Journal of Scientific & Industrial
Research Vol. 80, May 2021, pp. 438- 448

Recurrence Quantification Analysis of EEG signals for Children with ASD

R Menaka[1]*, M Thanga Aarthy[2], Renuka Mahadev Chavan[2], R C Perumal[3] and Mahima S Menon[3]

[1]Centre for Cyber Physical Systems, [2]Department of Electronics Engineering, Vellore Institute of Technology, Chennai 600 127, India

[3]Department of Speech Language and Hearing Sciences, Sri Ramachandra Institute of Higher Education and Research, Chennai 600 116, India

Received 25 August 2020; revised 29 March 2021; accepted 23 April 2021

The present study aims at identifying the brain response for auditory/visual stimuli in typically developing (TD) and children with autism through electroencephalography (EEG). Early diagnoses do help in customized training and progressing the children in regular stream. To reveal the underlying brain dynamics, non-linear analysis was employed. In the current study, Recurrent Quantification Analysis (RQA) with varying parameters was analyzed. For better information retrieval, cosine distance metric is additionally considered for analysis and compared with other distance metrics in RQA. Each computational combination of RQA is measured and the responding channels were analyzed and discussed. It was observed that the FAN neighborhood with cosine distance parameters was able to discriminate between ASD and TD prominently.

Keywords: Autism Spectrum Disorder, Auditory/visual, Distance metric, Electroencephalogram, Fixed amount of nearest neighbor

Introduction

Autism Spectrum Disorder is a neuro developmental disorder that affects the child's ability to communicate and interact with the social world. The global median prevalence for Autism is 62 out of 10000.[1] The prevalence was observed to be quite high in developing countries like Srilanka (1.07%), Bangladesh (0.84%), and India (0.25%).[2] Diagnosis of ASD is quite challenging as the behavioral patterns of those kids are not obvious and it often requires series of clinical consultations with the physicians till three years of age. ASD is typically diagnosed with subjective assessment techniques like 3di, Childhood Autism Rating Scale, Autism Spectrum Disorder – Observation for children, etc.[3] These high-quality comprehensive assessments for the diagnosis of ASD involve huge manual effort and often result in increased waiting time for assessment.[4] This introduces the need for the design and development of objective assessment techniques for ASD.

ASD Diagnosing Methods

The objective assessment usually involves Electro-encephalogram (EEG)[5], functional Magnetic Resonance Imaging (f-MRI)[6], Eye-tracking[7], etc. Out of these modalities, EEG performs better in observing brain dynamic changes across time frequency and space.[8] Early diagnoses are always a challenge for children with ASD. Feature extraction methods such as Statistical[9], Spectral or Non-linear domain were employed in various research. Recurrence Quantification Analysis is considered best with threshold neighborhood selection and maximum distance metrics.[6,7,9,10] EEG spectral techniques were efficient in analyzing sleep disorder[11], motor imagery-based brain-computer interface[12], brain injury assessment.[13] Compared to statistical and spectral the abnormal dynamics can be well analyzed with non-linear analysis of EEG signal.[14]

In recent years, different non-linear approaches have been proposed to evaluate transitions, laminar, and chaotic behaviors. These behaviors characterize the underlying dynamics of the brain signal. Riley *et al*. explained contemporary methods in nonlinear analysis and their application in recurrence and cross recurrence analysis.[15] Other nonlinear approaches with EEG signals include autism detection, absence epilepsy compared with autism EEG patterns.[9,16]

Recurrence plot emergence dates back to 1987, wherein Eckman *et al*. computed the dynamic parameters from nonlinear time series. This method acts well for both short and non-stationary data but with the consideration of certain pitfalls.[17]

RQA is one of the best approaches to understand the chaotic property of brain signals.[18] Different

*Author for Correspondence
E-mail: menaka.r@vit.ac.in

recurrence measures were extracted such as RR, DET, L, Lmax, ENTR, LAM, TT, and Trend.[19] Other types of research using RQA were investigating emotions[20], burst suppression detection[21], ASD identification[22], pre-seizure states[23], left and right-hand movements[24], epilepsy biomarker[25], gaze, and patterns in shuffled text.[26,27]

The process which occurs in nature has a distinct recurrent behavior and can achieve an arbitrary close state after a period of time.[28] This can be visualized through the recurrence plot in RQA.

Through RQA one can extract the dynamics of the autism in children by quantifying the system's repeatability, complexity, and local dynamic stability through several variables. To attain a distinctive recurrence plot, some parameter selection is crucial.[21] This research aims at finding suitable parameters for differentiating ASD and TD children.

The present work focuses on the non-linear analysis of EEG signals using RQA. In this research, a deeper analysis of recurrence plot (RP) features was done with three different distance metrics such as Euclidean, Maximum, and cosine, and two neighborhood selections such as False Nearest Neighbor [FAN] and threshold method. The features extracted for analysis are Recurrence rate, Determinism, entropy, and average diagonal length. The significant feature value changes between children with ASD and TD were analyzed and discussed. Among various combinations, FAN neighborhood selection with cosine distance metrics provides the best discriminating features.

Materials and Methods

EEG data were acquired from 5 [3M, 2F] typically developing children and 5 [3M, 2F] autistic kids in the age group of 3–7 years. In this study, EEG records were collected from an acquisition system: Natus Nihon Ohden MEB9000 version 05-81, with a sensitivity of 7 microvolts. The data acquisition was performed using international 10-20 electrode systems as presented in Fig. 1. The EEG signals were recorded from 22 channels with a sampling frequency of 500 Hz and filtered with a low pass filter and high pass filter at a frequency range of [0.53-60 Hz].

These electrodes are Fp2-F4(ch1), F4-C4(ch2), C4-P4(ch3), P4-O2(ch4), Fp2-F8(ch5), F8-T4(ch6),T4-T6(ch7), T6-O2(ch8), Fp1-F3(ch9), F3-C3(ch10), C3-P3(ch11), P3-O1(ch12), Fp1-F7(ch13), F7-T3(ch14), T3-T5(ch15), T5- O1(ch16) as shown in Fig. 1.

The parents/guardians of the participants involved in this study were advised to refer and sign a consent form. The study protocol was based on the manual "Ethics of Scientific Research" at the Universidad Autonoma de Quer etaro. The subjects were seated in normal light and quiet rooms. The distance between the subject's eye and the 32' monitor was 55 cm depending upon the height. The Ag/AgCl electrodes were then positioned on the scalp using a conductive gel and tapes. The subject was made to watch an animated cartoon video for 10 minutes. The workflow of the research is as shown in Fig. 2. The child was made to sit in front of a visual screen. The electrodes were then placed on the scalp using a conductive gel and tapes. The child was made to watch a Cartoon with audio for 8 minutes.

Preprocessing

After filtering out the signal at a frequency range of 0.53 to 60 Hz, the ocular artifacts in the EEG signal were removed by the threshold method. The threshold is set based on the average amplitude of the eye blink signal. The eye blink signal was observed for 10 seconds with the eye open and eye close event.

