

COMPLEXITY MEASURE FOR RR INTERVALS, DESCRIPTOR FOR ARRHYTHMIA CLASSIFICATION

Dragoş AROTĂRIŢEI^{1,*}, Cristian ROTARIU², Alexandru PĂSĂRICĂ³

¹⁾ PhD, Prof., Department of Biomedical Sciences, "Grigore T. Popa" University of Medicine and Pharmacy, Iasi, Romania

²⁾ Ass. Prof., Department of Biomedical Sciences, "Grigore T. Popa" University of Medicine and Pharmacy, Iasi, Romania

³⁾ PhD, Department of Telecommunications, Technical University of Iasi, Iasi, Romania

Abstract: *The complexity measure is associated with a certain level of randomness for a given discrete sequence of symbols. This assertion was applied to a RR sequence of intervals extracted from preprocessed ECG in order to classify different types of arrhythmias. Three complexity measures were used: sample entropy (the most known), Lempel-Ziv (single level and multilevel) and a less known one, T-Code. The expert system is made by a binary tree classifier that discriminate the class to what arrhythmia belongs. A comparison with other method in term of performance is made also.*

Key words: *complexity measure, randomness, RR sequence, ECG, binary decision tree.*

1. INTRODUCTION

Atrial fibrillation (AF) can be detected by various methods and techniques. Most of them use the RR intervals (the interval from two successive Rs peaks from electrocardiogram – ECG) but other methods have been proposed also. Among them, the linear/nonlinear methods using Heart Rate Variability (HRV), Spectral analysis, fractals and wavelets, Poincaré map, are the most used in published papers. The algorithms and are tested using customized records that are not often accessible to other researches but the most common benchmark used is MIT-BIH arrhythmia database.

The usage of P waveform extracted from ECG signal even very efficient in arrhythmia detection as some authors claimed but it is not too easy to be applied in wireless patient monitoring using microcontrollers or wearable devices. 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.

In [2], the authors proposed a methodology based on generalized discriminant analysis in order to reduce the number of selected features and after that, the classification is made using support vector machine for six types of ARH. The authors reported very good results, a precision of 98.51% for classification of atrial fibrillation types [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), evolutionary algorithms (genetic algorithms – GA, differential genetic algorithms – DGA) and clustering algorithms are the most popular. A fusion between some of the methods can take the advantage of intelligent reasoning as in fuzzy systems and power of calculus from neural networks. The results, combinations NN-fuzzy system proved to successful solution for many classification problems that are intractable by using other classical methods.

A simple feedforward neural network (NN) was used in [3] as classifier for a set of three AF classes: sinus rhythm, supraventricular extra systoles and/or ventricular extra systoles. The changes QRS rhythm was detected by applying the Discrete Fourier Transform (DFT) and classification by NN conducts to good results, sensitivity and specificity greater than 98% [4].

Dimensionality of the vector of features can improve in many cases the precision and accuracy of classification but the calculus for all the features can increase the calculus effort and some application can process these information only offline. A variant that used short-time generalized dimensions for cardiac rhythm characterization was proposed in [5]. To classify different type of arrhythmias a new fuzzy Kohonen network has been proposed in [5].

In offline analysis, the annotated ECG digitized signals or RR interval string are usually linked together in a very large series of values. A sliding window of a fixed length is used to scan the string of values in order to detect the AF and to classify of what type of ARH belong the record from window. One of the most important tasks for AF detection is he parameter set for classifier that (usually there are simply thresholds with binary decision) involved in rules or other parameters like window length [6]. Different methods are used most of them based on empirical evaluation, statistical assessments (e.g. ROC curves) or exhaustive combination of intervals that split the range of parameters in a finer divisions [6].

* Corresponding author: str. Kogalniceanu, nr. 9–13, Iasi, Romania,
Tel.: +40.232.213.573;
Fax: +40.232.213.573
E-mail addresses: dragos.arotaritei@gmail.com (D. Arotaritei).

The evaluation for all combination of steps can be very consuming time and moreover, it is possible that the solution to be near optimal if divisions are not enough fine.

The most common linear methods used in HRV are (based on RR interval): NN50, pNN50, RMSSD, SDNN, SDDSD, SDANN, LTI and STV. Other combination of methods as TPR (Turning Point Ratio) and Shannon Entropy along with RMSSD have been proposed in [6]. Detailed explanations of notation and a formulas present above are given in [6].

