

Nonlinear Analysis of Heart Sounds for the Detection of Cardiac Disorders Using Recurrence Quantification Analysis

Shadi Ghiasi, Mostafa Abdollahpur, Nasimalsadat Madani, Ali Ghaffari

Cardiovascular Research Group (CVRG), Department of Mechanical Engineering at K. N. Toosi University of Technology, Tehran, Iran

Abstract

Introduction: Auscultation of heart sounds using digital stethoscope technology is an effective method for the diagnosis of various cardiovascular disorders. The characteristics of phonocardiography (PCG) signals can be represented by developing a computer based algorithm as a complementary tool to facilitate clinicians.

Aims: This paper aims to address an effective feature extraction and classification technique to improve the detection of two common categories of cardiac arrhythmias, Mitral Valve Prolapse (MVP) and Coronary Artery Disease (CAD) using the heart sound recordings from PhysioNet/Computing in Cardiology 2016 challenge.

Methods: Band pass Butterworth filtering is employed to pre-process PCG signals. After detecting heart beats using common segmentation algorithms, beats which are less contaminated by noise are selected using a combination of time and time-frequency analysis. Since heart murmurs are engendered from various abnormalities in the heart, they show different characteristics. To characterize these complexities in each abnormality, Recurrence Quantification Analysis is applied. In this method the Recurrence plots which represent the mechanisms underlying the heart dynamics are quantified into recursive parameters. These set of nonlinear features show a great discrimination between classes. The resulting features are transformed to a new space using Fisher's discriminant analysis (FDA), a dimensionally reduction technique, to reduce the computational complexity of the algorithm. The classification process is in two stage. First, the new dimension-reduced feature vector is the input of the fuzzy C-means clustering (FCM) to predict normal and abnormal class. Then, the obtained 10 dimensional feature vector from RQA analysis feeds into a pattern recognition artificial neural network (ANN) to classify CAD and MVP recordings

Results: Using PCG signals of CAD, MVP and normal recordings from PhysioNet/Computing in Cardiology 2016, normal signals are easily discriminated from abnormal recordings and for classifying CAD from MVP recordings the sensitivity, specificity and accuracy of 0.853, 0.844 and 0.848 are achieved on the test set, respectively.

1. Introduction

Cardiovascular diseases have endangered humankind's health for many years.

An effective method to investigate the mechanical function of heart is auscultation of heart sounds. By using digital stethoscopes clinician have the power to analyze phonocardiogram signal (PCG) to discriminate normal heart sounds from pathological ones. This biosignal contains valuable information for the detection of heart valve disorders such as mitral valve prolapse (MVP) and coronary artery disease (CAD). These arrhythmias lead to the existence of murmurs in PCG signal which can be detected by computerized signal processing algorithms. Due to variety in heart murmurs and complex pattern of them, differentiating those thorough conventional time and / or frequency approaches is a problematic task. Therefore, a method that can investigate the nonlinear dynamics underlying heart murmurs is needed. Recurrence quantification analysis (RQA) is regarded as an effective method to examine the system behavior in phase space [2]. Many researchers have used this method to understand the complex behavior of biological signals. In [3], a vector of four features extracted by RQA method showed a great discrimination between innocent murmurs and murmurs caused by aortic stenosis. In another research [4], the identification of heart arrhythmia from heart sound is done based on RQA and Kolmogorov Entropy. These researches which are based on characterizing the complexity of the signal provide a significant basis for the detection of heart arrhythmias.

Numerous algorithms have been proposed in the literature for the classification of physiological signals. Among them artificial neural networks (ANN) and algorithms based on fuzzy logic such as fuzzy c-means clustering (FCM) have been successful for the detection of cardiovascular diseases. The combination of ANN and FCM was employed in [5] for the classification of ECG signals.

