Available online at www.jbr-pub.org.cn

Open Access at PubMed Central

The Journal of Biomedical Research, 2020 34(3): 151–161

Original Article

JBR

EEG epileptic seizure detection and classification based on dual-tree complex wavelet transform and machine learning algorithms

Itaf Ben Slimen[1,✉], Larbi Boubchir[2], Zouhair Mbarki[1], Hassene Seddik[1]

[1]*Centre de Recherche et de Production Research Lab., Ecole Nationale Supérieure des Ingénieurs de Tunis, University of Tunis, Tunis 1008, Tunisia;*
[2]*Laboratoire d'Informatique Avancée de Saint-Denis Research Lab., University of Paris 8, Saint-Denis, Cedex 93526, France.*

Abstract

The visual analysis of common neurological disorders such as epileptic seizures in electroencephalography (EEG) is an oversensitive operation and prone to errors, which has motivated the researchers to develop effective automated seizure detection methods. This paper proposes a robust automatic seizure detection method that can establish a veritable diagnosis of these diseases. The proposed method consists of three steps: (i) remove artifact from EEG data using Savitzky-Golay filter and multi-scale principal component analysis (MSPCA), (ii) extract features from EEG signals using signal decomposition representations based on empirical mode decomposition (EMD), discrete wavelet transform (DWT), and dual-tree complex wavelet transform (DTCWT) allowing to overcome the non-linearity and non-stationary of EEG signals, and (iii) allocate the feature vector to the relevant class (*i.e.*, seizure class "ictal" or free seizure class "interictal") using machine learning techniques such as support vector machine (SVM), *k*-nearest neighbor (*k*-NN), and linear discriminant analysis (LDA). The experimental results were based on two EEG datasets generated from the CHB-MIT database with and without overlapping process. The results obtained have shown the effectiveness of the proposed method that allows achieving a higher classification accuracy rate up to 100% and also outperforms similar state-of-the-art methods.

Keywords: electroencephalography, epileptic seizure detection, feature extraction, dual-tree complex wavelet transform, machine learning

Introduction

According to the World Health Organization, epilepsy affects about 1% to 2% of the world's population; about 65 million people have epilepsy[1–2]. Different examinations are used to confirm the epilepsy diagnosis, such as electroencephalogram (EEG) signals recorded from the patient's scalp, which is a collection of potential signals from many electrodes and they represent the spatial distribution of the potential fields in the brain. EEG is an invaluable tool for characterization of the spatial-temporal

✉Corresponding author: Itaf Ben Slimen, Electric Engineering department, Centre de Recherche et de Production Research Lab., Ecole Nationale Supérieure des Ingénieurs de Tunis, University of Tunis, street Taha Hussein Montfleury, Tunis 1008, Tunisia. Tel: +216-27-542-572, E-mail: itafslimen@gmail.com.

Received 15 February 2019, Revised 30 November 2019, Accepted 11 February 2020, Epub 24 April 2020

CLC number: R742.1, Document code: A

The authors reported no conflict of interests.

https://doi.org/10.7555/JBR.34.20190026

dynamics of neuronal activity in the brain; hence, diagnosis by manual observation of these signals is a very sensitive task that resulted in practical tools for the research area dedicated to the diagnosis of seizure detection[3].

More than 120 million articles cited for epileptic seizure detection proved that this subject is still topical and should be addressed. Various epileptic seizure detection methods have been proposed in literature, and a brief review of some techniques related to our work was conducted and introduced in the next section. Gotman and Gloor were among the first to detect seizures using EEG signals[4]. The first machine learning approach was developed using the CHB-MIT database[5]. Some researchers have used a large database, extracted from Freiburg and CHB-MIT databases for seizure detection, and their proposed method allows achieving a high detection performance up to 100%[6]. Some proposed methods are based on artificial neural networks allowing achieving a perfect classification rate for detecting epileptic seizures[7]. Some are based on support vector machine (SVM) with time-frequency features and wavelet variance features allowing achieving correct classification rates up to 100%[8]. Another seizure detection method is based on dual-tree complex wavelet transformation (DTCWT) and Fourier features allowing achieving a correct classification rate up to 100%[9].

Various approaches for epileptic seizure detection have been proposed in literature. The processing of EEG signals requires that these data must be noiseless to establish reliable disease diagnosis; therefore, a filter step is necessary for this purpose. Some researchers proposed to use the Savitzky-Golay filter and have its performance verified[10], while others proposed a multi-scale principal component analysis (MSPCA)[11]. In general, all had adopted the objectives of the extraction characteristics and their classification, in order to differentiate between classes (*i.e.*, seizure free: "interictal" class and seizure: "ictal" class). Frequency domain analysis[11] and other time-frequency approaches, such as the short-term Fourier transform[12] and the wavelets[13,14–15] have often been used to extract discriminant characteristics representing the EEG signals. Advanced detection of epileptic seizures used a feature extraction technique based on wavelets[16]. Some features have shown high performance for characterization changes in EEG signals such as entropy, energy[17–18], principal component analysis (PCA)[19], and independent component analysis (ICA).

Based on these previous studies, the discrete wavelet transform (DWT) is one of the most widely used techniques for extracting EEG signal characteristics. For the same purpose, the DTCWT technique[20], applies filter banks to handle the wavelet domain. This technique was recently applied for epileptic seizure detection which demonstrated its performance for extracting the characteristics of the time-frequency domain of EEG signals[21]. After extracting features, the next step is to employ classification methods in order to determine the class of input features. Various classification methods are available. These range from rule-based decision-making[4] and SVM[22] to artificial neural network (ANN). For example, the classification of EEG signals was investigated using k-means classifier and multilayer perceptron neural network model[14]. A supervised automatic learning method was proposed for detecting seizures using linear discriminant analysis (LDA)[23].

