ApEn-Based Epileptic EEG Classification Using Support Vector Machine

Hardika B. Gabani[(✉)] and Chirag N. Paunwala

Dr. R.K. Desai Marg, Opp. Mission Hospital, Athwalines, Surat, Gujarat, India
hardikagabani62@gmail.com, chirag.paunwala@scet.ac.in

Abstract. The ElectroEncephaloGram (EEG) signal plays an important role to identifying the disorder of epilepsy. Epilepsy is a neurological disorder which is an unexpected electrical disturbance of the brain. Due to which nerve cell activity in the brain becomes disrupted, causes people to have a "Seizure". Now a day researchers are working and focusing on an automatic analysis of EEG signal to classify the Epilepsy. The EEG signal recording system generate very lengthy data. So, classification of epilepsy seizure requires a time-consuming process. This paper proposes SVM (Support Vector Machine) based automatic epilepsy seizure classification system that uses ApEn (Approximation Entropy). ApEn is reducing the patient data size without any loss of patient data so; we can easily classify the epilepsy seizure. ApEn is statistical parameters that measure the current amplitude value of an EEG signal based on its previous amplitude value. In this paper, we measure sensitivity, specificity, and accuracy using SVM classifier. The overall values as high as 100% can be achieved using the proposed system to differentiate epileptic state (Seizure class) out of normal state (Non-seizure Class) using time domain method.

Keywords: ElectroEncephaloGram (EEG) signal
Epilepsy seizure classification · Approximate Entropy (ApEn)
SVM (Support Vector Machine)

1 Introduction

Electrical activity is occurring in different brain regions which are determined by the EEG signal and we can also determine the relative positions and strengths of it. The abnormal electrical activities fetched by using EEG are called epilepsy seizures. Approximately 50 million people have epilepsy seizure worldwide [1]. Possible causes of epilepsy include brain injury, metabolic disturbances, alcohol or drug abuse, brain tumors, and genetic disorders.

In the small time period, epileptic seizure can't be predicted in most of the cases. For classification purpose, continuous recording of the EEG is required. Sometimes EEG recording takes very large time duration. It may be up to one week or two weeks. As the traditional methods are monotonous and slow. In past few years, automated epilepsy seizure classification systems have been developed [2]. The proposed work is an automatic epileptic EEG classification system using SVM and feature extraction and reduction by using Approximate Entropy (ApEn).

As shown in below figure we give the EEG signal at the input side. ApEn technique [3] is used to extract the features of the signal. Extracted features are then apply to the classifier to classify seizures or non-seizures data (Fig. 1).
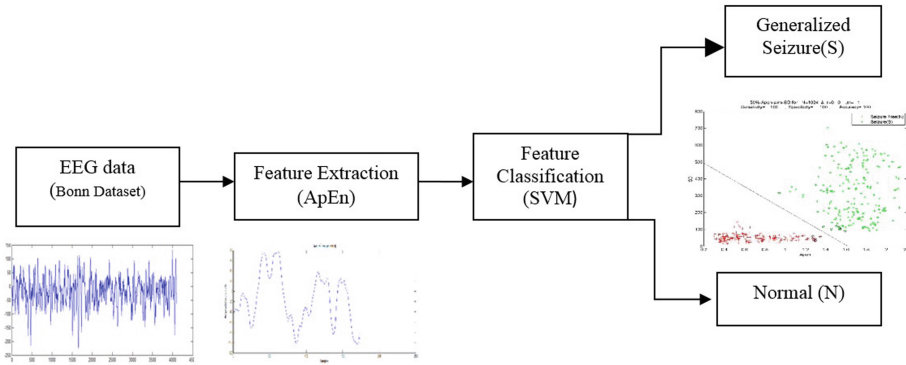

Fig. 1. Block diagram

Programmed examination and finding of epilepsy in view of EEG recordings is begun in the mid-1970s. Today, PC-based examination addresses two problems: Epilepsy seizure classification and EEG analysis. Many feature extraction techniques have been used for the classification of Epilepsy seizure. SVM (Support Vector Machine) based classification system for epilepsy seizure have been proposed by many researcher. The research based on nonlinear parameters has been found clinically fruitful for classification of Epilepsy seizure.

