Cardiac Arrhythmia Discrimination Using Evolutionary Computation

Juan Francisco Martín-García[1], Inmaculada Mora-Jiménez[1], Arcadio García-Alberola[2], José Luis Rojo-Álvarez[1,3]

[1]Department of Signal Theory and Communications, University Rey Juan Carlos, Madrid, Spain
[2]Arrhythmias Unit, University Hospital Virgen de la Arrixaca, Murcia, Spain
[3]Prometeo, Electric and Electronic Department, Universidad de las Fuerzas Armadas ESPE, Ecuador

Abstract

The use of Implantable Cardioverter Defibrillators (ICD) for cardiac arrhythmia treatment implies a search for efficiency in terms of discrimination quality and computational complexity, given that improved efficiency will automatically turn into more effective therapy and longer battery lifetime. In this work, we applied evolutionary computation to create classifiers capable of discriminating between ventricular and supraventricular tachycardia (VT/SVT) in episodes registered by ICDs. Evolutionary computation comprises several paradigms emulating natural mechanisms for solving a problem, all of them characterized by a population of individuals (possible solutions) which evolve generation after generation to provide fitter solutions. Genetic programming was the paradigm chosen here because its solutions, coded as decision trees, can be both computationally simple and clinically interpretable.

For the experiments, we considered electrograms (EGM) from episodes registered by ICDs in spontaneous/induced tachycardia, previously classified as VT/SVT by clinical experts from several Spanish healthcare centers. Training data were 38 real-valued samples, arranged as the concatenation of two beat segments: a sinus rhythm template immediately previous to the arrhythmic episode (basal reference), and the arrhythmic episode template. Several low complexity trees provided low error rates and allowed physiological interpretation. The best tree yielded an error rate of 1.8%, with both sensitivity and specificity above 98%. This solution compares two samples from the end of the arrhythmic pulse with another two samples from the sinus rhythm, pointing out to a relevant discrimination role of the lasting EGM.

1. Introduction

Since they were first introduced as a way of treatment for malignant arrhythmias, implantable cardioverter defibrillators (ICD) have undergone constant evolution, either by the introduction of new functionalities (from first-generation non-programmable ICDs to fourth-generation bicameral devices) or by the improvement of their discrimination criteria. There are in the literature several complex discrimination methods, such as the Prediction Error Criterion [1], the Correlation Waveform Analysis [2], the QRS Width Criterion [3], or the spectral analysis of the QRS complex [4]. Though elaborated methods usually provide low error rates, they force the device to perform complex calculations usually lacking physiological meaning. In this paper, we propose the use of evolutionary computation to design classifiers driven by data (voltages associated to beat segments registered by ICDs). The goal is to find computationally simple and physiologically meaningful discrimination criteria which may help the specialists to discern between ventricular tachycardia (VT) and supraventricular tachycardia (SVT).

The most natural way to search for automatic problem solvers seems to be the use of mechanisms similar to the most powerful problem solvers found in nature, that is, the human brain and the evolutionary process (which created human brains) [5]. Evolutionary computing methods appear as a machine learning approach to the second option. Evolutionary computation [5] is an artificial intelligence discipline emulating natural mechanisms to achieve the solution of a problem. Despite evolutionary computation is an umbrella term including different paradigms, the whole family can be characterized by a population consisting of individuals (possible solutions) which evolve generation after generation to provide better solutions. In this work, we have considered the paradigm of genetic programming (GP) [6] because it can yield solutions both computationally simple and clinically interpretable.

Solutions in GP are coded as tree-like structures representing syntactic expressions,: operators are the nodes of the tree, data features and constants make for leaves, and branches connecting nodes and leaves determine the operator composition. The natural mechanisms applied on individuals of one generation (so-called parents) to provide next generation are inspired in biological processes,

thus receiving the names of their biological analogues, namely, crossover and mutation. The crossover mechanism consists basically on the sub-trees swapping between two parents, while mutation implies a random change on just one individual. Remarkably, mutations should occur with lower probability than crossover, in order to properly emulate the biological process. The population evolves from one generation to the next one by evaluating a certain 'fitness' function on the individuals and favoring the best-performing ones.

The remaining of this paper is structured as follows. Section 2 details the dataset and presents the GP tool used in this work. Section 3 reports the results, and Section 4 states the conclusions.

