Biomedical Signal Processing and Control 59 (2020) 101921

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

Seizures classification based on higher order statistics and deep neural network

Rahul Sharma [a,*], Ram Bilas Pachori [b], Pradip Sircar [a]

[a] Department of Electrical Engineering, Indian Institute of Technology Kanpur, Kanpur, UP 208016, India
[b] Discipline of Electrical Engineering, Indian Institute of Technology Indore, Indore, MP 453552, India

ARTICLE INFO

Article history:
Received 7 October 2019
Received in revised form 14 February 2020
Accepted 26 February 2020

Keywords:
EEG
Seizure
Higher order statistics
Autoencoder
Deep neural network

ABSTRACT

The epileptic seizure is a transient and abnormal discharge of nerve cells in the brain that leads to a chronic disease of brain dysfunction. There are various features-based seizures classification algorithms listed in the literature. But, there is no standardized set of attributes that can perfectly capture the relevant information regarding the signal dynamics. In this paper, a computationally-fast seizure classification algorithm is presented. The obtained results through the proposed algorithm are consistent and repeatable. This paper describes an automated seizures classification technique using the nonlinear higher-order statistics and deep neural network algorithms. The sparse autoencoder based deep neural network is used to extract the essential structural details from the third-order cumulant coefficients matrix. The proposed algorithm achieves a reliable classification accuracy for both categories, i.e., binary classes and three-classes of electroencephalogram (EEG) signals with the softmax classifier. The proposed study is simulated on the publicly-available Bonn university EEG database. The achieved results show the effectiveness of the proposed algorithm for seizures classification.

1. Introduction

The human brain is the cluster of billions of nerve cells or neurons. These neurons are connected through a complex network that controls most of the brain activities. The synchronized firing of the neurons causes involuntary and temporary disturbance of regular neural activity, which can lead to seizures [1]. The recurrent unprovoked seizures are known as epilepsy. Despite the fact that the seizure is the predominant symptom of epilepsy, experiencing a seizure does not necessarily mean that a person has epilepsy. There are many causes of seizure, depending on where the disturbance first started and how far it spreads inside the brain [2,3]. The World Health Organization (WHO) report states that more than 1% of humans are affected by epilepsy every year, and out of them, nearly 90 percent of people belong to developing countries [4].

Despite the rapid advances of neuroimaging techniques such as the positron emission tomography (PET), magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), computerized tomography (CT), etc., the electroencephalogram (EEG) is the most commonly used tool for monitoring the brain activities associated with seizures. It provides an excellent temporal resolution that analyzes large-scale brain activities for short duration and vice-versa. The visual analysis of long-term EEG signal is very error-prone and time-consuming, and it also may lead to false key-point detection. Therefore, various feature dependent algorithms are proposed to capture the underlying non-linearity and complexity of EEG signals.

In this paper, an automated seizures classification algorithm is presented that utilizes the concepts of the higher-order statistics (HOS) and the deep neural network. The third-order cumulant (ToC) explores the nonlinear variations of studied signals into two-dimensional space. This higher-space representation preserves the original as well as the harmonic information of the EEG signals. The proposed algorithm is a data-dependent algorithm that automatically learns time-varying signal characteristics from data itself. A sparse autoencoder is used to extract the essential structural details from the ToC coefficients matrix.

The layout of the article is as follows. Section 2 contains the literature review of the seizures classification algorithms based on EEG signals. A brief overview of the database is given in Section 3. The systematic description of the proposed methodology, which involves formulations of HOS and autoencoder, is presented in Section 4. The results and discussion have been provided in Section 5, whereas Section 6 includes the conclusion of this work.

* Corresponding author.
E-mail addresses: rahuls@iitk.ac.in (R. Sharma), pachori@iiti.ac.in (R.B. Pachori), sircar@iitk.ac.in (P. Sircar).

https://doi.org/10.1016/j.bspc.2020.101921
1746-8094/

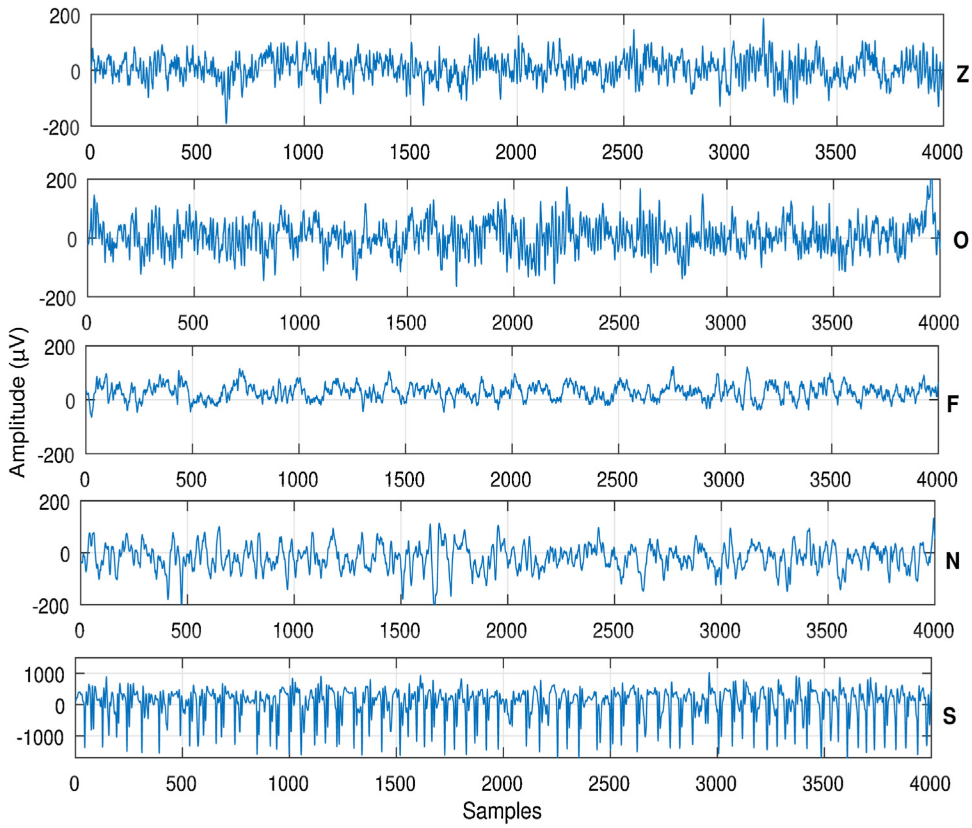

Fig. 1. EEG signals from different sets of Bonn university dataset (Z, O, F, N, and S).

2. Related works

The literature reveals that numerous methods such as linear prediction [5,6], autoregression (AR) model [7–9], spectral estimation [10–12], etc., are extensively implemented for seizures classification based on labeled EEG signals. The EEG signals exhibit non-stationary and nonlinear behavior that do not adequately analyze by the above-defined algorithms. Hence, various time-frequency (TF) methods are proposed to analyze non-stationary EEG signals [13–16]. The TF methods decompose the studied signal on time and frequency axes simultaneously that provides the information regarding the energy distribution over different frequencies present in the signals.

The wavelet transform is frequently used to analyze the time-variation characteristics of EEG signals. It provides a scale-time variant decomposition of a non-stationary signal that makes it feasible to capture transient behavior of the signal. Many researchers have proposed several variants of wavelet transform such as the discrete wavelet transform (DWT) [17–23], mixed band wavelet transform [24], multiwavelet transform [25], dual-tree complex wavelet transformation (DT-CWT) [26–29], wavelet packet entropy [30,31], empirical wavelet transform (EWT) [32], analytic TF flexible analysis wavelet transform (ATFAWT) [33], tunable Q-wavelet transform (TQWT) [34–36], and more, for automated seizures classification.

