

## Research Article

# Patient Specific Seizure Prediction System Using Hilbert Spectrum and Bayesian Networks Classifiers

Nilufer Ozdemir<sup>1</sup> and Esen Yildirim<sup>2</sup>

<sup>1</sup> Iskenderun Vocational School, Mustafa Kemal University, 31200 Iskenderun, Hatay, Turkey

<sup>2</sup> Department of Computer Engineering, Mustafa Kemal University, 31200 Iskenderun, Hatay, Turkey

Correspondence should be addressed to Esen Yildirim; esenyildirim@gmail.com

Received 1 April 2014; Revised 8 July 2014; Accepted 28 July 2014; Published 27 August 2014

Academic Editor: Thomas Desaive

Copyright © 2014 N. Ozdemir and E. Yildirim. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The aim of this paper is to develop an automated system for epileptic seizure prediction from intracranial EEG signals based on Hilbert-Huang transform (HHT) and Bayesian classifiers. Proposed system includes decomposition of the signals into intrinsic mode functions for obtaining features and use of Bayesian networks with correlation based feature selection for binary classification of preictal and interictal recordings. The system was trained and tested on Freiburg EEG database. 58 hours of preictal data, 40-minute data blocks prior to each of 87 seizures collected from 21 patients, and 503.1 hours of interictal data were examined resulting in 96.55% sensitivity with 0.21 false alarms per hour, 13.896% average proportion of time spent in warning, and 33.21 minutes of average detection latency using 30-second EEG segments with 50% overlap and a simple postprocessing technique resulting in a decision (a seizure is expected/not expected) every 5 minutes. High sensitivity and low false positive rate with reasonable detection latency show that HHT based features are acceptable for patient specific seizure prediction from intracranial EEG data. Time spent for testing an EEG segment was 4.1451 seconds on average, which makes the system viable for use in real-time seizure control systems.

## 1. Introduction

Epilepsy is a neurological disorder that affects about 50 million people, in a wide age range from neonatal to elderly, around the world [1]. It is characterized by spontaneous recurrent seizures. A seizure occurs as a result of an excessive electrical discharge in a group of brain cells. Although it lasts for a short duration, a few seconds or minutes and rarely longer, a seizure is known to be handicapping for an epilepsy patient due to its unpredictable nature. However, patients with medically controllable epilepsy might sustain a normal life if they were able to predict an upcoming seizure in a reasonable time period. A recent report released by the World Health Organization shows that 70% of epilepsy patients respond to medication and can completely take their seizures under control [2]. Therefore prediction of an upcoming seizure could improve the life quality of the majority of epileptic patients by providing the time to take necessary

precautions to prevent or decrease the possible severe effects of the seizure.

EEG is a widely accepted method for diagnosis of epilepsy. Recently, a number of studies have shown that EEG recordings also carry important information minutes prior to the seizure onset to distinguish between preictal and interictal states [3–15]. Besides, technological advances in digital EEG and implantable devices for seizure control [16, 17] have increased the attention to algorithms for automated epileptic seizure prediction systems based on EEG data with high sensitivity and specificity [4, 9, 18]. Most of the studies use analysis techniques for feature extraction and machine learning algorithms to classify the features extracted from EEG segments into interictal or preictal states. Various feature extraction methods are used, such as spectral power [9], wavelet coherence [19], wavelet energy and entropy [13], short-time Fourier transform [20], mean phase coherence [10], and empirical mode decomposition (EMD) [21–23].

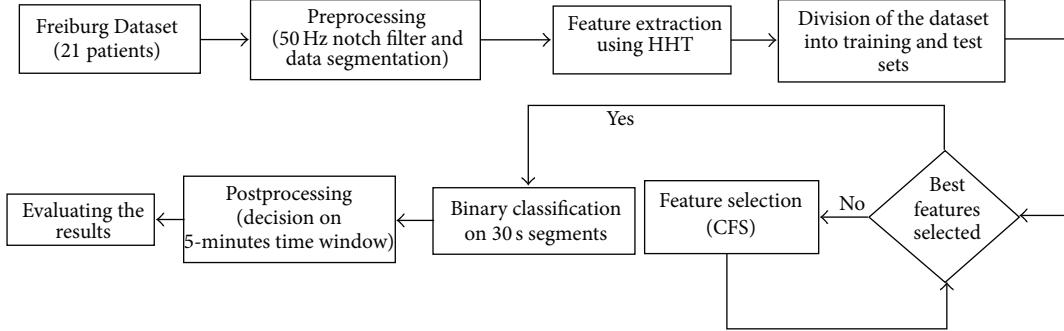


FIGURE 1: Flow chart of the proposed seizure prediction system.

Neural networks [20, 24] and support vector machines [9, 21, 25] are mostly preferred machine learning algorithms to classify the features extracted from EEG segments. Recently, Bayesian based methods are also employed in seizure prediction systems [26, 27].

EEG signals are spontaneous electrical brain activities that exhibit dynamic, stochastic, nonlinear, nonstationary, and also complex behaviour. Traditional data analysis techniques assume the time series to be linear and/or stationary. A relatively new signal processing technique, known as Hilbert-Huang transform (HHT), is proposed by Huang et al. in 1998 [28], specifically for analysing data from nonlinear and nonstationary processes. HHT was shown to be a powerful tool to examine biosignals, such as EEG and ECG [29]. Moreover, a recent study was conducted on a method for implementation of EMD for computationally efficient and accurate real-time analysis and the method's efficiency was shown for EEG and ECG datasets [30].

In this paper, we present a patient specific algorithm for seizure prediction based on HHT for feature extraction and Bayes networks for classification.

