ARTIFICIAL INTELLIGENCE BASED AUTOMATIC WHEEZE DETECTION

[1] **Abhishek Pati Tripathi,** Pg scholar, Department of E.C.E, VIKAS GROUP OF INSTITUTIONS, Email id: tripathi.abhishekpati@gmail.com**.**
[2] **Yarram Uma Maheswari,** Assistant professor, Department of E.C.E, VIKAS GROUP OF INSTITUTIONS, Email id: umaecestaff@gmail.com.
[3] **V L N Phani Ponnapalli,** Associate Professor, Department of E.C.E, VIKAS COLLEGE OF ENGINEERING AND TECHNOLOGY, Email id: pvlnphani0454@gmail.com.

Abstract
Wheezing is one of the most prominent symptoms for pulmonary attack. Hence, wheezing detection has attracted a lot of attention in recent years. However, there is a dearth of a reliable method that can automatically detect wheezing events during each respiration phase in presence of several concurrent sounds such as cough, throat clearing, and nasal breathing. In this paper, we develop a model called WheezeD which, to the best of our knowledge, represents the first step towards developing a computational model for respiration phased based wheeze detection. Wheeze 2D has two components, first, we develop an algorithm to detect respiration phase from audio data. We, then transform the audio into 2-D spectro-temporal image and develop a convolutional neural network (CNN) based wheeze detection model. We evaluate the model performance and compare them to conventional approaches. Experiments on a public dataset show that our model can identify wheezing event with an accuracy of 82.508%, specificity of 83.69%, and sensitivity of 96.24%, which is over 10% improvement in performance compared to the best accuracy reported in the literature by using traditional machine learning models on the same dataset. Moreover, we discuss how WheezeD may be used towards assessment and computation of patient severity.

1 Introduction

Asthma could be defined more as a syndrome characterized by several different phenotypes Therefore, one of the possible definitions describing the characteristics of the disease and unifying more different definitions could define asthma as chronic inflammatory disease characterized by acute variable onset of symptoms (coughing, air deficiency, chest tightening) with bronchoconstriction (clinical definition) reversible and passes spontaneously or under the impact of therapy (pharmacological definition), followed by bronchial hyperactivity on different stimulants (functional definition) and the inflammation of different stage, duration and difficulty (biological definition). Cough variant asthma (CVA) is defined as a phenotype of asthma, which characterized by cough as the sole symptom and airway hyperreactivity (AHR). Corrao and colleagues first defined "cough variant asthma" as AHR, chronic cough and absence of wheezing . The authors agree that CVA and classic asthma have the same pathophysiological and immunological mechanisms, so CVA is considered a precursor of classic asthma A 5-year-old boy presented to the clinic because of prolonged dry coughing with no history of wheezing. Because boy could not do spirometry, a forced oscillation technique was made. The total respiratory resistance was decreased by −20.4% after beta-2-agonist inhalation. At the first visit, 2-week therapy of inhaled beta-2-agonist was started. This treatment was clearly effective against his cough. The CVA was diagnosed, and his treatment with leukotriene receptor antagonist (Montelukast) and LABA (tulobuterol patch) was started for next 8 weeks. Eight months later, boy has the same symptoms. The same treatment was restarting. Three years later, boy has another episode of a dry cough with no complaints of wheezing. A physician confirmed a wheeze during expiration by auscultation. The treatment with inhaled steroid (Fluticasone), LABA (Salmeterol) and leukotriene receptor antagonist (Montelukast) was started. Over time, after boy developed recurrent wheezing, the diagnosis of asthma was set. A 64-year-old female presented to the clinic as a self-referral complaining of a persistent cough." She said that the symptoms last for almost 17 years. The patient had diagnosed seasonal rhinosinusitis with positive skin prick test. Previous evaluations were all unremarkable. She underwent a methacholine challenge test. Spirometry showed increase in FEV1 with a 13% change from baseline. The patient was diagnosed with CVA and therapy with a combination of medium dose inhaled steroid and long-acting beta-2-agonist (Mometasone/Formoterol) was started. 32-year-old women presented with an intermittent non-productive hacking cough that had lasted several days." Her medical history was unremarkable, and previous evaluations were normal. Results of a methacholine challenge test showed severe airway hyperreactivity. The patient was diagnosed with CVA, and bronchodilator with ICS treatment was started. The prevalence of CVA is unknown, and from these cases it can be noticed that patients with chronic cough, as the only symptom, remain unrecognized as asthma for a long-time period. The isolated cough is less common than other clinical manifestations of classic asthma. Diagnosis of CVA may prove to be a challenge for the physicians. Therefore, evaluation results of patients with CVA are usually normal (spirometry, skin prick test, chest radiography, blood test). Previous clinical history is also normal in these patients. Clinical feature of CVA is a good response to bronchodilator and ICS therapy. Studies have shown that the ICS therapy in CVA patients prevents the development of classic asthma. Namely, it has been noticed that an average of 30% of patients with CVA without treatment develop classic asthma with wheezing. A smaller number, about 10% of patients with CVA and with adequate therapy (bronchodilator, ICS or Montelukast) develop classic asthma. A good response of chronic cough to the therapy with ICS cannot be used to distinguish other cough present diseases (atopic cough, non-asthmatic eosinophilic bronchitis) from CVA. It should be emphasized that in patients with chronic cough, a diagnostic evaluation for asthma should be performed.

