Biomedical Signal Processing and Control 39 (2018) 94–102

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction

Emina Alickovic [a], Jasmin Kevric [b], Abdulhamit Subasi [c,*]

[a] Linkoping University, Department of Electrical Engineering, Linkoping, 58183, Sweden
[b] International Burch University, Faculty of Engineering and Information Technologies, Francuske Revolucije bb. Ilidza, Sarajevo, 71000, Bosnia and Herzegovina
[c] Effat University, College of Engineering, Jeddah, 21478, Saudi Arabia

ARTICLE INFO

Article history:
Received 5 December 2015
Received in revised form 21 March 2017
Accepted 20 July 2017

Keywords:
Electroencephalography (EEG)
Epilepsy
Seizure detection and prediction
Multiscale PCA (MSPCA)
Discrete wavelet transform (DWT)
Empirical mode decomposition (EMD)
Wavelet packet decomposition (WPD)

ABSTRACT

This study proposes a new model which is fully specified for automated seizure onset detection and seizure onset prediction based on electroencephalography (EEG) measurements. We processed two archetypal EEG databases, Freiburg (intracranial EEG) and CHB-MIT (scalp EEG), to find if our model could outperform the state-of-the art models. Four key components define our model: (1) multiscale principal component analysis for EEG de-noising, (2) EEG signal decomposition using either empirical mode decomposition, discrete wavelet transform or wavelet packet decomposition, (3) statistical measures to extract relevant features, (4) machine learning algorithms. Our model achieved overall accuracy of 100% in ictal vs. inter-ictal EEG for both databases. In seizure onset prediction, it could discriminate between inter-ictal, pre-ictal, and ictal EEG with the accuracy of 99.77%, and between inter-ictal and pre-ictal EEG states with the accuracy of 99.70%. The proposed model is general and should prove applicable to other classification tasks including detection and prediction regarding bio-signals such as EMG and ECG.

1. Introduction

Epilepsy is a neurological disorder affecting over 50 million people worldwide. The archetypal modality for studying the human brain activity and brain-related disorders is electroencephalography (EEG). The need for an automated detection technique becomes more evident as there are no strong differences between seizure and seizure-free EEG recordings. From this vantage point, every third epileptic patient cannot be effectively cured by existing treatments, such as anti-epileptic drugs and surgeries. Patient's everyday activity is negatively influenced by the unpredictable nature of epileptic seizures, which, moreover, increases a risk of severe injuries. Thus, a patient's quality of life could be considerably improved if we can develop an effective alarm system for upcoming seizures [1]. This study suggests such (possible) system.

In order to evaluate the performance of such systems, the interval-based and segment-based paradigms are considered [2,3]. The former is characterized by sensitivity and false detection rate – FDR (or false prediction rate, FPR, in case of seizure prediction), whereas the latter is evaluated according to the sensitivity and specificity values. Sensitivity and specificity are expressed in percentages, while FDR (FPR) represents the number of false detections (predictions) per hour. Many interval-based approaches also suggest the latency as a measure. However, the development of seizure onset and termination detector is not an objective of the present study. In addition, the aim is not the development of interval-based seizure prediction system either. Therefore, the FPR criterion will not be used for the performance evaluation as in [4–6]. The objective of this study is the development of an effective segment-based approach for classifying EEG signals that can be utilized in designing the automated interval-based seizure prediction (or onset detection) systems.

Electrophysiological studies usually include EEG to monitor the neural (brain) responses. It should be noted that intracranial EEG (iEEG) produces brain signals of the better quality, but its less attractive side is its invasiveness. Contrary to iEEG, scalp EEG became more attractive, but potentially useful information may be lost due to the lower signal quality. This implies that it is reason-

* Corresponding author.
E-mail addresses: emina.alickovic@liu.se (E. Alickovic), jasmin.kevric@ibu.edu.ba (J. Kevric), absubasi@effatuniversity.edu.sa (A. Subasi).

http://dx.doi.org/10.1016/j.bspc.2017.07.022
1746-8094/

able to have a model for the seizure onset detection and prediction applicable to both modalities. We evaluate our models on both modalities: iEEG modality used in Freiburg data recordings and scalp EEG modality used in CHB-MIT data recordings.

The first seizure prediction approach using Freiburg database were designed by adjusting thresholds for particular features extracted from (intracranial) EEG segment, generating an alarm if the features violated an absolute or adaptive value [1]. In [7], a plethora of univariate and bivariate features were investigated for the use in threshold-optimized prediction methods. As various features hold discriminative information related to different cerebral states (inter-ictal, pre-ictal, ictal, and post-ictal), many machine learning algorithms were normally implemented to enhance seizure prediction rates [8–11]. In [8], six different types of neural network architectures were compared by using 14 features extracted from EEG of two patients to classify brain states into four classes: inter-ictal, pre-ictal, ictal and post-ictal. The accuracies of up to 99% were achieved. Tafreshi et al. [9] analyzed 19 patients from Freiburg database and achieved average success rate of 89.68% by combination of Empirical Mode Decomposition (EMD) features and AR model coefficients. Another but more successful approach using EMD features was presented in [11]. EMD features were combined with discrete cosine transformation (DCT) features and then classified by least square support vector machine (SVM) to achieve average accuracy of 99.1%. Aarabi & He [10] presented a rule-based seizure prediction system for focal neocortical epilepsy using 5 univariate and one bivariate feature to achieve sensitivity and specificity of 90.2% and 97% respectively.

The first machine learning approach has been developed by using the CHB-MIT database [12] is reported in [13]. The subject-oriented approach detected the onset of 96% of 173 test seizures in interval-based assessment, with latency of 3 s and false detection rate (FDR) of 2 false detections per hour. Other studies have been published [14,15] to improve the onset detection performance presented in [13]. On the other hand, there were some studies that tried to pull a mark of separation between seizure and seizure-free activity using CHB-MIT database. In [16], an automated epileptic seizure detection using wavelet based feature extraction technique is evaluated on 23 patients with 195 seizures with a 96.5% classification accuracy. A supervised machine learning method for seizure detection using multiple subject records is presented in [17]. A few subject-oriented seizure detection approaches developed on Freiburg EEG database have been discussed and explained in [18]. Specificity and sensitivity values above 90% were reported in majority of these studies. An efficient seizure detection approach was developed in [19], achieving specificity and sensitivity of 99.82% and 87.5% respectively. In addition, differential windowed variance (DWV) algorithm have been successfully combined in an automatic detection of seizure onset on Freiburg dataset in [20]. Sensitivity of 91.525%, average delay of 9.2 s after the onset, and FDR of 3/24 h were achieved. Eight novel empirical measures have been introduced to avoid false detections. Liu et al. [21] developed wavelet-based automatic seizure detection method with effective features and support vector machine for classification. A post-processing step was performed on the raw classification results to get more accurate results achieving a sensitivity of 94.46%, a specificity of 95.26%, and a FDR of 0.58/h for seizure detection in Freiburg EEG dataset.