## 2. Materials and Methods

In this paper we propose a method that includes preprocessing of intracranial EEG (iEEG) data for noise removal and obtaining data segments for sliding window analysis, feature extraction using Hilbert-Huang transform, feature selection using correlation based feature selection algorithm, binary classification by Bayesian networks, and a simple postprocessing algorithm to remove spurious detections. Outline of the system is given in Figure 1. System is trained and tested on the Freiburg database. Test datasets are not used in classifier modelling.

**2.1. EEG Database.** The Freiburg EEG database, which is available online (<https://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database>) by request, is used for training and testing the proposed system. This database contains iEEG recordings collected from 21 patients with medically intractable focal epilepsy during invasive presurgical epilepsy monitoring at the Epilepsy Center of the University Hospital of Freiburg, Germany. Recordings were acquired

at 256 Hz (at 512 Hz for interictal recordings of patient 12) sampling rate with 16 bit A/D converter. Among the strip and grid electrodes used for recordings, six electrodes were identified by a certified epileptologist by visual inspection. First three electrodes (focal electrodes) are selected close to the region where the seizure occurs or the region where early ictal activity is detected, and the other three (extrafocal electrodes) are selected from the regions distal to the seizure focus or the regions in which an ictal activity is not observed. A total of 87 seizures, 504 h of interictal, and 73 h of preictal or ictal data are available in the database. At least 50 minutes of preictal data for each seizure and approximately 24 hours of iEEG data for each patient without seizure activity are available. Broader information can be obtained from [31]. Sample recordings of preictal, ictal, and interictal activities, extracted IMFs, and corresponding Hilbert spectrums for patient 2 are shown in Figure 2.

**2.2. Preprocessing: Artifact Removal and Segmentation.** We started with removing the artifacts in iEEG data, which were detected by an epileptologist visually and provided in the database in a text file. Then we filtered the data by a 50 Hz notch filter to remove the power line noise. For the analysis, we preferred moving window analysis technique and segmented the multichannel iEEG data into 30-second-long (7680 data points) windows with 50% overlap. For each window, we extracted HHT based features to feed the classifier to make a decision.

**2.3. Hilbert Huang Transform.** HHT is the designated name for an empirical and adaptive method which combines EMD, introduced by Huang et al. in 1996 [32], and very well known Hilbert spectral analysis (HSA). HHT is a time-frequency analysis technique in which the basis functions are data driven, in contrast to conventional data analysis techniques, such as Fourier transform and wavelet transform, which represent the signal by predetermined basis functions.

In this method, frequency component for each time point, called instantaneous frequency, is considered. Instantaneous frequency (IF) provides important information about the frequency content of the nonstationary signals. IF is easily obtained for complex valued signals by differentiating the phase of the signal with respect to time. However, it is

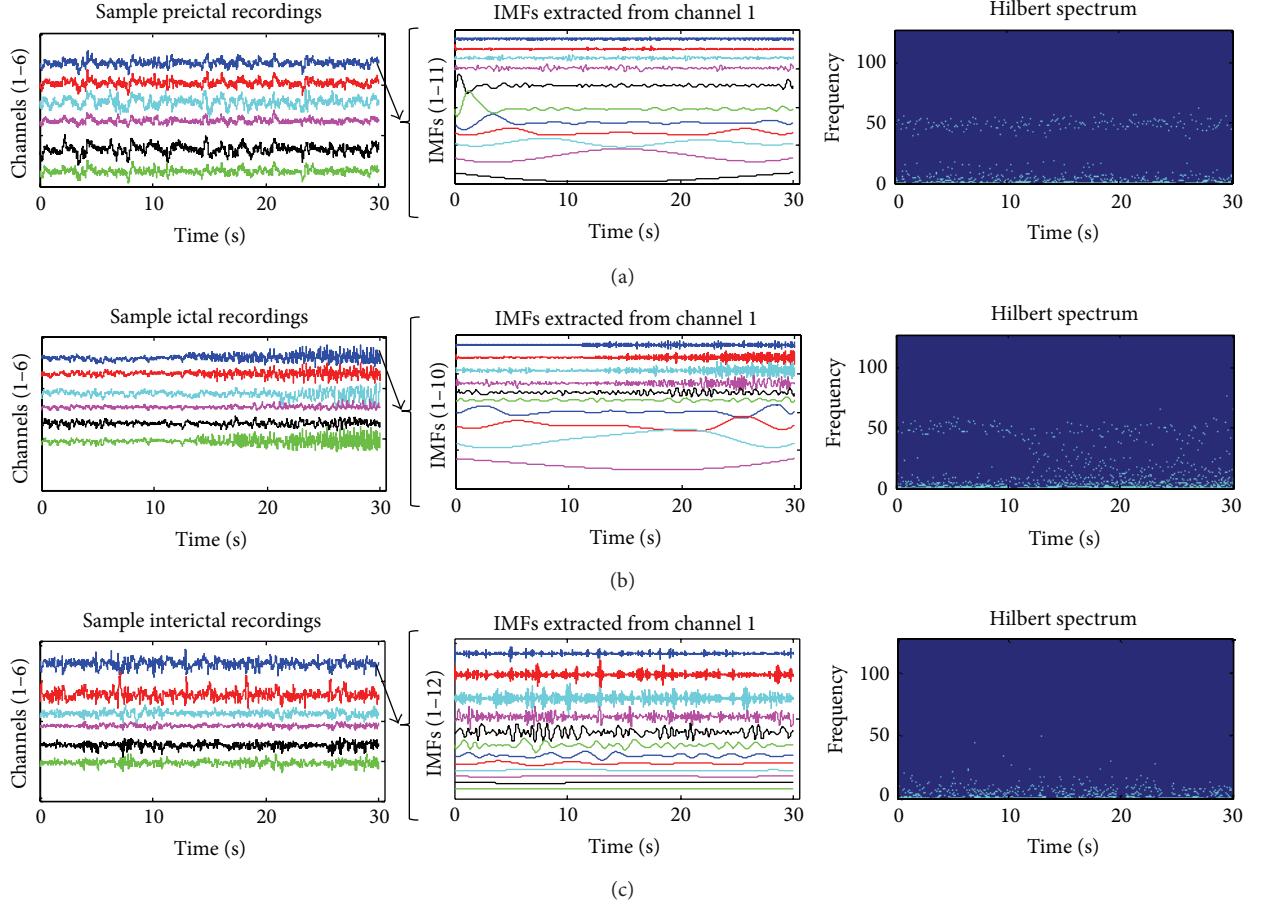


FIGURE 2: Sample EEG recordings, extracted IMFs and Hilbert spectrums for preictal activity (a), ictal activity (b), and interictal activity (c).