The goal of this work is to build a reliable and accurate seizure detection method to facilitate and accomplish the diagnosis of epilepsy. Artifact eliminator is a valuable aid for diagnosis in EEG signal, using the Savitzky-Golay filter and the MSPCA. To overcome the non-linearity and non-stationary of EEG signals, the time-frequency techniques have been developed to eliminate the strict restrictions of classical frequency methods. Three popular signal decomposition techniques chosen in this work include empirical mode decomposition (EMD), DWT, and DTCWT. After crossing through a step of denoising the artifacts and extracting the frequency bands using EMD, DWT, and DTCWT, statistical values were used to extract the relevant characteristics of these frequency bands. Then three classifiers *k*-nearest neighbor (*k*-NN, LDA, and SVM were established to distinguish between classes (*i.e.*, interictal and ictal).

The choice of MSPCA and Savitzky-Golay for artifact removal is justified by their proven superiority in different biomedical signals kinds, such as ECG[24], EMG[25], and EEG[26]. The plethora of EMD, DWT, and DTWCT signal decomposition methods used in different fields of analysis justifies the choice in this work. This step is followed by the extraction of the statistical characteristics to extract the necessary information with reduction of used data dimensions. The aforementioned modulators were combined with three popular machine-learning techniques to find a powerful model for detecting epileptic seizures. Therefore, this work aims to develop an EEG signal classification system for seizure detection using CHB-MIT (EEG scalp) databases, and to ensure the robustness of the proposed model in which two databases have been extracted and realized based on EEG segments with and without overlapping procedure. The results proved that the proposed model is suitable for epileptic seizure detection. ***Fig. 1*** illustrates the work plan presented in this study.

Materials and methods

The lack of visual difference between the free seizure and seizure EEG signals makes the epileptic seizure detection a difficult task. Therefore, three proposed steps were employed, with EEG signal denoising techniques followed by features extraction methods to enable faster and more effective seizure detection. Hence, the classification using the extracted characteristics is performed using different machine learning. The EEG database used in our study is presented with a brief description.

EEG database

All tests and simulations were conducted based on the database collected at the Boston Children's Hospital (CHB) and used to validate the experimental results. This "CHB-MIT" database includes 23 different subsets containing EEG records of pediatric patients. It contains 182 seizures. All signals were sampled at 256 Hz per second with a resolution of 16 bits. Most files contain 23 EEG signals. The international 10-20 EEG electrodeposition and nomenclature system was used for these recordings[5].

EEG signal pre-processing

Biomedical signals such as EEG were contaminated by artifacts and noise, and thus a pre-processing step is needed to clean the data. Many signal-filtering algorithms have been proposed to improve the performance of health devices (EEG, ECG, EMG...) besieged by artifacts. Savitzky-Golay filter and multi-scale principal component analysis have been used in recent research and their performance has been verified.

Savitzky-Golay filter

During the pre-processing of biomedical signals, the most challenging problem is to extract high resolution EEG signals from noisy measurements and retain the EEG waveform sharpness. The spectral quality of a signal [signal-to-noise (S/N) ratio] may be improved by increasing filter width or smoothing the signal multiple times. Hence, The Savitzky-Golay filter is a low-pass filter used for attenuation noise, without damaging or destroying the spectral properties of data. This filter is defined as a weighted moving average that performs a least-squares-fit convolution procedure.

Therefore, the S–Golay algorithm is characterized by the selected data interval, in other words, the window size, a polynomial function is fitted to the chosen size window and the data point (noisy signal) is calculated using the polynomial coefficients. The set of data $n\{x_j, y_j\}$ points, where $j=1,2,\ldots,n$; x is an independent variable and y is the observed value, can be represented with convolution coefficients called c and C_i and expressed by:

$$Y_j = \sum_{i=-\frac{c-1}{2}}^{i=\frac{c-1}{2}} C_i y_{j+1} \frac{c+1}{2} \leqslant j \leqslant n - \frac{c-1}{2} \quad (1)$$

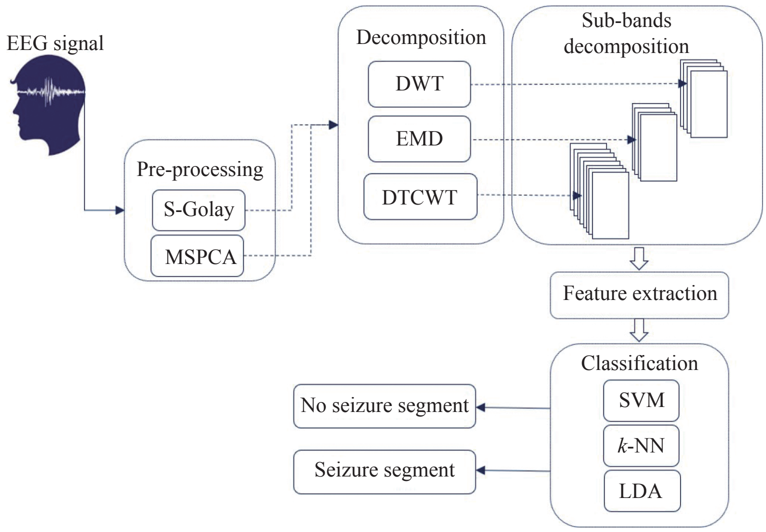

Fig. 1 **Diagram for proposed seizure detection approach.** MSPCA: multi-scale principal component analysis; DWT: discrete wavelet transform; EMD: empirical mode decomposition; DTCWT: dual-tree complex wavelet transformation; SVM: support vector machine; *k*-NN: *k*-nearest neighbor; LDA: linear discriminant analysis.

Hence, S-Golay filter requires the inputs which include: the input noisy signal (x), the polynomial order (k) and its frame size (f). One of the advantages of this filter is to retain distribution characteristics, which are often flattened with other smoothing methods[10,27].

Multi-scale principal component analysis (MSPCA)

The MSPCA approach is proposed to achieve the objective of eliminating noise without sacrificing a significant amount of fast changes in the signal and to provide a dimension reduction with select components relevant in the different scales. The advantages of this method come from the combination of the PCA, which is responsible for extracting the relationship between several variables, and the wavelets, that divide the stochastic processes of the determinists and approximately decorrelate the auto-correlation between the variables. The MSPCA calculates the PCA of all wavelet coefficients for all wavelet decomposition levels, and then establishes a relationship between the results obtained and the suitable levels[24]. ***Fig. 2*** illustrates the steps of the MSPCA method and the following algorithm explains it:

1) Perform wavelet decomposition at level $\boldsymbol{I}$ for each column of data matrix X.