The aforementioned studies suggest that an automated system for seizure onset detection and prediction can be designed. However, there is still room to investigate whether a different model could carry out seizure detection and prediction with higher performances in terms of statistical measures such as accuracy, sensitivity, and specificity. In addition, can shorter time interval (shorter segments) result in a comparable or a higher performance? Thus, the contribution of this study lies in the development of a model for seizure onset detection and prediction with very high confidence. This finding has implications for general design principles of epilepsy-based systems.

In order to cope with nonlinear and non-stationary signals, such as EEG, the classical frequency methods have rather strict restrictions. Therefore, time–frequency techniques have been developed to eliminate these restrictions. Such techniques for signal decomposition include Empirical Mode Decomposition (EMD), Discrete Wavelet Transform (DWT), and Wavelet Packet Decomposition (WPD). The model suggested in the present study consist of four modules: (1) multiscale principal component analysis (MSPCA) to remove artefact contaminated parts from EEG measurements, (2) three different decomposition methods (EMD, DWT and WPD) to find the most suitable set of frequency bands, (3) statistical values (lower and higher order statistics) to extract the relevant features from EEG frequency bands decomposed with EMD, DWT and WPD, and (4) machine learning methods (classifiers) to discriminate between different states (inter-ictal, pre-ictal and ictal). The rationale to select MSPCA for artefact removal (de-noising) is that its proven superiority when applied to different biomedical signals, such as ECG [22–24], EMG [25], EEG [18]. The rationale to select the suggested three decomposition methods (EMD, DWT and WPD) is plethora of their application in the different fields. The rationale to extract statistical features is to capture important information while keeping the low data dimensions. The selected classifiers are well-known classifiers with wide range of applications. We checked the aforementioned module combinations with four machine learning techniques to find the best system for seizure detection and prediction.

Hence, the aim of this study is to develop a segment-based system for classification of EEG signals that can be applied in automated interval-based seizure prediction (or onset detection) systems by using two omnipresent and archetypal EEG databases: Freiburg (iEEG) and CHB-MIT (scalp EEG). Our findings clearly indicate that the models suggested in the preset study are suitable for automated seizure onset detection and prediction systems.

The rest of this article is organized in the following way. Section 2 provides the materials and methods employed in this study. It explains databases used in this study, de-noising module, feature extraction and dimension reduction methods. In Section 3, EEG signal classification methods are shortly explained. The experimental results are presented in Section 4, whereas Section 5 concludes the paper.

2. Materials and methods

2.1. Experimental setup

2.1.1. Freiburg and Physionet CHB-MIT EEG databases

Freiburg EEG data was recorded at the Epilepsy Center of the University Hospital of Freiburg during the period of invasive presurgical epilepsy monitoring. The Freiburg EEG database is composed of invasive EEG recordings of 21 patients suffering from medically intractable focal epilepsy. Neurofile NT digital video EEG system with 128 channels, 256 Hz sampling rate, and a 16 bit analogue-to-digital converter was used to sample the EEG data. Each patient had between two to five seizures and at least 24 h of seizure-free (inter-ictal) recordings. Every patient's data is organized into "ictal" and "inter-ictal" datasets. The former contains seizure files and at least 50 min of pre-ictal data, whereas the latter holds one day of seizure-free EEG-recordings [26].

CHB-MIT Dataset consists of 23 different subsets containing EEG records from 22 different pediatric patients. This dataset contains 182 seizures. Generally, each of these digitized records is one hour long. Sampling frequency is 256 Hz with 16-bit resolution. All seizure starts and ends were confirmed by board-certified electroencephalographer who analyzed all EEG records manually [27,28].

2.1.2. EEG database preparation

We used a rectangular window with length of 2048 samples (=8 s) for extraction of EEG segments from both datasets ("inter-ictal" and "ictal") of each subject. Approximately two segments per hour were extracted from inter-ictal dataset, producing 1200 inter-ictal segments. As Freiburg EEG inter-ictal dataset is provided as a set of 1-h-long EEG segments, we extracted approximately two 8-s-long EEG segments from the 1-h-long segment. Since the onset times of each seizure were known, a minimum number of rectangular windows was used to cover all 87 seizure activities, producing 1200 ictal segments. No overlapping was implemented while extracting ictal segments. Since every seizure includes at least 50 min of pre-ictal data, we extracted one 8-s long EEG segment to produce 1200 pre-ictal segments. Finally, we created our own dataset containing 3600 EEG segments having length of 8 s and belonging to three different classes: inter-ictal, pre-ictal and ictal, what allowed us to assess seizure detection and prediction models. Similar procedure was followed to create our own dataset for CHB-MIT database. In total 1000 inter-ictal, 1000 ictal and 1000 pre-ictal seizures segments were extracted. 8-s long pre-ictal segments were extracted for the period between 30 min and 15 min prior to each seizure onset.

Apart from the 3-class dataset setup, it may be useful to prepare two more datasets that carry only two classes. Besides inter-ictal EEG segments, one of those datasets holds ictal EEG, while the other one includes pre-ictal EEG segments. This approach makes the former dataset suitable for the seizure detection, whereas the latter is appropriate for seizure prediction task. Different window sizes in the analysis of EEG signals were used, usually, not going with more than 4000 samples. We have chosen the window size of 2048 samples or 8 s as it has shown better performance than shorter windows.

2.2. Module 1: multiscale principal component analysis (MSPCA)

Principal Component Analysis (PCA) is a method to change input matrix $I_{n\times m}$ into $I_{n\times m} = LS^T$, where L are PC loadings, S are PC scores, n and m are the number of measurements and variables, by relating the variables as linear weighted sums. The advantage of the PCA is its ability to extract necessary information from input data and to decorrelate variables based on extracted information. Wavelets are used to extract deterministic features from stochastic processes and approximately decorrelate autocorrelation between signals. In this work, I is the input matrix containing EEG data, n is the number of observations (3600 for Freiburg dataset and 3000 for CHB-MIT dataset), m is the number of variables (2048 for both datasets). MSPCA combines these advantages of PCA and wavelets. MSPCA calculates PCA of all wavelet coefficients for all wavelet decomposition levels and after that it relates the outcomes at adequate levels [29,25,18,22,23]. MSPCA is used for denoising and the procedure is given in [29].

The usage of MSPCA may not be clear at first. But we know that EEG recordings may contain "noisy" data, i.e., artefact-contaminated data, which is one of the leading obstacles in EEG data analysis. A standard approach is to apply multichannel analysis techniques such as independent component analysis (ICA). In our previous study on EEG data [18], we have proven the superiority of MSPCA in isolating noise components. This finding conjecture, therefore, suggests the usage of MSPCA in de-noising stage of the model. Thus, we now use MSPCA to remove noise (artefact-contaminated) components from EEG signals [30].