Apart from the wavelet transform, the empirical mode decomposition (EMD) is another nonlinear algorithm that is used for seizures classification [37–42]. It is a data-dependent TF decomposition approach that decomposes the studied signal into sub-signals known as the intrinsic mode functions (IMFs). However, the low-energy components of nonlinear signals are not precisely extracted by the EMD, and this leads to the loss of the local information. Contrary to the EMD, the variational mode decomposition (VMD) algorithm can extract the relevant low-energy components information from the non-stationary signals [43,44]. It decomposes the studied signal into a fixed number of band-limited IMFs (BLIMFs) for exploring the underlying non-linearity and non-stationarity present in the EEG signals. For identifying seizures classes, various features are measured from the BLIMFs that are classified with different classifiers.

It can be noticed that the general procedure of the above-defined algorithms is to decompose the EEG signal to various stages, and different attributes are measured from each level to attain better classification accuracy. There is no standardized set of attributes that can perfectly represent the signal dynamics. Moreover, the irrelevant and redundant attributes increase the feature space that degrades the accuracy of the method, also increase the overfitting risk. Hence, the existing seizures classification methods give rise to computational complexity and features sustainability issues.

The authors used various machine learning algorithms for automated seizures classification based on EEG signals [45–47]. These learning algorithms not only reduce the feature-dependency problem but also reduce the computational complexity by introducing different network architecture. Acharya et al. [45] proposed a 13-layer deep convolutional neural network (CNN) algorithm, while Zhou et al. [46] introduced a three-layer CNN to classify ictal, preictal, and seizure classes. In this sequel, Husseni et al. [47] used long short-term memory (LSTM) based recurrent neural network (RNN) for seizures classification. By introducing more number of layers in the network, the model is well trained that may give better prediction accuracy. But, on other hand it increases the computational complexity due to overfitting and raises vanishing gradient problem.

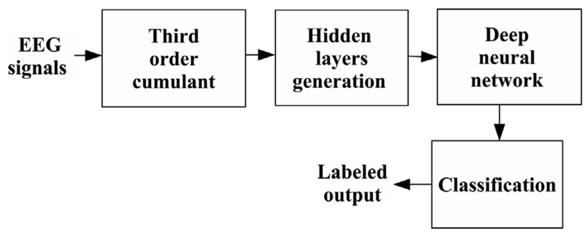

Fig. 2. Block diagram of proposed algorithm.

3. Dataset

The proposed algorithm is simulated on the publicly-available Bonn university EEG signals database [48], which includes recordings of normal, seizure, and seizure-free cases. It consists of five sets, namely, Z, O, F, N, and S, with each set containing 100 single-channel EEG signals of duration 23.6 s. These EEG signals are selected from the pool of 128 EEG signals after visual inspection by the neuroexperts. These signals are recorded at a 173.61 Hz sampling rate using 12-bit resolution and a band-pass filter of 0.53–40 Hz is used to remove the artifacts caused by muscle activities.

The EEG signal sets Z and O (normal) are recorded extracranially from five healthy persons during the relaxed state with eyes closed (Set Z) and eyes open (Set O). However, the remaining sets F, N, and S, are recorded intracranially from five epileptic patients, who had achieved complete seizure control after resection of one of the hippocampal formations, which was therefore correctly diagnosed to be the seizure generating area.

Sets F and N (interictal) contained only activities measured during seizure-free intervals. N is a seizures-free signal set which is recorded from the hippocampal formation of the opposite hemisphere of the brain while seizure-free signals set F is recorded from the epileptogenic zone. The signals set S (ictal) only contained seizure activity. Fig. 1 depicts the EEG signals from different sets. In this study, the four cases have been considered for two binary, a three, and a five-classes EEG signals classification, which are as follows:

Case-1: Z – S
Case-2: ZOFN – S
Case-3: ZO – FN – S
Case-4: Z – O – F – N – S

4. Proposed methodology

The proposed algorithm is a data-dependent algorithm that reduces computational complexity and features sustainability. Instead of several stages of signal analysis, only a single stage computation of the ToC is sufficient to capture the underlying nonlinear signal dynamics, and that reduces the computational complexity of the algorithm. Fig. 2 shows the block diagram of the proposed algorithm.

The existing deep neural network architectures have many limitations such as vanishing gradient, overfitting, computational time, etc., which restrict the network performance [49,50]. These issues are mainly generated when the parameters (weights and biases) are not learned adequately by the back-propagation algorithm [51]. These problems can be easily suppressed by changing the network architecture.

In place of the conventional neural network, an autoencoder generated two hidden layers neural network is proposed to classify seizures classes. The autoencoder removes the overfitting problem as it trains some of the randomly selected nodes rather than the entire network. Also, it reduces the dimension of the analysis data that allows the loss function error can reach till deeper nodes which can solve the vanishing gradient problem [52]. Before describing the proposed algorithm, first, we review the properties of the ToC and operating principles of the autoencoder and softmax classifier.

4.1. Third order cumulant

In literature, the HOS is frequently used in various signal processing applications, such as EEG signal classification [53,54], transient signal analysis [55], focal EEG signal detection [56], automated glaucoma detection [57], human emotion recognition [58], etc. The ToC is used to explore the nonlinear variations of studied signals into higher dimensional space, which not only preserves the original signal information but also consists of harmonic information at higher dimension.

Let (Y_j) be a zero-mean stationary discrete-time random process. The ith order cumulant can be defined as follows:

$$\begin{aligned} C_Y^i[\ell_1, \ell_2, \ldots, \ell_{i-1}] &= M_{i,Y}[\ell_1, \ell_2, \ldots, \ell_{i-1}] \\ &\quad - m_{i,G}[\ell_1, \ell_2, \ldots, \ell_{i-1}] \end{aligned} \tag{1}$$

where $M_{i,Y}$ is the ith order moment function of random process (Y_j) and $m_{i,G}$ is the ith order moment function of an equivalent Gaussian random process. Note that the ith order cumulant is a function of $(i-1)$ lags $\ell_1, \ell_2, \ldots, \ell_{i-1}$ for a random stationary process whereas for a non-stationary random process, it will be a function of j as well as $(i-1)$ lags. The number of lags determines the order of complexity of the analysis [59]. For orders $i = 1, 2,$ and 3, the cumulant can be defined as follows: First order cumulant: $C_Y^1 = M_{1,Y} = \mathcal{E}\{(Y_j)\}$

Second order cumulant:

$$\begin{aligned} C_Y^2[\ell_1] &= M_{2,Y}[\ell_1] - M_{1,Y}^2 \\ &= M_{2,Y}[-\ell_1] - M_{1,Y}^2 = C_Y^2[-\ell_1] \end{aligned} \tag{2}$$

ToC:

$$\begin{aligned} C_Y^3[\ell_1, \ell_2] &= M_{3,Y}[\ell_1, \ell_2] \\ &\quad - M_{1,Y}\left\{M_{2,Y}[\ell_1] + M_{2,Y}[\ell_2] + M_{2,Y}[\ell_1 - \ell_2]\right\} + 2M_{1,Y}^3 \end{aligned} \tag{3}$$

where $\mathcal{E}$ is the expectation operator. The HOS have some unique properties such as [60],

- The higher-order (three or more) cumulants of the Gaussian distributed process are equal to zero, that means $C_Y^i[\ell_1, \ell_2, \ldots, \ell_{i-1}] = 0$ for $i \geq 3$. Hence, the cumulants not only provide the amount of higher-order correlation but also measure the distance of the random process from the Gaussian process.
- The ToC is infinitely deferential and convex, which allow to analyzing the non-minimum phase and phase couple signals.
- The ToC can recognize Gaussian/non-Gaussian signals, linear/nonlinear systems, phase coupling, etc.