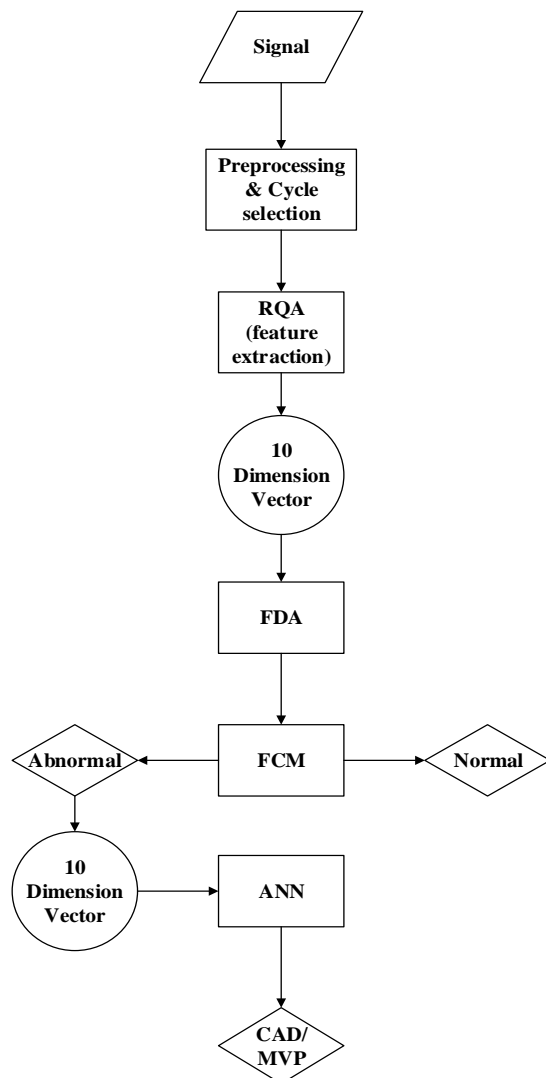
This paper presents a new method for the detection of two common heart disorders (CAD and MVP) from PCG signal. By combining nonlinear approach for feature extraction and using FCM as a classifier distinguishing these two arrhythmias from normal heart sounds are

obtained with high accuracy. This paper is organized in the following format. The preprocessing and cycle detection algorithms are described in section 2.2. Features extracted from nonlinear analysis and the new mapping space is proposed in section 2.4. Section 2.5 presents the classification procedure. Finally, the performance and a comprehensive discussion are given in section 3 and 4 respectively.

2. Materials and Methods

2.1. Database

PCG recordings of normal and pathological heart sounds exist in The PhysioNet/CinC Challenge 2016 which are collected at either clinical or nonclinical environment. Among these recordings 126 and 218 of PCG signals are annotated as MVP and CAD, respectively. More details about data can be found in [1].



2.2. Preprocessing and cycle detection

Each signal is down sampled to 1000 Hz. For the filtering, band pass filtering between 24 to 400 Hz is employed. Spike removal algorithm proposed in [6] is also applied. After the preprocessed signal is obtained, heart beats are separated using the state of the art segmentation algorithm proposed by [6]. In addition to heart beats the exact location of main heart sounds (S1 and S2) are detected. Since most of the recordings in database have heart cycles with noise or perturbation that could not be removed in preprocessing stage, an algorithm proposed in our previous research is applied [7]. The cycle quality assessment algorithm (CQA) allocates a quality to each cycle and keeps the high quality cycles for further processing.

2.3. Recurrence quantification analysis

One of the main properties of dynamical systems are recurrences which are related to the trajectory of a dynamic system. A method to visualize this behavior is recurrence plots (RP) which has the ability to map a trajectory of a system with high dimensional to a two dimensional matrix.

In order to extract useful information from these plots, recurrence quantification analysis (RQA) is used. This analysis provides features related to nonlinear dynamic behavior of time series and measures the complexity of them. Since the dynamic of PCG signals in the presence of different arrhythmias vary, this analysis is a good method to distinguish the normal and abnormal heart sound recordings. By the use of this analysis 10 features are extracted which are explained below. The first 5 features give information about the configuration and distribution of diagonal lines in a RP and the other 5 ones represent the structure behind the horizontal and vertical lines.

1. Recurrence rate: It represents the ratio of black recurrence points in the RP to the total number of points.
2. Determinism: It is the number of points which construct all diagonal lines.
3. The average length of diagonal lines
4. The length of the longest diagonal line
5. The Shannon entropy of the probability distribution of the diagonal line lengths
6. Laminarity: Histogram of vertical lines length
7. Trapping time: The average length of vertical lines
8. The length of maximal vertical line
9. Recurrence time of 1st type
10. Recurrence time of 2nd type

In the RQA theory there are some parameters that need to be chosen. The first one is embedding dimension which is chosen as 4 in this study. Time delay and the Euclidean threshold are also set as 4 and 0.8 respectively. More

detailed theory about RQA can be found in [2].

