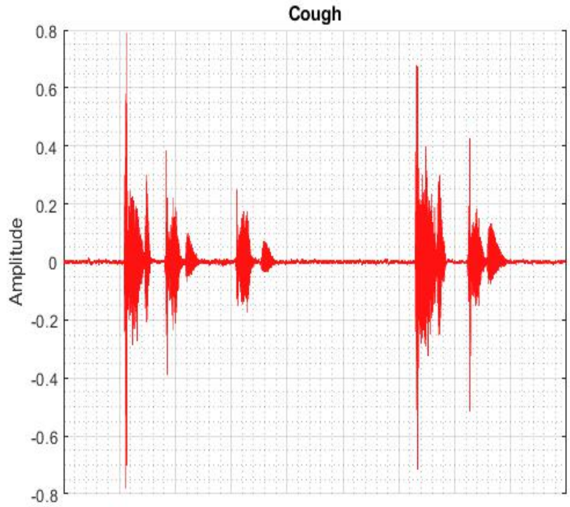

Fig 3.(a)

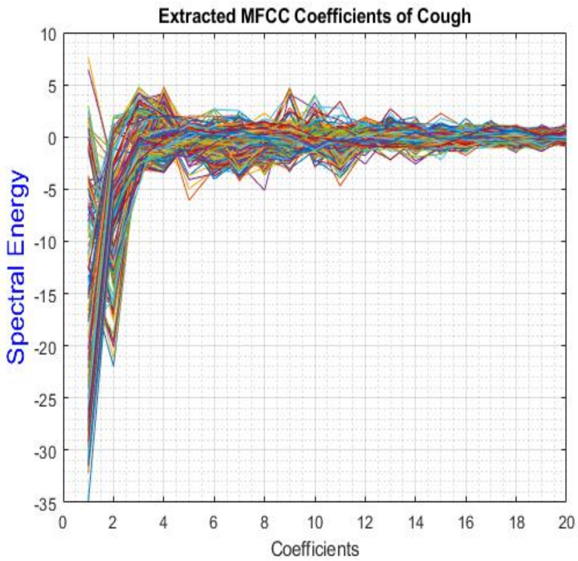

Fig 3.(b)

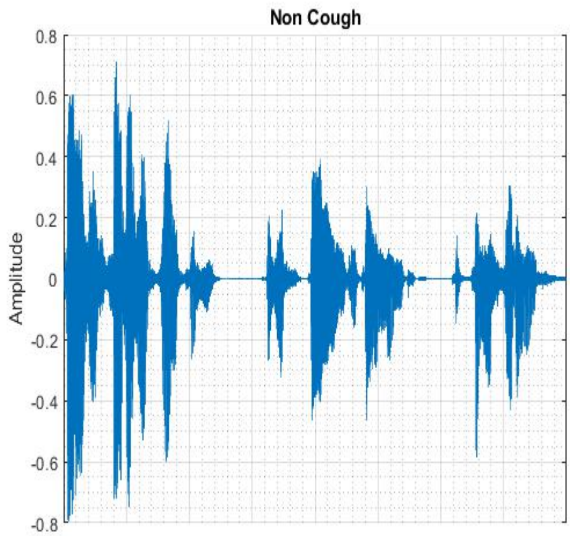

Fig 3.(c)

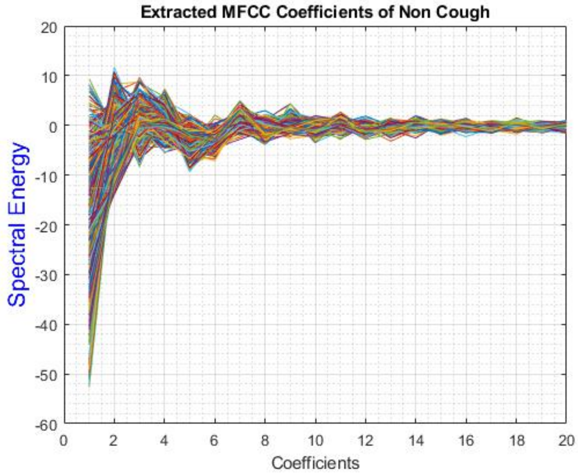

Fig 3.(d)

Fig 3.(a) Cough Signal (b) Extracted MFCC Feature coefficients of cough (c)Non Cough signal (d) Extracted MFCC Feature coefficients of Non cough

The Two figures i.e., Fig 3.c and Fig 3.d shows the Non-cough signal with respect to amplitude and its respective MFCC Feature set. In this MFCC-1 varies from -52dB to +10dB whereas MFCC-2 ranges from -15dB to +12dB. Most of MFCC-1 and 2 of both the Cough and Non-cough lie between the range -60dB to +10db(MFCC-1) and -20dB to +15dB(MFCC-2) as shown in Fig 4 and Fig 5 for multiple cough and noncough samples.

