A Method for Arrhythmic Episode Classification in ECGs Using Fuzzy Logic and Markov Models

MG Tsipouras[1,2], Y Goletsis[1,4], DI Fotiadis[1,2,3]

[1]Unit of Medical Technology and Intelligent Information Systems, Dept of Computer Science, University of Ioannina, GR 45110 Ioannina, Greece
[2]Biomedical Research Institute – FORTH, GR 45110 Ioannina, Greece
[3]Michaelidion Cardiology Center, GR 45110 Ioannina, Greece
[4]Dept of Economics, University of Ioannina, GR 45110 Ioannina, Greece

Abstract

A method for arrhythmic episode classification using only the RR-interval signal is presented. The method is based on fuzzy logic and Markov models, while classification is performed for nine categories of cardiac rhythms. A two-stage classifier is applied. In the first stage, a fuzzy system classifies the episode using the mean value and standard deviation of the RR-intervals. In the second, the RR-interval signal is transformed to character sequences, which are classified by Markov models. Two representation techniques are used for the extraction of the character sequences: symbolic dynamics and one based on the RR-interval length. The classification of an episode is achieved combining the outcomes of the two stages. The MIT-BIH arrhythmia database is used for the evaluation of the proposed method. The obtained results indicate high performance (accuracy 73%) in arrhythmic episode classification.

1. Introduction

Automatic classification of cardiac rhythms remains a vital problem in clinical cardiology, especially when it is performed in real time. Several researchers have addressed the problem of automatic classification of cardiac arrhythmias [1-6]. However, most of the approaches proposed in the literature deal with a limited number of arrhythmic types and process the entire ECG signal extracting several features from it, such as the P wave, which is an extremely time-consuming process and sometimes difficult due to the presence of noise.

This work proposes a method for the classification of nine cardiac rhythms (presented in Table 1), based on fuzzy logic and Markov models, using only the RR-interval signal, which can be easily and accurately extracted from the ECG. Two separate classifications are proposed, one based on fuzzy logic analysis, using the mean value and the standard deviation of the RR-intervals in an episode and a second one based on with Markov models using character sequences produced from the RR-interval signal using methods of symbolic representation. The results obtained by the two approaches are combined to achieve the final arrhythmia classification.

2. Methods

The proposed method consists of three steps: (a) fuzzy logic analysis, (b) symbolic representation and Markov model processing and (c) arrhythmic episode classification. The MIT-BIH arrhythmia database [7] is used for training and evaluation of the method.

2.1. Dataset

All records from the MIT-BIH arrhythmia database are included in the dataset used in our analysis. QRS detection [8] was applied to the ECG recordings to produce the RR-interval signals. However, the information carried by the RR-interval signal is insufficient to distinguish between certain types of cardiac rhythms, included in the database, therefore they are assumed to belong to a single rhythm category: normal sinus rhythm, paced rhythm, pre-excitation and sinus bradycardia are all classified as normal sinus rhythm, ventricular bigeminy and atrial bigeminy are classified as bigeminy, ventricular tachycardia and supraventricular tachyarrhythmia are classified as tachycardia, atrial flutter and atrial fibrillation are classified as atrial flutter/fibrillation. Table 1 presents the cardiac rhythms (categories) classified by the proposed method, the annotation used, the corresponding rhythm annotations of the MIT-BIH arrhythmia database and the number of episodes included in our dataset for each cardiac rhythm. The dataset was randomly divided into a training and a testing set, selecting half of the episodes from each rhythm category for training, while the remaining episodes were used for testing.

2.2. Fuzzy logic analysis

Fuzzy logic analysis is used for the first level classification of the cardiac rhythms. For every rhythm category the mean value and the standard deviation of the RR-intervals of each episode included in the category (only in the training set) are calculated. Then, the parameters M_i and S_i are calculated for each rhythm category i, as follows:

$$M_i = \frac{1}{N_i}\sum_{j=1}^{N_i} mean_j \ , \tag{1}$$

$$S_i = \frac{1}{N_i}\sum_{j=1}^{N_i} stdev_j \ , \tag{2}$$

where N_i is the number of episodes included in the i^{th} rhythm category, $mean_j$ is the mean value and $stdev_j$ is the standard deviation of the j^{th} episode. Also $min(M_i)$, $M_i - stdev(M_i)$, $M_i + stdev(M_i)$, $max(M_i)$, $min(S_i)$, $S_i - stdev(S_i)$, $S_i + stdev(S_i)$ and $max(S_i)$ are calculated. These values are used to create two fuzzy membership functions. Having an episode E to be classified, Eq. (3) presents the fuzzy membership function according to the mean value M of the RR-intervals of E. w_i^M expresses the similarity (membership) of E to the category i based on M. In a similar manner w_i^S, which is the membership degree of E in the i^{th} rhythm category based on the standard deviation S, is calculated ($w_i^M, w_i^S \in [0,1]$). The classification, from the fuzzy system, is then achieved computing $w_i = w_i^M * w_i^S$, where w_i is a measure of the membership of E in the i^{th} rhythm category.