ambiguous for real valued signals. A possible way to compute IF for a real valued signal,  $g(t)$ , is to define an analytic signal,  $\tilde{g}(t)$ , by determining the complex conjugate of the signal,  $\hat{g}(t)$ , using the Hilbert transform [33]. Hilbert transform is simply defined as the convolution between the signal and the Hilbert transformer,  $1/(\pi t)$ . Instantaneous frequency,  $w(t)$ , is defined as the derivative of the phase function with respect to time;  $w(t) = d\theta(t)/dt$ , where  $G(t)$  is the instantaneous amplitude and  $\theta(t)$  is the phase function of the analytical signal. Although the problem of obtaining the phase function is solved by the use of Hilbert transform, application of the Hilbert transform is limited to monocomponent signals [34], in which there is only one frequency component or a narrow range of frequencies varying as a function of time. After the introduction of EMD, which is used to break down a multicomponent signal into its components called intrinsic mode functions (IMF), benefits of the method became applicable for multicomponent signals. EMD is based on the local characteristics of the data instead of predefined basis functions. Therefore it is highly efficient and suitable for analysing nonlinear and nonstationary signals. Process of decomposing a signal into its IMFs is pretty well defined in the literature [28, 32].

Assuming that we have a real valued signal,  $x(t)$ , EMD process is applied to the signal to decompose the signal into

its IMFs. Considering the signal is decomposed into  $N$  IMFs, the original signal can be expressed as the sum of IMFs,  $g_i(t)$ ,  $i = 1, \dots, N$ , and a residue,  $r_N$ . Consider

$$x(t) = \sum_{i=1}^N g_i(t) + r_N. \quad (1)$$

After obtaining IMFs, Hilbert transform is applied to extracted IMFs in order to obtain the time-frequency-energy distribution of the signal. After removing the residue, the original signal is now expressed as

$$x(t) = \text{Re} \left\{ \sum_{i=1}^N G_i(t) e^{j \int w_i(t) dt} \right\}. \quad (2)$$

Equation (2) enables us to observe the signal's energy distribution as a function of time and frequency in a 3D plot which is designated as the Hilbert spectrum, that is,  $H(w, t)$ . Sample iEEG segments, IMFs extracted from these segments and their Hilbert spectrums are shown in Figure 2.

**2.4. Feature Extraction.** 30-second-long iEEG segments for all six channels are decomposed into IMFs via EMD separately. Employing all the extracted IMFs, Hilbert spectrums of the signals are obtained. Then, the total energy,  $E_{\text{Tot}}$ , energy

in  $j$ th frequency band, namely,  $E_j$ , where  $j = \{\delta(0.5\text{--}3\text{ Hz}), \theta(4\text{--}7\text{ Hz}), \alpha_1(8\text{--}9\text{ Hz}), \alpha_2(10\text{--}12\text{ Hz}), \beta_1(13\text{--}17\text{ Hz}), \beta_2(18\text{--}30\text{ Hz}), \gamma(31\text{--}40, 41\text{--}50, 51\text{--}90, \text{ and } 91\text{--}128\text{ Hz})\}$ , and contribution of  $E_j$  to the total energy, that is,  $E_j/E_{\text{Tot}}$ , are computed. Note that wide gamma band was split into four bands.

Assuming that  $g_i^k(t)$  is the  $i$ th IMF for the  $k$ th channel,  $i = 1, 2, \dots, N$ ,  $k = 1, 2, \dots, 6$ ,  $H_i^k(w, t)$  is the associated Hilbert spectrum and  $H^k(w, t)$  is the Hilbert spectrum of the EEG data from  $k$ th channel.  $E_j^k$  is the energy in  $j$ th frequency band and  $E_{\text{Tot}}^k$  is the marginal energy for the  $k$ th channel. Besides, groupiness factors of the energy in each frequency band are also used as features. Groupiness factor is frequently used in oceanography studies to examine wave grouping characteristics. Funke and Mansard [35] defined groupiness factor of a time series (or its spectral density) to describe the degree of grouping activity as the standard deviation of the time series over the mean of the time series. Groupiness factor based on instantaneous energy obtained by HHT was defined in [36]. Following this approach, in this study, we used groupiness factor values, standard deviation of the energy in previously defined frequency bands over the mean of the same energy values, as features. Energy in  $j$ th frequency band from channel  $k$  was used to compute  $GF_j^k$ . Listed features are estimated as in (3). Consequently, 31 features are obtained for each channel. Combining these features for each segment, finally we obtain 186 features for each window. Consider

$$\begin{aligned} H^k(w, t) &= \sum_{i=1}^N H_i^k(w, t), \\ E_j^k(t) &= \int_{w \in j} (H^k(w, t))^2 dw, \\ E_j^k &= \int_t E_j^k(t) dt, \\ E_{\text{Tot}}^k &= \int_t \int_w (H^k(w, t))^2 dw dt, \\ GF_j^k &= \frac{\sigma_j^k}{\bar{E}_j^k}. \end{aligned} \quad (3)$$

**2.5. Feature Selection.** Feature selection is a technique that removes irrelevant and/or redundant features from the whole feature set and selects the most relevant ones to build a model with a better performance. Filters [37] and wrappers [38] are two general evaluation strategies for feature selection.

