Computers in Biology and Medicine 87 (2017) 271–284

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Classification of ECG heartbeats using nonlinear decomposition methods and support vector machine

Kandala N.V.P.S. Rajesh *, Ravindra Dhuli

School of Electronics Engineering, VIT University, Vellore 632014, India

A R T I C L E I N F O

Keywords:
ECG signal
Arrhythmia
Empirical mode decomposition
Ensemble empirical mode decomposition
Classification
SMO-SVM

A B S T R A C T

Classifying electrocardiogram (ECG) heartbeats for arrhythmic risk prediction is a challenging task due to minute variations in the amplitude, duration and morphology of the ECG signal. In this paper, we propose two feature extraction approaches to classify five types of heartbeats: normal, premature ventricular contraction, atrial premature contraction, left bundle branch block and right bundle branch block. In the first approach, ECG beats are decomposed into intrinsic mode functions (IMFs) using ensemble empirical mode decomposition (EEMD). Later four parameters, namely the sample entropy, coefficient of variation, singular values, and band power of IMFs are extracted as features. In the second approach, the same features are computed from IMFs extracted using an empirical mode decomposition (EMD) algorithm. The features obtained from the two approaches are independently fed to the sequential minimal optimization-support vector machine (SMO-SVM) for classification. We used two arrhythmia databases for our evaluation: MIT-BIH and INCART. We compare the proposed approaches with existing methods using the performance measures given by the average values of (i) specificity, (ii) sensitivity, and (iii) accuracy. The first approach demonstrates significant performance with 98.01% sensitivity, 99.49% specificity, and 99.20% accuracy for the MIT-BIH database and 95.15% sensitivity, 98.37% specificity, and 97.57% accuracy for the INCART database.

1. Introduction

The functioning of the human body is under the profound influence of the heart, which plays a significant role in the cardiovascular system. Any functional limitations in the heart can lead to cardiovascular diseases (CVDs). CVDs are increasing the mortality rate worldwide, especially in the low and middle-income countries [43]. Sudden cardiac arrest due to cardiac arrhythmia is one of the major concerns in these countries. Cardiac arrhythmias occur due to improper electrical conduction or impulse formation in the heart, which can affect heart morphology or disrupt the rate of a regular heartbeat. An arrhythmia can cause abnormal heartbeats such as ectopic and bundle branch block beats. Ectopic beats are formed because of improper electrical impulse formation. Ectopic beats are further classified as premature ventricular and atrial contractions. These beats will disturb the heart's regular rhythm and can induce serious arrhythmias such as ventricular or atrial fibrillation. Bundle branch blocks impede the normal pathway of electrical impulses through the conduction system to the ventricles. This causes asynchronous ventricular contractions and heart function deterioration, which may lead to life-threatening situations [52]. These situations are prompting researchers to investigate both detection and classification methods for cardiac arrhythmias. A sequence of heartbeats can be mapped to an arrhythmia; hence, an important step in detection and classification of arrhythmias is heartbeats characterization [61].

An electrocardiogram (ECG) is a popular diagnostic tool for examining the heart's electrical activity. An ECG signal represents the heart muscle's depolarization and repolarization actions as heartbeats. Manual observation of subtle changes between and within the heartbeats in an ECG is a tedious job. Therefore, by recognizing various kinds of heartbeats, computer-aided diagnosis (CAD) plays an important role in detection and classification of cardiac arrhythmias. This will help cardiologists to monitor physiological conditions of heart activity at regular intervals. This work investigated a method for classification of heartbeats that could help in detecting various types of arrhythmias. The importance of CAD in cardiac arrhythmia detection and classification was demonstrated in previous studies [29,40,41]. An appropriate choice of features and classifiers would improve the classification method. Heartbeat classification methods can be broadly categorized into four groups based on the selection of features [39], viz., time-domain, frequency-domain, time-frequency and nonlinear methods.

* Corresponding author.
E-mail addresses: kandala.rajesh2014@gmail.com (K.N.V.P.S. Rajesh), ravindradhuli@vit.ac.in (R. Dhuli).

http://dx.doi.org/10.1016/j.compbiomed.2017.06.006
Received 14 February 2017; Received in revised form 3 June 2017; Accepted 3 June 2017

0010-4825/

In the time-domain, linear prediction [36], ECG morphology, R-R intervals [13] and linear discriminant analysis (LDA) [63] are used. In frequency-domain, feature extraction methods based on Fourier transform [44], subband decomposition [2] and non-parametric power spectral density (PSD) [30] were explored for heartbeat discrimination. Time-frequency approaches are popular because of their multi-resolution analysis capabilities. Afonso and Tompkins [1] used short time Fourier transform (STFT) and Wigner-Ville distribution (WVD) to discriminate shockable rhythms from non-shockable cardiac rhythms. Wavelets are also widely used in this context [3,27,31,42]; however, the reasons behind rhythmic and wave shape changes in ECG are complicated. Deviation from the rhythmic activity in the ECG signal is called arrhythmia. Arrhythmias are considered being caused by pathological conditions, such as disturbances in the autonomic nervous system, dysfunctions of natural pacemakers, or failures in the electrical conduction pathway in the myocardium. The ECG signal results from combining signals from mutually related and inherently nonlinear biological systems. Hence, to analyze this, we need a decomposition technique that is nonlinear and adaptive.