2.Literature survey

It is a chronic inflammatory disease of the respiratory airway and can be hyper-responsiveness to a variety of stimuli [2]. The asthmatic patient suffers attacks such as coughing, dyspnea, and the main manifestation is wheezing [3]. Sounds generated during breathing can be a good source of information on lung's health [4]. Any characteristic changes of the normal lung sounds can imply a diseased condition that probably is invading the lung. Each type of disease is different from each other and the variation can be ascertained from sound characteristic, pitch, amplitude, frequency, duration, etc. [5]. With regard to asthma, symptoms originating from the wall oscillations of narrowed airways at critical flow rates causes wheeze to occur [3]. Wheeze is one of the adventitious sounds present in lung that is clinically defined as abnormal. The presence of wheeze, its location, duration and its relation to the respiratory cycle can be very useful to assist the physician as it has become a crucial practice in diagnosing and managing a number of pulmonary pathologies such as chronic obstructive pulmonary disease (COPD), bronchiolitis and commonly asthma [6]. Wheezes are continuous adventitious sounds that are superimposed on the normal breath sounds. According to the American Thoracic Society (ATS), the word "continuous" can be defined as the duration of the wheeze that is longer than 250 ms. The ATS also defines wheezes as high-pitched continuous sounds with a dominant frequency of 400 Hz or more. Wheezes can be detected and classified based on the frequency characteristics of its sinusoidal waves that justifies the musical character of the wheeze [7]. Conventionally, stethoscope is used to diagnose and monitor wheezes in asthmatic patients. Although it is well known that auscultation with stethoscope is reliable, fast and non-invasive, continuous monitoring of the respiration condition is impossible [8, 9]. Due to increasing number of asthmatic patients at present, there is a growing demand for automatic monitoring of the wheeze to assist the physicians in diagnosing and monitoring the patient. For asthmatic patients, continuous and automatic monitoring is essential as the daily symptoms can provide crucial information to the medical diagnosis [8]. Therefore, the electronic stethoscope, which is capable of recording and storing lung sounds, is available for many years now. This stethoscope can not only store the data obtained, but such data can be retrieved in future to aid in the interpretation of disease by medical personnel [10]. However, the problem seems to amplify as different physicians interpret the lung sounds differently. To overcome these problems, computerized approach has been developed over the past three decades for automated wheeze detection [11]. It is a bit time-consuming, but low in cost and reliable. Many researchers were involved in developing and improving these automated systems and many have succeeded in their research. A survey of literature shows that the main methodologies can roughly be classified into two categories: Fourier peaks detection and spectrogram image analysis. Exisiting Approach The Short-Time Fourier Transform (STFT) (or short-term Fourier transform) is a powerful general-purpose tool for audio signal processing [7,9,8]. It defines a particularly useful class of time-frequency distributions [4] which specify complex amplitude versus time and frequency for any signal. We are primarily concerned here with tuning the STFT parameters for the following applications: Approximating the time-frequency analysis performed by the ear for purposes of spectral display. Measuring model parameters in a short-time spectrum. In the first case, applications of audio spectral display go beyond merely looking at the spectrum. They also provide a basis for audio signal processing tasks intended to imitate human perception, such as auditory scene recognition [26] or automatic transcription of music [15]. Examples of the second case include estimating the decay-time-versus-frequency for vibrating strings [2,8] and body resonances [1,19], or measuring as precisely as possible the fundamental frequency of a periodic signal [1,6] based on tracking its many harmonics in the STFT [64]. An interesting example for which cases 1 and 2 normally coincide is pitch detection (case 1) and fundamental frequency estimation (case 2). Here, ``fundamental frequency'' is defined as the lowest frequency present in a series of harmonic overtones, while ``pitch'' is defined as the perceived fundamental frequency; perceived pitch can be measured, for example, by comparing to a harmonic reference tone such as a sawtooth waveform. (Thus, by definition, the pitch of a sawtooth waveform is its fundamental frequency.) When harmonics are stretched so that they become slightly in harmonic, pitch perception corresponds to a (possibly non-existent) compromise fundamental frequency, the harmonics of which ``best fit'' the most audible overtones in some sense. The topic of ``pitch detection'' in the signal processing literature is often really about fundamental frequency estimation, and this distinction is lost. This is not a problem for strictly periodic signals.