2) For $1 \leqslant i \leqslant I$, perform the PCA of the detailed matrix G_iX, choose an appropriate number of significant principal components p_i or suppress the detail.

3) Perform the PCA of the approximation matrix H_IX and select a suitable number of principal components p_{I+1}.

4) The main characteristics of the original matrix X put in a new matrix $\dot{X}$, by inverting the DWT^T wavelet transformation.

5) Finally, execute the PCA of $\dot{X}$[28].

EEG signal decomposition

To characterize the EEG signal and determine its class, some methods have often been used to extract discriminant characteristics representing EEG signals from frequency analysis and time-frequency approaches. These, such as DWT, DTCWT, and EMD, are recently used in the literature and will be used in this work and represented in this section with their proven superiority to reach this goal.

Empirical mode decomposition (EMD)

EMD is a procedure used to decompose signals into different simple oscillatory modes called intrinsic function mode (IMF). Hence, the signal $x(t)$ can be represented by the sum of these IMFs components and by a residue as follows[29]:

$$x(t) = \sum_{m=1}^{M} D_m(t) + r_M(t) \tag{2}$$

where M is the number of IMFs and $r_M(t)$ is the residue. The IMFs of a signal must satisfy two necessary conditions. Firstly, the difference between the total numbers of extrema and zero-crossings is at most one. Secondly, in all points, the average value of the envelope defined by the local maxima and that by the local minima is zero or very close to zero[30]. The following steps represent the EMD algorithm:

1) Extract local maxima and local minima from signal of $x(t)$,

2) Connect between all minima and all maxima to obtain two envelopes $e_{min}(t)$ and $e_{max}(t)$,

3) Compute the average between them as: $m(t) = \frac{[e_{min}(t) + e_{max}(t)]}{2}$,

4) Extract the detail from $x(t)$ as: $d(t)=x(t)-m(t)$

5) Examine, if $d(t)$ is an IMF or not by verifying the above conditions for IMFs,

6) Reiterate on the residual $d(t)$.

After having the first IMF $D_1(t)$ of the signal, to get other IMFs, generate the residual $r(t)$ as follows: $r(t)=x(t)-D_1(t)$. The process described above will be continued until IMF can no longer be obtained from the final residue.

Discrete wavelet transform (DWT)

DWT is one of the most efficient time-frequency

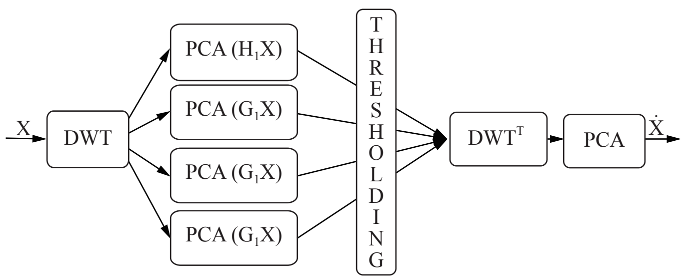

Fig. 2 **The procedure for MSPCA.** DWT: discrete wavelet transform; PCA: principal component analysis.

tools for non-stationary signal analysis. The DWT procedure makes it possible to present an input signal in sets of functions, called wavelets, by scaling and shifting the mother function into wavelets, that is to say, to decompose a signal into two components called sub-bands: a low frequency component called "approximation coefficients A", and a high frequency component called "detailed coefficients D". The low frequency component can further be decomposed into approximations and details. After each decomposition, the frequency resolution is doubled and the temporal resolution is halved. After choosing the decomposition level j, the components of the signal $s(k)$ can be represented by the following expression and ***Fig. 3***[31]:

$$D_j(i) = \sum_k s(k) \cdot h(2 \cdot i - k) \tag{3}$$

$$A_j(i) = \sum_k s(k) \cdot l(2 \cdot i - k) \tag{4}$$

Regarding ***Fig. 3***, h and g are low-pass and high-pass filters, respectively. s is the input original signal where $Ca1$, $Ca2$, and $Ca3$ represent the approximation coefficients of $h(1)$, $h(2)$ and $h(3)$ and $Cd1$, $Cd2$ and $Cd3$ represent the detail coefficients of $g(1)$, $g(2)$ and $g(3)$ at level 1, 2 and 3, respectively.

Dual-tree complex wavelet transform (DTCWT)

DTCWT, a DWT enhancement, uses double wavelet filters low and high at each scale to obtain two parts: a real and an imaginary complex wavelet coefficient. This transformation is invariant by directional and selective shifting in two or more dimensions, which are very important in applications such as pattern recognition and signal analysis.

Hence, the DTCWT can be represented by the following expression[20]:

$$\psi(t) = \psi_h(t) + i\psi_g(t) \tag{5}$$

$$\varphi(t) = \varphi_h(t) + i\varphi_g(t) \tag{6}$$

where the complex wavelet is $\psi(t)$ while the complex scaling function is $\varphi(t)$. With reference to ***Fig. 4***, s represents the input signal, and the output coefficients of the top tree h and those of the bottom tree g are considered as the real and imaginary parts,

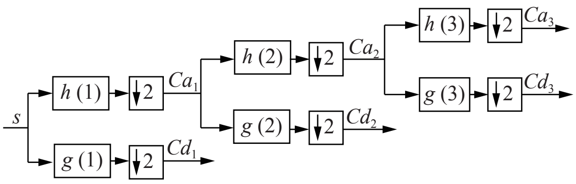

Fig. 3 **DWT decomposition tree.**

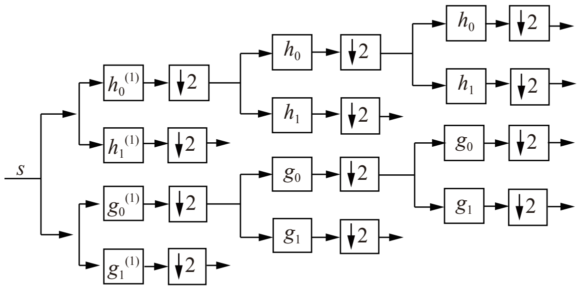

Fig. 4 **An illustrative scheme of the DTCWT at level 3.**

respectively. Therefore, $h_0^{(1)}, h_0$, $g_0^{(1)}$, and g_0 are a low-pass filter and $h_1^{(1)}$, $g_1^{(1)}$, h_1, and g_1 are high-pass filters. Consequently, the corresponding wavelets $\psi_h(t)$ and $\psi_g(t)$ form approximately a Hilbert pair, while the scaling function $\varphi_g(t)$ is approximately the Hilbert transform of $\varphi_h(t)$.

