Computers in Cardiology/PhysioNet Challenge 2004 : AF Classification Based on Clinical Features

M Lemay[1], Z Ihara[1], JM Vesin[1], L Kappenberger[2]

[1] Signal Processing Institute, EPFL, Lausanne, Switzerland
[2] Service of cardiology, CHUV,Lausanne, Switzerland

Abstract

The Computers in Cardiology / Physionet Challenge 2004 deals with the classification of ECG signals from AF patients into three categories: types N, S and T corresponding to AF episodes terminating never, soon and immediately, respectively. In our study, different features were used, extracted by the experienced clinician among the authors (LK) on the supplied training set. Algorithms were developed to quantify these features from provided ECG data. A Support Vector Machine was used to classify these features. In this paper, we present our method, results and conclusion about this clinically-oriented approach.

1. Introduction

Atrial fibrillation (AF) is the most common type of human arrhythmia and it is responsible for about one third of hospitalizations for arrhythmia problems. The diagnosis of AF has been assessed for years by visual inspection of the surface electrocardiogram (ECG). In 2001, the Computers in Cardiology conference proposed a challenge devoted to the clinical AF problem. The challenge was to predict the onset of paroxysmal atrial fibrillation by an automated method. Some methods used were based on the number and timing of atrial premature complexes[1], on the quantification of patterns of heart rate dynamics[2], on analysis of the incidence of premature atrial complexes and P-wave variability [3] or on the analysis of the R-R time series[4].

For the 2004 Workshop, CinC and PhysioNet have continued in the AF area with a new challenge which consists in the prediction of spontaneous termination of AF. The automated method has to classify three different types of AF: non-terminating AF (Group N), AF that terminates one minute after the end of the record (Group S) and AF that terminates immediately after the end of the record (Group T). A learning set of 30 records (two unspecified simultaneously recorded leads each) of 60 seconds (10 labelled records from each group) and two test sets were available. The test set A contains 30 records of groups N and T. The test set B contains 20 records of groups S and T. These test sets have to be classified correctly.

We decided to approach the challenge from a clinician's point of view. One of us (LK) identified different observations that characterized the disorganization of F-waves. Next, we developed the processing tools to quantify these observations. A support vector machine (SVM) technique was used to solve the classification problem.

2. Methods

We applied systematically a baseline removal step consisting of highpass filtering (0.5 Hz) on each signal. An accurate segmentation is crucial in our approach because most of our observations are based on time segments where only atrial activity (AA) is present. These segments include fibrillation waves (F-waves) or organised P-waves. The identification procedure first detects the R-waves using the first and second derivative of the signal. Q-, S- and T-waves are located by time oriented operations. Specifically, the T-waves are detected using the time course of root-mean-square (RMS) values of both leads.

When AF converts to sinus rhythm, the completely unstructured F-waves tend to be more organized and structured than P-waves. This behavior is detected by different clinical observations which lead to the features used in this study. These are presented in the next 7 subsections.

2.1. F-wave polarity

The first observation is the main polarity of the F-waves in the AA. When AF is to terminate soon, P-waves tend to reappear and F-waves are more organized. This translates into a stable (positive or negative) F-wave polarity which will be absent in a non-terminating AF. As such the clinician tries to identify F-wave asymmetry. Figure 1 shows a typical example where no obvious dominant
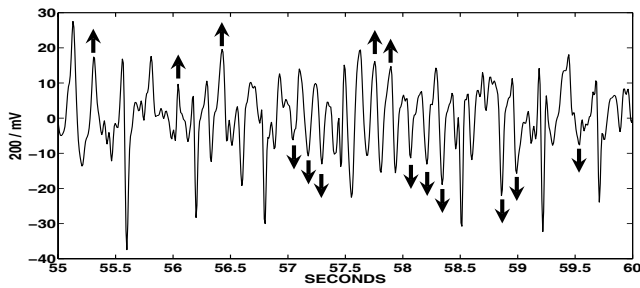
polarity can be found.

