Detection of AF and Other Rhythms Using RR Variability and ECG Spectral Measures

Lucia Billeci[1], Franco Chiarugi[2], Magda Costi[3], David Lombardi[3], Maurizio Varanini[1]

[1] Institute of Clinical Physiology, National Research Council of Italy (IFC-CNR), Pisa, Italy
[2] Institute of Computer Science, FORTH, Heraklion, Crete, Greece
[3] Cardioline S.p.A., Trento, Italy

Abstract

Aims: Atrial fibrillation (AF) is one of the principal cause of mortality in elderly, thus its detection is extremely clinically relevant. The aim of this study was to classify short, single lead, ECG recordings, as atrial fibrillation, normal sinus rhythm, other type of rhythms or noisy signal. Methods: First, we extracted, both from the ECG signals and from the RR interval series, about fifty features characterizing these four classes. Then, we applied the stepwise linear discriminant analysis for dimensionality reduction selecting a subset of thirty discriminating features. A Least Squares Support Vector Machine (LS-SVM) classifier using these features was tuned and trained on the dataset of the Physionet/Computing in Cardiology Challenge 2017. Results: The LS-SVM classifier provided, on the hidden test set of the Challenge, an official final score F1= 0.81, obtaining the twelfth place in the ranking of results with only 2 cents from the best (0.83). Conclusions: This approach seems promising in particular in detecting atrial fibrillation. Further work is needed to improve the discrimination of other rhythms and noisy signals.

1. Introduction

Atrial fibrillation (AF) is the most common cardiac arrhythmia and cause of mortality in elderly. It is characterized by a disorganized electrical activity in the atria and rapid circulating waves of abnormal electrical signals continuously stimulate the atrium instead of the sinus node normally stimulating the atrium [1].

Different algorithms have been designed for detection and classification of AF. They are based on time-frequency analysis of ECG [2], RR intervals analysis [3,4], short term analysis of heart rate variability (HRV) [5] and sequential analysis to check the absence of P wave [6, 7]. Algorithms based on P-waves performs poorly in presence of noise as these waves are prone to contamination with motion and noise artefacts [8]. Thus, most of the more recent approaches for AF detection are based on RR analysis. To improve accuracy and specificity in AF detection, several neural networks approaches have been recently proposed [9–12]. In particular, Support Vector Machine (SVM) classifier has been commonly employed, has it gives promising results in various medical diagnostics [13]. Most of the algorithms implemented aim at distinguish AF from both normal sinus rhythm. However, it is also important discriminating AF from rhythms with frequent ectopic beats and noise. Indeed, these signals can confound usual strategies for numerical algorithms to detect AF [14].

The aim of this study was to propose an approach for classification short ECG recording as AF, normal rhythms, other type of rhythms and noise, by applying LS-SVM to features extracted both from the ECG signals and from the RR interval series. This algorithm was trained and tested on Physionet/Computing in Cardiology Challenge 2017 database [15]. The code was submitted for the Open-Source Challenge call.

2. Methods

2.1. ECG processing

Artifact canceling was obtained by comparison of ECG with a median filtered signal (60ms window) [16]: the ECG values whose absolute difference from the filtered ones exceeded a threshold were replaced with the average of the values before and after them. Baseline wander was estimated applying a linear phase low pass filter with cut-off frequency at 3 Hz and detrended signal was obtained as difference. The resulting signal was then upsampled to 1200 Hz to allow a better localization of QRSs.

QRS detection was performed by a threshold on the absolute amplitude of a filtered derivative signal. This threshold was updated at each new detection and was changed with the temporal distance from the previous QRS. The fiducial point of each QRS was selected as the time occurrence of the maximum (minimum) of the signed derivative

 ISSN: 2325-887X DOI:10.22489/CinC.2017.344-144

signal. The beginning and the end of QRS were estimated by the crossing of derivative through 0.25 threshold.

AF may manifest on the ECG signal as a weak oscillation with frequency in the range 2-10Hz. This oscillatory component may be hidden, in the ECG spectrum, by the high QRST power, therefore a QRST cancelling procedure was applied. This was based on approximating each QRST by Singular Value Decomposition (SVD) method. The signal around each QRS was weighted by a trapezoidal window and stored in the columns of the matrix X, which was decomposed by SVD. A matrix Xr was then rebuilt from the SVD decomposition using a reduced number of eigenvectors (2 or 3). Its columns contains only the signal components which are powerful and synchronous thus they approximated the original signal interval around the QRS. A signal containing almost only the ventricular origin component was obtained by unweighting the estimated QRS segments and connecting them with a straight line. This signal was subtracted from the original ECG obtaining a residual signal where the AF component was enhanced.