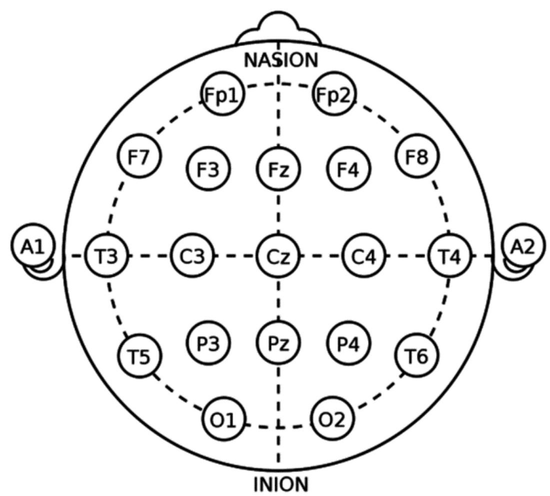

Fig. 1 — Electrode positions selected for this research [10-20 system]

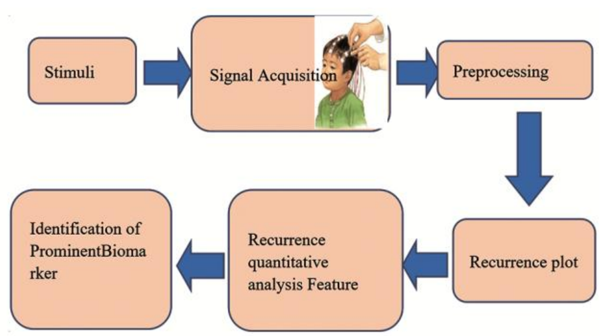

Fig. 2 — Block diagram of the study. M-Male, F-Female

After thresholding, the resultant signal is further considered for analysis.

Recurrence Quantitative Analysis

In recent times, RQA emerged as a powerful method for identifying the minor changes in nonlinear or non-stationary patterns of the time series. The significant part of RQA, is the visualization of the transitions in the recurrence plot (RP) which denotes the dynamic behavior of the time series. The recurrence plot and various parameter selections are explained in detail in the following sections.

The primary step in RQA is the phase space reconstruction and construction of RP. RP in the phase space gives information about the recurrence of the dynamic system states.

Recurrence Plot

A recurrence plot (RP) provides the visualization of complex dimensional phase spaces in a two-dimensional plot. The underlying dynamics of the system can be characterized by the visual appearance of the recurrence plot. A recurrence event is plotted based on, Eq. (1) and Eq. (2), for each sample combination i and j of time series x at a specified threshold distance ε (neighborhood size) are stored in an N x N matrix to construct a RP.[21]

$$R_{i,j}^{m,\varepsilon} = \Theta\left(\varepsilon_i - \| x_i - x_j \| \right); \quad \ldots(1)$$

where ε_i is the threshold value, m - embedding dimension

$\| . \|$ is the norm, and Θ is the Heaviside function.

$$\left\{ \begin{matrix} 1 : \| Xi - Xj \| < \varepsilon \\ Ri, j = 0 : \text{Otherwise } i, j = 1,2, .. N \end{matrix} \right\} \quad \ldots(2)$$

where Coordinates (X_i, X_j) whose distance is greater and lesser than ε falls into the category of 0 and 1. These are considered as non-recurrent and recurrent values.[5]

In practice, the state of a chaotic system such as an EEG signal would not recur precisely as the previous state. But it approaches the original state arbitrarily close. Therefore, a recurrence is defined as a state $\mathbf{x_j}$ is sufficiently close to $\mathbf{x_i}$. This means that those states $\mathbf{x_j}$ that fall into an m-dimensional neighborhood of size ε centered at x_i are recurrent. These $\mathbf{x_j}$ are called recurrence points. In Eq. (1), this is simply expressed by the Heaviside function.[5] For phase space reconstruction time-delayed embedding method is used by (Takens, 1981). This nonlinear method infers that one can recover the multidimensional dynamics from a one-dimensional time-series by plotting that time-series against itself at a certain time delay τ.[24]

Consider one-dimensional time series as in Eq. (3):

$$x = (x_1, x_2, \ldots.. x_n) \quad \ldots(3)$$

The estimation of the suitable dimensionality D and proper time-delay τ at which values of x must be shifted for each reconstructed dimension.[7]

The reconstructed phase space of the initial coordinate is expressed as Eq. (4),

$$V_1 = (x_1, x_{1+\tau}, x_{1+2\tau}, \ldots\ldots.. x_{1+(D-1)\tau}) \quad \ldots(4)$$

We can maximally construct $V_{n-(D-1)\ \tau}$ such coordinates which result in recurrence plot.

The combination of parameters for obtaining a recurrence plot and its features which were considered for analysis are depicted in the flow chart as shown in Fig. 3.

The trajectories observed preserve the fundamental information of the nonlinear dynamic system when it satisfies the condition m>2d+1 in which d represents the dimension of the attractor and m is the embedding dimension.[5] Embedding dimension and time delay have to be selected appropriately. Finding an optimal embedding parameter is a challenge to date. There exist some techniques to obtain an optimal embedding parameter. Initially embedding parameter was increased from a lower value up to an invariant solution. Another way is based on orthogonal direction but Liangyue *et al.* explain another method to determine the possible embedding dimension that is based on time delay and combination of false neighbor method.[29] Other methods use the ratios of the distances between the same neighboring points.[5] Recurrence plots are always used to find hidden correlations in complex dynamic systems. Joseph *et al.* discussed a recurrence plot that does no significant changes in most of the dataset used with the change in embedding dimension.[29] In this study, after preprocessing, the input signal by a notch filter and manually selecting the artifact-free EEG signal from the obtained input signal. The embedding dimension and time delay are ideally fixed as 3 and 1. Various parameters are considered to visualize the variation in recurrence features and recurrence plots. The parameters selected are fixed amount of nearest neighbors [FAN] with neighbor selection as 50 and threshold method [threshold is considered as 40% of the standard deviation of the EEG signal]. The various