Figure 1. Second lead of n04, instable peak direction

Mathematically, Features 1 and 2 (one for each lead) were expressed by the skewness of the AA segments:

$$skewness = \frac{\sum_{i=1}^{N} \left(x_{(i)} - \mu_x\right)^3}{(N-1)\,\sigma^3}, \qquad (1)$$

where $x_{(i)}$ are the samples over all AA segments, μ_x and σ the estimated mean and standard deviation of these samples and N is the total of samples $x_{(i)}$ in all segments.

2.2. R-R intervals

When the ventricular rate is high, AF is more likely to terminate soon because the heart is thought to be unable to sustain a high ventricular rate in AF for a long period of time. We quantified this observation by the mean of R-R intervals, which gave Feature 3. Moreover, there is often a marked change (positive or negative) in ventricular rythm prior to AF termination. This was quantified using a simple R-R interval analysis: the difference between the mean of R-R intervals between the first and the last ten seconds of the record, which gave Feature 4. Such extended R-R analysis has previously been reported in [2, 4].

2.3. F-wave peak intervals

When AF reverts to sinus rhythm, these intervals decrease and multiple F-waves reduce to an individual P-wave. First, a lowpass filter with a cutoff frequency of 8 Hz was applied. This cutoff frequency is appropriate for the identification of F-waves: it is high enough not to disturb the shape of F-waves and low enough to produce a single peak for each wave. The QRST detector gave the AA segments. Next, a simple peak detector (the same as for the QRST detector) was applied to those segments. Figure 2 shows a typical exemple. The bold dashed lines represent the filtered AA segments and the black dots corresponds to the identified F-wave peaks. This observation was quantified by the average number of positive peaks per time unit (Feature 5).
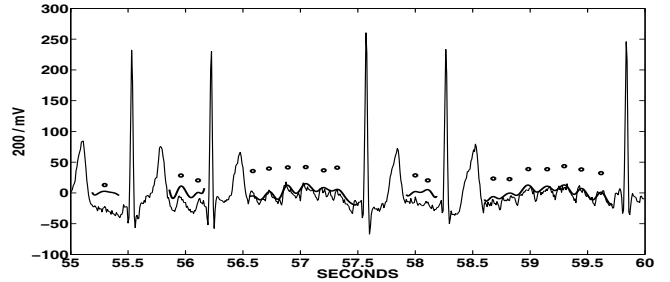

Figure 2. First lead of n07, AA intervals

2.4. Atrial activity amplitude

When the AF is more organized in the atrial tissue, the resulting F-waves in the ECG are higher. To have a correct estimate of the AA amplitude, we used the F-wave detector algorithm described above. The identified peaks and the minima between these peaks were used to compute an average AA amplitude on both leads, which gave Features 6 and 7.

2.5. Low-frequency modulation of atrial activity amplitude

The next observation relates to the possible low-frequency modulation in the amplitude of AA. It has been observed empirically that prior to AF termination F-wave amplitude oscillates slowly. Just before AF stops, slow modulations appear in the AA amplitudes. A good example is the record t07 as shown in Figure 3 where this amplitude modulation is clearly visible.
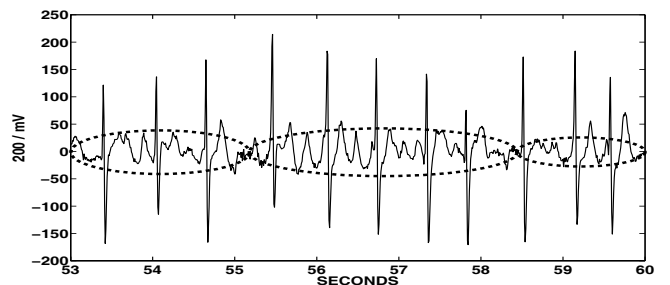