The ToC plots (contour and mesh plots) of signals from each set (Z, O, F, N, and S) can be visualized in Fig. 3. It can be noticed that the ToC has various symmetries. The analysis space contains lots of redundant and repeated information that can be defined as,

$$\begin{aligned} &C(y[\lambda], y[\lambda+\ell_1], \ldots, y[\lambda+\ell_{i-1}]) \\ &= C(y[\lambda+\ell_1], y[\lambda], \ldots, y[\lambda+\ell_{i-1}]) \\ &= C(y[\lambda+\ell_{i-1}], \ldots, y[\lambda+\ell_1], y[\lambda]) \end{aligned} \tag{4}$$

In this study, the lower oblique attributes are only considered for further analysis. It contains all the information about the distribution of ToC, and it is computationally efficient.

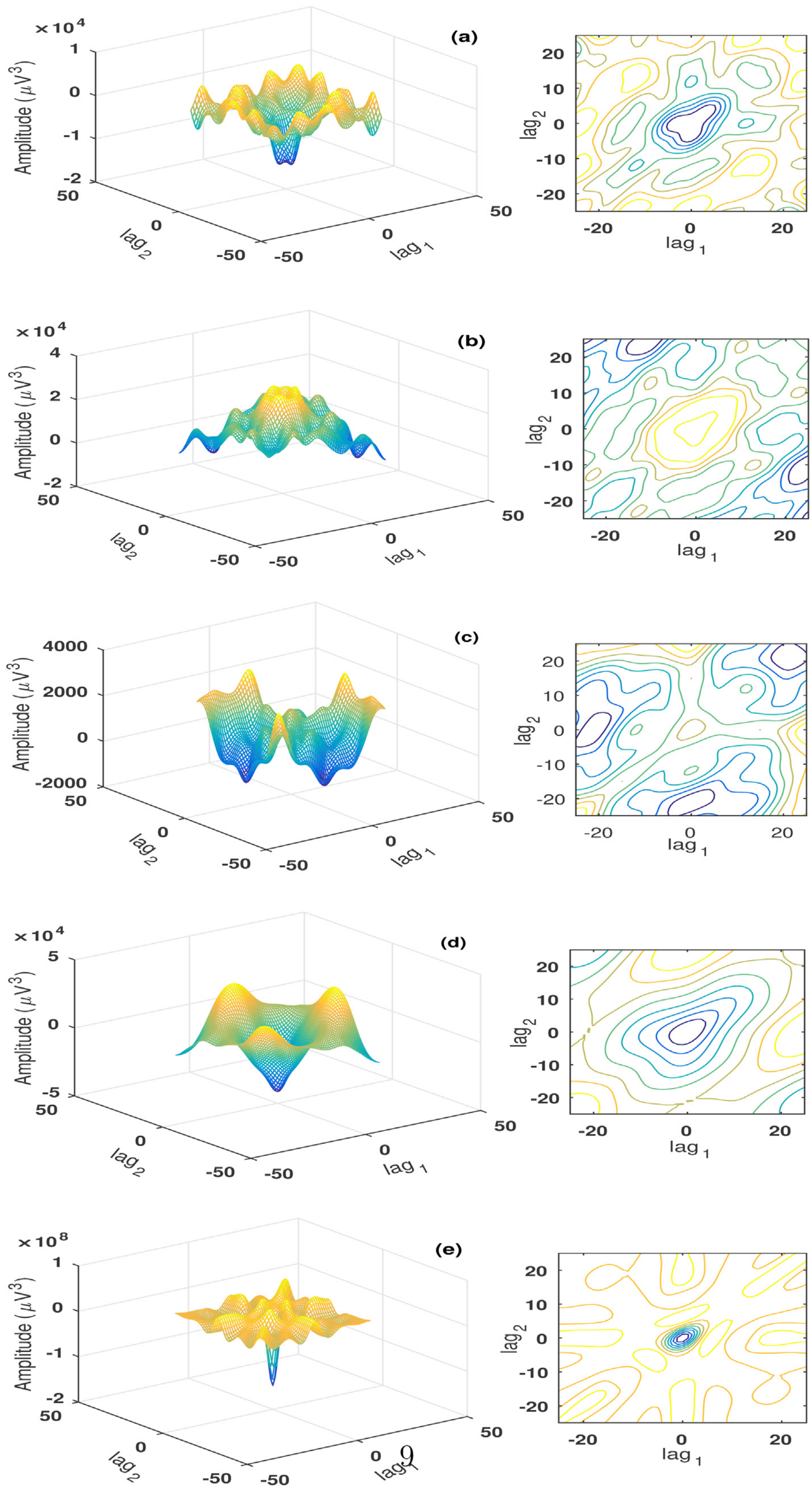

Fig. 3. ToC plots (mesh and contour plots) of EEG signals from each set (Z, O, F, N, and S).

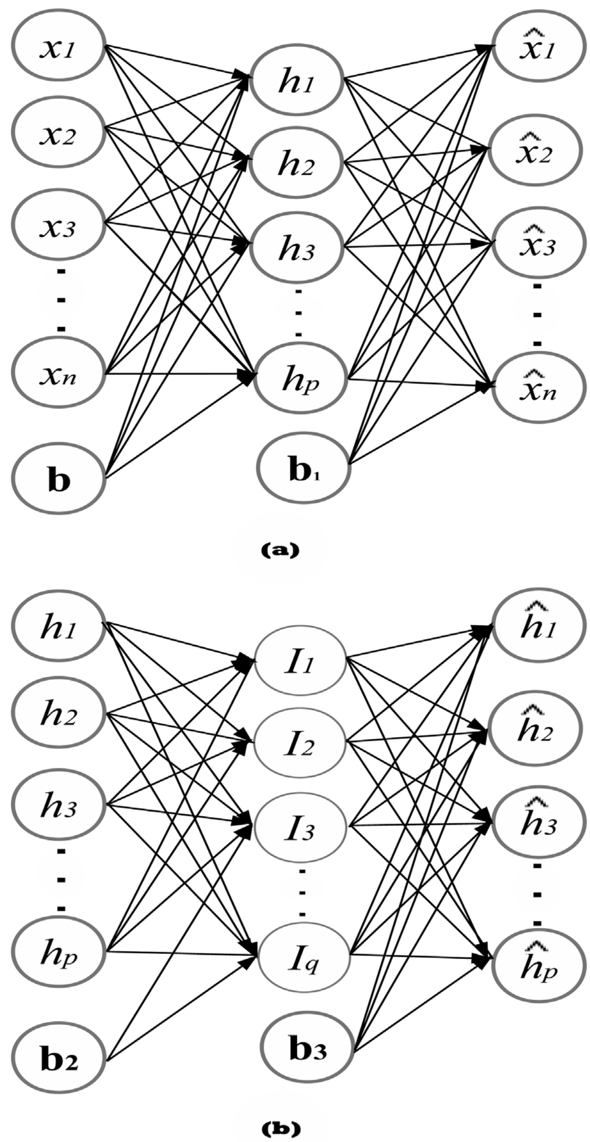

Fig. 4. Single layer autoencoder: (a) first hidden layer information $\mathbf{h} \in \Re^{p}$ is generated from the ToC coefficients $\mathbf{x} \in \Re^{n}$, (b) second hidden layer information $\mathbf{I} \in \Re^{q}$ is measured from the data of the first hidden layer $\mathbf{h} \in \Re^{p}$.

