Heart Sound Classification using Convolutional Neural Networks

Sumanta Banerjee, Astha Gupta

Abstract- In the proposed model, a method is proposed to classify heart sounds using Convolutional Neural Networks. The dataset was obtained from kaggle and was in the form of .wav audio files. It was then pre-processed and divided into four categories – Artifact, Extrasystole, Murmur, and Normal. A sampling technique was used to divide the data into training and validation sets. To train the model, the data was required in form of image files. So, the .wav files were converted into spectrograms using a conversion algorithm. The model was then trained on the training data and then validated to obtain the training and validation accuracy. Maximum accuracy obtained was 96.24% for the training data 89.36% on the validation data.

1. INTRODUCTION

One of the major health issues the world is facing today is that of heart related diseases. Cardiovascular diseases continue to take more lives than any other disease worldwide. Early diagnosis and increasing the accessibility of healthcare facilities especially in developing nations can help us minimize the number of deaths caused by heart diseases.

While checking for any symptoms of heart disease, the first step is listening to the heartbeat sounds using a stethoscope and checking for any unusual sounds (auscultation). A healthy heart produces two sounds in a rhythm- a lub and a dub. One heartbeat is comprised of one lub and one dub. These are called the first heart sound (S1) and the second heart sound (S2). These sounds are made by the closing of two valves of the heart- the atrioventricular valves and the semilunar valves.

If any unusual sound is detected during this method, then some techniques are used to determine the cause and nature of the sound. This involves the use of phonocardiography along with either electrocardiography or echocardiography. Phonocardiography is a method in which the heart sounds are recorded using an electronic stethoscope. The recordings are then used to create a plot of the heart sounds.

Electrocardiography (ECG) is the most common method for diagnosing heart diseases. This involves placing electrodes on different locations of the body and then tracing the electric signals of the heart and plotting an electrocardiogram. Another technique used is that of echocardiography. It relies on the transduction of sound waves into electric signals to find and record information about the structure of the heart. It uses the ability of sound waves to penetrate human tissues. Many other techniques like the stress test are also used for a more detailed analysis of any disease.

The problem with these conventional methods is that an expert is required to be able to analyse the output and diagnose the disease. Another disadvantage is that these machines are expensive hence they are not easily accessible everywhere. So, there is a need for a method that is can overcome these shortcomings. The aim is to find a method that is self-sufficient, affordable and can accurately diagnose heart diseases based on the heart sound.

The proposed model acts as a basic diagnostic method that classifies heart sounds into one of the four categories namely Artifact ,Extrasystole , Murmur , and Normal.

The Artifact category consists of a range of different sounds like feedback squeals and noises. This category is the most different from the remaining three. No discernible heart sounds are present in the Artifact sounds.

The Extrasystole heart sounds are basically some extra sounds that appear occasionally. An extrasystole sound is basically an out of rhythm heart beat. These are easy to identify as they have an extra or a skipped heartbeat. It happens frequently in children and usually does not imply any heart disease. But in some cases, an extrasystole heart sound in adults may be an indication of a serious heart disease and it is very important that it is diagnosed early.

Murmur sounds indicate a whooshing, roaring or squishing fluid noise in between heartbeats. This is caused by the blood flow through or near the heart. It can be an indication of many heart diseases and so, it is important that it is detected early.

Normal heartbeats are those which consist only of S1 and S2. These indicate a healthy heart.

2. RELATED WORK

Many studies have been done on the classification of heart sounds before. Since the advent of neural network, studies involving it have risen due to its capabilities. Researchers have done supervised and unsupervised learning on heart sounds to either find patterns in the data or classify the data available to them.

Simarjot Kaur Randhawa et al[1] proposed a method to classify heartsounds using multi-modal features.Their method used PCG signals to classify them into three categories namely normal, systolic murmur signal and diastolic murmur signal. The classifiers used in their study were k-NN (k nearest neighbour), fuzzy k-NN and ANN (Artificial Neural Network). They achieved an accuracy of 99.6%.

In 2010, Ari et al[2] classified heart sounds into normal and abnormal using least square support vector machine (LVSSM) as a classifier using wavelet-based feature set.

In 2013, Singh et al[3] proposed a method to distinguish between normal and abnormal heart sounds using feature extraction. No ECG gating was used and a new feature 'mean12' i.e. maximum of mean in systolic region and diastolic region was proposed.

In 2017, Maryam Hamidi et al[4] proposed a method to classify heart sound signal using curve fitting and fractal dimension. In the first proposed method, curve fitting is used to achieve the information contained in the sequence of heart sound signal. In the second method, the powerful features extracted by MFCC2 are fused with the fractal features by stacking.

Grzegorz Redlarski, Dawid Gradolewski, and Aleksander Palkowski[5] proposed a classifier that utilized Linear Predictive Coding coefficients for feature extraction. The classifier was built upon combining Support Vector Machine and modified Cuckoo Search algorithm. This model showed an improvement in the performance of the diagnostic system, in terms of accuracy, complexity and range of distinguishable heart sounds. For all considered cases including simultaneous identification of twelve different heart sound classes, the model was able to achieve an accuracy of more than 93%.

