Analysis of Phonocardiogram Signal Using Deep Learning

T. Shanthi, R. Anand, S. Annapoorani, and N. Birundha

Abstract Phonocardiogram (PCG) plays an important role in the initial diagnostic screenings of patients to assess the presence of cardio-vascular abnormalities. It is also used to complement the ECG-based cardiac diagnosis for detecting cardio-vascular abnormalities. This task has been proposed to classify the heart sounds by performing deep learning technique known as CNN (Convolutional neural network). Two types of datasets were collected from the clinical environment and used as an input to this project. The experimental results show CNN provides the better results in the detection of abnormal heart sounds with good accuracy.

1 Introduction

The human heart is the most important organ in the body and supplies blood to all parts of the body. During the pumping action of the heart, the electrical and mechanical activities are done and therefore result in the flow of blood throughout the body. Phonocardiogram is a graph that represents the recording of heart sounds and murmurs obtained using stethoscope and plays a vital role in detecting abnormality and therefore it is used as the input data for our algorithm. The sounds that are obtained from a PCG are due to the vibration that occurs during the cardiac cycle of the heart. There are two sounds, namely, S1 and S2 that occur during the heart function. The first sound (S1) is produced when atrioventricular valves (tricuspid and mitral) close at the beginning of systole and the second heart sound S2 occurs when the aortic valve and pulmonary valve (semilunar valves) close at the end of systole. Sometimes there may be a chance of abnormality in the heart functions so it may result in the production of some abnormal heart sounds such as artifact, extra

T. Shanthi (✉) · R. Anand · S. Annapoorani · N. Birundha
Sona College of Technology, Salem, India
e-mail: shanthi@sonatech.ac.in

R. Anand
Sri Eshwar College of Engineering, Coimbatore, India

systole and murmur. Phonocardiogram is preferred because of its simplicity and cost efficiency. In this project the CNN model is built with the help of keras library and classified the heart sounds into normal or abnormal. Under abnormal-artifact, murmur, extra systole heart sounds were detected in the early stage by employing deep learning algorithm.

2   Literature Survey

In the past few years many researchers have used phonocardiogram [1] in different methods to diagnose the heart disease. They used Short time Fourier transform (STFT) based spectrograms to learn typical patterns of normal and abnormal PCG signals. They underwent three different studies to generate spectrogram and to perform CNN to detect the normal and abnormal signals.

In the article [2] this study they used a one-dimensional neural network based on phonocardiogram data to create an automated classification method. For the classification of PCG data they have proposed an intelligent CNN.

In this paper [3], the datasets taken from 2016 PhysionetCinc challenge PCG signals are obtained from patients with the help of a condenser microphone mounted on a stethoscope and followed by the process of amplification and filtering. Using audio jack the signals are passed to the laptop. There are separated as 4 types namely S1, Systole, S2 and Diastole using Hidden Semi Markov Model (HSMM). Using Radial Basis Function a support vector classifier was trained on 190 recordings.

In the article [4] two different heart sounds were taken from Physionet which was classified as normal and abnormal using Convolutional Neural Network. In this method, they have adopted some Preprocessing methods to nullify the effect of noises in heart sounds. By using resampled energy method, the heart sounds are segmented as S1 and S2. The classification was done using CNN.

In this paper [5], the 2016 Physio Net /Cinc challenge database has been used to validate their algorithm. The heart sound recordings had been collected from the clinical or non-clinical environment. The various algorithms are RNN, LSTM, GRU, B-RNN, B-LSTM and CNN. The input layer is given as raw signal. CNN has given them the best results when compared to other methods. In this paper [6] They used three types of datasets and analyzed the performance using CNN, RNN, LSTM, GRU and got better results using FFT.

In this paper [7], two types of datasets were used for heart sound classification from clinical environment. The input was given as raw input signal and it passed to the recurrent hidden layers such as RNN, LSTM and GRU. The experiment used a Tensorflow deep learning framework that supports graphics processing units(GPU). The LSTM has several parameters such as learning rate, memory blocks, number of hidden layers, number of epochs and so on. LSTMs gave better results than RNNs but required more training costs compared to GRU.

In the article [8] the database was collected from various clinical and non-clinical environment. The data was collected from both normal and pathological patients and also collected commonly at aortic area, pulmonic area, tricuspid area and mitral area. The classification was based on Artificial Neural Network, SVM, Hidden Markov Model and clustering. The heart sounds were classified as normal, abnormal. All heart sound recordings are divided into two types based on expert classification of normal and abnormal. Heart valve defects are mitral valve, aortic stenosis, and valve surgery. They have used hand correction in their method of project.

In this paper [9] digital recording of the heart sound is classified as normal signal, systolic murmur signal and diastolic murmur signal. Many features are extracted for classification. The classification which is used in their study is k-NN, fuzzy k-NN and Artificial Neural Network (ANN). In this k-NN and fuzzy k-NN have highest accuracy. The heart sound was recorded using electronic stethoscope and signal is displayed on the computer. The heart sounds were classified as S1, S2, S3 and S4. Using Wave pad sound editor in NCH Software the signals were first divided. The accuracies are different for different classification.

