RESPIRATORY SOUND ANALYSIS FOR EARLY DETECTION OF COPD(CHORONIC OBSTRUCTIVE PULMONARY DISEASE) USING DEEP LEARNING

|  |  |  |
| --- | --- | --- |
| Pathan Shohaib Khan  Department of Electronics and Communication Engineering,  Madanapalle Institute of Technology  & Science,  Madanapalle , Andhra Pradesh, India,  19691A04H9@mits.ac.in | Aavula Siseendri  Department of Electronics and Communication Engineering,  Madanapalle Institute of Technology  & Science  Madanapalle , Andhra Pradesh, India,  19691A04I2@mits.ac.in,  Dr.K.Sathesh Ph.D. | Somisetty Siva Dinesh  Department of Electronics and Communication Engineering,  Madanapalle Institute of Technology  & Science,  Madanapalle , Andhra Pradesh, India, 19691A04I3@mits.ac.in |

Department of Electronics and Communication Engineering,

Madanapalle Institute of Technology & Science,

Madanapalle , Andhra Pradesh, India, drsatheshk@mits.ac.in

***Abstract:***

**In recent years, technologies like machine learning and deep learning have made a significant contribution to the provision of assistive solutions to the difficulties encountered in the medical field** **Additionally, they increase prediction accuracy for quick and early illness identification by using medical imaging and auditory analysis. Medical professionals are appreciative of such technology support because it enables them to cope with an increasing number of patients. This is due to the fact that there is a lack of qualified human resources.**

**In recent years, technologies like machine learning and deep learning have played a crucial role in providing assistive solutions to the difficulties faced by a medical field. In this project we have developed a CNN model which can able to early detect the COPD disease. Using medical imaging and audio analysis, they also improve predictive accuracy for prompt and early disease detection clinical professionals are inviting such innovation help as it gives a assistance to them in adapting to additional patients.**

**Tobacco use, a genetic disorder (alpha-1-antitrypsin deficiency), pollution in the air, and other factors are known to contribute to COPD. The field of early or timely COPD detection is still in its infancy and has not yet achieved 100% accuracy.**

***Keywords: envelop correlation coefficient (ECC), multiple-inputmultiple- output (MIMO), DGS, 5G antenna, millimetre wave.***

I INTRODUCTION

1. Chronic Obstructive Pulmonary Disease:

Men and women worldwide are affected by the common, preventable, and treatable chronic lung disease known as Chronic Obstructive Pulmonary Disease (COPD). Airflow restriction in the lungs is caused by abnormalities in the lungs' small airways. The narrowing of the airways is caused by a number of processes. Parts of the lung may be destroyed, the airways may be blocked by mucus, and the airway lining may become inflamed and swollen.

Other terms for COPD include chronic bronchitis and emphysema. Emphysema often indicates the devastation of the little air sacs at the end of the airways in the lungs. The symptoms of chronic bronchitis, an airway inflammation, include phlegm production and a persistent cough. can COPD and asthma, which can cause coughing, wheezing, and breathing difficulties, can affect a person.

Chronic obstructive pulmonary disease (COPD), the third biggest cause of mortality globally in 2019, claimed 3.23 million lives. Nearly 90% of COPD fatalities in adults under the age of 70 occur in low- and middle-income (LMIC) nations. Despite the fact that there is currently no known cure for COPD, early diagnosis and treatment are essential to delay the onset of symptoms and reduce the incidence of flare-ups.Graphical user interface

Description automatically generated with medium confidence

Fig. 1.1: COPD and Symptoms

Figure 1.1 shows the various symptoms which are related to the COPD disease, From middle age on, common COPD symptoms appear. 2 Symptoms include breathing difficulties or chronic coughs frequently accompanied by phlegm, tiredness (Figure 1.1)

