Computer Science, Engineering and Technology
Vol: 2(1), March 2024
REST Publisher; ISSN: 2583-9179 (Online)
Website:https://restpublisher.com/journals/cset/
DOI: https://doi.org/10.46632/cset/2/1/1

Classifying Abnormalities In Heart Beat Sound

***Y. Befina, A. Chandrasekar**
St.Joseph's College of Engineering Chennai, Tamil Nadu, *India.*
*Corresponding author Email: befina20@gmail.com

Abstract: *Heartbeat sounds play a major role in the detection of various diseases such as heart disease, hyperthyroidism, and high blood pressure. in an early stage. In the proposed method various abnormal and healthy heartbeat audio signals are given as input and the features are extracted using MFCC (Mel-frequency cepstral coefficients). Then, a Deep Learning Approach is applied in which the MFCC audio signals are sent to the CNN (Convolutional Neural Network) + LSTM (Long Short-Term Memory) model to extract and classify the abnormal heartbeat sounds with its associated disease. Thus, a system for classifying abnormal heartbeat sounds is proposed by leading a way to identify and diagnose diseases that are identified using heartbeat sounds in an early stage.*
Keywords: *Heartbeat audio signals, Deep Learning, Mel-frequency cepstral coefficients, Convolutional Neural Network, Long Short-Term Memory*

1. Introduction

During the cardiac cycle the cardiac valves open and close producing heartbeat sounds. Heartbeat sounds play a major role in the detection of various diseases that are identified using heartbeat sounds in an early stage. Evaluation of the heart sounds produced by the beating heart and the blood flow that results from it is a useful diagnostic technique for various diseases such as heart disease, hyperthyroidism, and high blood pressure. There are a few types of heartbeat sounds such as Normal, Murmur, Extrasystole, and Extra heart sounds. Heart sounds that are normal and healthy fall into the normal group. An abnormal heart murmur indicates complications such as heart failure. An extra heart sound in some situations is an important sign of high blood pressure. Extrasystole can happen normally in an adult and can be very common in children. In some situations, extrasystoles can be dangerous as they can indicate hyperthyroidism. With reference to the papers for the classification of heartbeat sounds it has been found that deep learning techniques performed better than machine learning techniques [1]. The existing papers performed the classification using single neural networks providing maximum and minimum accuracies based on the neural network. This motivated to consider two neural networks combined to produce maximum accuracy for classifying the abnormalities in the heartbeat sound with its associated disease. The neural network used in the proposed method is CNN+LSTM. The pre-processing step in the proposed method is carried out with MFCC (Mel-Frequency Cepstral Coefficient) in which the power spectrum of the input audio signal is identified. Data augmentation of the extracted features are further done by adding noise, shifting, stretching time and pitch to it. The pre-processed data is then used to implement in the neural network. CNN's primary benefit is its ability to learn from unprocessed data, eliminating the need for manual feature engineering or pre-processing. Additionally, lowering the dimensionality and complexity of the input data improves the speed and efficiency of inference and training. By finding patterns and retaining pertinent information, LSTMs effectively increase performance. As a result, the CNN may concentrate on extracting spatial information from the LSTM network's output for use in classification and other applications. In section I, it contains the overview of the paper. Section II contains the methodology of the paper. Section III contains the result and discussion of the paper. Section IV contains the conclusion and Section V contains the future development of the paper.

2. Methodology

The system architecture of the proposed method as shown in Fig.1 explains that the input audio signal data is taken as wav format from kaggle sponsored by Pascal, the dataset contains both normal and abnormal heart beat sounds. Features of the input audio signals are extracted using MFCC. The extracted MFCC are pre-processed by adding noise, shifting, applying time stretch and pitch shifting and then they are passed to two deep learning neural network for extraction and classifying the heartbeat sound with its associated diseases. The deep learning neural networks used here is CNN+LSTM where CNN plays a role of extraction and LSTM plays the role of classification. The model is then evaluated and passed as the final output. The final output from the model is the classification system of abnormal heartbeat sounds.

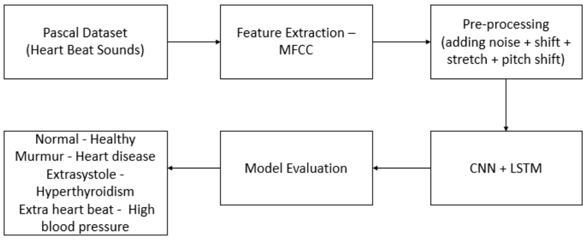


FIGURE 1. System Architecture

Data Understanding: The data is collected from kaggle in which a challenge was sponsored by Pascal, where the heartbeat sounds are collected from general public and hospitals using iStethoscope and digital stethoscope respectively. The aim is to classify the heartbeat sound abnormalities. The audio files are represented in wav format (Waveform format). As shown in Fig.2. The dataset consists of 40 artifacts(sounds other than heartbeat) , 351 normal, 129 murmur , 46 extrasystole and 19 extra heartbeat sound. The heartbeat sounds are in first heart sound S1 and second heart sound S2 pattern they are monitored to check the heartbeat signal. If the S1 and S2 are in regular pattern then they are called as normal heartbeat sounds. If S1and S2 together if they are repeated then there is an abnormality in the heartbeat sound resulting murmur heartbeat sound. If S1 and S2 in a separated manner if they are repeated then they are extrasystole and extra heartbeat sounds.

