Computers in Biology and Medicine 75 (2016) 190–202

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Proposition of novel classification approach and features for improved real-time arrhythmia monitoring

Yoon Jae Kim [a], Jeong Heo [a], Kwang Suk Park [b,c], Sungwan Kim [b,c,*]

[a] Interdisciplinary Program for Bioengineering, Graduate School, Seoul National University, Seoul 08826, Republic of Korea
[b] Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul 03080, Republic of Korea
[c] Institute of Medical and Biological Engineering, Medical Research Center, Seoul National University, Seoul 03080, Republic of Korea

ARTICLE INFO

Article history:
Received 29 January 2016
Received in revised form
3 June 2016
Accepted 7 June 2016

Keywords:
Arrhythmia
Electrocardiography
Support vector machine
Taguchi method
Heart rate variability

ABSTRACT

Arrhythmia refers to a group of conditions in which the heartbeat is irregular, fast, or slow due to abnormal electrical activity in the heart. Some types of arrhythmia such as ventricular fibrillation may result in cardiac arrest or death. Thus, arrhythmia detection becomes an important issue, and various studies have been conducted. Additionally, an arrhythmia detection algorithm for portable devices such as mobile phones has recently been developed because of increasing interest in e-health care. This paper proposes a novel classification approach and features, which are validated for improved real-time arrhythmia monitoring.

The classification approach that was employed for arrhythmia detection is based on the concept of ensemble learning and the Taguchi method and has the advantage of being accurate and computationally efficient. The electrocardiography (ECG) data for arrhythmia detection was obtained from the MIT–BIH Arrhythmia Database ($n=48$). A novel feature, namely the heart rate variability calculated from 5 s segments of ECG, which was not considered previously, was used. The novel classification approach and feature demonstrated arrhythmia detection accuracy of 89.13%. When the same data was classified using the conventional support vector machine (SVM), the obtained accuracy was 91.69%, 88.14%, and 88.74% for Gaussian, linear, and polynomial kernels, respectively. In terms of computation time, the proposed classifier was 5821.7 times faster than conventional SVM. In conclusion, the proposed classifier and feature showed performance comparable to those of previous studies, while the computational complexity and update interval were highly reduced.

1. Introduction

Arrhythmia is a collective term for the cardiac rhythm that occurs due to abnormal electrical activity in the heart and deviates from the normal sinus rhythm. Various types of arrhythmias exist, and many of them have no distinct symptoms. When symptoms are present, arrhythmias may include palpitations, light-headedness, fainting, shortness of breath, or chest pain. Whereas most arrhythmias are not serious, some predispose a person to complications such as stroke or heart failure. In extremely serious cases, arrhythmias such as ventricular fibrillation may result in cardiac arrest or death. Therefore, arrhythmia detection becomes an important issue and various studies have been conducted [1–17]. Previous studies used machine-learning approaches such as support vector machines (SVMs) [8,9,15,16], artificial neural networks (ANNs) [4,7,16], fuzzy logic [7], and decision trees [1,13]. Various features also have been proposed to increase the detection accuracy. Not only time-domain features such as *R* peak amplitude, RR-interval, PR interval, and the width of QRS complexes, but also frequency-domain features contribute to accurate arrhythmia detection. Recently, an arrhythmia detection algorithm for portable devices such as mobile phones has been developed because interest for e-health care is increasing [1,13–16]. In this study, a novel classification approach and features are proposed, and validated for improved real-time arrhythmia monitoring.

Abbreviations: SVM, support vector machine; ANN, artificial neural network; LSE, least squared error; HRV, heart rate variability; ECG, electrocardiography; Ust-HRV, ultra-short term HRV; SDRR, standard deviation of RR-intervals; RMSSD, square root of the mean of squares of the successive differences between adjacent RR-intervals; SDSD, standard deviation of the successive differences between adjacent RR-intervals greater than 20 ms; pRR20, proportion of intervals presenting time duration differences between adjacent RR-intervals greater than 20 ms; pRR50, proportion of intervals presenting time duration differences between adjacent RR-intervals greater than 50 ms; SNR, signal-to-noise ratio; ROC, receiver operating characteristic; AUC, area under the curve; PSO, particle swarm optimization

* Correspondence to: Department of Biomedical Engineering, Seoul National University College of Medicine, Seoul 03080, Republic of Korea.
E-mail addresses: kyj182731@naver.com (Y.J. Kim), hjeong20@bmsil.snu.ac.kr (J. Heo), kspark@bmsil.snu.ac.kr (K.S. Park), sungwan@snu.ac.kr (S. Kim).

