

Atrial Fibrillation Screening Through Combined Timing Features of Short Single-Lead Electrocardiograms

Manuel García¹, Juan Ródenas¹, Raúl Alcaraz¹, José J Rieta²

¹ Research Group in Electronic, Biomedical and Telecomm. Eng., Univ. of Castilla-La Mancha, Spain

² BioMIT.org, Electronic Engineering Department, Universitat Politècnica de Valencia, Spain

Abstract

Atrial fibrillation (AF) is the most common cardiac arrhythmia, as well as a growing healthcare burden worldwide. It is often asymptomatic and usually starts with very brief episodes, thus making its early detection an interesting challenge. For that purpose, the present work introduces a novel method exploiting the variability presented both by ventricular and atrial activities reflected on the surface electrocardiogram (ECG). Thus, time series from the RR intervals and the fibrillatory waves morphology contained by the TQ intervals are first generated and, then, their regularity is estimated making use of the Coefficient of Sample Entropy (COSEn). The collected information is finally combined through a multi-class support vector machine (SVM) approach to discern among short episodes of AF, normal sinus rhythm (NSR) and other rhythms (OR). The algorithm has been validated in the context of the PhysioNet Computing in Cardiology Challenge 2017, thus reporting a global F_1 measure of 0.73 for the training set and 0.71 for the testing group. Nonetheless, to evaluate the method in a common scenario for previous works, the widely used MIT-BIH AF database has also been considered. A notably higher F_1 score of 0.87 has been provided in this case. The significantly different balance between the number of AF, NSR and OR recordings in both databases could justify the obtained outcomes.

1. Introduction

Atrial Fibrillation (AF) is the most common type of arrhythmia in the general population and is becoming a major cardiovascular challenge in modern world [1]. It is defined as a “tachyarrhythmia characterized by predominantly uncoordinated atrial activation with consequent deterioration of atrial mechanical function” by the American College of Cardiology, the American Heart Association and the European Society of Cardiology [2]. However, the pathophysiological mechanisms causing and maintaining this abnormal atrial activation are still not fully known,

thus making its diagnosis and therapy highly complex and poorly efficient [2]. Indeed, the management of AF nowadays involves a significant part of the resources from any Health Service in the developed countries [3]. Therefore, reducing the prevalence of this arrhythmia is a current priority [4], being its early detection a key aspect to avoid its natural progression and final perpetuation [5].

However, given that AF usually starts with very brief episodes and is often asymptomatic [6], its first signs are not easy to find [5]. For this purpose, a variety of automatic algorithms dealing with the surface electrocardiogram (ECG) have been proposed in the last years. Most of them exploit the two alterations provoked by the arrhythmia on the ECG. On the one hand, AF is characterized by a rapid atrial activity, whose rates can vary between 240 and 540 activations/min [2]. Because the atrioventricular (AV) node is permanently shelled by all of them, a fast and irregular ventricular response can be observed on the ECG. Clearly, this behavior contrasts with the completely regular pattern found in the RR interval series during normal sinus rhythm (NSR). On the other hand, the uniform P-wave related to atrial depolarization during NSR is also replaced by a series of low amplitude and variable morphology fibrillatory waves (f -waves) [2].

The main goal of this work is to quantify and combine information from the variability presented both by the RR intervals and the f -waves to discern between short ECG recordings of AF, NSR and other rhythms (OR) in the context of the PhysioNet Computing in Cardiology Challenge 2017. More precisely, the Coefficient of Sample Entropy (COSEn) is used to compute the regularity of the time series comprised from the RR intervals as well as from the morphological variations of the TQ intervals.

2. Materials

To validate the proposed detector of AF, the freely available database for the PhysioNet Computing in Cardiology Challenge 2017 was used [7]. It contains 8528 single-lead ECGs lasting from 9 to 60 seconds for training and 3658 recordings of similar lengths for testing. All the sig-

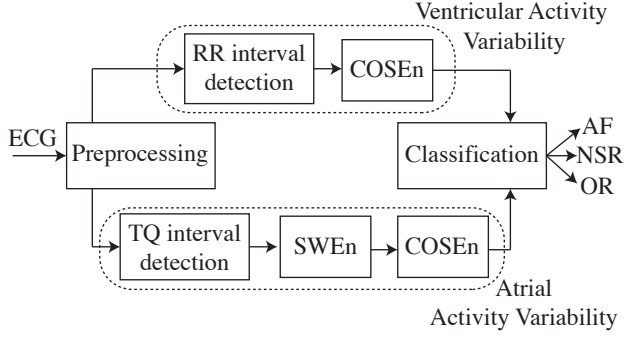


Figure 1. Block diagram describing the main steps for the proposed algorithm to discern among short ECG recordings of AF, NSR and OR.

nals were sampled at 300 Hz and were band-pass filtered. Whereas the test set is private for the purpose of scoring during the Challenge, the learning set is public and contains 5050, 738 and 2456 episodes of NSR, AF and OR, respectively, as well as 284 noisy intervals.

