Accepted Manuscript

Title: Automatic Seizure Detection Based on Nonlinear Dynamical Analysis of EEG Signals and Mutual Information

Authors: Behnaz Akbarian [1], Abbas Erfanian [1]

[1] Iran Neural Technology Research Centre, Department of Biomedical Engineering, Iran University of Science and Technology (IUST)

Number of text pages: 36

Number of figures: 5

Number of tables: 3

Running title: Seizure Detection by Nonlinear Analysis

Corresponding author at: Department of Biomedical Engineering, Iran University of Science and Technology, Tehran, Iran, Tel.: +98-21-77240465; fax: +98-21-77240253, *E-mail addresses*: erfanian@iust.ac.ir (A. Erfanian)

To appear in: Basic and Clinical Neuroscience

Received date: 2017/03/3
Revised date: 2018/05/28
Accepted date: 2017/10/4

Please cite this article as:
Akbarian, Z. Erfanian, A. (In Press). Automatic Seizure Detection Based on Nonlinear Dynamical Analysis of EEG Signals and Mutual Information. *Basic and Clinical Neuroscience*. Just Accepted publication Oct. 4, 2017

ABSTRACT

In this paper, nonlinear dynamical analysis based on recurrence quantification analysis (RQA) has been employed for characterizing the nonlinear EEG dynamics. RQA can provide some useful quantitative information on the regular, chaotic or stochastic property of the underlying dynamics. We use the RQA-based measures as the quantitative features of the nonlinear EEG dynamics. Mutual information (MI) is used to find the most relevant feature subset from RQA-based features. The selected features are fed into an artificial neural network for grouping of EEG recordings to detect ictal, interictal, and healthy states. The performance of the proposed procedure is evaluated using a database for different classification cases. The combination of five selected features based on MI achieved 100% accuracy, which demonstrates the superiority of the proposed method.

Key Words: Epilepsy, Mutual Information, Nonlinear Analysis, Recurrence Quantification Analysis, Seizure Detection.

1. Introduction

Epilepsy is a chronic neurological disorder that can cause recurrent seizures and characterized by sudden, excessive, disorder, hypersynchronous, and localized electrical discharges of a group of neurons in the brain that temporarily can change brain functions, i.e., transient impairments of sensation, altered state of consciousness or loss of awareness, focal involuntary movements or convulsions (Osorio, Zaveri, Frei, & Arthurs, 2011). Sudden and recurrent seizures can have significant effect on the life of a patient. It is clear that reliable real-time detecting the occurrence of seizures could significantly improve the therapeutic potentials such as "closed-loop" therapies. In closed-loop therapies, electrical stimulation, drug infusion, cooling, or biofeedback may be delivered in response to a seizure detection (Ramgopal et al., 2014). Patients with epilepsy are usually treated with antiepileptic drugs (AEDs) to control their seizures (Wlodarczyk, Palacios, George, & Finnell, 2012); accurate real-time detection of seizures is critical to reduce the side effects by on demand delivering of AEDs during the preictal case with short-acting drugs.

A conventional technique for diagnosis and analysis of epilepsy is the long-term EEG recording for several days and then visual inspection of EEG recordings by human specialists. To reduce the burden of time-consuming inspection, a robust real-time seizure-detection system could facilitate long-term monitoring and localization of the epileptogenic zone (i.e., the brain zone that can generate seizures), which is helpful in presurgical evaluations. Accordingly, there is currently a strong demand for developing automatic seizure detection systems. A seizure detection system should be able to identify the occurrence of seizures from the ongoing EEG or intracranial EEG and this can be achieved by classification of the brain signals. Different approaches have been proposed to deal with the automatic seizure detection. The key components of seizure detection are feature extraction from brain electrical activity and classification of the extracted features. So far, different approaches based on time-domain analysis, frequency-domain analysis, and information theory have been used for feature extraction from brain electrical activity (Thomas et al., 2013; Yang et al., 2013; Page et al., 2015; Peker, Sen, & Delen, 2016).

Empirical mode decomposition (EMD) have been used for extracting features from the intrinsic mode functions (IMFs) of EEG signals for seizure detection (Pachori, 2008; Oweis & Abdulhay, 2011; Pachori & Bajaj, 2011; Bajaj & Pachori, 2012; Alam & Bhuiyan, 2013; Riaz et al., 2016). The mean frequency measure of IMFs has been used as a feature to recognize the difference between seizure (ictal) and seizure-free (interictal) EEG signals (Pachori, 2008). In (Oweis & Abdulhay, 2011), the weighted frequency of IMFs has been used as the feature set for discriminating healthy EEG from epileptic seizure EEG signals. The area measurement of the analytic IMFs has been also used as a feature set for discriminating healthy from the epileptic seizure (Pachori & Bajaj, 2011). Bajaj & Pachori (2012) used the amplitude and frequency modulation bandwidths of the analytic IMFs as the feature set for classifying seizure and nonseizure EEG signals. The higher order moments, including variance, kurtosis, and skewness, extracted from the IMFs of the EEG signals were used as the features for classification of various cases including healthy, interictal, and ictal; healthy and seizure; nonseizure and seizure; and interictal and ictal (Alam & Bhuiyan, 2013). Recently, spectral centroid, coefficient of variation, and the spectral skew of the IMFs have been used for feature extraction to detect epileptic seizure (Riaz et al., 2016).

A series of studies focused on nonlinear dynamical analysis of EEG signals to extract features for detection of epilepsy (Srinivasan, Eswaran, & Sriraam, 2007; Chen et al., 2011; Niknazar et al., 2013; Yaylali, Koçak, & Jayakar, 1996; Cerf, El Amri, El Ouasdad, & Hirsch, 1999; Adeli, Dastidar, & Dadmehr, 2007; Dastidar, Adeli, & Dadmehr, 2007; Iasemidis et al., 2003; Drongelen et al., 2003; Easwaramoorthy & Uthayakumar, 2011; Zhou, Liu, Yuan, & Li, 2013; Zabihi et al., 2016; Thomasson, Hoeppner, Webber, & Zbilut, 2001; Li, Ouyang, Yao, & Guan, 2004; Ouyang, Li, Dang, & Richards, 2008; Niknazar et al., 2013). These features include approximate entropy (ApEn) (Srinivasan, Eswaran, & Sriraam, 2007; Chen et al., 2011; Niknazar et al., 2013), correlation dimension (Yaylali, Koçak, & Jayakar, 1996; Cerf, El Amri, El Ouasdad, & Hirsch, 1999; Adeli, Dastidar, & Dadmehr, 2007; Dastidar, Adeli, & Dadmehr, 2007), Lyapunov exponent (Niknazar et al., 2013; Adeli, Dastidar, & Dadmehr, 2007; Dastidar, Adeli, & Dadmehr, 2007; Iasemidis et al., 2003), Kolmogorov entropy (Drongelen et al., 2003), fractal dimension (Niknazar et al., 2013; Easwaramoorthy & Uthayakumar, 2011), lacunarity (Zhou, Liu, Yuan, & Li, 2013), and features extracted from Poincaré section (Zabihi et al., 2016) as well as recurrence quantification analysis (RQA) (Niknazar et al., 2013; Thomasson, Hoeppner, Webber, & Zbilut, 2001; Li, Ouyang, Yao, & Guan, 2004; Ouyang, Li, Dang, & Richards, 2008; Niknazar et al., 2013).