The HRV records are measured frequently using RR interval and the randomness of sequences can indicate an AF. HRV itself has a certain level of randomness, so a discrimination of different level of randomness must be enough accurate to produce a satisfactory classification results (e.g. an acceptable level of sensitivity and specificity as statistical measure of classifier performance).

There are very few papers that deal with complexity measure for time series to identify arrhythmic pulse patterns. The Lempel-Ziv algorithm was used to identify seven patterns in traditional Chinese pulse diagnostic (TPCD), but the classification algorithms proposed in [7] is very complex and difficulties to implement are evident.

Three types of rhythm are analyzed in [8]; types that belong to nonlinear processes and the measure of complexity are associated using Lempel-Ziv algorithm [8]. The maximum specificity and sensitivity is aimed by authors to be 100% in 7th stage of computation.

EEG complexity was used as measure of depth of patient's anesthesia in [9]. Depending on anesthetic technique, the sensitivity and specificity were between 93% and 100%.

2. ECG PREPROCESSING

A major source of possible errors in ARH detection and ARH classification are outliers and ectopic beats. The ECG signals are usually filtered by hardware methods and the basic a notch filter for 50 Hz frequency is used to remove to remove the power line interference.

The RR intervals are extracted from all ECG database using the well known Pan-Tompkins algorithm [11] and linked together.

The signal is first filtered from ectopic beats. Ectopic beats are often the source of false detection of AF events but in some cases can be a valuable source of information for medical diagnosis. Taking into account the advantages and disadvantages of presence of ectopic beats can conduct in some situation that it better to remove them, and it is our case.

The problem of ectopic beats is itself a subject of research and there are different approaches that deal with this subject [12–17]. Ectopic beats (ectopic rhythm) are small irregular changes in normal heart rhythm due to some premature or extra heartbeats. The most common ectopic beats are: premature atrial contraction (PAC), premature ventricular contraction (PVC) and extrasystole.

In what follows the method proposed in [12] will be adopted, adapted to our purpose using RR records. The premature beats are identified by a short-long sequence,

premature and compensatory pause. The $RR[i]/RR[i - 1]$ ratio (where i represent the index of RRs) is calculated. The Perc99, Perc25 and Perc1 percentiles are used to discriminate ectopic beats from normal small fluctuations that are present due to physiologic variability [12]. After identification of ectopic beats, the RRs segments are removed from total RRs sequence.

The RR sequence can have also outliers [18], abnormal values that are greater than the normal RR range of values. Two main methods are actually used: the outliers are deleted from sequence and the outliers are replaced with a value that depends on the values that are borders of outliers. Both methods are approximations or the real RRs sequence, the second one being preferred in order to avoid the possibility to introduce high frequencies when signal is reconstructed (aliasing phenomenon). The outlier is detected when a value of RR is greater than a threshold $th_{outlier}$ find empirical form median value of all RRs.

3. COMPLEXITY MEASURE

The complexity measure associated with a certain level of randomness for a given discrete sequence of symbols. Based on this assumption, depending on how the randomness is defined, the complexity can be used for evaluate the level irregularities in electrocardiogram that are usually the major source of different types of arrhythmias.

It should be noted that other types of measure of randomness are proposed in literature, most of them based on statistical methods, like TPR (turning point ratio) proposed by [19] but they are not connected with complexity of information present in coded strings.

All the complexity measures operated over a finite set of strings that are coded using threshold level in order to convert into sequence with less number of symbols using a predefined alphabet. The most common conversion is based on two symbols alphabet, usually $A = \{ '0', '1' \}$ producing binary strings. There are multilevel conversion also, based on alphabet with more than two symbols but there are only few papers that deal with it [20].

The main reason for implementation of a single level conversion is the fact that the multilevel variant involves a much more complicated algorithm and calculus and as sequel, the implementation is very difficult.

Three types of complexity are investigated: sample entropy (the most known), Lempel-Ziv (single level and multilevel) and a less known one, T-Code. The authors propose in this paper a novel usage of T-Code, as index for arrhythmia detection on ECG database, an arrhythmia database, 48 files, 109236 records, 13 classes used for arrhythmia classification.

