

Artificial Intelligence Methods Applied to Parameter Detection of Atrial Fibrillation

D Arotaritei and C Rotariu

Department of Biomedical Sciences, “Grigore T. Popa” University of Medicine and Pharmacy, Iasi, Romania

E-mail: cristian.rotariu@umfiasi.ro

Abstract. In this paper we present a novel method to develop an atrial fibrillation (AF) based on statistical descriptors and hybrid neuro-fuzzy and crisp system. The inference of system produce rules of type if-then-else that are extracted to construct a binary decision system: normal of atrial fibrillation. We use TPR (Turning Point Ratio), SE (Shannon Entropy) and RMSSD (Root Mean Square of Successive Differences) along with a new descriptor, Teager-Kaiser energy, in order to improve the accuracy of detection. The descriptors are calculated over a sliding window that produce very large number of vectors (massive dataset) used by classifier. The length of window is a crisp descriptor meanwhile the rest of descriptors are interval-valued type. The parameters of hybrid system are adapted using Genetic Algorithm (GA) algorithm with fitness single objective target: highest values for sensibility and sensitivity. The rules are extracted and they are part of the decision system. The proposed method was tested using the Physionet MIT-BIH Atrial Fibrillation Database and the experimental results revealed a good accuracy of AF detection in terms of sensitivity and specificity (above 90%).

1. Introduction

There are various methods proposed to detect atrial fibrillation (AF). The linear/nonlinear methods using Heart Rate Variability (HRV), Poincaré map, Spectral analysis, fractals and wavelets are the most used in publications.

A knowledge-based method is proposed in [1] using arrhythmic (ARH) episode detection for six rhythm types. The algorithm implies many rules and thresholds and a finite automaton is used for detection and classification. The algorithm reaches 98% accuracy for ARH classification and 94% accuracy for ARH classification and detection using MIT-BIH arrhythmia database. A methodology based on generalized discriminant analysis and for reduction of number of features is proposed in [2]. Artificial intelligence offers a variety of methods, tools and algorithms than can overcome some of the drawbacks of the classic methods. Among these methods, neural networks (NN) and evolutionary algorithms (especially genetic algorithms – GA) are the most popular. Some of the methods take the advantage of intelligent reasoning as in fuzzy systems and some combinations NN-fuzzy system proved to be successful solution for classification problems. A feed forward NN was used in [3] as classifier for AF in three classes: sinus rhythm, supraventricular extra systoles and/or ventricular extra systoles. Fourier transform was used to observe the changes in QRS rhythm and NN was used to classify the rhythm in three classes with sensitivity and specificity > 98% [4]. Short-time generalized dimensions for cardiac rhythm characterization was proposed in [5].



In this paper we propose an evolutionary for parameter optimization using genetic algorithms (GA) for a situation inspired by [6].

2. Data Pre-processing

The most common linear methods used in HRV are based on RR intervals. Other combination of methods as TPR (Turning Point Ratio) and Shannon Entropy along with RMSSD have been proposed in [6]. We will use the definition of TPR, SE and RMSSD as they have been proposed in [6] along with a new descriptor, Teager-Kaiser (TK) energy in order to improve the accuracy of detection. TK energy has been used in very few papers related to HRV analysis [7]. TK is a nonlinear energy and its value increases if tone of signal increases at the same amplitude. The higher is frequency, the higher is the value of TK operator. We propose to use the formula given in [7], with regularization depending on length of window. We denote by x a sequence of n values.

$$TK\{x(k)\} = x^2(k) - x(k-1) \cdot x(k+1) \quad (1)$$

The average energy over a window of length L is defined by:

$$ATK_L = \left(\sum_i TK_i \right) / n^2 \quad (2)$$

where L is the window length, and n is the number of values in L window. The bigger is number of i 's the higher is the sum ATK_L . The signals are first filtered from ectopic beats, reducing the source of false detection of AF.

The tradeoff between benefits and disadvantages in our case showed that is better to remove them. There are some proposals to filter ectopic beats and outliers ([6], [8], and [9]). We have been chosen the method from [6] but for treatment of outliers we proposed to replace the outlier the average value of left and right valid RR instead of remove them as in [6].

After filtration, all the RR segments are not linked together as in [5], but the files that contains the RR segments are individual records for the same data test. The set of parameters are focused on length of window L because an optimal window can produce a relevant value of randomness of time series that can be interpreted as AF [10]. A set P of k parameters over a window L can be defined as $P = [L, p_1, p_2, \dots, p_k]$. Optimization of some parameters can be conflictual that is if optimization of objective using one parameter can produce a decrease of optimized value for other parameters. A tradeoff using optimized curve of values can be used to choose a correct tradeoff (e.g. Pareto front).