Mathematical Definition of the STFT
The usual mathematical definition of the STFT is [9]

HIDDEN MAKOV MODEL

The theoretical foundation of HMM is established by Baumet al., which is promoted by Rabiner and others, which has strong modeling ability to the time domain signal so that it has become a research hotspot [16, 25]. It has been successfully used in speech recognition, behavior recognition, and character recognition and fault diagnosis. HMM is a powerful statistical tool on describing discrete time data samples, it is a double random process, one of which is the Markov chain, which is a basic stochastic process to describe the transfer of state. Another stochastic process describes the statistical correspondence between States and observations. Because the state cannot be seen directly, HMM is a random process to perceive the existence of the state and its characteristics. A hidden Markov model can be defined by:

{S} –a set of state including an initial observation state S_M and a hidden state S_N;
A –the transition probability matrix, A = a_{ij},where a_{ij} is the transition probability of taking the transition from state i to j;

B –the output probability matrix, B = $b_j(O_k)$ for discrete hidden Markov model or B = $b_j(x)$ a continuous hidden Markov model, where O_k stands for a discrete observation symbol, and x stands for continuous observations of k–dimensional random vectors.

If the initial state distribution $\pi = \{\pi_i\}$, the complete parameter set of the hidden Markov model can be expressed compactly as $\lambda = \{N, M, A, B, \pi\}$.

Model initialization.
The initialization of hidden Markov model is to confirm the original value of hidden Markov model a set of five parameters $\lambda = (N, M, A, B, \pi)$. The sequence of driving style and braking behavior constitute a Markov chain. The number of hidden states (driving style) is 3, which are aggressive, moderate and mild; the number of observation status (braking behavior) is 2, which are general breaking and emergency breaking. And it is necessary to consider two factors to observe the number of sequences: the recognition accuracy and the recognition duration of the algorithm. If the sequence is too long, then the matrix dimension is too high, recognition duration is reduced with amount of calculation; if the sequence is too short, it is difficult to reflect the relationship between each node of the Markov chain, which would result in lower recognition rate. According to set the sequence of different length, it can be found that when the sequence length is 8, the fitting effect of the model is the best [24]. For the observation sequence length T, which should be considered whether it can describe the braking characteristics sufficiently, T is fixed as 8 in the experiment. Eight consecutive braking characteristics constitute an observation sequence. Traditional incident clearance time studies rely on statistical models with rigorous assumptions [26]. It is generally believed that the initial parameter selection has little effect, which can be randomly or evenly selected [27]. The initial state probability distribution vector can be selected as: $\pi_0 = (1, 0, \cdots, 0)'$. the initial probability distribution initialization of matrix A is uniform distribution, the scale of matrix A is 3 × 3, $A_0 = [a_{ij}]_{3\times3}$, $a_{ij} = 1/3$. The initial probability distribution of mixed matrix B is confirmed by different braking characteristics prior probability. According to the selection of the number of hidden states N and the number of observations M, the scale of matrix B is 3 × 2, and the initial probability distribution of matrix B is uniform distribution, $B_0 = [b_{ij}]_{3\times2}$, $b_{ij} = 1/3$. After stipulation of the initial value of hidden Markov model parameters, the initial model of the hidden Markov model λ_0 can be obtained.