2.2. Feature extraction and selection

The extracted features can be categorized in three types: 1) computed on the ECG signal, 2) derived from the RR series and 3) obtained combining QRS morphology and rhythm. ANOVA was used for feature pre-selection and tuning, then the stepwise discriminant analysis with Rao' V criterion for feature inclusion was applied to select the final feature set (30 features).

In the following, a generic description of the selected features is reported, for details see the submitted code:

1) the power of the signal obtained as the difference between the preprocessed and the original ECG signal; the spectral power and the peak in the 4-10 Hz band of the residual signal obtained by canceling the QRST complexes; the max, the mean and the standard deviation of the QRS width, obtained as the difference between the QRS offset and onset provided by the QRS detector; features extracted from the SVD of a matrix whose columns contain time aligned QRS samples 150ms before and 200 ms after the QRS reference point;

2) the mean, the min and the max value of the RR intervals; the mean value of the RR intervals below the threshold of 0.7 s and the mean value of RR intervals above the threshold of 1.2 s; the root mean square of the successive differences (RMSSD) [4]; the mean of the absolute weighted successive difference (Mawsd); the coefficient of sample entropy (CoSEn) [14]; the turning point count (TPC) [4]; the Katz Fractal Dimension (KFD) [17];

3) the ratio between the QRS amplitude and the beat prematurity where the beat prematurity was computed in two ways: ratio between the actual RR value and the trimmed mean on a moving window of 5 RR values (AtypBeatPr); ratio between the actual RR value and the successive RR value (AtypBeatPr1).

Some of the extracted measures were log-transformed to get a more symmetrical distribution.

2.3. Classification

In this study, we applied the LS-SVM classifier proposed by Suykens [18] which is derived from the Vapnik's SVM classifier [19]. The binary LS-SVM is formulated as

$$\underset{\mathbf{w},b,\xi_i}{min}\left(\|\mathbf{w}\|^2+\gamma\sum_1^n \xi_i^2\right)$$

subject to the equality constraints

$y_i\left[\mathbf{w}^T\varphi\left(\mathbf{x}_i\right)+b\right]=1-\xi_i$ for i=1,..,n

where $\mathbf{w}\in\mathbb{R}^n$ are the hyperplane coefficients, $\mathbf{x}_i$ is of the i-th feature vector, y_i is the target class, $b\in\mathbb{R}$ and $\gamma\in\mathbb{R}_0^+$. The function $\varphi:\mathbb{R}^m\rightarrow\mathbb{R}^n$ maps the feature $\mathbf{x}_i\in\mathbb{R}^m$ to the high dimensional space $\mathbb{R}^n$. The regularization parameter γ weights the sum of the squared classification error ξ^2 such that misclassifications can be tolerated. This parameter trades off classification errors versus a smooth decision surface.

LS-SVM involves a least squares cost function with equality constraints so the solution can be obtained by solving a system of linear equations in the transformed space. Defining a positive definite kernel

$k\left(\mathbf{x}_i,\mathbf{x}_j\right)=\varphi\left(\mathbf{x}_i\right)^T\varphi\left(\mathbf{x}_j\right)$

the LS-SVM classifier formulation results:

$$y\left(\mathbf{x}\right)=sign\left[\sum_{1}^{n}\alpha_i y_i K\left(\mathbf{x},\mathbf{x}_i\right)+b\right]$$

The chosen kernel function $K\left(\cdot,\cdot\right)$ was the radial basis function (RBF): $K\left(\mathbf{x},\mathbf{x}_i\right)=\exp\left[-\left\|\mathbf{x}-\mathbf{x}_i\right\|^2/\sigma^2\right]$

The multiclass categorization problem is solved by a set of binary classifiers. We chose the "one-versus-one" coding [20], consisting in a set of $m\left(m-1\right)/2$ binary classifiers, each discriminating between two classes. We used the LS-SVM toolbox [18] (LS-SVMlab) which provides a tuning function aimed at optimizing, for every binary classifier, the regularization parameter γ and the parameter σ^2 of the RBF kernel. Crossvalidation (10 fold) was applied.

3. Results

3.1. Feature discriminant power

According to the ANOVA analysis, in the multiclass comparison, the CoSEn (F=1708) and the Mawsd (F=1533) were the features with the most discriminant power.

In particular, performing every between classes ANOVA, Mawsd had the maximal F in discriminating normal and AF rhythms (F=8136) while the CoSEn was the most powerful in discriminating between AF and other rhythms (F= 2143) as shown in the histograms in Figure 1 and Figure 2. KFD was the most powerful in the discrimination between normal and other rhythms (F=1705) as observed in Figure 3. In all the figures, the areas of the histograms for each class are normalized to one, so the amplitudes estimate the class-conditional probability densities.

