Artificial Intelligence in Medicine (2008) **44**, 51—64

ARTIFICIAL INTELLIGENCE IN MEDICINE

http://www.intl.elsevierhealth.com/journals/aiim

Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal

Babak Mohammadzadeh Asl [a,*], **Seyed Kamaledin Setarehdan** [b], **Maryam Mohebbi** [a]

[a] *Department of Biomedical Engineering, Tarbiat Modares University, Tehran, Iran*
[b] *Control and Intelligent Processing Center of Excellence, Faculty of Electrical and Computer Engineering, University of Tehran, P.O. Box 14395/515, Tehran, Iran*

Received 14 July 2007; received in revised form 24 April 2008; accepted 28 April 2008

KEYWORDS
Arrhythmia classification;
Generalized discriminant analysis;
Heart rate variability;
Nonlinear analysis;
Support vector machine

Summary

Objective: This paper presents an effective cardiac arrhythmia classification algorithm using the heart rate variability (HRV) signal. The proposed algorithm is based on the generalized discriminant analysis (GDA) feature reduction scheme and the support vector machine (SVM) classifier.
Methodology: Initially 15 different features are extracted from the input HRV signal by means of linear and nonlinear methods. These features are then reduced to only five features by the GDA technique. This not only reduces the number of the input features but also increases the classification accuracy by selecting most discriminating features. Finally, the SVM combined with the one-against-all strategy is used to classify the HRV signals.
Results: The proposed GDA- and SVM-based cardiac arrhythmia classification algorithm is applied to input HRV signals, obtained from the MIT-BIH arrhythmia database, to discriminate six different types of cardiac arrhythmia. In particular, the HRV signals representing the six different types of arrhythmia classes including *normal sinus rhythm*, *premature ventricular contraction*, *atrial fibrillation*, *sick sinus syndrome*, *ventricular fibrillation* and 2° *heart block* are classified with an accuracy of 98.94%, 98.96%, 98.53%, 98.51%, 100% and 100%, respectively, which are better than any other previously reported results.
Conclusion: An effective cardiac arrhythmia classification algorithm is presented. A main advantage of the proposed algorithm, compared to the approaches which use the ECG signal itself is the fact that it is completely based on the HRV (R—R interval) signal which can be extracted from even a very noisy ECG signal with a relatively high

* Corresponding author. Tel.: +98 912 4715235; fax: +98 21 88633029.
E-mail address: b.mohammadzadeh@modares.ac.ir (B.M. Asl).

doi:10.1016/j.artmed.2008.04.007

accuracy. Moreover, the usage of the HRV signal leads to an effective reduction of the processing time, which provides an online arrhythmia classification system. A main drawback of the proposed algorithm is however that some arrhythmia types such as left bundle branch block and right bundle branch block beats cannot be detected using only the features extracted from the HRV signal.

1. Introduction

Heart diseases are a major cause of mortality in the developed countries. Many different instruments and methods are developed and being daily used to analyze the heart behavior. One of the relatively new methods to assess the heart activity and to discriminate different cardiac abnormalities is to analyse the so-called heart rate variability (HRV) signal. HRV signal, which is generated from electrocardiogram (ECG) by calculating the inter-beat intervals, is a nonlinear and nonstationary signal that represents the autonomic activity of the nervous system and the way it influences the cardiovascular system. Hence, measurement and analysis of the heart rate variations is a non-invasive tool for assessing both the autonomic nervous system and the cardiovascular autonomic regulatory system. Furthermore, it can provide useful information about the current and/or the future heart deficiencies [1]. Therefore, HRV analysis can be considered as an important diagnostic tool in cardiology.

Several methods have been proposed in the literature for automatic cardiac arrhythmia detection and classification. Some examples of the techniques used include: threshold-crossing intervals [2], neural networks [3—10], wavelet transforms [11], wavelet analysis combined with radial basis function neural networks [12], support vector machines [13], Bayesian classifiers [14], fuzzy logic combined with the Markov models [15], fuzzy equivalence relations [16], and the rule-based algorithms [17]. Most of these studies [2—6,11—13] are based on the analysis of the ECG signal itself. In most methods, the various features of the ECG signal including the morphological features are extracted and used for classification of the cardiac arrhythmias. This is a time-consuming procedure and the results are very sensitive to the amount of the noise.

An alternative approach would be to extract the HRV signal from the ECG signal first by recording the R—R time intervals and then processing the HRV signal instead. This is a more robust method since the R—R time intervals are less affected by the noise. Different HRV signal analysis methods for cardiac arrhythmia detection and classification were introduced in the past. Tsipouras and Fotiadis [8] proposed an algorithm based on both time and time—frequency analysis of the HRV signal using a set of neural networks. Their method could only classify the input ECG segments as "normal" or "arrhythmic" segments without the ability to identify the type of the arrhythmia. Acharya et al. [16] employed a multilayer perceptron (MLP) together with a fuzzy classifier for arrhythmia classification using HRV signal. They could classify the input ECG segments into one of the four different arrhythmia classes. In [17], Tsipouras et al. proposed a knowledge-based method for arrhythmia classification into four different categories. The main drawback of their algorithm was the fact that the atrial fibrillation, which is an important life-threatening arrhythmia, was excluded from the ECG database.

In this paper a new arrhythmia classification algorithm is proposed which is able to effectively identify six different and more frequently occurring types of cardiac arrhythmia. These arrhythmias are namely the normal sinus rhythm (NSR), the premature ventricular contraction (PVC), the atrial fibrillation (AF), the sick sinus syndrome (SSS), the ventricular fibrillation (VF) and the 2° heart block (BII). The proposed algorithm is based on the two kernel learning machines of the *generalized discriminant analysis* (GDA) and the *support vector machine* (SVM). By cascading SVM with GDA, the input features will be nonlinearly mapped twice by radial basis function (RBF). As a result, a linear optimal separating hyperplane can be found with the largest margin of separation between each pair of arrhythmia classes in the implicit dot product feature space.

GDA is a data transformation technique which was first introduced by Baudat and Anouar [18]. It can be considered as a kind of generalization to the well-known *linear discriminant analysis* (LDA) algorithm and has become a promising feature extraction scheme [19—24] in recent years. The main steps in GDA are to map the input data into a convenient higher dimensional feature space F first and then to perform the LDA algorithm on the F instead of the original input space. By GDA therefore, both dimensionality reduction of the input feature space and selection of the useful discriminating features can be achieved simultaneously.

SVM, which was first proposed by Vapnik [25], has been considered as an effective classification