Discriminative feature extraction

The classification requires a necessary step whose characteristic extraction aims to obtain the most important information and preserve the small size of the data. Therefore, after signal decomposition using the DWT, EMD and DTCWT methods, the most prominent statistical features have been used for this study to obtain the class of the input EEG signal:

Mean of coefficients in each sub-band:

$$\mu = \frac{1}{M} \sum_{j=1}^{M} y_j \tag{7}$$

Average power of coefficients in each sub-band:

$$\lambda = \sqrt{\frac{1}{M} \sum_{j=1}^{M} y_j^2} \tag{8}$$

Standard deviation of coefficients in each sub-band:

$$\sigma = \sqrt{\frac{1}{M} \sum_{j=1}^{M} \left(y_j - \mu\right)^2} \tag{9}$$

Signal classification

Frequency band decomposition is performed to represent an EEG signal in feature vector and to obtain the most important information of the data to determine its class. The identification of signals relied on classification algorithms to estimate the desired class. For this purpose, three popular classifiers are used, namely LDA, k-NN, and SVM, which are briefly explained in this section.

Linear discriminant analysis (LDA)

LDA, also called Fisher's LDA, is a simple technique used in machine learning, statistics, and pattern recognition to find a linear separating two or more classes, and in general, it gives a good class estimation. Fisher's LDA requires very little computation. LDA aims to use hyperplanes to separate data from different classes, and the separator hyperplane is obtained by looking for the projection that maximizes the distance between classes and minimizes the interclass variance between them. For a class number N (N>2), several hyperplanes are used[32–33]. In this work, the linear discriminant type was used for regularized linear discriminant analysis.

Support vector machine (SVM)

With excellent performance results, SVM is applied in various fields. Contrary to LDA, which constructs a probabilistic model for each class using all data points, SVM is to minimize a linear separation error by focusing on the neighboring points of the linear separation surface, that is to say, the points located far from the separation surface are ignored in the learning process. SVM is based on the Vapnik-Chervonenkis theory and the principle of structural risk minimization. Its purpose is to determine the minimizing of the training set error by maximizing the boundary of the separating hyper-plane between the data[34–35]. In this work, the Kernel function used to differentiate between the two classes is linear kernel with the Kernel scale parameter chosen to be 1.

k-nearest neighbor (k-NN)

The *k*-NN is a machine-learning technique used for the classification task, which aims to find the nearest k-neighbors among the learning set, and the categories of *k* closest neighbors, which are used to weight the candidates in the category. The closest neighbors to the sample are expressed as distance such as Euclidean distance, Chebyshev distance, and Manhattan distance. The performance of *k*-NN algorithm strongly depends on two factors: an appropriate similarity function and an appropriate value for *k*. If *k* is too large, large classes will overwhelm small classes, whereas if *k* is too small, the advantage of the *k*-NN algorithm cannot be traced[36–37,19]. In this research, 3-NN algorithm (3 nearest neighbors) and Euclidean distance metric were used to classify the EEG signals.

Database preparation

The CHB-MIT database is one of the EEG databases used in recent research for epileptic seizure detection. To prove the effectiveness of the proposed method, two EEG datasets have been extracted from the CHB-MIT database. Hence, the extraction of shorter EEG segments based on no overlapping process represents the first used database, while the second based on overlapping process. The segment extraction process of these two datasets is clearly explained in this section. The performance evaluation criteria to assess the performance and the effectiveness of the proposed method and the experimental results obtained will be shown as follows.

The developed seizure detection model is based on EEG segment extraction with an 8-second rectangular window. The seizure time is known and given in the description database for each patient. The first extracted dataset contains 400 EEG segments extracted approximately from all the recorded signals in the CHB-MIT database with 200 no seizure segments ("interictal") and 200 seizure segments ("ictal").

According to a previous study[6], the choice of the 8 seconds window size (with 2 048 samples) provides a more robust performance than that of shorter or longer windows. To prove the effectiveness of the proposed seizure detection model, the overlapped database has been created with an overlapping process based on the first extracted database that will be presented as follows.

Overlapping procedure

While the first database was created (with 400 segments), to evaluate the performance and better prove the effectiveness of the proposed seizure detection method, another database will be created based on the first realized with an overlapping process. Hence, new segments will be created by overlapping between two segments belonging to the same EEG signal in the same class. Let suppose that:

- N is the segment number of the first database already created (without overlapping) called B. $N1$ and $N2$ is the number of segments in "interictal" class $B1$ and "ictal" class $B2$ respectively,

- Nc is the segment number of the second database with overlapping process called Bc. $Nc1$ and $Nc2$ are the numbers of segments in "interictal" class $Bc1$ and "ictal" class $Bc2$ respectively.

Therefore, the overlap process is presented as follows:

- n represents the segment samples number (2 048), i is segment index (i^{th} segment), p is the point offset for overlap, fe is the sampling frequency 256 Hz, and sc is the segment seconds number (8 seconds).

Example: let suppose i and $i+1$ are two segments of the same class to be overlapped where $X1_i$, $X2_i$, $X3_i$, $X4_i$, $X5_i$, $X6_i$, $X7_i$, $X8_i$ and $X1_{i+1}$, $X2_{i+1}$, $X3_{i+1}$, $X4_{i+1}$, $X5_{i+1}$, $X6_{i+1}$, $X7_{i+1}$,$X8_{i+1}$, are represent the sample of segments of i and $i+1$, respectively, and j the new obtained segment with a gap between the segment is $p=6$.

$$i = [X1_i, X2_i, X3_i, X4_i, X5_i, X6_i, X7_i, X8_i]$$

$$i+1 = [X1_{i+1}, X2_{i+1}, X3_{i+1}, X4_{i+1}, X5_{i+1}, X6_{i+1}, X7_{i+1}, X8_{i+1}]$$

$$j = [X3_i, X4_i, X5_i, X6_i, X7_i, X8_i, X1_{i+1}, X2_{i+1}]$$

Finally, the new segment can be presented by the following expression:

$$j = \left\{ i\left[\frac{n}{sc} \times (sc-p) : n \times sc\right]; i+1\left[1 : \frac{n}{sc} \times (sc-p)\right]\right\} \tag{10}$$

$$j = \{ i[fe \times (sc-p) : n \times sc]; i+1[1 : fe \times (sc-p)]\} \tag{11}$$

The size of new database (overlapped) will be increased compared to the first database.