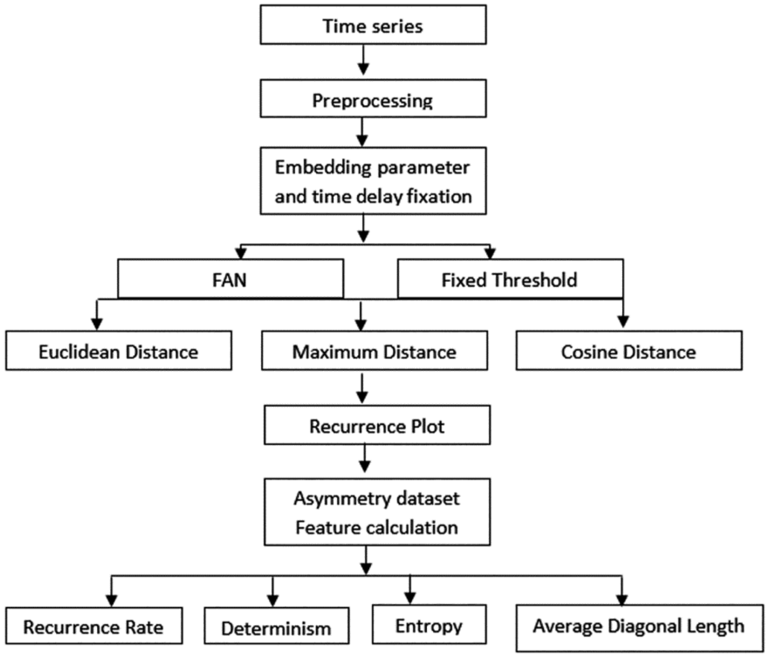

Fig. 3 — Flow for obtaining Recurrence plot and Recurrence Features

distance metrics within the parameters considered are Euclidean distance, maximum distance, and cosine distance.

The combinations considered are

1. FAN, Euclidean, Recurrence features
2. FAN, Maximum, Recurrence features
3. FAN, Cosine, Recurrence features
4. Threshold, Euclidean, Recurrence features
5. Threshold, Maximum, Recurrence features
6. Threshold, Cosine, Recurrence features

By considering the above combinations, the one which provides the efficient result in discriminating ASD and TD are analyzed. The efficiency of each method is explained as follows. In this research, the following recurrent features obtained from diagonal line histogram features are considered such as recurrence rate, determinism, entropy, and average diagonal length.

Parameter Selection

The process that occurs in nature has a distinct recurrent behavior. All these processes attain a state arbitrarily close to the original state after a period of time.[5] This can be visualized through a recurrence plot. To attain a distinctive recurrence plot, the parameters as embedding dimension, neighborhood selection, and distance metric need to be selected carefully.[5] For this analysis, we have considered the embedding dimension (m) as 3, the time delay as 1, for the recurrence plot using FAN the T value as 10, and threshold value as 40% of the standard deviation of the time series while using fixed threshold method. Recurrence Quantification Analysis (RQA) depends on factors such as recurrence point density. The diagonal structures are computed from the histogram of diagonal lines. The vertical line structures are computed from the histogram of the vertical lines. These structures give a quantified analysis rather than visual transitions. The parameters considered were explained in Table 1

Results

The results acquired from 5 children with ASD and 5 TD children are as follows. The mean value is calculated by considering each channel of all the TD and ASD children.

Results for various combinations of distance metrics and neighborhood selection are shown in the following tables and box plots. The recurrence plot obtained from ASD and TD children is shown below.

As there appeared no distinguishing features in the recurrence plot shown in Fig. 4. The patterns which are visualized in ASD/TD recurrence plot do not have any coordination with another ASD/TD child. Hence based on the visual observation a promising result cannot be concluded. In the future, with the help of learning algorithms, a prominent classification result can be obtained. Recurrence features were considered for further analysis.

Table 1 — RQA Parameters

Neighborhood Selection	
1. Fixed Amount of Nearest Neighbours (FAN)	This neighborhood selection method provides a fixed amount of closest states **x**j defined by parameter 'T'. With such a neighborhood, the radius changes for each **x**i and Ri,j ≠ Rj;i because the neighbourhood of **x**i and **x**j can be different. This property of FAN leads to an asymmetric RP
2. Fixed radius / threshold method (ε):	This uses fixed radius ε_i which ensures that Ri,j = Rj,i , i.e. a symmetric RP. ε chosen with two conditions such as <10% of maximum phase space diameter and 20% and 40% of standard deviation
Distance/Norm Types	
1. L2- Norm (Euclidian distance):	$ED = \sqrt{(x2-x1)^2+(y2-y1)^2}$ …(5) It is less prone to noise. If the points $(x_1,y_1)$and (x_2,y_2) are in 2-dimensional space, then the Euclidean distance between them is represented with Eq. (5)
2. L∞ - Norm (Maximum distance):	L_∞-norm is usually used because of it is simple computation and its property of independence to the phase space dimension.[5] The infinity norm (also known as the max norm, or uniform norm) is defined as the maximum of the absolute value as shown in Eq. (6) $\max\{\lvert x_i \rvert : i = 1, 2, \ldots, n\}$ …(6) where xis the input signal, i is the index and n is the length of the signal
3. Cosine metric	Cosine distance between two vectors of coordinates (a,b,c) and (x,y,z) is defined in Eq. (7) as: $Cosine\ Distance = 1 - \frac{ax+by+cz}{\sqrt{abs(a)^2+abs(b)^2+abs(c)^2}\sqrt{abs(x)^2+abs(y)^2+abs(z)^2}}$.. .(7) Cosine distance is the complement in positive space. It is widely used in information retrieval in high dimension positive spaces
Recurrence Features	
1. Recurrence rate (RR)	The equation for Recurrence Rate (RR) is given in Eq. (8) as: $RR = 1/N^2 \sum_{I,j=1}^{N} R_{i,j}$ …(8) Where $R_{i,j}$ is the recurrence plot and N is the size of the recurrence plot The recurrence rate is a measure of the probability that the exact state would occur.
2. Determinism (DET)	DET denotes the predictability of recurring patterns in the dynamic system. Regular patterns have higher DET values whereas chaotic patterns have lower DET values. The equation for Determinism is given in Eq. (9) as $DET = \sum_{l=lmin}^{N} l.P(l) / \sum_{l=1}^{N} l.P(l)$ …(9) where P(l) is the histogram of the diagonal line with length l.
3. Entropy (ENTR)	Entropy denotes the intricacy of the deterministic structure in the system. Entropy is defined in Eq. (10) as – $Entropy = - \sum_{l=lmin}^{N} P(l) . \ln(P(l))$ …(10) Lesser entropy value specifies the lower complexity and better regularities in signal and vice versa.
4. Avg. diagonal length (L_{avg})	P(l) is the histogram of the diagonal lines with length l. It is the computation of the longest diagonal line parallel to the main diagonal line. Average diagonal Length is defined in Eq. (11) as – $L_{avg} = \sum_{l=lmin}^{N} l.P(l) / \sum_{l=1min}^{N} P(l)$ …(11)

The performance of RQA based classification depends on the choice of parameters like embedding dimension, time delay, neighborhood selection, and distance metric. Different experiments were conducted by varying methods for neighborhood selection and distance metrics.

Analysis of FAN Neighborhood Selection and Distance Metrics:

In this section, the performance analysis of FAN neighborhood selection with different measures like RR, Determinism, ADL, and Entropy was analyzed. RQA based measures with these neighborhood techniques are performed and the resultant observations are highlighted in Tables 2 and 3. It is to be noted that, each of this neighborhood selection technique is analyzed with three distance metrics namely Euclidean distance, Maximum distance, and Cosine distance.

