Middle-East Journal of Scientific Research 12 (10): 1403-1412, 2012
ISSN 1990-9233
© IDOSI Publications, 2012
DOI: 10.5829/idosi.mejsr.2012.12.10.118

Entropy-Based Algorithm to Detect Life Threatening Cardiac Arrhythmias Using Raw Electrocardiogram Signals

Chandrakar Kamath

Department of Electronics and Communication, Manipal Institute of Technology, Manipal-576104

Abstract: Ventricular tachycardia or fibrillation (VT-VF) as fatal cardiac arrhythmia is the major cause leading to sudden cardiac death. It is crucial for the patient to receive immediate medical intervention when either VT or VF occurs. The aim of this study is to investigate the possibility of predicting ventricular arrhythmia from electrocardiogram (ECG) signals. We use symbolic dynamics together with an entropy based complexity measure to discriminate between normal sinus rhythm (NSR) and life threatening arrhythmias like, ventricular tachycardia/fibrillation (VT/VF). The statistical analyses show that either, Renyi or Shannon Entropy measure is found to have potential in discriminating normal and VT/VF subjects and thus can significantly add to the prognostic value of traditional cardiac analysis. However, it is found that Renyi entropy outperforms Shannon entropy. The receiver operating characteristic curve (ROC) analysis confirms the robustness of this new approach and exhibits an average sensitivity of about 91.8% (94.9%), specificity of about 95.4% (97.5%), positive predictivity of around 96.1% (95.6%) and accuracy of about 93.4% (96.6%), with Renyi entropy to distinguish between normal and VT (VF) subjects. The presented method is simple, but with high detection quality, computationally fast and well suited for real time implementation in automated external or implantable cardioverter-defibrillators.

Key words: Complexity measures · Electrocardiogram · Renyi entropy · Shannon Entropy · Symbolic dynamics · Ventricular tachycardia · Ventricular fibrillation

INTRODUCTION

Ventricular tachycardia (VT) and ventricular fibrillation (VF) are life threatening cardiac arrhythmias [1]. Despite numerous recent advances in the field of medicine, Ventricular tachycardia/ fibrillation (VT/VF) has been difficult to manage with in clinical practice and mortality rate has remained high. It is crucial for the patient to receive immediate medical intervention when either VF or VT occurs. As a consequence the development of new noninvasive methods and measures of mortality risk in VT/VF, including sudden cardiac death, is still a major challenge. For this reason, a number of quantitative analysis techniques for ECG arrhythmia detection have been proposed [1-3]. Sequential hypothesis testing of binary sequences has been employed to detect ventricular fibrillation [4]. Though the method shows an improvement over previous methods, the accuracy is not high enough for clinical applications. Gustavo Santos used regularized least squares technique with a radial basis kernel to predict ventricular arrhythmia [5]. Power spectrum of raw ECG signals and power spectrum of beat-to-beat intervals were tried as input vectors to a neural network. Though the approach is said to lead to more accurate risk stratification, no details of accuracy and other measures are available.

Besides manual defibrillation by an emergency paramedic in the recent years, bystander defibrillation with automated external defibrillators (AEDs) has also been recommended for resuscitation. A reliable automated classification system combined with computationally fast real time implementable algorithm can resolve this issue. This work is an attempt to develop such an automated computationally fast system to discriminate between normal and Ventricular tachycardia/ fibrillation subjects.

Physiological data more often show complex structures which can not be quantified or interpreted using linear methods. The classical nonlinear methods suffer from the disadvantage of dimensionality. Further, there are not enough samples in the time series to arrive at a reasonable

Corresponding Author: Chandrakar Kamath, Department of Electronics and Communication, Manipal Institute of Technology, Manipal-576104, Permanent Address: H.No. 107, Ananthnagar, Manipal- 576104, Tel: +910820 4294139, Mob: +91-9880172507.

estimate of the nonlinear measures. From this point of view it is advisable to resort to methods which can quantify system dynamics even for short time series, like the symbolic dynamics. The prime advantages of symbolic dynamics are the following: If the fluctuations of the two data series are governed by different dynamics then the evolution of the symbolic sequences is not related. The resulting symbolic sequences histograms give a reconstruction of their respective histories and provide a visual representation of the dynamic patterns. In addition, they may be used as a basis to build statistics to compare the regions that show different dynamical properties and indicate which patterns are predominant. Thus methods of symbolic dynamics are useful approaches for classifying the underlying dynamics of a time series. Parameters of time domain and frequency domain often leave these dynamics out of consideration. Besides computational efficiency, symbolic methods are also robust when noise is present. The process of symbolization can be used to represent any possible variation over time, depending on the number of symbols and the sequence lengths used. This is a very powerful property because it does not make any assumptions about the nature of the signals/ patterns (e.g., it works equally well for both linear and nonlinear phenomena).