A different approach for classification was proposed by Christine C. Balili, Ma. Caryssa C. Sobrepena and Prospero C. Naval[6] in a paper published in 2015. The proposed model used wavelet analysis and random forest classifiers for the classification process. The heart sounds were segmented through detection of the S1 and S2 heart sounds using Shannon Energy. image- and frequency-based features derived from Discrete and Continuous Wavelet Transforms were used as feature vectors for the random forest classifier that categorized heart sounds into

Normal, Murmur, Extrasystole, and Artifact. The segmentation process produced lower errors compared to related literature using the same dataset.

Jonathan Rubin et al[7] proposed the development of an algorithm for automatic classification of heart sound phonocardiogram waveforms as normal, abnormal or uncertain. The approach had three major steps- heart sound segmentation, transformation of one-dimensional waveforms into two-dimensional time frequency heat map representations using Mel-frequency cepstral coefficients and finally, the classification of MFCC heat maps using deep convolutional neural networks.

A convolutional neural networks model was proposed by TianyaLi,Chunmei Qing and Xiang Tian[8] which worked on the dataset from the Pascal heart sound classification workshop and attempted to classify the heart sounds into four categories namely Normal,Murmur, Extra Heart Sound and Artifact. This model used Cross-cutting and Framing to pre-process the data and Butterworth Filter and Fast Fourier Transform to manage the heart sounds signals. The proposed model input the captured frequency features for the classification. 98% global identification rate was achieved by this model. An important feature that distinguished the model from the rest was that it overcame the uncertainty in selecting features, unlike traditional classification models that used feature selection and then designing the classifier.

3. IMPLEMENTATION

The dataset of size 111 MB contained 832 files divided into two categories- Category A and Category B. Both categories contained some labelled and some unlabelled audio files with .wav extension. These were the heart sounds collected from various sources.
Steps involved-

3.1 Pre-processing-

To train the model by means of supervised learning, a total of 305 files were selected from the dataset and divided them into two batches-one for training and one for validation. A combination of random oversampling of minority class and random under-sampling of majority class was used to deal with the imbalance in the dataset. The training batch contained 225 and the validation batch had 80 files, all in their respective folders.

3.2 Conversion into Spectrograms-

The proposed model aims to build a Convolutional Neural Network. So the data to be fed into the model had to be in the form of images. For the conversion of an audio file into an image file, it was cross checked that each was a mono stereo file. Conversion of audio files to image files was achieved by formulating an algorithm that would take each file in a folder as input and then convert that .wav file into a spectrogram. A spectrogram is a graphical representation of the spectrum of the frequencies of sound.

When the algorithm was run on all the wav files, we got spectrograms of size 64x64 for all the four classes in the Validation and Training batches. Some of those are displayed below.

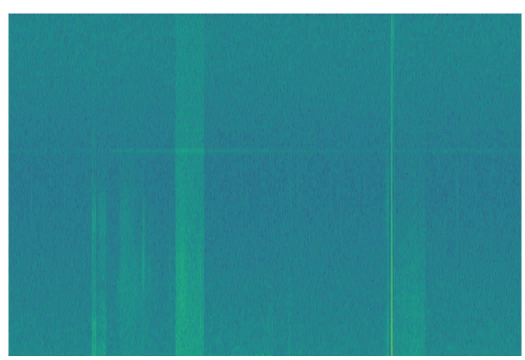

Fig 1:Artifact

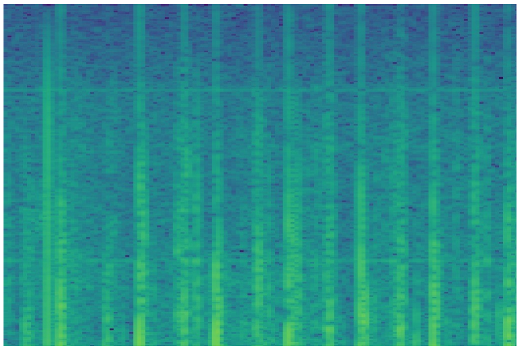

Fig 2: Extrasystole

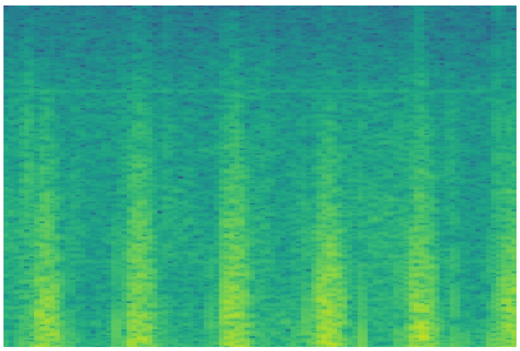

Fig 3: Murmur

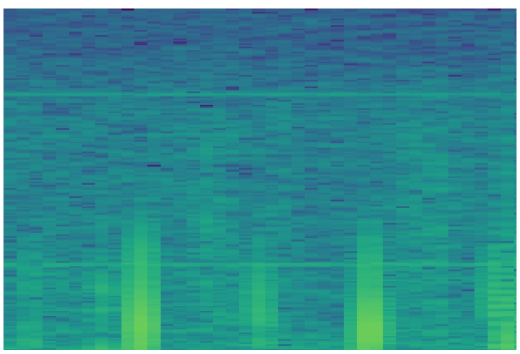

Fig 4: Normal

3.3 Building the CNN model-

The Sequential model chosen contained 10 layers with one input layer, one output layer and 8 hidden layers.