2.4. Nonlinear analysis of heart sounds

After segmenting heart sound signals into heart beats each heart beat is subjected to the RQA algorithm. Since the heart sound recordings related with CAD and MVP arrhythmias are different in dynamics compared to normal signals, this analysis can capture their differences outstandingly. Also distinguishing the CAD and MVP groups is possible with RQA. This inter dynamic discrimination is best captured with the diagonal features.

2.5 Classification

The feature matrix obtained for each patient is denoted as $F \in R^{m \times n}$ where m is the number of heart cycles excluding those noisy beats detected by CQA algorithm for each heart sound recording and n is the dimension of feature vector which is 10 in this study. The vector $F_j \in R^n$, belongs to each beat. In order not to miss the valuable information existing in each 10 features obtained by RQA

all these features are used to train one of the classifiers. On the other hand, a different representation of features may contain valuable information. Therefore, two different approaches for classification are considered.

1. ANN: An equal portion of normal and MVP signals are used to train a supervised neural network. After averaging all F_j s for a signal the 10 dimensional feature vector for each signal is fed to this classifier. This network has one hidden layer with 44 neurons and random order weight/bias is used as the training function.

2. FCM: The advantage of FCM compared to other clustering methods is that each data belongs to each class by means of a membership function. The feature matrix used in this stage is another representation of the feature vector F obtained from RQA and is used as the input of this unsupervised classifier. A feature reduction technique known as FDA (Fisher Discriminant Analysis) is applied to map the 10 dimensional feature vector to a 2 dimensional space where the features are no longer mutually correlated.

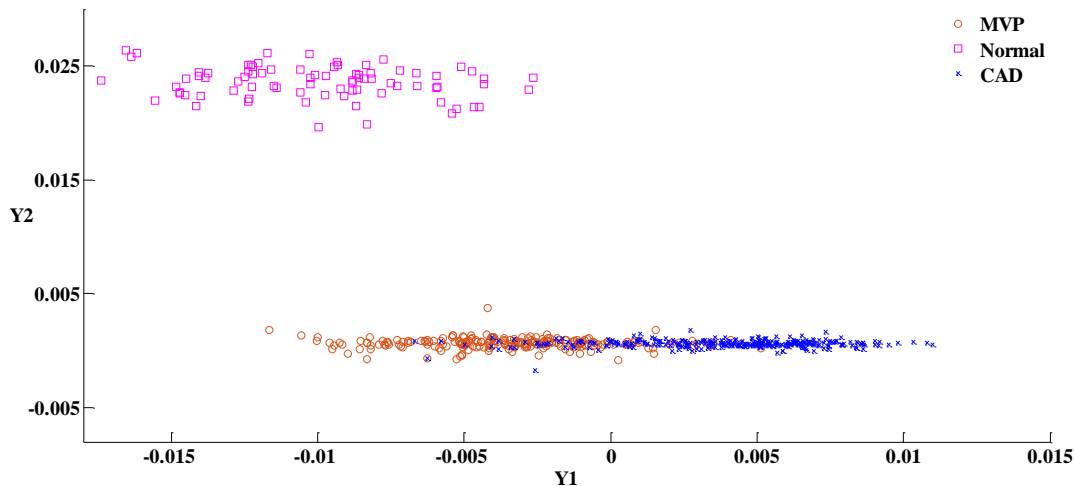


Figure 1. The distribution of features of normal, CAD and MVP heart sound recordings in the new mapped space. Y1 and Y2 represent the first and second dimensions of the new space after the 10 dimensional feature vector is fed into the feature reduction algorithm.

By combining these two approaches the classification procedure is as follows: first of all the MVP and CAD signals are regarded as one class, the abnormal class, and the FCM clustering decides whether a heart sound signal is abnormal or normal. After the normal heart sounds are separated with high accuracy, the remaining signals are fed into the ANN classifier which discriminates CAD from MVP.