3.1. Sample entropy

Sample entropy (SampEn) is a measure of complexity used especially for discrete time series that represent physiological signals [21, 22]. Different from approximate entropy (ApEn), SampEn does not count the self-similar patterns.

For a time series $X = (x_1, x_2, \dots, x_n)$ let us denote by A the number of template vectors that matches of length $m + 1$ and B the number of template vectors that matches

of length m . A sub-vector of length m that belong to vector X can be defined as $X_{i,m} = (x_i, x_{i+1}, \dots, x_{i+m-1})$. The matching is usually based on a metric distance that is usually Chebyshev distance (d_c) but other metrics (Euclidean, Minkowski, etc.) can be used. $SampEn$ is the negative value of natural logarithm of two numbers A and B [22]:

$$SampEn = -\log \frac{A}{B}; \quad (1)$$

$$d_c(X_{i,m}, X_{j,m}) = \max_{k=1, \dots, m} (x_{i+k}, x_{j+k}); \quad (2)$$

$$d_c(X_{i,m}, X_{j,m}) < r. \quad (3)$$

In eq. (3) it is denoted by r the tolerance for m points in direct relation with the length of vectors that are compared for match. Length of windows and distance measure (d_c) for comparison are the main parameters that are passed to all implementations that calculate $SampEn$ for time series including biomedical signal processing [23].

3.2. Lempel-Ziv complexity analysis

Lempel-Ziv complexity analysis is used to evaluate the randomness of a finite discrete time series. Lempel-Ziv algorithm has as main result a complexity counter $c(n)$ which is the number of distinct patterns in a time series sequence [24–25].

The first step in to calculate $c(n)$ is the transformation of a numerical sequence into a symbolic sequence among the most popular one is 0–1, the level of separation being a simple threshold S_d . If the value of signal is greater or equal with S_d , the symbol is 1, otherwise the symbol is 0. One good choice for S_d that practically is valid for all the complexity measurement algorithms is the median value of numerical time series [26]. The symbolic sequence obtained after coding is then parsed from left to right to get to distinct words. The approach is consistent with older results published in [27].

The implementation of Lev-Ziv algorithm follows the steps from [24]. Let us denote by $S = s_1 s_2 \dots s_n$ a finite length of symbols, $S(i,j) = s_i s_{i+1} \dots s_j \in S$, $i < j$ one substring that start at i point and end at position j , and a null set $S(i,j) = \{\}$. Let Q and R substrings of type $S(i,j)$, QR the concatenation of them and QRD a sequence after last character of QR is deleted (D is the notation for deleted). The string S is parsed from left to right in a scan procedure. It is denoted by $B(S)$ the set of basic words. A substring $S(i,j)$ is compared with the strings from $B(S)$, substrings up to $j - 1$, that is $S(i, j-1)$. If $S(i, j)$ is present, no new component is present, $B(S(i, j-1))$ is updated to $B(S(i, j))$, and $S(i, j)$ is updated to $S(i, j+1)$, and the process is repeated until the end of string. If $S(i, j)$ is not present, a new component is found and a dot is placed after $S(j)$. These dots mark the new distinct string found in the parsing process. The process repeats until $j = n$, the length of string S . The start process begins with $S(1,1)$, the first symbol in the string S . For example, the string 0111010100010 is parsed as 0·1·11·010·100·010.

For 0–1 sequence, the normalized complexity Lempel-Ziv measure is given by:

$$C(n) = c(n)/(c(n)/\log_2(n)). \quad (4)$$

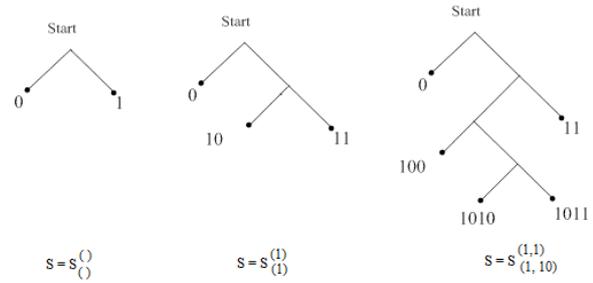


Fig. 1. Intermediate T-code, set for $S_{(1,10)}^{(1,1)} = \{0, 11, 1100, 1010, 1011\}$

For string 0111010100010, the complexity measure is calculated for $c(n) = 5$ and $n = 13$, so $C(n) = 1.7079$.