Table 1. Cardiac rhythms classified by the proposed method.

Cardiac rhythms	Anno-tation	MIT-BIH annotation	Number of episodes
Normal sinus rhythm	N	N, P, PREX, SBR	695
Idioventricular rhythm	IV	IVR	4
Nodal rhythm	ND	NOD	36
2° heart block	BII	BII	5
Trigeminy (ventricular or atrial)	TG	T	83
Bigeminy (ventricular or atrial)	BG	B, AB	224
Tachycardia (ventricular or supraventricular)	TC	VT, SVTA	87
Ventricular flutter/fibrillation	VF	VFL	6
Atrial flutter/fibrillation	AF	AFL, AFIB	123

2.3. Methods

Symbolic representation is used to produce sequences of characters (symbols) from the RR-interval signal. Two techniques are used. The first technique is symbolic dynamics (SD), a non-linear technique, used to measure the complexity of the signal [9]. SD transforms the RR-interval signal into a sequence of symbols D_j, reducing the samples of the signal to few symbols, and therefore, simplifying the modelling of its dynamical behaviour. It is based on the transformation presented in Eq. 4, where RR_j is the duration of the j^{th} RR-interval, μ is the mean

$$w_i^M = \begin{cases} 0, & M < min(M_i) \\ \dfrac{M - min(M_i)}{(mean(M_i) - stdev(M_i)) - min(M_i)}, & min(M_i) \le M < mean(M_i) - stdev(M_i) \\ 1, & mean(M_i) - stdev(M_i) \le M \le mean(M_i) + stdev(M_i) \\ \dfrac{max(M_i) - M}{max(M_i) - (mean(M_i) + stdev(M_i))}, & mean(M_i) + stdev(M_i) < M \le max(M_i) \\ 0, & max(M_i) < M \end{cases} \tag{3}$$

$$D_j = \begin{cases} a & 0 < RR_j \le (1-a)\mu \\ b & (1-a)\mu < RR_j \le \mu \\ c & \mu < RR_j \le (1+a)\mu \\ d & (1+a)\mu < RR_j \end{cases}, \quad (4)$$

value of the RR-interval signal and a is a parameter quantifying the standard deviation of the RR-interval signal, set to 0.05. The alphabet consists of four symbols $\{a,b,c,d\}$.

The second approach transforms the RR-interval signal into a sequence of symbols, L_j, and is based only on the RR-interval duration:

$$L_j = \begin{cases} a & RR_j \le 0.3 \\ b & 0.3 < RR_j \le 0.4 \\ c & 0.4 < RR_j \le 0.5 \\ \dots & \dots \\ y & 2.6 < RR_j \le 2.7 \\ z & 2.7 < RR_j \end{cases}, \quad (5)$$

where RR_j is the duration of the j^{th} RR-interval. The alphabet used consists of 26 symbols $\{a,b,c,\dots,y,z\}$.

Two sets of Markov models are created, one using the B_D symbol sequences (i.e. using symbolic dynamics for transformation of the RR-interval signal into symbol sequence) and one using the B_L symbol sequences (i.e. using the RR-interval length for transformation of the RR-interval signal into symbol sequence). The Markov models are crated using the episodes of the training set, measuring the a-priori transition probability from each symbol to all others, given a specific rhythm category:

$$p^i_{s_1s_2} = \frac{\text{\# of transitions from } s_1 \text{ to } s_2 \text{ in category } i}{\text{total \# of transitions in category } i}. \quad (6)$$

This corresponds to i $M \times M$ matrices, where i is the number of rhythm categories (in our case nine) and M is the number of symbols used (4 for B_D and 26 for B_L). The probability to produce a given symbol sequence B (B_D or B_L) from model i (i.e. belongs to i rhythm category) is defined as:

$$p^i_D(B_D) = \prod_{j=1}^{l_B-1} p^i_{s_j s_{j+1}} \text{ for the } D_j \text{ symbol sequences} \quad (7)$$

$$p^i_L(B_L) = \prod_{j=1}^{l_B-1} p^i_{s_j s_{j+1}} \text{ for the } L_j \text{ symbol sequences} \quad (8)$$

where l_B is the length of B (number of symbols).

2.4. Arrhythmic episode classification

Arrhythmic episode classification is performed as follows:

1. For a new episode E, w_i, $i = 1,\dots,9$ are calculated.
2. The RR-intervals of episode E are transformed into a sequence of symbols using one of the proposed representations (B_D or B_L).
3. The probabilities $p^i_D(B_D)$ or $p^i_L(B_L)$, $i = 1,\dots,9$ are calculated.
4. The episode is classified into the category c_D or c_L according to the representation selected:

$$c_D = \max_{i=1,\dots,9}(w_i * p^i_D(B_D)), \quad (9)$$

$$c_L = \max_{i=1,\dots,9}(w_i * p^i_L(B_L)). \quad (10)$$

3. Results

Both representation techniques were tested. The results are shown in Table 2, in terms of sensitivity, specificity and positive predictive value for all nine rhythm categories. The results indicate that the proposed method performs well in classifying cardiac rhythms. The accuracy of the method is 72.2% for the first symbolic representation (SD) and 72.7% for the second one.