The Lyapunov exponent [4–6] provides significant details about changes in EEG activity in turn facilitating early detection of epilepsy. The correlation dimension [7] is useful to measure correlation which quantifies complex neural activity of human brain. During epileptic seizure, the value of ApEn has been found to exhibit strong relationship with synchronous discharge of large groups of neurons. The features obtained from complexity analysis and spectral analysis of EEG signals has been effectively used for diagnosis of epilepsy [8]. Recently, the ApEn (Approximate Entropy) [3] based methods have been developed for analyzing linear signals for classification of epileptic seizures in epilepsy seizure [9, 13]. The MEAN frequency parameter of IMFs has been proposed to discriminate well between seizure and seizure-free EEG signals. For classification between healthy and epileptic EEG signals, weighted frequency has been found to be some parameter [10]. Analysis of normal and epileptic seizure EEG signals by using area measured from the trace of analytical signal representation of Intrinsic Mode Function (IMF) has been proposed in [11]. The area parameter and mean frequency of IMFs computed using Fourier–Bessel expansion used for epileptic seizure classification in EEG signals [12]. Also, IMFs of EEG signals have been used for recognition of epileptic seizure [13].

2 Proposed Algorithm

2.1 ApEn (Approximate Entropy) Based Feature Extraction

An ApEn is a technique used to quantify the amount of regularity and the unpredictability of fluctuations over time-series data [3].

(1) Let EEG signal with N data points $X = [x(1), x(2), x(3), \ldots, x(N)]$.
(2) Let $x(i)$ be a subsequence of X such that $x(i) = [x(i), x(i + 1), x(i + 2),\ldots, x(i + m - 1)]$ for $1 \leq i \leq N - m$, where m represents the number of samples used for the prediction.
(3) To reduce the noise, filter with level r is represented as, $r = k * SD$ for $k = 0, 0.1, 0.2, 0.3,\ldots, 0.9$
Where SD is the standard deviation of X.
(4) Let $\{x(j)\}$ represent a set of subsequence's obtained from $x(j)$ by varying j from 1 to N. Each sequence $x(j)$ in the set of $\{x(j)\}$ is compared with $x(i)$ and, in this process, two parameters, namely, $C_i^m(r)$ and $C_i^{m+1}(r)$ are defined as follows:

$$C_i^m(r) = \frac{\sum_{j=1}^{N-m} k_j}{N - m} \tag{1}$$

Where,

$$k = \begin{cases} 1, & \text{if } |x(i) - x(j)| \text{ for } 1 \leq j \leq N - m \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

$$C_i^{m+1}(r) = \frac{\sum_{j=1}^{N-m} k_j}{N - m}$$

(5) Finally, we get Approximation Entropy,

$$\text{ApEn}(m, r, N) = \frac{\sum_{i=1}^{N-m} \ln(c_i^m(r))}{N - m} - \frac{\sum_{i=1}^{N-m} \ln(c_i^{m+1}(r))}{N - m} \tag{3}$$

Approximation entropy value extracted from the different size of data frames is shown in Table 1. Now as from the algorithm of ApEn, m is sample value varies from 1 to 3 and for particular m, we are using 10 values of entropy as mentioned in Table 2. Table 3 shows the reduction of the 4,09,700 sample to small sample size.