Wrappers search through the space of possible features and evaluate each feature subset using classifier accuracy as the objective function. Filter objective functions, on the other hand, are usually simpler measures such as information gain ratio or feature-feature and feature-class correlation. Although best performance is obtained using exhaustive search through all possible feature combinations, it is unfeasible and time consuming. Therefore, suboptimal but faster

TABLE 1: Most selected features selected by CFS.

Channel	Feature
Channel 4	Total energy in $\gamma$ (91–128)
Channel 2	Total energy in $\gamma$ (41–50)
Channel 4	Total energy in $\gamma$ (51–90)
Channel 6	Total energy in $\gamma$ (91–128)
Channel 4	Total energy in $\gamma$ (41–50)
Channel 5	Total energy in $\gamma$ (51–90)
Channel 6	Total energy in $\gamma$ (41–50)
Channel 6	Total energy In $\gamma$ (51–90)
Channel 5	Total Energy in $\gamma$ (91–128)
Channel 1	Total energy in $\gamma$ (41–50)
Channel 1	Total energy in $\gamma$ (91–128)
Channel 2	Total energy in $\alpha_1$ (8–9)
Channel 3	Total energy in $\gamma$ (91–128)
Channel 2	Total energy in $\gamma$ (51–90)
Channel 2	GF in $\gamma$ (91–128)
Channel 5	Total energy in $\gamma$ (41–50)
Channel 6	Total energy in $\beta_2$ (18–30)

search functions, such as hill-climbing, genetic, best first, and random, are usually chosen.

In this study, we employed a filtering type of feature selection, correlation based feature selection (CFS) [39] with best first search algorithm. The fundamental idea behind this method is that a good feature set consists of the features that are correlated with the class but are not correlated with each other. In our study, the number of features is reduced from 186 to 14.49 on the average ( $\sigma = 6.95$ ). Table 1 shows the most selected features that are selected more than 15 times in 87 trials by CFS.

**2.6. Train-Test Files.** In our study we used Freiburg database for training and testing the system. Dataset used for testing is left completely out of training dataset in order to present more realistic sensitivity and specificity values. In the database there are 87 seizures for 21 patients. In order to test the system for all seizures, 87 patient specific training and testing datasets are constructed considering the number of seizures the patient has and the number of preictal files recorded for the same patient. Assuming that  $i$ th patient has  $N_i$  seizures and  $M_i$  files containing interictal recordings,  $N_i$  disjoint datasets are obtained as shown in Figure 3. In the train-test process, set  $k$  ( $k = 1, 2, \dots, N_i$ ) is left for testing and the remaining datasets are combined for training assuring that the test datasets are not used in classifier modelling. Therefore, contiguous recordings, which are highly correlated and might result in overoptimistic specificity results, were included either in training or in testing sets. Note that interictal files contain about 1-hour-long recordings ( $\mu_{\text{filelength}} = 58.58$  minutes,  $\sigma_{\text{filelength}} = 6.97$  minutes).

**2.7. Bayesian Networks Classifiers.** Bayesian networks have become increasingly popular and were applied in various different areas, for example, biology, financial applications,

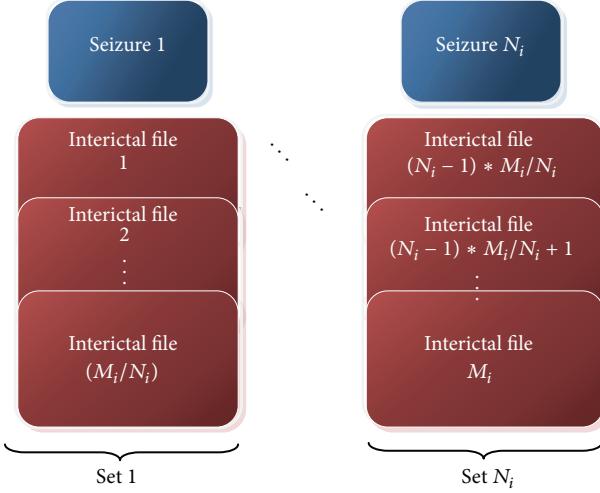


FIGURE 3: Division of dataset to be used for training and testing for  $i$ th patient with  $N_i$  seizures.

causal learning, computer games, computer vision, data mining, medicine, natural language processing, and speech recognition [40]. Bayesian network method [41] is a directed acyclic graph (DAG) that represents the joint probability distribution over a set of random variables and searches for the causal connections between variables using probability and graph theory [42].

Considering a finite set of random variables,  $X = \{X_1, \dots, X_n\}$ , a Bayesian Network,  $B$ , for  $X$  is a directed acyclic graph. A DAG includes nodes, which correspond to the random variables, and edges which represent the conditional dependencies. Each node is annotated with a conditional probability distribution, that is,  $\Theta$ , which represents  $P(X_i | \text{par}(X_i))$  where  $\text{par}(X_i)$  denotes the parents of  $X_i$ . Finally the Bayesian Network represents a unique joint probability distribution over  $X$ :

$$\begin{aligned} P_B(X_1, \dots, X_n) &= \prod_{i=1}^n P_B(X_i | \text{par}(X_i)) \\ &= \prod_{i=1}^n \Theta_{X_i|\text{par}(X_i)}. \end{aligned} \quad (4)$$

To use a Bayesian Network as a classifier, we define the variable set as  $X = \{X_1, \dots, X_n, C\}$ , where  $X_i$  are the attributes and  $C$  is the class variable. A Bayesian Network,  $B$  with joint probability distribution  $P_B(X_1, \dots, X_n, C)$  is formed and for a given set of attributes  $x_1, \dots, x_n$ , the class label  $c$ , which maximizes the posterior probability  $P_B(c | x_1, \dots, x_n)$ , is obtained based on the network  $B$ .