http://dx.doi.org/10.1016/j.compbiomed.2016.06.009
0010-4825/

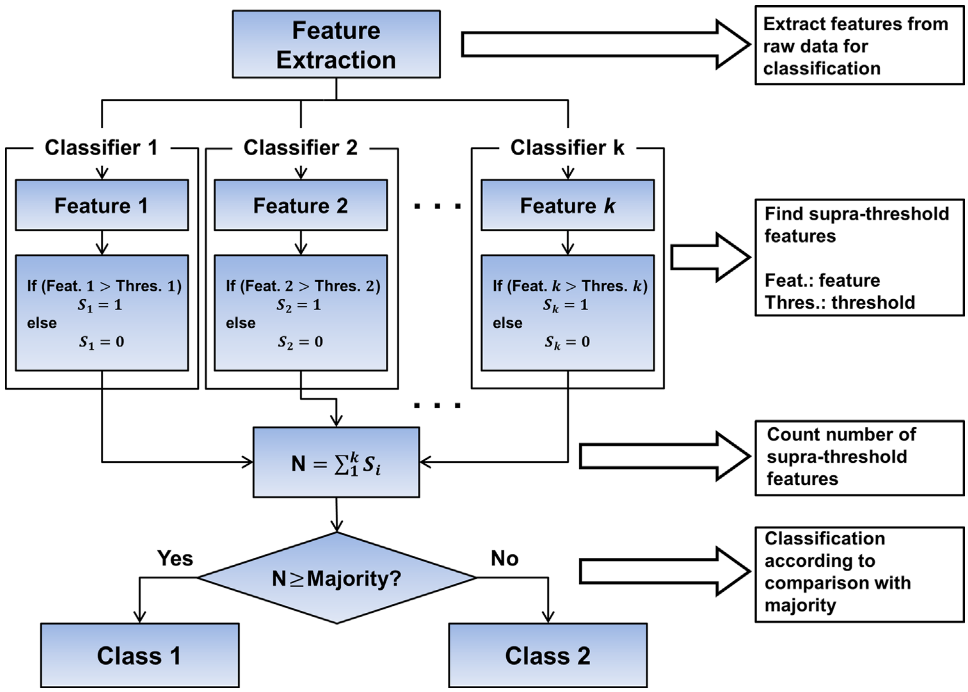

Fig. 1. The simplest form of majority voting for two-class classification. Feat.: feature; Thres.: threshold.

The classification approach that was employed for arrhythmia detection is based on the concept of ensemble learning and the Taguchi method and has the advantage of being accurate and computationally efficient. Ensemble approaches such as boosting, majority voting, decision templates, least squared error (LSE) based-classifier weighting, and hierarchical combination were proposed previously [18–24]. These methods improve classification performance by combining several types of classifiers with relatively weak performance. The method of majority voting is one of the most intuitive ensemble combination techniques; this method has been used for various purposes such as disease detection. It is a simple but powerful method that improves performance by aggregating individual classifiers of uncorrelated errors [22]. In the current study, one of the simplest forms of majority voting, which has extremely low computational complexity, is proposed. As the simplest form of majority voting for binary classification, it judges a person as having a disease when the number of features with supra-threshold values is above a specific integer (conventionally half of the total number of features). The flowchart of this classification approach is shown in Fig. 1. This simple approach does not guarantee performance comparable to other machine-learning approaches owing to the simplicity of its algorithm. However, the optimal combination of thresholds can provide improved performance. Thus, the Taguchi method is used to optimize the thresholds without full factorial experiment [25–27]. The method has been applied to optimize other pattern recognition approaches such as ANN and fuzzy logic [28–31].

The performance of the proposed classification approach is evaluated by comparing its performance in terms of arrhythmia detection with that of another machine learning classifier, conventional SVM. The classification is achieved by using heart rate variability (HRV) features. According to previous studies, HRV is an important index to detect arrhythmia [3,4,8,9,11,12,14,15]. Time and frequency analysis of HRV have been used in these previous studies and demonstrated high performance [3,4,8,9,12,14,15]. Especially for portable real-time arrhythmia monitoring in daily life, HRV based on the RR-interval is important because it is relatively difficult to obtain other features such as R peak amplitude, PR-intervals, and shape of QRS complexes owing to serious motion artifacts. Additionally, time-domain HRV features can be calculated from other types of bio-signals such as photoplethysmography, which can be estimated from light and the camera of a mobile phone [14,15,17]. This method therefore enables arrhythmia detection using a portable device.

Although previous studies based on HRV exhibited notable results, most arrhythmia detection algorithms using HRV require an electrocardiography (ECG) segment that contains more than 32 RR-intervals [3,4,8,9,14]. Thus, a previous study tried to determine whether ultra-short term HRV (Ust-HRV) has the ability to detect atrial fibrillation [12]. The study showed that Ust-HRV based features calculated from a 10 s segment of ECG contains atrial fibrillation that significantly ($p < 0.05$) deviates from those of a normal sinus rhythm. However, the analysis of Ust-HRV was limited to atrial fibrillation. Furthermore, an arrhythmia detection algorithm using Ust-HRV has not yet been implemented to the best of our knowledge. In the current study, Ust-HRV calculated from a 5 s segment of ECG is used for arrhythmia detection for the purpose of a reduced update interval of monitoring. The target arrhythmia types contain not only atrial fibrillation but also various types of arrhythmia beats and rhythms, even though binary classification is conducted. Five types of HRV features are used in the current study. They are the standard deviation of RR-intervals (SDRR), the square root of the mean of squares of the successive differences between adjacent RR-intervals (RMSSD), the standard deviation of the successive differences between adjacent RR-intervals (SDSD), the proportion of intervals presenting time duration differences between adjacent RR-intervals greater than 20 ms (pRR20), and the proportion of intervals presenting time duration differences between adjacent RR-intervals greater than 50 ms (pRR50).

2. Method for optimizing the thresholds for majority voting

The determination of a globally optimal combination of thresholds of several features is best achieved by performing an