Table 1. Frame size

Size of frame (N)	No. of frame per each time-series
173	4097/173 = 23
256	4097/256 = 16
512	4097/512 = 8
1024	4097/1024 = 4
2048	4097/2048 = 2

Table 2. Number of entropy value per each time-series

Size of frame (N)	No. of frame per each time-series	No. of entropy values per each time-series
173	23	30 * 23 = 690
256	16	30 * 16 = 480
512	8	30 * 8 = 240
1024	4	30 * 4 = 120
2048	2	30 * 2 = 60

Table 3. Reduction of sample size

Size of frame (N)	Final data after apply Apen (4097 * 100 = 4,09,700)
173	690 * 100 = 69000
256	480 * 100 = 48000
512	240 * 100 = 24000
1024	120 * 100 = 12000
2048	60 * 100 = 6000

2.2 Support Vector Machine (SVM)

We map input patterns into a higher dimensional feature space through using SVM (Support Vector Machine). In this high dimensional feature space, linear decision surface constructed. So, SVM is a linear classifier in the parameter space [15].

Let we take m dimensional training data set $xi = (1,..., M)$ and their class labels be yi, where $yi = 1$ and $yi = -1$ for positive and negative classes respectively. In particular input space, linear separable data then the following decision function can be determined as,

$$D(x) = w^t g(x) + b \tag{4}$$

Maps x into the l-dimensional space, we use $g(x)$ is a mapping function. B is a scaler and w is the vector in 1-dimensional space. If we separate data linearly, the decision function satisfies the following condition given below:

$$Yi = (w^t g(xi) + b) > = 1 \tag{5}$$
$$\text{Where, } i = 1,..., M$$

For an infinite number of decision functions I is linearly separable in the feature space then it satisfy Eq. (5). So, we require that the hyper-plane that have the largest margin between positive and negative class. The $D(x)/\|w\|$ is margin that contain minimum distance from the separating hyper-plane to the input data.

Assume that the margin is ρ, the following condition is to be satisfied

$$\frac{YiD(xi)}{\|w\|} \geq \rho$$

$$\text{Where, } i = 1, \ldots, M \tag{6}$$

The product of ρ and $\|w\|$ is fixed

$$\rho\|w\| = \mathbf{1} \tag{7}$$

In order to obtain the optimal separating hyper-plane with contain maximum margin, w with the minimum $\|w\|$ that satisfying Eq. (6) found. From Eq. (7), this equations are solving this optimization problem. Minimizing Yi,

$$Yi = (w^t g(xi) + b) > = \mathbf{1} \tag{8}$$

We introduce slack variable ξ, When training data are not linearly separable into Eq. (5) as follows subject to the constraints:

$$Yi = (w^t g(xi) + b) > = \mathbf{1} - \xi i$$

$$\xi i \geq 0 \text{ for } i = 1, \ldots, M \tag{9}$$

The optimal separating hyper-plane is determined so that the maximization of the margin and the minimization of the training error achieved. Minimizing

$$\frac{1}{2} w^t w + \frac{c}{2} \sum_{i=1}^{n} \xi_i^p \tag{10}$$

Subject to the constraints:

$$Yi = (w^t g(xi) + b) > = \mathbf{1} - \xi i$$

$$\xi i \geq 0 \text{ for } i = 1, \ldots, M \tag{11}$$

Where C is a parameter that determines the trade-off between the maximum margin and the minimum classification error and p is 1 or 2. When $p = 1$, the SVM is called L1 soft margin SVM (L1-SVM), and when $p = 2$, L2 soft margin SVM (L2-SVM). In the conventional SVM, optimal separating hyper-plane obtained by solving the above quadratic programming problem. In this empirically and optimal results achieved using Radial Basis Function (RBF).

In first experiment, all 100 time-series of N and S is taken for training and testing. For frame size 173, entropy values are 690 for each time-series, so if we take 100 time-series, entropy values would be 69000 for one class and it is double (138000) by considering both seizure and non-seizure class. These procedures followed for all four features. Entropy values of both classes S and N for training and testing dataset for all frames is shown in Table 4 (Fig. 2).

