Physica A 509 (2018) 56–65

Contents lists available at ScienceDirect

Physica A

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

Early prediction of paroxysmal atrial fibrillation based on short-term heart rate variability

Ali Narin [a], Yalcin Isler [b], Mahmut Ozer [a], Matjaž Perc [c,*]

[a] Bülent Ecevit University, Department of Electrical and Electronics Engineering, Zonguldak, Turkey
[b] Izmir Katip Celebi University, Department of Biomedical Engineering, Cigli, Izmir, Turkey
[c] Faculty of Natural Sciences and Mathematics, University of Maribor, Slovenia

ARTICLE INFO

Article history:
Received 6 June 2018
Available online xxxx

Keywords:
Paroxysmal atrial fibrillation
Heart rate variability
Genetic algorithm
Wavelet packet transform
Early prediction

ABSTRACT

Atrial fibrillation (AF) is the most common arrhythmia type and its early stage is paroxysmal atrial fibrillation (PAF). PAF affects negatively the quality of life by causing dyspnea, chest pain, feeling of excessive fatigue, and dizziness. In this study, our aim is to predict the onset of paroxysmal atrial fibrillation (PAF) events so that patients can take precautions to prevent PAF events. We use an open data from Physionet, Atrial Fibrillation Prediction Database. We construct our approach based on the heart rate variability (HRV) analysis. Short-term HRV analysis requires 5-minute data so that each dataset was divided into 5-minute data segments. HRV features for each segment are calculated from time-domain measures and frequency-domain measures using power spectral density estimations of fast Fourier transform, Lomb–Scargle, and wavelet transform methods. Different combinations of these HRV features are selected by Genetic Algorithm and then applied to *k*-nearest neighbors classification algorithm. We compute the classifier performances by the 10-fold cross-validation method. The proposed approach results in 92% sensitivity, 88% specificity and 90% accuracy in the 2.5–7.5 min time interval priors to PAF event. The proposed method results in better classification performance than the similar studies in literature. Comparing the existing studies, we propose that our approach provide better tool to predict PAF events.

1. Introduction

A healthy heart beats at 60–80 times per minute [1]. Electrical impulses from the sinoatrial (SA) node spread curvaceously to stimulate the atria and initiate contractions for the healthy beat. However, random and multiple impulses are produced in patients with atrial fibrillation (AF), in addition to the impulses from the SA node [2]. These impulses cause fibrillation instead of normal contractions of the atrium.

There are three types of AF: Paroxysmal AF (PAF), Persistent AF, and Chronic AF. PAF is the first-stage of AF. In this situation, AF starts suddenly and continues up to a week. If it is determined as soon as possible, the complications and the progress of this situation can be avoided [3].

AF affects negatively the quality of life by causing dyspnea, chest pain, feeling of excessive fatigue, and dizziness [4,5]. Moreover, AF increases the risk of stroke five times, the risk of death (due to stroke) two times. Consequently, the patient care

* Corresponding author.
E-mail addresses: alinarin45@gmail.com (A. Narin), islerya@yahoo.com (Y. Isler), mahmutozer2002@yahoo.com (M. Ozer), matjaz.perc@uni-mb.si (M. Perc).

https://doi.org/10.1016/j.physa.2018.06.022
0378-4371/

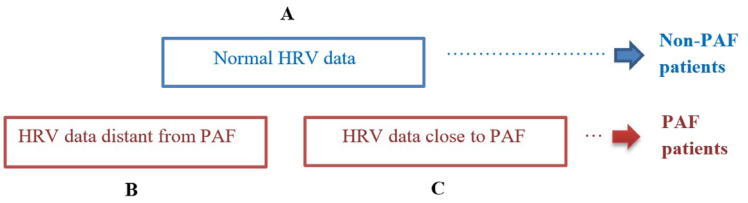

Fig. 1. Datasets used in the study: (A) Normal subjects (B) HRV data distant from PAF (C) HRV data close to PAF.

costs increase 1.5 times [3]. AF is very common among heart diseases with the incidence of 1–2% of the general population. It is estimated that 2.7–6.1 million Americans and over 6 million Europeans suffer from this situation [3,6]. Furthermore, the incidence of AF is around 5–15% at the age of 80s while this is less than 0.5% at the age of 50s, which means the incidence of AF drastically increases by the age [3].