4.Proposing Scheme

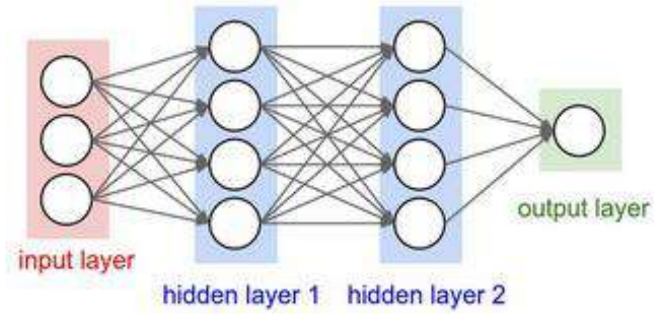

Figure 1: An example of the traditional neural network

Neural networks are the techniques of machine learning. They are just like the neural networks in biology. There are many neurons and many connections between neurons. Figure 1 is an example of the neural network. The white circles represent neurons and the arrows represent the connections between neurons. Note that the connections are directed, therefore we use arrows to represent it. In this section, we will introduce what the neural network is.

First, we need to know what the neuron is. In biology, neurons have inputs, thresholds, and output. If the input voltage is larger than threshold, the neuron will be activated and a signal is transmitted to output. Note that the neuron might have many inputs but there is only one output signal. The operation model of the neuron in machine learning is very like the one in biology. They also have the inputs and outputs. Despite the neuron’s output is connected to many neurons in Figure 1, the value of the outputs are the same. Of course, there are some differences of them. Instead of the threshold, the “neuron” in machine learning use a function to transfer the inputs to the output. There are many choices of the activation function. We often choose it as the sigmoid function $\sigma(x)$.

$$\sigma(x) = \frac{1}{1+e^{-x}} \qquad (1)$$

The sigmoid function is very similar to the step function, which acts similar to thresholding. When x is a large positive number, the output of the sigmoid function is near to 1. When x is much smaller than 0, the output is near to zero. We can see these factst in Figure 2. Another good property is that the sigmoid function is continuous and differentiable. So we can apply some mathematics on it.

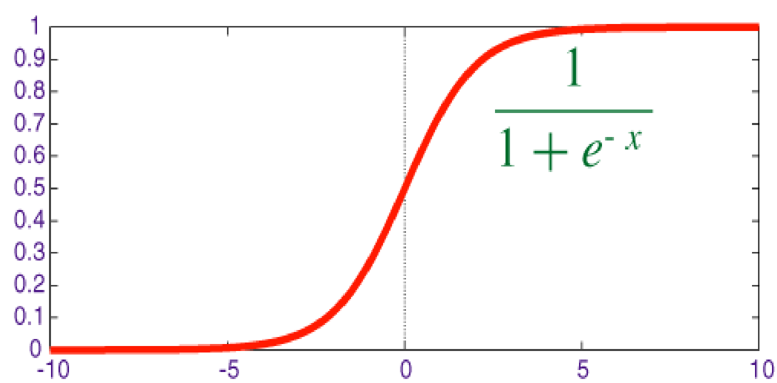

Figure 2: The sigmoid function [7]