4. Discussion and conclusions

The proposed method classifies arrhythmic episodes into nine rhythm categories through a two-stage classification: (a) fuzzy logic analysis and (b) Markov models. The obtained results indicate high performance in arrhythmic episode classification. Selected approaches in the literature indicate better results [2-6], which, however, are justified due to: (i) the application in very small datasets compared to the 24 hours of ECG recordings included in the MIT-BIH arrhythmia database and (ii) the classification in a maximum of four (4) categories instead of nine (9). Therefore, the results are good enough considering the multi-class nature of the classification problem.

A thorough review of the episode classifications indicates that: (a) there is high correlation between the normal category (N) and categories ND, TG and AF, for both symbolic representations, therefore there is a decrease in the positive prediction value for these categories; (b) although the accuracy is almost identical

Table 2. Classification results for cardiac rhythms.

Rhythm category	Symbolic dynamics: Sensitivity %	Symbolic dynamics: Specificity %	Symbolic dynamics: Positive predictive value %	RR-interval length: Sensitivity %	RR-interval length: Specificity %	RR-interval length: Positive predictive value %
N	67.8	93.4	92.5	72.4	92.0	91.6
IV	100	79.1	50.0	100	81.2	66.7
ND	72.2	78.9	24.1	77.8	81.3	36.8
BII	66.7	85.4	100	100	85.4	100
TG	78.6	84.3	46.8	69.0	84.5	67.4
BG	83.9	88.9	90.4	75.9	83.3	92.4
TC	81.8	88.1	65.5	77.3	81.4	64.2
VF	66.7	88.7	66.7	66.7	83.0	50.0
AF	64.5	91.6	45.5	64.5	85.5	32.5
Accuracy %		72.2			72.7	

for both B_D and B_L sequences the sensitivity and specificity for each rhythm category varies. Further testing of the proposed method is necessary to identify which of the two symbolic representations results in higher performance.

An efficient method for arrhythmic episode classification in ECGs is proposed. The method is based on the analysis of the RR-interval signal using fuzzy logic and Markov models. The method is advantageous compared to methods presented in the literature because: (a) it provides classification for a large number of cardiac rhythms (nine); (b) it uses only the RR-interval signal, which can be extracted with high accuracy even for noisy ECG recordings, while the extraction of all other ECG features or any other type of ECG analysis is seriously affected by noise; (c) it performs in real time. The proposed method classifies pre-determined arrhythmic episodes (i.e. the start and the end of the episodes are determined). Therefore, an extension of the present work must include automatic detection of episodes.

References

[1] Ham FM, Han S. Classification of cardiac arrhythmias using fuzzy ARTMAP. IEEE Trans Biom Eng 1996;43:425-430.

[2] Chen SW, Clarkson PM, Fan Q. A robust sequential detection algorithm for cardiac arrhythmia classification. IEEE Trans Biom Eng 1996;43:1120-1125.

[3] Khadra L, Al-Fahoum AS, Al-Nashash H. Detection of life-threatening cardiac arrhythmias using wavelet transformation. Med Biol Eng Comp 1997;35:626-632.

[4] Zhang XS, Zhu YS, Thakor NV, Wang ZZ. Detecting ventricular tachycardia and fibrillation by complexity measure. IEEE Trans Biom Eng 1999;46:548-555.

[5] Al-Fahoum AS, Howitt I. Combined wavelet transformation and radial basis neural networks for classifying life-threatening cardiac arrhythmias. Med Biol Eng Comp 1999;37:566-573.

[6] Owis MI, Abou-Zied AH, Youssef AM, Kadah YM. Study of features based on nonlinear dynamical modeling in ECG arrhythmia detection and classification. IEEE Trans Biom Eng 2002;49:733-736.

[7] MIT-BIH Arrhythmia Database CD-ROM. Third Edition, 1997, Harvard-MIT Division of Health Sciences and Technology.

[8] Hamilton PS, Tompkins WJ. Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. IEEE Trans Biom Eng 1986;33:1157-1165.

[9] Vallverdu M, Palacios M, Caminal P. Modeling the dynamics of the heart rate variability by hidden markov models. In: Murray A, editor. Computers in Cardiology 2003. IEEE, 2003;30:461-464.

Address for correspondence.

Dimitrios I. Fotiadis
Unit of Medical Technology and Intelligent Information Systems, Dept. of Computer Science, University of Ioannina Campus, P.O. BOX 1186, GR 45110 Ioannina, Greece.
fotiadis@cs.uoi.gr