Classification of AF and Other Arrhythmias from a Short Segment of ECG Using Dynamic Time Warping

Maria Tziakouri[1], Costas Pitris[1,2], Christina Orphanidou[2,3]

[1]Department of Electrical and Computer Engineering, University of Cyprus
[2]Kios Research and Innovation Center of Excellence, University of Cyprus
[3]Oxygen Research Ltd, Limassol, Cyprus

Abstract

Atrial Fibrillation(AF) is a major public health risk but its identification is challenging because it may be episodic and non-symptomatic. Automatically identifying episodes of AF from a short segment of ECG would, thus, be beneficial. As a response to the Physionet/Computing in Cardiology Challenge 2017 we have implemented a three-stage classifier which can classify segments of ECG into Noisy, Normal, AF or Other Rhythm. We employ a state-of-the-art SQI for identifying noisy segments and then learn two different Support Vector Machine (SVM) classifiers using features extracted from the ECG. The features used are derived using a template matching approach via Dynamic Time Warping and using the statistical characteristics of the R-R intervals. Our average F1 score on the validation set was 0.66.

1. Introduction

Atrial fibrillation (AF) is the most common abnormal heart rhythm which occurs when rapid, disorganized electrical signals cause the heart's two upper chambers (the atria) to "fibrillate", i.e., to contract very fast and irregularly when moving blood into the ventricles [1]. While AF is not a life-threatening arrhythmia, it has been associated with an increased risk for heart failure, stroke and mortality [2]. From a socioeconomic perspective, AF has also been associated with permanent disability, cognitive disturbance, hospitalization and absence from work [3]. Advances in the treatment of cardiac disease and the improved ability to diagnose AF have led to an increase in the reported prevalence of AF; the number of patients with AF in Europe is projected to be 14-17 million by 2030 with the number of new cases per year estimated at 120000-215000 [3]. In the US, the prevalence of AF in 2030 is projected to be 12.1 million [4], making it an important public health problem and significant cause of healthcare expenditure in the western world.

A major challenge in identifying incidences of AF is that it is often episodic; AF incidences are often paroxysmal, short-lived and asymptomatic. The ability to correctly detect and monitor AF using only short-segments of Electrocardiogram (ECG) signals is, thus, desirable, and would result in improved patient care and health outcomes.

The present work describes an algorithm designed to address the topic of "AF classification using a short single-lead ECG recording", which was the Physionet/Computing in Cardiology Challenge 2017, described in [5].

2. Methods

Data were provided by the Physionet/Computing in Cardiology Challenge 2017. Data comprised of ECG recordings, collected using the AliveCor device. The training set made available to participants contained 8,528 single lead ECG recordings lasting from 9s to just over 60s. A test set consisting of 3,658 ECG recordings of similar lengths will be used for the final scoring but is not available to participants until after the completion of the Challenge. ECG recordings were sampled as 300 Hz and were band-pass filtered by the recording device. The ground truth classification was also provided. Data were classified in the four following categories: Normal, AF, Other (abnormal) rhythm and Noisy.

2.1 Proposed algorithm

As a pre-processing step, we initially apply a Savinsky-Golay smoothing filter to the ECG segments. The algorithm then classifies the short ECG segments into the four categories using a three-stage classification scheme:
I) Stage 1, Noisy/Clean: a state-of-the-art ECG signal quality index (SQI) is applied to all segments to determine if they are "acceptable" or "unacceptable" quality-wise. The algorithm, uses a set of decision rules, based on physiological viability and then proceeds to a template matching step where an average QRS template is extracted in each segment which is then used to calculate an average correlation coefficient of the template with each individual QRS complex in the ECG segment. The original algorithm, proposed in [6], performs quality analysis in 10s segments, it was therefore adapted and optimized for use in segments

 ISSN: 2325-887X DOI:10.22489/CinC.2017.348-295

of variable length. Segments which are determined to be "unacceptable" after application of the SQI are classified as "Noisy". Segments which are determined to be "acceptable" are classified as "Clean" and proceed to the next classification step.

II) Stage 2, Normal/Abnormal: in this step, segments labelled as either AF or Other rhythm are collapsed into one group, collectively labelled as Abnormal. Using seven features extracted from the ECG segments, a Support Vector Machine (SVM) classifier is trained which is used to classify segments as Normal or Abnormal. Segments which are determined to be Abnormal after this step proceed to the next classification step.