Table 4. Number of entropy value for testing

Sr. no.	Time-series of N and S	Frame size	No. of entropy values for training	No. of entropy values for testing
1.	200	173	1,38,000	1,38,000
2.	200	256	96,000	96,000
3.	200	512	48,000	48,000
4.	200	1024	24,000	24,000
5.	200	2048	12,000	12,000
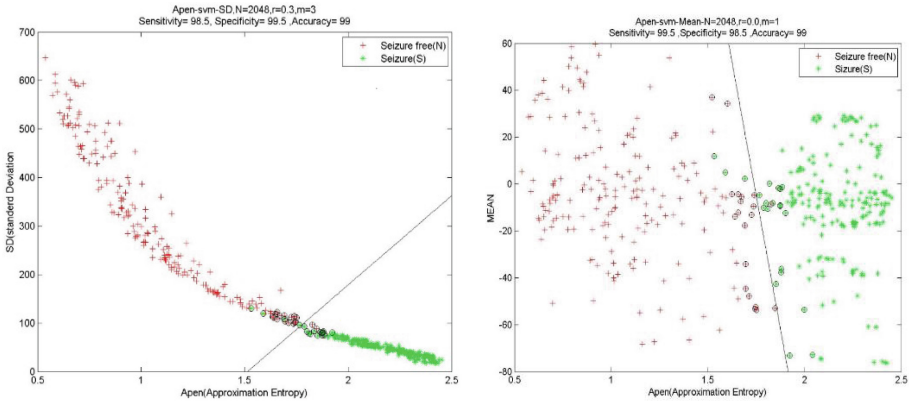
Fig. 2. ApEn for N and S file set for (a) $N = 2048$, $m = 3$, $r = 0.3$ with SD, (b) $N = 2048$, $m = 1$, $r = 0.0$ with mean

For frame size $N = 2048$ and $m = 3$ and $r = 0.3$, gets optimum accuracy for ApEn. From that, we get highest accuracy 99.00% of the feature ApEn with SD for experiment and frame size $N = 2048$ and $m = 1$ and $r = 0.3$, gets optimum accuracy for ApEn. From that, we get highest accuracy 99.00% of the feature ApEn with Mean for experiment.

For training purpose, all 50-time-series data for N and S taken and 100 time-series data, taken for testing. Entropy values of both classes S and N for training and testing dataset for all frames as shown in Table 5.

Table 5. Entropy value after 50% training

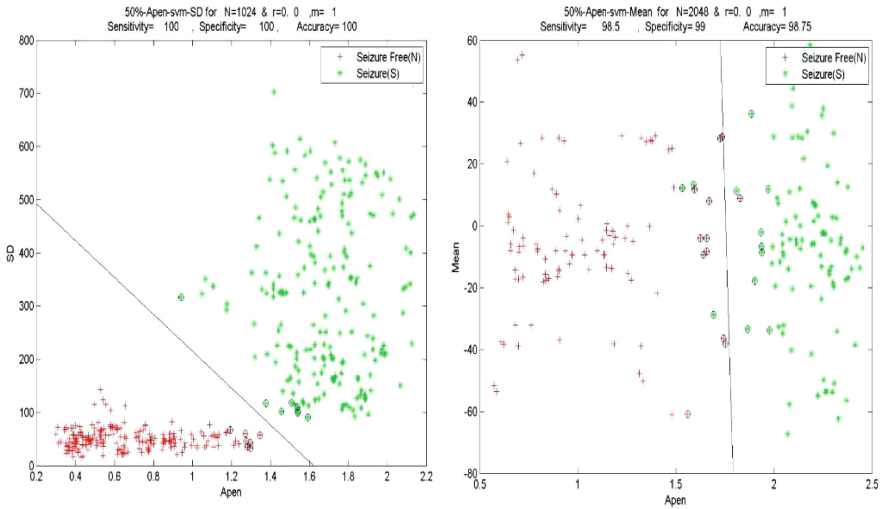| Sr no.	Frame size	Time-series of N and S for training	Time-series of N and S for testing	No. of entropy values for training
1.	173	100	200	69,000
2.	256	100	200	48,000
3.	512	100	200	24,000
4.	1024	100	200	12,000
5.	2048	100	200	6000