Based on the literature survey it is evident that CNN has achieved better results in many of the cases. In our proposed method we have adopted CNN with additional features to achieve better classification accuracy.

3 Methodology

The main objective of our project is the detection of cardiac abnormalities at an early stage with the deep learning model using PCG signals. In order to meet the objective, took PCG signals as an input collected from the clinical environment as SET A and SET B, done some preprocessing techniques that help in enhancing the quality of the heart sounds and undergone CNN model building with appropriate layers considering some optimization parameters. Trained and tested the model as shown in Fig. 1.

3.1 Input Description

The input to this project is the Phonocardiogram signal. The PCG Dataset was collected from two sources, namely, Set A and Set B. Both the datasets were collected from the clinical environment. Set A was collected from the public via the iStethoscope Pro iPhone app and Set B was collected from hospitals using the Digital Stethoscope DigiScope. Set A has 4 classes namely artifact, murmur, extrasystole and normal and Set B has 3 classes namely murmur, extrasystole and normal.

Fig. 1 Block diagram of the proposed work

```
                        U(6)

              U(5)      O(6)

              O(5)      O(5)      SU(3)

              O(3)      O(3)      O(3)
```
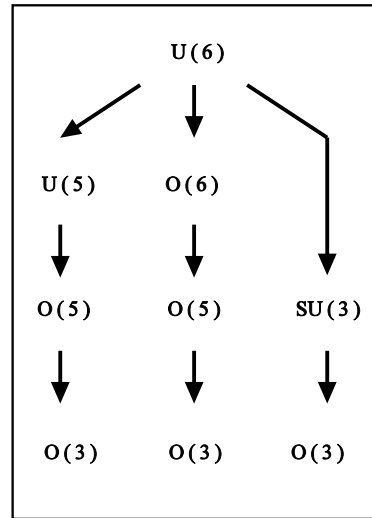
3.2 Preprocessing Technique

The preprocessing technique used here is the normalization of the audio files (.wav format). It is helpful in getting the audio with highest quality with fixed padding. The features of heart sounds are extracted. The task is to extract the best features from the heart sounds within audio data.

3.2.1 Use of Librosa

In preprocessing technique, librosa is used. Librosa is a python package which is used for audio and music analysis. It provides retrieval systems for heart sound files which is collected from the clinical environment. Kaiser Fast extraction is also used here to load the librosa package to the wave file.

3.2.2 Padding Audio File

The audio file (.wav format) is padded to a duration of 12 s and sampling rate of 160000 at a fixed rate because each wave has different durations and sampling rate so here it is padded with the audio file to have a fixed duration. The normalized audio file is received from this padding which can be used as input for further process.

3.3 Dataset Description

There are two datasets considered for this project and they are preprocessed with the techniques mentioned above. The datasets are Set A and Set B. Set A has four classes namely artifact, murmur, normal and extrasystole. Set B has three classes namely murmur, normal and extra systole. The Set A dataset is loaded with each class separately by mentioning its directory and a unique label for all the four classes. Set B dataset is also loaded with each class separately by mentioning its directory and a unique label for all the three classes shown in Tables 1 and 2.

3.4 Deep Learning Architecture

In this project, we considered CNN as our architecture since it works well with the classification of heart sounds into four classes namely the artifact, murmur, normal and extrasystole.

3.4.1 Convolutional Neural Network (CNN)

The Convolutional Neural Network also known as CNN/ConvNet is used in this project to classify the heart sound wave. The input wave has been preprocessed and taken into consideration by the CNN model. To build the CNN model, we utilized the keras library. It includes the sequential model, layers, optimizers, callbacks and regularizers.

Table 1 Set A database summary

Class name	Class label
Artifact	0
Murmur	1
Normal	2
Extra systole	3

Table 2 Set B database summary

Class name	Class label
Murmur	1
Normal	2
Extra systole	3

Table 3 Hyper parameters of the model

Hyperparameters	CNN Model
Batch size	32
Optimizer	Adam
No of epochs	100

3.4.2 Building CNN Model

The model is built by the simple Sequential architecture. Since this project is based on the analysis of Phonocardiogram heart sounds we considered the sequential model. Conv1D is used here. This layer creates a convolutional kernel with an input layer and a convolution in a single spatial(or time) dimension to generate a tensor for the output. Filters and kernel size have been defined according to the layers. We have considered RELU as the activation function as it is the most used activation functions in CNN. The kernel regularizer used here is the L2. It applies a L2 regularization penalty. Maxpool1D is used and it performs operation on down sampling the input representation by taking the maximum value over a spatial window of pool size. The window will be shifted by strides mentioned in the algorithm. Batch normalization is designed to automatically standardize the inputs to a layer in the model. Used dropouts to avoid over fitting of the model. The Global Average Pooling1D (GAP) layer is used to minimize overfitting by reducing the total number of parameters in the model. Dense layer used here is 4 (since number of classes = 4). Softmax is an activation function that is used as the last layer in the model. In this project we used 21 layers to build the CNN model.