In spite of all these numerous approaches for feature extraction, a major challenge to classify the electrical brain activity to detect epilepsy is the feature selection from a large number of available EEG features. Searching important and relevant features is essential to improve the accuracy, efficiency, and generalization of a classification process. There have been a few studies on the feature selection for seizure detection (Alessandro et al., 2003; Temko et al., 2011; Wang & Lyu, 2015; Zhang & Parhi, 2016). Genetic algorithm (Alessandro et al., 2003), recursive feature elimination (Temko et al., 2011; Wang & Lyu, 2015), and Fisher's linear discriminant analysis combined with the branch and bound algorithm (Zhang & Parhi, 2016) were employed to select EEG features for epileptic seizure detection.

In this paper, a feature selection algorithm, which is based on mutual information (MI) estimates (Kwak & Choi, 2002; Peng, Long, & Ding, 2005) is used for seizure detection. MI is a nonparametric measure of the dependence between random variables and is always non-negative. In terms of MI, the aim of the feature selection is to find features from a large feature set which jointly has the largest dependency on the target class. The original features were extracted from RQA of the EEG signals. The RQA of the EEG signals is used to characterize the nonlinear EEG dynamics and extract appropriate features for automatic seizure detection. We extend the previous RQA-based features and introduce different RQA measures, which are important for measuring the complexity.

2. Dataset

The dataset provided by Dr. R. Andrzejak (http://epileptologie-bonn.de/cms/front_content.php?idcat=193&lang=3&changelang=3) is used in this study. The dataset consists of 500 single-channel EEG segments, each one having a duration of 23.6 s, and is categorized into five subsets (marked as sets A-E) while each subset contains 100 EEG segments. The subsets A and B have been recorded from five healthy candidates with their eyes open and closed, respectively, using the standard 10–20 electrode arrangement. The subsets C and D contain EEG signals recorded during interictal intervals from the epileptogenic region and the hippocampal formation of the opposite hemisphere, respectively. The subset E includes EEG segments corresponding to seizure attacks, recorded using all the electrodes. The subsets A and B have been recorded extracranially, whereas subsets C, D, and E have been recorded intracranially. The EEG signals were recorded in a digital format at the sampling rate of 173.61 Hz and were band-pass filtered between 0.53 and 60 Hz. Fig. 1 demonstrates the typical EEG signals from each subset.

Fig. 1

3. Methods

3.1. Feature Extraction Procedure

The original features were extracted from the RQA of the EEG signals. The first step in RQA is the reconstruction of the phase space trajectory and construction of the recurrence plot (RP). RP is a technique, which can visualize the recurrence of the system states of a dynamical system in the phase space.

3.1.1. Phase Space Reconstruction

An important step in the analysis of any dynamical system is the reconstruction of its phase space. The phase space of a dynamical system is a space, which shows all states of a system, whereas each state of the system corresponds to one unique point in the phase space. Phase space is a geometrical representation of system dynamics. A frequently used method for the phase space reconstruction is the Taken's time delay method (Marwan, Romano, Thiel, & Kurths, 2007). According to the Taken's theorem, the dynamics of time series $(u_1, u_2, \dots, u_N)$ can be embedded in a *m*-dimensional phase space by the vector as follows

$$x_t = \left(u_t, u_{t+\tau}, \dots, u_{t+(m-1)\tau}\right) \tag{1}$$

where τ and *m* are the time delay and the embedding dimension, respectively. In order to fully capture the dynamics, an appropriate time delay and the embedding dimension should be chosen. A proper time delay is the first local minimum of the MI function (Fraser & Swinney, 1986). Cao proposed a method to define the minimum embedding dimension from a scalar time series (Cao, 1997). The method is started with a low value of the embedding dimension *m* and then increasing it until the number of false neighbors' reduces to zero. In this paper, we used MI and Cao' methods to approximate the time delay and embedding dimension, respectively.

3.1.2. Recurrence Plot (RP)

Recurrence is a substantial nature of dynamical systems (Marwan, Romano, Thiel, & Kurths, 2007). Eckmann et al. introduced a method to visualize the recurrences of dynamical systems called RP (Marwan, Romano, Thiel, & Kurths, 2007). To construct the RP, a symmetrical $N \times N$ array called recurrence matrix R is computed as follows:

$$R_{i,j}(\varepsilon) = \Theta\left(\varepsilon - \left\|\vec{x}_i - \vec{x}_j\right\|\right) \tag{2}$$

where N is the number of intended states $\vec{x}$, $\Theta(x)$ is the Heaviside function (i.e., $\Theta(x) = 0$ if $x < 0$ and $\Theta(x) = 1$ otherwise), ε is the threshold distance, and $\|\cdot\|$ is a norm, Thus recurrence matrix is a matrix consisting of ones and zeros. To calculate recurrence matrix, a suitable norm has to be selected. In this paper, we used Euclidean norm for calculating the distance between two states. RP of each dynamical system has its own topology. For example, RP related to periodic systems has uncut and long diagonal lines. The vertical distance between these diagonal lines indicates the period of the fluctuations. The RP of chaotic system also has diagonal lines, which are shorter than periodic systems with certain vertical distances. But, vertical distances in chaotic systems are not as regular as in the periodic systems. The RP of the uncorrelated stochastic signal consists of many single black points.

3.1.3. Recurrence Quantification Analysis (RQA)

To quantify the structures in RPs, several measures of complexity have been proposed. These measures are known as RQA (Marwan, Romano, Thiel, & Kurths, 2007) and are based on the recurrence point density, the diagonal and vertical line structures, recurrence time, and recurrence network.

3.1.3.1. Recurrence Point Density Measure

$$RR = \frac{1}{N^2} \sum_{i,j=1}^{N} R_{i,j}(\varepsilon) \tag{3}$$

The measure RR describes the probability that a state recurs to its ε-neighborhood in phase space.

3.1.3.2. Diagonal Line Measures

These measures are calculated from the histogram $P(\varepsilon, l)$ of diagonal lines with a length of l. The RP of stochastic systems has none or short diagonal lines structure and more single points, while deterministic systems are characterized by longer diagonal lines and less single isolated recurrence points**.**

Determinism (DET): The *determinism* (or predictability) of a system can be measured by the ratio of recurrence points that form diagonal structures (of at least length l_{min}) to all recurrence points:

$$DET = \frac{\sum_{l=l_{min}}^{N} l \cdot P(l)}{\sum_{l=1}^{N} l \cdot P(l)} \tag{4}$$

where l_{min} is the least length and $P(l)$ is the frequency distribution of the length *l* of the diagonal structures in the RP. In this paper, we select $l_{min} = 2$.