3.3. T – code complexity measure

T-complexity is based on T-codes, a method of self-synchronizing code initially introduced by Tichener [5]. T-code augmentation starts with copying the symbols of the alphabet in the first level of augmentation. The next steps are constructed recursively by deleting a chosen T-prefix and append it to codewords [28–30]. The inverse operation, T-decomposition parses the sequence in T-prefixes, the number of prefixes being used in complexity calculation.

The T-code is applied to substrings using a finite length window W . It is denoted by A a finite alphabet set, uw the concatenation of two strings u and w , and uk the concatenation of k copies of string u . The series of T-code is constructed using a recursive formula [30]:

$$S_j = \bigcup_{i=1}^{k_j} \{P_j^i s \mid s \in S_{j-1} \setminus \{P_j\}\} \cup \{P_j^{k_j+1}\}. \quad (5)$$

where $S_0 = A$, P_j – prefix, P_j – string, k_j – positive integer number, and s – string. The common notation used in tree is $S_{(P_1, P_2, \dots, P_n)}^{(k_1, k_2, \dots, k_n)}$. If the finite alphabet is $A = \{0,1\}$, an example of T-augmentation as a tree is given in Fig. 1.

Using a parsing algorithm named T-decomposition, for a given string the sequence is parsing in the form:

$$s = P_n^{k_n} P_{n-1}^{k_{n-1}} \dots P_1^{k_1} A; \quad (6)$$

$$tc = \sum_{j=1}^n \log_2(k_j + 1). \quad (7)$$

The T-complexity of string s as it is given by eq. (6) is defined as in eq. (7).

4. EXPERIMENTAL RESULTS

A selection of records from arrhythmia database was used, consisting of 48 files, 109236 records, 47 records removed after ectopic detection, 13 classes used for arrhythmia classification $C_{ARRH} = \{!, ', /, 'A', 'E', 'F', 'J', 'L', 'N', 'Q', 'R', 'S', 'V', 'a', 'e', 'f, 'j\}$.

The ECG signal is preprocessed before to extract RR intervals. The ectopic beats were filtered using method described in [6] and the outliers were detected and eliminated using a method inspired by [31]. Based on statistical analysis, if an absolute RR value is greater than $1.75 \cdot RR_{\text{median}}$, the value is considered outlier and it is eliminated.

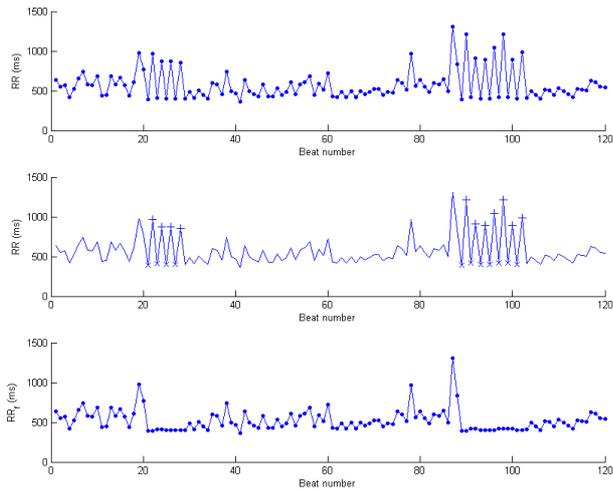


Fig. 1. Beat sequence from 8219 MIT AFIB database with ectopic beats and replacement of them by average values.

The RR intervals are linked together in long time series string. Different from [6], the removed ectopic beat is replaced by average value of left and right border of ectopic beat.

The codification is made using a variable threshold that split the values of input string RR in sequence of '0' and '1' and a variable length of sliding window W (number of samples) that extracts a sequence to calculate its complexity. The RR range values is $RR \in [78, 2451]$ msec, with threshold step $th = 124$ msec equally spaced in range, and the length of window is set to $W = \{16, 24, 32, 40, 48, 56, 64, 128, 196, 256, 360, 400\}$. The greater is window, the better accuracy is supposed to be acquired but if the window is too large, the computational effort is substantially increased (time of computation increase also) and the efficiency of solution decrease also.