2. Materials and Methods

We considered electrograms (EGM) from episodes registered by ICDs (sampling rate 128 Hz, 8-bit precision) in spontaneous and induced tachycardia [7, 8], collected in six spanish healthcare centers. The EGM source was HVA/HVB for all episodes, which were classified as VT or SVT by clinical experts. Features from every episode consisted of 38 real-valued samples associated to electric voltages in milivolts (mV). These samples were arranged as the concatenation of two beat segments, namely, a sinus rhythm template immediately previous to the arrhythmic episode (basal reference) and the arrhythmic episode template. Every beat segment was composed of 19 samples. Episodes were split into two independent datasets: one dataset to design the classifier (train dataset) and the another one to provide its generalization behaviour (test dataset). Figure 1 depicts some statistics for the feature space, taking into account the class (VT/SVT) and the dataset (train/test).

Attending to the usual criteria for arrhythmia discrimination in ICDs, the approach of this work can be considered as based on the association-dissotiation of atrium and ventricle (as opposed to those based on the morphological analysis, or on the cycle length).

The training dataset was composed by 38 SVT and 68 VT episodes from 16 patients with ICD (Medtronic Micro Jewel, models 7218, 7219, 7221, 7223 and 7271). The origin of the SVT episodes included effort tests, both post-implant tests and tracking checks carried out from one to three months after the intervention. VTs were associated both to spontaneous and induced episodes. The test dataset (used just for evaluation) was composed by spontaneous episodes (284 SVT and 1057 VT) from 28 patients with bicameral ICD (Medtronic Micro Jewel, model 7271).

We selected the GPLAB tool [9], a GP framework designed for Matlab, to perform our experiments. After exploring different values for crossover and mutation rates, we decided to use the default GPLAB configuration of adaptive operator rates, as exposed in [10] for genetic algorithms. The fitness function was the absolute error between the output provided by the tree and the desired output (true class). Desired outputs were coded as '1' (SVT) and '0' (VT). A threshold of '0.5' on the output provided by the tree was considered for classification purposes.

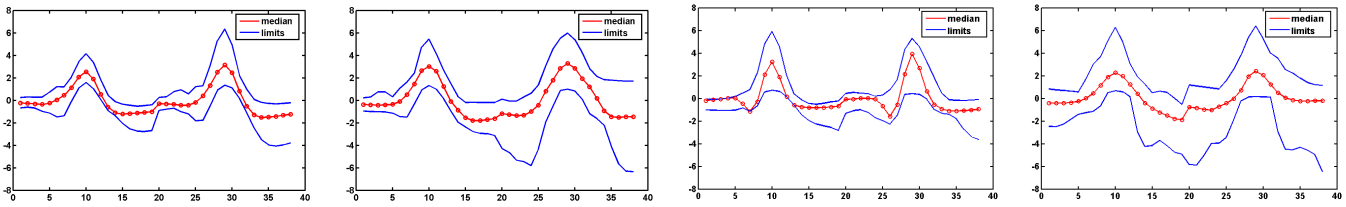
The evolutionary process was explored through 44 experiments, where the main differences were the set of operators used as tree nodes and the constraints for the leaves. Regarding the operators, different combinations of logical (e.g., and, or, not), arithmetical (e.g., plus, minus, times), comparison (e.g., equal, greater than, lower than) and trigonometrical (e.g, sine, cosine) functions where applied. As for the leaves, certain constants (zero, random number) were allowed, in addition to the 38 features defining the episode. A population size of 100 individuals was considered in all cases, with the maximum number of generations set to 50. For each experiment, 50 runs with different initial populations were performed.

3. Results

We present in this section the results obtained when applying GP to find criteria to discern between VT and SVT episodes, in order to be employed in embedded devices. Experiments provided decision trees with different computational complexity (i.e., function evaluations, directly related to the number of nodes) and error rates. Thus, error rates in the test dataset ranged from values around 2% to more than 35%. Experiments considering trigonometrical and logical functions provided the worst results. The use of constants in the leaf nodes did not provide a significant worsening neither improvement in the error rate.

Histograms in Figure 2 illustrate the variety of results provided by GP on the 50 runs when using two families of operators as nodes: Fam-A (operators 'equal', 'greater than' and 'times'), which provided fairly reasonable results; and Fam-B, which provided the best results by completing the set of previous operators ('equal', 'greater than', 'times', 'non equal', 'lower than', 'lower or equal', 'greater or equal','plus', 'minus'). Performance measurements used for comparison were sensitivity, specificity and error rate, all of them on the test dataset. Results with the remaining families of nodes are worse or similar to those presented in Figure 2.