Mean diagonal line length (L): A diagonal line of length *l* shows that a segment of the trajectory is partly close during *l* time step to another segment of the trajectory at a different time. These diagonal lines indicate the divergence of the trajectory segments. The average time that two segments of the trajectory are close to each other can be measured by the average diagonal line length, and can be considered as the mean prediction time (Marwan, Romano, Thiel, & Kurths, 2007):

$$L = \frac{\sum_{l=l_{min}}^{N} l \cdot P(l)}{\sum_{l=l_{min}}^{N} P(l)} \tag{5}$$

The main diagonal is not considered for calculation L.

Maximal diagonal line length (L_{max}): L_{max} is the length of the longest diagonal line in the RP that is parallel to the main diagonal (Marwan, Romano, Thiel, & Kurths, 2007). The main diagonal line is not considered for calculation of L_{max}. The exponential divergence of the phase space trajectory is measured by L_{max}. The shorter diagonal lines indicate faster trajectory divergence.

Entropy of the diagonal line lengths ($ENTR$): The entropy of the length of diagonal lines is calculated as follows:

$$ENTR = -\sum_{l=l_{min}}^{N} p(l) \ln p(l) \tag{6}$$

where $p(l)$ is the probability that a diagonal line has exactly length l. Complexity in the RP can be measured by $ENTR$. The small value of $ENTR$ indicates strong regularity and less complexity and the large value indicates significant fluctuations.

3.1.3.3. Vertical Line Measures

The chaos–chaos, order-chaos, and chaos-order transitions can be found by vertical line measures (Marwan et al., 2002). Hence, these measures are appropriate for investigating the intermittency and short and non-stationary data series.

Laminarity (LAM): The ratio of the recurrence points forming the vertical structures to the entire set of recurrence points is defined as the LAM as follows:

$$LAM = \frac{\sum_{v=v_{min}}^{N} v\, P(v)}{\sum_{v=1}^{N} v\, P(v)} \tag{7}$$

where the $P(v)$ is the histogram of the vertical lines with length v and v_{min} is a minimal length.

Trapping time (TT): TT is the average length of vertical structures as follows

$$TT = \frac{\sum_{v=v_{min}}^{N} vP(v)}{\sum_{v=v_{min}}^{N} P(v)} \tag{8}$$

and describes the mean time that the system remains in a state.

Maximal vertical line length (V_{max}): This measure is the length of the longest vertical line in the RP:

$$V_{max} = max(\{v_i; i = 1, \dots, N_v\}) \tag{9}$$

where N_v is the total number of vertical lines in RP.

3.1.3.4. Recurrence Time based Measures

Three RQA measures based on recurrence time statistics have been proposed for detecting the transitional signals in noisy and nonstationary environments (Gao et al., 2003). These measures are called the first type (T_1) and the second type (T_2) of recurrence time and recurrence period density entropy ($RPDE$).

First type (T_1) *and second type* (T_2) *of recurrence time:* To define the 2nd type of recurrence time, consider a scalar time series $\{u(i), i = 1,2, \dots\}$ and corresponding reconstructed trajectory in *m*-dimensional phase space as $x_t = (u_t, u_{t+\tau}, \dots, u_{t+(m-1)\tau})$. An arbitrary reference point (x_0) on the reconstructed trajectory is selected, then a neighborhood of radius *r* for reference point $B_r(x_0) = \{x: \|x - x_0\| \leq r\}$ is defined. The set of points consisting of the first trajectory point getting inside the neighborhood from outside are defined as recurrence points of the 2nd type (Fig. 2). The trajectories that remain inside the neighborhood for a while, produce a sequence of points that are called the sojourn points (white circle in Fig. 2). The set of the recurrence points of the second type as well as the sojourn points constitute the recurrence points of the first type. If the recurrence points are defined as $S = \{x_{t1}, x_{t2}, \dots, x_{ti}, \dots\}$, then the corresponding recurrence time T is $\{T(i) = t_{i+1} - t_i, i = 1,2, \dots\}$.

Fig. 2

Recurrence period density entropy (RPDE): RPDE is a measure that can describe the complexity of a signal and determine the periodicity of a signal (Mukherjee et al., 2015). A system with periodic behavior has a *RPDE* with a value close to 0, whereas a system with chaotic behavior has a $RPDE$ close to 1 (Nair & Kiasaleh, 2014). The $RPDE$ can be computed as follows:

$$RPDE = -(lnT_{max})^{-1} \sum_{t=1}^{T_{max}} P(t) lnP(t) \quad (10)$$

where T_{max} and $P(t)$ are the largest recurrence value and the recurrence period density function, respectively.

3.1.3.5. Recurrence Network Analysis (RNA) based Measure

RNA measure is the so-called network transitivity ($Trans$) and is based on the adjacency matrix *A* elements (Webber & Marwan, 2015), defined as:

$$Trans = \frac{\sum_{i,j,\mathrm{k}=1}^{N} A_{i,j} A_{j,k} A_{k,i}}{\sum_{i,j,k=1}^{N} A_{i,j} A_{k,i}} \tag{11}$$

where $A = R - I$. *Trans* reflects network complexity and distinguishes between regular and irregular dynamics.

3.2. Feature Selection Based on Mutual Information

The relevance between two variables can be measured by MI. A formalism for quantifying MI is Shannon's information theory. Assume *X* is a random variable that represents continuous-valued random feature vector, and *C* is a discrete-valued random variable that represents the class labels, the MI between two variables *X* and *C* are calculated as follows:

$$I(X; C) = \sum_{c \in C} \int_x p(c, x) \log \frac{p(c,x)}{p(c)p(x)} dx \tag{12}$$

where $p(c, x)$ is the joint probability density function of *x* and *c*, $p(x)$ and $p(c)$ are the marginal probability density functions of *x* and *c*, respectively. A large value of the MI between two random variables indicates that two variables are closely related. If two random variables are strictly independent, the MI is zero. In terms of MI, the optimal feature selection requires selecting a feature set *f* with *m* features, which jointly have the largest dependency on the target class *C* (i.e., maximal dependency). That is, we seek:

$$\max D(f, c), \quad D = I(f; C) \tag{13}$$

$$I(f,C)=\sum_{c\in C}\int K\int p(f_1 K f_m)\log\frac{p(f_1 \Lambda f_m, c)}{p(f_1 \Lambda f_m)p(c)}df_1 \Lambda df_m$$

However, it requires an accurate estimation of the underlying probability density functions (pdfs) of the data and the integration on these pdfs. Moreover, due to the tremendous computational requirements of the method, the practical applicability of the above solution to the problems requiring a large number of features is limited. To overcome this problem, a heuristic method proposed in (Peng, Long, & Ding, 2005), which is based on minimal-redundancy-maximal-relevance (mRMR) framework. It was proven that mRMR criterion is equivalent to maximal dependency (13) if one feature is added at one time (Peng, Long, & Ding, 2005). This criterion is given by:

$$J=\left\{I(f_i;c)-\beta\sum_{f_s\in S}I(f_i;f_s)\right\} \tag{14}$$

According to this criterion, Term $I(f_i;c)$ indicate dependency between a new feature f_i and the target class that should be maximized (i.e., $\max_i I(f_i;c)$) and the term $\sum_{f_s\in S}I(f_i;f_s)$ indicates the dependency of the new feature with the already selected features. This term should be minimized (i.e., $\min_i \sum_{f_s\in S}I(f_i;f_s)$). The parameter β is the redundancy parameter, which regulates the relative importance of the MI between the new feature and the already selected features with respect to the MI with the output class.