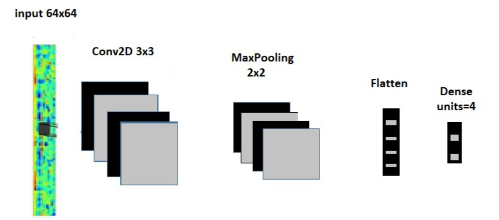

Fig 5: Layers in the CNN model

The first layer was a Convolutional 2D layer with 'Relu' as the activation function. This layer acted as the input layer and took images of size 64x64 pixels as input. This was followed by a max pooling layer to extract the important features from the previous layer with a (2,2) pool size.
Four more layers consisting of two max convolution and two max pooling layers in alternate were added next. The activation function again, was 'Relu' and the pool size was (2,2).

To convert the multi-dimensional data into a one- dimensional form, a flatten layer was added followed by a dense layer consisting of 128 units. A dropout layer was also used to ensure that the model was not overfitting the data.

The final layer was a Dense layer which utilised 'Sigmoid' as the activation function. This was done to ensure that the output values for each of the four classes lied between 0 and 1, giving a better idea of the probability.

The proposed model was then compiled, and the training data and the validation data was used to train and validate the model. The number of epochs was finalised as 22 and the number of steps for each epoch was set as 225.

3.4 Testing new data-

Labelled files which were not included in either the training or the validation dataset were fed into the model as new data and the model predicted the class in which the heart sound belonged to accurately.

4. OBSERVATIONS AND RESULT

Different accuracies were obtained with different number of epochs. There was a necessity to find a perfect balance between different constraints to get the maximum possible accuracy.On increasing the number of epochs from 1 to 21 there was a gradual increase in the accuracy. But when more than 21 epochs were run, it was found that the accuracy remained constant or kept fluctuating. Overfitting is a possible explanation for this behaviour.

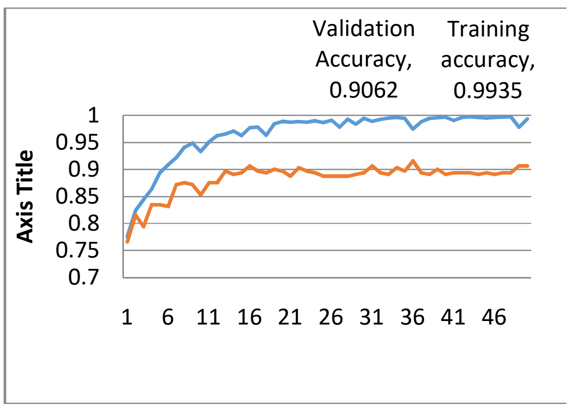

Fig 6: Dependency of accuracy on the number of epochs run

Finally, it was found that the maximum validation accuracy achieved was 90.62%.

5. CONCLUSION

This model is a stepping stone towards a more efficient and effective method to teach a machine how to accurately classify heart sounds. Further work is required on this model for it to be fit for practical applications. First and foremost, the accuracy needs to be increased which can be achieved by using higher quantity of quality data for the model. Furthermore, machines with higher computation power are required for a faster implementation of the model to ensure a much faster prediction.

If above-mentioned goals are achieved then it is possible to use this model to build an application which can be used to diagnose heart diseases accurately, thus saving time and effort.

6. REFERENCES

[1] Simarjot Kaur Randhawa and Mandeep Singh. "Classification of Heart Sound Signals using Multi-modal features", International Symposium on Computer Vision and Internet.

[2] S. Ari, K. Hembram, G. Saha. "Detection of cardiac abnormality from PCG signal using LMS based least square SVM classifier"

[3] Cheema A, Singh M, "Steps Involved in Heart Sound Analysis- A Review of Existing Trends", International Journal of Engineering Trends and Technology 2013.

[4] Maryam Hamidi, Hassan Ghassemian, Maryam Imani. "Classification of heart sound signal using curve fitting and fractal dimension", Elsevier 2017.

[5]GrzegorzRedlarski, DawidGradolewski, and AleksanderPalkowski "A System for Heart Sounds Classification"

[6] Christine C. Balili ; Ma. Caryssa C. Sobrepena ; Prospero C. Naval "Classification of heart sounds using discrete and continuous wavelet transform and random forests"

[7] Jonathan Rubin ; Rui Abreu ; Anurag Ganguli ; SaigopalNelaturi ; Ion Matei ; Kumar Sricharan "Classifying heart sound recordings using deep convolutional neural networks and mel-frequency cepstral coefficients"

[8] Tianya Li, Chunmei Qingand Xiang Tian "Classification of Heart Sounds Based on Convolutional Neural Network"