2.3. Module 2: EEG decomposition methods

2.3.1. Empirical mode decomposition (EMD)

Empirical Mode Decomposition (EMD) is a signal decomposition method to decompose a non-linear and non-stationary signal into its oscillatory modes, called intrinsic mode functions (IMFs). Every IMF satisfies two conditions [31]:

- The difference between the total number of extrema (local minima and local maxima) and total number of zero-crossings is zero or one;
- At any point, the mean value of the envelope defined by the local maxima and the envelope defined by the local minima is zero or very close to zero [32].

2.3.2. Discrete wavelet transform (DWT)

Wavelets are an effective time-frequency tools for the analysis of non-stationary signals. Discrete Wavelet Transform (DWT) is a procedure *for the decomposition of input signal $u(t)$($u(t)$ is EEG signal in this work)* into sets of function, called wavelets, by scaling and shifting of mother wavelet function. Consequently, the decomposition i.e. set of wavelet coefficients are formed. To accomplish this, the signal $u(t)$ can be reconstructed as linear combination of wavelets and weighting wavelet coefficients. To correctly reconstruct the signal $u(t)$, it is necessary to obtain sufficient number of wavelet coefficients during DWT decomposition [33]. After every decomposition into approximation coefficients A and detailed coefficients D, time resolution is halved and frequency resolution is doubled. Hence, we formed different sub-bands of the EEG signal which are approximation coefficients A and detail coefficients D of the EEG signal.

2.3.3. Wavelet packet decomposition (WPD)

The wavelet packet decomposition (WPD) extends the capabilities of the DWT. Whereas DWT decomposes the approximations records only, WPD does the decomposition of both approximation and detail records into sublevels [34–36]. Alternatively, WPD can be thought of as a continuous-time wavelet decomposition sampled at various frequencies at each scale level. Hence, WPD delivers better frequency resolution for the signal being decomposed. Another benefit of the WPD is that it represents the reconstruction of the original signal by combining various decomposition levels [37]. If WPD has k levels, size of different set of wavelet coefficient will be 2^k, while in DWT it will be $k+1$. Daubechies4 (db4) mother wavelet function is used for WPD in this study, whereas the number of decomposition levels is chosen to be 4.

2.4. Module 3: feature extraction and selection

In this study six different statistical features were selected for EEG classification in three different domains, i.e., EMD, DWT and WPD, aiming at decreasing the dimensionality of data set(s) [38–40]. The rationale to use signal statistics or to extract statistical features is to capture important information while keeping the low data dimensions. These statistical features (SFs) are:

1) Mean of coefficients' absolute values in every sub-band,
2) Average power of the coefficients in every sub-band,
3) Standard deviation of the coefficients in every sub-band,
4) Ratio of absolute mean values of adjacent sub-bands;
5) Skewness in every sub-band;
6) Kurtosis in every sub-band;

In this study, based on previous works done before and our experience, three different experiment setups were conducted. In the first experiment, SF 1, 2, 3, and 4 were used as to classifier and set of these features is called **F1**. In the second experiment, SF 1, 3, 5, and 6 were used as inputs for classifiers and set of these feature is called **F2**. In the third experiment, all statistical features mentioned previously were used as inputs for classifier and set of these feature is referred to as **F3**. Performances of these setups were compared. *In brief, F1 consists of SF1, SF2, SF3 and SF4, F2 consists of SF1, SF3, SF5 and SF6 and F3 consists of SF1-SF6 (all features).* For this purpose, the broad study is implemented by means of detailed and point of reference EEG signal database, where the values of these statistical features were computed for a sizeable quantity of EEG segments in DWT, WPD and EMD domains.

As the sampling rates for both databases are equal to 256 Hz, five level-decomposition was used in MSPCA to generate detail and approximation signals. The Kaiser rule was responsible for eliminating the number of loadings [18]. After the de-noising (artefact removal) procedure with MSPCA algorithm, discriminative features were extracted from these “cleaned” EEG signals. Prior to feature reduction, de-noised EEG signals are decomposed into sub-band signals using DWT, WPD, and EMD methods. Daubechies4 (db4) mother wavelet function was used to decompose the de-noised EEG signals using both DWT and WPD methods. The number of decomposition levels in DWT was selected to be 6, whereas that the same parameters in WPD was chosen as 4. This combination of parameters resulted in $k + 1 = 6 + 1 = 7$ sub-band signals for DWT, and $2^k = 2^4 = 16$ sub-band signals for WPD. In EMD, we were always able to generate at least 7 IMFs. Counting residue signal, every de-noised EEG signals was decomposed into 8 signals. Although Residue represents the low-pass component and IMF1 holds highest frequencies (IMF2–IMF7 are band-pass signals in between), the exact frequency ranges cannot be given as in DWT and WPD cases.

Totally, six different statistical features were extracted from sub-band signals generated from EEG signals using DWT, WPD, and EMD. It means that every EEG signal is characterized by a feature vector containing $7 \times 6 = 42$ features for DWT, $16 \times 6 = 96$ features for WPD, and $8 \times 6 = 48$ features for EMD. The case when all six features were used is denoted as F3. We also created additional two cases where we picked 4 out of 6 features and designated them as F1 and F2.

The selection of scale levels for DWT and WPD was predominantly based on the overall number of extracted features in the resultant feature vector. It would seem more appropriate to choose 3 levels in WPD algorithm which would result in 8 sub-bands and be comparable to other two methods. However, opting for 4 levels did considerably improve the performance in our testing, although the number of features was doubled. Going for 5 levels in WPD would make the resulting feature vector four times bigger than in other two methods, making the comparison between them quite unfair.

In addition, even the selection of more decomposition levels in DWT did noticeably increase the accuracy of our systems. The choice of 7 levels would match 7 IMFs produced by the EMD method but would also result in unnecessary decomposition of the approximation signal.

2.5. Module 4: EEG signal classification

2.5.1. Random forest (RF)

Random Forest (RF), proposed by Breiman [41], is novel, fast, highly accurate, noise resistant classification method. Bagging and random feature selection are combined together in RF. Every tree in the forest is influenced by the values of random vectors sampled separately and has identical distribution as any other tree in the forest [41]. RF consists of massive number of decision trees where decision tree selects their separating features from bootstrap training set S_i where i represent ith internal node. Trees in RF are grown by means of Classification and Regression Tree (CART) method with no pruning. As number of trees in the forest turns into massive number, generalization error will also increase until it converges to some boundary level [41]. More details about RF can be found in [41]. In this study, the number of trees was set to 10.