3.3. Classification of EEG Features

Each EEG segment was split into 16 blocks of 1.475 s duration. Original features were formed from each block. Thus, 1600 feature vectors were constructed from each EEG subset. Then, the MI-based feature selection process was carried out to select optimal feature vector. For classification of the selected features, a two-layer feed-forward neural network was employed to perform the classification. The scaled conjugate gradient algorithm was used to train the network using the selected feature vectors. The number of neurons in the hidden layer was 20 and the output layer equal to the number of classes. Repeated random subsampling for evaluation, whereas during each repeat, 60%, 5%, and 35% of the feature vectors are randomly selected for training, validation, and testing, respectively. The evaluation procedure was repeated 20 times and the mean and standard deviation were calculated. Classifications were executed using the well-known MATLAB software package.

For the EEG dataset described in Section 2, five different cases of classification were considered. The cases were selected due to their clinical relevance and their wide usage by the researchers (Alam & Bhuiyan, 2013; Riaz et al., 2016). In Case I, the sets A and B were grouped as *healthy* class, the sets C and D were grouped as *interictal* class, and the set E was recognized as *ictal* class. In Case II, the sets A, D, and E were considered as *healthy*, *interictal*, and *ictal* classes, respectively. In Case III, the sets A and E were classified as *healthy* and *ictal* classes, respectively. In Case IV, the sets A, B, C, and D were grouped as *nonseizure* class and the set E as *seizure* class. In Case V, the first class consisted of the set D as *interictal* class and the second class included the set E as *ictal* class.

4. Results and Discussion

4.1. RP of the EEG signals

Fig. 3 shows examples of RP of the EEG recordings corresponding to healthy (A and B), interictal (C and D), and ictal (E) conditions. It is observed that there is vertical and horizontal line structure in the RP of the healthy subject (Fig. 3 (a) and (b)). The rectangles formed by the vertical and horizontal lines indicate that the system trapped in a state and does not change or changes very slowly for some time. The vertical structures in the RP of EEG indicate intermittency and laminar. Interesting observation is the white band structures during seizure-free (Fig. 3 (c) and (d)). White area or bands correspond to sudden changes in the dynamic as well as extreme events (Webber & Marwan, 2015). During a seizure, diagonal lines and checkerboard structures are observed in RP (Fig. 3 (e) and (f)). These structures indicate the system with periodic or quasi-periodic behavior (Webber & Marwan, 2015). The results demonstrate that the RP can visualize the dynamic changes of the EEG signals during different brain states.

Fig. 3

4.2. MI based Feature Selection

Fig. 4 shows the results of feature selection using mRMR for the Case I. It is observed that L_{max} is the first relevant feature that is selected (Fig. 4 (a)). According to mRMR criterion, the feature that has maximum MI with the class labels is selected as the first relevant feature. As already mentioned, L_{max} is a RQA measure is based on the diagonal lines structures and indicates repeating recurrences within a state. The diagonal lines are long for periodic signals and short for chaotic signals (Webber & Marwan, 2015). The second, third, fourth, and fifth selected features are V_{max}, RR, $RPDE$ and DET, respectively (Fig. 4 (b)-(e)).

Table 1 summarizes the results of feature selection using mRMR for different cases of classification. It is observed that in all cases, L_{max} is the first feature that is selected. Moreover, V_{max} is also selected in all cases.

Fig. 4

Table 1

4.3. Classification

The classification accuracy for different RQA-based features is shown in Fig. 5. It is observed that L_{max} and *Trans* features provide a high classification accuracy for the Cases III, IV, and V with respect to the cases I and II. This is because the EEG data grouped into two classes in the Cases III, IV, and V while the classification Cases of I and II have three classes. Moreover, diagonal structures within the RP reflect the system with periodic and quasi-periodic behavior and the *Trans* feature can distinguish regular from irregular dynamics. The L could discriminate accurately the Case III which contains only *healthy* and *ictal* classes.

Fig. 5

The average of overall detection accuracy, using selected features by mRMR algorithm and the feature vectors used in (Niknazar et al., 2013) is shown in Table 2. The results show that the average of detection accuracy is 100% using only five selected features for all cases. In (Niknazar et al., 2013), RQA was applied on the EEG recordings provided by Dr. R. Andrzejak as in the current study. The RQA-based features (i.e., DET, L, L_{max}, $ENTR$, LAM, TT) of the original signal and their subbands (i.e., delta, theta, alpha, beta, and gamma) were used for classification. The overall accuracy was 89.50% and 98.67% using the RQA-based features of the original signal (i.e., six features) and a combination of the original signal and subbands (i.e., 36 features), respectively.

Table 2

The classification accuracies obtained in this study and in the previous studies are summarized in Table 3. Only the previous studies that used the data set provided by Dr. R. Andrzejak were considered for comparison to provide a fair comparison.

Table 3

5. Conclusion

There is significant interest in developing accurate automatic seizure detection. The classification of EEG into healthy, ictal, and interictal EEGs is the main goal of seizure detection. Two of the major components of a classification process are the feature extraction and feature selection. Different linear approaches have been proposed for time series analysis of EEG signal and extraction features. However, the linear approaches ignore the underlying nonlinear EEG dynamics. The complex nonlinear EEG dynamics show different transitions between regular, laminar, and chaotic behaviors. The knowledge of these transitions is necessary for characterizing the underlying dynamics. A very useful nonlinear approach for measuring the complexity of a nonlinear dynamical system is RQA. Up to now, different RQA measures, including RR, DET, L, L_{max}, $ENTR$, LAM, TT, and $trend$ have been used as the features of the EEG signal for seizure detection. In the current study, different RQA measures, including V_{max}, T_1, T_2, $RPDE$ and $Trans$ have been introduced as the features for the EEG classification. These measures are very important for detecting the dynamic transitions and measuring the complexity.

Moreover, a systematic approach based on MI has been proposed to select the most relevant features. The first selected feature in all cases was L_{max}. Deterministic processes have longer diagonals and less single, isolated recurrence points, whereas chaotic signals cause the short diagonal lines. The diagonal lines for periodic signals are long and for stochastic signals are absent (Webber & Marwan, 2015).