Fig. 3. ApEn for N and S file set after 50% training and testing (a) $N = 1024$, $m = 1$, $r = 0.0$ with SD, (b) $N = 2048$, $m = 1$, $r = 0.0$ with mean

The above Fig. 3 are for all optimum results of experiment feature dataset as shown in above Table 5. The figure shows the SVM classification for the seizure and normal class using radial basis kernel function. Where seizure is denoted by * and normal by +. The line is describing linear classification of the dataset. The o describes wrongly classify data points of opposite class.

2.3   Performance Parameters

2.3.1   Standard Deviation

Quantify the amount of variation or dispersion of a set of data values by using standard deviation. The standard deviation of a random variable like,

(1) Statistical population,
(2) Data set, or probability distribution is the square root of its variance [15].

2.3.2   Mean

The Mean is also called as a arithmetic mean of a sample. It is usually denoted by x. The x is the sum of the signals sampled values divided by the number of items in the sample [15].

2.3.3 Sensitivity

$$\text{Sensitivity} = \frac{\text{No. of true positive detected data points}}{\text{total no. of positive data points}} \tag{12}$$

Sensitivity considered for detection of seizure data [16].

2.3.4 Specificity

$$\text{Specificity} = \frac{\text{No. of true negative detected data points}}{\text{total no. of negative data points}} \tag{13}$$

Specificity considered for detection of non-seizure data [16].

2.3.5 Accuracy

$$\text{Accuracy} = \frac{(\text{TP}) + (\text{TN})}{\text{total no. of data points}} \tag{14}$$

TP = No. of true positive detected data points
TN = No. of true negative detected data points [16].

3 Experimentation Results

In our work, we have extracted the features from the EEG signal and classification done using SVM classifier in to two class seizure-free and seizure patient data. ApEn values is measure in form of m, r, and N. The values of m, r, and N are as follows:

(1) Number of Samples (m) = 1, 2, 3;
(2) Normalization Ratio (r) = 0%–90% of SD of the data sequence in increments of 10%;
(3) Frame Size (N) = 173, 256, 512, 1024 and 2048.

Approximation Entropy is extracted along with SD and mean. The randomness of EEG signal were extracted in the features, based on different size of frame (N), number of samples values (m) and normalized ratio (k). From the set of features, ApEn with SD and mean, are used for classification using the SVM classifier.

We have used BONN dataset for EEG signals which is publicly available online and described in Andrzejak et al. [17]. The EEG dataset contains both seizures and non-seizures. The Bonn dataset consists five subsets (Z, O, N, F, and S) each containing 100 single-channel EEG signals, each signal of 23.6 s in duration with the sampling rate of 173.61 Hz.

EEG recordings of five healthy volunteers with eyes open (Z) and closed (O) have been recorded on the surface, using standard electrode placement scheme. The signal F and S are seizure free subset. These two are recorded in seizure-free intervals from five patients in the epileptogenic zone (F-Seizure free) and from the hippocampal formation

of the opposite hemisphere of the brain (N-seizure free). The set S is contained seizures signal which gives an ictal activity by using with the same 128-channel amplifier system with an average common reference all EEG signals are recorded. In the proposed work classification of the N (Seizure free class) and S (Seizure class) is done by using ApEn (Approximate Entropy) feature extraction and reduction and SVM as
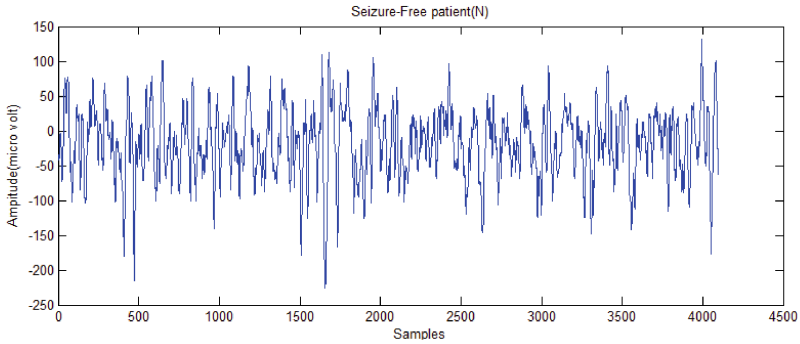