1. Deep Learning:

Deep learning is a form of mimic of the human brain, much like artificial neural networks are. A branch of machine learning called "deep learning" is solely dependent on neural networks. The concept of deep learning is not new. It has been around for some time. Because the initiative now has access to more data and processing power than it had in the past, it is becoming popular. Computing power is growing rapidly, similar to the last 20 years. There are several different deep neural network topologies, including Deep Neural Networks, It is a neural network that incorporates a particular amount of complexity, therefore there are many hidden layers between the input and output layers. They are exceptional. Diagram, text

Description automatically generated

Fig. 1.2: Deep Learning Working Steps

Figure 1.2 depicts that to find the best answer, must first understand the problem at hand.in this projectmust also determine whether deep learning is a viable option. The second step is to determine the pertinent facts that should be prepared in accordance with the problem at hand. Third, pick the right Deep Learning algorithm. Fourth, while training the dataset, an algorithm should be utilized. Fifth, the dataset should be subjected to final testing. All the steps which has shown in the figure 1.2.1 should be followed to get the best solution to a problem using deep learning techniques.

II DESIGN

A system that uses sounds from the lungs to identify COPD is whatin this projecthave proposed. In the event of a heart attack, asthma, COPD, etc., the sound that is produced by the body's internal organs is very different. Robotized distinguishing such sounds to order if a individual is vulnerable to COPD is a too efficient, self-disturbing strategy for both the patient and the specialist. The specialists can involve the framework for affirmed discovery of COPD.

Convolutional neural networks or multi-layered feed-forward neural networks are constructed by stacking a number of invisible layers on top of one another in a certain order. The sequential design enables CNN to learn hierarchical properties. In a CNN, activation layers are frequently followed by convolutional layers, with some of them coming after grouping layers and hidden layers. The pre-processing required for a ConvNet, which is comparable to the corresponding arrangement of neurons in the human brain, was inspired by the organisation of the Visual Cortex.

1. Data collection:

The process of obtaining information from all pertinent sources in order to address a statistical research issue is known as data collection. Reviewing the solution to the issue is beneficial. Thanks to the data collecting techniques, a person can come up with an answer to the relevant query. Most organisations gather data in order to deduce future trends and probability. After the data has been gathered, it is required to organise it.Graphical user interface

Description automatically generated with low confidence Fig, 2.1. Patient Record data

from figure 2.1 shows the extracted records of all the subjected patients whether they are affected by COPD or not. Among 126 subjects 1,864 are crackles, 886 are wheezes, and 506 are both crackles and wheezes. Experts in the respiratory field annotated the cycles as consisting of crackles, wheezes, a combination of the two, or no accidental respiratory sounds. The chest locations from which the recordings were obtained are also provided. The recordings were made with a variety of different pieces of equipment and lasted anywhere from 10 seconds to 90 seconds. Commotion levels in some breath cycles is high, which recreate genuine circumstances.

Chart, bar chart

Description automatically generated Fig. 2.2. Graphical representation of patient data

From Figure 2.2 is plotted using the matplotlib library which shows the graphical representation of the patient diagnosis data. "Data" is the primary source of the data collection methods. Primary data and secondary data are the two categories of data. The primary reason that data collection is crucial to any research or business procedure is that it assists in determining numerous crucial aspects of the company, particularly performance. As a result, the procedure for gathering data is crucial to each and every stream. Unstructured data and numerous irregularities were present in the dataset.in this projectused the Python library Librosa to trim and padding the audio files to a length of 20 seconds in order to normalize the data.

Graphical user interface, text, application

Description automatically generated

Fig 2.3. Processing of audio signals for feature extraction

From Figure 2.3 shows the processing of the audio signals, every audio signal is processed for and make ready for the feature extraction.