2.5.2. Support vector machine (SVM)

Support vector machine (SVM) technique was first proposed by [42] and [43]. This technique founded wide range of application in different areas where classification, regression, prediction, estimation and forecasting is to be performed due to its excellent performance results. SVM is based on Vapnik– Chervonenkis (VC) theory and Structural Risk Minimization (SRM) principle. It intends to determine the minimizing training set error by maximizing the boundary among separating hyper-plane and the data. Usage of convex quadratic programming is one of the major advantages of SVM, as convex quadratic programing results in the global minima only (cannot be trapped in local minima) [43,22,40]. Different kernel functions, such as RBF kernel function, polynomial kernel function, normalized polynomial kernel function [44,22,39], are proposed in the broad literature on SVM. The value for complexity constant c was set to 100, whereas the values of σ and γ parameters of Puk kernel were selected as 1.

2.5.3. Multilayer perceptron (MLP)

Artificial neural network (ANN) is computational emulation of human nervous system consisting of vast amount of simple, extremely interconnected processing components, known as neurons. All neurons are interconnected with synapses having variable weights. Due to this fact, ANN can be understood as parallel distributed processing system. Multilayer Perceptron (MLP) is a noteworthy type of ANN consisting of input layer, at least one hidden layer and output layer. EEG input signal is transmitted across the MLP neural network in onward direction, on layer-by-layer base. All processing of the data from input layer is done in the hidden layer and transmitted to output layer. Number of neurons in the hidden layer depends on data being studied. Lacking or extreme number of neurons in the hidden layer can initiate drawbacks with generalization and overfitting. Each neuron results in output based on transfer function, called activation function, which is linear or nonlinear mapping of input to output [45,22]. In this paper, the model with two hidden units was used, and the tolerance was set to 10^{-6}.

2.5.4. K-Nearest neighbor (k-NN)

The k-nearest neighbor machine learning tool can be used for the classification task as well. Sample's nearest neighbors are expressed in the form of Euclidean distance. If it is assumed that random sample w is described with feature vector as $\langle q_1(w), q_2(w), \ldots, q_k(w)\rangle$ where $q_i(x)$ stands for ith value of sample w. Usually, all features vectors are normalized to have zero mean and variance equal to 1 because possible scenario is that these feature vectors are gauged in unequal units. Then, the distance between two samples w_a and w_b is expressed in the form of different distance measures types, such as Euclidean distance, Chebyshev distance, Manhattan distance etc. In this study, Euclidean distance measure, which is most commonly used measure type, was employed. More information about k-NN can be found in [46,47]. 1-NN algorithm (1 nearest neighbor) with Euclidean distance measure was used in this research.

3. Results and discussion

The first driving goal of this study is the development of segment-based model for classification of EEG signals that can be applied in automated interval-based seizure prediction (or onset detection) systems by using two archetypal EEG databases: Freiburg and CHB-MIT. However, all recordings were not employed in training and testing phases of our proposed approach. Consequently, shorter EEG segments were extracted from these databases, which can open the opportunity for development of the practical and computationally efficient seizure detection and prediction systems.

The pre-ictal and ictal recordings constitute small amount of data when compared to the inter-ictal recordings. Therefore, a classification system developed using this unbalanced dataset would report high total accuracy (due to good prediction of inter-ictal class) even in case of low prediction of preictal and inter-ictal class [48]. Down-sampling the class with larger proportions represents one paradigm for circumventing the class unbalance [49]. In this paper, segments from the pre-ictal and inter-ictal data were randomly chosen to produce comparable number of ictal, inter-ictal, and pre-ictal segments.

3.1. Performance evaluation criteria

In this study, different machine learning methods were employed for EEG signal classification. In order to determine performances of suggested approaches, 10-fold cross-validation (CV) method is used.

In 10-fold CV, dataset is randomly separated into 10 different mutually exclusive folds having the same sizes. Nine (9) folds are used for training and remaining one (1) fold is used for testing. No fold is used for validation purposes. This procedure is iterated 10 times. At the end of each iteration, individual accuracy is computed. The average of 10 obtained individual accuracies is CV accuracy.

Overall accuracy (or CV accuracy) is used as the performance evaluation criteria for 3-class dataset. For the two-class datasets, the performance of the classification task can also be expressed using sensitivity and specificity values. Specificity shows the accuracy of a classifier when recognizing inter-ictal EEG, or it can serve as a measure of false detections/predictions. However, sensitivity represents the accuracy of a classifier in detecting the class of interest. The class of interest in seizure detection task is the ictal class, whereas in seizure prediction this class is pre-ictal class.

The aforementioned three performance evaluation criteria measures can also be expressed by the number of true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN) as:

$$\mathit{Sensitivity} = \frac{\mathrm{TP}}{\mathrm{TP}+\mathrm{FN}} \times 100\% \tag{1}$$

$$\mathit{Specificity} = \frac{\mathrm{TN}}{\mathrm{TN}+\mathrm{FP}} \times 100\% \tag{2}$$

$$\mathit{Accuracy} = \frac{\mathrm{TP}+\mathrm{TN}}{\mathrm{TP}+\mathrm{FN}+\mathrm{TN}+\mathrm{FP}} \times 100\% \tag{3}$$

3.2. Experimental results

The aim of this study is to develop segment-based model for the classification of EEG signals that can be applied for automated interval-based seizure prediction (or onset detection) systems by using two popular EEG databases. Hence, the contribution of this study is to find a best model by comparing three different widely used feature extraction methods namely EMD, DWT and WPD; six different statistical features for dimension reduction and four commonly used machine learning methods namely ANN, k-NN, SVM and RF.

In order to compare the performance results of the feature extraction methods average accuracy was calculated and the results of all experiments are summarized in Tables 1–6. The values for parameters of these classification techniques were selected as explained in Section II because they produced the best and robust results in our testing. Tables 1–6 provide average accuracy results only, whereas *sensitivity* and *specificity* values are stated through-

Table 1
Overall accuracy (%) using Freiburg EEG database.

Three-class case				
Signal Process	Predict. Tools	F1	F2	F3
EMD	k-NN	63.6	55.7	60.9
	MLP	83.8	83.5	85.6
	SVM	84.1	78.3	80.3
	RF	88.2	88.7	90.4
DWT	k-NN	87.1	79.8	84.1
	MLP	86.1	83.5	93.3
	SVM	96.8	86.1	95.1
	RF	98.2	96.4	98.4
WPD	k-NN	97.5	93.8	98.1
	MLP	82.0	92.1	95.5
	SVM	99.61	96.9	99.4
	RF	99.4	99.1	**99.66**

Table 2
Overall accuracy (%) using CHB-MIT EEG database.

Three-class case				
Signal Process	Predict. Tools	F1	F2	F3
EMD	k-NN	83,97	80,13	87,2
	MLP	86,97	83,37	89,67
	SVM	91,6	85,1	91,43
	RF	89,27	92,5	92,83
DWT	k-NN	98,6	96,4	98,6
	MLP	99,2	98,9	99,33
	SVM	99,4	99,0	99,4
	RF	98,93	98,53	99,4
WPD	k-NN	99,4	97,4	99,2
	MLP	99,7	99,53	99,7
	SVM	**99,77**	99,17	99,67
	RF	99,6	99,1	99,4

Table 3
Overall accuracy (%) for seizure detection using Freiburg EEG database.