Symbolic time series analysis has found application for the past few decades in the field of complexity analysis, including astrophysics, geomagnetism, geophysics, classical mechanics, chemistry, medicine and biology, mechanical systems, fluid flow, plasma physics, robotics, communication and linguistics [6]. To be specific, in medicine, various implementations of symbolic sequences have been used to characterize encephalography (EEG) signals to understand the interaction between brain structures during seizures [7]. Under mechanical systems, symbolic methods were applied to combustion data from internal combustion engines to study the onset of combustion instabilities [8] and in multiphase flow data-symbolization were found to be useful in characterizing and monitoring fluidized-bed measurement signals [9]. Symbolic dynamics, as an approach to investigate complex systems, has found profound use in the analysis of heart rate variability signals [10-14]. There are many ways symbolic dynamics can be used for analysis of time series and all of them require coding i.e. converting the time series into symbolic series. The differences in symbolic methods are usually in their coding procedure or used complexity indices. In this contribution we employ static transformation with multiple partitions [6] as the symbolic method and Renyi entropy and Shannon Entropy [15-17] as measures of complexity of the symbolic sequences to classify ECG signals obtained from standard Holter recordings from PhysioNet database [18] into normal and VT/VF subjects. Static transformation with multiple partitions preserves more information and is preferred when the time series, like ECG, is nonstationary. The rationale behind the application of Renyi entropy and Shannon entropy complexity is that it is suitable for short-term segments of the ECG signal. Receiver operating characteristic (ROC) plots were used to evaluate the ability of the complexity measures to discriminate normal from VT/VF subjects. It is found that Renyi Entropy yielded excellent results better than those with Shannon entropy, with an average sensitivity of about 91.8% (94.9%), specificity of about 95.4% (97.5%), positive predictivity of around 96.1% (95.6%) and accuracy of about 93.4% (96.6%), with Renyi entropy to distinguish between normal and VT (VF) subjects. The presented method is simple, computationally fast, has high detection quality and hence, is well suited for real time implementation in automated external or implantable cardioverter-defibrillators.

METHODS AND MATERIALS

ECG Records: All the ECG records used are from the benchmark PhysioNet databases [18]. The work involved 18 ECG records from normal sinus rhythm (NSR) database (nsrdb) and ECG records of 35 subjects who experienced episodes of sustained ventricular tachycardia, ventricular flutter and ventricular fibrillation (VT/VF) from Creighton University ventricular tachyarrhythmia database (cudb). The NSR database includes 5 men, aged 26 to 45 years and 13 women, aged 20 to 50 years. The age and gender of subjects in VT/VF database are not available. For sake of comparison and validation, the NSR database was divided into two groups, first with 9 ECG records (Normal-1) and second, also, with 9 ECG records (Normal-2). Likewise, the VT/VF database was divided into two groups, first with 15 ECG records (VT/VF-1) and second also with 15 ECG records (VT/VF-2). From each record the modified limb lead II was only considered for analysis. The resolution is 200 samples per mV for nsrdb and 400 samples per mV for cudb. The sampling frequency of normal sinus rhythm signal from NSR is 128 Hz and that of VT/VF signal from cudb is 250 Hz. Since the sampling frequency does influence upon the calculated indices it is necessary to have the same sampling frequency for all the records. For this reason ECG signals from NSR database are first re-sampled at 250 Hz. Then each record is divided into segments of equal time duration (14 sec), with 3500

samples/ segment in both normal sinus rhythm and VT/VF database. A total of 1000 segments from normal sinus rhythm and from VT/VF data base, each, are analyzed. All the records are normalized before analysis. Also all the signals from both database are filtered using an 8-point moving average filter to remove high-frequency noise.

Symbolic Dynamics

Static Transformation and Dynamic Transformation: Symbolic dynamics, as an approach to investigate complex systems, facilitates the analysis of dynamic aspects of the signal of interest. The concept of symbolic dynamics is based on a coarse-graining of the dynamics [5]. That is the range of original observations or the range of some transform of the original observations such as the first difference between the consecutive values, is partitioned into a finite number of regions and each region is associated with a specific symbolic value so that each observation or the difference between successive values is uniquely mapped to a particular symbol depending on the region into which it falls. The former mapping is called static transformation and the latter dynamic transformation. Thus the original observations are transformed into a series of same length but the elements are only a few different symbols (letters from the same alphabet), the transformation being termed symbolization. A general rule of thumb is the partitions must be such that the individual occurrence of each symbol is equiprobable with all other symbols or the measurement range covered by each region is equal. This is done to bring out ready differences between random and nonrandom symbol sequences. The transformations into symbols have to be chosen context dependent. For this reason, we use complexity measures on the basis of such context-dependent transformations, which have a close connection to physiological phenomena and are relatively easy to interpret. This way the study of dynamics simplifies to the description of symbol sequences. Some detailed information is lost in the process but the coarse and robust properties of the dynamic behavior is preserved and can be analyzed [5].

Symbol-Sequences from Static Transformation of ECG Signals: In this study, we use static transformation approach for the symbolic dynamics [6]. For example, given the ECG signal time series $x_1, x_2, x_3, \ldots x_N$. In the static transformation, assuming uniform quantization, the full range of the series is spread over $\xi = 6$ symbols with a resolution of $\delta = (x_{max}-x_{min})/\xi$, where, x_{max} and x_{min} are respectively the maximum and minimum of the series, x. After quantization the series x becomes a new series $x_\xi = \{x_\xi(i), i=1,2, \ldots N\}$ of integer values ranging from 0 to ξ -1. Then this series is transformed into a new series, $x_{\xi,L} = \{x_{\xi,L}(i), i=1,2, \ldots N\}$, depending on a sequence of patterns of L delayed samples, where, $x_{\xi,L}(i) = \{ x_\xi(i), x_\xi(i-1), x_\xi(i-2), \ldots x_\xi(i-L+1)\}$. The number of possible $x_{\xi,L}(i)$ is ξ^L.