4.2. Autoencoder

The autoencoder is an unsupervised neural network that is trained using supervised learning methods. It recreates the input at the output terminal by imposing constraints on the network parameters such as weight matrix and bias vector [61]. The autoencoder has two components, an encoder, and a decoder. Each of these has separate element-wise activation functions ψ and φ, respectively. For a single hidden layer, the encoder process any input random vector, $\mathbf{x} \in \Re^{n}$ and maps it to the hidden layer $\mathbf{h} \in \Re^{p}$, as,

$$\mathbf{h} = \psi(\mathbf{W}\cdot\mathbf{x} + \mathbf{b}) \tag{5}$$

where $\mathbf{h}$ is the vector of latent variables. $\mathbf{W} \in \Re^{p\times n}$ and $\mathbf{b}_{p\times 1}$ are randomly initialized weight matrix and bias vector, respectively. These random parameters are updated iteratively through the backpropagation algorithm. Similarly, the decoder maps $\mathbf{h} \in \Re^{p}$ to the approximate input vector $\hat{\mathbf{x}} \in \Re^{n}$, i.e.

$$\hat{\mathbf{x}} = \varphi(\mathbf{W}'\cdot\mathbf{h} + \mathbf{b}') \tag{6}$$

where $\mathbf{W}' \in \Re^{n\times p}$ and $\mathbf{b}'_{n\times 1}$ are weight matrix and bias vector respectively, at decoder end. Hence, the cost error function $\phi(\mathbf{e})$ of an autoencoder network is given as [62]

$$\begin{aligned} \phi(\mathbf{e}) &= \frac{1}{2}\|\mathbf{x} - \hat{\mathbf{x}}\|^2 + \Omega(\mathbf{h}, \mathbf{x}) \\ &= \frac{1}{2}\left\|\mathbf{x} - \varphi\left(\mathbf{W}'\left(\psi\left(\mathbf{W}\cdot\mathbf{x} + \mathbf{b}\right)\right) + \mathbf{b}'\right)\right\|^2 + \Omega(\mathbf{h}, \mathbf{x}) \end{aligned} \tag{7}$$

where the first term is the loss function, the second term reflects regularization, i.e., the sparsity penalty term with the weight decay parameter λ which can be defined as follows:

$$\Omega(\mathbf{h}, \mathbf{x}) = \lambda \sum_{i=1}^{p} \|\nabla_{x_i} \cdot h_i\|^2 \tag{8}$$

where ∇_x is the difference operator defined over ith node random variable x. The regularization term fits the model to learn a function that does not change much even when $\mathbf{x}$ is slightly changed [63,64].

In the proposed algorithm, two hidden layers deep neural network is used for seizures-labeled EEG signals classification. Fig. 4 shows the architecture of two different hidden-layers that are generated from the autoencoder network. The first hidden layer information $\mathbf{h} \in \Re^{p}$ is generated from the lower oblique of the ToC coefficients $\mathbf{x} \in \Re^{n}$, while the second hidden layer information $\mathbf{I} \in \Re^{q}$ is measured from the data of the first hidden layer $\mathbf{h} \in \Re^{p}$, by minimizing their respective cost error functions. The analogy of hidden layers data generation is shown by Fig. 4(a) and (b), respectively.

The Kruskal–Wallis test has been performed to analyze the significance of the hidden layer information [65]. Fig. 5 shows the boxplots of each hidden layer output for all four considered cases, along with its p-value. The p-value illustrates the standard significance of the measured features, which is lower the p-value higher the significance. It can be noticed that the p-values of the first hidden layer are less. This indicates that the hidden layer may not able to extract significant structural information from the ToC coefficients. Hence, to refine the features space, the other set of information $\mathbf{I} \in \Re^{q}$ is extracted from the hidden layer data.

It can be observed that the p-values for the second hidden layer data are much lesser than the p-values for the first hidden layer data. Therefore, the attributes of the second hidden layer having much more significance than the first hidden layer attribute; also, it reduces the dimension of the studied data $n > p > q$. This justifies the proposed architecture of the deep neural network. If we further generate data from second layer information, the p-value becomes very high that data is not practically acceptable.

The entire seizures-labeled ToC coefficients data is divided as 75% for training while remaining 25% for testing. The combined network is trained on, rather than merely acting as an identity function, and the softmax classifier is used to classify the unknown EEG signal class.

4.3. Softmax classifier

The proposed classification algorithm is a multiclass classification algorithm to classify binary and three-categories seizure classes. A softmax classifier is used as a last layer of the proposed network. It gives the actual probabilities score for each labeled class

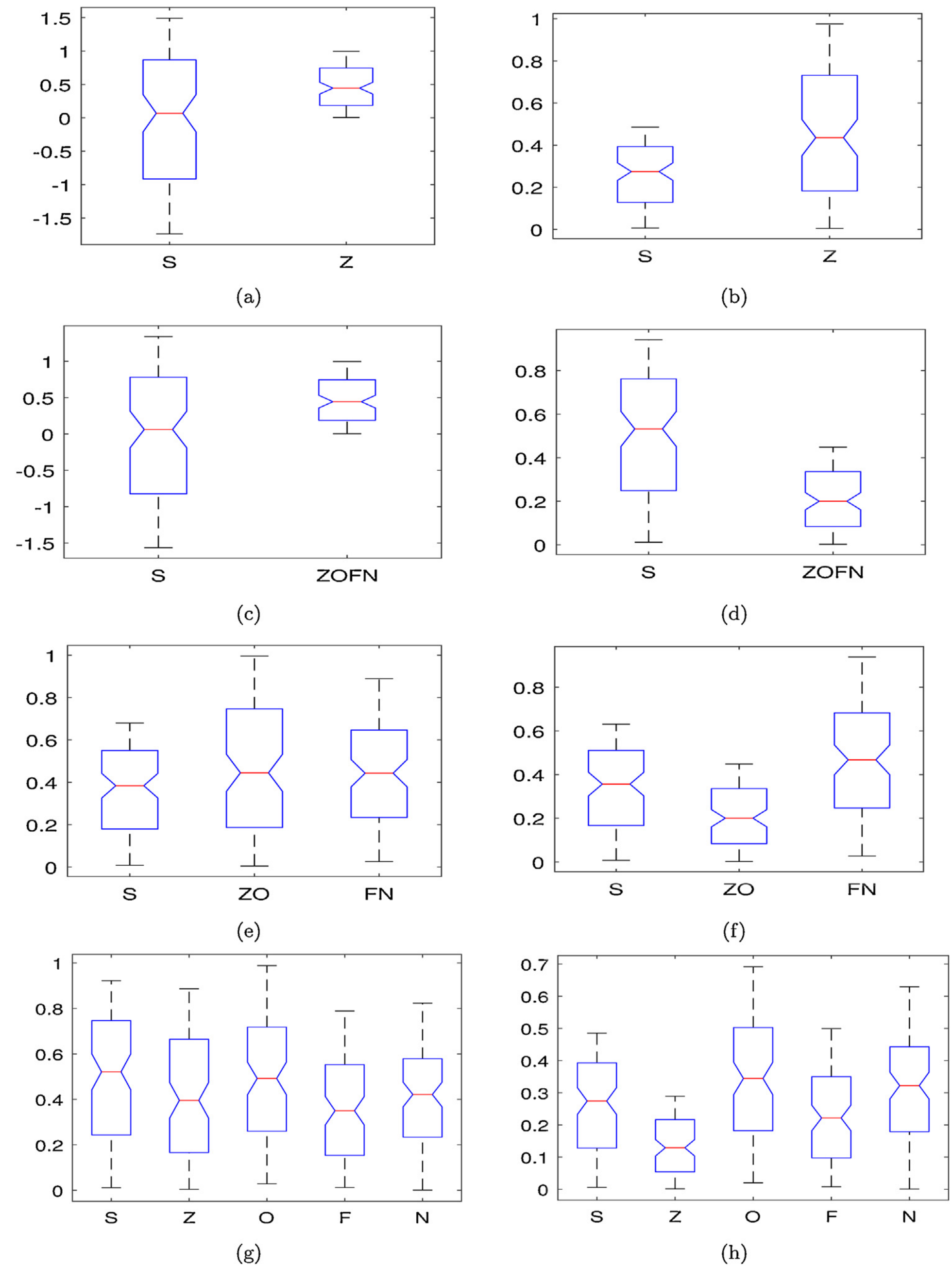

Fig. 5. Boxplots of autoencoder generated hidden layer data. Case-1: (a) layer-1 ($p = 0.004$), (b) layer-2 ($p = 1.09665 \times 10^{-10}$), Case-2: (c) layer-1 ($p = 0.001$), (d) layer-2 ($p = 1.39994 \times 10^{-9}$), Case-3: (e) layer-1 ($p = 0.045$), (f) layer-2 ($p = 2.042 \times 10^{-8}$), Case-4:(g) layer-1 ($p = 0.019$), (h) layer-2 ($p = 2.1431 \times 10^{-5}$).