Seizure detection				
Signal Process	Predict. Tools	F1	F2	F3
EMD	k-NN	75.1	72.0	74.5
	MLP	93.7	93.5	95.5
	SVM	92.4	87.7	90
	RF	94.3	95.0	96.3
DWT	k-NN	99.95	99.95	99.95
	MLP	99.95	99.91	99.91
	SVM	99.95	99.95	99.95
	RF	99.95	99.95	99.95
WPD	k-NN	100	100	100
	MLP	99.95	100	99.95
	SVM	99.95	99.95	99.95
	RF	99.95	100	100

Table 4
Overall performance results (%) for seizure detection using CHB-MIT EEG database.

Seizure detection				
Signal Process	Predict. Tools	F1	F2	F3
EMD	k-NN	89,35	82,55	94,90
	MLP	94,55	90,00	96,90
	SVM	95,35	91,40	97,50
	RF	93,75	96,55	96,90
DWT	k-NN	100,00	100,00	100,00
	MLP	100,00	100,00	100,00
	SVM	100,00	100,00	100,00
	RF	100,00	100,00	100,00
WPD	k-NN	100,00	100,00	100,00
	MLP	100,00	100,00	100,00
	SVM	100,00	100,00	100,00
	RF	100,00	100,00	100,00

Table 5
Overall accuracy (%) for seizure prediction using Freiburg EEG database.

Seizure prediction				
Signal Process	Predict. Tools	F1	F2	F3
EMD	k-NN	71.2	63.3	73.29
	MLP	87.2	85.1	88.29
	SVM	83.8	80.4	85.0
	RF	90.8	89.6	89.4
DWT	k-NN	75.5	69.9	80.3
	MLP	88.5	86.7	94.8
	SVM	92.6	79.6	95.2
	RF	97.5	94.7	97.7
WPD	k-NN	97.2	90.7	96.0
	MLP	98.2	98.2	98.8
	SVM	99.0	95.5	99.3
	RF	**99.5**	98.8	99.4

Table 6
Overall performance results (%) for seizure prediction using CHB-MIT EEG database.

Seizure prediction				
Signal Process	Predict. Tools	F1	F2	F3
EMD	k-NN	88,65	93,30	96,45
	MLP	94,80	95,95	97,55
	SVM	97,05	97,55	98,35
	RF	96,30	96,80	96,35
DWT	k-NN	98,20	94,30	98,40
	MLP	98,80	98,65	99,20
	SVM	99,20	98,70	99,40
	RF	98,40	97,90	98,40
WPD	k-NN	98,50	95,55	98,40
	MLP	99,65	99,20	99,60
	SVM	**99,70**	99,40	**99,70**
	RF	99,30	98,15	99,15

out the analysis and discussion of the experimental results. From these tables, it can be seen that WPD was superior when compared to DWT and much better once compared to EMD.

3.2.1. System performance using all three classes

Overall results for classification of EEG signals into inter-ictal, pre-ictal and ictal class for the two datasets are shown in Tables 1 and 2. We can see that the WPD outperformed other two feature extraction methods, i.e. DWT and EMD. For both databases, RF and SVM classifiers resulted in high classification accuracy rates (difference between highest results for RF and SVM is only 0.05%). Results obtained in this study show that inclusion of all features does not necessarily improve classification accuracy rates. As it can be seen from Table 1, even when less number of features are used, drop in classification accuracy rate is negligible, from 99.66% to 99.1% for RF classifier. Classification accuracy of our proposed system (MSPCA + WPD + F3 + RF) with three classes (inter-ictal, pre-ictal and ictal periods) for Freiburg dataset is 99.66%.

In the case of children epilepsy, SVM resulted in slightly higher results when compared to other three classifiers, as it can be seen from Table 2. It gave higher classification accuracy rate of 99.77% when F1 features are used. For this database, there is no important change between three experiments where different features were employed. For classification of children EEG signals the proposed model is MSPCA + WPD + F1 + SVM.

3.2.2. Seizure detection (ictal vs inter-ictal)

After all performance evaluations were completed having three different feature extraction methods, three different feature set combinations and four different classifiers, the finest detector by its closeness to the ideal performance was selected (100% of sensitivity, specificity and accuracy). In the seizure detection system, this study achieves remarkable rates of 100% for sensitivity, specificity and accuracy. This is indeed an astonishing performance resulted from WPD feature extractor.

Table 3 illustrates the performance results of seizure detection task using Freiburg dataset. From Table 3, it can be seen that WPD gave rather high classification results when compared to DWT and EMD for Freiburg dataset. Furthermore, it can be seen that, k-NN results in 100% sensitivity, specificity and accuracy rates for all three different sets of features selected, but RF and MLP resulted in 100% sensitivity, specificity and accuracy rates when different feature sets were employed. For seizure detection task in Freiburg dataset, we propose MSPCA + WPD + F1(or F2 or F3) + k-NN since this system can also perform classification task with highest accuracy. Based on achieved outcomes, it is easy to conclude that the suggested system is robust and reliable.

Table 4 illustrates the performance results of seizure detection task using CHB-MIT dataset. It can be seen from Table 4 that, DWT and WPD gave 100% classification accuracy with all classifiers for CHB-MIT dataset. For seizure detection task in CHB-MIT dataset, all classifiers achieve 100% accuracy with DWT and WPD, hence it can be easily concluded that the suggested system is reliable and accurate.

3.2.3. Seizure prediction (pre-ictal vs inter-ictal)

The experimental results for the seizure prediction task are presented in Tables 5 and 6. From these tables, it can be seen that the best model is based on WPD since WPD outperformed DWT and EMD. Concerning different classifier methods selected for Freiburg database, RF resulted in highest performances when F3 was selected as feature set with sensitivity 99.92%, specificity 99.17%, and accuracy 99.5%. The proposed system for automated seizure prediction is MSPCA + WPD + F1 + RF.

Concerning different classifiers selected for CHB-MIT database, it can be seen from Table 6 that SVM resulted in slightly higher performance compared to other three classifiers when F1 or F3 were selected as feature set. The sensitivity is 99.6%, the specificity is 99.8% and the accuracy is 99.7% for F1 and the sensitivity is 99.8%, the specificity is 99.6% and the accuracy is 99.7% for F3. Since F1 is smaller set and hence computation time is less, and the proposed system for automated seizure prediction in children seizure prediction is MSPCA + WPD + F1 + SVM.

3.3. Discussion

Let us first discuss the performance of EMD method which lags behind WPD and DWT in all assessments and configurations. Using EMD technique, we were always able to generate at least 7 IMFs from all EEG segments of both databases used in this paper. On the other hand, some EEG segments generated more than 10 IMFs, raising a question of which frequency ranges have been occupied by each IMF. In other words, does the first IMF (which represents the highest frequency component) comprise the same frequencies for all EEG segments? The same inquiry applies for other IMFs also.