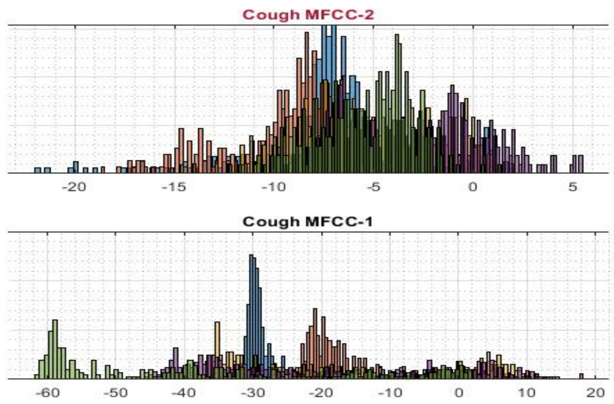

Fig 4.

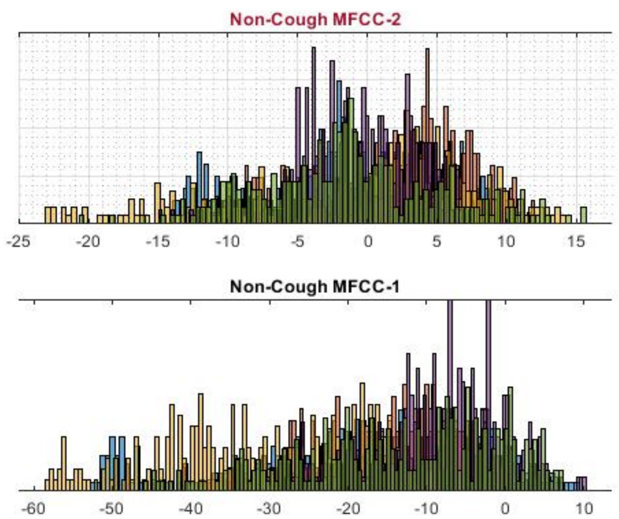

Fig 5.

For representation , a dataset was created with 10 cough samples and 5 Non-cough samples. First 10 samples of cough and cough were taken and $D_v(1)$ and $D_v(2)$ was calculated for finding the sensitivity. The resulting values are plotted as bar graphs in Fig 6.

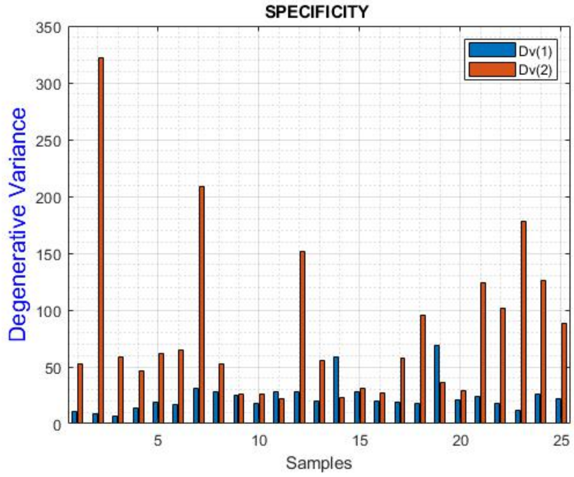

Fig 6.

We can see that among 25 combinations of variance from the dataset, 2 combinations of samples were mismatched or termed as False Negative (FN).

Again 5 Cough and Non-Cough samples were taken and the variance was calculated and plotted in Fig 7. Two values result as False Positive(Fp).

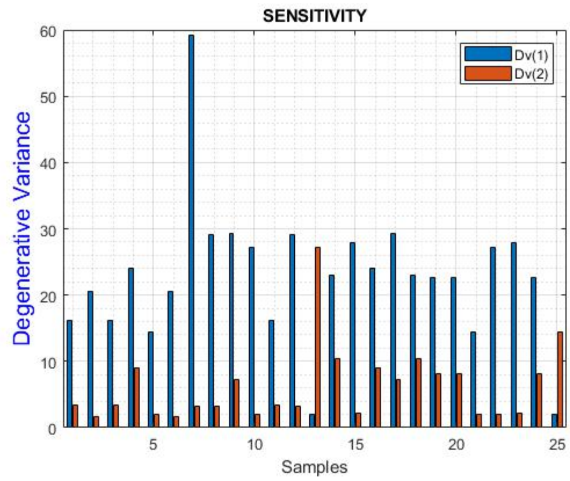

Fig 7.