and predicts the output (most probable class) based on the probability distribution. The probabilistic output of a softmax classifier for any labeled kth class is defined in terms of the multiclass labeled input data set $\mathbf{X}$ and weight matrix $\mathbf{W}$ as follows [64]:

$$\begin{aligned} P(\mathbf{Y}=k|\mathbf{X}) &= \frac{\exp(\mathbf{W}_k^T \cdot \mathbf{X})}{\sum_{i=1}^{N} \exp(\mathbf{W}_i^T \cdot \mathbf{X})} \\ &= \frac{1}{\sum_{i=1}^{N} \exp[(\mathbf{W}_i - \mathbf{W}_k)^T \cdot \mathbf{X}]} \end{aligned} \quad (9)$$

where $\cdot$ indicates the dot product and N represents the total number of classes.

4.4. Proposed algorithm for seizures classification

The basic steps of the proposed algorithm can be listed as follows:

- Analyze the raw EEG signals in two-dimensional space using the ToC.
- The two-dimensional ToC space contains various symmetries. Hence, the lower oblique attributes of the ToC are selected for further processing.
- Generate two hidden layers using the sparse autoencoder network. The first hidden layer is generated from the lower oblique ToC coefficients, and the second hidden layer is measured from the information of the first hidden layer.

- Apply two hidden-layers deep neural network to compute the seizures labeled sensitive structural information that is classified with the softmax classifier.

5. Results and discussion

In this paper, a sparse autoencoder together with ToC based data-dependent algorithm is proposed for automated seizures classification. The nonlinear dynamics of the studied EEG signals can easily be captured with the ToC, as shown in Fig. 3. The distribution of ToC for the seizures (S) EEG signal is more compact, and it contains a very high value of order 10^8 at the origin, then becomes zero towards higher lag space. However, the plots of ToC in cases of non-seizures (Z, O, F, and N) classes are found to be very different. Hence, one can achieve 100% classification accuracy between the seizures and non-seizures categories.

The ToC distribution of the normal EEG signals with eyes-open (O) is positive at the origin, whereas, the remaining ToC plots have negative values. Moreover, the plot has very less amplitude as compare to the ToC magnitude plot of the seizures EEG signals. The ToCs of the seizure-free (F, N) EEG signals are more distributed at higher lags space, which indicates the higher degree of nonlinear variation. The ToCs are negative and uniformly distributed up to some certain lags, whereas the distributions are different from other cases.

The ToC distribution of normal EEG signals (Z, O) contains several maximum and minimum locations. Variations are less for the ToC of seizure-free (F, N) EEG signals, and there is a distinct minimum for the ToC of seizures (S) EEG signal. Hence, S is easily discriminated against compared to other seizures-classes yielding 100% classification accuracy. There is some small similarity between the normal eyes-closed (Z) and seizures-free (N) EEG signals. Thus, some loss of discrimination accuracy has occurred, and we achieved 99.6% classification accuracy for three-class, while 97.2% discrimination accuracy for five-class seizures cases.

5.1. Performance measures

Table 1 shows the confusion matrix of various considered categories of seizures. The proposed algorithm has obtained 100% for binary, 99.6% for three-classes, and 97.2% for five-classes, classification accuracy to discriminate seizures-labeled EEG signals. The performance of the proposed algorithm is also validated on the basis of various performance parameters, namely, the sensitivity (SEN), specificity (SPE), total accuracy (TA), kappa value (κ), random accuracy (RA), and the Matthews correlation coefficient (MCC) [66]. The mathematical expressions of these measurements are given in terms of the true-positive (TP), false-positive (FP), true-negative (TN), and false-negative (FN), defined as follows:

$$\mathrm{SEN} = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{10}$$

$$\mathrm{SPE} = \frac{\mathrm{TN}}{\mathrm{TN} + \mathrm{FP}} \tag{11}$$

$$\mathrm{TA} = \frac{\mathrm{TP} + \mathrm{TN}}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \tag{12}$$

$$\kappa = \frac{\mathrm{TA} - \mathrm{RA}}{1 - \mathrm{RA}} \tag{13}$$

$$\mathrm{RA} = \frac{(\mathrm{TN} + \mathrm{FP}) \times (\mathrm{TN} + \mathrm{FN}) + (\mathrm{FN} + \mathrm{TP}) \times (\mathrm{FP} + \mathrm{TP})}{(\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN})^2} \tag{14}$$

$$\mathrm{MCC} = \frac{\mathrm{TP} \times \mathrm{TN} - \mathrm{FP} \times \mathrm{FN}}{\sqrt{(\mathrm{TP} + \mathrm{FP}) \times (\mathrm{FN} + \mathrm{TP}) \times (\mathrm{FP} + \mathrm{TN}) \times (\mathrm{TN} + \mathrm{FN})}} \tag{15}$$

The performance parameters for all four considered cases are listed in Table 2 that shows the effectiveness of the proposed seizures classification algorithm.

5.2. Discussion

Literature reveals that various algorithms are proposed to classify different seizures classes based on the EEG signals. In this paper, the recent year's seizures classification algorithms are considered to compare the performance of the proposed algorithm. The achieved results are compared with the results of the existing seizures classification algorithms, as shown in Table 3.

Peker et al. [26] used the DT-CWT to decompose the EEG signals, and various features are measured from each decomposed sub-signal. These parameters are classified with the complex-valued neural network (CVNN). They achieved 99.50%, 99.15%, and 97.79% classification accuracy for binary classes (S-Z, S-NFZO) and three-classes (S-FN-ZO), respectively. Swami et al. [29] employed the over-decimated DT-CWT and measured features such as the Shannon entropy, energy, and more from each sub-band. These labeled-features are discriminated with the general regression neural network (GRNN) classifier. They got 100% classification accuracy for Z-S, while 95.25% discrimination accuracy for S-ZOFN. Zamir [67] proposed four linear least squares-based pre-processing (LLSP) methods to analyze EEG signals. Four sinusoidal models are used to represent the studied signal, and a polynomial function is used to mapping the amplitude of these models. The LLSP1 and LLSP3 models measured promising features and achieved 100% classification accuracy with the logistic and LazyIB1 classifiers in seizures classification cases.

The TQWT is also used for seizures classification [34–36]. Hassan et al. [34] measured spectral slope, spectral spread, spectral centroid, and spectral flatness, whereas the multi-scale k-nearest neighbor (KNN)-entropy is measured from each decomposed sub-band in [35]. They obtained 100%, 99.60%, and 98.40% discrimination accuracy with the bootstrap aggregating algorithm whereas 100%, 99%, and 98.60% discrimination accuracies are achieved with the support vector machine (SVM) classifier in the mentioned classes respectively [35]. Patidar and Panigrahi [36] measured the Kraskov entropy as a feature from each labeled sub-signal and classify them with the least square SVM (LS-SVM) classifier. They got 97.75% discrimination accuracy to discriminate seizures and non-seizures EEG signals.