It is observed from Tables 1 and 2 that, determinism did not exhibit a significant difference

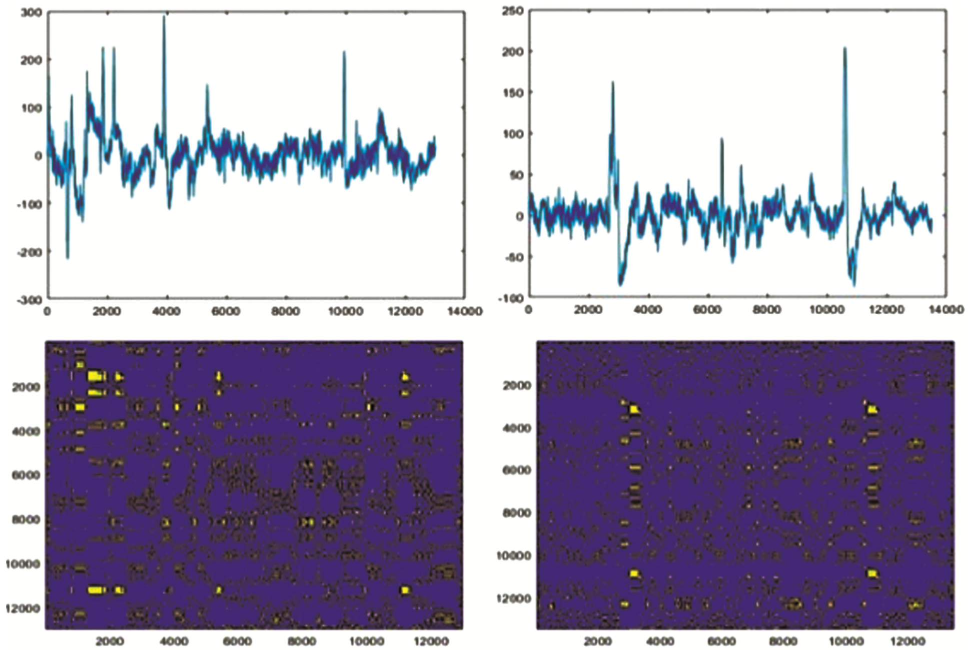

Fig. 4 — Recurrence plot of ASD and TD using threshold method [threshold selection 40% of the standard deviation of signal]

Table 2 — Features acquired from children with autism for various distance metrics [Euclidean, maximum, cosine] with neighborhood selection FAN

Euclidean Distance				Maximum Distance			Cosine Distance		
	AUTISM			AUTISM			AUTISM		
Channels	Determinism	Entropy	Average length	Determinism	Entropy	Average length	Determinism	Entropy	Average length
Fp2-F4	0.6191	0.8263	2.7768	0.673	0.9323	3.0119	0.3820	0.5005	2.4209
F4-C4	0.6155	0.8155	2.7373	0.6970	0.9712	3.042	0.4025	0.5251	2.4531
C4-P4	0.6362	0.8453	2.7918	0.6060	0.9399	3.0046	0.3902	0.5156	2.4620
P4-O2	0.6338	0.8204	2.706	0.6959	0.9580	3.0147	1.1562	0.503	2.4201
Fp2-F8	0.672	0.8963	2.8414	0.6803	0.9409	3.0325	0.4412	0.5681	2.4865
F8-T4	0.5717	0.7234	2.5815	0.5972	0.7704	2.7003	0.4313	0.5058	2.4357
T4-T6	0.5247	0.6556	2.5025	0.5564	0.7041	2.5746	0.3962	0.5001	2.3612
T6-O2	0.5811	0.7350	2.5931	0.6467	0.8539	2.8431	0.3847	0.4951	2.3698
Fp1-F3	0.66	0.8619	2.7917	0.6879	0.9362	2.9716	0.4193	0.5404	2.4464
F3-C3	0.6617	0.8622	2.7807	0.7061	0.9698	3.0168	0.4363	0.5659	2.4938
C3-P3	0.6940	0.9314	2.909	0.7457	1.0589	3.1835	0.4509	0.587	2.5244
P3-O1	0.5688	0.7161	2.5687	0.6114	0.7971	3.0257	0.3853	0.4901	2.3487
Fp1-F7	0.6971	0.9373	2.9103	0.7217	1.0135	3.1333	0.4400	0.5676	2.491
F7-T3	0.5574	0.7030	2.5786	0.6235	0.8131	2.7621	0.4351	0.5556	2.4423
T3-T5	0.5180	0.6513	2.5141	0.6005	0.7766	2.7057	0.4973	0.5413	2.4161
T5-O1	0.5418	0.6778	2.5307	0.6215	0.8125	2.7529	0.3889	0.4968	2.3650

between ASD and TD for different distance metrics. Also, Entropy yielded a notable difference between ASD and TD groups. Entropy, which quantifies the degree of irregularity, seems to be higher for ASD than TD. The magnitude of ADL measures how often how a particular pattern repeats in a 1D signal. Higher the magnitude of ADL, the lesser the repetitive behavior of the pattern. It is to be noted that the range of values obtained for the ADL parameter is comparatively greater for ASD children when compared to the TD group. Out of the three distance measures, Cosine distance obtained maximum difference for the difference between ASD and TD groups. The box plot analysis for ADL and Entropy measures with Cosine distance metrics is presented in Fig. 5.

The boxplot figures presented below with the neighborhood parameter selected as FAN and various

Table 3 — Features acquired from typically developing children for various distance metrics [Euclidean, maximum, cosine] with neighborhood selection FAN

Neighborhood Selection: Fixed Amount of Nearest Neighbors									
	Euclidean Distance			Maximum Distance			Cosine Distance		
	TYPICALLY DEVELOPING			TYPICALLY DEVELOPING			TYPICALLY DEVELOPING		
Channels	Determinism	Entropy	Average length	Determinism	Entropy	Average length	Determinism	Entropy	Average length
Fp2-F4	0.6162	0.7951	2.7290	0.6405	0.8479	2.8610	0.4157	0.5284	2.3977
F4-C4	0.6119	0.7840	2.6584	0.652	0.8674	2.8224	0.396	0.5064	2.3796
C4-P4	0.6280	0.7829	2.6847	0.6499	0.8713	2.8485	0.3779	0.4897	2.389
P4-O2	0.5508	0.6734	2.5285	0.5855	0.7509	2.6524	0.3265	0.4205	2.2659
Fp2-F8	0.6374	0.8255	2.7536	0.6749	0.9098	2.9495	0.3796	0.4101	2.3459
F8-T4	0.4808	0.6125	2.4832	0.5285	0.6769	2.5936	0.3524	0.4471	2.2798
T4-T6	0.4433	0.5498	2.4060	0.4869	0.6085	2.4577	0.3466	0.4393	2.2671
T6-O2	0.5144	0.6462	2.4977	0.5489	0.6938	2.5727	0.3294	0.4185	2.2412
Fp1-F3	0.5785	0.7368	2.6493	0.6208	0.8143	2.8071	0.3687	0.4707	2.3259
F3-C3	0.6456	0.7742	2.6394	0.6222	0.8210	2.8248	0.3757	0.4848	2.3691
C3-P3	0.5864	0.7912	2.6917	0.6258	0.8394	2.8173	0.4100	0.5303	2.4443
P3-O1	0.5004	0.6251	2.4604	0.5492	0.6971	2.5894	0.3257	0.4173	2.2532
Fp1-F7	0.6394	0.8310	2.7872	0.6769	0.9164	2.9971	0.3678	0.4713	2.3328
F7-T3	0.5555	0.7015	2.5871	0.5958	0.7737	2.7361	0.3532	0.4492	2.2885
T3-T5	0.5024	0.6247	2.4957	0.5477	0.6885	2.5878	0.3584	0.4559	2.2984
T5-O1	0.4775	0.5908	2.4270	0.5646	0.6512	2.8483	0.3456	0.4398	2.2756