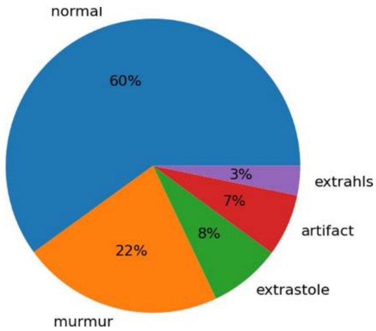


FIGURE 2. Exploration of the dataset

***Feature Extraction – MFCC*:** Feature extraction from audio data is done by computing Mel Frequency Cepstral Coefficients (MFCCs) to create 2D image-like patches as shown in Fig.3.

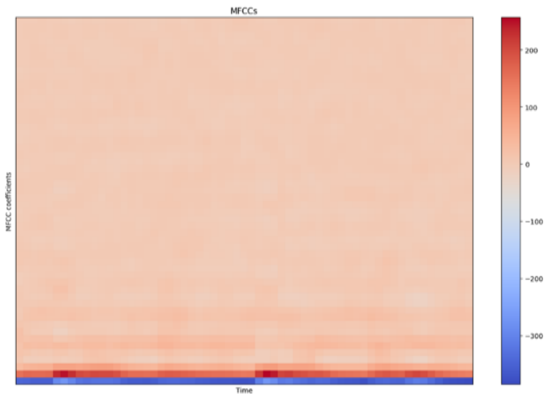


FIGURE 3. MFCC representation

The audio signal given to the MFCC will first break the audio signal waves into an overlapping frame. The audio signals are then passed to Fast Fourier Transform where the window length is given for segmentation of the frames to extract the features. MFCC captures the shape of the power spectrum of sound signal. They are obtained by first utilizing a method such as the Discrete Fourier Transform (DFT) to convert the raw audio signal into a frequency domain, and then using the mel-scale a filter bank to simulate how the human ear perceives sound frequency.

Therefore 52 MFCC features from each audio data were created using 512 hop length and segmentation of frames using windows and passed to DFT thus filtering of the audio data takes place.

Pre-processing: Pre-processing or Data Augmentation step involves a set of techniques that are applied to a raw data that are coherent to be given as input to deep learning neural networks. To make the system robust by exposing the model with different range of inputs, we can inject noise to the audio data, shift the time of the audio, change the pitch and speed of the audio. NumPy is helpful in injection of the noise and shifting time whereas librosa is useful in processing audio data. The first step for pre- processing used here is to add noise to the 52 features extracted audio data from MFCC and see how the models perform under these noisy conditions. The next step is to apply shift to both the extracted audio data and added noise that is applied, time stretch is applied to the audio data and finally, pitch shift is applied to it.

CNN + LSTM: A CNN+LSTM network is used for extraction and classification of audio data. CNN works with sequential data by extracting feature of the input. It can learn features from both spatial and time dimensions. An LSTM network works with sequential data by looping leading to coherent classification of the audio signals.

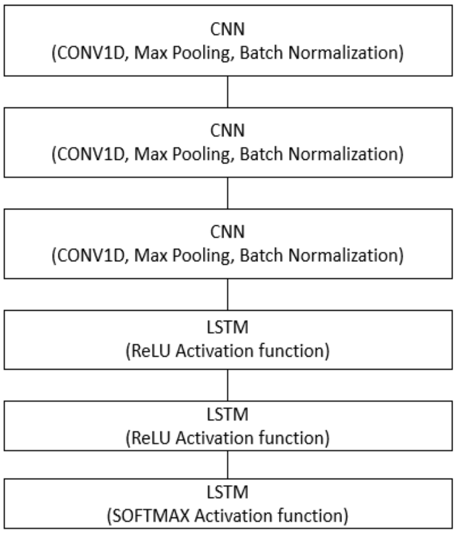


FIGURE 4. CNN+LSTM Architecture

CNN+LSTM network as shown in Fig.4 use 3 convolutional and 3 LSTM layers to perform extraction and classification. To train a CNN+LSTM network with audio data, we can extract audio based spectrograms such as MFCC to be given as input to the neural network. Sequence of input layer and input size are given to the neural network. Activation functions such as ReLU and SoftMax layer are included.CNN includes convolutional 1D layer with input shape (52,1), Max pooling layer is given with pool size 2 and a batch normalization is added. Totally 3 CNN layers were added. Three layers of LSTM which includes two ReLU activation functions and one SoftMax activation function. All the LSTM layers included dense and dropout layers.