3.4.3 Model Fitting

To fit the model, first we complied the model with the loss named binary cross entropy and the metrics we considered are accuracy, precision, recall and F1 score. Optimizer used here is the Adam. The batch generator function is defined initially to make use in fitting the model to 100 epochs with 1000 steps per epoch The batch size considered here is 32. The annealer used here is the learning rate scheduler. The model is fitted by using all those parameters as shown in Table 3.

3.4.4 Saving the Model

The model is saved, and it is used to plot the graph to check the accuracy and validation accuracy.

3.5 Result and Observation

The model is plotted with respect to the training accuracy and validation accuracy. Then the model is tested with test data to observe the classification of the heart sounds into four categories as normal, murmur, extrasystole and artifact. The model showed 79% accuracy in validation and accuracy of 85% as shown in Fig. 2. The comparison of other existing methods was shown in Table 4.

The performance of the model was analyzed using confusion matrix plot and achieved better classification accuracies for the multiple class (artifact, murmur, normal, extrasystole) classification problem as shown in Fig. 3.
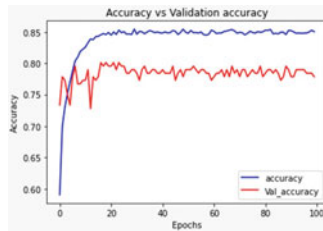

Fig. 2 Accuracy versus validation accuracy

Table 4 Performance metrices

Author method	Accuracy	Precision	Recall	F1 score
P. Gopika et al. CNN	0.80	0.85	0.79	0.82
P. Gopika et al. RNN	0.6	0.61	0.6	0.61
P. Gopika et al. LSTM	0.76	0.84	0.74	0.75
P. Gopika et al. GRU	0.78	0.8	0.78	0.78
Sujadevi VG RNN	0.53	0.55	0.53	0.54
Sujadevi VG LSTM	0.76	0.83	0.76	0.76
Sujadevi VG GRU	0.75	0.78	0.76	0.76
Proposed method	0.85	0.83	0.80	0.81

Fig. 3 Confusion matrix
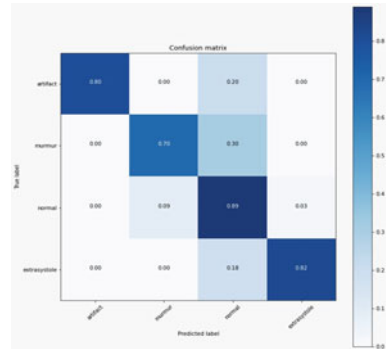
3.6 Conclusion

The phonocardiogram signals (heart sounds) play a major role in the detection of cardiac abnormalities that have undergone some preprocessing techniques like the normalization of audio waves and Kaiser fast technique to load the librosa to the wave file. Then the preprocessed Heart sound files have been loaded separately and it is gone for the CNN model building and the model is saved, got an accuracy of 84% and validation accuracy of 79% and got the results in an effective way. The proposed work gave good results with CNN model by doing simple preprocessing techniques which will be helpful for the detection of cardiac abnormalities at an early stage.

References

1. Khane KN, Khan FA, Abid A, Olmez T, Dokur Z, Khandakar A, Chowdhury MEH, Khan MS (2020) Deep learning based classification of unsegmented phonocardiogram spectrograms leveraging transfer learning. arXiv:2012.08406
2. Dhabale P, Salunke V, Vidhate K, Yalawar P, Joshi H (2020) Heart anomaly detection using deep learning approach supported signal analysis. Int J Eng Res Technol (IJERT) 09(09). (September 2020)
3. Kumar SS, Vijayalakshmi K (2020) Coronary artery disease detection from PCG signals using time domain based automutual information and spectral features. In: 2020 international conference on computing, electronics & communications engineering (iCCECE), Aug 17 2020. IEEE, pp 69–74
4. Deperlioglu O (2019) Classification of segmented phonocardiograms by convolutional neural networks. In: BRAIN. Broad research in artificial intelligence and neuroscience, vol 10, no 2, pp 5–13
5. Sujadevi VG, Soman KP, Vinayakumar R, Prem Sankar AU (2019) Anomaly detection in phonocardiogram employing deep learning. In: Computational intelligence in data mining. Springer, Singapore, pp 525–534
6. Gopika P, Sowmya V, Gopalakrishnan EA, Soman KP (2019) Performance improvement of deep learning architectures for phonocardiogram signal classification using fast fourier transform. In: 2019 9th international conference on advances in computing and communication (ICACC), pp 290–294. https://doi.org/10.1109/ICACC48162.2019.8986216

7. Sujadevi VG, Soman KP, Vinayakumar R, Sankar AP (2017) Deep models for phonocardiography (PCG) classification. In 2017 international conference on intelligent communication and computational techniques (ICCT) 22 Dec 2017. IEEE, pp 211–216
8. Clifford GD, Liu C, Moody B, Springer D, Silva I, Li Q, Mark RG (2016) Classification of normal/abnormal heart sound recordings: the PhysioNet/computing in cardiology challenge 2016. In: 2016 Computing in cardiology conference (CinC). IEEE, pp 609–612
9. Randhawa SK, Singh M (2015) Classification of heart sound signals using multi-modal features. Proc Comput Sci 58:165–171