A picture containing table

Description automatically generated

Fig 2.4 All features that are extracted

In the Figure 2.4: all the features which need to be extracted from the audio series of the audio signals has been shown in the above figure, every feature is stored in the form of keyvalue pairs (Object) It depicts a sound's short-term power spectrum. A powerful set of features for analyzing music whose pitches can be categorized are chroma-based features, also referred to as "pitch class profiles" like the onesin this projectmentioned earlier. Chroma is an excellent feature for our user case due to the fact that the pitch of the respiratory sounds also varies quite significantly. Due to their resistance to dynamics, timbre, and articulation, CENS features are frequently utilized in audio matching and retrieval tasks. In order to maintain consistency across the features,in this projectassigned a value of 40 to each feature, similar to n\_mfcc bins in chroma features and n\_mfcc bins in mfcc.

III IMPLEMENTATION AND RESULTS

The number of nodes in each layer is determined by the filter parameter. The individual layer sizes increase from 16, 32, 64, to 128 while the kernel size option specifies the kernel window size. This results in a 2 2 filter matrix. As seen in Table 3, the first layer will be given an input shape. The input shape's values would be (40, 862, 1), which would stand for the number of MFCCs at 40, the total number of frames, including padding, at 862, and the mono audio structure at 1. The information is sent to the MaxPooling2D layer of (Pooling Size: 2) via the Convolution2D layer, which has 16 filters and a kernel size of: 2. After then, to avoid overfitting the data, this project established a 20% dropout rate. The data are sent to a MaxPooling2D layer (Pooling size: 2) after the dropout and again to a Convolution layer with 32 filters and a kernel size of 2 (relu). Once more, there is a MaxPooling2D layer with a pooling size of: 2) and a 20% dropout rate, as well as a Convolution2D layer with 64 filters and a kernel size of: 2, relu). The MaxPooling2D layer, a Convolution2D layer with 32 filters and a kernel size of 2, relu, receives the data next. 2), followed by dropouts (20%). ReLu is the name of the activation function that is utilised to produce the activation map from the convolutional layer. The remaining information is then transmitted to the GlobalAveragePooling2D layer (127), where it is flattened and split into COPD and non-COPD outputs. It has been noticed that audio inputs often include more noise than visual inputs do. In this project, the Adam optimisation method was employed for optimising.

1. Implementation of key functions :

The audio sample that has been provided is broken up into two sets of data: training data and testing data. A number of procedures, including Normalization, Feature Extraction, and Augmentation, are carried out on the training data prior to the training of the model. After that, the model stores the modified features of the training data. The testing information, moreover, is gone through a comparable series of cycles. It makes use of the trained model that was developed for CNN's COPD classification and prediction. A spectrogram is helpful in portraying the sign concerning time, recurrence, and extent. A spectrogram is a two-dimensional representation of audio, with time and magnitude serving as the two dimensions and colors serving as the third. A distinct spectrogram is also produced when respiratory diseases are present, as they are musical anomalies that can be identified using appropriate auscultation methods. Mel-Spectrograms, MFCC, and Chromogram are the three main categories of our extracted features. The MFCC accurately depicts the short-time power spectrum envelope, which shows the shape of the vocal tract. A powerful way to represent an audio sample are chroma features, which are also referred to as pitch-class features.

1. Output Screens and Result Analysis:

The output of the project such as the data acquisition, preprocessing of data, feature extraction and feature extraction process has been shown below. The patient data and the data processing which has shown in the figure 1 and figure 2 and the audio data processing samples as shown in the above figures are used fo the feature extraction of the audio signals as shown in the below figures. The object contains the audio\_series as the first element followed by the file\_name of the audio signal and all the fearures associated with the audio signal as the mfcc, melspectrogram, chroma\_stft, chroma\_cqt, chroma\_cens, augumented, patient\_id, rec\_instrument, all these features are stored in the form of the object in the python as an key value pairs, because it is very useful to extract the desired data efficiently. Now this extracted data from the audio signals are ready for the detailed feature extraction, all the details of the feaure extraction has been carried out The extracted and preprocesssed data is useful for the feature extractiona as shown in the figure 3 where proessing has been carried out

Table

Description automatically generated

Fig 3.1. All extracted features from the audio signals

As shown in the Figure 3.1 all the features are successfully extracted,in this projecthave used the LIbrosa library form the python which is very useful to extract the features form any audio signals and the features are extracted and by using the matplot lib library all the features are shown as in the above figure.