III) Stage 3, *AF/Other rhythm:* in the final classification step, we revert to the original labelling of AF/Other rhythm. Using the same seven features as in the previous classification step, a second Support Vector Machine (SVM) classifier is trained which is used to classify segments labeled Abnormal in the previous step, as AF or Other rhythm.

A flow chart of the proposed classification scheme is shown on figure 1.

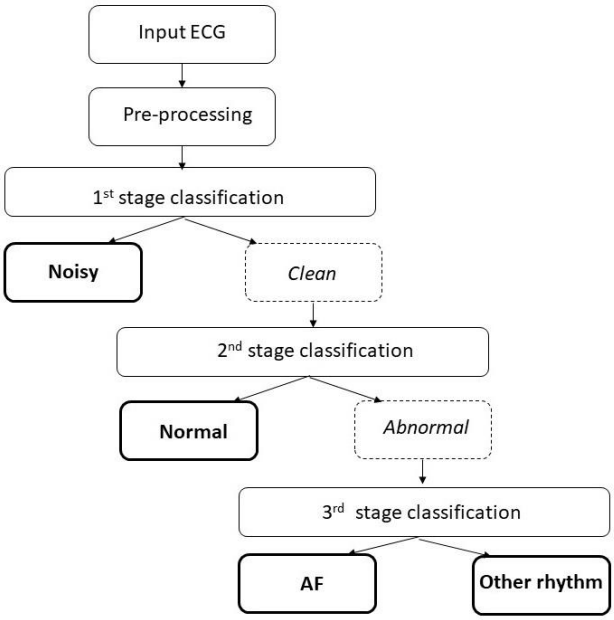

Figure 1. Flow chart of proposed classification scheme.

The features used in the 2nd and 3rd stage classifiers are explained next.

2.2 Features

Seven features were extracted from the ECG segments which were then fed into the two different SVM classifiers to perform classification steps 2 and 3. Two categories of features were used: two features are based on a template matching scheme employing Dynamic Time Warping. Five additional features are based on the statistical characteristics of the R-R intervals (the time-intervals between successive R-peaks).

2.3 Template matching via Dynamic Time Warping

Dynamic Time Warping (DTW) is a technique for aligning two time-series of different length, using a non-linear transformation. Given two time series P and Q of lengths $i = 1 \dots N$ and $j = 1 \dots M$, respectively, a $N \times M$ matrix is constructed where element (i,j) contains the distance of points $d(p_i, q_j)$. After using a piecewise linear approximation algorithm to transform P and Q to short line sequences, $d(p_i, q_j)$ is calculated as the absolute difference between the slopes in each short line. A cumulative distance measure $c_{i,j}$ (also termed the *alignment* or *warping* cost), is then calculated, defined as

$$c_{i,j} = min \begin{cases} c_{i-1,j} + d(p_i, q_j) l(p_i) \\ c_{i-1,j-1} + d(p_i, q_j)(l(p_i) + l(q_j)) \\ c_{i,j-1} + d(p_i, q_j) l(q_j) \end{cases}$$

where $l(p_i)$ and $l(q_j)$ are the duration of line p_i and q_j [7]. The optimal warping path is chosen as the one which minimizes the cumulative distance and it is then applied to Q as a non-linear transformation to obtain the best alignment with P.

Template-matching approaches via DTW have been used in the past in [7] and [8] for quality assessment of the the Photoplethysmogram (PPG) by extracting a template pulse waveform from segments of PPG and aligning them with each pulse waveform in the same segment using DTW. In [7], the average correlation coefficient of each pulse waveform with the template was calculated to obtain a quality index. In [8], the alignment cost was taken as a quality index since the more similar two time-series are, the "cheaper" it will be to align them. Dissimilarity between a template and an individual pulse waveform was taken as an indication of the presence of noise.

For the detection of arrhythmias, to take into consideration the variability in the duration between successive R-peaks, we employed a template matching approach via DTW by extracting a *two-beat* template from each ECG segment, calculating its alignment cost with each two-beat waveform in the same segment and taking the average and standard deviation of the alignment costs. After performing R-peak detection using the Hamilton and Tompkins algorithm [9], the median R-R interval in the segment is calculated. The two-beat template is the extracted by taking the segment starting at half the median R-R interval before the first R-peak until half the median R-R interval after the second R-peak. The duration of the template is, thus, approximately equal to two R-R