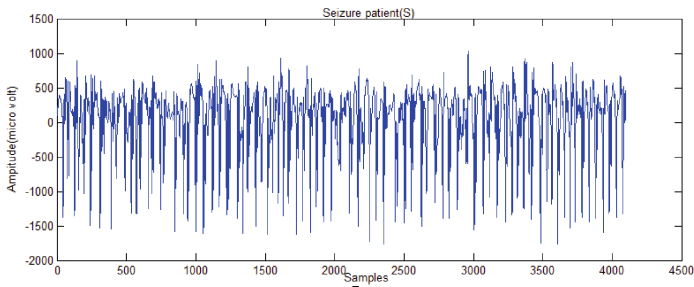**Fig. 4.** N (Seizure-free or Normal EEG) Patient Class

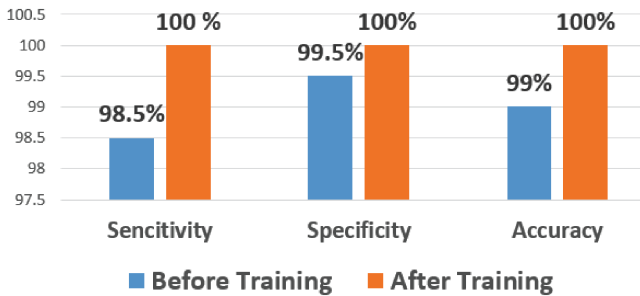**Fig. 5.** S (Seizure or Epileptic EEG) Patient Class

Fig. 6. Classification accuracy, sensitivity and specificity at before training and after 50% training for N (Seizure free) and S (Seizure Class)

classifier. Figures 4 and 5 show N (Seizure free-patient) and S (Seizure-patient) EEG signals, respectively. It is containing only seizure impulse. Here each dataset contain 100 time-series. Each signal contains 4097 samples (Fig. 6).

For all 100 EEG data sets, 50 data sets are used for training and the others are used for testing using SVM classifier. SVM classifier is used to classify unknown data properly. The highest accuracy is 100% for the feature set ApEn with SD for frame size $N = 1024$, sample value $m = 1$ and normalization ratio $r = 0.0$. In the proposed method accuracy achieved up to 100% for the feature set ApEn with SD. For training and testing purpose we get different accuracy, sensitivity and specificity as shown in graph.

Table 6. Comparison of methodology for same dataset

Methodology	Subset	Classification accuracy (%)
Permutation entropy (PE) and SVM classifier [14]	N and S	88.83
Empirical mode decomposition (EMD) [16]	N and S	95.33
Clustering and SVM classifier [18]	S and N	97.69
SODP and artificial neural network (ANN) classifier [15]	S and N	97.75
Empirical mode decomposition (EMD) and phase space representation (PSR) [19]	N and S	98.67

As shown in Table 6 all the papers are worked on Bonn dataset and they achieved maximum accuracy is 98.67%. In the proposed method accuracy achieved up to 100% for the feature set ApEn with SD.

4   Conclusion

We have extracted the features from the EEG signal and classification done using SVM classifier in to two class seizure and normal. Approximation entropy is extracted along with SD and mean. The randomness of EEG signal were extracted in the features, based on different size of frame (N), no. of samples values (m) and normalized ratio (r). From the set of features ApEn with SD, ApEn with mean, were used for classification using the SVM classifier. The highest classification accuracy is 100% for N and S class.