3. Optimization using genetic algorithms (GA)

The selected parameters are those from [6], L – length of window, P_{th1} – threshold from where the segment is considered to be AF (number of AF segments/number of total segments), $RMSt$ – threshold of Root Mean Square of Successive Differences, $TPRt$ threshold of Turning Point Ratio, SEt - threshold of Shannon Entropy and TKt - threshold of Teager-Kaiser energy. We must remark that usual method used in [6] is to split each interval in a number of slots and use all the combination to choose an optimal one that is a large number of repeated calculi.

The first question is if an individual HRV descriptor is enough for a good detection of AF, that is the sensitivity Se and specificity Sp are both enough close to 1.0 for a correct classification.

The main classification problem is that there is overlapping of N (Normal) values over AF (Atrial Fibrillation) values. A selection of type below is hard to be successfully defined. In these conditions, taking into account more descriptors can be a solution: it is possible that for a window of given length, an overlapping value of one descriptor that conducts to an imprecise decision to be solved using another descriptor for that decision is precise. This study will be developed in the further research.

The parameters used by the developed classifier are given by: $P = [L, P_{thresh}, mRMSSD_{thresh}, TPR_{thresh}, SE_{thresh}, ATK_{thresh}]$, where $mRMSSD = RMSSD/n$, a normalized RMSSD value.

We propose to use GA [11] for optimization in order to make both o more efficient and more compact procedure, easier to use for more than 6 variables. The operations are the known ones: selection using tournament method, crossover (in two points) and fitness calculation ([11]-[12]). The initialization of population with individuals is made using random generator. The last term, ATK is appending to improve the performance of decision system [6].

IF ($mRMSSD > mRMSSD_{\text{thresh}}$) AND (TPR in confidence interval TPR_{thresh}) AND ($SE > SE_{\text{thresh}}$)
 AND ($ATK > ATK_{\text{thresh}}$) THEN (segment $\in$ AF, atrial fibrillation)
 ELSE (segment $\in$ N, normal).

The rule from above decision system is similar to [5] except the last decision $ATK > ATK_{\text{thresh}}$.

4. Extraction of rules used in neuro-fuzzy inference system

Neuro-fuzzy systems are known to be universal approximators [13]. We propose to use ANFIS [14] as classifier for two classes of parameters: the class that identify the AN and the class for normal stage. The neuro-fuzzy inference system is used after that to extract fuzzy rules of type if-then-else and the membership functions are chosen to be triangular. In our approach, the output is a crisp value that are interpreted as discriminator between normal and AF segment.

There are many papers that propose to extract rules from structures that are proposed by artificial intelligence approach [15]. Decision trees are also used in medical decision system and in medical decision support system (DSS). Construction of decision tree can be based on crisp or fuzzy approach. Automatic rules extraction from ANFIS fuzzy inference system is proposed to be developed in this paper. The output y is set to < 0.5 if segment of RR is normal and output y is set to ≥ 0.5 if segment of RR is associated with AF. In training stage, the values of thresholds for classification are set to 0.4 and 0.6 respectively.

5. Experimental results

The MIT-BIH Atrial Fibrillation database was used in our experiments. We select 9 files due to clarity of annotation of AF/N segments and the fact that record are complete with known start and no unread blocks. In the first approach we used the decision rules from [5], but the thresholds are identified using a systematic algorithm, heuristic optimization using GA. GA proved to be a good choice for optimization problem, single objective or multi-objective. The parameters are set to be limited by lower values [32, 0.1%, 0.01, 0.01%, 0.01, 0.01] and upper values [480, 99.99%, 0.999, 99.99%, 0.999, 0.999]. The chromosome coded for first value is mapping linearly into integer values between 32 and 480. The objective function is the maximum Se and Sp that is, these variables must have the closet values to 1.0. There are clearly to objective, possible contradictory, so the Pareto front offers to us a possibility to choose optimal values for both variables.

We start with a population of 50 individuals and after 40 iterations the solution is obtained. The multi-objective function is minimization of both functions $1-Se$ and $1-Sp$. The results obtained are $Se=90.17$ and $Sp = 90.12$, are slightly less than those reported in [6], but there are using a less number of files. The values are acceptable for a good decision system. Moreover, the method can be extended to test more parameters avoiding possible explosive number of combinations as in [5].