Another difference is the weight. The weights describe that how much each input affects the neuron. That is, we will not just put every inputs into the activation function. The value of activation function's input is the linear combination of the inputs. The mathematical representation is as follows:

$$\sigma(w_1x_1 + w_2x_2 + \cdots + w_N x_N) \qquad (2)$$

where N is the amount of the inputs, w_i are weights of x_i, and σ( ) is the activation function. However, there is a problem of it! We reduce the amount of inputs to 1 and change the weight to observe how weights influence on the output. The result is shown in **Error! Reference source not found.**(a). One can see that 0 can be viewed as the threshold to determine whether the output is near to 0 nor near to 1. However, how do we modify the model if we want to change the threshold to a value other than 0? In this case, we add a bias θ to achieve that so that we can shift the sigmoid function. The result of the sigmoid function with bias is shown in **Error! Reference source not found.**(b). So the new relation is revised as follows:

$$\sigma(\theta + w_1x_1 + w_2x_2 + \cdots + w_N x_N) \qquad (3)$$

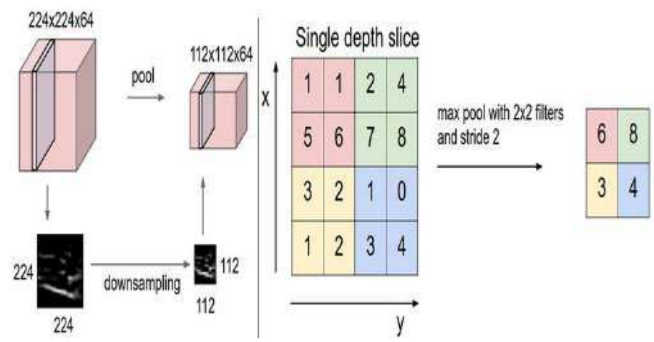

Figure 3: A simple example of the pooling layer

RESULTS
Normal signal

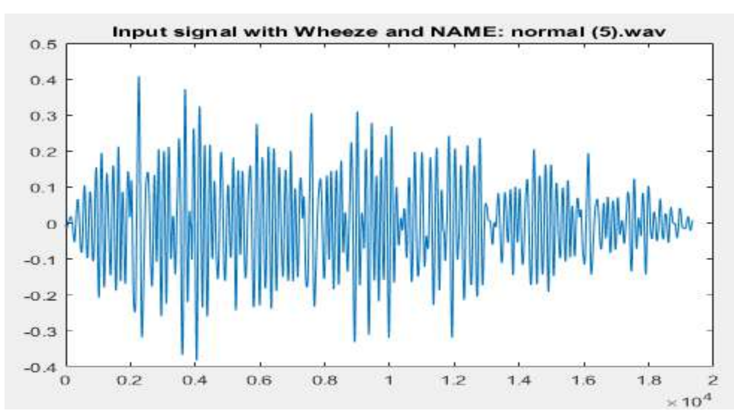

Figure 4: Wheeze signal from database name normal (5).wave

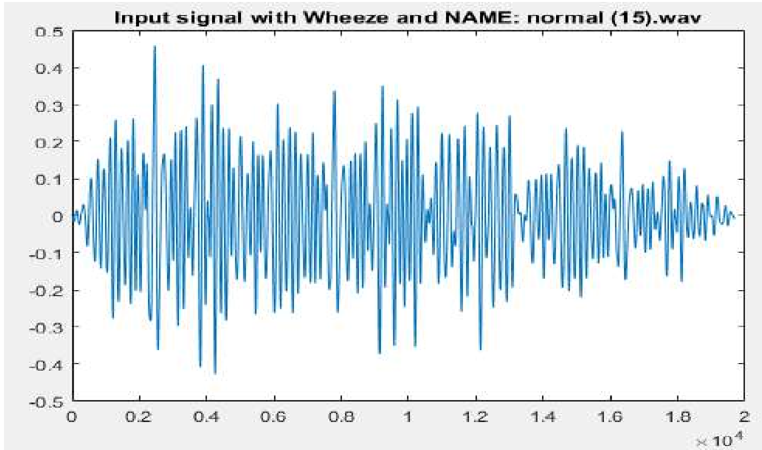

Figure 5: Wheeze signal from database name normal (15).wave

Abnormal Signals

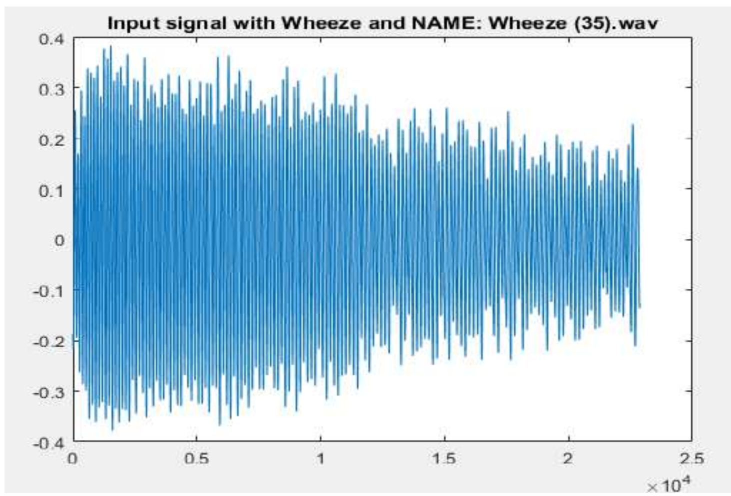

Figure 6: Wheeze signal from database name wheeze (35).wave