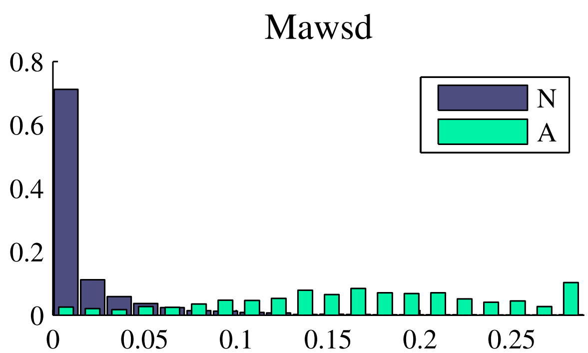

Figure 1. Histograms of Mawsd for normal (N) and atrial fibrillation (A) rhythms.

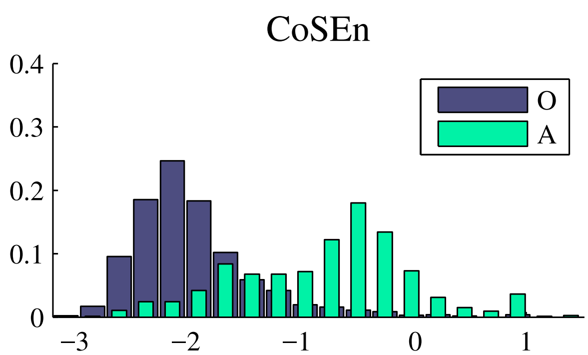

Figure 2. Histograms of CoSEn for atrial fibrillation (A) and other (O) rhythms.

3.2. Least Squares SVM

The LS-SVM classifier provided the following performance on the training set (before the final relabeling of the dataset records [15]): F1 Normal rhythm= 0.94; F1 AF rhythm: 0.91; F1 Other rhythm= 0.86; Global F1= 0.90 [15]; (F1 Noisy rec= 0.78). The confusion matrix obtained on the training set is reported in Table 1.

Our software was uploaded to the challenge website for testing. The global F1 score, on the hidden test data set, after the final relabeling, resulted 0.81 obtaining the twelfth place in the ranking of results with only 2 cents from the best (0.83).

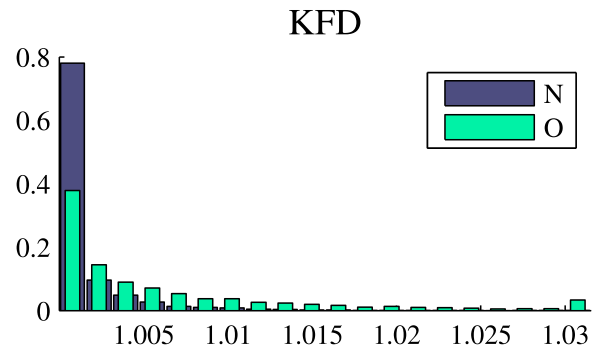

Figure 3. Histograms of KFD for normal (N) and other (O) rhythms.

Table 1. Confusion matrix on training set. N: normal, A: atrial fibrillation, O: other, P: noisy, e: estimated.

	eN	eA	eO	eP
N	4946	6	85	13
A	32	658	43	5
O	474	33	1939	10
P	66	5	13	200

4. Discussion and conclusion

In this study, we proposed an approach for discriminating between normal, AF, other rhythms and noisy ECG records. We extracted measures from the overall ECG signal, from each QRS and from the RR series for a better characterization and discrimination of the different rhythms.

The high efficiency of CoSEn in detecting AF, which is related to the increased atrial signal irregularity, confirms previous studies [12,21]. Moreover, we observed that CoSEn is the most powerful feature in multi-class discrimination and in the discrimination between AF and other rhythms.

We also introduced a novel feature, the Mawsd, which showed the highest discriminant power in the discrimination between AF and normal rhythms and a high discriminant power in multi-class discrimination.

The problem of discriminating between normal and other rhythms was the most complex. Among the features computed, the KFD appeared to be the most discriminant one, which could be due to the higher fractal properties of RR series with ectopic beats [22].

The application of the LS-SVM classifier on the training set provided quite high performance. In particular, as it can be observed from the confusion matrix, the classifier seems particularly efficient in discriminating between normal from AF rhythms and AF from other rhythms while it is critical in discriminating normal from other rhythms.

The marked decreasing in performance, between the training set (F1=0.90) and the hidden test set (F1=0.81), points out the poor generalization properties and the overfitting of our estimated classification model. In particular, a strong decrease of F1 index occurs for the Other rhythm (training set F1= 0.86; test set F1= 0.72, before relabeling).

It is well known that SVM with Kernel extension, working in a higher-dimensional feature space, may suffer by an increasing of the generalization error. Moreover, in LS-SVM approach, sparseness of α_i is lost and all the training observations are considered as support vectors.

In the future, the performances of the proposed algorithm could be improved in three ways. First, more discriminant features could be introduced focused on the discrimination between normal and other rhythms. Second, implementing LS-SVM pruning algorithms to decrease the generalization error. Third, testing other classification approach with higher generalization capabilities.