$$x_\xi(i) = \begin{cases} 0 & x_i <= x_{max} \text{ and } x_i >= x_{max} - \delta \\ 1 & x_i <= x_{max} - \delta \text{ and } x_i >= x_{max} - 2\delta \\ 2 & x_i <= x_{max} - 2\delta \text{ and } x_i >= x_{max} - 3\delta \\ 3 & x_i <= x_{max} - 3\delta \text{ and } x_i >= x_{max} - 4\delta \\ 4 & x_i <= x_{max} - 4\delta \text{ and } x_i >= x_{max} - 5\delta \\ 5 & x_i <= x_{max} - 6\delta \text{ and } x_i >= x_{max} - 6\delta \end{cases} \quad (1)$$

There are several quantities that properly characterize such symbol strings. In this work we primarily investigate the average frequency distribution (relative frequencies) of each of the patterns/ symbols from the alphabet {0, 1, 2, 3, 4, 5} for all the signal segments of each class, plot the symbol histogram and perform pattern classification.

Shannon Entropy and Renyi Entropy Complexity Measures: Claude Shannon defined entropy as a measure of uncertainty associated with a random variable. The Shannon entropy is a decreasing function of scattering of the random variable and is a maximal when all the outcomes are equally probable. It is a measure that can be used globally, on the whole data, or locally, to evaluate the probability density distributions around some points. This notion of entropy can be generalized to provide additional information about importance of specific events, for example outliers or rare events.

$$S = - \Sigma_i p_i \log_2 (p_i) \text{ for } 1 \le i \le n \quad (2)$$

Where p_i is the probability of occurrence of an event or feature value x_i being an element of the event (feature) X that can take values $\{x_1, x_2, \ldots, x_n\}$. A larger value of S implies higher complexity and a smaller value implies a lower complexity.

We use, in this work, modified Shannon entropy [17] defined as

$$H_s = - \Sigma_i p_i \log (p_i) / \log (N_{obs}) \text{ for } 1 \le i \le n \quad (3)$$

Where p_i is the normalized probability of the i^{th} symbol sequence and N_{obs} is the number of possible sequences which are actually observed in the data. Note that the normalization is with respect to Shannon entropy for a completely random process (one in which all sequences

are equiprobable). The advantage of this normalization is to bring down the bias on the statistics due to finite size of the data sets. This implies that the modified Shannon entropy will converge to 1 as the data approaches true randomness and for non-random data this value will be $0 \leq H_s \leq 1.0$ and a lower H_s implies more deterministic structure.

Further, Shannon entropy implicitly assumes certain tradeoff between contributions from main mass of the distribution and the tails. In some cases there is a need to distinguish weak signals overlapping with the much stronger one. This problem can be remedied by using entropies that depend on the powers of the probability, as in $\Sigma_i\, p(x_i)^{\alpha}$ for $1 = i = n$. Renyi entropy belongs to this category and is defined as

$$I_{\alpha} = \log (\Sigma_i (p_i)^{\alpha}) / (1-\alpha) \text{ for } 1 = i = n \qquad (4)$$

α is a continuous parameter and $\alpha \Sigma 1$. For $\alpha = 1$, this entropy reduces to Shannon entropy. If α is a large positive value, the measure becomes more sensitive to events that occur often, while if α is a large negative value, the measure becomes more sensitive to events that occur seldom. In this work we have chosen $\alpha = 0.25$.

Renyi entropy has a higher dynamic range than Shannon entropy over a range of scattering conditions. Nevertheless Shannon entropy is used in a variety of applications.

***t*-Tests and Receiver Operating Characteristic (ROC) Analysis:** Individual and pair-wise significance tests (Student's *t*-tests) are used to evaluate the statistical differences between the Shannon Entropy values of normal and VT/VF groups. If significant differences between groups are found, then the ability of the non-linear analysis method to discriminate normal from VT/VF subjects is evaluated using receiver operating characteristic (ROC) plots in terms of C-statistics. ROC curves are obtained by plotting sensitivity values (which represent the proportion of the patients with diagnosis of VT/VF who test positive) along the y axis against the corresponding (1-specificity) values (which represent the proportion of the correctly identified normal subjects) for all the available cutoff points along the x axis. Accuracy is a related parameter that quantifies the total number of subjects (both normal and VT/VF) precisely classified. The area under ROC curve (AUC), also called C-statistic, measures this discrimination, that is, the ability of the test to correctly classify those with and without the disease and is regarded as an index of diagnostic accuracy. The optimum threshold is the cut-off point in which the highest accuracy (minimal false negative and false positive results) is obtained. This can be determined from the ROC curve as the closet value to the left top point (corresponding to 100% sensitivity and 100% specificity). A C-statistic value of 0.5 indicates that the test results are better than those obtained by chance, where as a value of 1.0 indicates a perfectly sensitive and specific test.

RESULTS AND DISCUSSION

To test for significance of Renyi entropy measure and efficacy of the multiple partitioning of ECG using static transformation, first we compare the entropy measures of the ECG data from normal and VT/VF subjects of Group-I and show that Renyi entropy outperforms Shannon entropy. Next, we validate our approach conducting another case study on normal and VT/VF subjects from Group-II. Both Renyi and Shannon Entropy are analyzed from segments of 3500 samples and averaged to obtain mean values for the entire recording period.