On the other hand, because the MIT-BIH AF database has been widely used to test most of the previously published AF detectors [8], it has also been considered in the study. This dataset is freely available at PhysioNet [9] and contains 23 fully-annotated 10 hour-length ECG recordings, mainly obtained from paroxysmal AF patients. However, two signals (*04936* and *05091*) were discarded because they were mistakently annotated [8]. Nonetheless, more than 1 million of beats completely balanced from NSR and AF were analyzed. Additionally, about 15,000 beats of OR from atrial flutter and AV junctional rhythm were also considered. Note that the recordings were acquired with a sampling rate of 250 Hz, but they were up-sampled to 300 Hz.

3. Methods

The proposed algorithm to detect automatically short episodes of AF, NSR and OR is graphically summarized in Fig. 1. As can be seen, the ECG is firstly preprocessed and, then, the variability both from RR and TQ intervals is computed through parallel ways. Finally, the information gained from both cases is combined in a classification step to assign a potential class to the signal. More details about each step are provided next.

3.1. Data preprocessing

Every signal was typically preprocessed to reduce noise and any nuisance interference. Thus, the baseline wander was removed making use of an IIR high-pass filtering with a 0.5 Hz cut-off frequency. Moreover, high frequency noise and powerline interference were reduced with an IIR

low-pass filtering with a 50 Hz cut-off frequency. Both filters were designed using a Chebyshev window with a relative sidelobe attenuation of 40 dB and applied to the ECG signal in a forward/backward fashion. Thereafter, the R-peaks were detected by using a phasor transform-based algorithm [10]. Remark that this algorithm has proven an ability higher than 99% to identify every R-peak regardless of the beat morphology [10]. Thus, ventricular ectopic beats were also identified and considered for the study.

3.2. Ventricular activity variability

The RR intervals were computed from the difference between consecutive R-peaks and, then, their variability was computed making use of the COSEn. This index was proposed some years ago by Lake & Moorman to discern between AF and OR from short ECG intervals, obtaining an accuracy about 90% [11]. Mathematically speaking, it can be defined as

$$\text{COSEn} = -\ln\left(\frac{A}{B}\right) + \ln(2r) - \ln(\mu), \quad (1)$$

where A and B denote the total number of RR patterns of length $m + 1$ and m , respectively, that match within a certain tolerance r . The parameter μ is the mean length of the analyzed RR intervals, thus providing information about the average heart rate. Note that COSEn was here computed making use of the parameters specified by Lake & Moorman to optimize its performance [11].

3.3. Atrial activity variability

To characterize the f -waves morphology, every TQ interval was firstly detected by using the subsequent R-peak as a reference and, then, characterized into the wavelet domain, such as in some previous works [12, 13]. More precisely, the Stationary Wavelet Entropy (SWEn) was obtained by decomposing the signal into different time-frequency scales and computing the common Shannon Entropy from their relative energy distributions [12]. In this way, organized waveforms (such as the P-waves) provide low values of SWEn and, contrarily, complex morphologies (such as the f -waves) high values. Note that the parameters used to compute the SWEn were 4 levels of decomposition and a fourth-order Daubechies function as mother wavelet. Finally, the variability of the obtained time series was also computed through the COSEn. In this case, after some experimental tests, the used parameters were $m = 2$ and $r = 0.35$.

3.4. Classification

The information obtained from the variability both for atrial and ventricular components was finally combined by

means of a multi-class support vector machine (SVM) to classify each ECG recording. Given that three classes were available, i.e. AF, NSR and OR, the Crammer and Singer approach was used to optimize simultaneously the two required binary classifications, thus maximizing the margins from each class to the remaining ones [14]. This classifier was selected due to its interesting ability of generalization, even when unbalanced datasets are analyzed [15].

3.5. Performance assessment

According to the specific rules for the PhysioNet Computing in Cardiology Challenge 2017, the discriminant ability of the proposed algorithm has been summarized with the average of the F_1 measures obtained for each group, i.e.

$$F_1 = \frac{F_{1\text{NSR}} + F_{1\text{AF}} + F_{1\text{OR}}}{3}. \quad (2)$$

The F_1 score is a measure of accuracy which combines both sensitivity and positive predictivity. Thus, for a specific class c , it can be defined as

$$F_{1c} = \frac{2 \cdot \text{TP}_c}{2 \cdot \text{TP}_c + \text{FP}_c + \text{FN}_c}, \quad (3)$$

where TP_c is the number of true positives, FP_c is the number of false positives and FN_c is the number of false negatives. Note that c could be NSR, AF or OR in this study.