3. RESULT AND DISCUSSION

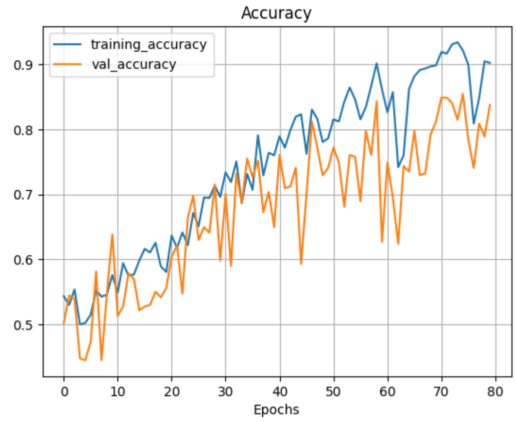
In this study as shown in TABLE I the classification of abnormal heartbeat sound is examined using CNN+LSTM. It was also compared with CNN+GRU in order to identify the best suitable deep learning model for the classification. After examining the accuracies, it has been found that CNN+LSTM is best suitable deep learning neural network.

TABLE 1. ACCURACY RESULTS

NEURAL NETWORK	ACCURACY
CNN + LSTM	83%
CNN+GRU	71%

The accuracy graph of the CNN+LSTM model as shown in Fig.5 consists of both training and validation accuracies that have been performed with 80 Epochs and the loss graph of the model as shown in Fig.6 has been obtained while performing with 80 Epochs.The model evaluation is implemented by the confusion matrix as shown in Fig.7 which is helpful in model evaluation. Confusion matrices are generated by a binocular network, where the vertical coordinate corresponds to ground truth labels and the horizontal coordinate corresponds to prediction results labels.

The model is finally integrated to a web application as given in Fig.8 in which classification of the abnormal heartbeat sound takes place.



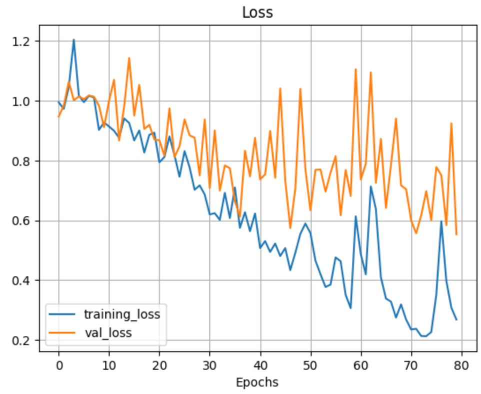


FIGURE 5. Accuracy graph **FIGURE 6.** Loss Graph

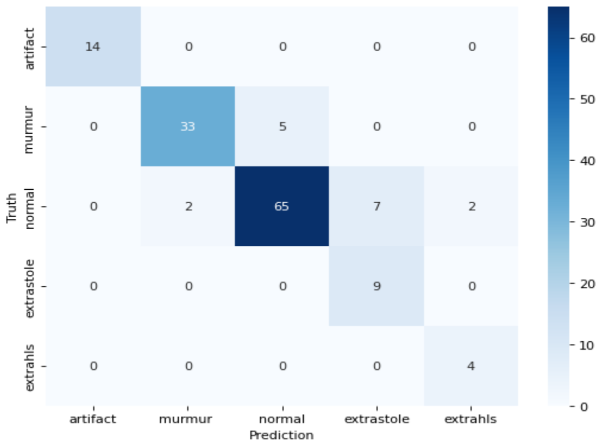


FIGURE 7. Confusion Matrix

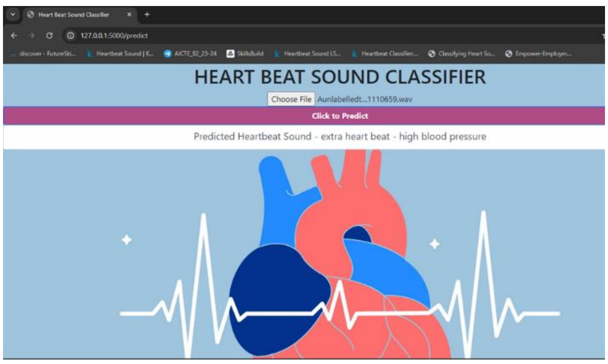


FIGURE 8. Classification Web Application

4. CONCLUSION

In this study coherent feature extraction from the given dataset has been done using MFCC to extract discriminant feature and pre-processing of augmenting the data to work even in noisy conditions has been done. Integrating of two neural networks such as CNN and LSTM for extraction and classification of the heartbeat sound respectively has been done. The extracted feature of the audio by MFCC is taken as

the input to the model which is further trained, validated and tested. The model is finally deployed in web for making an end-to-end deep learning model where the web application will be helpful in predicting and classifying the abnormalities in heartbeat sound. The study can be further improved by training the model using different heartbeat sound datasets that contains heartbeat sounds in different locations.