The sampling frequency of NSR database is 128 Hz and that of VT/VF database is 250 Hz. Since the sampling frequency does influence upon the calculated indices it is necessary to have the same sampling frequency for all the records. For this reason ECG signals from NSR database are first re-sampled at 250 Hz. For sake of comparison and validation, as mentioned earlier, the normal sinus rhythm database (NSR) was divided into two groups, first with 9 ECG records (Normal-1) and second with 9 ECG records (Normal-2). Likewise, the VT/VF database was divided into two groups, first with 10 ECG records (VT/VF-1) and second also with 10 ECG records (VT/VF-2). Then each record is divided into segments of equal time duration (14 sec), with 3500 samples/ segment in both normal sinus rhythm and VT/VF database. A total of 1000 segments from normal sinus rhythm and from VT/VF data base, each, are analyzed. Renyi and Shannon Entropy analysis is applied to these segments from both the groups to decide whether a particular segment belongs to normal, or VT/VF group. Static transformation as given in Eq. (1) is first applied on each segment to arrive at a symbol string with a range of six possible symbols {0, 1, 2, 3, 4, 5} and both Renyi and Shannon entropies are computed. This is repeated for all the segments of all the three classes.

Figure 1 shows the characteristics of three types of ECG signals, NSR episode, VT episode and VF episode. All the signals are plotted with respect to same time scale (in samples). It can be observed that the widths of the QRS complexes are different in the three signals. For NSR the QRS width is usually in the range 0.06-0.1 sec and for VT the QRS complex is much wider (> 0.1 sec).

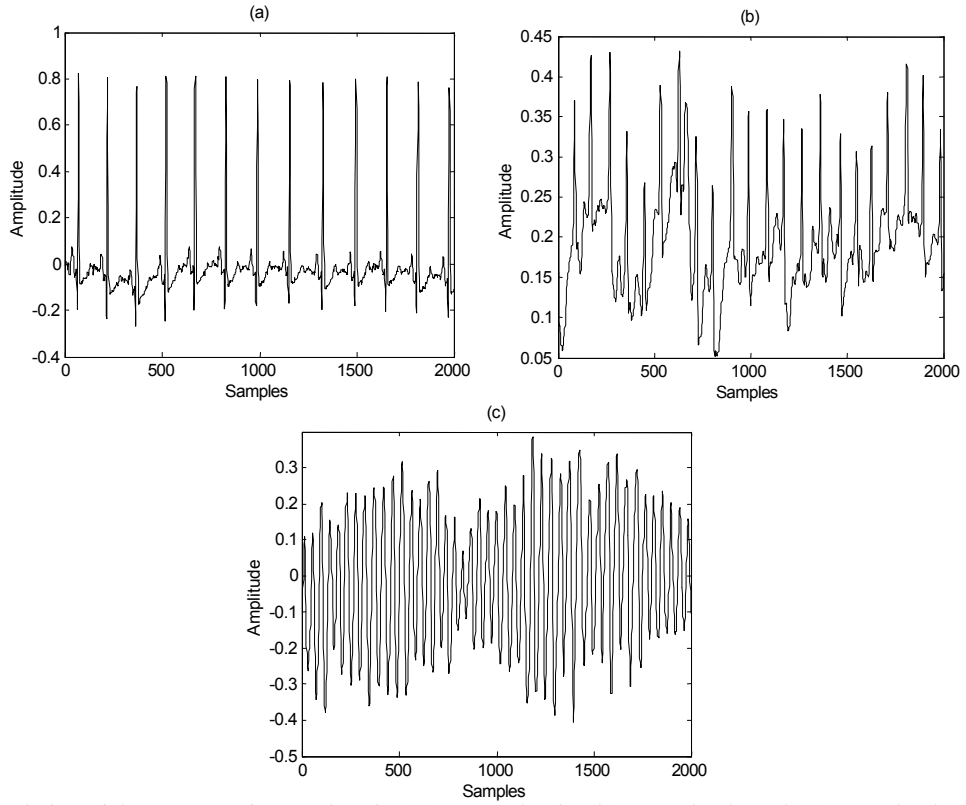

Fig. 1: Characteristics of three types of ECG signals. (a) NSR episode, (b) VT episode and (c) VF episode.