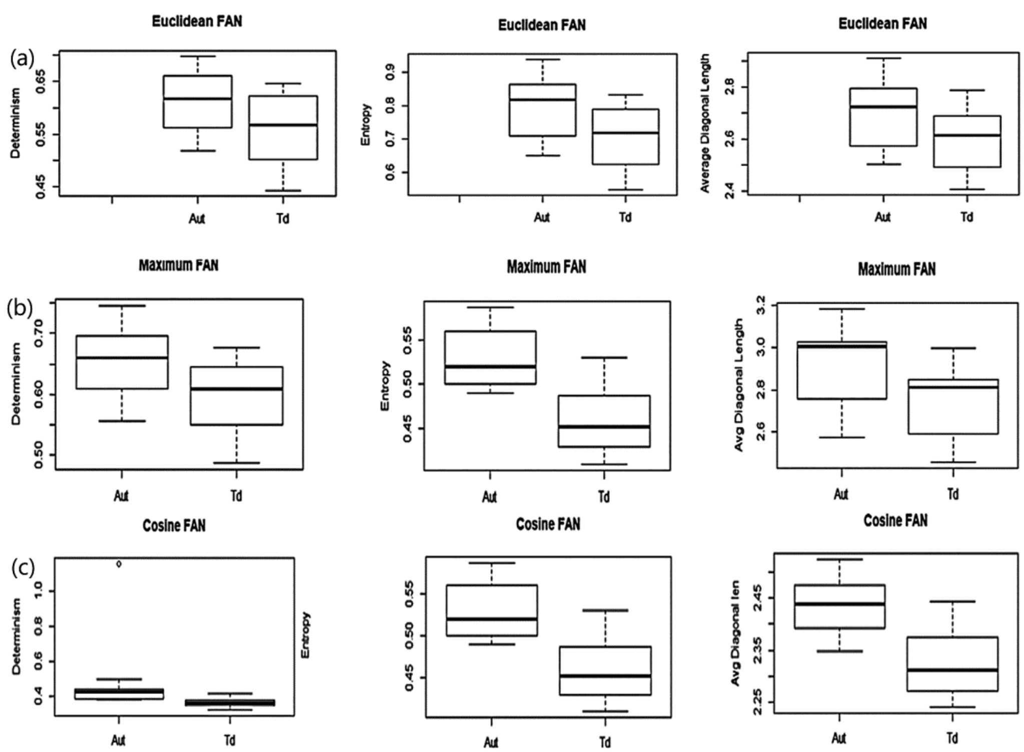

Fig. 5 — Comparison Box plot between ASD and TD for various distance metrics with neighborhood selection FAN

distance metrics such as Euclidean, cosine, and maximum distance.

The determinism, entropy, and average diagonal length of autism and typically developing using Euclidean distance, maximum distance, and cosine distance with FAN are presented in Fig. 5. This cosine distance with FAN combination provides better discrimination between ASD and TD for entropy and average diagonal length. A wide variation in the interquartile range and the median value in the box plot between children with ASD and TD can act as a distinguishing feature in identifying the prominent biomarker. From the results above it is inferred that FAN using cosine distance is best suited for discriminating EEG signal with the RQA feature. In summary, the cosine distance metric performed well in differentiating the different RQA parameters between ASD and TD groups. The responding channels are highlighted in Table 3 and explained in detail later.

Analysis of Threshold Neighborhood Selection and Distance Metrics:

In this section, the performance analysis of threshold neighborhood selection with different measures like RR, Determinism, ADL, and Entropy was analyzed. RQA based measures with these neighborhood techniques are performed and the resultant observations are highlighted in Tables 4 and 5. It is to be noted that, each of this neighborhood selection technique is analyzed with three distance metrics namely Euclidean distance, Maximum distance, and Cosine distance. This provides a better picture of the best combination and responding channels. In this, HTD means channels that have higher feature value in typically developing compared to children with autism whereas HA means channels responding well in children with autism than children with TD. If HTD is marked that means the responding channels in the next column show the channels which have a higher value in TD compared to ASD for the particular feature and combination and vice versa for HA. In Table 5 the feature selection technique using Neighborhood Component Analysis (NCA) is represented. Based on the feature weight, it can be concluded whether the feature is relevant or irrelevant. When the feature weight is closer to zero then the concerned feature is irrelevant. In this table, the Cosine FAN RQA combination was showing good results compared to other combinations.

The diagonal recurrence features such as determinism, entropy, and average diagonal length using neighborhood selection threshold and various distance metrics such as Euclidean, maximum, and cosine are shown in Table 6 for children with ASD. The above features are compared with Table 7 to visualize the discriminating relationship between ASD and TD. The result of the threshold as neighborhood selection is compared with the FAN.

Table 4 — Responding channels for each combination of FAN neighborhood selection

Combination	HTD	HA	Responding channels [Descending order]
FAN, Euclidean, Entropy			P4-O2, T4-T6, F8-T4, C3-P3, FP1-F3, T5-O1, P3-O1, T6-O2, Fp1-F7
FAN, Maximum, Entropy			P4-O2, C3-P3, T5-O1, T6-O2, F3-C3, T4-T6, FP1-F3, P3-O1, F8-T4, T3-T5, F4-C4
FAN, Cosine Entropy			FP2-F8, F7-T3, FP1-F7, P4-O2, T3-T5, T6-O2, P3-O1, F3-C3, FP1-F3, T4-T6, F8-T4, T5-O1, C3-P3
FAN, Euclidean, ADL			C3-P3, P4-O2, FP1-F3, F3-C3, FP1-F7, P3-O1, T5-O1
FAN, maximum, ADL			P3-O1, P4-O2, C3-P3, T6-O2, F4-C4, F3-C3, FP1-F3
FAN, Cosine, ADL			P4-O2, FP2-F8, F8-T4, T6-O2, FP1-F3, F3-C3, C3-P3, FP1-F7, F7-T3

Table 5 — Feature selection using NCA

S. No	RQA combinations	Feature weight		
		DET	ENTR	ADL
1	Euclidean FAN	1.135	4.992	2.363
2	Maximum FAN	0.0581	0.0504	0.1082
3	Cosine FAN	2.107	5.152	3.690

As shown in Table 7, the Features extracted were not as discriminating as seen using FAN and other distance metrics. Cosine distance metric with threshold is showing enough differences when compared to another distance metric. This may be due to the threshold selection for the acquired signal but it was selected based on the standard threshold selection [40% of standard deviation]. In the future with the various threshold value, it may provide a good result. In summary, compared to other distance metric cosine combinations is providing better discriminating results. The box plot of the results is shown below in Fig. 6.

