RESEARCH ARTICLE OPEN ACCESS

Detection of Lung Disease Using MFCC AND CNN

V.Subapriya*, Dr.K.L.Shunmuganathan**, Kypa Sandhil ***, Naresh Mayya***, Gopireddy Tejeswarreddy ***
*(Faculty, Department of CSE, AVIT, Chennai
Email: subapriya.cse@avit.ac.in)
** (Principal, AVIT, Chennai)
***(Student,Department of CSE, AVIT, Chennai)

--************************----------------------------------

Abstract:

Lungs sounds has been used as trustworthy tool for diagnosis of various diseases for centuries. The growth in science and technology has developed different tools for identification of different types of diseases related to lungs. However we cannot fully depend on these tools for identification of various lungs related problems as using these tools is more or less subjective. In a study conducted among the physicians in US it was found that the doctors were unable to characterize and identify only 70% to 80% of wheeze sounds in a series of recording. In developed and under developing countries there is a scarcity of trained doctors. In such cases the accuracy of sound identification is much worse which will lead to misdiagnose of pulmonary diseases and their proper lungs sounds. Hence there is a need for technique which will classify various abnormal sounds such as wheeze and crackles against the normal one. In this paper we have proposed a system to classify the lung sounds into normal,wheeze, crackle and both that are taken by an electronic stethoscope. The proposed method achieves an accuracy of 80% and 66% for training and testing respectively.

Keywords — **Respiratory cycles, classes, MFCC,CNN**

--************************----------------------------------

I. INTRODUCTION

Auscultation is the way toward listening to the organ sounds in the human body by utilizing a stethoscope. It has been a viable apparatus for the conclusion of lung issue [10]. This procedure for the most part depends on the physicans knowledge and expertise. Utilizing a stethoscope, for diagnosing anomalies in the respiratory system, doctors may hear ordinary breathing sounds, diminished or missing breath sounds, and strange breath sounds. Though there is advancement in technologies this basic method of auscultation is still widely in practice today as auscutlation via using a stethoscope is not only cost effective but also non invasive.In the last 35 years,various machine learning and signal processing techniques has been applied on the breath sound signals as interest in these field has been growing in the recent years.

A.*ANALYSIS USING AUSCULTATION*

The stethoscope is an acoustic medical device for auscultation, or checking out the internal traces of an animal or human body. It has a little circle framed resonator that is set against the skin, and a few chambers related with two earpieces. A 2012 research paper asserted that the stethoscope, when contrasted with the other restorative hardware, had the most noteworthy positive sway on the apparent reliability of the specialist seen with it[12]. Studies have demonstrated that the auscultation expertise (i.e., the capacity to make an analysis dependent on what is heard through a stethoscope) has been in decrease for quite a while, with this end goal medicinal teachers are attempting to restore it.

B.Analysis using electronic stethoscope

In this technique the range of respiratory sounds were plotted and doctor outwardly analyze the respiratory sounds variation from the norm. The sounds are recorded using electronic stethoscope and are stored in the computer in .wav format. Then some analysis is applied on these files and a spectogram representation is generated. The lung issue are distinguished by utilizing various qualities of the range. This technique requires the specialists with earlier learning of typical and anomalous range of lung sound which will rely on the experience and aptitude of doctor. However these method are mostly used by students practicing medicine in-order to properly identify the behaviour of different sounds such as normal, wheeze and crackle.

C. Limitations of Existing System

From the literature, we came to know that traditional evaluation using electronic and conventional stethoscope has some problems. Electronic stethoscope are prone to electronic interference from other devices. One requires a prior knowledge of spectrogram representa-tion in-order to identify the abnormalities present in lung sound. This method fails to classify different types abnormal sounds in lungs. While usage of conventional stethoscope for auscultation does not only requires ones knowledge and experience but also the physicians hearing ability which degrades with age.

II. PROBLEM DEFINITION

Lungs sounds has been used as a diagnostic tool from centuries. Whenever any patient complaint about respiratory distress the basic approach is to listen to lung sounds through stethoscope .Thus listening of lung sounds through stethoscope is an important step for identification of various pulmonary diseases. Based on this decision the doctor can decide further treatment and can provide further medication. However sometimes auscultation can results in high monetary cost as auscultation is a subjective method for pulmonary disease identification. As auscultation is subjective method training physician is a difficult task . Thus because of these issues there is a need to develop better tools in order to train physicians and a standard for abnormal lung sound is desperately required. With this advancement in technology the trainee physicians will be able to identify the disease more accurately and can decide the further treatment.