Another test used the first five parameters without ATK. The results are slightly less performant, $Se=90.01$ and $Sp = 88.76$, and the proposed improvement by adding ATK proved that the new descriptor drive on to a better results. In this case, 21 windows of length 320 (the optimal one) were misclassified and as follow the performance is decreased.

6. Conclusions

We proposed to use GA to optimize values of parameters for detection of AF. A new parameter, Teager operator is introduced to optimize the detection. The selection of parameters is based on assumption that in AF, somewhat randomness characteristics are developed in RR records. We

proposed to use a systematic method to optimize the parameter thresholds for a decision system. Parameters can increase with minor modification of the algorithm that is, adding new descriptors supposed to be in two limits to detect AF, a simple new parameter is added in GA and the simulation software can run with minimal modifications.

The rules are however not optimized in the case of ANFIS inference that is the universe of discourse for each parameter is split in the same number of membership functions (e.g. 3). The algorithm for pruning rules and number of membership function will be developed in the further research.

Acknowledgement

The work has been carried out within the program Joint Applied Research Projects, funded by the Romanian National Authority for Scientific Research (MEN – UEFISCDI), contract PN-II-PT-PCCA-2013-4-0761, no. 21/2014 (SIACT).

References

- [1] M.G. Tsipouras, D.I. Fotiadis, and D. Sideris, An arrhythmia classification system based on the RR-interval signal, *Artificial Intelligence in Medicine*, vol. 33, pp. 237-250, 2005.
- [2] B. M. Asl, S. K. Setarehdan, and M. Mohebbi, "Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal, *Artificial Intelligence in Medicine*," vol. 44, pp. 51-54, 2008.
- [3] T.F.Yang, B. Devine, and P.W. Macfarlane, "Artificial neural networks for the diagnosis of atrial fibrillation," *Med Biol Eng Comput.*, vol. 32, no. 6, pp. 615-619, 1994.
- [4] K. Minami, H. Nakajima, and T. Toyoshima, Real-time discrimination of ventricular tachyarrhythmia with Fourier transform neural network, *IEEE Trans Biomed Eng*, vol. 46, no. 2, pp. 179-185, 1999.
- [5] Y. Wang, Y.S. Zhu, N.V.Thakor, and Y.H. Xu, A short-time multifractal approach for arrhythmia detection based on fuzzy neural network, *IEEE Trans Biomed Eng*, vol. 48, no. 9, pp.989-995, 2001.
- [6] S. Dash, K.H. Chon, S. Lu, E.A. Raeder, Automatic Real Time Detection of Atrial Fibrillation, *Annals of Biomedical Engineering*, vol. 37, issue 9, pp. 1701-1709, 2009.
- [7] M. D'Alessandro, et. all., "Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: A report of four patients," *IEEE Transactions on Biomedical Engineering*, vol. 50, no. 5, pp. 603-615, 2003.
- [8] R. Karlsson, R. Hörnsten, A. Rydberg, and U. Wiklund, Automatic filtering of outliers in RR intervals before analysis of heart rate variability in Holter recordings: a comparison with carefully edited data, *BioMedical Engineering OnLine*, pp. 1:12, 2012.
- [9] N. Storck, M. Ericson, L. Lindblad, and M. Jensen-Urstad M, Automatic computerized analysis of heart rate variability with digital filtering of ectopic beats, *Clin. Physiol.*, vol. 21, no. 1, pp. 15-24, 2001.
- [10] J.H. Zar, Serial Randomness of measurements: nonparametric testing, In: *Biostatistical Analysis*, Prentice-Hall, pp. 587-588, 1985.
- [11] Z. Michalewicz, *Genetic Algorithms + Data Structures = Evolution Programs*, Springer, 1996.
- [12] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Reading, MA: Addison-Wesley Professional, 1989.
- [13] Castro, J.L., Delgado, M.: Fuzzy systems with defuzzification are Universal approximators. In: *IEEE Transactions on Systems, Man and Cybernetic*, vol. 26, no.1, pp. 149-152 (1996)
- [14] J.-S. Roger Jang, ANFIS: Adaptive-network-based fuzzy inference system, *IEEE Transactions on System, Man, and Cybernetics*, No. 5, Vol. 23, PP. 665-685, May 1993.
- [15] S. Kulluk, L. Özbakır, A. Baykasoglu, Fuzzy DIFACONN-miner: A novel approach for fuzzy rule extraction from neural networks, *Expert Systems with Applications* vol. 40 pp. 938–946, 2013.