Bhardwaj et al. [41] used the EMD for automated detection of epileptic seizures based on the EEG signals. The features are measured from each IMF, and the constructive genetic programming (CGP) technique is used to classify the seizures-labeled EEG signals. The CGP introduced the constructive crossover and constructive subtree mutation operators that remove the destructive nature of conventional genetic programming. They obtained 98.64%, 99.12%, and 98.33% discrimination accuracy for three cases. Zhang and Chen [68] employed the local mean decomposition (LMD) theorem to decompose the studied signal into a series of product functions (PFs). Subsequently, some temporal and non-linear attributes are computed from the first five PFs. These features are used as the feature vectors representation of the EEG signals. The labeled-feature vectors are input to five classifiers, namely, the SVM, KNN, SVM optimized by genetic algorithm (GA-SVM), backpropagation neural network (BPNN), and linear discriminant analysis (LDA). The one-way analysis of variance is carried out for significance testing, and only those features are selected that are having the p-value lower than 0.001. They achieved 100% for Z-S, 98.86% for ZOFN-S, and 98.4% for ZO-FN-S classification accuracy.

Tiwari et al. [69] employed the difference of Gaussian (DoG) filter at multiple scales to extract key-points features of the EEG signals. The one-dimension local binary pattern (LBP) algorithm

Table 1
Confusion matrix of various classes of seizures.

True labels																
		S	Z		S	FNSO		S	FN	ZO		S	F	N	Z	O
Test labels	S	100	0	S	100	0	S	100	0	0	S	100	0	0	0	0
	Z	0	100	FNZO	0	400	FN	0	198	2	F	0	95	2	2	1
							ZO	0	0	200	N	0	2	97	1	0
											Z	0	2	1	96	1
											O	0	1	1	0	98

Table 2
Measured performance parameters.

Measurements	TA (%)	SEN (%)	SPE (%)	RA	κ	MCC
Z-S	100	100	100	0.5	1	1
ZOFN-S	100	100	100	0.68	1	1
		100	100	0.0756	0.9957	1
ZO-FN-S	99.6	100	100	0.0579	0.9958	0.9917
		100	99.33	0.0577	0.9958	0.9917
		96	99	0.6812	0.9561	1
		98	99.5	0.6801	0.9750	0.97
Z-O-F-N-S	97.2	95	98.75	0.68	0.9375	0.9417
		97	99.25	0.6788	0.9564	0.9623
		100	100	0.1	0.1	1

Table 3
Comparison table of seizures classification algorithms.

Authors (year)	Methods and Features	Classifier	Classes	Accuracy (%)
Peker et al. (2016) [26]	DT-CWT	CVNN	S-Z	99.50
			S-NFZO	99.15
			S-FN-ZO	97.79
Zamir (2016) [67]	LLSP1	Logistic, LazyIB1	Z-S	100
	LLSP3	Logistic, LazyIB1	ZOFN-S	100
			Z-O-F-N-S	100
Hassan et al. (2016) [34]	TQWT	Bootstrap aggregating	Z-S	100
			ZOFN-S	99.60
			ZO-FN-S	98.40
Bhardwaj et al. (2016) [41]	EMD	CGP	Z-S	98.64
			ZOFN-S	99.12
			ZO-FN-S	98.33
Zhang &Chen (2016) [68]	LMD	GA-SVM	Z-S	100
			ZOFN-S	98.86
			ZO-FN-S	98.4
Swami et al. (2016) [29]	DT-CWT	GRNN	S-Z	100
			S-ZOFN	95.24
Tiwari et al. (2016) [69]	1D-LBP	SVM	S-ZOFN	99.31
			S-ZO-FN	98.80
Chen et al. (2017) [23]	DWT	NN/SVM	Z-S	100/78
		NN	ZO-FN-S	100
		NN/SVM	ZOFN-S	100/93.6
Bhati et al. (2017)[70]	WFB	MLP	S-NF	99.66
Bhattacharyya et al. (2017) [35]	TQWT	SVM	S-FN-ZO	98.60
			S-Z	100
			S-FNZO	99
Zhang et al. (2017) [44]	VMD		ZO-FN-S	97.352
Bhattacharyya et al. (2017) [32]	EWT	Random forest	S-FN-ZO	99.4
Sharma et al. (2017) [33]	ATFAWT	LS-SVM	S-Z	100
			S-ZOFN	99.2
Patidar and Panigrahi (2017) [36]	TQWT	LS-SVM	S-ZOFN	97.75
Acharya et al. (2018) [45]	13-layer deep CNN	–	ZO-FN-S	88.70
Zhou et al. (2018) [46]	3-layer CNN	–	ZO-FN-S	93
			S-Z	100
Hussein et al. (2018) [47]	LSTM cells based RNN	Softmax	S-FN-ZO	100
			S-F-N-Z-O	100
			S-Z	100
Proposed work	ToC and deep neural network	Softmax	S-FN-ZO	99.6
			S-FNZO	100
			S-F-N-Z-O	97.2

is applied at these detected key-points, and the histogram of the LBPs are used as features. The SVM classifier with the radial basis function (RBF) kernels is used to distinguish both binary and three seizures classes. Chen et al. [23] used the Daubechies-4, Coiflet-3, Haar, and Symmlet-4 wavelet transforms to split the EEG signals up to six levels without downsampling. They compute the fast Fourier transform (FFT) on each decomposed level, and the magnitudes of the FFT are considered as features. They achieved 100% classification accuracy in all cases with the nearest neighbor (NN) classifier and got 78% and 93.6% accuracy for Z-S and ZOFN-S, respectively, with the SVM classifier. The three-band synthesis wavelet filter bank (WFB) technique is also proposed to discriminate seizures classes [70]. The semidefinite relaxation and the nonlinear least square method are used to generalize the WFB. They obtained 99.66% classification accuracy for S-NF classes using the multi-layer perception (MLP) classifier.

Bhattacharyya et al. [32] used the EWT to explored the studied EEG signals and measured the amplitudes and instantaneous frequencies of the signal at adaptive frequency scales. They evaluated their proposed algorithm with six classifiers and achieved 99.41% average classification accuracy with the random forest classifier. The ATFAWT is also proposed for seizures classification [33]. They obtained 100% and 99.2% classification accuracy to discriminate S-Z and S-ZONF, EEG signals. Zhang et al. [44] executed the VMD method to resolve the studied raw EEG recordings into fifteen BLIMFs, and the logarithmic operation is subsequently performed on frequency localized sub-signals. The Burg method is employed to compute the AR model parameters of these sub-signals. Various statistical parameters, namely, average, maximum, length, minimum, energy, skewness, variance, and kurtosis, are measured from the AR model coefficients. They obtained 97.352% classification accuracy to classify ZO-FN-S, EEG signals. The CNN is also employed for seizures classification.

Acharya et al. [45] obtained 88.70% seizures classification accuracy to discriminate the ictal, preictal, and seizure classes with the 13-layers deep CNN. The labeled EEG signals are possessed by the *z*-score normalization algorithm, and the resultant features are employed to the one-dimensional deep-CNN. The proposed CNN architecture has different layers, namely, five convolutional, five max-pooling, and three fully connected layers. Zhou et al. [46] proposed a three-layer CNN that consists of a convolutional layer, a pooling layer and a fully connected layer. The time and frequency-domain studied EEG signals are directly inputted to the defined CNN. They obtained 93% classification accuracy with frequency-domain signals, while achieved 47.9% discrimination accuracy with time-domain signals, to classify ZO-FN-S. Hussein et al. [47] first divided the non-stationary EEG signals into short-length segments. They proposed an RNN architecture that contains memory cells known as the LSTM cells. These EEG segmented data are feed to the proposed network that learns the robust features of underlying nonlinear variation of studied signals. They achieved 100% seizures classification accuracy for all cases with the softmax classifier.