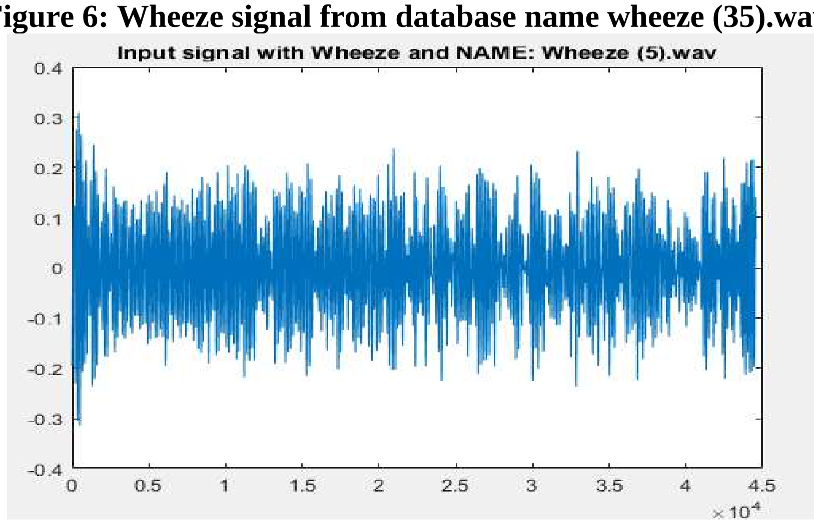

Figure 7: Wheeze signal from database name wheeze (5).wave

STFT BASED SPECTRUM FOR BOTH NORMAL AND ABNORMAL

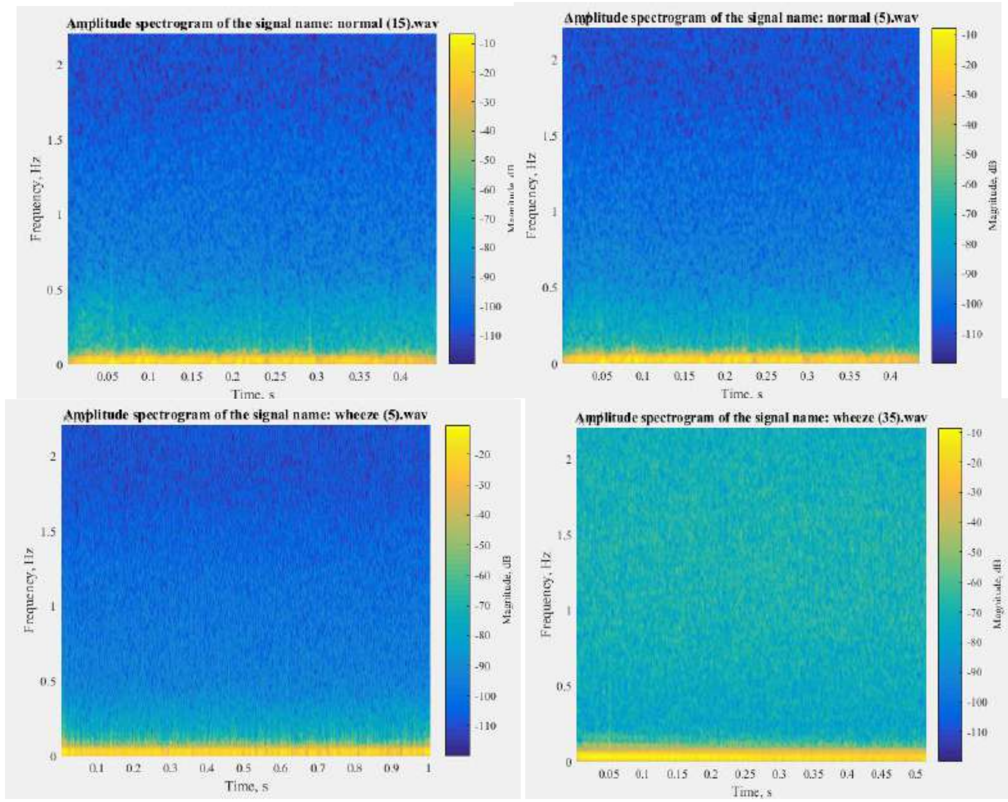

Figure 8: Spectrum analysis from STFT for both normal and abnormal signals

TABLE 1: Accuracy, Execution time, Sensitivity and Specificity for HMM classifiers and CNN classifiers

	HMM	CNN
TESTING EXECUTION TIME	2.59sec	**1.52sec**
SENSITIVITY	59.83%	**82.508%**
SPECIFICITY	62.48%	**83.69%**
ACCURACY	64.5%	**96.24%**

CONCLUSION FUTURE SCOPE

Following are the few limitations which may lead to interesting research opportunities. First, we use limited dataset of wheezing sounds in this work, which in turn restricts the model to two con-volutional layers.

In future, large amount of data can be collected in order to build deeper and more robust networks. Moreover, we did not explore LSTM based model, which we believe may improve the detection taking into consideration the temporal signature of wheeze in the acoustic data. Second, the detection method will help us collect long term longitudinal wheeze data from patients, which may be useful in assessing triggers of asthma or COPD ex-acerbation in the wild. Our hope is that, respiratory phase based wheezing detection, like WheezeD can be utilized towards assess-ment of severity.

ISSN: 0972-2750
Vol-12 Issue-12 No.01: 2022