The proposal of algorithms is tested and the performance is compared with other algorithms from literature in Table 1.

The implementation of code is complex for all three cases, so the computational time is also important for all algorithms. Computational times for different windows length for all three algorithms are showed in Figs. 2–4.

Table 1

Title of table Performance of proposed solution in comparison with other algorithms

Algorithm	Performance over all records	Number of classes
Lempel-Ziv algorithm [8]	100%	3
Autoregressive modeling GLM [32]	93.2% – 100%	6
Prony algorithm [33]	95.24% – 97.78%	3
SVM (Support Vector Machine) [34]	99.307% – 99.883%	6
PPM algorithm [35]	91,74% – 99,37%	4
Sample Entropy (proposed)	99.5% – 100%	13
Lempel-Ziv algorithm (proposed)	97% – 100%	13
T-code (proposed)	98% – 100%	13

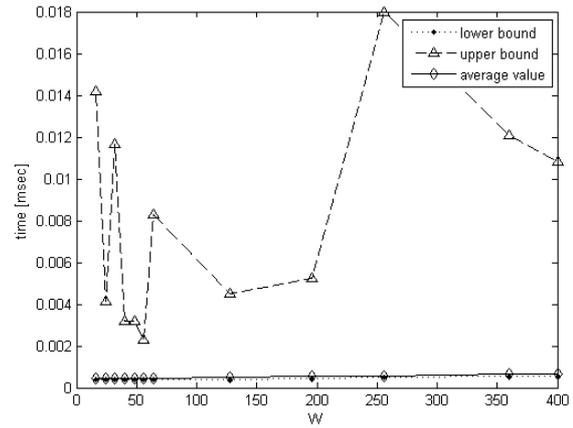


Fig. 2. Computational time for different windows length, T-code complexity analysis.

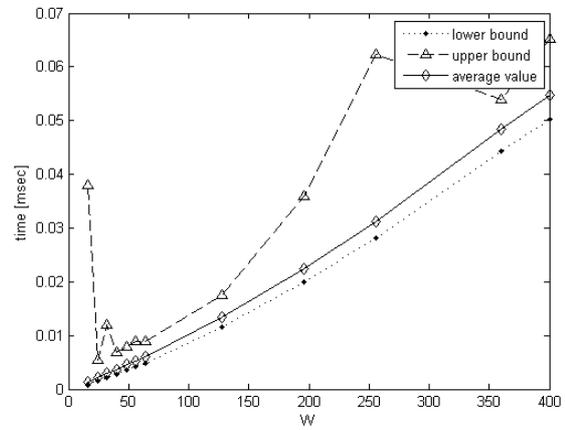


Fig. 3. Computational time for different windows length, Lempel-Ziv complexity analysis (exhaustive).

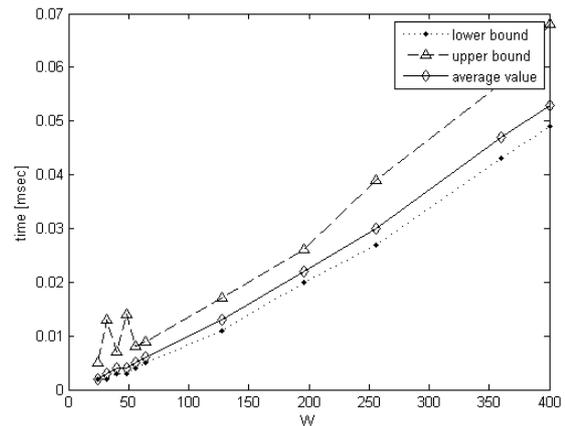


Fig. 4. Computational time for different windows length, Sample Entropy complexity analysis.