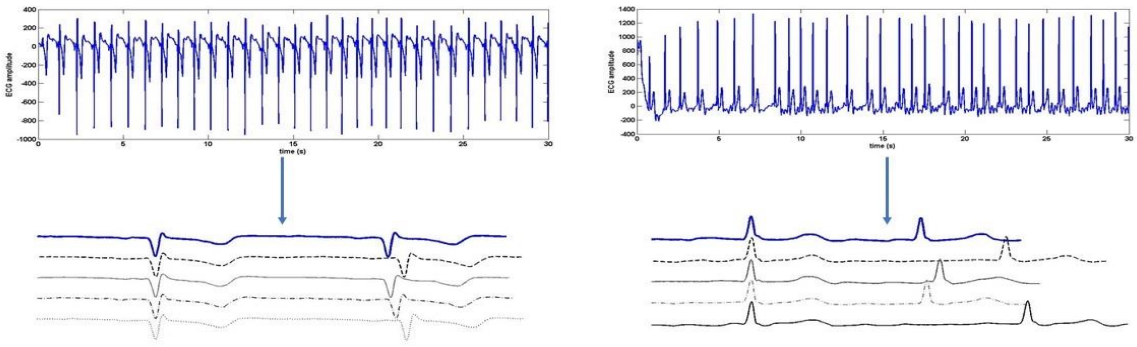

Figure 2. 30s ECG segments with Normal rhythm (left) and Atrial Fibrillation (right) from the training set with corresponding two-beat template (top waveform) and example two-beat templates from the same segment. In the Normal rhythm segment the extracted two-beat waveforms would have a small DTW alignment cost whereas in the AF case the alignment cost would be higher due to the variability in the length between successive beats.

intervals. The alignment cost of this template with each two-beat segment (taking it in the same way as the template around each two consecutive R-peaks) is then calculated after applying DTW. Example templates are shown on Figure 2. The basis of this approach is that a segment in normal sinus rhythm will have a low alignment cost since the QRS morphology and R-R intervals will be regular. However, a segment with arrhythmias will have variable R-R intervals, as well as variability in the QRS morphology, (e.g., in the morphology of the P-wave) so it will have a higher alignment cost. After calculating the alignment cost of each two-beat segment with the template, two features are extracted from each segment to be used in our classification scheme:

D1: the mean of the alignment costs within a segment of ECG.

D2: the variance of the alignment costs within a segment of ECG.

2.4 Statistical characteristics of the R-R intervals

Measures of variability in the R-R interval are well-established features for differentiating between normal and arrhythmic segments of signal. Our algorithm uses the following five features:

S1: Standard deviation of successive differences in R-R intervals (SDSD). This is a standard measure of variability in the R-R intervals.

S2: Skewness of distribution of R-R intervals: skewness is the third standardised moment of a probability distribution and measures how symmetric the distribution is. The presence of an arrhythmia will result in the presence of outliers in the distribution of R-R intervals which will cause asymmetry in the distribution. Skewness is given by

$$\hat{S} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \widehat{x})^3}{\hat{\sigma}^3} \qquad (1)$$

where x_i is the discrete signal (the R-R intervals in this case), $\hat{\mu}$ and $\hat{\sigma}$ are the empirical estimates of the mean and standard deviation of the distribution of x_i and N the number of data points in the signal.

S3: Kurtosis of the distribution of R-R intervals: kurtosis is the fourth standardised moment of a probability distribution and measures how sharp the peak of a distribution is. Similarly to skewness, the presence of an arrhythmia will result in the presence of outliers in the distribution of R-R intervals which will flatten the distribution. Kurtosis is given by

$$\widehat{K} = \frac{\frac{1}{N}\sum_{i=1}^{N}(x_i - \widehat{x})^4}{\hat{\sigma}^4} \qquad (2)$$

where parameters are defined as in S2.

While S2 and S3 rely on the assumption of normally distributed R-R intervals we observed that in some cases, the skewness and asymmetry of the distribution pointed to the shape of a gamma-distribution. We therefore additionally include as features the shape (k) and scale (θ) parameters of the gamma distribution which would presumably differ between normal ECGs and ECGs with arrhythmias.

These parameters are derived by fitting a gamma-distribution to the distribution of R-R intervals in each segment, where the probability density function of the gamma-distribution is given by

$$f(x;k,\theta) = \frac{x^{k-1}e^{-\frac{x}{\theta}}}{\theta^{\kappa}\Gamma(k)} \quad \text{for } x > 0 \text{ and } k,\theta > 0 \qquad (3)$$

Where $\Gamma(k)$ is the gamma function value at k.

The final features included in the classification algorithms are, thus:

S4: The shape parameter k of the gamma-distribution
S5: The scale parameter θ of the gamma-distribution

2.5 Support Vector Machine Classifiers

For the two different classification steps, features D1, D2, S1, S2, S3, S4 and S5 were fed into an SVM classifier with a radial basis kernel function (with a σ=0.5 which was optimized using trial and error).