During the interictal state, the EEG signals have lower amplitude and are less rhythmic and more irregular in morphology. During ictal state, an abrupt change in the amplitude, frequency, and morphology of the EEG signals occurs, and rhythmicity increases and a synchronization of activity occurs happens across widespread areas of the cerebral cortex. Therefore, diagonal lines can provide a suitable measure for prediction of rhythmic and periodic EEG patterns. As it can be seen in Fig. 3 (e) and (f), the RP of the brain signals during ictal state has the checkerboard structures indicating periodic behavior whereas such structures have not been observed during healthy and interictal states.

The second selected feature is V_{max}, which indicates the vertical line structure in the RP. RP of the healthy signal (Fig. 3 (a) and (b)) contains vertical and horizontal lines that form rectangles. This structure indicates that some states do not change or change slowly for some time (laminar states) or the process is halted at a singularity in which the dynamic is stuck in paused states.

Another selected feature is *DET*, which is a measure of *determinism.* In the seizure state, excessive synchronization of large neuronal populations occurs, leading to a hypersynchronous state which implies an increasing determinism of EEG data. Therefore, *DET* can be a suitable measure for seizure detection. $Trans$ and $RPDE$ are other complexity measures which are selected as the features.

The results of this study show that a robust accurate seizure detection with a short period of time (1.475 s) can be obtained using the proposed method. The method could distinguish *healthy*, *ictal*, and *interictal* states with 100% classification accuracy.

Acknowledgement

This work was supported by the Iran University of Science and Technology, Tehran, Iran under Grant 48.M.111194.

Conflict of interest

There are no conflicts of interest for the authors of this study.

References

Adeli, H., Ghosh-Dastidar, S., & Dadmehr, N. (2007). A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy. IEEE Transactions on Biomedical Engineering, 54(2), 205-211.

Alam, S. S., & Bhuiyan, M. I. H. (2013). Detection of seizure and epilepsy using higher order statistics in the EMD domain. IEEE Journal of Biomedical and Health Informatics, 17(2), 312-318.

Bajaj, V., & Pachori, R. B. (2012). Classification of seizure and nonseizure EEG signals using empirical mode decomposition. IEEE Transactions on Information Technology in Biomedicine, 16(6), 1135-1142.

Cao, L. (1997). Practical method for determining the minimum embedding dimension of a scalar time series. Physica D: Nonlinear Phenomena, 110(1-2), 43-50.

Cerf, R., Amri, M. El., Ouasdad, E. H., & Hirsch, E. (1999). Non-linear analysis of epileptic seizures I. Correlation-dimension measurements for absence epilepsy and near-periodic signals. Biological cybernetics, 80(4), 247-258.

Chen, T. J., Chiueh, H., Liang, S. F., Chang, S. T., Jeng, C., Hsu, Y. C., & Chien, T. C. (2011). The implementation of a low-power biomedical signal processor for real-time epileptic seizure detection on absence animal models. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 1(4), 613-621.

D'Alessandro, M., Esteller, R., Vachtsevanos, G., Hinson, A., Echauz, J., & Litt, B. (2003). Epileptic seizure prediction using hybrid feature selection over multiple intracranial EEG electrode contacts: a report of four patients. IEEE Transactions on Biomedical Engineering, 50(5), 603-615.

Easwaramoorthy, D., & Uthayakumar, R. (2011). Improved generalized fractal dimensions in the discrimination between healthy and epileptic EEG signals. Journal of Computational Science, 2(1), 31-38.

Fraser, A. M., & Swinney, H. L. (1986). Independent coordinates for strange attractors from mutual information. Physical review A, 33(2), 1134-1140.

Gao, J. B., Cao, Y., Gu, L., Harris, J. G., & Principe, J. C. (2003). Detection of weak transitions in signal dynamics using recurrence time statistics. Physics Letters A, 317(1), 64-72.

Ghosh-Dastidar, S., Adeli, H., & Dadmehr, N. (2007). Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection. IEEE transactions on biomedical engineering, 54(9), 1545-1551.

Ghosh-Dastidar, S., Adeli, H., & Dadmehr, N. (2008). Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection. IEEE Transactions on Biomedical Engineering, 55(2), 512-518.

Güler, N. F., Übeyli, E. D., & Güler, I. (2005). Recurrent neural networks employing Lyapunov exponents for EEG signals classification. Expert systems with applications, 29(3), 506-514.

Guo, L., Rivero, D., Dorado, J., Munteanu, C. R., & Pazos, A. (2011). Automatic feature extraction using genetic programming: An application to epileptic EEG classification. Expert Systems with Applications, 38(8), 10425-10436.

Guo, L., Rivero, D., Dorado, J., Rabunal, J. R., & Pazos, A. (2010). Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks. Journal of neuroscience methods, 191(1), 101-109.

Iasemidis, L. D., Shiau, D. S., Chaovalitwongse, W., Sackellares, J. C., Pardalos, P. M., Principe, J. C., Carney, P. R., Prasad, A., Veeramani, B., & Tsakalis, K. (2003). Adaptive epileptic seizure prediction system. IEEE transactions on biomedical engineering, 50(5), 616-627.

Iscan, Z., Dokur, Z., & Demiralp, T. (2011). Classification of electroencephalogram signals with combined time and frequency features. Expert Systems with Applications, 38(8), 10499-10505.

Kumar, Y., Dewal, M. L., & Anand, R. S. (2014). Epileptic seizure detection using DWT based fuzzy approximate entropy and support vector machine. Neurocomputing, 133, 271-279.

Kwak, N., & Choi, C. H. (2002). Input feature selection for classification problems. IEEE Transactions on Neural Networks, 13(1), 143-159.

Li, X., Ouyang, G., Yao, X., & Guan, X. (2004). Dynamical characteristics of pre-epileptic seizures in rats with recurrence quantification analysis. Physics Letters A, 333(1), 164-171.

Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. Physics reports, 438(5), 237-329.

Marwan, N., Wessel, N., Meyerfeldt, U., Schirdewan, A., & Kurths, J. (2002). Recurrence-plot-based measures of complexity and their application to heart-rate-variability data. Physical review E, 66(2), 026702.

Mukherjee, S., Palit, S. K., Banerjee, S., Ariffin, M. R. K., Rondoni, L., & Bhattacharya, D. K. (2015). Can complexity decrease in congestive heart failure?. Physica A: Statistical Mechanics and its Applications, 439, 93-102.

Naghsh-Nilchi, A. R., & Aghashahi, M. (2010). Epilepsy seizure detection using eigen-system spectral estimation and Multiple Layer Perceptron neural network. Biomedical Signal Processing and Control, 5(2), 147-157.

Nair, A., & Kiasaleh, K. (2014). Function mapped trajectory estimation for ECG sets. Biomedical Engineering Letters, 4(3), 277-284.

Niknazar, M., Mousavi, S. R., Motaghi, S., Dehghani, A., Vahdat, B. V., Shamsollahi, M. B., Sayyah, M. & Noorbakhsh, S. M. (2013). A unified approach for detection of induced epileptic seizures in rats using ECoG signals. Epilepsy & Behavior, 27(2), 355-364.