In this work, we used WEKA implementation of Bayesian networks [43, 44] for classification.

**2.8. Postprocessing.** The ultimate goal of the study is to define a system which could predict an upcoming seizure with sufficient time to prepare for the seizure or to take the necessary precautions to prevent it. The system could be used

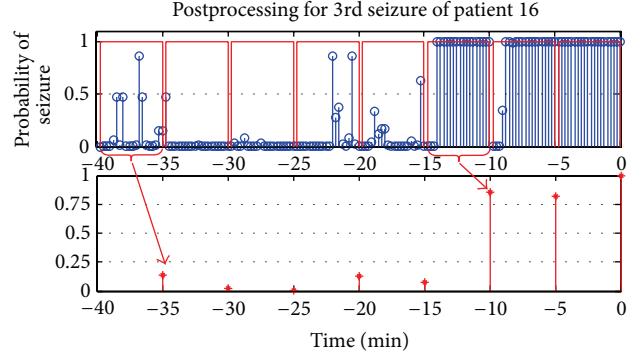


FIGURE 4: An example of postprocessing procedure shown for 3rd seizure of patient 16. Upper row shows the probability of having a seizure in near future obtained every 15 seconds (red rectangles show 5-minute windows). Lower row shows the mean probabilities of having a preictal state in 5-minute time intervals.

in or with an implantable device and triggers an alarm if a seizure is expected. In presented system, such a decision is made every 15 seconds. Considering the negative effects of hearing an alarm every 15 seconds when a seizure is expected, very simple postprocessing is performed by making a decision every 5 minutes depending on the classification probabilities in this time window. The average probability of having a preictal state is computed for every 5-minute decision window. This average probability value is used for thresholding in performance assessment. For instance, for a threshold value of 0.25, a preictal classification is made only if the average probability in the five-minute decision window is higher than 0.25 and no seizure is expected in near feature (interictal classification) otherwise. Figure 4 shows an example of the postprocessing applied on 40-minute-long preseizure data. Figure 5 shows sample demonstration of thresholding process for preictal and interictal recordings.

### 3. Results and Discussion

We tested our patient specific algorithm for all patients in the Freiburg dataset. We examined a total of 58 hours (87 seizures \* 40 min/seizure) of preictal data and 503.1 hours of interictal data. Results are presented in sensitivity, mean detection latency, number of false positives per hour (FPs/h), and time spent in warning (FP%) measures to show the performance of the algorithm. Detection latency is defined as the time between the alarm and the seizure onset.

Sensitivity is calculated as the ratio of correctly predicted seizures to all seizures. A correct prediction is assessed in the sense of seizure prediction horizon (SPH) and seizure occurrence period (SOP). SOP is the time period during which the seizure onset is expected and SPH is defined as the minimum time window between the alarm and the beginning of the SOP [45]. In case of an alarm, the system holds on for a time period of SPH + SOP; if a seizure occurs in this process prediction is recorded as a true positive and as a false positive otherwise. In this study we define SPH as 5 minutes and SOP as 35 minutes. A prediction is accepted as correct only if the