References

1. Lehnertz, K., Mormann, F., Kreuz, T., Andrzejak, R.G., Rieke, C., David, P., Elger, C.E.: Seizure prediction by nonlinear EEG analysis. IEE Eng. Med. Biol. Mag. **22**(1), 57–63 (2003). Article in IEEE engineering in medicine and biology magazine, Research gate January 2003
2. Mc Grogan, N.: Neural network detection of epileptic seizures in the electroencephalogram (1999). http://www.newox.ac.uk/~nmcgroga/work/transfer

3. Srinivasan, V.: Approximate entropy-based epileptic EEG detection using artificial neural networks. IEEE Trans. Inf. Technol. Biomed. **11**(3), 288–295 (2007)

4. Güler, N.F., Übeyli, E.D., Güler, I.: Recurrent neural networks employing Lyapunov exponents for EEG signal classification. Exp. Syst. Appl. **29**(3), 506–514 (2005)

5. Foster, I., Kesselman, C., Nick, J., Tuecke, S.: The physiology of the grid: an open grid services architecture for distributed systems integration. Technical report, Global Grid Forum (2002)

6. Übeyli, E.D.: Lyapunov exponents/probabilistic neural networks for analysis of EEG signals. Exp. Syst. Appl. **37**(2), 985–992 (2010)

7. Accardo, A., Affinito, M., Carrozzi, M., Bouquet, F.: Use of the fractal dimension for the analysis of electroencephalographic time series. Biol. Cybern. **77**(5), 339–350 (1997)

8. Liang, S.F., Wang, H.C., Chang, W.L.: Combination of EEG complexity and spectral analysis for epilepsy diagnosis and seizure detection. EURASIP J. Adv. Sig. Process. **2010** (2010). Article ID 853434

9. Pachori, R.B.: Discrimination between ictal and seizure-free EEG signals using empirical mode decomposition. Res. Lett. Sig. Process. **2008** (2008). Article ID 293056

10. Oweis, R.J., Abdulhay, E.W.: Seizure classification in EEG signals utilizing Hilbert-Huang transform. BioMed. Eng. OnLine **10**, 38 (2011)

11. Pachori, R.B., Bajaj, V.: Analysis of normal and epileptic seizure EEG signals using empirical mode decomposition. Comput. Meth. Programs Biomed. **104**(3), 373–381 (2011)

12. Bajaj, V., Pachori, R.B.: EEG signal classification using empirical mode decomposition and support vector machine. In: Proceedings of the International Conference on Soft Computing for Problem Solving, AISC 131, 20–22 December 2011, Roorkee, India, pp. 623–635 (2011)

13. Li, S., et al.: Feature extraction and recognition of ictal EEG using EMD and SVM. Comput. Biol. Med. **43**(7), 807–816 (2013)

14. Nicolaou, N., Georgiou, J.: Detection of epileptic electroencephalogram based on permutation entropy and support vector machines. Elsevier Trans. Exp. Syst. Appl. **39**, 202–209 (2012)

15. Pachori, R.B., Patidar, S.: Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode function. Elsevier Trans. Comput. Meth. Program. Biomed. **113**, 494–502 (2014)

16. Bajaj, V., Pachori, R.B.: Classification of seizure and non seizure EEG signals using empirical mode decomposition. IEEE Trans. Inf Technol. Biomed. **16**(6), 1135–1142 (2012)

17. Andrzejak, R.G., et al.: 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** (2001). Article ID 061907

18. Siulya, Li, Y., Wen, P.: Clustering technique-based least square support vector machine for EEG signal classification. Elsevier Trans. Comput. Meth. Program. Biomedic. **104**, 358–372 (2011)

19. Sharma, R., Pachori, R.B.: Classification of epileptic seizures in EEG signals based on phase space representation of intrinsic mode functions. Elsevier Trans. Exp. Syst. Appl. **12**, 1106–1117 (2015)