It becomes even more interesting when we compare the power spectral densities (PSD) of IMF1 from two EEG segments: one generating only 7 IMFs, and the other having 10 IMFs (Fig. 1). Although the two PSDs are very similar, the first IMF excludes frequencies beyond 60 Hz. Since the sampling frequency of EEG signals used is 256 Hz (thus the highest captured frequency of the signal is 128 Hz), IMF1 omits very significant portion of these EEG signals.

Comparatively, the highest frequency component after the application of DWT is the detail D1 which contains the frequency range of 64–128 Hz, shown also in Fig. 2. The same figure represents the difference between the highest frequency signals generated by EMD and DWT. Although the maximum and minimum value of PSD for IMF1 and D1 differ, it is obvious that EMD ignores a component of the EEG signal which is very pronounced by D1. After the

Table 7
Comparison table for seizure detection using Freiburg EEG database.

Authors	Signal Processing and Features	Sens. (%)	Spec. (%)	Acc. (%)
Raghunathan et al. [19]	Wavelet filter + Coastline + Variance	87.5	99.82	x
Aarabi et al. [3]	Entropy + Dom. Frequency + Avg. amplitude + Rhythmicity	68.9	97.8	x
Yuan et al. [2]	Fractal intercept + relative fluctuation index	91.72	94.89	x
Patnaik et al. [54]	Statistical parameters from the DWT coefficients	91.29	99.19	x
Tafreshi et al. [51]	EMD + Hilbert Transform + wavelet	x	x	95.42
Kevric & Subasi [18]	MSPCA + Eigenvector Pseudospectrum	99.8	99.4	99.59
This work	MSPCA + WPD + statistics	100	100	100

Table 8
Comparison table for seizure detection using CHB-MIT EEG database.

Authors	Signal Processing and Features	Sens. (%)	Spec. (%)	Acc. (%)
Fergus et al. [17]	Band-pass filter + Peak and Median Frequency, RMS, Entropy + LDA	88	88	x
Raffiudin et al. [16]	Wavelet-based features + statistics + IQR + MAD	x	x	96.5
This work	MSPCA + WPD + statistics	100	100	100
	MSPCA + DWT + statistics			

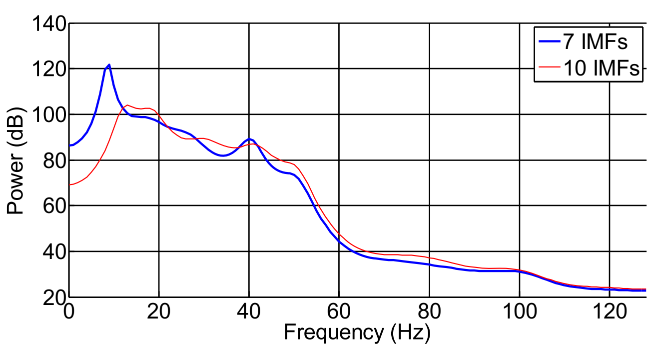

Fig. 1. PSD for IMF1 using Eigen decomposition method with 10 complex sinusoids for two inter-ictal EEG segments with different number of generated IMFs.

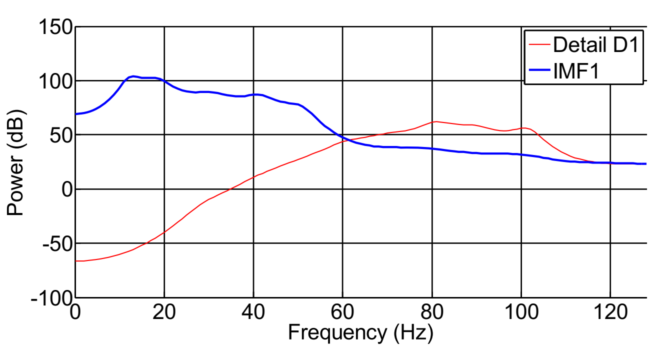

Fig. 2. PSD for IMF1 and D1 using Eigen decomposition method with 10 complex sinusoids for the inter-ictal EEG segment with 10 generated IMFs.

application of WPD with 4 levels, the detail D1 is shared amongst 8 different sub-bands generated by WPD. Extracting features from these bands must have been decisive as WPD outperformed DWT in almost all scenarios.

Another important difference between EMD and wavelet based methods lies in the process of feature extraction. Namely, the product of EMD methods are IMF signals of the same length as the original EEG signals (no reduction in the number of samples/features occurred). On the other hand, the direct product of wavelet-based methods (DWT and WPD) are the coefficients characterizing different sub-bands, whose number is much less than the number of samples of EEG signals. In the EMD case, the six features were extracted from the IMF signals, whereas in wavelet-based methods we calculated the features from the resulting coefficients directly.

Among three different scenarios (seizure detection, seizure prediction, and the case with the three classes), seizure detection represents the easiest task indicated by total accuracies of 100% in Tables 3 and 4 in several assessments using wavelet-based methods. Even in case with the three classes, the detection of ictal events was 100%, and none of the inter-ictal and pre-ictal samples were recognized as being ictal. These results suggest that ictal events are distinguishable from non-ictal events (inter-ictal and preictal) due to the spiky nature of ictal EEG and contribution of noise removal carried out by MSPCA.

Tables 7 and 8 show that our findings are one of the first that achieve overall accuracy of 100% in classifying EEG signals for epileptic seizure detection using EEG segments from one large database rather than [50]. Even the features extracted from IMFs by EMD resulted in very high accuracy for seizure detection: 96.3% (Freiburg database), and 97.5% (CHB-MIT database) using all six features, the highest results currently found in literature using Freiburg database [51], and even outperforming an approach developed on small EEG dataset [52]. Table 9 shows that our proposed prediction system outperformed previous systems found in up-to-date studies.

When comparing the highest obtained results between the seizure prediction and three-class case, we can conclude that introducing the ictal class did not significantly affect the overall accuracy. The drop in total accuracy for CHB-MIT database was from 99.77 to 99.70%, whereas the same drop for Freiburg database was from 99.66 to 99.5%. In other words, the capability of recognizing pre-ictal and inter-ictal events remained rather constant regardless of ictal class taking the part in developing the classifier. Since the interpretation of ictal events is not highly necessary when predicting the upcoming seizures, only inter-ictal and pre-ictal data can be considered when developing epileptic seizure predictors.

The following contributions and originalities can be highlighted based on the results of our work:

- The effective classification of challenging ictal, inter-ictal, and pre-ictal EEG signals has been achieved regardless of seizure type, epileptic focus, patient's age, and EEG acquisition methods used (scalp or intracranial electrodes),
- High frequency ranges of EEG signals hold discriminative features for separating ictal, inter-ictal, and pre-ictal events,
- One way of improving the performance of EMD technique, especially for seizure prediction, is to utilize Ensemble EMD method, where the number of IMFs is determined by the length (number of samples) of the signal [53].