The corresponding discriminating channels and their feature weight are presented in Table 8 and Table 9 below. It is clearly shown in Table 8 that other than ADL remaining features were not relevant for classifying ASD and TD.

The recurrence rate, determinism, entropy, and average diagonal length of autism and typically developing using Cosine distance and threshold are presented in Fig. 6. Compared to Euclidean and maximum distance metric with threshold combination, cosine distance with threshold provides better discriminating results between ASD and TD.

Table 6 — Features acquired from children with autism for various distance metrics [Euclidean, maximum, cosine] with neighborhood selection threshold

	Euclidean Distance				Maximum Distance				Cosine Distance			
	AUTISM				AUTISM				AUTISM			
Channels	Recurrence Rate	Determinism	Entropy	Average length	Recurrence Rate	Determinism	Entropy	Average length	Recurrence Rate	Determinism	Entropy	Average length
Fp2-F4	0.1911	0.998	3.8921	22.7763	0.3405	0.9988	4.2841	39.4013	0.5326	0.9997	4.0866	60.846
F4-C4	0.2139	0.9999	4.0606	25.0275	0.3302	0.9991	4.4461	45.137	0.5191	1	4.3803	92.2562
C4-P4	0.2047	0.9988	3.3561	12.8404	0.3172	0.9984	3.9127	22.0608	0.4859	0.9991	4.0551	37.5488
P4-O2	0.1702	0.9944	2.9977	9.8220	0.2600	0.9946	3.4618	16.1979	0.4877	0.9979	3.8505	31.5566
Fp2-F8	0.2159	0.9991	3.8827	21.8865	0.3442	0.9991	3.3652	38.5330	0.5363	0.9997	4.1951	59.6709
F8-T4	0.1559	0.9723	2.2036	5.6106	0.2271	0.9783	2.5964	8.2961	0.4855	0.9957	3.1986	14.9221
T4-T6	0.1366	0.9487	1.8761	4.5797	0.1994	0.9608	2.2149	6.2427	0.4525	0.9929	3.0436	12.754
T6-O2	0.1684	0.9926	2.7432	7.8957	0.2517	0.9935	3.2268	12.7825	0.4783	0.9967	3.5273	21.8844
Fp1-F3	0.2138	0.9976	3.4634	14.6064	0.3315	0.9977	3.9069	25.3202	0.4992	0.9983	3.7949	35.0872
F3-C3	0.3134	0.9989	3.8827	21.3457	0.4721	0.9989	4.3497	42.0969	0.4908	0.9984	3.9671	52.0228
C3-P3	0.3691	0.9997	4.4616	36.1434	0.5333	0.9997	4.9904	82.2408	0.525	0.9989	4.3078	96.225
P3-O1	0.1360	0.9550	1.9508	4.7964	0.1984	0.9662	2.3012	6.6067	0.4583	0.9944	3.1808	12.9895
Fp1-F7	0.1880	0.9988	3.7094	18.5975	0.2963	0.9988	4.1464	31.7165	0.5029	0.9996	4.2358	49.2634
F7-T3	0.1855	0.9901	2.6469	7.4221	0.2758	0.9922	3.1219	11.8416	0.4714	0.9975	3.4539	18.4805
T3-T5	0.1603	0.9563	1.9937	4.9856	0.2344	0.9659	2.3408	7.0037	0.4762	0.9948	3.192	12.7614
T5 – O1	0.1487	0.9654	2.0963	5.2602	0.2173	0.9741	2.4723	7.4948	0.4734	0.9964	3.2186	16.4287

Table 7 — Features acquired from typically developing children for various distance metrics [Euclidean, maximum, cosine] with neighborhood selection Threshold

	Euclidean Distance				Maximum Distance				Cosine Distance			
	TYPICALLY DEVELOPING				TYPICALLY DEVELOPING				TYPICALLY DEVELOPING			
Channels	Recurrence Rate	Determinism	Entropy	Average length	Recurrence Rate	Determinism	Entropy	Average length	Recurrence Rate	Determinism	Entropy	Average length
Fp2-F4	0.2722	0.9984	3.6554	18.5665	0.4053	0.9980	3.9341	35.8593	0.4668	0.9957	3.1871	17.0675
F4-C4	0.1886	0.9983	3.4251	13.7407	0.2935	0.9978	3.9309	23.7979	0.4945	0.9982	3.7842	26.7289
C4-P4	0.1657	0.9959	2.8085	8.0939	0.2505	0.9971	3.3392	12.9106	0.4799	0.9983	3.8373	22.263
P4-O2	0.1755	0.9895	2.7197	8.0107	0.2563	0.9904	3.1654	13.1453	0.4543	0.9916	3.1708	14.423
Fp2-F8	0.2333	0.9984	3.6282	19.1816	0.3546	0.9976	3.8748	35.6903	0.4824	0.9983	3.2251	28.6786
F8-T4	0.2404	0.9965	3.3416	13.2642	0.3542	0.9964	3.8043	24.7069	0.4647	0.9922	3.1071	18.2625
T4-T6	0.1457	0.9650	2.0802	5.1829	0.2113	0.9751	2.4675	7.3177	0.4604	0.9943	3.219	15.2706
T6-O2	0.1809	0.9899	2.7831	8.4945	0.2645	0.9907	3.2164	14.1086	0.4438	0.9876	2.921	12.0352
Fp1-F3	0.2691	0.9982	3.6782	17.9633	0.4105	0.9981	4.0865	33.7845	0.4732	0.9966	3.4065	19.8862
F3-C3	0.1763	0.9956	2.8564	8.4689	0.2656	0.9960	3.3668	13.7720	0.4707	0.9974	3.6455	19.5143
C3-P3	0.1626	0.9938	2.7017	7.5158	0.2432	0.9959	3.2207	11.8439	0.477	0.9981	3.6314	17.598
P3-O1	0.1732	0.9774	2.3465	6.2282	0.2505	0.9830	2.7672	9.4542	0.4444	0.9875	2.9587	12.1599
Fp1-F7	0.2458	0.9977	3.6138	19.2698	0.3690	0.9978	3.9175	36.0333	0.4767	0.9962	3.3453	19.9735
F7-T3	0.2090	0.9942	3.2106	13.1292	0.3151	0.9948	3.4761	22.5466	0.4631	0.9971	3.1929	14.82
T3-T5	0.1410	0.9521	1.9566	4.9084	0.2065	0.9645	2.3004	6.7734	0.4473	0.9907	2.8365	10.0788
T5 – O1	0.1501	0.9551	1.978	4.9340	0.2179	0.9677	2.344	6.8578	0.4315	0.9843	2.6999	9.3536

The changes with three distance metrics with FAN and threshold are depicted in Table 4 and Table 9. Both the features are presenting good discrimination with cosine metrics between ASD and TD. The discriminating channels for the presented stimuli are segregated as a table based on the percentage difference is shown in Table 4 and Table 9 above.