For instance, severity may be associated with the wheezing duration, rate and diurnal pattern (for e.g., wheezing at night time) [22]. Moreover, severity can be determined by the greater level of obstruction of upper airways, and wheezing during inspiration is a defined surrogate for that [18]. Hence wheezing during inspiration phase may be assessed to be more severe than during expiration. Interestingly, we note that WheezeD presents the fundamental attributes required (respiration phase detection and wheezing detection with much better performance than existing works) to assess wheezing severity.

REFERENCES

[1] Accessed Aug, 2018. CDC Asthma. https://www.cdc.gov/nchs/fastats/asthma.htm
[2] Accessed Aug, 2018. CDC COPD. https://www.cdc.gov/nchs/fastats/copd.htm
[3] Accessed Aug, 2018. SoundSnap. https://www.soundsnap.com/
[4] Manuel Abella, John Formolo, and David G Penney. 1992. Comparison of the acoustic properties of six popular stethoscopes. The Journal of the Acoustical Society of America 91, 4 (1992), 2224–2228.
[5] Jont B Allen and Lawrence R Rabiner. 1977. A unified approach to short-time Fourier analysis and synthesis. Proc. IEEE 65, 11 (1977), 1558–1564.
[6] Justice Amoh and Kofi Odame. 2015. DeepCough: A deep convolutional neural network in a wearable cough detection system. In Biomedical Circuits and Systems Conference (BioCAS), 2015 IEEE. IEEE, 1–4.
[7] Henry Gong et al. 1990. Wheezing and asthma. (1990).
[8] Song Han, Huizi Mao, and William J Dally. 2015. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. arXiv preprint arXiv:1510.00149 (2015).
[9] PixSoft Inc. Accessed Aug, 2018. R.A.L.E Lung Sound. http://www.rale.ca/
[10] Richard S Irwin, Peter J Barnes, FRS Deputy, and Helen Hollingsworth. 2013. Evaluation of wheezing illnesses other than asthma in adults. UpToDate. Waltham: UpToDate (2013).