An accurate classification of cardiac rhythms would be important in the clinical practice for the implementation of the specific treatment.

Acknowledgements

This study was supported by Bando FAS Salute Sviluppo Toscana 2014-PANACEE Project (No. 6247) and by Cardioline S.p.A., Trento, Italy.

References

[1] Markides V, Schilling R. Atrial fibrillation: classification, pathophysiology, mechanisms and drug treatment. Heart 2003;89:939–943.

[2] Chiarugi F, Varanini M, Cantini F, Conforti F, Vrouchos G. Noninvasive ecg as a tool for predicting termination of paroxysmal atrial fibrillation. IEEE Transactions on Biomedical Engineering Aug 2007;54(8):1399–1406. ISSN 0018-9294.

[3] Tateno K, Glass L. A method for detection of atrial fibrillation using RR intervals. In Computers in Cardiology, volume 27. 2000; 391–394.

[4] Dash S, Chon K, Lu S, Raeder A. Automatic real time detection of atrial fibrillation. Ann Biomed Eng 2009; 37:1701–1709.

[5] Salahuddin L, Jeong MG, Kim D. Ultra short term analysis of heart rate variability using normal sinus rhythm and atrial fibrillation ECG data. In IEEE 9th Int. Conf. E-health Networking, Application and Services. 2007; 240–243.

[6] Stridth M, Sornmo L. Shape characterization of atrial fibrillation using time-frequency analysis. In Computers in Cardiology, volume 29. 2002; 17–20.

[7] Clavier L, Boucher J, Lepage R, Blanc J, Cornily J. Automatic P-wave analysis of patients prone to atrial fibrillation. Med Biol Eng Comput 2002;40:63–71.

[8] Larburu N, Lopetegi T, Romero I. Comparative study of algorithms for atrial fibrillation detection. In Computing in Cardiology. 2011; 265–268.

[9] Kara S, Okandan M. Atrial fibrillation classification with artificial neural networks. Pattern Recognition 2007; 40:2967–2973.

[10] Mohebbi M, Ghassemian H. Detection of atrial fibrillation episodes using SVM. In 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Aug 2008; 177–180.

[11] Asgari S, Mehrnia A, Moussavi M. Automatic detection of atrial fibrillation using stationary wavelet transform and support vector machine. Comput Biol Med 2015;60:132–142.

[12] Kennedy A, Finlay D, Guldenring D, Bond R, Moran K, McLaughlin J. Automated detection of atrial fibrillation using R-R intervals and multivariate-based classification. J Electrocardiol 2016;49:871–876.

[13] Nayak J, Naik B, Behera H. A comprehensive survey on support vector machine in data mining tasks: Applications & challenges. International Journal of Database Theory and Application 2015;8:169–186.

[14] Carrara M, Carozzi L, Moss T, De Pasquale M, Cerutti S, Ferrario M. Heart rate dynamics distinguish among atrial fibrillation, normal sinus rhythm and sinus rhythm with frequent ectopy. Physiological Measurement 2015;36:1873–1888.

[15] Clifford GD, Liu C, Moody B, Lehman LWH, Silva I, Li Q, Johnson AE, Mark RG. AF classification from a short single lead ecg recording: the physionet/computing in cardiology challenge 2017. In Computing in Cardiology, volume 44. Ieee, (In Press), 2017; .

[16] Varanini M, Tartarisco G, Billeci L, Macerata A, Pioggia G, Balocchi R. An efficient unsupervised fetal qrs complex detection from abdominal maternal ecg. Physiological Measurement 2014;35:1607–1619.

[17] Katz M. Fractals and the analysis of waveforms. Comput Biol Med 1988;18(3):145–156.

[18] Suykens J, Van Gestel T, De Brabanter J, De Moor B, Vandewalle J. Least Squares Support Vector Machines. World Scientific, 2002.

[19] Vapnik V. Pattern recognition using generalized portrait method. Automation and Remote Control 1963;774–780.

[20] Hastie T, Tibshirani R. Classification by pairwise coupling. The Annals of Statistics 1998;26:451–471.

[21] DeMazumder D, Lake D, Cheng A, Moss T, Guallar E, Weiss R, Jones S, Tomaselli G, Moorman J. Dynamic analysis of cardiac rhythms for discriminating atrial fibrillation from lethal ventricular arrhythmias. Circ Arrhythm Electrophysiol 2013;6:555–561.

[22] Saeed M. Fractals analysis of cardiac arrhythmias. TheScientificWorldJOURNAL 2005;5:691–701.

Address for correspondence:

Maurizio Varanini
Institute of Clinical Physiology, National Research Council of Italy (CNR), via Moruzzi 1, 56127, Pisa, Italy
maurizio.varanini@ifc.cnr.it