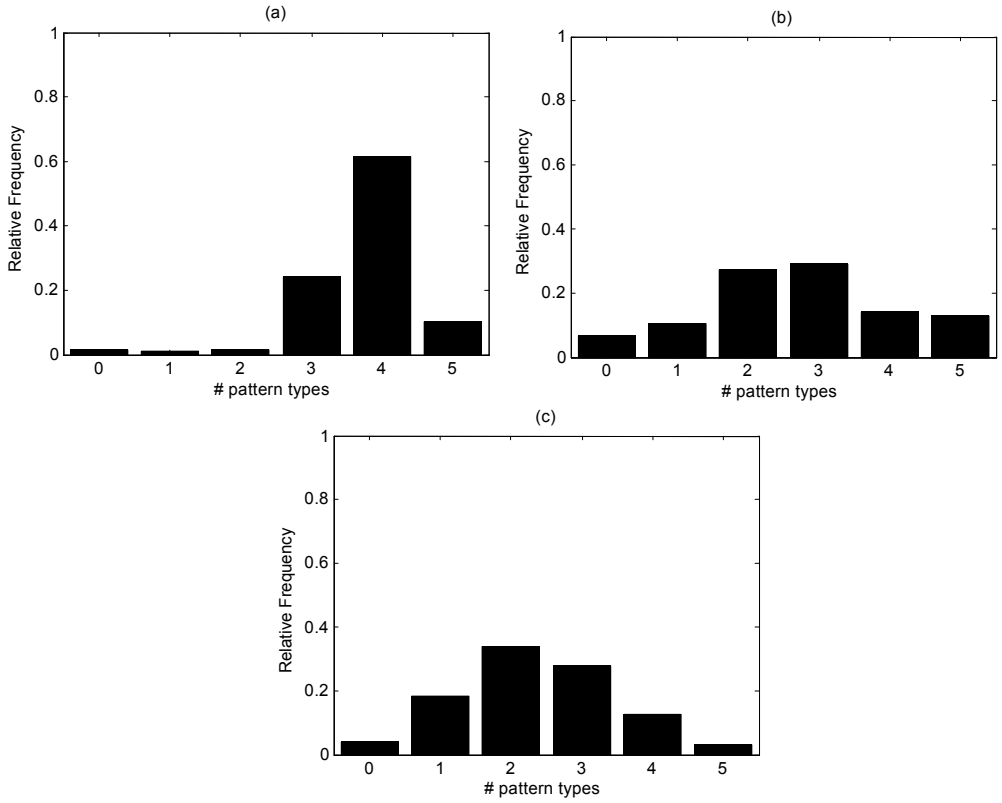

Fig. 2: Symbol/ pattern type histogram of (a) NSR episode, (b) VT episode and (c) VF episode.

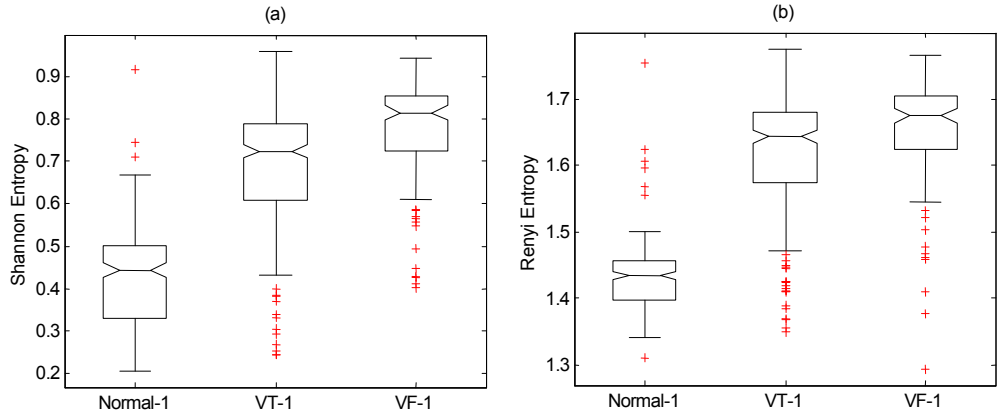

Fig. 3: (a) The distribution of Shannon entropy values (with static transformation) using Box-whiskers plots (with outliers) for normal, VT and VF subjects from Group-I. (b) The distribution of Renyi entropy values (with static transformation) using Box-whiskers plots (with outliers) for normal, VT and VF subjects from Group-I.

In VF, no QRS complexes are seen. Further, in the case of NSR P waves are normal, while in the case of VT/ VF no P waves are seen. Figure 2 shows averaged symbol histograms for NSR, VT and VF segments of the three classes. The relative frequency distribution of patterns for the three cases is found to be distinctly different. Comparison among the different distributions from Figure 2 reveals that for NSR subjects the descending order of distribution of symbols is 4, 3, 5, 0, 2 and 1. For VT subjects the descending order of distribution of symbols is 3, 2, 4, 5, 1 and 0 and for VF case it is 2, 3, 1, 4, 0 and 5. The distribution of Shannon Entropy values, with static transformation of ECGs, for the NSR, VT and VF groups (Group-I) are shown using Box-whiskers plots in Figure 3(a). The boxes (inter-quartile range) of normal and VT/VF subjects are non-overlapping. But, the box of VT and the upper whisker of normal classes overlap. Also, the upper whisker of normal and the lower whisker of VF classes overlap. This plot shows that Shannon Entropy can be used to distinguish between normal and VT/VF subjects. The results of statistical analysis of non-paired Student's *t*-test for normal, VT and VF groups of Group-I are depicted in Table 1. All values are expressed as mean±Standard Deviation (median) [95% Confidence Interval]. For normal subjects, we find the following Shannon Entropy (mean±S.D.): 0.4218±0.1055. For VT subjects we find the following Shannon Entropy (mean±S.D.): 0.6910±0.1406, different from normal. For VF subjects we find the following Shannon Entropy (mean±S.D.): 0.7748±0.1163, different from normal. These distributions show that Shannon Entropy is sufficient to distinguish between normal and VT/VF subjects. It is found that Shannon Entropy for VT/VF group is always larger than that of the normal group. This implies a decrease in the regularity of VT/VF group compared to normal group. Of course, experimental studies are necessary to confirm the mechanisms behind the decrease in the regularity of signals in VT/VF subjects. The distribution of Renyi Entropy values, with static transformation of ECGs, for the NSR, VT and VF groups (Group-I) are shown using Box-whiskers plots in Figure 3(b). The boxes (inter-quartile range) of normal and VT/VF subjects are non-overlapping. Only the upper whisker of normal and the lower whisker of VT classes overlap. This plot shows that Renyi Entropy can also be used to distinguish between normal and VT/VF subjects. The results of statistical analysis of non-paired Student's *t*-test for normal, VT and VF groups of Group-I are depicted in Table 3. All values are expressed as mean ± Standard Deviation (median) [95% Confidence Interval]. For normal subjects, we find the following Renyi Entropy (mean±S.D.): 1.4295±0.0495. For VT subjects we find the following Renyi Entropy (mean±S.D.): 1.6222±0.0878, different from normal. For VF subjects we find the following Renyi Entropy (mean±S.D.): 1.6542±0.0788, different from normal. These distributions show that Renyi Entropy is also sufficient to distinguish between normal and VT/VF subjects. It is found that Renyi Entropy for VT/VF group is always larger than that of the normal group. This implies a decrease in the regularity of VT/VF group compared to normal group. Of course, experimental studies are necessary to confirm the mechanisms behind the decrease in the regularity of signals in VT/VF subjects.