Figure 3. First lead of t07, AA amplitude periodicity

To quantify this observation, we applied first a bandpass filter to the signal with cutoff frequencies at 1 and 8 Hz. The filtering below 1 Hz is used to eliminate any residual baseline drift. After the filtering, QRST (for the AA segments) and F-wave detectors were applied to identify the maximum values of AA. To minimize the frequency power of the QRST, a saturation operation was applied to

the signal. This operation is an hyperbolic tangent one:

$$signal = AA_{max} * tanh(2 * signal/AA_{max}), \qquad (2)$$

where AA_{max} is the maximum value in the AA segment. A morphological envelope detector was used to extract the upper and lower envelopes. The final amplitude estimate was obtained by the subtraction of the upper envelope from the lower one. The power spectrum density of this amplitude signal was estimated. Features 8 and 9 used are the total power below 0.6 Hz on both leads.

2.6. Similarity between atrial activities in different leads

Another observation is a measure of the similarity between AAs in the two leads. During non-terminating AF, the F-waves in the two leads are quite dissimilar, possibly due to the large number of wavefronts propagating in the atrial tissue. Before AF returns to sinus rhythm, F-wave shapes become similar reflecting structurisation of AF waves. Both leads of records s02 (learning set) display this similarity (Figure 4). From 55 to 58.5 seconds, the F-waves are correlated and after 58.5 seconds, the F-waves are anti-correlated.
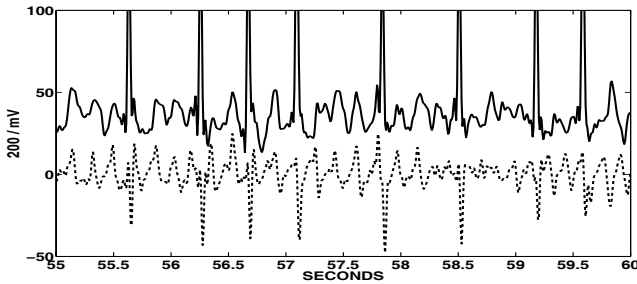

Figure 4. Leads of s02, Similarities among AAs

Mathematically, this observation can be expressed as the normalized cross-covariance:

$$\frac{\sum_{i=1}^{N_j}(x_{ij} - \mu_{x_j})(y_{ij} - \mu_{y_j})}{\sqrt{\sum_{i=1}^{N_j}(x_{ij} - \mu_{x_j})^2 * \sum_{i=1}^{N_j}(y_{ij} - \mu_{y_j})^2}}, \qquad (3)$$

where x_{ij} and y_{ij} are the ith samples on leads 1 and 2 in the jth segments. μ_{x_j} and μ_{y_j} are the estimated means of the jth segment for each of the two leads and N_j is the total of samples in j^{th} segment, N_j is the same for both leads. The cross-covariance values for each AA segment were averaged, which gave Feature 10.

2.7. High-frequency power in atrial activity segments

The final observation is the power of the AA segments in the high-frequency range. It is assumed that the F-waves close to conversion in a P-wave have more high-frequency components than the F-waves in non-terminating AF. The quantification used was the average of the estimated spectral density above 20 Hz across the AA segments on both leads (Features 11 and 12).

2.8. SVM Classification

From the 7 clinical observations, we obtained 12 features: one feature for the F-wave peak intervals, one for the similarity between atrial activities in different leads and two features for each of the other observations. A Support Vector Machine (SVM) approach was used to classify these features. The SVM tries to separate the data by an optimal hyperplane. Optimisation is obtain by maximizing the margin (maximizing the distance between the hyperplane and the nearest data point of each class). If the linear boundary is inapropriate, the SVM can map the input data into a high dimensional space. We used the Support Vector Machine Toolbox [6] to implement the SVM algorithm with a linear kernel function and a upper bound at infinity. The normal vector of the hyperplane for the N-T classification was : [0.2428 0.7196 5.5665 -0.8466 2.6906 4.6257 0.2393 -0.9238 -3.4938 1.9830 -2.8672 -0.8543]; for the S-T classification, it was : [7.5222 17.3131 10.3262 -3.8598 7.2648 -8.5767 4.3471 -3.0488 2.3011 4.0013 -0.2104 -15.0128] where each axis of the space dimension is represented by one of the 12 features. For more details on SVM methods, see [5]. Last, we set the time intervals parameter of the ECG signals mentioned in Section 3.