Niknazar, M., Mousavi, S. R., Vahdat, B. V., & Sayyah, M. (2013). A new framework based on recurrence quantification analysis for epileptic seizure detection. IEEE journal of biomedical and health informatics, 17(3), 572-578.

Orhan, U., Hekim, M., & Ozer, M. (2011). EEG signals classification using the K-means clustering and a multilayer perceptron neural network model. Expert Systems with Applications, 38(10), 13475-13481.

Osorio, I., Zaveri, H. P., Frei, M. G., & Arthurs, S. (Eds.). (2016). Epilepsy: the intersection of neurosciences, biology, mathematics, engineering, and physics. CRC press.

Ouyang, G., Li, X., Dang, C., & Richards, D. A. (2008). Using recurrence plot for determinism analysis of EEG recordings in genetic absence epilepsy rats. Clinical neurophysiology, 119(8), 1747-1755.

Oweis, R. J., & Abdulhay, E. W. (2011). Seizure classification in EEG signals utilizing Hilbert-Huang transform. Biomedical engineering online, 10(1), 38.

Pachori, R. B. (2008). Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. Research Letters in Signal Processing, 2008, 14.

Pachori, R. B., & Bajaj, V. (2011). Analysis of normal and epileptic seizure EEG signals using empirical mode decomposition. Computer Methods and Programs in Biomedicine, 104(3), 373-381.

Page, A., Sagedy, C., Smith, E., Attaran, N., Oates, T., & Mohsenin, T. (2015). A flexible multichannel EEG feature extractor and classifier for seizure detection. IEEE Transactions on Circuits and Systems II: Express Briefs, 62(2), 109-113.

Peker, M., Sen, B., & Delen, D. (2016). A novel method for automated diagnosis of epilepsy using complex-valued classifiers. IEEE journal of biomedical and health informatics, 20(1), 108-118.

Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. IEEE Transactions on pattern analysis and machine intelligence, 27(8), 1226-1238.

Ramgopal, S., Thome-Souza, S., Jackson, M., Kadish, N. E., Fernández, I. S., Klehm, J., Bosl, W., Reinsberger, C., Schachter, S., & Loddenkemper, T. (2014). Seizure detection, seizure prediction, and closed-loop warning systems in epilepsy. Epilepsy & behavior, 37, 291-307.

Riaz, F., Hassan, A., Rehman, S., Niazi, I. K., & Dremstrup, K. (2016). EMD-based temporal and spectral features for the classification of EEG signals using supervised learning. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24(1), 28-35.

Srinivasan, V., Eswaran, C., & Sriraam, N. (2007). Approximate entropy-based epileptic EEG detection using artificial neural networks. IEEE Transactions on information Technology in Biomedicine, 11(3), 288-295.

Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. Expert Systems with Applications, 32(4), 1084-1093.

Subasi, A., & Gursoy, M. I. (2010). EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Systems with Applications, 37(12), 8659-8666.

Temko, A., Nadeu, C., Marnane, W., Boylan, G., & Lightbody, G. (2011). EEG signal description with spectral-envelope-based speech recognition features for detection of neonatal seizures. IEEE Transactions on Information Technology in Biomedicine, 15(6), 839-847.

Thomas, E. M., Temko, A., Marnane, W. P., Boylan, G. B., & Lightbody, G. (2013). Discriminative and generative classification techniques applied to automated neonatal seizure detection. IEEE journal of biomedical and health informatics, 17(2), 297-304.

Thomasson, N., Hoeppner, T. J., Webber, C. L., & Zbilut, J. P. (2001). Recurrence quantification in epileptic EEGs. Physics Letters A, 279(1), 94-101.

Tzallas, A. T., Tsipouras, M. G., & Fotiadis, D. I. (2009). Epileptic seizure detection in EEGs using time–frequency analysis. IEEE transactions on information technology in biomedicine, 13(5), 703-710.

Übeyli, E. D. (2009). Statistics over features: EEG signals analysis. Computers in Biology and Medicine, 39(8), 733-741.

Van Drongelen, W., Nayak, S., Frim, D. M., Kohrman, M. H., Towle, V. L., Lee, H. C., McGee, A. B., Chico, M. S., & Hecox, K. E. (2003). Seizure anticipation in pediatric epilepsy: use of Kolmogorov entropy. Pediatric neurology, 29(3), 207-213.

Wang, D., Miao, D., & Xie, C. (2011). Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection. Expert Systems with Applications, 38(11), 14314-14320.

Wang, N., & Lyu, M. R. (2015). Extracting and selecting distinctive EEG features for efficient epileptic seizure prediction. IEEE journal of biomedical and health informatics, 19(5), 1648-1659.

Webber, C. L., & Marwan, N. (2015). Recurrence Quantification Analysis. Springer International Publishing, Switzerland.

Wlodarczyk, B. J., Palacios, A. M., George, T. M., & Finnell, R. H. (2012). Antiepileptic drugs and pregnancy outcomes. American Journal of Medical Genetics Part A, 158(8), 2071-2090.

Yang, C., Jeannès, R. L. B., Bellanger, J. J., & Shu, H. (2013). A new strategy for model order identification and its application to transfer entropy for EEG signals analysis. IEEE Transactions on Biomedical Engineering, 60(5), 1318-1327.

Yaylali, I., Koçak, H., & Jayakar, P. (1996). Detection of seizures from small samples using nonlinear dynamic system theory. IEEE Transactions on Biomedical Engineering, 43(7), 743-751.

Zabihi, M., Kiranyaz, S., Bahrami Rad, A., Katsaggelos, A. K., Gabbouj, M., & Ince, T. (2016). Analysis of high-dimensional phase space via Poincaré section for patient-specific seizure detection. IEEE Transactions on Neural Systems and Rehabilitation Engineering, 24(3), 386-398.

Zhang, Z., & Parhi, K. K. (2016). Low-complexity seizure prediction from iEEG/sEEG using spectral power and ratios of spectral power. IEEE transactions on biomedical circuits and systems, 10(3), 693-706.

Zhou, W., Liu, Y., Yuan, Q., & Li, X. (2013). Epileptic seizure detection using lacunarity and Bayesian linear discriminant analysis in intracranial EEG. IEEE Transactions on Biomedical Engineering, 60(12), 3375-3381.

Tables/Figure Legends

Table 1. Selected features using mRMR algorithm for difference cases of classification.

Table 2. The average of classification accuracy (± Standard Deviation) for different number of selected features.

Table 3. A comparison of the results obtained by the proposed method and others′ methods.

Figure 1. Examples of EEG signals from each of the five subsets A, B, C, D, and E.

Figure 2. A schematic of the recurrence points of the second type (solid circles) and the sojourn points (open circles).

Figure 3. Recurrence plot of a block of subsets A (a), B (b), C (c), D (d), and E (e-f).

Figure 4. Feature selection process using mRMR algorithm for the Case I. Selection of the first (a), second (b), third (c), fourth (d), and fifth (e) feature.

Figure 5. Classification accuracy of different features for different cases of classification (Case I: Dark blue, Case II: Blue, Case III: Green, Case IV: Red, Case V: Brown).