The proposed algorithm has the following advantages:

- The proposed algorithm is feature-independent.
- The algorithm is fully automated that reduces human error.
- It is easy to implement and immune to the Gaussian noise.

We plan to study the potential of using HOS and other deep leaning architectures for the diagnosis and prediction of other brain disorders such as Parkinson's disease, epilepsy detection, sleep disorders, and more.

6. Conclusion

The EEG signals have nonlinear and non-stationary behavior. The dynamics of these signals are not interpreted by visual inspection. In this article, a nonlinear method is used that captures the subtle information about the underlying dynamics of the EEG signals. The sparse autoencoder is introduced to automatically identify epileptic seizure attacks based on the ToC of the EEG signals. The initial features are extracted from the lower diagonal matrix of the ToC of the EEG signal that are further analyzed by a deep neural network. These hidden-layers are generated from the sparse autoencoder network that is initially trained separately using the greedy algorithm. The proposed method obtained up to 100% seizures classification accuracy. The achieved results show that the proposed algorithm is able to capture the hidden information of the EEG signals that leads to better classification accuracy.

Declaration of Competing Interest

None.

References

[1] D.S. Bassett, O. Sporns, Network neuroscience, Nat. Neurosci. 20 (3) (2017) 353.

[2] S. Pati, A.V. Alexopoulos, Pharmacoresistant epilepsy: from pathogenesis to current and emerging therapies, Cleve. Clin. J. Med. 77 (7) (2010) 457–567.

[3] V.K. Jirsa, W.C. Stacey, P.P. Quilichini, A.I. Ivanov, C. Bernard, On the nature of seizure dynamics, Brain 137 (8) (2014) 2210–2230.

[4] WHO, Epilepsy report by World Health Organization, 2015 http://www.who.int/en/news-room/fact-sheets/detail/epilepsy.

[5] S. Altunay, Z. Telatar, O. Erogul, Epileptic EEG detection using the linear prediction error energy, Expert Syst. Appl. 37 (8) (2010) 5661–5665.

[6] V. Joshi, R.B. Pachori, A. Vijesh, Classification of ictal and seizure-free EEG signals using fractional linear prediction, Biomed. Signal Process. Control 9 (2014) 1–5.

[7] E.D. Übeyli, Least squares support vector machine employing model-based methods coefficients for analysis of EEG signals, Expert Syst. Appl. 37 (1) (2010) 233–239.

[8] H. Khamis, A. Mohamed, S. Simpson, Seizure state detection of temporal lobe seizures by autoregressive spectral analysis of scalp EEG, Clin. Neurophysiol. 120 (8) (2009) 1479–1488.

[9] S.-H. Kim, C. Faloutsos, H.-J. Yang, Coercively adjusted auto regression model for forecasting in epilepsy EEG, Comput. Math. Methods Med. 2013 (2013).

[10] K. Polat, S. Güneş, Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform, Appl. Math. Comput. 187 (2) (2007) 1017–1026.

[11] K. Polat, S. Güneş, Artificial immune recognition system with fuzzy resource allocation mechanism classifier, principal component analysis and FFT method based new hybrid automated identification system for classification of EEG signals, Expert Syst. Appl. 34 (3) (2008) 2039–2048.

[12] O. Faust, U.R. Acharya, L.C. Min, B.H. Sputh, Automatic identification of epileptic and background EEG signals using frequency domain parameters, Int. J. Neural Syst. 20 (02) (2010) 159–176.

[13] K. Samiee, P. Kovacs, M. Gabbouj, Epileptic seizure classification of EEG time-series using rational discrete short-time Fourier transform, IEEE Trans. Biomed. Eng. 62 (2) (2015) 541–552.

[14] A.T. Tzallas, M.G. Tsipouras, D.I. Fotiadis, Automatic seizure detection based on time–frequency analysis and artificial neural networks, Comput. Intell. Neurosci. 2007 (2007).

[15] Z. Iscan, Z. Dokur, T. Demiralp, Classification of electroencephalogram signals with combined time and frequency features, Expert Syst. Appl. 38 (8) (2011) 10499–10505.

[16] C. Guerrero-Mosquera, A.M. Trigueros, J.I. Franco, A. Navia-Vázquez, New feature extraction approach for epileptic EEG signal detection using time–frequency distributions, Med. Biol. Eng. Comput. 48 (4) (2010) 321–330.

[17] H. Ocak, Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy, Expert Syst. Appl. 36 (2) (2009) 2027–2036.

[18] L. Guo, D. Rivero, J. Dorado, J.R. Rabunal, A. Pazos, Automatic epileptic seizure detection in EEGs based on line length feature and artificial neural networks, J. Neurosci. Methods 191 (1) (2010) 101–109.

[19] A. Subasi, M.I. Gursoy, EEG signal classification using PCA, ICA, LDA and support vector machines, Expert Syst. Appl. 37 (12) (2010) 8659–8666.

[20] U. Orhan, M. Hekim, M. Ozer, EEG signals classification using the k-means clustering and a multilayer perceptron neural network model, Expert Syst. Appl. 38 (10) (2011) 13475–13481.

[21] M. Li, W. Chen, T. Zhang, Classification of epilepsy EEG signals using DWT-based envelope analysis and neural network ensemble, Biomed. Signal Process. Control 31 (2017) 357–365.

[22] L. Guo, D. Rivero, J. Dorado, C.R. Munteanu, A. Pazos, Automatic feature extraction using genetic programming: An application to epileptic EEG classification, Expert Syst. Appl. 38 (8) (2011) 10425–10436.

[23] G. Chen, W. Xie, T.D. Bui, A. Krzyżak, Automatic epileptic seizure detection in EEG using nonsubsampled wavelet-Fourier features, J. Med. Biol. Eng. 37 (1) (2017) 123–131.

[24] S. Ghosh-Dastidar, H. Adeli, N. Dadmehr, Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection, IEEE Trans. Biomed. Eng. 54 (9) (2007) 1545–1551.

[25] L. Guo, D. Rivero, A. Pazos, Epileptic seizure detection using multiwavelet transform based approximate entropy and artificial neural networks, J. Neurosci. Methods 193 (1) (2010) 156–163.