4. Results

Although the ECG recordings proposed for the PhysioNet Computing in Cardiology Challenge 2017 presented different lengths, they were completely analyzed without any kind of segmentation. Thus, the F_1 measures obtained from the training and testing sets are displayed in Table 1. As can be observed, very similar outcomes were noticed for both cases, thus suggesting a successful training of the proposed method. Moreover, it is also surprising that the F_1 measures for the OR group were notably lower than for the remaining ones.

On the other hand, given that long episodes of NSR, AF and OR can be found in the MIT-BIH AF database, 15 beat-length ECG intervals were analyzed with the proposed algorithm. In this case, a global F_1 value of 0.87 was obtained, $F_{1\text{NSR}}$, $F_{1\text{AF}}$ and $F_{1\text{OR}}$ being 0.96, 0.86 and 0.79, respectively.

5. Discussion and conclusions

The classification results obtained for the two analyzed databases have shown a similar behavior among groups. Thus, for both cases the NSR and OR episodes were classified with the highest and lowest accuracies, respectively.

Table 1. Classification results obtained by the proposed methodology from the training and testing sets of the PhysioNet Computing in Cardiology Challenge 2017 database.

Set	F_1	$F_{1\text{NSR}}$	$F_{1\text{AF}}$	$F_{1\text{OR}}$
Training	0.73	0.85	0.76	0.60
Testing	0.71	0.85	0.76	0.52

However, the absolute values of F_1 measures have been notably higher for the MIT-BIH AF database than for the dataset proposed for the PhysioNet Computing in Cardiology Challenge 2017. In our view, the relevant differences characterizing both databases could explain this outcome.

On the one hand, whereas the MIT-BIH AF presents a balanced number of beats in AF and NSR, the dataset for the Challenge 2017 shows a significantly higher number of NSR recordings. This aspect could somewhat bias the classification, since a good global result could be obtained if a classifier is overtrained to identify NSR. Unfortunately, many AF episodes would then be lost and the algorithm would be far from reaching its main goal. Maybe, a weighted F_1 score would be a fairer measure of diagnostic accuracy and could provide a more comparable classification outcome among different databases, regardless of their balance between NSR, AF and OR recordings.

On the other hand, while ECG segments with the same length have been considered from the MIT-BIH AF database, signals with a duration ranging from 9 to 60 seconds have been analyzed from the Challenge dataset. Clearly, the shorter the recording length, the lower the available information for its classification, thus making this task more difficult [11]. In fact, many previous works have shown that proper identification of AF on ECG recordings shorter than about 10 beats is highly complex [8, 11, 13]. Therefore, the shortest signals in the Challenge database, which could not even contain 10 beats, render more challenging the detection of the different rhythms.

Finally, another relevant confounding factor is the OR group composition for both datasets. To this respect, only a few beats from atrial flutter and AV junctional rhythm were considered from the MIT-BIH AF database, because no more rhythms were found in the recordings. Contrarily, information about which rhythms were included in the OR group of the Challenge database is completely unknown. Moreover, after reviewing carefully the signals for this group, numerous examples of episodes very close to SNR or AF have been found labelled as OR. A couple of examples in this line are shown in Fig. 2. Within this context, it is very difficult to develop an AF detector only based on physiological observations. It is highly probable that an advanced classifier fed with dozens or hundreds of vari-

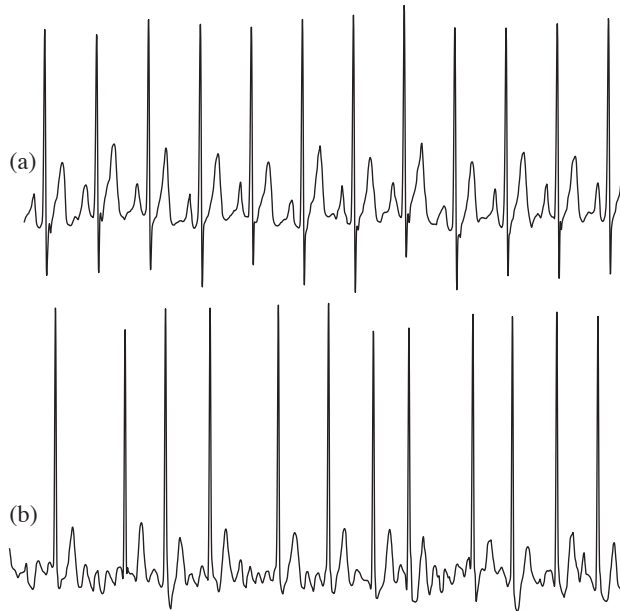


Figure 2. ECG segments extracted from signals included in the OR group but presenting rhythms very close to (a) NSR (recording A03496) or (b) AF (recording A00326).

ables extracted from the ECG signals will obtain a good classification result with this database, but even so, it is reasonable to think that the Challenge database is still not sufficiently prepared to tackle the problem of the automatic detection of AF.