Table 1. Selected features using mRMR algorithm for difference cases of classification.

Case	Case I	Case II	Case III	Case IV	Case V
1st selected feature	L_{max}	L_{max}	L_{max}	L_{max}	L_{max}
2nd selected feature	V_{max}	V_{max}	DET	V_{max}	V_{max}
3rd selected feature	RR	LAM	V_{max}	RR	LAM
4th selected feature	PRDE	T1	T1	PRDE	Trans
5th selected feature	DET	T2	TT	DET	T1

Table 2. The average of classification accuracy (± Standard Deviation) for different number of selected features.

Case / RQA-based features	signal	I	II	III	IV	V
3 selected features (mRMR)	Original	87.91±4.39	98.61±4.79	99.85±0.64	99.6±0.43	98.9±0.34
4 selected features (mRMR)	Original	97.59±1.52	99.28±1.52	100	100	100
5 selected features (mRMR)	Original	100	100	100	100	100
DET, L, L_{max}, $ENTR$, LAM, TT (features used in [29])	Original	89.5±1.72	-	-	-	-
	Delta band	67.46±2.68	-	-	-	-
	Theta band	77.53±2.43	-	-	-	-
	Alpha band	63.73±3.35	-	-	-	-
	Beta band	82.73±2.26	-	-	-	-
	Gamma band	86.6±7.5	-	-	-	-
	Original + subbands	98.67±0.52	-	-	-	-

Table 3. A comparison of the results obtained by the proposed method and others' methods.

Case / Authors	Case 1	Case II	Case III	Case IV	Case V	Number of feature	Block Duration	method
Peker et al. (2016)	98.28%	99.30%	100	99.33%	-	5	23.6	DTCWT-CVANN
Shafiul Alam et al. (2013)	80%	100%	100%	100%	100%	3	1.475	EMD-higher order moments- neural network
Riaz et al. (2016)	94%	91%	99%	98%	96%	6	23.6	EMD-Based Temporal and Spectral Features
Srinivasan et al. (2007)	-	-	100	-	-	1	2.95	ApEn- neural network
Ghosh-Dastidar et al. (2007)	-	96.7%	-	-	-	9	23.6	Mixed-band wavelet-chaos, Levenberg-Marquardt backpropagation NN
Niknazar et al. (2013)	98.67%	-	-	-	-	36	23.6	RQA on EEG signal and its wavelet-based sub-bands-ECOC
Kumar et al. (2014)	-	-	100%	97.38%	95.85%	6	23.6	DWT-Fuzzy ApEn-SVM
Guo et al. (2011)	-	93.5%	99.2%	-	-	3	23.6	Genetic algorithm-KNN
Orhan et al. (2011)	95.6%	96.67%	100%	99.6%	-	56 56 4 18	23.6	DWT-K-means clustering-probability distribution-MLPNN
Iscan et al. (2011)	-	-	100%	-	-	10	1.475	combined time and frequency features
Wang et al. (2011)	-	-	99.44%	-	-	4	1.475	wavelet packet entropy-hierarchical EEG classification
Naghsh-Nilchi et al. (2010)	-	97.49%	-	-	-	27	23.6	Eigen-system spectral estimation-MLPNN
Subasi et al. (2010)	-	-	100%	-	-	24	2.95	DWT-PCA, ICA, LDA and SVM
Guo et al. (2010)	-	-	99.6%	97.77%	-	5	23.6	DWT-line length feature-MLPNN
Ubeyli (2009)	-	96.33%	-	-	-	9	1.475	DWT-Lyapunov exponents, Eigenvector-MLP
Tzallas et al. (2009)	-	100%	100%	-	-	3	23.6	Time-Frequency analysis- neural network
Ghosh-Dastidar et al. (2008)	-	96.6%	-	-	-	9	23.6	Wavelet-chaos, PCA-NN
Subasi (2007)	-	-	94.5%	-	-	16	2.95	DWT-mixture of expert model
Guler et al. (2005)	-	96.79%	-	-	-	4	1.475	Lyapunov exponent-Recurrent neural network
This work	100%	100%	100%	100%	100%	5	1.475	RQA, mutual information, neural network

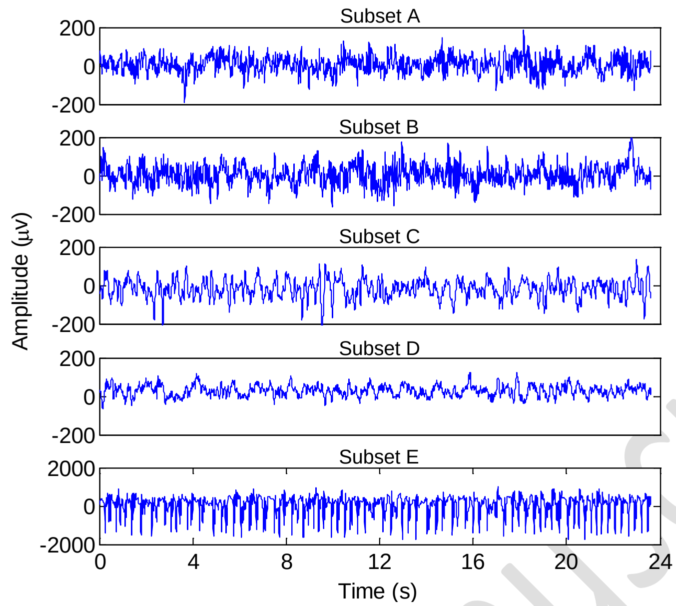

Figure 1. Examples of EEG signals from each of the five subsets A, B, C, D, and E.

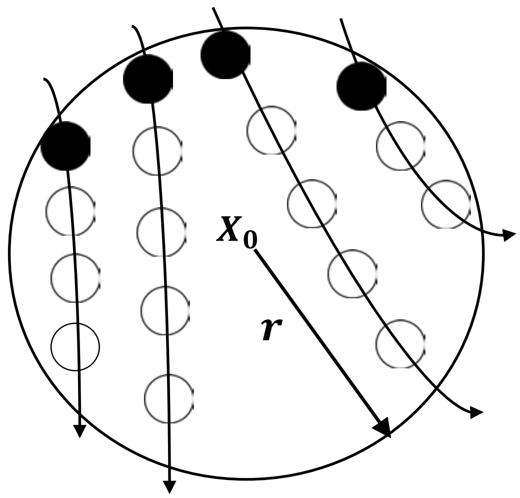

Figure 2. A schematic of the recurrence points of the second type (solid circles) and the sojourn points (open circles).