Text

Description automatically generated

Fig 3.2. Pretraining of the model

The Figure 3.2 shows us the pretraining of the model by using our dataset for different features of the audio signals individually. As for every feature the accuracy to find out the chronic obstructive pulmonary disease (COPD) is calculated as all the features has very good accurec to find out the COPD in the patient among all the features mfcc spectrogram has the good accuracy to know the COPD disease mfcc is the main feature which can be considered to find the disease present in the patient easily. The Delta-Delta MFCC, which enhances speaker verification, is one of the most recent MFCC implementations.A picture containing text

Description automatically generated

Fig 3.3. Training of features

The Figure 3.3 shows the training of the features mfcc, melspectrogram, and chroma stft.

A screenshot of a computer

Description automatically generated with medium confidence

Fig 3.4 Accuracy of every feature

The above figure shows the accuracy related to every feature of the audio signal among all the features mfcc has good accuracy.

Chart, bar chart

Description automatically generated

Fig 3.5. Loss and Accuracy.

The figure 3.5 shows the loss vs Accuracy representation, andin this projectcan observe that mfcc has less loss as compared to the other features of the audio signals During testing of data for every feature deep analysis has been performed like for every feature of the audio signal Confusion matrix (post-Augmentation) has been designed, Precision, Recall, f1-score and support form the feature has been successfully extracted The below are the detailed analysis of all the features of the audio signals A confusion matrix, also called an error matrix, is a process that helps evaluate and predict the effectiveness of a classification model. Confusion matrices allow you to see different errors that can occur when making predictions. ROC stands for Receiver Operating Characteristic. A ROC curve is a useful visual tool for analyzing two classification models. ROC curve emerges from signal detection theory.

Graphical user interface, application

Description automatically generated

Fig 3.6. Detailed analysis of mfcc feature

The figure 3.6 shows the confusion matrix and ROC curve for COPD for the mfcc feature. Mel Frequency Cepstral Co-efficient (MFCC) is an inner audio illustration layout which is straightforward to paintings on. This is much like JPG layout for pictures.

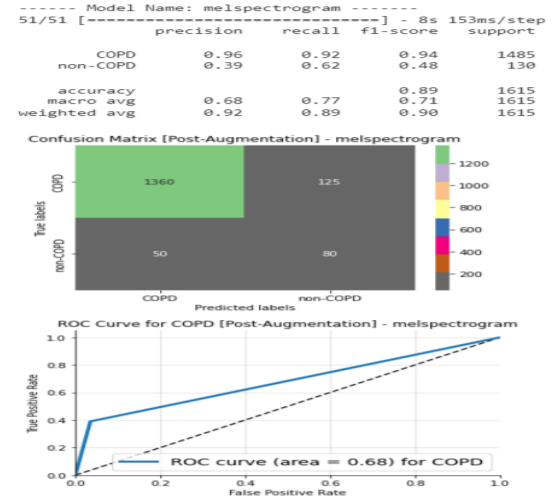


Fig 3.7. Detailed analysis of Mel spectrogram

The figure 3.7 shows the confusion matrix and ROC curve for COPD for the Mel spectrogram feature. An item of kind Mel Spectrogram represents an acoustic time-frequency illustration of a sound.

Graphical user interface, application

Description automatically generated

**Figure 3.8:** Detailed analysis of Chroma stft feature

The Figure 3.8 shows the confusion matrix and ROC curve for COPD for the chroma stft feature, Chroma STFT The Chroma cost of an audio essentially constitutes the depth of the twelve specific pitch lessons which might be used to observe music. They may be hired withinside the differentiation of the pitch elegance profiles among audio signals.