The computational time is considered for a number of 256×2^{10} number of windows of length that belong to $W = \{16, 24, 32, 40, 48, 56, 64, 128, 196, 256, 360, 400\}$, windows constructed from random strings of '1' and '0'. It can be seen that for a reasonable window length (around 100 RRs sequence), the upper bound of T-Code is the most time consuming. But in average

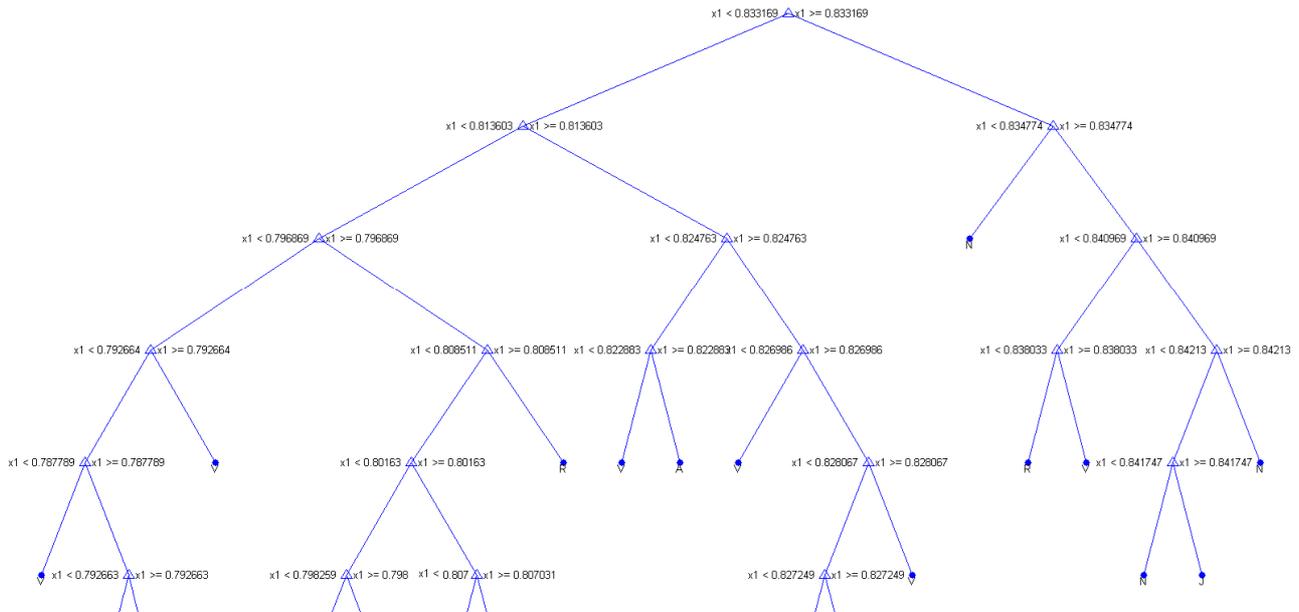


Fig. 5. A zoom from classification tree (right part from root node) for the maximum performance of classification and minimum number of nodes, T-code, $W = 240$.

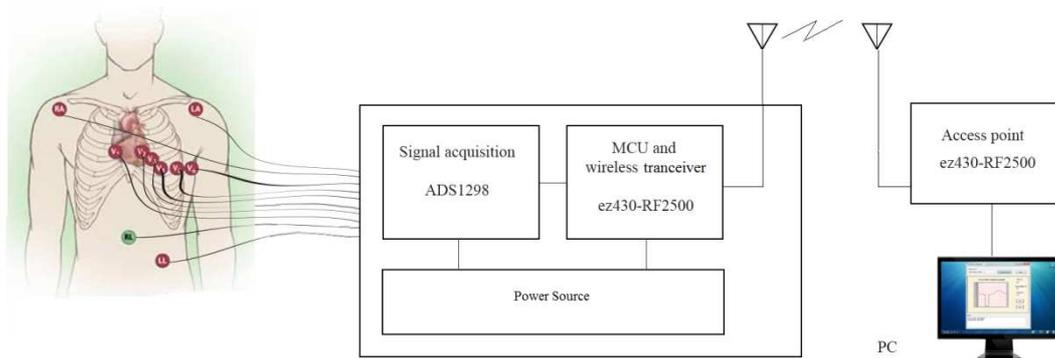


Fig. 6. The block diagram of the wireless device for ECG acquisition in order to detect and classify different types of arrhythmias.

T-code is most advantageous in rapport with Lempel-Ziv and Sample Entropy complexity analysis.

A tree classifier was used having the input only one variable and the output the classes C_{ARH} . The most performing tree was selected as a three which has the best performance of classification and a minimum complexity structure that is the three which has the minimum number of nodes (Fig. 5). In Fig. 5 the performance for T-code complexity analysis was represented. The trees for other two methods (Lempel-Ziv and Sample Entropy) have the number of nodes in approximatively same range of values as the T-code case.