Table 8 — Feature selection using NCA

S. No	RQA combinations	Feature weight			
		RR	DET	ENTR	ADL
1	Euclidean FAN	0.0012	0.0052	0.0002	0.7904
2	Maximum FAN	0	0.0015	0.0026	0.8523
3	Cosine FAN	0	0.0005	0.0021	0.9485

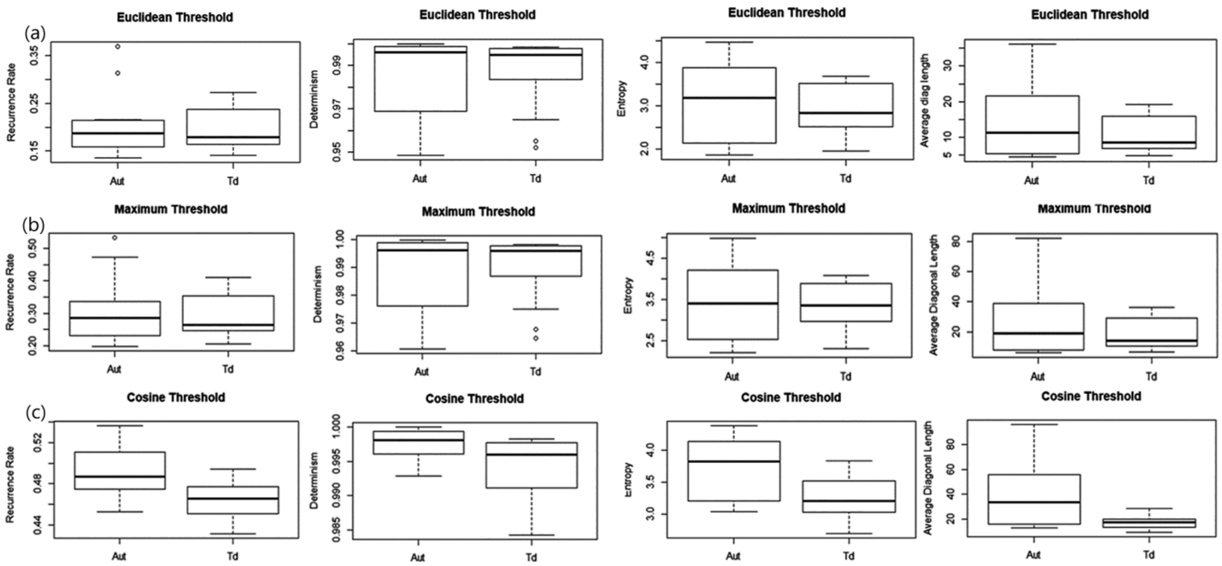

Fig. 6 — Comparison Box plot between ASD and TD for various distance metrics [Euclidean, Maximum and cosine] with neighborhood selection threshold method

Table 9 — Responding channels for each combination of threshold neighborhood selection

Combination	HTD	HA	Responding channels [Descending order]
Threshold, Euclidean, Rec Rate			F8-T4, FP2-F4, FP1-F7, P3-O1, FP1-F7, FP1-F3, FP2-F4
Threshold, Euclidean, Rec Rate			C3-P3, F3-C3, C4-P4, F4-C4, T3-T5
Threshold, Maximum, Rec Rate			F8-T4, P3-O1, FP1-F7, FP1-F3, FP2-F4, F7-T3
Threshold, Maximum, Rec Rate			C3-P3, F3-C3, C4-P4, T3-T5, F4-C4
Threshold, Cosine, Rec Rate			FP2-F4, FP2-F8, C3-P3, T5-O1, P4-O2, T6-O2, T3-T5, FP1-F3, FP1-F7
Threshold, Euclidean, entropy			F8 -T4, P3-O1, F7-T3, T4-T6
Threshold, Euclidean, entropy			C3-P3, F3-C3, C4-P4, F4-C4, P4-O2
Threshold, Maximum, entropy			F8-T4, FP2-F8, T4-T6, P3-O1, F7-T3
Threshold, Maximum, entropy			C3-P3, F3-C3, C4-P4, F4-C4
Threshold, Cosine, entropy			T4-T6, F8-T4, P3-O1
Threshold, Cosine, entropy			FP2-F8, FP2 - F4, FP1-F7, P4-O2, T6-O2, T5-O1, F4-C4, T3-T5
Threshold, Euclidean, ADL			F8-T4, F7-T3, P3-O1, FP1-F3
Threshold, Euclidean, ADL			C3-P3, F3-C3, F4-C4, C4-P4, P4-O2, FP2-F4
Threshold, Maximum, ADL			F8-T4, T4-T6, FP1-F3 P3-O1, F7-T3
Threshold, Maximum, ADL			C3-P3, F3-C3, F4-C4, C4-P4, P4-O2
Threshold, Cosine, ADL			F8-T4, T4-T6, F7-T3
Threshold, Cosine, ADL			C3-P3, F3-C3, P4-O2, FP2-F4, F4-C4, C4-P4, F3-C3, FP1-F7

This provides a better picture of the best combination and responding channels. In this research, it is concluded that the cosine distance metric is well suited for this EEG analysis. This may be due to cosine distance judgment capability of orientation and not magnitude and also its cohesion within clusters.

Conclusions

Autism is a heterogeneous disorder that affects 1–2% of the population. In this work, variation of EEG signal between children with ASD and children with TD are observed. Different parameters such as Fixed amount of nearest neighbor, threshold, and different distance metrics such as Euclidian distance, maximum distance, and cosine distance are used to obtain the features such as recurrence rate, determinism, entropy, and average diagonal length. Each combination provides the best result in some of the electrode regions for certain features. There are no wide changes visualized in determinism using any combination. The discussion for other features with various combinations was discussed. Among these cosine distance provides a better result in combination with FAN as well as a threshold with wider variations. It is observed that the values for all the four features were higher for the child with ASD compared to the typically developed child with FAN selection. This may be due to the repetitive behavior of the ASD child. In the combination of threshold method as neighborhood selection, there appear some channels showing higher value for TD whereas some channels show higher value in TD for various distance metrics. The EEG channels responding to the presented audio/visual stimuli are FP1, FP2, F4, F8, T4, T6, O1, O2, C4, and P4. This study could be used well for interventional studies and the level of improvement in training, as each lobe has pertained to a particular function. If there appears a lower value in the particular channel about a specific lobe, training for the autistic child can be changed based on the lag to improve the particular lobe function. The improvement of that particular lobe can be seen through these feature values.