Graphical user interface, text, application, email

Description automatically generated

Fig 3.9. Individual Features Accuracy

Figure 3.9 shows that mfcc feature has 97% accuracy, Mel spectrogram and chroma stft has 90.3% and 93.5% accuracy.

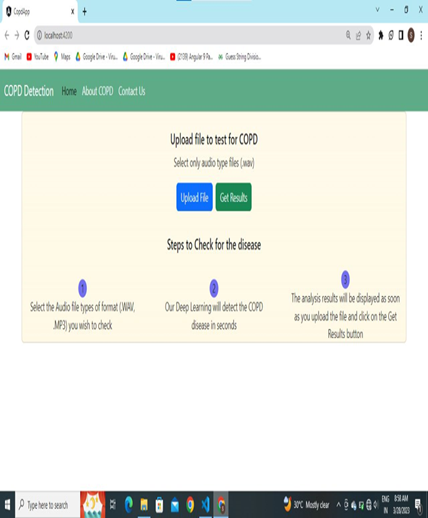
Text

Description automatically generated

Fig 3.10. Feature Characteristics

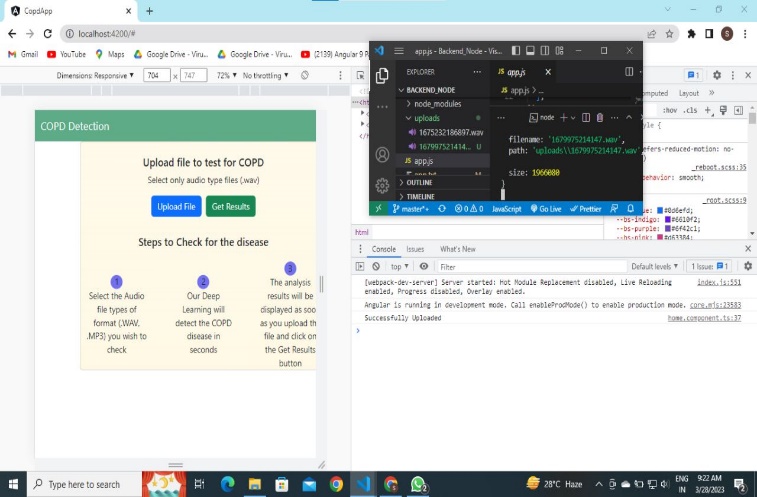
Figure 3.10 in this projecthave all the feature characteristics like sensitivity, specificity, auc, icbhi score of the individual feature of the audio signal, andin this projectcan see that the specificity of the feature has been increased to 99% so that the model can be more reliable to get the output.

Thus, the implementation of key functions of the project has been performed in this section successfully like preprocessing of audio signals, feature extraction and the feature accuracy of the individual feature has been carried out, the various method of implementation which are used to get the output and analysis has been discussed in this module. Various libraries of python which are used in the project for the purpose of the implementation of key functions has been discussed like Librosa, matplotlib etc. and also about various features information has been provided successfully.