Table 9
Comparison table for seizure prediction.

Authors	Database	Signal Processing and Features	Sens. (%)	Spec. (%)	Acc. (%)
Aarabi & He [10]	Freiburg	Combination of 5 univariate and one bivariate measure	90.2	97	x
Parvez et al. [11]	Freiburg	EMD + Direct Cosine Transform	X	x	99.1
Tafreshi et al. [9]	Freiburg	AR + EMD	X	x	89.68
Costa et al. [8]	Freiburg	14 features: different energies, correlation, Lyapunov exponent	X	x	99
This work	Freiburg	MSPCA + WPD + statistics	99.91	99.16	99.5
Qidwai et al. [55]	CHB-MIT	Variance + Entropy	x	x	90
This work	CHB-MIT	MSPCA + WPD + statistics	99.6	99.8	99.7

- The proposed approaches based on MSPCA de-noising method and WPD decomposition techniques combined with RF or SVM classifier can now be utilized in interval-based seizure prediction (or seizure onset detection) system.

4. Conclusion

The aim of this study is to develop a segment-based model for the classification of EEG signals that can be applied in an automated interval-based seizure prediction (or onset detection) systems by using two popular large EEG databases: Freiburg and CHB-MIT. Furthermore, two large benchmark databases were used to prove efficiency of our proposed seizure detection and prediction algorithms. A broad study was conducted by employing WPD, DWT and EMD feature extraction methods to examine the ability of proposed algorithms to detect and predict seizure automatically. Four different classifiers were employed to test the efficiency of three proposed feature extraction methods. As a feature extractor, we propose WPD since it outperformed other two methods in many cases. In terms of the classifiers, we propose RF and SVM since these two classifiers results in very high classification accuracies. For classification of EEG signal into inter-ictal, pre-ictal and ictal for adults, we propose system composed of MSPCA + WPD + RF (overall classification accuracy is 99, 66%) and for children we propose similar system (only difference is classifier) MSPCA + WPD + SVM (overall accuracy is 99.77). For seizure detection task, we propose a system composed of MSPCA + WPD + RF and this system results in 100% sensitivity, specificity and accuracy rates. For seizure prediction for adults, we propose system as MSPCA + WPD + RF and this system results in 99,5% accuracy rate and for seizure detection for children, we propose system MSPCA + WPD + SVM and this system achieves accuracy of 99, 7%. Experimental results demonstrate that our proposed systems outperform the current proposed methods and systems for seizure detection and prediction found in up-to-date literature in terms of sensitivity, specificity and accuracy.

References

[1] B. Schelter, M. Winterhalder, T. Maiwald, A. Brandt, A. Schad, J. Timmer, A. Schulze-Bonhage, Do false predictions of seizures depend on the state of vigilance? a report from two seizure-Prediction methods and proposed remedies, Epilepsia 47 (12) (2006) 2058–2070.
[2] Q. Yuan, W. Zhou, Y. Liu, J. Wang, Epileptic seizure detection with linear and nonlinear features, Epilepsy Behav. 24 (2012) 415–421.
[3] A. Aarabi, R. Fazel-Rezai, Y. Aghakhani, A fuzzy rule-based system for epileptic seizure detection in intracranial EEG, Clin. Neurophysiol. 120 (2009) 1648–1657.
[4] P. Mirowski, D. Madhavan, Y. LeCun, R. Kuzniecky, Classification of patterns of EEG synchronization for seizure prediction, Clin. Neurophysiol. 120 (11) (2009) 1927–1940.
[5] L. Chisci, A. Mavino, G. Perferi, M. Sciandrone, C. Anile, G. Colicchio, F. Fuggetta, Real-time epileptic seizure prediction using AR models and support vector machines, IEEE Trans. Biomed. Eng. 57 (no. 5) (2010).
[6] Y. Park, L. Luo, K. Parhi, T. Netoff, Seizure prediction with spectral power of EEG using cost-sensitive support vector machines, Epilepsia (2011) 1761–1770.
[7] F. Mormann, R.G. Andrzejak, C.E. Elger, K. Lehnertz, Seizure prediction: the long and winding road, Brain 130 (2007) 314–333.
[8] R.P. Costa, P. Oliveira, G. Rodrigues, B. Leitao, A. Dourado, Epileptic seizure classification using neural networks with 14 features, in: I. Lovrek, R.J. Howlett, L.C. Jain (Eds.), KES 2008 Part II, LNAI 5178, Springer-Verlag, Berlin Heidelberg, 2008, pp. 281–288.
[9] A. Tafreshi, A. Nasrabadi, A. Omidvarnia, Empirical mode decomposition in epileptic seizure prediction, IEEE International Symposium on Signal Processing and Information Technology (2008).
[10] A. Aarabi, B. He, A rule-based seizure prediction method for focal neocortical epilepsy, Clin. Neurophysiol. (2012) 1111–1122.
[11] M.Z. Parvez, M. Paul, M. Antolovich, Detection of pre-stage of epileptic seizure by exploiting temporal correlation of EMD decomposed EEG signals, J. Med. Bioeng. 4 (2) (2014) 110–116.
[12] CHB-MIT Scalp EEG Database, [Online]. Available: http://physionet.org/physiobank/database/chbmit/.
[13] A. Shoeb, J. Guttag, Application of machine learning to epileptic seizure detection, in: Proc. of the 27th International Conference on Machine Learning, Haifa, Israel, 2010.
[14] Y.U. Khan, O. Farooq, P. Sharma, N. Rafiuddin, Latency study of seizure detection, Adv. Comput. Sci., Eng. & Appl., AISC 166 (2012) 129–136.
[15] S. Nasehi, H. Pourghassem, Patient-Specific epileptic seizure onset detection algorithm based on spectral features and IPSONN classifier, 2013 International Conference on Communication Systems and Network Technologies (2013).
[16] N. Rafiuddin, Y.U. Khan, O. Farooq, Feature extraction and classification of EEG for automatic seizure detection, International Conference on Multimedia, Signal Processing and Communication Technologies (2011).
[17] D. Fergus, A.J. Hussain, D. Al-Jumeily, An advanced machine learning approach to generalised epileptic seizure detection, Intell. Comput. Bioinf., LNBI 8590 (2014) 112–118.
[18] J. Kevric, A. Subasi, The effect of multiscale PCA de-noising in epileptic seizure detection, J. Med. Syst. 38 (no. 10) (2014).
[19] S. Raghunathan, A. Jaitli, P.P. Irazoqui, Multistage seizure detection techniques optimized for low-power hardware platforms, Epilepsy Behav. 22 (2011) S61–S68.
[20] K.K. Majumdar, P. Vardhan, Automatic seizure detection in ECoG by differential operator and windowed variance, IEEE Trans. Neural Syst. Rehabil. Eng. 19 (no. 4) (2011) 356–365.
[21] Y. Liu, W. Zhou, Q. Yuan, S. Chen, Automatic seizure detection using wavelet transform and SVM in long-Term intracranial EEG, IEEE Trans. Neural Syst. Rehabil. Eng. 20 (no. 6) (2012) 749–755.
[22] E. Alickovic, A. Subasi, Effect of Multiscale PCA de-noising in ECG beat classification for diagnosis of cardiovascular diseases, Circuits, Syst. Signal Process. 34 (no. 2) (2015) 513–533.
[23] E. Alickovic, A. Subasi, Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier, J. Med. Syst. 40 (108) (2016) 1–12.
[24] E. Alickovic, Z. Babic, The effect of denoising on classification of ECG signals, XXV International Conference on Information, Communication and Automation Technologies (ICAT), Sarajevo (2015).
[25] E. Gokgoz, A. Subasi, Effect of multiscale PCA de-noising on EMG signal classification for diagnosis of neuromuscular disorders, J. Med. Syst. 38 (no. 4) (2014) 1–10 (31: April).
[26] Uni-Freiburg, Seizure Prediction Project Freiburg, 2011. [Online]. Available: https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database. [Accessed 2 October 2011].
[27] A. Shoeb, Application of Machine Learning to Epileptic Seizure Onset Detection and Treatment, 2009.
[28] A.L. Goldberger, L.A.N. Amaral, L. Glass, J.M. Hausdorff, P.C. Ivanov, R.G. Mark, J.E. Mietus, G.B. Moody, C.-K. Peng, H.E. Stanley, PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals, Circulation 101 (no. 23 (13 June)) (2000) 215–220.
[29] B.R. Bakshi, Multiscale PCA with application to multivariate statistical process monitoring, AIChE J. 44 (7) (July 1998) 1596–1610.
[30] J. Kevric, A. Subasi, Comparison of signal decomposition methods in classification of EEG signals for motor-imagery BCI system, Biomed. Signal Process. Control 31 (2017) 398–406.
[31] N.E. Huang, Z. Shen, S.R. Long, M.C. Wu, H.H. Shih, Q. Zheng, N.C. Yen, C.C. Tung, H.H. Liu, The empirical mode decomposition and the Hilbert spectrum fornonlinear and non-stationary time series analysis, Proc. R. Soc. Lond. A 454 (1971) (1998) 903–995.
[32] G. Rilling, P. Flandrin, P. Gonçalves, On empirical mode decomposition and its algorithms, IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP-03) Grado (2003).