3. Results

This section discusses the overall performance of the proposed method on heart sound signals provided by PhysioNet/CinC Challenge 2016 [1]. Among all signals in database there exists 126, 218 and 3000 PCG signals of MVP, CAD and normal class, respectively. In order to have fair proportion between PCG signals of different classes we separated 300 normal PCG signals randomly. Half of the signals from MVP and CAD class are used to train the ANN classifier. Since the FCM classifier is an

unsupervised learning method there is no need to separate data for train and test. Therefore all of the signals are used to test the FCM classifier.

Figure 1 shows the distribution of CAD, MVP and normal recordings when they are mapped into the two dimensional space. It is clear that the normal signals can be separated with the high accuracy but there is some conflict in discriminating CAD from MVP recordings. The results of classifying these two groups by the ANN classifier are shown in Table 1. The scores are represented for both the training and test set that we separated.

Table 1. Results of ANN classifier

Method	Dataset	Sensitivity	Specificity	accuracy
ANN	Training set	90.1	89.8	89.9
	Test Set	85.3	84.4	84.8

4. Discussion

A method for classifying three classes of heart sound recordings is presented in this paper. The two abnormal classes are CAD and MVP which are well identified by the existence of murmurs in systolic and diastolic segments of the PCG signal. Since the noises are combined with the valuable frequency information of systolic and diastolic phases, it is so important to have a good preprocessing algorithm. Also the selection of high quality cycles for further feature extraction process is also needed.

In case of PCG signals, not all the information can be obtained by time and time-frequency features. The best way to capture the underlying dynamic of signals is applying nonlinear feature extraction methods. RQA algorithm has been proved to obtain high accuracy in terms of arrhythmia detection. But discriminating CAD and MVP recordings is a challenging task since the identifiers are quite similar. But the RQA analysis acquired a high accuracy in discriminating them.

The identification of these two type of abnormal recordings from normal PCG signals is easily done with the use of a new feature space where the first two dimensions show a good discernment. The ANN pattern recognition classifier is used with the selected parameters since it proved a high performance in case of PCG signals in our previous research [6].

Since there is no specific algorithm for selecting the parameters in the RQA theory these parameters should be selected by trial and error. It is suggested to test a range of these parameters and find the optimum ones in this case. However, the selected parameters resulted in a high

accuracy in classifying CAD and MVP signals.

References

- [1] Liu C Y, Springer D B, Li Q, Moody B, Juan R A, Chorro F J, Castells F, Riog J M, Silva I, Johnson A E W, Syed Z, Schmidt S E, Panadaniil C D, Hadjileontiadis L, Naseri H, Moukadem A, Dieterlen A, Brandt C, Tang H, Samieinasab M, Samieinasab M R, Sameni R, Mark R G and Clifford G D. An open access database for the evaluation of heart sound algorithms. *Physiol Meas* 2016; 37 (in press).
- [2] Marwan N, Romano MC, Thiel M, Kurths J. Recurrence plots for the analysis of complex systems. *Physics reports*. 2007 Jan 31;438(5):237-329.
- [3] Ahlstrom C, Höglund K, Hult P, Häggström J, Kvarn C, Ask P. Distinguishing innocent murmurs from murmurs caused by aortic stenosis by recurrence quantification analysis. *Proceedings of the Enformatika Transactions on Engineering, Computing and Technology*. 2006 Dec:1305-5313.
- [4] Liang QZ, Guo XM, Zhang WY, Dai WD, Zhu XH. Identification of heart sounds with arrhythmia based on recurrence quantification analysis and Kolmogorov entropy. *Journal of Medical and Biological Engineering*. 2015 Apr 1;35(2):209-17.
- [5] Dallali, A., Kachouri, A. and Samet, M., 2011. Fuzzy c-means clustering, Neural Network, wt, and Hrv for classification of cardiac arrhythmia. *ARPN Journal of Engineering and Applied Sciences*, 6(10), p.2011.
- [6] Springer DB, Tarassenko L, Clifford GD. Logistic regression-hsmm-based heart sound segmentation. *IEEE Transactions on Biomedical Engineering*. 2016 Apr;63(4):822-32.
- [7] Abdollahpur M, Ghaffari A, Ghiasi S, Mollakazemi MJ. Detection of pathological heart sounds. *Physiological measurement*. 2017 Jun 8.

Address for correspondence

Shadi Ghiasi
 Mechanical engineering department, K.N.Toosi
 University of Technology, Tehran, Iran
 Shadi1337@gmail.com