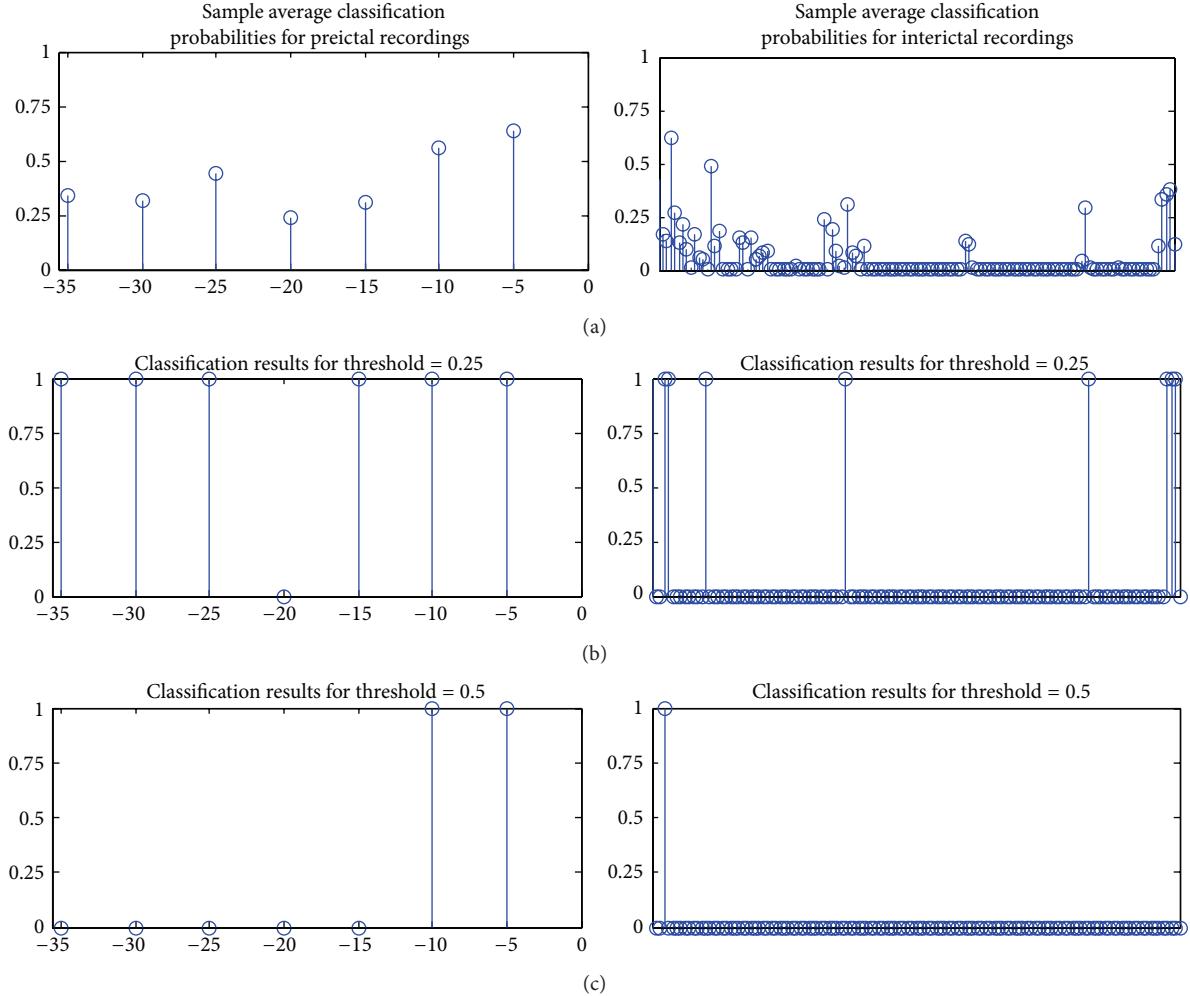


FIGURE 5: Sample demonstration of thresholding. Upper row shows the probability of having a seizure in near future. Middle row shows the classification results using 0.25 as threshold. For this preictal recordings, seizure is accepted, predicted 35 minutes prior to the seizure; that is, detection latency is 35 minutes. For the interictal recordings number of false positives is recorded as 5; consecutive alarm in a SPH + SOP window is counted as one. Lower row shows the classification results using 0.5 as threshold and the detection latency now 10 minutes and number of false positives is 1.

seizure occurs in a time period of 35 minutes starting five minutes after the alarm. Therefore if an alarm is triggered the patient will have at least five minutes to prepare but at most 40 minutes to wait for the seizure onset. Mean detection latency is recorded as the average time period between the alarms of true predictions and the following seizure onsets.

In a realistic seizure prediction system, false positives are inevitable. Number of excessive false positives, however, would decrease the credibility of the system. Therefore, an acceptable prediction system should provide a low number of false positives while resulting in high sensitivity. Number of false positives per hour (FPs/h) is a frequently used measure to evaluate the seizure prediction systems. FPs/h is computed as the division of the number of false predictions to the total interictal time examined, in hours, in system evaluation. Time spent in warning is also provided for each patient. A sample set of classification results for interictal recordings is

given in Figure 7. Note that in this representation number of FPs is 2 and the time spent in warning is 120 minutes.

Figure 6 shows the sensitivity versus FP/h values for different threshold values. Threshold value of 0.99 in post-processing yields 59.77% sensitivity with a FP/h rate of 0.048. Sensitivity quickly increases to 88.51% for a FP/h value of 0.147. Best sensitivity (96.55%) for the lowest FP/h (0.205) value is obtained for the threshold value of 0.25. For this point on the curve, a preictal classification is made only if the average classification probability in a five-minute time interval is higher than or equal to 0.25; that is, at least half of the windows are classified as ictal with a probability of 50% or more on the average.

Table 2 shows the seizure prediction results, corresponding to the FP/h value of 0.205, using the whole feature set and using only the features that are selected to be relevant and necessary using CFS algorithm. Note that CFS is employed

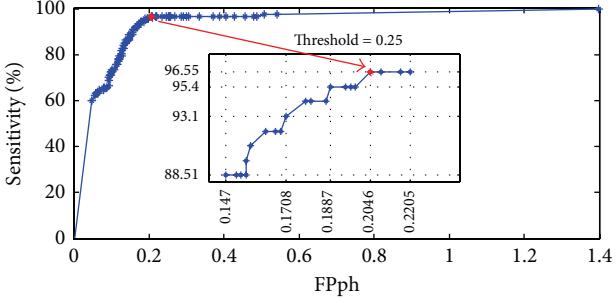


FIGURE 6: Sensitivity versus FPs/h values. Best sensitivity (96.55%) for lowest FP/h (0.2046) point is marked with a red dot.

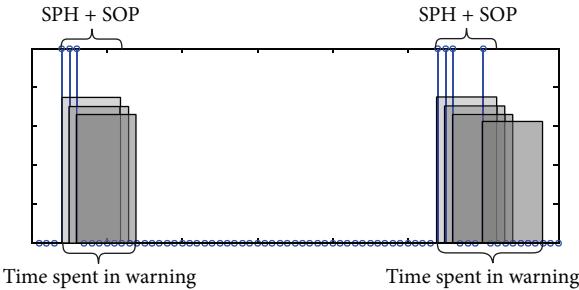


FIGURE 7: Classification results for interictal recordings of 6th patient's first test set and graphical representation for time spent in warning.

independently for each trial. CFS dramatically decreased the number of features to be used for seizure prediction which shows that the feature sets chosen in the beginning are highly correlated with each other. Another result of the correlated features is that there were 11 features that are not selected in any trial. These features are listed in Table 3. Features selected more than 15 times by CFS algorithm in all trials (87 trials; one for each seizure) are shown in Table 1. Table shows that the total energy values in different frequency bands, mostly in  $\gamma$  band, are effective features in seizure prediction. However, the number of times the features used in the trials shows that there is no feature that is dominantly used in the trials. Mostly chosen feature,  $E_{\gamma_4}^4$ , was selected in 31 trials. Besides, features selected for each patient are noticeably different from each other and therefore starting feature selection with a large feature set is reasonable for a better chance of finding the best feature set for the patient, although the final feature set is dramatically reduced after CFS.