The best solution (tree) found is shown in Figure 3 (a) and corresponds to Fam-B. This simple tree is constituted just for three nodes ('times','minus' and 'lower than') and four leaves (two samples from the sinus beat and another two ones from the arrhythmic beat). Specifically, it compares the difference of two voltages extracted from the end of the arrhythmic beat ($X_{33} - X_{22}$) with the product of two voltages extracted from the beginning of the sinus beat ($X_{12} * X_3$). The error rate provided by this tree was 4.7% in



(a) (b) (c) (d)

Figure 1. Representation of the highest/median/lowest voltage associated to the features of every class (VT/SVT) and dataset (train/test): (a) SVT-train; (b) VT-train; (c) SVT-test; (d) VT-test. The abscissa axis corresponds to the sample (feature) number, while the ordinate represents voltage in mV units. Median voltages are depicted in red dotted lines, while highest and lowest voltages are defined by blue lines.

train and 1.8% in test; specificity and sensitivity on the test dataset were 98.20% and 98.24%, respectively. It seems reasonable to state that this result is not due to overfitting, since the error rate for training is higher that for testing. For a better solution interpretation, Figure 3 (b) depicts two waveforms, one associated to VT and the another one to SVT. Voltages in this figure correspond to the median value for each feature and type of tachycardia. The analysis of this simple solution is in accordance with previous studies, which had shown the importance of the initial segment of the EGM for discerning between VT and SVT [7, 8], but no other criteria had been focused on the final segment of the rhythm.

Other solutions providing test error rates around 5% are also composed of simple operations, easily implemented by visual methods (e.g., comparison and subtraction). This shows that the interpretability of the result is not limited to the specific case presented in previous figures, but occurs commonly. Figure 4 corresponds to another tree in Fam-B providing an error rate of 4.5%. Note that this classifier also compares two values: one voltage of the sinus rhythm template and the difference of two voltages in the arrhythmic template. Referring again to Figure 3 (b), we observe that this solution compares the first part of the arrythmic episode with the end of the sinus rhythm.

Individuals with an error rate around 10% seem to classify incorrectly a common subset of SVT episodes. It is possible that this subset of episodes share some common characteristics that differentiate them from other SVT episodes. Specifically, it is suggested that this subset could correspond to branch block episodes, characterized by a wider QRS-complex, what usually leads to misclassification in some of the most used discriminators too.

To get an idea of the quality of the performance provided by the GP scheme, a variety of well-known machine learning techniques was considered for comparison: k-Nearest Neighbours (k-NN), Linear Logistic Model, Multi-Layer Perceptron (MLP), and Support Vector Machines (SVM) with different kernels (gaussian and polynomial). These classifiers were designed using the Weka sofware (available at http://www.cs.waikato.ac.nz/ml/weka) and the datasets described in Section 2. The test error rates provided by these approaches ranged from $2.6\%-$MLP to $11.8\% - k$-NN (SVM: 3.1% ; Linear Logistic Model: 6.8%). Note that the best results provided by an adequate GP technique are comparable to the best results provided by other machine learning approaches.

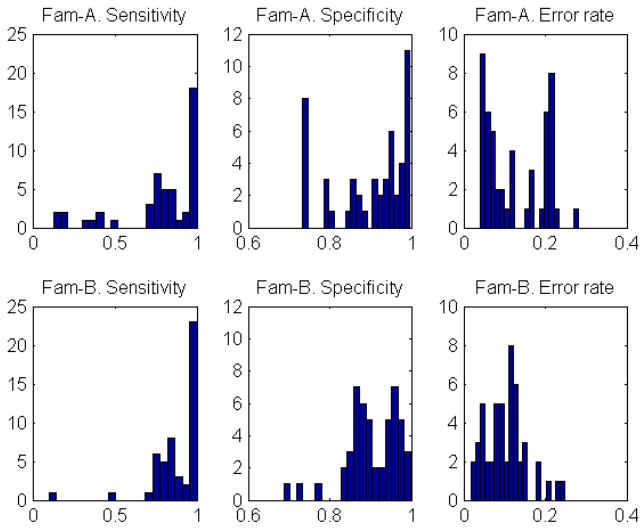


Figure 2. Histograms of three performance measurements (sensitivity, specificity and error rate) for two families of experiments on 50 runs: Fam-A (upper panel) and Fam-B (lower panel).

4. Discussion and Conclusions

This article shows that the use of evolutionary computing techniques may be a valid option when searching for simple and physiologically meaningful criteria for arrhythmia discrimination. It has also provided some examples of classifiers yielding error rates comparable to those obtained by more complex procedures in the literature (MLP or SVM, among others), with the advantage of physiolog-