It is clearly that for any of these complexity methods, even it is very good the size of decision tree is great. The simple usage of one value makes difficult an efficient implementation in real-time the decision at wearable device level (Fig. 6). As sequel, the microcontroller will send only RR raw data and the detection of arrhythmia the class to which belong the arrhythmia will be made remote, by PC access point (Fig. 6).

7. CONCLUSIONS

In this paper the possibility of arrhythmia classification using a single discriminant value given by complexity analysis of codified RRs raw sequence was investigated. The results proved that all the three methods (Sample Entropy, Lempel-Ziv and T-code) are very good in percent of correct classifications for 13 types of arrhythmias.

In order to reduce the numbers of nodes in decision tree, an extended vector with other simple discriminants (easily computable) can be one solution. This aspect is the subject of future research.

ACKNOWLEDGMENTS: The work has been funded by 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, D. Sideris, *An arrhythmia classification system based on the RR-interval signal*, Artificial Intelligence in Medicine, Vol. 33, 2005, pp. 237–250.
- [2] B. M. Asl, S. K. Setarehdan, M. Mohebbi, *Support vector machine-based arrhythmia classification using reduced features of heart rate variability signal*, Artificial Intelligence in Medicine, Vol. 44, 2008, pp. 51–54.
- [3] T.F. Yang, B. Devine, P.W. Macfarlane, *Artificial neural networks for the diagnosis of atrial fibrillation*, Med. Biol. Eng. Comput., Vol. 32, No. 6, 1994, pp. 615–619.
- [4] K. Minami, H. Nakajima, T. Toyoshima, *Real-time discrimination of ventricular tachyarrhythmia with Fourier transform neural network*, IEEE Trans Biomed Eng, Vol. 46, No. 2, 1999, pp. 179–185.
- [5] Y. Wang, Y.S. Zhu, N.V. Thakor, Y.H. Xu, *A short-time multifractal approach for arrhythmia detection based on fuzzy neural network*, IEEE Trans Biomed Eng, vol. 48, no. 9, 2001, pp. 989–995.
- [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, 2009, pp. 1701–1709.
- [7] L. Xu, D. Zhang, K. Wang, L. Wang, *Arrhythmic Pulses Detection Using Lempel-Ziv Complexity Analysis*, EURASIP Journal on Applied Signal Processing, Vol. 2006, pp. 1–12.
- [8] X.-S. Zhang, Y.-S. Zhu, N. V. Thakor, Z.-Z. Wang, *Detecting Ventricular Tachycardia and Fibrillation by Complexity Measure*, IEEE Transactions on Biomedical Engineering, Vol. 46, No. 5, 1999, pp. 548–555.
- [9] X.-S. Zhang, R. J. Roy, *EEG Complexity as Measure of Depth of Anesthesia for Patients*, IEEE Transactions on Biomedical Engineering, Vol. 48, No. 12, 2001, pp. 1424–1433.
- [10] L. Sarlabous, A. Torres, *Multistate Lempel-Ziv (MLZ) Index Interpretation as a Measure of Amplitude and Complexity Changes*, 31st Annual International Conference of the IEEE EMBS Engineering in Medicine and Biology Society, pp. 4375–4378, Minneapolis, Minnesota, USA, September 1999.
- [11] J. Pan, W.J. Tompkins, *A real-time QRS detection algorithm*, IEEE Trans Biomed Eng., Vol. 32, No. 2, 1985, pp. 230–236.
- [12] 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, 2009, pp. 1701–1709.
- [13] J. Mateo, A.M. Torres, A. Aparicio, J.L. Santos, *An efficient method for ECG beat classification and correction of ectopic beats*, Computers and Electrical Engineering, Vol. 53, 2016, pp. 219–229.
- [14] A. Martínez, R. Alcaraz, J.J. Rieta, *Ventricular activity morphological characterization: Ectopic beats removal in long term atrial fibrillation recordings*, Computer Methods and Programs in Biomedicine, Vol. 109, 2013, pp. 283–292.