[33] P. Ghorbanian, D.M. Devilbiss, A.J. Simon, A. Bernstein, T. Hess, H. Ashrafiuon, Wavelet transform EEG features of alzheimer's disease in activated states, in: 34th Annual International Conference of the IEEE EMBS, San Diego, California USA, 2012.
[34] M. Unser, A. Aldroubi, A review of wavelets in biomedical applications, Proceedings of the IEEE (1996) 626–638.
[35] I. Daubechies, The wavelet transform, time-frequency localization and signal analysis, IEEE Trans. Inf. Theory (1990) 961–1005.
[36] R.E. Learned, A.S. Willsky, A wavelet packet approach to transient signal classification, Appl. Comput. Harmon. Anal. (1995) 265–278.
[37] Y. Kutlu, D. Kuntalp, Feature extraction for ECG heartbeats using higher order statistics of WPD coefficients, Comput. Methods Programs Biomed. (2012) 257–267.
[38] A. Kandaswamy, C.S. Kumar, R.P. Ramanathan, S. Jayaraman, N. Malmurugan, Neural classification of lung sounds using wavelet coefficients, Comput. Biol. Med. vol. 34 (no. 6) (2004) 523–537.
[39] A. Subasi, EEG signal classification using wavelet feature extraction and a mixture of expert model, Expert Syst. Appl. 32 (2007) 1084–1093.
[40] A. Subasi, Medical decision support system for diagnosis of neuromuscular disorders using DWT and fuzzy support vector machines, Comput. Biol. Med. 42 (8) (2012) 806–815.
[41] L. Breiman, Random forests, Mach. Learn. 45 (2001) 5–32.
[42] B.E. Boser, I.M. Guyon, A training algorithm for optimal margin classifiers, Fifth Annual Workshop on Computational Learning Theory, Pittsburgh : ACM (1992).
[43] V. Vapnik, The Nature of Statistical Learning Theory, Springer-Verlag, New York, 1995.
[44] W. Chun-mei, Z. Chong-ming, Z. Jun-zhong, Z. Jian, Performance evaluation for epileptic electroencephalogram (EEG) detection by using Neyman?Pearson criteria and a support vector machine, Physica A vol. 391 (2012) 1602–1609.
[45] S. Haykin, Neural Networks A: Comprehensive Foundation, 2nd ed., Prentice Hall PTR Upper Saddle River, NJ, USA, 1998.
[46] T.M. Mitchell, Machine Learning, McGraw-Hill Science/Engineering/Math, 1997.
[47] D.W. Aha, D. Kibler, M.K. Albert, Instance-Based learning algorithms, Mach. Learn. 6 (1991) 37–66.
[48] C.A. Teixeira, B. Direitoa, M. Bandarabadi, M.L.V. Quyen, M. Valderrama, B. Schelter, A. Schulze-Bonhage, V. Navarroe, F. Sales, A. Douradoa, Epileptic seizure predictors based on computational intelligence techniques: a comparative study with 278 patients, Comput. Methods Programs Biomed. (2014) 324–336.
[49] X. Qiao, Y. Liu, Adaptive weighted learning for unbalanced multicategory classification, Biometrics 65 (1) (2009) 159–168.
[50] R. Andrzejak, K. Lehnertz, F. Mormann, C. Rieke, P. David, C. 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 (no. 6) (2001).
[51] A.K. Tafreshi, A.M. Nasrabadi, A.H. Omidvarnia, Epileptic seizure detection using empirical mode decomposition, in: IEEE International Symposium on Signal Processing and Information Technology, 2008. ISSPIT 2008, Sarajevo, 2008.
[52] R.J. Martis, U.R. Acharya, A. Tan, R. Petznick, C.K. Chua, E.Y.K. Ng, L. Tong, Application of empirical mode decomposition (emd) for automated detection of epilepsy using eeg signals, Int. J. Neural Syst. 22 (6) (2012).
[53] Z. Wu, N. Huang, Ensemble empirical mode decomposition: a noise-assisted data analysis method, Adv. Adapt. Data Anal. 1 (1) (2009) 1–41.
[54] L.M. Patnaik, O.K. Manyam, Epileptic EEG detection using neural networks and post-classification, Comput. Methods Programs Biomed. (2008) 100–109.
[55] U. Qidwai, A.S. Malik, M. Shakir, Embedded fuzzy classifier for detection and classification of preseizure state using real EEG data, The 15th International Conference on Biomedical Engineering (2014).