III. PROBLEM DEFINITION

The Proposed system consist of classification of the data-set into four classes: normal, wheeze, crackle and both(crackle and wheeze). Further transforming the data-set(.wav files) to a consistent 22kHz frequency which were varying previously.Then floating point time series array are extracted from each .wav files which helps in splitting up the wav files (data-set) into respiratory cycles with the help of annotation.txt files. Then these breath cycles are fed to MFCC feature module to extract MFCC coefficient and a data-frame with MFCC feature and their respective label is generated. Further fed the MFCC coefficient array into the CNN classifier to distinguish normal,wheeze,crackle, and both(crackle and wheeze).The data-set was splited into 80:20 ratio for training model on the 80 % of the data while testing the model o the remaining 20% of the data. The system block diagram of the proposed system is shown below in the figure Fig.1

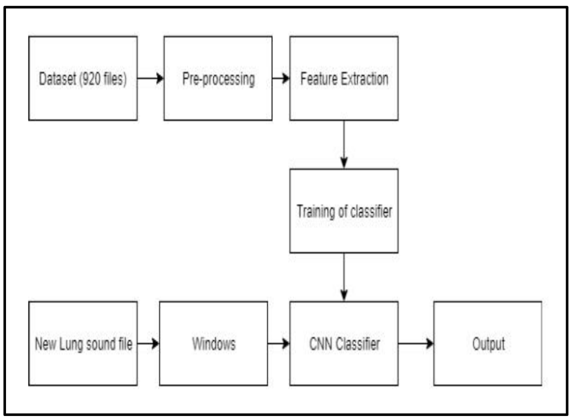

Fig. 1 Architectural View of Proposed system

The Respiratory Sound Database created by the two research teams in Portugal and Greece. It consists of 920 annotated recordings of varying length - 10s to 90s. These recordings were taken from 126 subjects. There are a total recordings consisting of 6898 respiratory cycles - 886 for wheezes, 1864 for crackles and 506 for both wheezes and crackles. The data includes both clean respiratory sounds together with noisy recordings that simulate the real life conditions. The patients span all age groups - children, adults and elderly. This data-set includes, 920 .wav sound files, 920 annotation and .txt files which specify the patient number, recording index, chest location , acquisition mode and recording equipment. This could be illustrated in Table 1 and Figure 2.

TABLE 1
NUMBER OF RESPIRATORY CYCLES

Sound label	Number of cycles
Normal	3642
Crackle	1864
Wheeze	886
Crackle and Wheeze	506

Figure 2(i) represents normal, 2(ii) represents wheeze, which are continuous high pitch with frequency>200 Hz, while duration>250 ms. Wheeze are whistling harmonic sounds caused by vibration of narrow airways as air passes through it. Figure 2(iii) represents crackle, which are discontinuous low pitch with frequency range between 200-2000 Hz, while duration<30 ms. Crackles are rapid rattling/cracking sounds and Figure 2(iv) represents both.