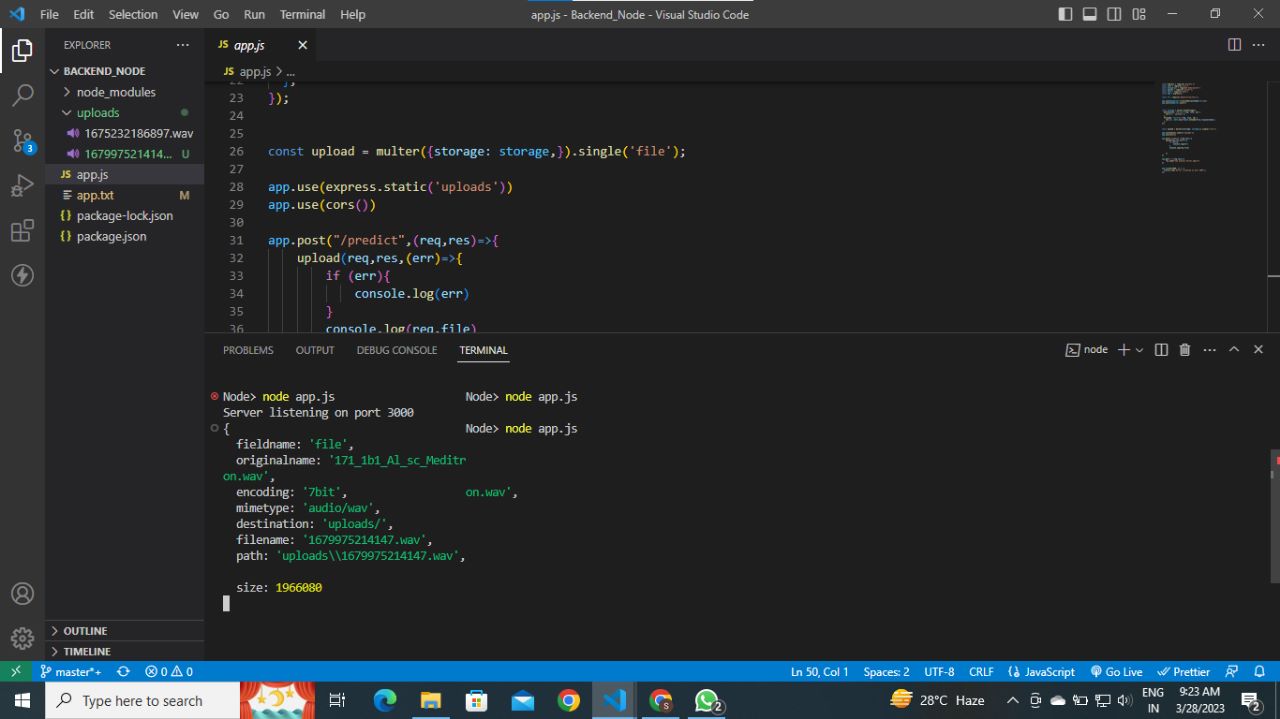
**Figure 3.11:** Frontend view of the web Application

Figure 3.11 On the frontend view of the application users will be able to upload an audio file using a file upload component. Once the user has uploaded the file, the application will process the audio and analyze it for signs of COPD (Chronic Obstructive Pulmonary Disease). The analysis will be performed by a backend system that uses advanced algorithms and machine learning techniques to identify COPD markers in the audio file. The results of the analysis will then be displayed on the frontend view of the application, providing the user with an accurate assessment of their condition. It's important to ensure that the results displayed on the frontend are clear and easy to understand. This could be achieved by presenting the results in a graphical format or providing a summary that highlights the key findings. In addition, it's crucial to ensure that the application is designed with user privacy and security in mind. The audio files uploaded by users should be stored securely and all data transmissions should be encrypted to prevent unauthorized access. Finally, it's essential to ensure that the content in the application is original and free from plagiarism. This can be achieved by using reliable sources for information and ensuring that all content is written from scratch



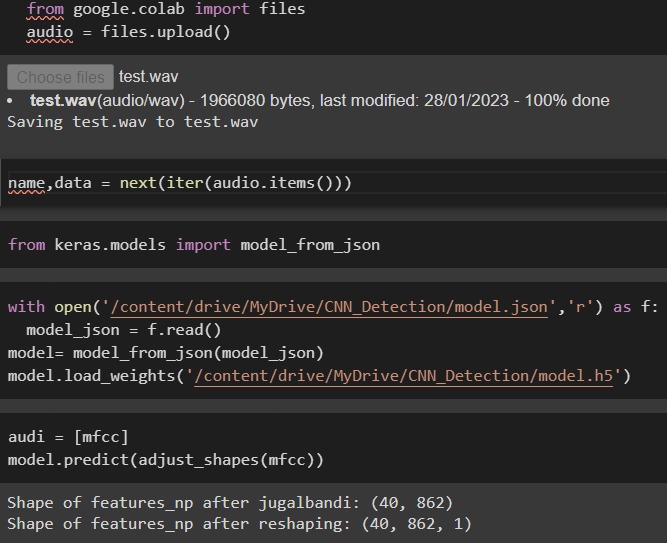
**Figure 3.12:** Frontend view user uploading audio file