Researchers attempted to decompose ECG into various modes using some basis functions [26,40]. Nonlinear features, like higher-order cumulants [42,64] and approximate entropy [34] are extracted from these modes. However, these basis functions are not dynamic and thus unable to meet morphological changes. There is a need for an adaptive mechanism to decompose the ECG signal. A popular analysis tool known as ensemble empirical mode decomposition (EEMD) [59] is used in this work. EEMD decomposes non-stationary signals originating from nonlinear systems into intrinsic mode functions (IMFs), depending on the signal characteristics, without depending on any prior basis. These modes are regular and spread across the whole time span with similar scales [12]; therefore, EEMD will give the subtle information of a signal, in terms of modes. EEMD has shown its potentiality in a variety of applications, including fault diagnosis in mechanical applications [15,32,33], ECG filtering [10] and seismic signal analysis [58]. So far, EEMD is not exploited in ECG signal classification. In this paper, we attempt to classify five types of morphological heartbeats (normal, PVC, APC, LBBB and RBBB) based on their EEMD features. Classifier selection is crucial in the detection and classification problem; different classifiers were used in the literature, including self-organizing maps (SOM) [7], k-nearest neighbor [26], ANN's [21,47] and cluster analysis [62]. In this paper, we used a sequential minimal optimization-support vector machine (SMO-SVM) for classification. Martis et al. [39] reported that in certain cases, nonlinear methods perform well even under noisy conditions [28,54]; therefore, we tested the performance of our method in the presence of noise.

This paper is structured as follows. Section 2 provides information regarding the dataset used in this work also presents the proposed method including pre-processing, feature extraction and classification. Description about the experiment and discussion on the obtained results are presented in Section 3. Conclusion is presented in Section 4.

2. Methodology

For analyzing the proposed method, we chose the database from the Massachusetts Institute of Technology-Beth Israel Hospital (MIT-BIH) arrhythmia database [45]. In this database, several types of heartbeats obtained from male and female patients are stored in 48 records. Each record duration is 30:06 min sampled at 360 Hz. From the collected data, five types of heartbeats: normal, PVC, APC, LBBB and RBBB are used for classification. Details of these data are provided in Table 1. The proposed methodology is elucidated below.

Table 1
Number of beats used in this work.

Type of ECG beat	Number of beats	Record name
Normal (N)	2000	100,101,108,112
Premature Ventricular Contraction (PVC)	2000	106,107,200,201
Atrial Premature Contraction (APC)	2000	100,101,103,108,112,114,116, 118,121,124,200,201,202,205, 207,209,213,215,219,220,222, 223,228,231,232,233
Left Bundle Branch Block (LBBB)	2000	109,111,207,214
Right Bundle Branch Block (RBBB)	2000	118,207,212

In general, automatic heartbeat classification system comprises three building blocks: pre-processing, feature extraction and classification as shown in Fig. 1. We will discuss the details of the approaches and also the theoretical background of the employed techniques.

2.1. Pre-processing

Pre-processing consists of denoising and segmentation parts.

2.1.1. Denoising

For an accurate diagnosis, clinicians always prefer noise free ECG signals. However, in real time scenario, ECG recordings have inherent artifacts. Prominent among them are high frequency noise due to power line interference and low frequency noise such as baseline wander due to motion or respiration of the patient. Hence, it is advisable to first denoise the ECG records. For denoising the filtering technique proposed by Ref. [5] is used with slight modification.

This method consists four steps:

1. Eliminate the mean bias from the noisy ECG signal.
2. Moving average filter with order five (cut-off frequency = 24 Hz) is used for eliminating the high frequency components due to power line interference and muscle noise.
3. High-pass filter with cut-off frequency 1 Hz is used for baseline wander suppression.
4. Additional low-pass filter with cut-off frequency 45 Hz is used to further suppress any left out high frequency artifacts, since most of the ECG signal energy lies between 0.5 and 45 Hz [60].

Segmentation follows the denoising mechanism.

2.1.2. Segmentation

ECG signal obtained after denoising has to be segregated into individual heartbeats. Two major approaches available for this purpose are (1) detecting the QRS complex, (2) using the annotation files provided by the experts.

QRS complexes can be detected using various algorithms such as [24,38,46,49].

In our analysis, we make use of MIT database annotation files [11,14,29]. To obtain a heartbeat, a window of length 300 is applied on ECG signal. These beats from each class are presented in Fig. 2 with the corresponding time axis. By observing Fig. 2, large morphological variations among and within the normal and abnormal heartbeats can be observed. A major problem for the automated ECG beat classification is uncertain ECG morphology. It can be different even within a patient record. But, there is a possibility that it is similar to a different type of ECG beat [48]. This situation can be handled by choosing appropriate features.

2.2. Feature extraction

With the help of feature extraction, we can represent a large data in a few samples, often called, features. Choosing appropriate features will improve the performance accuracy. As shown in Fig. 1, we implement feature extraction in two stages.

In stage one, ECG beats are decomposed into IMFs using EEMD and EMD.

In stage two, sample entropy, the coefficient of variation, singular values and band power values are calculated from IMFs for subsequent analysis.