$N = N1 + N2$ and $Nc = Nc1 + Nc2$; hence, $Nc1 = (N1 \times 2) - 1$ and $Nc2 = (N2 \times 2) - 1$.

$$Nc = [(N1 \times 2) - 1] + ((N2 \times 2) - 1) = 2N1 + 2N2 - 2 \tag{12}$$

Therefore, the overlapped database size will be 798 segments, with 399 segments in both classes and each length segment 8 seconds. The experimental results of the proposed approach using the two extracted databases will be presented by following the performance criteria to validate the model accuracy.

Feature selection process

The sampling rates of the databases are 256 Hz; the polynomial order was chosen to be 2 on S-Golay filter and the frame length was selected to be 17, proved to be the parameter' choice[10]. In MSPCA, a five-level decomposition was used to generate the sub-bands signals. The Kaiser rule was responsible for eliminating the number of loadings[26]. After the artifact removal procedure with S-Golay and MSPCA, the features have been extracted from the sub-band decomposed signals using EMD, DWT, and DTCWT methods. The Daubechies wavelets of different orders (2, 3, 4, 5, and 6) are investigated for the analysis of epileptic EEGs. This family of wavelets has orthogonal property and is efficient as a filter implementation[38]. The fourth order Daubechies wavelet is found to be most appropriate for analysis of EEG data[6,39]. Hence, DB4 wavelet function was used to decompose the EEG signals with DWT. The decomposition levels in all methods were selected to be 3. Hence, 4, 8 and 3 sub-band signals were obtained using DWT, DTCWT, and EMD respectively. Therefore, in each sub-band signals three statistical characteristics have been extracted to obtain 12, 24, and 9 features for DWT, DTCWT, and respectively. However, although the number of extracted features was considerably doubled in DTCWT, the performance in the testing step improved.

Performance evaluation criteria

To prove the robustness of the proposed method, cross-validation (CV) is a procedure mainly applied to estimate the proficiency of machine learning algorithm. CV is used to estimate how the model should work when used to make predictions of unused data at the step of preparation and model learning. In a 10 times CV, the database is randomly separated into 10 different mutually exclusive folds of the same size, where nine folds will be used to training the data and the remaining one will be used to test and predict the classification. The CV accuracy is expressed by:

$$CV = \frac{1}{k}\sum_{i=1}^{k} A_i \tag{13}$$

where represents the number of folds used, and A_i is the accuracy measure of each, $i = 1, \ldots, k$. In addition, there are three performance measures to evaluate machine learning methods using true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN), which can be expressed by:

$$Sensitivity(\%) = \frac{TP}{TP+FN} \times 100 \tag{14}$$

$$Specificity(\%) = \frac{TN}{TN+FP} \times 100 \tag{15}$$

$$Accuracy(\%) = \frac{TP+TN}{TP+FN+TN+FP} \times 100 \tag{16}$$

Results

The previous sections have shown that the first goal in this work is to find a model to classify EEG signal segments to develop automatic seizure detection systems using the children's EEG database CHB-MIT. However, shorter EEG segments with 8 seconds have been extracted from this database that makes it possible to develop a practical and robust seizure detection system to diagnose epilepsy. Ictal record signals represent a small amount of data compared to interictal record signals. Therefore, a subsampling paradigm of the class with larger proportions (interictal) is used to circumvent the imbalance between the two classes[40], in order to have better accuracy of seizure detection. In this work, segments of interictal data were randomly selected to produce a comparable and balanced number of ictal and interictal segments.

Based on the segment model, the purpose of this study was to develop an EEG signal classification model. Therefore, the contribution is to have a powerful and suitable model using three popular methods to extract discriminant characteristics such as EMD, DWT, and DTCWT, followed by a dimension reduction employing three different statistical characteristics and three commonly used methods of automatic learning such as LDA, *k*-NN, and SVM.

To prove the performance results of the average precision, extraction methods were evaluated and the results are presented in detail and illustrated in ***Fig. 5***. The parameter values of all selected techniques are explained in the preview section because they produce the most robust results. In ***Fig. 5***, the plots are shown in the left column present the S-Golay results to apply a series of data points to denoise the signals without deforming its input, while the plots are shown in the right column present the MSPCA results, which shows the advantage of PCA in extracting necessary information and the effectiveness of extracting deterministic features from stochastic processes using wavelets. With the two filter processes compared, it can be found that MSPCA outperformed the S-Golay filter in the removal of artifacts; MSPCA achieved a 100% accuracy with the three classification methods while S-Golay 98%.

With the three EEG signal decomposition methods compared, the results indicate that using the two denoising methods, DTCWT was superior to DWT and much better than EMD where with S-Golay (DTCWT, 99%; DWT, 94%; and EMD, 91%). ***Fig. 5*** also shows that the DTCWT gave a fairly high classification result with 100% accuracy classification rates, which is attributed to the approximate shift invariant property, a very important factor in epileptic seizure detection. On aspects of the performance of feature extraction methods, the experimental results show the superiority of the standard deviation and the