Although both, Shannon and Renyi entropy measures, perform well in separating normal from VT/VF groups, comparing paired *t*-test results (*p*-value and tstat) from Tables 2 and 4, it is found that Renyi entropy

Table 1: Descriptive results of Shannon entropy analysis for Group-I. All values are expressed as mean±SD (median) [95% CI]. (non-paired Student's *t*-test; $p < 0.0001$)

Subject	Shannon entropy
NSR	0.4218±0.1055 (0.4425) [0.4084 0.4352]
VT	0.6910±0.1406 (0.7235) [0.6748 0.7072]
VF	0.7748±0.1163 (0.8145) [0.7551 0.7944]

Table 2: *p*-values and tstat (test statistic) values of paired *t*-test for Shannon entropy analysis of normal and VT/VF subjects from Group-I.

Subject	VT	VF
NSR	$p= 2.0200\text{x}10^{-89}$; tstat= -24.5327	$p= 5.0048\text{x}10^{-102}$; tstat= -30.0882

Table 3: Descriptive results of Renyi entropy analysis for Group-I. All values are expressed as mean±SD (median) [95% CI]. (non-paired Student's *t*-test; $p < 0.0001$)

Subject	Renyi entropy
NSR	1.4295±0.0495 (1.434) [1.4232 1.4358]
VT	1.6222±0.0878 (1.643) [1.6121 1.6322]
VF	1.6542±0.0788 (1.675) [1.6409 1.6676]

Table 4: *p*-values and tstat (test statistic) values of paired *t*-test for Renyi entropy analysis of normal and VT/VF subjects from Group-I

Subject	VT	VF
NSR	$p= 8.6245\text{x}10^{-118}$; tstat= -30.2884	$p= 1.4715\text{ x}10^{-116}$; tstat= -33.9877

outperforms Shannon entropy. This finding is substantiated using ROC plots, which are shown in Figure 4(a), for Shannon entropy and in Figure 4(b), for Renyi entropy, respectively, with normal and VT (shown by solid line) and normal and VF (shown by dash-dot line). It is found, comparing both the figures that, Renyi entropy performs better than Shannon entropy. With Shannon entropy, in Figure 4(a), it is found that (i) for normal and VT separation, the area under the curve (AUC) is 0.92553 with sensitivity = 87.7%, specificity = 93.8%, positive predictivity = 94.5% and accuracy = 90.4% and (ii) for normal and VF separation, the area under the curve (AUC) is 0.97038 with sensitivity = 95.6%, specificity = 97.1%, positive predictivity = 97.5% and accuracy = 96.6%. With Renyi entropy, in Figure 4(b), it is found that (i) for normal and VT separation, the area under the curve (AUC) is 0.94962 with sensitivity = 91.8%, specificity = 95.4%, positive predictivity = 96.1% and accuracy = 93.4% and (ii) for normal and VF separation, the area under the curve (AUC) is 0.97074 with sensitivity = 94.9%, specificity = 97.5%, positive predictivity = 95.6% and accuracy = 96.6%. Comparing these measures, it is obvious that using Renyi entropy for complexity measure has an advantage over the usual Shannon entropy.

Finally, we validate our approach conducting another case study on normal and VT/VF subjects from Group-II. The results of statistical analysis of non-paired Student's *t*-test for normal and VT/VF groups of Group-II are depicted in Table. 5. All values are expressed as mean±Standard Deviation (median) [95% Confidence Interval]. For normal subjects, we find the following

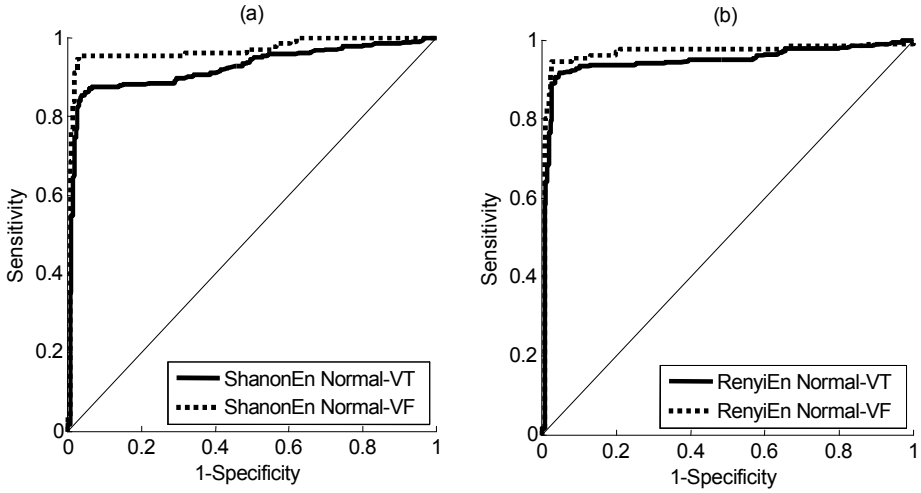

Fig. 4: (a) ROC curves with Shannon entropy, for normal and VT (solid line) and normal and VF (dash-dot line). (b) ROC curves with Renyi entropy, for normal and VT (solid line) and normal and VF (dash-dot line). The diagonal line (dotted line) from 0,0 to 1,1 represents ROC curve that can not discriminate between normal and VT/VF from Group-I.