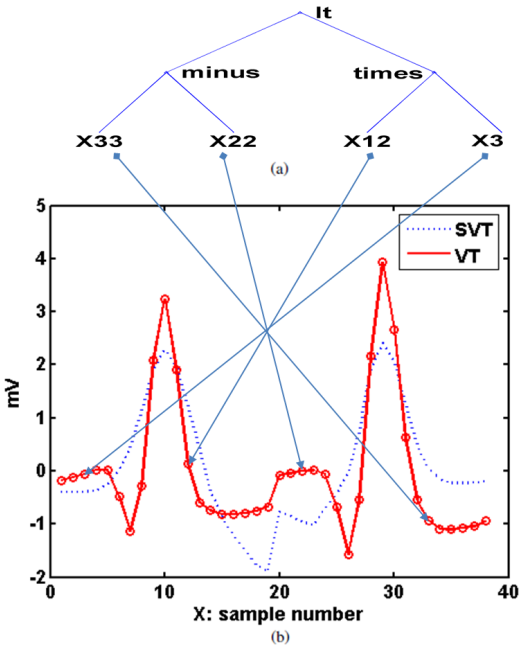


Figure 3. Best tree found, confronted with related samples: (a) Decision tree standing for the formula "SVT := (X33-X22)<(X12*X3)"; (b) Median value for samples of the SVT (dotted blue) and VT (red) templates.

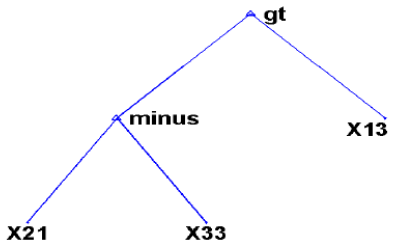


Figure 4. Decision tree standing for the formula "SVT := (X21-X33)>X13". Error rate of 4.5%.

ical interpretability.

Though the training set used in this work have data collected from induced episodes, the performance turned out to be quite good even when tested with just spontaneous episodes. The best classifier found in this work yielded an error rate of 1.8%, with both sensitivity and specificity greater than 98%, thus obtaining a very balanced discrimination. Experiments showed that, in general, the GP technique was able to provide a successful generalization with interpretability when operators are adequate. On the other hand, the intuition that a set of misclassified SVT episodes both with the GP solutions and with other criteria are associated to branch blocking paves the way for a new research line determining the differences between SVTs with and without branch blocking.

Finally, the qualitative analyses carried out on some GP solutions show that the information for arrhythmia discrimination is not only at the beginning of the ventricular activation, but also at the final segment of the sinus rhythm. This result had not been previously indicated in the literature before, and we hope it will open new ways for further improvement of arrhythmia discrimination criteria.

Acknowledgements

This work has been partly supported by Spanish Research Projects TEC2010-19263, TEC2013-48439-C4-1-R, and by Prometeo Project of the Secretariat for Higher Education, Science, Technology and Innovation of the Republic of Ecuador.

References

[1] Finelli CJ, Jenkins JM, DiCarlo LA. Detection and identification of cardiac arrhythmias using an adaptive, linear-predictive filter. Computers in Cardiology 1993; 177–80. IEEE Computer Society Press.

[2] Stevenson SA, Jenkins JM, DiCarlo LA. Analysis of the intraventricular electrogram for differentiation of distinct monomorphic ventricular arrhythmias. Pacing and Clinical Electrophysiology 1997; 20: 2730–8.

[3] Klingenheben T, Sticherling C, Skupin M, Hohnloser SH. Intracardiac QRS electrogram width–an arrhythmia detection feature for implantable cardioverter defibrillators: exercise induced variation as a base for device programming. Pacing and Clinical Electrophysiology 1998; 21(8): 1609–17.

[4] Minami K, Nakajima H, Toyoshima T. Real-time discrimination of ventricular tachyarrhythmia with Fourier-transform neural network. IEEE Transactions on Biomedical Engineering 1999; 46(2): 179–85.

[5] Eiben AE, Smith JE. Introduction to Evolutionary Computing. Springer, 2003.

[6] Koza JR. Genetic programming: A paradigm for genetically breeding populations of computer programs to solve problems. Technical Report. Stanford University. Stanford, CA, USA, 1990.

[7] Rojo-Álvarez JL, Arenal-Maíz A, Artés-Rodríguez A. Discriminating between supraventricular and ventricular tachycardias from EGM onset analysis. IEEE Engineering in Medicine and Biology Magazine 2002; 21(1): 16–26.

[8] Rojo-Álvarez JL, Arenal-Maíz A, Artés-Rodríguez A. Support vector black-box interpretation in ventricular arrhythmia discrimination. IEEE Engineering in Medicine and Biology Magazine 2002; 21(1): 27–35.

[9] Silva S, Almeida J. GPLAB - A genetic programming toolbox for MATLAB. In Proc from the Nordic MATLAB Conference 2003; 273–278. Copenhagen, Denmark.

[10] Lawrence D. Adapting operator probabilities in genetic algorithms. Proc 3rd Intl Conf on Genetic Algorithms 1989; 61–9.