Acknowledgements

Research supported by grants TEC2014–52250–R and DPI2017–83952–C3 MINECO/AEI/FEDER, UE.

References

- [1] Chugh SS, Roth GA, Gillum RF, Mensah GA. Global burden of atrial fibrillation in developed and developing nations. *Glob Heart* Mar 2014;9(1):113–9.
- [2] Kirchhof P, Benussi S, Kotecha D, Ahlsson A, Atar D, Casadei B, Castella M, Diener HC, et al. 2016 ESC Guidelines for the management of atrial fibrillation developed in collaboration with EACTS. The Task Force for the management of atrial fibrillation of the European Society of Cardiology (ESC). Developed with the special contribution of the European Heart Rhythm Association (EHRA) of the ESC. *Europace* November 2016;18(11):1609–1678.
- [3] Sheikh A, Patel NJ, Nalluri N, Agnihotri K, Spagnola J, et al. Trends in hospitalization for atrial fibrillation: epidemiology, cost, and implications for the future. *Prog Cardiovasc Dis* 2015;58(2):105–16.
- [4] Van Wagoner DR, Piccini JP, Albert CM, Anderson ME, Benjamin EJ, et al. Progress toward the prevention and treatment of atrial fibrillation: A summary of the Heart Rhythm Society Research Forum on the Treatment and Prevention of Atrial Fibrillation, Washington, DC, December 9–10, 2013. *Heart Rhythm* Jan 2015;12(1):e5–e29.
- [5] Amara W, Montagnier C, Cheggour S, Boursier M, Gully C, Barnay C, Georger F, Deplagne A, Fromentin S, Mlotek M, Lazarus A, Taïeb J, SETAM Investigators. Early detection and treatment of atrial arrhythmias alleviates the arrhythmic burden in paced patients: The SETAM study. *Pacing Clin Electrophysiol* May 2017;40(5):527–536.
- [6] Stachon P, Ahrens I, Faber T, Bode C, Zirlik A. Asymptomatic atrial fibrillation and risk of stroke. *Panminerva Med* Dec 2015;57(4):211–5.
- [7] Clifford G, Liu C, Moody B, Lehman L, Silva I, Li Q, Johnson A, Mark R. AF classification from a short single lead ECG recording: The Physionet Computing in Cardiology Challenge 2017. *Computing in Cardiology 2017*;
- [8] Zhou X, Ding H, Wu W, Zhang Y. A real-time atrial fibrillation detection algorithm based on the instantaneous state of heart rate. *PLoS One* 2015;10(9):e0136544.
- [9] Goldberger AL, Amaral LA, Glass L, Hausdorff JM, Ivanov PC, Mark RG, Mietus JE, Moody GB, Peng CK, Stanley HE. PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals. *Circulation* Jun 2000;101(23):E215–20.
- [10] Martínez A, Alcaraz R, Rieta JJ. Application of the phasor transform for automatic delineation of single-lead ECG fiducial points. *Physiol Meas* Nov 2010;31(11):1467–85.
- [11] Lake DE, Moorman JR. Accurate estimation of entropy in very short physiological time series: the problem of atrial fibrillation detection in implanted ventricular devices. *Am J Physiol Heart Circ Physiol* Jan 2011;300(1):H319–25.
- [12] García M, Ródenas J, Alcaraz R, Rieta JJ. Application of the relative wavelet energy to heart rate independent detection of atrial fibrillation. *Comput Methods Programs Biomed* Jul 2016;131:157–68.
- [13] Ródenas J, García M, Alcaraz R, Rieta JJ. Wavelet entropy automatically detects episodes of atrial fibrillation from single-lead electrocardiograms. *Entropy* 2015;17:6179–6199.
- [14] Crammer K, Singer Y. On the algorithmic implementation of multiclass kernel-based vector machines. *J Mach Learn Res* 2001;2:265–292.
- [15] Eitrich T, Lang B. Efficient optimization of support vector machine learning parameters for unbalanced datasets. *J Comput Appl Math* 2006;196:425–436.

Address for correspondence:

Manuel García Teruel
 E.I. Industriales, Campus Universitario, 02071, Albacete, Spain.
 Phone: +34–967–599–200 Ext. 2488
 Fax: +34–967–599–224
 e-mail: manuel.garcia@uclm.es