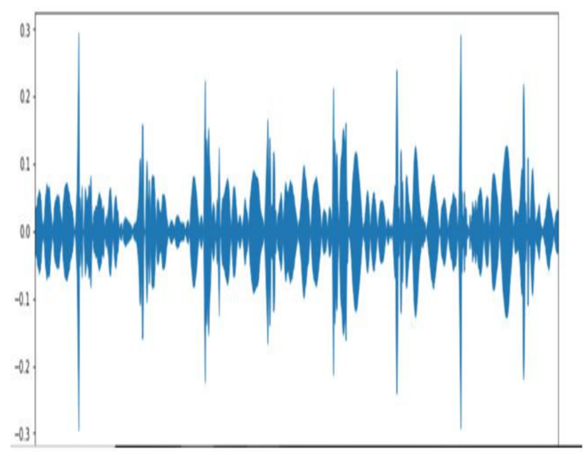

Fig 2 (i) Lung Sound of Normal Person

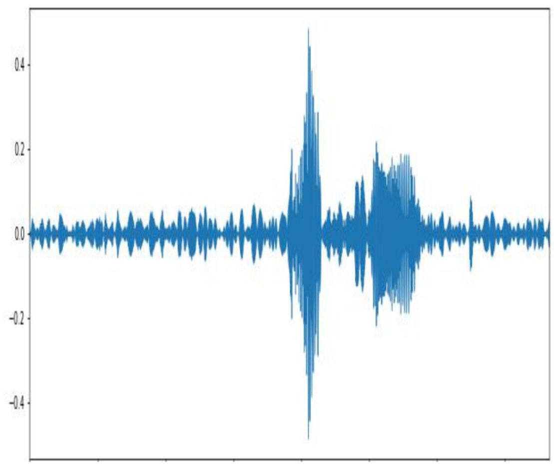

Fig 2 (ii) Lung Sound of Person Suffers From Wheezing

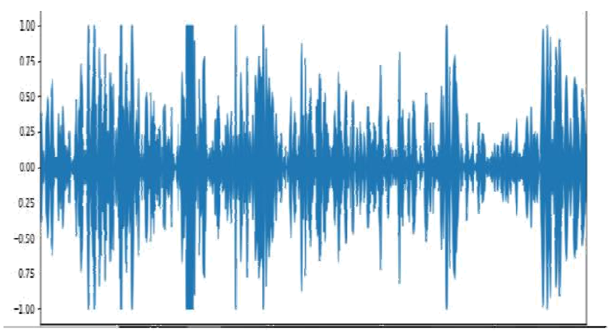

Fig 2 (iii) Lung Sound of Person with Crankle

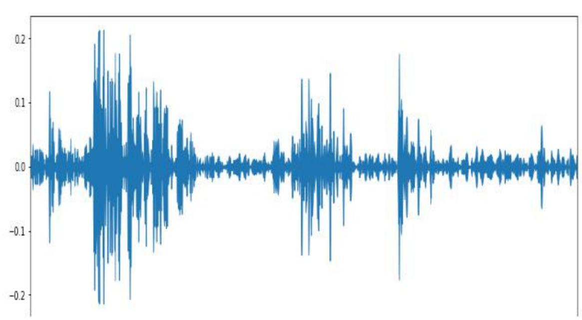

Fig 2 (iv) Lung Sound of Person with Crankle and Wheezing

III. PREPROCESSING

Pre-processing consists of extraction of information from the .wav and text files which if further used for training purpose. Which helps trains data-sets based on different classes. This step also includes down sampling the variable frequencies to a constant frequency of 22kHz followed by extracting time series array for .wav file and segmenting these data into respiratory cycles with the help of annotation .txt files. Further generated a dataframe with two columns, one for the respiratory cycle and second for the respective class label.

A.Implementation environment

1.Jupyter Notebook: Being an open source web application jupyter notebook permits to create and collaborate documents that contains equations, live code, text and visualizations. The areas where jupyter notebook is widely used are machine learning, statistical modeling, numerical simulation, data cleaning and transformation, etc.

2.NumPy: NumPy is a python library for multi-dimensional array and matrix processing, with the support of a substantial clusters of mathematical functions. The application areas of NumPy includes fourier transform, random number capabilities, and linear algebra. Libraries such as Tensor

Flow make use of NumPy in its inner layers for computation of Tensors.

3. Keras: Keras is a python based deep learning framework which is actually high level API of tensorflow. It runs on the top of tensorflow, theaon or CNTK. In keras building model is simple as stacking layers and connecting graphs.

4. Pandas: Pandas is a tool for data processing which help in data analysis. Pandas library is built on numpy package. It provides functions and methods to efficiently manipulate large dataset. It offers data structures and operations for manipulating numerical table and time series, which is panel data.

501. Matplotlib: Matpoltlib is a Python based library for data visualization. It is generally used when the programmer wants to visualization the pattern present in the data. It is an open-source drawing library which supports creation of graphs and plots in two dimensional and three dimensional. The different types of graphs sup-ported by matplotlib are scatterplot, density plot, marginal histogram, etc.