3. Results

We evaluated the performance of the proposed system on the validation set provided by the Physionet/ Computing in Cardiology Challenge 2017. The results were assessed in terms of the F1 score which is a measure of classification performance in the absence of True Negatives (TN). The assessment strategy is explained in detail in [5]. Table 1 shows the F1 scores obtained for the four types of signals

Table 1. F1 scores on validation data

Type	F1 score
Normal	0.7759
AF	0.7347
Other rhythm	0.4404
Noise	0.7660

4. Discussion

We have presented an algorithm for detecting AF and other arrhythmias from short segments of ECG which relies on a three-stage classification scheme and employs several features which are based on a template matching approach using Dynamic Time Warping (DTW) and low- and high-order statistics of the distribution of R-R intervals. Our algorithm performed satisfactorily for detecting noisy segments and segments of normal rhythm and AF but underperformed for detecting Other rhythms. Differentiating between the different abnormal rhythms (AF/Other rhythms) is difficult, especially since the nature of the different rhythms is not clearly known. It is possible that building a system for differentiating between AF and a specific other type of arrhythmia would perform well if the differences between the presentation of the two arrhythmias on the ECG were clearly understood such that appropriate features were selected. Additionally, it would be interesting to explore the performance of the proposed system for different durations of signal. All statistical features used for the classification steps of the algorithm might have better differentiating performance for longer segments of signal compared to smaller segments of signal where only a few R-R intervals are found. In general, a system which relied on a pre-determined length of window for processing would likely perform better. Lastly, DTW is a computationally intensive process which slows down the procedure. However, in practice the identification of arrhythmias from the ECG could be done with some time-delay without compromising the clinical requirements of the system.

References

[1] American Heart Association, What is Atrial Fibrillation, http://www.heart.org/HEARTORG/Conditions/Arrhythmia/AboutArrhythmia/What-is-Atrial-Fibrillation-AFib-or-AF_UCM_423748_Article.jsp#.WagGksgjHIU, accessed 30/08/2017.

[2] Lee J., Reyes B.A., McManus D. D, Mathias O., and Chon K. I. (2012), Atrial Fibrillation Detection using a Smart Phone, Proceedings of the 24th Annual International Conference of the IEEE EMBS, pp. 1177-1180.

[3] Zoni-Berisso M., Lercari F., Carazza T. and Domenicucci S., (2014), Epidemiology of atrial fibrillation: European perspective, Clinical Epidemiology, 6, pp. 213-220.

[4] Colilla S., Crow A., Petkun W., Singer D. E., Simon T. and Liu Xianchen (2013), Estimated of Current and Future Incidence and Prevalence of Atrial Fibrillation in the U.S. Adult Population, American Journal of Cardiology, 112(8), pp. 1142-7.

[5] Gari Clifford, Chengyu Liu, Benjamin Moody, Ikaro Silva,Qiao Li, Alistair Johnson, Roger Mark. AF Classification from a Short Single Lead ECG Recording: the PhysioNet Computing in Cardiology Challenge 2017. *Computing in Cardiology* (Rennes: IEEE), Vol 44, 2017 (In Press).

[6] C. Orphanidou, T. Bonnici, P. Charlton, D. Clifton, D. Valance and L. Tarassenko, "Signal-Quality Indices for the Electrocardiogram and Photoplethysmogram: Derivation and Applications to Wireless Monitoring," IEEE J. Biomed. Health Inform, vol. 19, no. 3, pp. 832-838, 2015.

[7] Li Q. and Clifford G. D.(2012), Dynamic Time Warping and Machine Learning for signal quality assessment of pulsatile signals, Physiological Measurement, 33, pp. 1491-1501.

[8] Sun X., Yang P., Zhang Y-T. (2012), Assessment of Photoplethysmogram Signal Quality using Morphology Integrated with Temporal Information Approach, Proceedings of the 34th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC 2012), pp. 3456-3459.

[9] P. S. Hamilton and W. J. Tompkins, "Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database.," IEEE Trans. Biomed. Eng., vol. 33, no. 12, pp. 1157–1165, 1986.

Address for correspondence.

Christina Orphanidou
Department of Electrical and Computer Engineering
University of Cyprus, 1 Panepistimiou Avenue , 2109 Aglanzia, P.O.Box. 20537
1678, Nicosia Cyprus
orphanid@ucy.ac.cy