Further implying this method with 1390 samples of cough and non-cough and nearly with more than 5000 combinations the sensitivity is averaged at 94.86% whereas the specificity is averaged at 88.2%.

Conclusion

In this paper, the proposed method for the identification of cough events is easy but gives an efficient output. The cough events are detected by the analysis carried out on the coefficients of the MFCC features. Variances calculated over the MFCC feature coefficients produce excellent results on par with the complex neural network architectures. Hence, we have proposed an efficient methodology for the evaluation of cough events in acoustic signals based on

MFCC Coefficients and Degenerative variance.Even though this method has excellent results it heavily depends on the preprocessing being performed strictly according to the method mentioned in the literature. Future work of this paper can be categorized into further improving the non-dependency of this proposed method on the preprocessing stage.

REFERENCES

[1] C. Infante, D. Chamberlain, R. Fletcher, Y. Thorat and R. Kodgule, "Use of cough sounds for diagnosis and screening of pulmonary disease," 2017 IEEE Global Humanitarian Technology Conference (GHTC), San Jose, CA, 2017, pp. 1-10, doi:10.1109/GHTC.2017.8239338.

[2] S. A. Walke and V. R. Thool, "Differentiating nature of cough sounds in time domain analysis," 2015 International Conference on Industrial Instrumentation and Control (ICIC), Pune, 2015, pp. 1022-1026, doi: 10.1109/IIC.2015.7150896.

[3] C. R. Rodriguez, D. Angeles, R. Chafloque, F. Kaseng and B. Pandey, "Deep Learning Audio Spectrograms Processing to the Early COVID-19 Detection," 2020 12th International Conference on Computational Intelligence and Communication Networks(CICN), Bhimtal, India, 2020, pp. 429-434, doi: 10.1109/CICN49253.2020.9242583.

[4] E. L. Chuma and Y. Iano, "A Movement Detection System Using Continuous-Wave Doppler Radar Sensor and Convolutional Neural Network to Detect Cough and Other Gestures," in IEEE Sensors Journal, doi: 10.1109/JSEN.2020.3028494.

[5] J. Amoh and K. Odame, "Deep Neural Networks for Identifying Cough Sounds," in IEEE Transactions on Biomedical Circuits and Systems, vol. 10, no. 5, pp. 1003-1011, Oct. 2016, doi: 10.1109/TBCAS.2016.2598794.

[6] G. Pinkas, Y. Karny, A. Malachi, G. Barkai, G. Bachar and V. Aharonson, "SARS-CoV-2 Detection From Voice," in IEEE Open Journal of Engineering in Medicine and Biology, vol. 1, pp. 268-274, 2020, doi: 10.1109/OJEMB.2020.3026468.

[7] R. X. Adhi Pramono, S. Anas Imtiaz and E. Rodriguez-Villegas, "Automatic Cough Detection in Acoustic Signal using Spectral Features," 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin,Germany, 2019, pp. 7153-7156, doi: 10.1109/EMBC.2019.8857792.

[8] C. Hoyos-Barceló, J. Monge-Álvarez, M. Zeeshan Shakir, J. -M. Alcaraz-Calero and P. Casaseca-de-la-Higuera, "Efficient k-NN Implementation for Real-Time Detection of Cough Events in Smartphones," in IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 5, pp. 1662-1671, Sept. 2018, doi: 10.1109/JBHI.2017.2768162.

[9] M. You et al., "Novel feature extraction method for cough detection using NMF," in IET Signal Processing, vol. 11, no. 5, pp. 515-520, 7 2017, doi: 10.1049/iet-spr.2016.0341.

[10] J. Liu et al., "Cough signal recognition with Gammatone Cepstral Coefficients," 2013 IEEE China Summit and International Conference on Signal and Information Processing, Beijing, 2013, pp. 160-164, doi: 10.1109/ChinaSIP.2013.6625319.

[11] B. Ferdousi, S. M. Ferdous Ahsanullah, K. Abdullah-Al-Mamun and M. N. Huda, "Cough detection using speech analysis," 2015 18th International Conference on Computer and Information Technology (ICCIT), Dhaka, 2015, pp. 60-64, doi:10.1109/ICCITechn.2015.7488043.

[12] T. H. Pingale and H. T. Patil, "Analysis of Cough Sound for Pneumonia Detection Using Wavelet Transform and Statistical Parameters," 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, 2017, pp. 1-6, doi: 10.1109/ICCUBEA.2017.8463900.