- [15] J. Mateo, P. Laguna, *Analysis of heart rate variability in the presence of ectopic beats using the heart timing signal*, IEEE Trans. Biomed. Eng., Vol. 50, No. 3, 2003, pp. 334–343.
- [16] K. Solem, P. Laguna, L. Sornmo, *An efficient method for handling ectopic beats using the heart timing signal*, IEEE Trans. Biomed. Eng., Vol. 53, No. 1, 2006, pp. 13–20.
- [17] D. Nabil, F. B. Reguig, *Ectopic beats detection and correction methods: A review*, Biomedical Signal Processing and Control, Vol. 18, 2015, pp. 228–244.
- [18] R. Karlsson, R. Hörnsten, A. Rydberg, 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, 2012, pp. 1–12.
- [19] J.H. Zar, *Biostatistical Analysis*, Prentice-Hall, Upper Saddle River, New Jersey, 1985, pp. 587–588.
- [20] L. Sarlabous, A. Torres, J.A. Fiz, J. Gea, J.B. Galdiz, R. Jane, *Multistate Lempel-Ziv (MLZ) index interpretation as a measure of amplitude and complexity changes*, Conf Proc IEEE Eng Med Biol Soc., 2009, pp. 4375–4378.
- [21] J.S. Richman, J.R. Moorman, *Physiological time-series analysis using approximate entropy and sample entropy*, American Journal of Physiology, Heart and Circulatory Physiology, Vol. 278, No. 6, 2000, pp. H2039–49.
- [22] D. E. Lake, J.S. Richman, M.P. Griffin, J. R. Moorman, *Sample entropy analysis of neonatal heart rate variability*, Am. J. Physiol. Regul. Integr. Comp. Physiol., Vol. 283, No. 3, 2002, pp. R789–97.
- [23] R. Alcaraz, J.J. Rieta, *A review on sample entropy applications for the non-invasive analysis of atrial fibrillation electrocardiograms*, Biomedical Signal Processing and Control, Vol. 5, 2010, pp. 1–14.
- [24] A. Lempel, J. Ziv, *On the Complexity of Finite Sequences*, IEEE Transactions on Information Theory, Vol. IT-22, No. 1, 1976, pp. 75–81.
- [25] J. Ziv, *Coding Theorems for Individual Sequences*, IEEE Transactions on Information Theory, Vol. IT-24, No. 4, 1978, pp. 405–412.
- [26] L. Xu, D. Zhang, K. Wang, L. Wang, *Arrhythmic Pulses Detection Using Lempel-Ziv Complexity Analysis*, EURASIP Journal on Applied Signal Processing, Vol. 2006, 2006, pp. 1–12.
- [27] A.N. Kolmogorov, *Three approaches to the quantitative definition of information*, Problems of Information Transmission, Vol. 1, pp. 1–7, 1965.
- [28] M.R. Titchener, *Generalized T-Codes: An Extended Construction Algorithm for Self-synchronizing Codes*, TAMAKI T-CODE PROJECT SERIES, Vol. 1, No. 4, 1995, pp. 1–8.
- [29] U. Günter, *Data compression and serial communication with generalized T-codes*, Journal of Universal Computer Science, Vol. 2, No. 11, 1996, pp. 769–795.
- [30] K. Hamano, H. Yamamoto, *A Randomness Test based on T-Complexity*, IEEE Trans. Fundamentals, Vol. E93, No. 7, 2010, pp. 1346–1354.
- [31] M. Karlsson, R. Hörnsten, A. Rydberg, 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 2012, pp. 11–22.
- [32] D. Ge, Narayanan Srinivasan, S. M. Krishnan, *Cardiac arrhythmia classification using autoregressive modeling*, BioMedical Engineering OnLine, 2002, pp. 1–5.
- [33] S.W. Chen, *Two-stage discrimination of cardiac arrhythmias using a total least squares-based prony modeling algorithm*, IEEE Trans Biomed Eng, Vol. 47, 2000, pp. 1317–1326.
- [34] M. H. Song, J. Lee, S. P. Cho, K. J. Lee, S. K. Yoo, *Support Vector Machine Based Arrhythmia Classification Using Reduced Features*, International Journal of Control, Automation, and Systems, Vol. 3, No. 4, 2005, pp. 571–579.
- [35] T.F.L. de Medeiros, et. al., *Heart arrhythmia classification using the PPM algorithm*, Biosignals and Biorobotics Conference ISSNIP, pp. 1–5, 2011.