Figure 3.12 When a user uploads an audio file in the frontend view of the application, a POST request is sent to the backend server built with Node.js and Express. This POST request contains the audio file in its binary form, which is then processed by the backend server. Upon receiving the audio file, the backend server saves the file in a designated folder, typically named "uploads" or a similar name, to ensure that the file can be accessed and analyzed at a later time. The server may also perform additional processing, such as converting the audio file to a specific format or extracting relevant data from the file. It is important to ensure that the backend server is designed to handle the file uploads securely and efficiently. The server should be configured to handle large file uploads and validate the input data to prevent malicious file uploads or other security risks. Once the audio file is stored in the backend, the server can perform the necessary analysis to identify any COPD markers in the audio file. The analysis may involve using complex algorithms and machine learning techniques to process the audio data and extract relevant features.This can be achieved by using encryption to protect the data during transmission and storage, implementing secure access controls to restrict access to the data, and following best practices for data management and storage.



**Figure 3.13**: Backend view

Figure 3.13 The storing of the audio file and all its content at the backend.



**Figure 3.14:** Stores the Model structure in model.json file

Figure 3.14 After training the machine learning model to identify COPD markers in audio data, the model structure is saved in a JSON file named model.json. This file contains the architecture of the model, including the number of layers, the type of layers used, and their parameters. In addition to the model structure, the trained model data is saved in a separate file named model.h5. This file contains the weights and biases learned by the model during the training process, which are used to make predictions on new audio data. When the user submits an audio file for analysis, the backend server loads the model structure from the model.json file and the model data from the model.h5 file. Using these files, the server can then process the audio data and make a prediction on whether or not the audio data contains COPD markers. It is important to ensure that the trained model and its associated files are stored securely and that appropriate access controls are in place to prevent unauthorized access. It is also important to ensure that the model files are regularly backed up to prevent data loss.

Graphical user interface, application

Description automatically generated

**Figure 3.15:** output of the patient

Figure 3.15 shows the output after analyzing the given data of patient.

IV. CONCLUSION:

proposed a straightforward and time-consuming CNN-based depth. a breathing sound support model for learning that can help medical practitioners recognise COPD. The Librosa machine was utilised in the trials. Use learning library tools like MFCC, Mel-Spectrogram, Chroma, Chroma (ConstantQ), Chroma CENS (Librosa), and ICBHI's 2017 Dataset Chroma Feature Overview to do in-depth analysis of breath sounds. (Accessed May 20, 2020. Base The studies we conducted for this project revealed that'mfcc' is more accurate. In contrast to all other elements of the Librosa machine learning package, COPD detection. Future iterations of this research may expand its capabilities to let medical professionals identify a variety of additional illnesses, such as the possibility of a heart attack or cardiac arrest based on the sound of the heartbeat, the diagnosis of asthma based on the sounds of the lungs, etc. Make present illness severity detection techniques more accurate.Applications are also accepted for this project several methods of data augmentation to enhance performance. Identifying COPD has become simpler because to our system's ability to perform (Radogna et al., 2019) method and respiratory monitoring system. The technology also offers greater privacy and is more attack-resistant. The most crucial thing is saving.

REFERENCES

1. Ahmed J, Vesal S, Durlak F, Kaergel R, Ravikumar N, Rémy-Jardin M, Maier A, Tolxdorff T, Deserno T, Handels H, Maier A. 2020. COPD classification in CT images using a 3D convolutional neural network. In: Maier-Hein K, Palm C, eds. Bildverarbeitung für Die Medizin 2020—Informatik Aktuell. Wiesbaden: Springer Vieweg, 39–45.