The number of studies related to pre-detection of PAF events has been increased for 2 decades [7–19]. Zong et al. [7] reported that the frequency of APC beats is a highlighting feature before PAF occurs of 30-min ECG signals. Langley et al. [8] have also estimated PAF by looking at the quantities of atrial ectopic and ventricular ectopic beats using 30-min RR data. They have stated that there is a significant increase in atrial ectopic beats before a PAF event [8]. Using 1-min, 5-min, 10-min, and 30-min RR data, Chazal and Henegham [9] have performed early estimation of PAF by using 1–6 correlation coefficients, time-domain measurements, frequency-domain measurements, and *P* waveform and spectral densities, and they found that the power spectral densities and P wave characteristics of RR intervals have distinctive features. In another study, Chesnokov [10] found that the spectral components increase statistically before PAF event while the sample entropy and approximate entropy values decrease. Mohebbi and Ghassemian [11] have used spectral, bispectral and nonlinear measurements obtained from 30-min heart rate variability (HRV) data. Their results indicate that the spectral powers in the LF and HF bands increase before PAF event. In bispectral measurements, phase couplings were observed in data distant from a PAF event, while the phase couplings decrease as the PAF event approaches. They noted that the Poincare measurements might be a critical PAF event indicator [11]. Boon et al. [12,13] investigated the 5-min, 10-min, 15-min, 20-min and 30-min segments of free ectopic data and noted that the performance decreases toward the shorter segments than the 30-min segments.

Another widely used method for predicting PAF in the literature is to examine *P*-wave s on Electrocardiography (ECG) [16–19]. In these studies, researchers have used *P*-wave duration, amplitude, *P*-wave change, spectral power intensities of *P*-wave change and non-linear measurements of *P*-wave. In particular, Alcaraz et al. [17] and Artuno Martinez et al. [18] have shown that *P*-wave is an effective for predicting PAF events beforehand on one-hour data segments.

In this study, we attempt to construct a complex expert system to predict better the onset of the PAF events based on the HRV so that patients can take precautions to prevent PAF events. For this aim, we use a free and open data from Physionet, Atrial Fibrillation Prediction Database (AFPDB). The database contains 30-minute ECG datasets from 49 normal subjects, 25 PAF patients having a PAF event just after recording the data, 25 PAF patients having no immediate experience after recording the data. Each dataset is divided into 5-minute data segments and then HRV features for each segment are calculated from time-domain measures and frequency-domain measures using power spectral density estimations of fast Fourier transform, Lomb–Scargle, and wavelet transform methods. Different combinations of these HRV features are selected by Genetic Algorithm (GA) and then applied to *k*-nearest Neighbors classification algorithm. Then, we compute the classifier performances by the 10-fold cross-validation method.

2. Materials and methods

2.1. Data

We use the Atrial Fibrillation Prediction Database (AFPDB), which is free and open to all researchers on Phsionet.org website [20]. All ECG datasets in the database are sampled by the sampling rate of 128 Hz, digitized by a resolution of 12 bits and included 30-minute ECG records. The database consists of two parts: 50 datasets from normal subjects and 50 datasets from patients with PAF. The PAF datasets are also divided into two parts: (1) 25 data just before a PAF event and (2) 25 data with no PAF events 45 min before or after the recording. The general representation of the data is given in Fig. 1.

Nonetheless, the dataset numbered 'n27' among normal datasets is excluded from the study because it has excessive noise. This case has been reported in similar studies [21]. As a result, in this study, 49 datasets from normal subjects, 25 datasets from PAF patients with an event and 25 datasets from PAF patients with no near event are included.

2.1.1. Data segmentation

Standards of HRV analysis were determined by the Task Force group in 1996, a 5-minute recording period is recommended for short-term HRV analysis and a 24-hour recording period is recommended for long-term HRV analysis [22]. In order to determine more precise time before the PAF event, short-term (5-minute) HRV analysis is preferred. Based on this approach, all 30-min segments of data were divided into 10 5-minute segments with 50% overlap as shown in Fig. 2.

For all the HRV data segments, we calculate the HRV features and the classifier performances.