Table 5: Descriptive results of Renyi entropy analysis for Group-II. All values are expressed as mean±SD (median) [95% CI]. (non-paired Student's *t*-test; $p < 0.0001$)

Subject	Renyi entropy
NSR	1.4678±0.0736 (1.467) [1.4595 1.4762]
VT	1.5896±0.0838 (1.589) [1.5811 1.5982]
VF	1.6737±0.0767 (1.688) [1.6621 1.6853]

Table 6: *p*-values and tstat (test statistic) values of paired *t*-test for Renyi entropy analysis of normal and VT/VF subjects from Group-II

Subject	VT	VF
NSR	p= 0; tstat= -19.7392	p= 1.7047x10^{-105}; tstat= -28.7390

Renyi Entropy (mean±S.D.): 1.4678±0.0736. For VT subjects we find the following Renyi Entropy (mean±S.D.): 1.5896±0.0838, different from normal. For VF subjects we find the following Renyi Entropy (mean±S.D.): 1.6737±0.0767, different from normal. The paired *t*-test results (*p*-value and tstat) are shown in Table 6. The ROC plots for Renyi entropies are shown in Figure 5. It is found that (i) for normal and VT separation, the area under the curve (AUC) is 0.87892 with sensitivity = 78.3%, specificity = 81.3%, positive predictivity = 83.8% and accuracy = 79.7% and (ii) for normal and VF separation, the area under the curve (AUC) is 0.94637 with sensitivity = 93.6%, specificity = 94.7%, positive predictivity = 96.3% and accuracy = 94.3%. The above results again substantiate our finding that Renyi entropy with multi-partition static transformation of ECGs outperforms Shannon Entropy with the same static transformation and that former is preferred to distinguish between normal and VT/VF subjects. The difference in accuracy and other measures of Group-II can be attributed to age differences and differing male-to-female ratios between groups I and II.

Figure 6 shows a synthesized ECG signal comprising NSR, VT and VF rhythms in sequence, each 3500 samples long, together with the corresponding Renyi entropy variation. Two empirically found thresholds (at 1.5 and 1.3 marked by horizontal solid lines) are used for separating NSR and VT/ VF rhythms.

The presented method is simple, computationally efficient and well suited for real time implementation in automated external or implantable defibrillators.

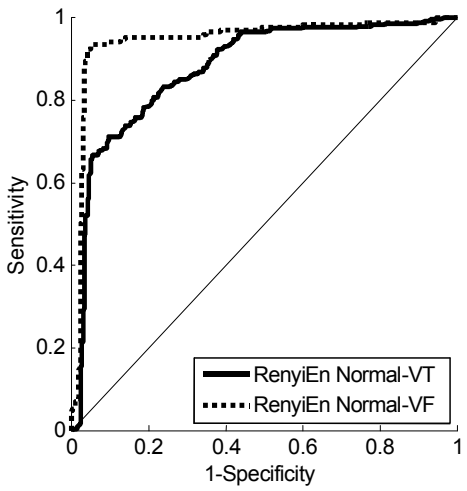

Fig. 5: ROC curves with Renyi entropy, for normal and VT (solid line) and normal and VF (dash-dot line). The diagonal line (dotted line) from 0,0 to 1,1 represents ROC curve that can not discriminate between normal and VT/VF from Group-II.

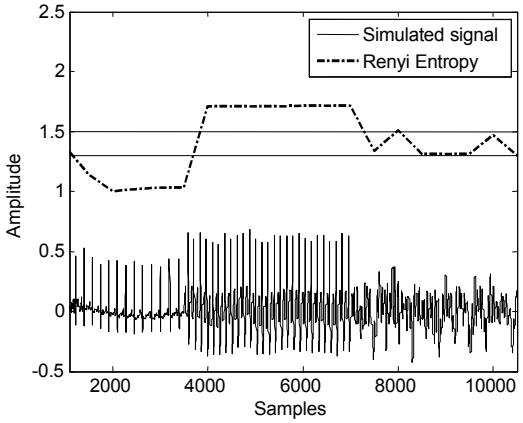

Fig. 6: Variation of Renyi entropy (dotted line) for a simulated signal (solid line) comprising NSR, VT and VF rhythms in sequence. The two horizontal solid lines mark the thresholds.

One limitation of the current study is the small sample size. Although we have reported symbolic dynamics to yield very sensitive measures based on the *p*-value generated from the *t*-statistics, factors like high variance, age differences and differing male-to-female ratios between groups will have an impact on the results when statistical analyses are carried out on small sample sizes. Nevertheless, the results of this study provide sufficient evidence to warrant the execution of larger studies that can provide more statistically robust confirmation of the application of symbolic dynamics as a reliable measure of cardiac health.