[26] M. Peker, B. Sen, D. Delen, A novel method for automated diagnosis of epilepsy using complex-valued classifiers, IEEE J. Biomed. Health Inform. 20 (1) (2016) 108–118.
[27] G. Chen, Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features, Expert Syst. Appl. 41 (5) (2014) 2391–2394.
[28] A.B. Das, M.I.H. Bhuiyan, S.S. Alam, Classification of EEG signals using normal inverse Gaussian parameters in the dual-tree complex wavelet transform domain for seizure detection, Signal, Image Video Process. 10 (2) (2016) 259–266.
[29] P. Swami, T.K. Gandhi, B.K. Panigrahi, M. Tripathi, S. Anand, A novel robust diagnostic model to detect seizures in electroencephalography, Expert Syst. Appl. 56 (2016) 116–130.
[30] D. Wang, D. Miao, C. Xie, Best basis-based wavelet packet entropy feature extraction and hierarchical EEG classification for epileptic detection, Expert Syst. Appl. 38 (11) (2011) 14314–14320.
[31] U.R. Acharya, S.V. Sree, A.P.C. Alvin, J.S. Suri, Use of principal component analysis for automatic classification of epileptic EEG activities in wavelet framework, Expert Syst. Appl. 39 (10) (2012) 9072–9078.
[32] A. Bhattacharyya, R.B. Pachori, A multivariate approach for patient-specific EEG seizure detection using empirical wavelet transform, IEEE Trans. Biomed. Eng. 64 (9) (2017) 2003–2015.
[33] M. Sharma, R.B. Pachori, U.R. Acharya, A new approach to characterize epileptic seizures using analytic time–frequency flexible wavelet transform and fractal dimension, Pattern Recogn. Lett. 94 (2017) 172–179.
[34] A.R. Hassan, S. Siuly, Y. Zhang, Epileptic seizure detection in EEG signals using tunable-Q factor wavelet transform and bootstrap aggregating, Comput. Methods Programs Biomed. 137 (2016) 247–259.
[35] A. Bhattacharyya, R.B. Pachori, A. Upadhyay, U.R. Acharya, Tunable-Q wavelet transform based multiscale entropy measure for automated classification of epileptic EEG signals, Appl. Sci. 7 (4) (2017) 385.
[36] S. Patidar, T. Panigrahi, Detection of epileptic seizure using Kraskov entropy applied on tunable-Q wavelet transform of EEG signals, Biomed. Signal Process. Control 34 (2017) 74–80.
[37] R.J. Oweis, E.W. Abdulhay, Seizure classification in EEG signals utilizing Hilbert–Huang transform, Biomed. Eng. Online 10 (1) (2011) 38.
[38] K. Fu, J. Qu, Y. Chai, Y. Dong, Classification of seizure based on the time–frequency image of EEG signals using HHT and SVM, Biomed. Signal Process. Control 13 (2014) 15–22.
[39] S. Li, W. Zhou, Q. Yuan, S. Geng, D. Cai, Feature extraction and recognition of ictal EEG using EMD and SVM, Comput. Biol. Med. 43 (7) (2013) 807–816.
[40] S.S. Alam, M.I.H. Bhuiyan, Detection of seizure and epilepsy using higher order statistics in the EMD domain, IEEE J. Biomed. Health Inform. 17 (2) (2013) 312–318.
[41] A. Bhardwaj, A. Tiwari, R. Krishna, V. Varma, A novel genetic programming approach for epileptic seizure detection, Comput. Methods Programs Biomed. 124 (2016) 2–18.
[42] F. Riaz, A. Hassan, S. Rehman, I.K. Niazi, K. Dremstrup, EMD-based temporal and spectral features for the classification of EEG signals using supervised learning, IEEE Trans. Neural Syst. Rehabil. Eng. 24 (1) (2015) 28–35.
[43] M.R. Kumar, Y.S. Rao, Epileptic seizures classification in EEG signal based on semantic features and variational mode decomposition, Cluster Comput. (2018) 1–11.
[44] T. Zhang, W. Chen, M. Li, AR based quadratic feature extraction in the VMD domain for the automated seizure detection of EEG using random forest classifier, Biomed. Signal Process. Control 31 (2017) 550–559.
[45] U.R. Acharya, S.L. Oh, Y. Hagiwara, J.H. Tan, H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals, Comput. Biol. Med. 100 (2018) 270–278.
[46] M. Zhou, C. Tian, R. Cao, B. Wang, Y. Niu, T. Hu, H. Guo, J. Xiang, Epileptic seizure detection based on EEG signals and CNN, Frontiers Neuroinform. 12 (2018) 95.
[47] R. Hussein, H. Palangi, R. Ward, Z.J. Wang, Epileptic Seizure Detection: A Deep Learning Approach, 2018, arXiv preprint arXiv:1803.09848.
[48] R.G. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C.E. Elger, Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: dependence on recording region and brain state, Phys. Rev. E 64 (6) (2001) 061907.
[49] C.M. Bishop, Pattern Recognition and Machine Learning, Springer Cambridge, UK, 2006.
[50] A. Van Ooyen, B. Nienhuis, Improving the convergence of the back-propagation algorithm, Neural Netw. 5 (3) (1992) 465–471.
[51] Y. Hirose, K. Yamashita, S. Hijiya, Back-propagation algorithm which varies the number of hidden units, Neural Netw. 4 (1) (1991) 61–66.
[52] P. Baldi, Autoencoders, unsupervised learning, and deep architectures, Proceedings of ICML Workshop on Unsupervised and Transfer Learning (2012) 37–49.
[53] R. Sharma, P. Sircar, R.B. Pachori, A new technique for classification of focal and nonfocal EEG signals using higher-order spectra, J. Mech. Med. Biol. 19 (01) (2019) 1940010.
[54] R. Sharma, P. Sircar, R.B. Pachori, Computer-aided diagnosis of epilepsy using bispectrum of EEG signals, in: Application of Biomedical Engineering in Neuroscience, Springer, 2019, pp. 197–220.
[55] J. Fonoliosa, C. Nikias, Wigner higher order moment spectra: definition, properties, computation and application to transient signal analysis, IEEE Trans. Signal Process. 41 (1) (1993) 245.
[56] R. Sharma, P. Sircar, R.B. Pachori, Automated focal EEG signal detection based on third-order cumulant function, Biomed. Signal Process. Control 58 (2020) 101856.
[57] R. Sharma, P. Sircar, R.B. Pachori, S.V. Bhandary, U.R. Acharya, Automated glaucoma detection using center slice of higher order statistics, J. Mech. Med. Biol. 19 (01) (2019) 1940011.
[58] R. Sharma, R.B. Pachori, P. Sircar, Automated emotion recognition based on higher order statistics and deep learning algorithm, Biomed. Signal Process. Control 58 (2020) 101867.
[59] D.R. Brillinger, An introduction to polyspectra, Annals Math. Stat. (1965) 1351–1374.
[60] C.L. Nikias, J.M. Mendel, Signal processing with higher-order spectra, IEEE Signal Process. Mag. 10 (3) (1993) 10–37.
[61] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion, J. Mach. Learn. Res. 11 (Dec) (2010) 3371–3408.
[62] W. Sun, S. Shao, R. Zhao, R. Yan, X. Zhang, X. Chen, A sparse auto-encoder-based deep neural network approach for induction motor faults classification, Measurement 89 (2016) 171–178.
[63] I. Goodfellow, Y. Bengio, A. Courville, Y. Bengio, Deep learning, vol. 1, MIT Press, Cambridge, 2016.
[64] J. Schmidhuber, Deep learning in neural networks: an overview, Neural Netw. 61 (2015) 85–117.
[65] W.H. Kruskal, W.A. Wallis, Use of ranks in one-criterion variance analysis, J. Am. Stat. Assoc. 47 (260) (1952) 583–621.
[66] D.M. Powers, Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Markedness and Correlation, 2011.
[67] Z.R. Zamir, Detection of epileptic seizure in EEG signals using linear least squares preprocessing, Comput. Methods Programs Biomed. 133 (2016) 95–109.
[68] T. Zhang, W. Chen, LMD based features for the automatic seizure detection of EEG signals using SVM, IEEE Trans. Neural Syst. Rehabil. Eng. 25 (8) (2016) 1100–1108.
[69] A.K. Tiwari, R.B. Pachori, V. Kanhangad, B.K. Panigrahi, Automated diagnosis of epilepsy using key-point-based local binary pattern of EEG signals, IEEE Journal of Biomedical and Health Informatics 21 (4) (2016) 888–896.
[70] D. Bhati, M. Sharma, R.B. Pachori, V.M. Gadre, Time-frequency localized three-band biorthogonal wavelet filter bank using semidefinite relaxation and nonlinear least squares with epileptic seizure EEG signal classification, Digital Signal Process. 62 (2017) 259–273.