5. FUTURE WORK

Further the study can be carried out by using different combinations of neural network to improve the accuracy of the model. The study can be enhanced by using many new datasets and pre-processing techniques.

REFERENCES

[1]. H. F. D. Ul Haq et al., "EfficientNet Optimization on Heartbeats Sound Classification," 2021 5th International Conference on Informatics and Computational Sciences (ICICoS), Semarang, Indonesia, 2021, pp. 216-221, doi: 10.1109/ICICoS53627.2021.9651818.

[2]. A. R. Jadhav, A. G. Ghontale and A. Ganesh, "Heart sounds segmentation and classification using adaptive learning neural networks," 2017 International Conference on Signal Processing and Communication (ICSPC), Coimbatore, India, 2017, pp. 33-38, doi: 10.1109/CSPC.2017.8305881.

[3]. H. Ryu, J. Park and H. Shin, "Classification of heart sound recordings using convolution neural network," 2016 Computing in Cardiology Conference (CinC), Vancouver, BC, Canada, 2016, pp. 1153-1156.

[4]. Hassaan Malik, Umair Bashir, Adnan Ahmad, Multi-classification neural network model for detection of abnormal heartbeat audio signals, Biomedical Engineering Advances, Volume 4, 2022, 100048, ISSN 2667-0992, doi: 10.1016/j.bea.2022.100048.

[5]. Ali Raza et al., “Heartbeat Sound Signal Classification using Deep Learning,” Sensors 2019, 19(21), 4819; doi: 10.3390/s19214819

[6]. Uddipan Mukherjee et al., “A Visual Domain Transfer Learning Approach for Heartbeat Sound Classification,” Arxiv 2021 , arxiv: abs/2107.13237v1

[7]. D. Kumar, P. Carvalho, M. Antunes, P. Gil, J. Henriques and L. Eugenio, "A New Algorithm for Detection of S1 and S2 Heart Sounds," 2006 IEEE International Conference on Acoustics Speech and Signal Processing Proceedings, Toulouse, France, 2006, pp. II-II, doi: 10.1109/ICASSP.2006.1660559.

[8]. Gomes, E. F., Bentley, P. J., Pereira, E., Coimbra, M. T., & Deng, Y. (2013, February). Classifying Heart Sounds-Approaches to the PASCAL Challenge. In HEALTHINF (pp. 337-340).

[9]. Shi-Wen Deng, Ji-Qing Han, Towards heart sound classification without segmentation via autocorrelation feature and diffusion maps, Future Generation Computer Systems, Volume 60,2016,Pages 13-21,ISSN 0167-739X, doi: 10.1016/j.future.2016.01.010.

[10]. Wenjie Zhang, Jiqing Han, Shiwen Deng, Heart sound classification based on scaled spectrogram and tensor decomposition, Expert Systems with Applications, Volume 84,2017,Pages 220-231,ISSN 0957-4174, doi: 10.1016/j.eswa.2017.05.014.

[11]. Alex Graves, Jürgen Schmidhuber, Framewise phoneme classification with bidirectional LSTM and other neural network architectures, Neural Networks, Volume 18, Issues 5–6, 2005, Pages 602-610, ISSN 0893-6080, doi: 10.1016/j.neunet.2005.06.042.

[12]. Groch MW, Domnanovich JR, Erwin WD. A new heart-sounds gating device for medical imaging. IEEE Trans Biomed Eng. 1992 Mar;39(3):307-10. doi: 10.1109/10.125016. PMID: 1555861.

[13]. Sepideh Babaei, Amir Geranmayeh, Heart sound reproduction based on neural network classification of cardiac valve disorders using wavelet transforms of PCG signals, Computers in Biology and Medicine,Volume 39, Issue 1, 2009, Pages 8-15, ISSN 0010-4825, doi: 10.1016/j.compbiomed.2008.10.004.

[14]. S.M. Debbal, F. Bereksi-Reguig, Automatic measure of the split in the second cardiac sound by using the wavelet transform technique, Computers in Biology and Medicine, Volume 37, Issue 3, 2007, Pages 269-276, ISSN 0010-4825, doi: 10.1016/j.compbiomed.2006.01.005.

[15]. Simarjot Kaur Randhawa, Mandeep Singh, Classification of Heart Sound Signals Using Multi-modal Features, Procedia Computer Science, Volume 58, 2015, Pages 165-171, ISSN 1877-0509, doi: 10.1016/j.procs.2015.08.045.