In the future, the transition in the recurrence plot has to be analyzed closely. The study has to be applied to a larger sample for validation. The vertical line features must be analyzed in the future with the euclidean and cosine distance metric as it provides the better result with the diagonal line features of RQA.

Acknowledgment

The authors acknowledge DST/TIDE for the sponsored research project [SEED/TIDE/092/2016].

References

1 Elsabbagh M, Divan G, Koh Y-J, Kim Y S, Kauchali S, Marcín C, Montiel-Nava C, Patel V, Paula C S, Wang C, Yasamy M T & Fombonne E, Global Prevalence of Autism and Other Pervasive Developmental Disorders, *Autism Res*, **5** (2012) 160–179.

2 Hossain M D, Ahmed H U, Uddin M M J, Chowdhury W A, Iqbal M S, Kabir R I, Chowdhury I A, Aftab A, Datta P G, Rabbani G, Hossain S W & Sarker M, Autism Spectrum disorders (ASD) in South Asia: a systematic review, *BMC Psychiatry*, **17** (2017) 1–41.

3 Sharma S R, Gonda X & Tarazi F I, Autism spectrum disorder: Classification, diagnosis and therapy, *Pharmacol Ther*, **190** (2018) 91–104.

4 Zwaigenbaum L & Penner M, Autism spectrum disorder: advances in diagnosis and evaluation, *Bmj*, **361** (2018) 1–15.

5 Webber C L & Marwan N, *Recurrence Quantification Analysis Theory and Best Practices* (Springer, Switzerland) 2015, 1–41.

6 Tripathy R & Acharya U R, Use of features from RR-time series and EEG signals for automated classification of sleep stages in deep neural network framework, *Biocybern Biomed Eng*, **38 (**2018) 890–902.

7 Pierece K, Marinero S, Hazing R, Mckenna B, Carter L, Banees & Malisa A, Eye tracking reveals abnormal visual preference for geometric images as an early biomarker of autism spectrum disorder subtype associated with increased symptom severity, *Biol Psychiatry*, **79** (2016) 657–666.

8 Bosl W J, Loddenkemper T & Nelson C A, Nonlinear EEG biomarker profiles for autism and absence epilepsy, *Neuropsychiatr Electrophysiol*, **3** (2017) 1–22.

9 Yuen C T, San W S, Rizon M & Seong T C, Classification of Human Emotions from EEG Signals using statistical features and neural network, *Int J Integr Eng*, **1** (2009) 1–6.

10 Fan J, Wade J W, Key A P, Warren Z E & Sarkar N, EEG-based affect and workload recognition in a virtual driving environment for ASD intervention, *IEEE Trans Biomed Eng*, **65** (2018) 43–51.

11 Siddiqui M M, Srivastava G & Saeed S H, Diagnosis of Insomnia sleep disorder using short time frequency analysisof PSD approach applied onEEG signal using ROC-LOC, *Sleep Sci*, **9** (2019) 186–191.

12 Padfield N, Zabalza J, Zhao H, Masero V & Ren J, EEG-based brain-computer interfaces usingmotor-imagery: techniques and challenges, *Sensors*, **19 (**2019) 1–34.

13 Goldfine A M, Victor J D, Conte1 M M, Bardin J C & Schiff N D, Determination of awareness in patients with severe brain injury using EEG power spectral analysis, *Clin Neurophysiol*, **122** (2011) 2157–2168.

14 Akbarian B & Erfanian A, Automatic seizure detection based on nonlinear dynamical analysis of EEG signals and mutual information, *Basic Clin Neurosci*, **9** (2018) 227–240.

15 Riley M A & Van O G C, *Tutorials in contemporary nonlinear methods for the behavioral sciences* (National Science Foundation Newyork) 2005, 353–400.

16 Kannathal N, Acharya U R, Lim C & Sadasivan P, Characterization of EEG—A comparative study, *Comput Meth Prog Bio*, **80** (2005) 17–23.

17 Li X, Ouyang G, Yao X & Guan X, Dynamical characteristics of pre-epileptic seizures in rats with recurrence quantification analysis, *Physics Letters A*, **333** (2005) 164–171.

18 Coco M I & Dale R, Cross-recurrence quantification analysis of categorical and continuous time series: an R package, *Front Psychol*, **5** (2013) 1–15.

19 Abdossalehi M & Nasrabadi A M, Evaluating the determinism of brain signals using recurrence chaotic features in positive, negative and neutral emotional states in the sources achieved from ICA algorithm, *ICNSJ* **4** (2017) 63–71.

20 Liang Z, Wang Y, Ren Y, Li D, Voss L, Sleigh J & Li X, Detection of burst suppression patterns in EEG using recurrence rate, *Sci*, **1** (2014) 1–11.

21 Eckmann J P, Kamphorst S O & Ruelle D, Recurrence plots of dynamical systems, *Europhys Lett*, **4** (1987) 973–977.

22 Bosl W J, T-Flusberg H & Nelson C A, EEG analytics for early detection of autism spectrum disorder: A data-driven approach, *Sci Rep*, **8** (2018), 1–20.

23 Ngamgaa E J, Bialonskib S, Marwana N, Kurthsa J, Geiere C & Lehnertze K, Evaluation of selected recurrence measures in discriminating pre-ictal and inter-ictal periods from epileptic EEG data, *Physics Letters A*, **16** (2016) 1419–1425.

24 Al-Suify M, Al-Atabany W & Eldosoky M A A, Classification of right and left hand movement using phase space and recurrence quantification analysis, *NRSC*, **35** (2018) 457–464.

25 Gruszczyńska I, Romuald M, Sobaniec P, Ż-Sobaniec M & Borowska M, Epilepsy identification based on EEG signal using RQA method, *Int J Adv Med Sci*, **64** (2019) 58–64.

26 Rooij M J W V & Wallot S, *Qualitative Psychology Nexus* (Die Deutsche Bibliothek, Germany) 2015, 69–87.

27 Anderson N N, Bischof W F & Kingstone A, Recurrence Quantification Analysis, (2014), from https://webdocs.cs.ualberta.ca/~wfb/varia/RQA-Introduction.pdf.

28 Iwanski J S & Bradley E, Recurrence plots of experimental data: To embed or not to embed?, *Chaos*, **8** (1998) 861–871.

29 Cao L, Practical method for determining the minimum embedding dimension of a scalar time series, *Physica D*, **110** (1997) 43–50.