[11] [11] Yann LeCun, Léon Bottou, Yoshua Bengio, and Patrick Haffner. 1998. Gradient-based learning applied to document recognition. Proc. IEEE 86, 11 (1998), 2278– 2324.
[12] P Mayorga, C Druzgalski, RL Morelos, OH Gonzalez, and J Vidales. 2010. Acoustics based assessment of respiratory diseases using GMM classification. InEngineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE. IEEE, 6312–6316.
[13] Ebrahim Nemati, Md Mahbubur Rahman, Viswam Nathan, and Jilong Kuang. 2018. Private Audio-Based Cough Sensing for In-Home Pulmonary Assessment using Mobile Devices. In EAI International Conference on Body Area Networks.
[14] Hans Pasterkamp, Steve S Kraman, and George R Wodicka. 1997. Respiratory sounds: advances beyond the stethoscope. American journal of respiratory and critical care medicine 156, 3 (1997), 974–987.
[15] Renard Xaviero Adhi Pramono, Stuart Bowyer, and Esther Rodriguez-Villegas. 2017. Automatic adventitious respiratory sound analysis: A systematic review. PloS one 12, 5 (2017), e0177926.
[16] Md Mahbubur Rahman, Ebrahim Nemati, Viswam Nathan, and Jilong Kuang. 2018. InstantRR: Instantaneous Respiratory Rate Estimation on Context-aware Mobile Devices. In EAI International Conference on Body Area Networks.

[17] RJ Riella, P Nohama, and JM Maia. 2009. Method for automatic detection of wheezing in lung sounds. Brazilian Journal of Medical and Biological Research 42, 7 (2009), 674–684.

[18] Malay Sarkar, Irappa Madabhavi, Narasimhalu Niranjan, and Megha Dogra. 2015. Auscultation of the respiratory system. Annals of thoracic medicine 10, 3 (2015), 158.

[19] Chang S Shim and M Henry Williams. 1983. Relationship of wheezing to the severity of obstruction in asthma. Archives of internal medicine 143, 5 (1983), 890–892.

[20] Germán D Sosa, Angel Cruz-Roa, and Fabio A González. 2015. Automatic detec-tion of wheezes by evaluation of multiple acoustic feature extraction methods and C-weighted SVM. In 10th International Symposium on Medical Information Processing and Analysis, Vol. 9287. International Society for Optics and Photonics, 928709.

[21] DR Taylor, ED Bateman, LP Boulet, HA Boushey, WW Busse, TB Casale, P Chanez, PL Enright, PG Gibson, JC De Jongste, et al. 2008. A new perspective on concepts of asthma severity and control. European Respiratory Journal 32, 3 (2008), 545–554.

[22] M Yumiko, S Kazuko, and N Asako. 2015. Pediatric respiratory severity score (PRESS) for respiratory tract infections in children. Austin Virol and Retrovirology 2, 1 (2015), 1009.

[23] Limpert R. Brake design and safety- 2nd edition. Braking. 1999 Dec;61–64.

[24] Cheng H, Zhang J. IV, GMM or likelihood approach to estimate dynamic panel models when either N or T or both are large. Journal of Econometrics.2015 Mar;187(1):312–322.

[25] Nikolaus M, Franz P. Maximum margin hidden Markov models for sequence classification. Pattern Recognition Letters.2016 Mar;1–4.