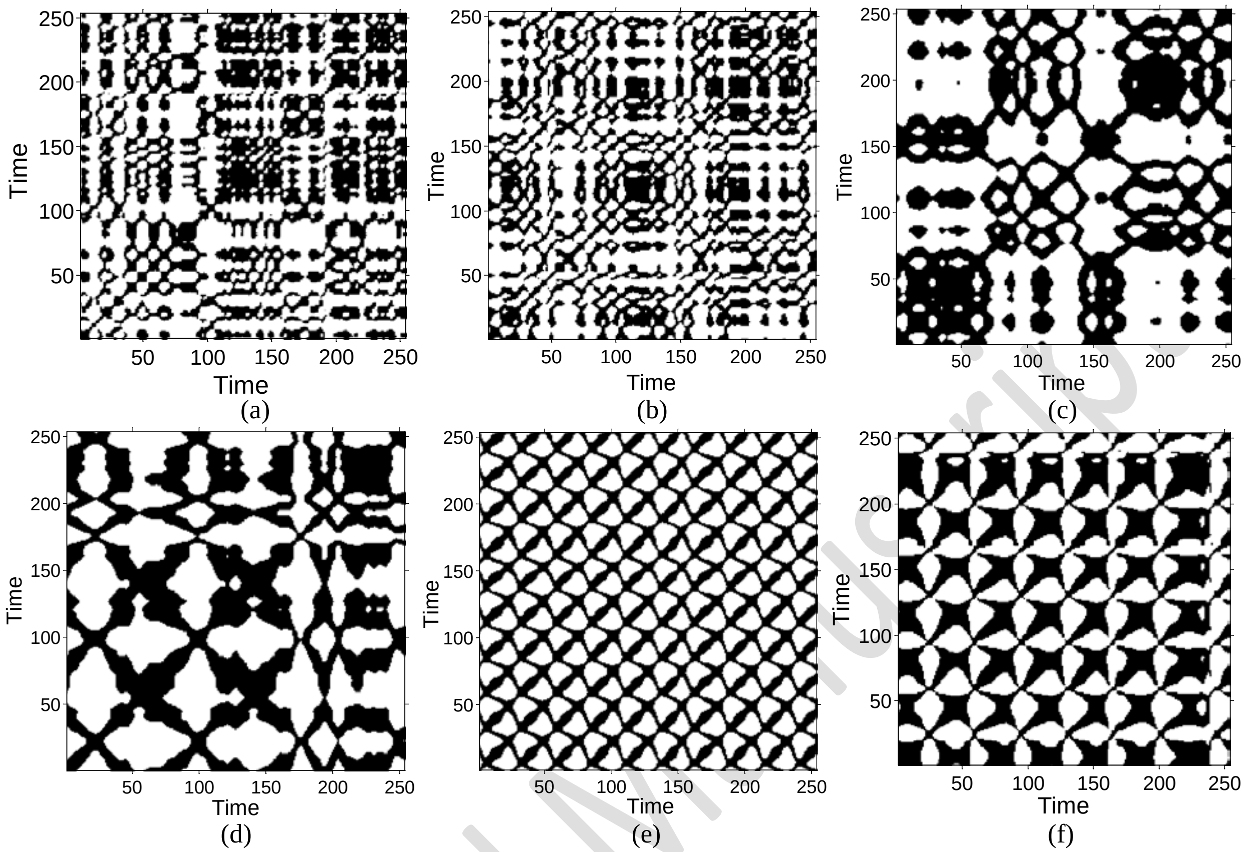

Figure 3. Recurrence plot of a block of subsets A (a), B (b), C (c), D (d), and E (e-f).

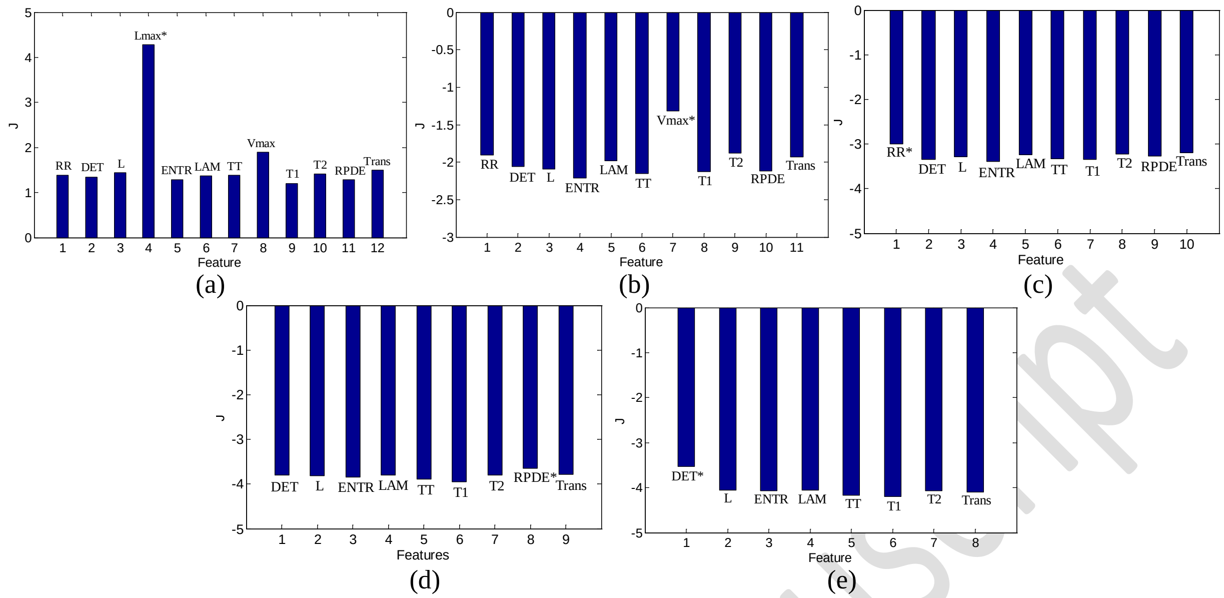

Figure 4. Feature selection process using mRMR algorithm for the Case I. Selection of the first (a), second (b), third (c), fourth (d), and fifth (e) feature.

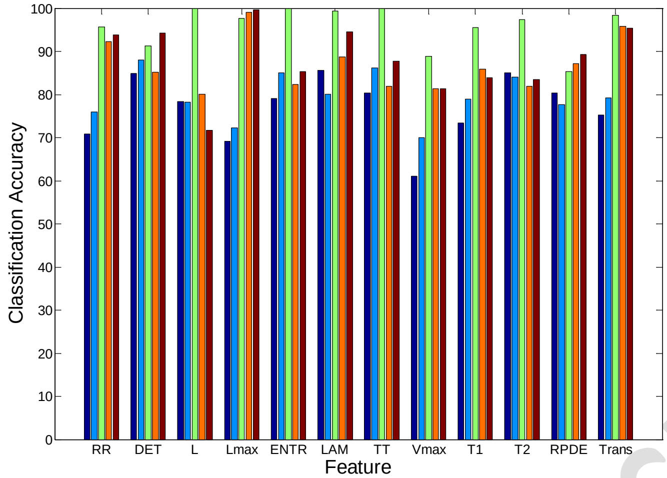

Figure 5. Classification accuracy of different features for different cases of classification (Case I: Dark blue, Case II: Blue, Case III: Green, Case IV: Red, Case V: Brown).