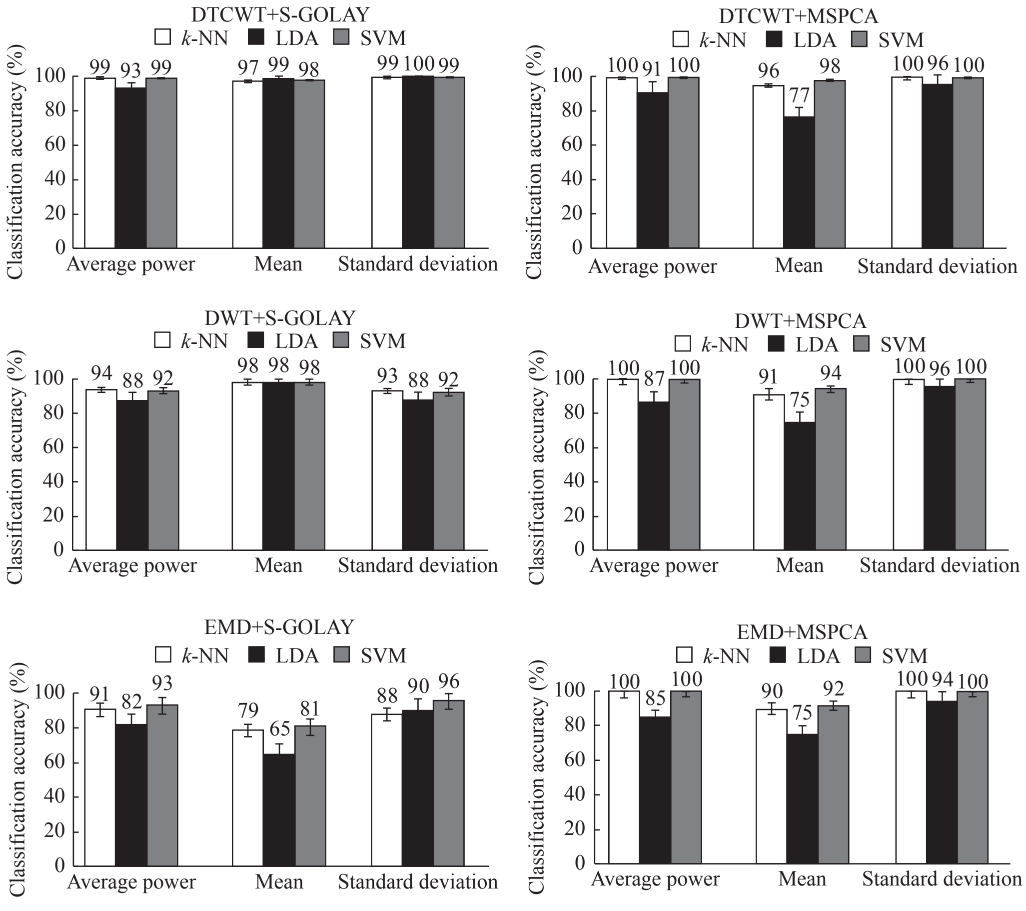

Fig. 5 **Classification accuracy of the proposed method using the signal decomposition transforms for feature extraction considered in this study (*i.e.*, EMD, DWT, DTCWT) with EEG signal pre-processing techniques: MSPCA (right column) and S-Golay (left column).**

average power, which both gave a rather high classification accuracy when compared to the mean of the sub-band features. Hence the mean static feature with DTCWT outperforms that with DWT, which ensures the high effectiveness of DTCWT. In addition, the result provides that *k*-NN and SVM gave 100% classification accuracy when compared to LDA.

Discussion

The overall results obtained in this study show that combination of the DTCWT with SVM and *k*-NN improves the accuracy rates of the classification especially by using the standard static deviation and average power characteristics with the two denoising processes. As shown in ***Fig. 5*** which (MSPCA+ DTCWT+STD+SVM), (MSPCA+DTCWT+STD+*k*-NN, (S-Golay+DTCWT+STD+SVM) and (S-Golay+ DTCWT+STD+*k*-NN achieve a 100% classification accuracy rate. Hence, it can be concluded that the proposed model is reliable and robust for seizure detection with the DTCWT combined with SVM and *k*-NN.

To compare the results of the proposed approach with seizure detection methods realized in literature, ***Table 1*** provides that the proposed approach has achieved an overall perfect accuracy of 100% in seizure detection using EEG segments.

The obtained results using the overlapped database ensure the reliability and stability of this seizure detection model as illustrated in ***Table 2*** and ***Fig. 6***. The combination of DTCWT and MSPCA offers advantages over the DWT and EMD and keeps the same 100% accuracy performance. The important difference between EMD and wavelet-based methods lies in the process of feature extraction. IMF signals are the product of EMD methods and have the same length as the original EEG signals, which have given no reduction in the number of samples. On the other hand, the coefficients characterizing different sub-bands produced by wavelet-based methods (DWT and DTCWT) have a much smaller number of samples than the EEG signals input. However, the DWT suffers from certain disadvantages such as the oscillatory nature, the aliasing of the spectrum with limited directional information, and the lack of lag invariance. The DTCWT has solved these problems by its non-oscillating property with approximate invariant shift amplitude where this property is very important for the detection of epileptic seizures in EEG signals.

In conclusion, we have proposed a robust automatic approach for epileptic seizure detection and classification that was assessed on two EEG datasets generated from the CHB-MIT database. The proposed approach is based on two denoising signal methods (*i.e.*, S-Golay and MSPCA) combined with three popular feature extraction techniques (*i.e.*, DTCWT, DWT, and EMD). Three different machine-learning techniques (namely SVM, *k*-NN, and LDA) have been used to distinguish between the seizure-free and seizure segments based on the feature extraction methods already mentioned. The experimental results have shown that the employment of DTCWT has been very successful because it outperforms the other two methods (DWT and EMD). The results obtained have also clearly shown that MSPCA+DTCWT+SVM and *k*-NN offers a perfect classification with 100% accuracy. The findings demonstrate the effectiveness of the proposed method compared to existing methods. After completing the seizure detection task, future studies may take more interest in the prediction of epileptic seizures.