1. Altan G, Kutlu Y, Allahwardi N. 2019. Deep learning on computerized analysis of chronic obstructive pulmonary disease. IEEE Journal of Biomedical and Health Informatics 24(5):1344–1350 DOI 10.1109/JBHI.2019.2931395.

1. Altan G, Kutlu Y, Pekmezci AÖ, Nural S. 2018. Deep learning with 3D-second order difference plot on respiratory sounds. Biomedical Signal Processing and Control 45:58–69 DOI 10.1016/j.bspc.2018.05.014.
2. Amaral JL, Lopes AJ, Jansen JM, Faria AC, Melo PL. 2012. Machine learning algorithms and forced oscillation measurements applied to the automatic identification of chronic obstructive pulmonary disease. Computer Methods and Programs in Biomedicine 105(3):183–193 DOI 10.1016/j.cmpb.2011.09.009.

1. Aykanat M, Kılıç Ö, Kurt B, Saryal S. 2017. Classification of lung sounds using convolutional neural networks. EURASIP Journal on Image and Video Processing 2017(1):65 DOI 10.1186/s13640-017-0213-2.

1. Badnjevic A, Gurbeta L, Custovic E. 2018. An expert diagnostic system to automatically identify asthma and chronic obstructive pulmonary disease in clinical settings. Scientific Reports 8(1):11645 DOI 10.1038/s41598-018-30116-2.

1. Bai Z, Li Y, Woźniak M, Zhou M, Li D. 2020. DecomVQANet: decomposing visual question answering deep network via tensor decomposition and regression. Pattern Recognition 110:107538 DOI 10.1016/j.patcog.2020.107538.

1. Bellos CC, Papadopoulos A, Rosso R, Fotiadis DI. 2014. Identification of COPD patients’ health status using an intelligent system in the CHRONIOUS wearable platform. IEEE Journal of Biomedical and Health Informatics 18(3):731–738 DOI 10.1109/JBHI.2013.2293172.

1. Chamberlain D, Kodgule R, Ganelin D, Miglani V, Fletcher RR. 2016. Application of semi-supervised deep learning to lung sound analysis. In: 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Piscataway: IEEE, 804–807.

1. Du R, Qi S, Feng J, Xia S, Kang Y, Qian W, Yao YD. 2020. Identification of COPD from multi-view snapshots of 3D lung airway tree via deep CNN. IEEE Access 8:38907– 38919 DOI 10.1109/ACCESS.2020.2974617.

1. Er O, Temurtas F. 2008. A study on chronic obstructive pulmonary disease diagnosis using multilayer neural networks. Journal of Medical Systems 32(5):429–432 DOI 10.1007/s10916-008-9148-6.

1. Fernandez-Granero MA, Sanchez-Morillo D, Leon-Jimenez A. 2018. An artificial intelligence approach to early predict symptom-based exacerbations of COPD. Biotechnology & Biotechnological Equipment 32(3):778–784 DOI 10.1080/13102818.2018.1437568.

1. ICBHI. 2017. ICBHI challenge. Available at https://bhichallenge.med.auth.gr/ICBHI\_2017\_ Challenge (accessed 21 May 2020).
2. Khatri KL, Tamil LS. 2018. Early detection of peak demand days of chronic respiratory diseases emergency department visits using artificial neural networks. IEEE Journal of Biomedical and Health Informatics 22(1):285–290 DOI 10.1109/JBHI.2017.2698418.
3. Kingma DP, Ba J. 2017. Adam: a method for stochastic optimization. arXiv. Available at <http://arxiv.org/abs/1412.6980>.
4. Librosa.2020. Feature extraction—Librosa 0.8.0 documentation. Available at https://librosa.org/ doc/latest/feature.html (accessed 20 May 2020).
5. Librosa. 2019. Librosa 0.8.0 documentation. Available at https://librosa.org/doc/main/generated/ librosa.feature.chroma\_stft.html (accessed 21 May 2020).