3. Results

3.1. Classification of N and T

For this classification, the SVM was trained on all records of the learning set, where S and T were grouped. The best result was obtained with ECG signals restricted to time intervals between 40 to 60 seconds.

On the test set A: A1, A3, A5, A6, A7, A8, A10, A12, A13, A15, A16, A19, A20, A21, A23, A26, A28 and A29 were classified as N type. A2, A4, A9, A11, A14, A17, A18, A22, A24, A25, A27 and A30 were classified as T type.

This classification resulted in a score of 20 out of 20 for the learning set and 20 out of 30 for the test set, which represents 80% of classification accuracy on learning set and test set put together.

3.2.   Classification of S and T

For this classification, the SVM was trained on the groups S and T of the learning set, where s09 was included in set T (in fact, the record s09 is in sinus rhythm!). The best result was obtained with ECG signals restricted to time intervals between 0 to 60 seconds.

On the test set B: B5, B6, B9, B10, B11, B15, B17, B18, B19 and B20 were clasified as S type. B1, B2, B3, B4, B7, B8, B12 and B13, B14 and B16 were classified as T type.

For this classification, we obtained a score of 20 out of 20 for the learning set and 12 out of 20 for the test set, which represents 80% of classification accuracy on learning set and test set put together.

4.   Discussion and conclusions

Our objective was to determine whether an approach based on clinician's expertise was able to predict the termination of AF. At the beginning, the clinician (LK) looked at the learning set and identified 7 observations that characterized the disorganisation of F-waves. These observations were subjectively classified into 3 categories. These categories were associated to the 3 different groups. For the test sets, the clinician applied the same classification and based his final decision on the majority of classified observations.

The automated method quantified these observations into 12 features. These features were confirmed by the clinician's opinion but the final classification processes (i.e. mental or SVM based) can hardly be compared. Our "clinical" features worked well on the learning set. This confirms the good fit between the clinical observation and our quantification procedures. Our classification was also able to separate each type of AF without any errors on the learning set. The overall 80% of classification accuracy was good, but the results on both test sets A and B (66.7% and 60%, respectively) are below our expectation. There are some explanations for this. The main reason is probably that there were not enough records in the test sets to represent all the characteristics of each of our quantified clinical observations. It maybe also that, whitout the possi-

bility to build a validation set due to the small size of the learning set, some overfitting took place. Despite this, our goal to build an automated method based on a clinician's approach was reached.

Acknowledgements

References

[1]   Zong W, Mukkalama R, Mark RG. A Methodology for Predicting Paroxysmal Atrial Fibrillation Based on ECG Arrythmia Feature Analysis. Computers in Cardiology 2001;28:125–128

[2]   Yang ACC, Yin HW. Prediction of Paroxysmal Atrial Fibrillation by Footprint Analysis. Computers in Cardiology 2001;28:401–404

[3]   Moody GB, Goldberger AL, McClennen S, Swiryn SP. Predicting the Onset of Paroxysmal Atrial Fibrillation: The Computers in Cardiology Challenge. Computers in Cardiology 2001;28:113–116

[4]   Thong T, McNames J, Aboy M, Goldstein B. Prediction of Paroxysmal Atrial Fibrillation by Analysis of Atrial Premature Complexes. IEEE Transactions on Biomedical Engineering April 2004;51-4:561–569

[5]   Gunn SR. Support Vector Machines for Classification and Regression. Technical Report - University of Southampton 1998

[6]   http://www.isis.ecs.soton.ac.uk/resources/svminfo/

Address for correspondence:

Mathieu Lemay
STI - ITS - LTS, EPFL, 1015 Lausanne, Switzerland
mathieu.lemay@epfl.ch