Table 1 **Comparison of the proposed method against some state-of-the-art methods**

Method	Feature extraction	Classifier	Accuracy
Fergus *et al*[23] (2014)	Band-pass filter, peak and median frequency, RMS, Entropy based-features	*k*-NN	88%
Raffiudin *et al*[16] (2011)	Wavelet-based features+statistics-based features+IQR+MAD	LDA	96%
Alickovic *et al*[6] (2018)	Multiscale PCA, EMD, DWT, WPD, statistics-based features	RF, SVM, MLP, and *k*-NN	99.7%
Swami *et al*[21] (2016)	DTCWT	GRNN	98%
Gandhi *et al*[18] (2011)	DWT, energy, standard deviation and entropy based-features	SVM and PNN	99%
Acharya *et al*[13] (2012)	Approximate entropy, sample entropy, HOS, decision tree, fuzzy-based features	GMM, *k*-NN, SVM, and RBPNN	99%
Swami *et al*[22] (2014)	WPT, energy, standard deviation and entropy based-features	SVM	99%
Proposed method	MSPCA, DTCWT	SVM and *k*-NN	100%

MSPCA: multi-scale principal component analysis; DWT: discrete wavelet transform; EMD: empirical mode decomposition; DTCWT: dual-tree complex wavelet transformation; SVM: support vector machine; *k*-NN: *k*-nearest neighbor; LDA: linear discriminant analysis.

Table 2 **Classification accuracy of the proposed overlapping method**

Overlapping			Features								
			Average power			Mean			Standard deviation		
			2	4	6	2	4	6	2	4	6
DTCWT	SVM	S-Golay	98	98	96	75	80	78	97	98	96
		MSPCA	100	100	100	87	71	71	100	100	100
	k-NN	S-Golay	98	97	96	86	88	89	99	98	96
		MSPCA	100	100	100	93	92	93	100	100	100
	LDA	S-Golay	93	92	93	70	81	71	99	93	92
		MSPCA	91	87	91	76	74	75	98	96	97
DWT	SVM	S-Golay	94	92	92	75	78	75	93	94	94
		MSPCA	100	100	99	76	71	72	100	100	99
	k-NN	S-Golay	95	93	96	84	89	90	95	92	96
		MSPCA	100	100	99	88	90	91	100	100	100
	LDA	S-Golay	98	90	89	80	81	81	91	91	92
		MSPCA	90	91	90	76	72	75	98	93	98
EMD	SVM	S-Golay	94	95	95	66	61	64	93	94	92
		MSPCA	100	100	99	65	70	63	100	100	99
	k-NN	S-Golay	94	94	94	75	76	84	96	92	94
		MSPCA	99	100	99	92	91	78	99	100	100
	LDA	S-Golay	80	81	83	50	62	60	88	90	91
		MSPCA	84	88	87	63	72	67	99	97	97

MSPCA: multi-scale principal component analysis; DWT: discrete wavelet transform; EMD: empirical mode decomposition; DTCWT: dual-tree complex wavelet transformation; SVM: support vector machine; *k*-NN: *k*-nearest neighbor; LDA: linear discriminant analysis.

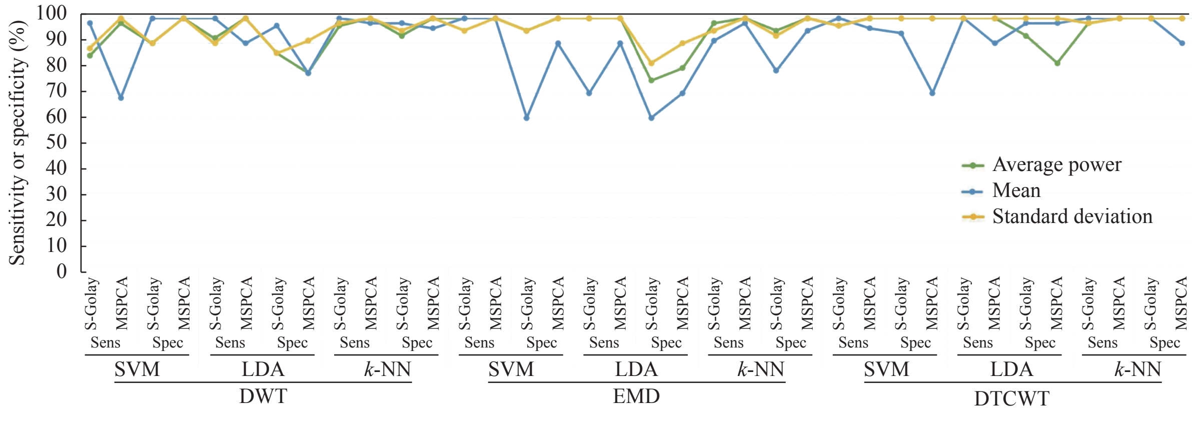

Fig. 6 **The sensitivity and specificity of the proposed method using the signal decomposition transforms for feature extraction considered in this study (*i.e.*, EMD, DWT, DTCWT) with EEG signal pre-processing techniques (MSPCA and S-Golay).** MSPCA: multi-scale principal component analysis; DWT: discrete wavelet transform; EMD: empirical mode decomposition; DTCWT: dual-tree complex wavelet transformation; SVM: support vector machine; *k*-NN: *k*-nearest neighbor; LDA: linear discriminant analysis; Sens: sensitivity; Spec: specificity.

References

[1] Acharya UR, Sree SV, Swapna G, et al. Automated EEG analysis of epilepsy: a review[J]. *Knowl-Based Syst,* 2013, 45: 147–165.

[2] Moshé SL, Perucca T, Ryvlin P, et al. Epilepsy: new advances[J]. *Lancet,* 2015, 385(9971): 884–898.

[3] Adeli H, Ghosh-Dastidar S. Automated EEG - based diagnosis of neurological disorders inventing the future of neurology[M]. New York: CRC Press, 2010: 71–75.

[4] Gotman J. Automatic detection of seizures and spikes[J]. *J Clin Neurophysiol,* 1999, 16(2): 130–140.