B.Feature extraction

Feature selection is an essential part for the success of identifying an event. Hence, we had evaluated features of the training set using MFCC .MFCC is a feature extraction method which is mainly used in acoustic signal processing. MFCCs are mel frequency cepstral coefficients. It converts the conventional frequency to mel scale by considering human perception of hearing sound for sensitivity at appropriate frequencies. Hence, MFCC scales the frequency in order to match the sound human can hear because human are better in identifying the changes in sound at lower frequencies. For example, the range of audible sound is 20Hz - 20000Hz. Consider a tune at 400Hz, now if the frequency is increased to 500Hz the differences in these frequencies can be identified by human ear. However if the frequencies is on higher side i.e. 1100Hz and another frequency at 1200Hz then human cannot perceive the difference between these as better as previous one, although the differnce is constant i.e. 100Hz. Librosa is a package in python for audio analysis. It creates building block for information retrieval in various audio signals. It also includes low level feature extraction such as mfcc, mel spectogram and tuning estimation. The values returned by librosa.feature. mfcc is used for feature extraction. The extracted features were given to cnn classifier to distinguish between normal, wheeze, crackle, wheeze and crackle.

C.Data Classification

Classification is a procedure to sort out data into an ideal and particular number of classes where we can allot label to each classes [11]. CNN is a artificial neural network which can detect pattern and make sense of them. It contains convolutional layer. CNN receives an input then transform the input in some way then output the transformed input onto next layer. This transformation is a convolutional operation. With each convolutional layer we need to specify the number of filters the layer should have. These filters actually detect the pattern.

Generally the activation function in CNN is a relu layer which is followed by some extra convolutions such as max pooling, fully connected and normalization layer.The last layer regularly involves back propagation so that end product is accurately weighed. The keras library is used to build CNN. The model type which we have used is sequential. It builds a model layer by layer. add() function is used to add layers in the model. The initial layers are CONv2D layer that deals with input and activation function for the layer are relu and softmax. The flatten layer serves as connection between convolution and dense layer. Dense layer is a standard layer type used as output layer. Hence after building the model the steps followed are compiling the model then training the model and using model to make predictions.

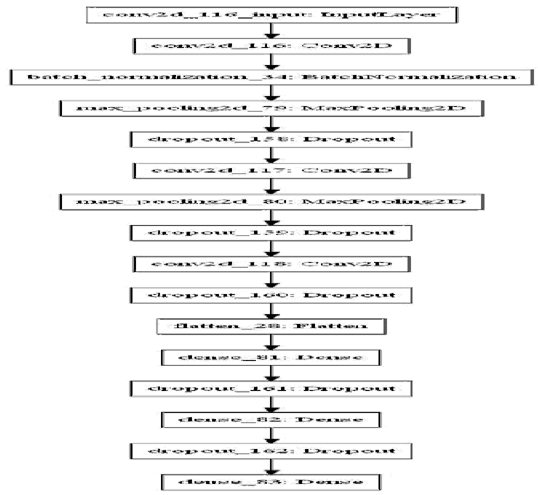

Figure 3: CNN layer diagram

IV. RESULT

Performance measures derived from confusion matrix are shown in tables. As data-set are divided into 80:20 ratio we had 1380 files for testing data and got following result shown in confusion matrix tested on 721 normal, 373 crackle, 174 wheeze, 112 wheeze and crackle out of which 630, 195, 24 and 16 are correctly predicted by model as normal, crackle ,wheeze ,both (crackle and wheeze) respectively. Hence giving accuracy of 85%, 64%, 40%, 36% for respective classes in recall or True positive rate column.Precision is calculated as number of correct positive prediction/positive prediction. Support shows number of data value that are present for evaluation. And F1 score is a mean harmonic of precision and recall. In some cases when imbalanced data is present f1 score help to choose best model suitable for prediction of all classes present in data-set. Similarly, Table of accuracy and confusion matrix are obtained for Random Forest Algorithm representing result for RF model. Hence the overall accuracy obtained by CNN model is 69% and that of Random Forest is 60%.