CONCLUSION

We apply entropy analysis on the static transformed nonstationary raw ECG time series from normal and VT/VF subjects. The quantified Renyi complexity measure of the resulting symbolic sequences are found to have potential in discriminating normal and VT/VF subjects and thus can significantly add to the prognostic value of traditional cardiac analysis. This complexity measure can easily be analyzed from the automated external or implantable defibrillator ECG recordings without time consuming preprocessing and hence, are appropriate for practical applications. Inappropriate defibrillator discharge or anti-tachycardiac pacing remains an important clinical problem in implantable cardioverter-defibrillator therapy as they lead to unnecessary pain and sometimes proarrhythmic effects. As an implication in real time applications the value for specificity is more important than the value for sensitivity and with this approach average specificity (about 97.0%) is a little more than sensitivity (about 94.0%). The usual nonlinear methods applied to time series (other than symbolic dynamics) usually demand long-term series, at times of 24 hours length. Using long-term record for this kind of analysis is not amenable. Although the ECG data we use contains both 30 minutes and 20 hours duration records, our method tests short-term segments, of the order of 14 sec duration. Hence the method is well suited for real time implementation in automated external or implantable cardioverter-defibrillators.

REFERENCES

1. Szi-Wen Chen, 2007. Complexity-Measure-Based Sequential Hypothesis Testing for Real-Time Detection of Lethal Cardiac Arrhythmias. EURASIP Journal on Advances in Signal Processing, 2007(20957): 1-8.
2. Chen, S.W., 2000. A two-stage discrimination of cardiac arrhythmias using a total least squares-based Prony modeling algorithm. IEEE Transactions on Biomedical Engineering, 47(10): 1317-1327.
3. Xu, L., D. Zhang, K. Wang and L. Wang, 2006. Arrhythmic pulses detection using Lempel-Ziv complexity analysis. EURASIP Journal on Applied Signal Processing, Article ID 18268: 1-12.
4. Pardey, J., 2007. Detection of Ventricular Fibrillation by Sequential Hypothesis Testing of Binary Sequences. Computers in Cardiology, 34: 573-576.
5. Gustavo Santos, 2006. Towards Ventricular Arrhythmia Prediction from ECG Signals. MIT Computer Science and Artificial Intelligence Laboratory, September 2006.
6. Daw, C.S., C.E.A. Finney, and E.R. Tracy, 2003. A review of symbolic analysis of experimental data. Review of Scientific Instruments, 74(2): 915-930.
7. Xu, J.H., Z.R. Liu and R. Liu, 1994. The measures of sequence complexity for EEG studies. Chaos, 4(11): 2111-2119.
8. Daw, C.S., 1998. Observing and modeling nonlinear dynamics in an internal combustion engine. Phys. Rev. Lett., 57(3): 2811-2819.
9. Finney, C.E.A., K. Nguyen, C.S. Daw and J.S. Halow, 1998. Symbol-sequence statistics for monitoring fluidization. Proceedings of the ASME Heat Transfer Division, pp: 405-411.
10. Kurths, J., A. Voss, P. Saparin, A. Witt, H.J. Kleiner and N. Wessel, 1995. Quantitative analysis of heart rate variability. Chaos, 5: 88-94.
11. Porta, A., G. D'Addio, G.D. Pinna, R. Maestri, T. Gnecchi-Ruscone, R. Furlan, N. Montano, S. Guzzetti and A. Malliani, 2005. Symbolic analysis of 24h Holter heart period variability series: Comparison between normal and heart failure patients. Computers in Cardiology, 32: 575-578.
12. Eleonora Tobaldini, Alberto Porta, Shun-Guang Wei, Zhi-Hua Zhang, Joseph Francis, Karina Rabello Casali, M. Robert Weiss, B. Robert Felder and Nicola Montano, 2009. Symbolic analysis detects alterations of cardiac autonomic modulation in Ventriclar tachycardia/ Ventricular fibrillation rats. Auton Neurosci., 150(1-2): 21-26.
13. Wessel, N., U. Schwarz, P.I. Saparin and J. Kurths. Symbolic dynamics for Medical data Analysis, pp: 46-62.
14. Voss, A., J. Kurths, H.J. Kleiner, A. Witt, N. Wessel, P. Saparin, K.J. Osterziel, R. Schurath and R. Dietz, 1996. The application of methods of nonlinear dynamics for the improved and predictive recognition of patients threatened by sudden cardiac death. Cardiovasc. Res., 31: 419-433.
15. Carlos Gómez and Roberto Hornero, 2010. Entropy and Complexity Analyses in Alzheimer's Disease: An MEG Study, The Open Biomedical Engineering Journal, 4: 223-235.

16. Yingjie Li, Shanbao Tong, Dan Liua, Yi Gai, Xiuyuan Wang, Jijun Wang, Yihong Qiu and Yisheng Zhu, 2008. Abnormal EEG complexity in patients with schizophrenia and depression. Clinical Neurophysiology, 119: 1232-1241.
17. Finney, C.E.A., K. Nguyen, C.S. Daw and J.S. Halow, 1998. Symbol-sequence statistics for monitoring fluidization. Proceedings of the ASME Heat Transfer Division, pp: 405-411.
18. Ary Goldberger, L., A.N. Luis Amaral, Leon Glass and M. Jeffrey, 2000. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng and H. Eugene Stanley. PhysioBank, PhysioToolkit and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals. Circulation, 101: e215-e220. (Available at http:// circ.ahajournals.org/ cgi/ content/ full/101/23/e215).