Results show that presented prediction system has very high sensitivity of 96.55 with a low rate of false positives, compared to the seizure prediction systems tested on the Freiburg EEG database. Earliest studies on Freiburg dataset resulted in less than 60% sensitivity [3, 31, 45], with controlled false positive rates. Average sensitivity for FP/h rate of 0.25 was reported as 57.95% in [31], and maximum %42 sensitivity was achieved for 0.15 FP/h value in [3, 45]. Later studies achieved better performances: 77.8% sensitivity for 9 patients [46], 71% sensitivity with no false positives for 15 patients [47], and 74.2% sensitivity with less than 0.2 FPs/h for all

patients [19]. In [11], 11 patients with focal neocortical epilepsy were examined using a rule based prediction system and 79.9% and 90.2% of the 49 seizures were predicted for 30 min and 50 min SOP values with 0.17 and 0.11 FPs/h, respectively. For the patients with 3 or more seizures, 95% sensitivity was achieved in [14]. The best results on Freiburg database are demonstrated in [9] for the same patients with 97.5% sensitivity (78 of the 80 seizures were successfully predicted) and FPs/h rate of 0.27 for 18 of the 21 patients. In the study linear features of spectral power and support vector machines were used along with bipolar preprocessing and Kalman filters for postprocessing. In this study, we examined all patients in the database and obtained comparable sensitivity (96.55%) with better FPs/h rate of 0.21 on the average. Besides the detection latency is presented as 33.21 minutes. Note that if patients 2, 8, and 13 were excluded, as in the aforementioned study, sensitivity would be equivalent (78 out of 80) to the sensitivity reported in [9] with a better FP/h rate of 0.22 and the detection latency would be 33.46 minutes.

In order to statistically validate the results, the prediction performance of the system is compared to a random predictor. Basically, a valid prediction method should be superior to an unspecific random predictor that makes no use of any information contained in the EEG data.

**3.1. Statistical Validation.** A prediction method is considered to perform above chance level only if the performance is shown to be statistically significant. We tested our seizure prediction against a chance predictor, which is described in [8]. The chance predictor is based on a Poisson process where the interval between two consecutive alarms follows an exponential distribution [5, 8]. Note that most of the seizure prediction systems use this random predictor to show the significance of the results presented [3, 5, 8–13].

Sensitivity of the chance predictor is defined as

$$S_{nc} = 1 - \exp(-\lambda_w (SPH + SOP) + (1 - e^{-\lambda_w SPH})), \quad (5)$$

where  $\lambda_w$  is the Poisson rate for the chance predictor which is calculated as

$$\lambda_w = -\frac{1}{(SPH + SOP)} \ln(1 - \rho_w), \quad (6)$$

where  $\rho_w$  is the proportion of time spent in warning.

To assess the significance of the system's sensitivity,  $P$  value of the null hypothesis should be lower than a predefined value, mostly chosen to be 0.05. Null hypothesis is stated as “*The sensitivity of the system presented is not different than the sensitivity of the chance predictor.*” Given that the algorithm under evaluation correctly predicts  $n$  of  $N$  seizures, the one sided  $P$  value is estimated as [8]

$$P = 1 - \sum_{k=0}^{n-1} \binom{N}{k} S_{nc}^k (1 - S_{nc})^{N-k}, \quad \text{for } \frac{n}{N} \geq S_{nc}. \quad (7)$$

Prediction results are shown to be significantly better than a chance predictor for all patients except patient 19.

TABLE 2: Evaluation of the proposed seizure prediction system.

Patient number	Number of Sz.	Interictal hours	CFS					No feature selection				
			Sens. (%)	Det. Lat. (min)	FPs/h	FP%	P value	Sens. (%)	Det. Lat. (min)	FPs/h	FP%	P value
1	4	23.91	100	35	0	0	0	100	27.5	0	0	0
2	3	23.89	100	35	0.08	5.178	0.0001	66.67	27.5	0	0	0
3	5	23.73	100	34	0	0	0	100	29	0.08	5.296	0
4	5	23.91	100	33	0.04	2.500	0	100	33	0.04	2.500	0
5	5	23.64	100	27	0.51	35.000	0.0031	80	23.75	0.63	40.625	0.0067
6	3	23.56	100	35	0.42	33.443	0.0272	100	33.33	0.30	21.967	0.0075
7	3	24.50	100	35	0	0	0	100	35	0.16	10.760	0.0009
8	2	24.06	100	20	0.21	17.105	0.023	100	22.5	0.21	14.803	0.0172
9	5	23.84	100	35	0	0	0	100	35	0.04	2.477	0
10	5	24.35	100	35	0.21	16.109	0.0001	100	35	0.33	27.964	0.001
11	4	23.02	75	35	0	0	0	75	35	0.22	18.301	0.015
12	4	24.63	100	35	0.04	2.462	0	100	35	0.12	9.846	0.0001
13	2	23.78	50	35	0	0	0	0	—	0	0	1
14	4	23.30	100	31.25	0.39	25.566	0.0027	75	30	0.56	41.100	0.1485
15	4	23.75	100	35	0.08	6.349	0	100	35	0.25	17.143	0.0005
16	5	23.92	100	29	0.42	24.768	0.0005	100	31	0.42	24.768	0.0005
17	5	23.98	100	35	0.29	18.154	0.0001	100	35	0.38	21.231	0.0002
18	5	24.79	100	35	0.20	13.773	0	100	35	0.40	22.455	0.0003
19	4	24.25	75	35	0.82	53.125	0.2931	75	35	0.78	53.125	0.2931
20	5	24.79	100	35	0.52	32.537	0.0021	80	35	0.65	48.358	0.1224
21	5	23.85	100	31	0.04	4.334	0	100	35	0.25	20.433	0.0002
	<b>87</b>	<b>503.45</b>	<b>96.55</b>	<b>33.21</b>	<b>0.21</b>	<b>13.896</b>	<b>0.000</b>	<b>90.8</b>	<b>32.53</b>	<b>0.28</b>	<b>19.347</b>	<b>0.000</b>

TABLE 3: Features never selected by CFS.