TABLE II
CNN CLASSIFICATION REPORT

	precision	recall	f1-score	support
normal	0.74	0.85	0.79	721
crackles	0.70	0.64	0.66	373
wheeze	0.54	0.40	0.46	174
wheeze and crackle	0.48	0.36	0.41	112

TABLE III
CNN CONFUSION MATRIX

	normal	crackle	wheeze	wheeze and crackle
normal	610	70	33	8
crackles	108	237	7	21
wheeze	79	10	70	15
wheeze and crackle	30	23	19	40

TABLE IV
RANDOM FOREST CLASSIFICATION REPORT

	precision	recall	f1-score	support
normal	0.62	0.87	0.73	714
crackles	0.56	0.45	0.50	380
wheeze	0.49	0.15	0.23	169
wheeze and crackle	0.74	0.12	0.21	117

TABLE V
RANDOM FOREST CONFUSION MATRIX

	normal	crackle	wheeze	wheeze and crackle
normal	623	78	11	2
crackles	200	171	7	2
wheeze	117	26	25	1
wheeze and crackle	63	32	8	14

V. CONCLUSION

The division of the data-set into different arrays containing characteristics of wheeze, crackle and normal is accomplished in-order to train our data-set. CNN method was used for classification for sounds into normal, wheeze and crackle and MFCC for feature extraction. MFCC based feature extraction method has been implemented and the model using CNN classifier has been built with keras library in-order to classify the lung sounds. Keras contains a number of activation and optimizer those can be used very easily in the model. This method provides a novel method to predict lung disease easily and accurately.

REFERENCES

[1]Palaniappan, R., Sundaraj, K., & Lam, C. K. "Reliable system for respiratory pathology classification from breath sound signals",International Conference on System Reliability and Science (ICSRS), 2016. doi:10.1109/icsrs.2016.7815855.

[2] Chamberlain, D., Kodgule, R., Ganelin, D., Miglani, V., and Fletcher, R. R. "Application of semi-supervised deep learning to lung sound analysis",38th Annual Inter- national Conference of the IEEE Engineering in Medicine and Biology Society (EMBC),2016. doi:10.1109/embc.2016.7590823.

[3] Aykanat, Murat, Özkan Kılıç, Bahar Kurt, and Sevgi Saryal. "Classification of lung sounds using convolutional neural networks."*EURASIP Journal on Image and Video Processing* , vol.65, no. 1,2017.

[4] Andrès, E et al. "Respiratory sound analysis in the era of evidence-based medicine and the world of medicine 2.0." *Journal of medicine and life* vol. 11,2,pp. 89-106,2018.

[5] Aras, S., Gangal, A.," Comparison of different features derived from mel frequency cepstrum coefficients for classification of single channel lung sounds", 40th International Conference on Telecommunications and Signal Processing (TSP), 2017. doi:10.1109/tsp.2017.8076002

[6]http://practicalcryptography.com/miscellaneous/machine-learning/guide-mel-frequency-cepstral-coefficients-mfccs/

[7]M. Bahoura, "Pattern recognition methods applied to respiratory sounds classification into normal and wheeze classes," Computers in Biology and Medicine, vol. 39, 2009, pp. 824-843.

[8]Y LeCun, K Kavukcuoglu, C Farabet, Convolutional networks and applications in vision. Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, IEEE, pp. 253–256, 2010.

[9] Pramono, R. X. A., Imtiaz, S.A., Rodriguez-Villegas,E." Evaluation offeatures for classification of wheezes and normal respiratory sounds. PLOS ONE, vol.14(3),e0213659.2019. doi:10.1371/journal.pone.0213659.

[10] BDCN Prasadl, PESNK Prasad, Y Sagar, An approach to develop expert systems in medical diagnosis using machine learning algorithms (asthma) and a performance study. IJSC vol.2(1),pp. 26–33,2011.

[11]Y LeCun, K Kavukcuoglu, C Farabet, Convolutional networks and applic ations in vision. Circuits and Systems (ISCAS), Proceedings of 2010 IEEE International Symposium on, IEEE, pp. 253–256, 2010

[12]Jiwa, Moyez; Millett, Stephan; Meng, Xingqiong; Hewitt, Vivien M. "Impact of the Presence of Medical Equipment in Images on Viewers' Perceptions of the Trustworthiness of an Individual On-Screen". Journal of Medical Internet Research,2012.