[5] Shoeb AH. Application of machine learning to epileptic seizure onset detection and treatment[D]. Cambridge: Harvard-MIT Division of Health Sciences and Technology, 2009: 157–162.
[6] Alickovic E, Kevric J, Subasi A. Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction[J]. *Biomed Signal Process Control,* 2018, 39: 94–102.
[7] Tzallas AT, Tsipouras MG, Fotiadis DI. Automatic seizure detection based on time-frequency analysis and artificial neural networks[J]. *Comput Intell Neurosci,* 2007, 2007: 80510.
[8] Xie SK, Krishnan S. Wavelet-based sparse functional linear model with applications to EEGs seizure detection and epilepsy diagnosis[J]. *Med Biol Eng Comput,* 2013, 51(1-2): 49–60.
[9] Chen GY. Automatic EEG seizure detection using dual-tree complex wavelet-Fourier features[J]. *Expert Syst Appl,* 2014, 41(5): 2391–2394.
[10] Acharya D, Rani A, Agarwal S, et al. Application of adaptive savitzky-golay filter for EEG signal processing[J]. *Perspect Sci,* 2016, 8: 677–679.
[11] Polat K, Günes S. Classification of epileptiform EEG using a hybrid system based on decision tree classifier and fast Fourier transform[J]. *Appl Mathem Comput,* 2007, 187(2): 1017–1026.
[12] Duque-Muñoz L, Espinosa-Oviedo JJ, Castellanos-Dominguez CG. Identification and monitoring of brain activity based on stochastic relevance analysis of short - time EEG rhythms[J]. *BioMed Eng OnLine,* 2014, 13: 123.
[13] Acharya UR, Sree V, Ang PCA, et al. Application of non-linear and wavelet based features for the automated identification of epileptic EEG signals[J]. *Int J Neural Syst,* 2012, 22(2): 1250002.
[14] Gandhi TK, Chakraborty P, Roy PG, et al. Discrete harmony search based expert model for epileptic seizure detection in electroencephalography[J]. *Expert Syst Appl,* 2012, 39(4): 4055–4062.
[15] Swami P, Gandhi TK, Panigrahi BK, et al. A comparative account of modelling seizure detection system using wavelet techniques[J]. *Int J Syst Sci: Oper Logist,* 2017, 4(1): 41–52.
[16] Rafiuddin N, Khan YU, Farooq O. Feature extraction and classification of EEG for automatic seizure detection[C]// Proceedings of 2011 International Conference on Multimedia Signal Processing and Communication Technologies. Aligarh: IEEE, 2011.
[17] Gandhi T, Panigrahi BK, Bhatia M, et al. Expert model for detection of epileptic activity in EEG signature[J]. *Expert Syst Appl,* 2010, 37(4): 3513–3520.
[18] Gandhi T, Panigrahi BK, Anand S. A comparative study of wavelet families for EEG signal classification[J]. *Neurocomputing,* 2011, 74(17): 3051–3057.
[19] Subasi A, Gursoy MI. EEG signal classification using PCA, ICA, LDA and support vector machines[J]. *Expert Syst Appl,* 2010, 37(12): 8659–8666.
[20] Selesnick WI, Baraniuk RG, Kingsbury NC. The dual-tree complex wavelet transform[J]. *IEEE Signal Process Mag,* 2005, 22(6): 123–151.
[21] Swami P, Gandhi TK, Panigrahi BK, et al. A novel robust diagnostic model to detect seizures in electroencephalography[J]. *Expert Syst Appl,* 2016, 56: 116–130.
[22] Swami P, Godiyal AK, Santhosh J, et al. Robust expert system design for automated detection of epileptic seizures using SVM classifier[C]//Proceedings of 2014 International Conference on Parallel, Distributed and Grid Computing. Solan: IEEE, 2014: 219–222.
[23] Fergus P, Hignett D, Hussain AJ, et al. An advanced machine learning approach to generalised epileptic seizure detection[C]//Proceedings of the 10th International Conference on Intelligent Computing. Taiyuan, China: 2014, Springer: 112–118.
[24] Alickovic E, Subasi A. Effect of Multiscale PCA de-noising in ECG beat classification for diagnosis of cardiovascular diseases[J]. *Circuits, Syst Signal Process,* 2015, 34(2): 513–533.
[25] Gokgoz E, Subasi A. Effect of multiscale PCA de-noising on EMG signal classification for diagnosis of neuromuscular disorders[J]. *J Med Syst,* 2014, 38(4): 31.
[26] Kevric J, Subasi A. The effect of multiscale PCA de-noising in epileptic seizure detection[J]. *J Med Syst,* 2014, 38(10): 131.
[27] Agarwal S, Rani A, Singh V, et al. EEG Signal enhancement using cascaded S-Golay filter[J]. *Biomed Signal Process Control,* 2017, 36: 194–204.
[28] Aminghafari M, Cheze N, Poggi JM. Multivariate denoising using wavelets and principal component analysis[J]. *Computat Statist Data Anal,* 2006, 50(9): 2381–2398.
[29] Pachori RB, Patidar S. Epileptic seizure classification in EEG signals using second-order difference plot of intrinsic mode functions[J]. *Comput Methods Programs Biomed,* 2014, 113(2): 494–502.
[30] Rilling G, Flandrin P, Gonçalvès P. On empirical mode decomposition and its algorithms[C]//Proceedings of IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing. Grado: IEEE, 2003.
[31] Alickovic E, Subasi A. Medical decision support system for diagnosis of heart arrhythmia using DWT and random forests classifier[J]. *J Med Syst,* 2016, 40(4): 108.
[32] Duda RO, Hart PE, Stork DG. Pattern classification[M]. 2nd ed. New York: Wiley, 2001.
[33] Fukunaga K. Introduction to statistical pattern recognition[M]. 2nd ed. San Diego: Academic Press, Inc, 1990.
[34] Boser BE, Guyon IM, Vapnik VN. A training algorithm for optimal margin classifiers[C]//Proceedings of the Fifth Annual Workshop on Computational Learning Theory. Pittsburgh: ACM, 1992.
[35] Vapnik VN. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995.
[36] Mitchell TM. Machine learning[M]. New York: McGraw-Hill, 1997.
[37] Aha DW, Kibler D, Albert MK. Instance-Based learning algorithms[J]. *Mach Learn,* 1991, 6(1): 37–66.
[38] Smit DJA, Boersma M, Schnack HG, et al. The brain matures with stronger functional connectivity and decreased randomness of its network[J]. *PLoS One,* 2012, 7(5): e36896.
[39] Metin A. Time frequency and wavelets in biomedical signal processing[M]. New York: Wiley-IEEE Press, 1998: 174–176, 207–210.
[40] Qiao XY, Liu YF. Adaptive weighted learning for unbalanced multicategory classification[J]. *Biometrics,* 2009, 65(1): 159–168.