Channel	Feature
Channel 1	$GF_{\alpha_1}, E_{\gamma_1}/E_{\text{Tot}}$
Channel 2	$E_{\gamma_1}/E_{\text{Tot}}$
Channel 3	$GF_{\beta_1}, GF_{\beta_2}$
Channel 4	$E_{\text{Tot}}, GF_{\theta}, GF_{\alpha_1}, E_{\alpha_1}/E_{\text{Tot}}$
Channel 5	$GF_{\alpha_1}, GF_{\gamma_1}$

## 4. Conclusions

In this paper, a seizure prediction system based on Hilbert-Huang transform and Bayesian classification is presented. The system, basically, has 4 steps after preprocessing, extracting the features using HHT, selecting the best features for each patient by CFS, classification by means of Bayesian networks, and applying a simple postprocessing algorithm which combines the individual probabilities, obtained after classification task, in a 5-minute window. Results show that such a system can successfully predict the seizures with high sensitivity, 96.55% (84 out of 87 seizures) and a low false positive rate of 0.21 per hour (about one false alarm every 4.88 hours on the average). Sensitivity results are validated to be significantly better than a chance predictor for all patients

except patient 19 for a significance level of 0.05. Excluding patient 19, the sensitivity would be 97.59% (81 of 83 seizures) for a FP/h rate of 0.173 (11.92%).

The system perfectly predicts the seizures with a FP/h rate of lower than 0.08 (1 FP in 12 hours) at least 31 minutes prior to the seizure onsets for nine patients (patients 1–4, 7, 9, 12, 15, and 21). Only 3 seizures (one for each patients 11, 13, and 19) were missed by the system. For four of the remaining patients, seizures were predicted at least 20 minutes before the seizure occurs with a chance of having only one false alarm in 3 hours to 5 hours.

Presented system is a patient specific seizure prediction system in which the system is trained and tested for each patient separately. By examining the features selected for each patient by CFS, there was no remarkable feature that is significant for all patients. Therefore we can conclude that the patterns followed by the extracted features for each patient are different from each other and each patient should be administered separately with the presented system. In general, however, the energy contribution of the frequency bands are shown to be the mostly selected features for the system. Note that the band that makes the contribution is different for each patient.

Performance of this system may be compared to other seizure prediction systems which used the Freiburg EEG

database for training and testing. Some algorithms resulted in low sensitivity (mostly less than 50%) [3, 31, 45]. Better results are presented in other studies; however, in most of the studies, results are given only for selected patients [9, 47]. The best results on Freiburg database are demonstrated in [9] with 97.5% sensitivity with a FPs/h rate of 0.27 for 18 of the 21 patients. In this study, we examined all patients of the database and obtained comparable sensitivity (96.55%) with a better FPs/h rate of 0.21 on the average, noting that the same sensitivity is obtained for a FPs/h rate of 0.22 for those 18 patients examined in [9]. Besides detection latency is also reported as 33.21 minutes on average.

The system was trained and tested on an average desktop computer and analysis is performed using MATLAB. Time spent for processing a 30-second EEG segment was 4.1451 seconds on average, which is much smaller than required time to process 30-second data with 50% overlap (15 seconds) which makes the system viable for use in a real-time system. Average time spent on each step is as follows: 0.0754 seconds for converting database file format (.asc) to MATLAB data file format (.mat), 3.7848 seconds for obtaining IMFs from filtered data, 0.1986 seconds for extracting features for all channels and combining them into a one feature set, and 0.0863 seconds for testing and classifying the segment using the model based on the training set and applying postprocessing.

This study shows that HHT is a sensible method to extract the relevant features in a seizure prediction system. The system may be improved by using different preprocessing algorithms, choosing the optimal channels for feature extraction and employing wrapper methods for feature selection to optimize the classifier performance.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

This work was supported by the Turkish Scientific and Technical Research Council (TUBITAK) under Project no. 109E223. The authors are very grateful to the Epilepsy Center of the University Hospital of Freiburg, Germany, for their consent to use the invasive EEG recordings in this work.

## References

- [1] T. Dua, H. M. De Boer, L. L. Prilipko, and S. Saxena, "Epilepsy care in the world: results of an ILAE/IBE/WHO Global Campaign Against Epilepsy survey," *Epilepsia*, vol. 47, no. 7, pp. 1225–1231, 2006.
- [2] WHO, "Programmes and projects," fact sheet no. 999, 2009.
- [3] M. Winterhalder, T. Maiwald, H. U. Voss, R. Aschenbrenner-Scheibe, J. Timmer, and A. Schulze-Bonhage, "The seizure prediction characteristics: A general framework to assess and compare seizure prediction methods," *Epilepsy and Behavior*, vol. 4, no. 3, pp. 318–325, 2003.
- [4] F. Mormann, T. Kreuz, C. Rieke et al., "On the predictability of epileptic seizures," *Clinical Neurophysiology*, vol. 116, no. 3, pp. 569–587, 2005.
- [5] B. Schelter, M. Winterhalder, T. Maiwald et al., "Testing statistical significance of multivariate time series analysis techniques for epileptic seizure prediction," *Chaos*, vol. 16, Article ID 013108, 2006.
- [6] S. Sanei and J. Chambers, Eds., *EEG Signal Processing*, Wiley, Chichester, UK, 2007.
- [7] J. Sackellares, "Seizure prediction," *Epilepsy Currents*, vol. 8, pp. 55–59, 2008.
- [8] D. E. Snyder, J. Echauz, D. B. Grimes, and B. Litt, "The statistics of a practical seizure warning system," *Journal of Neural Engineering*, vol. 5, no. 4, pp. 392–401, 2008.
- [9] Y. Park, L. Luo, K. K. Parhi, and T. Netoff, "Seizure prediction with spectral power of EEG using cost-sensitive support vector machines," *Epilepsia*, vol. 52, no. 10, pp. 1761–1770, 2011.
- [10] H. Feldwisch-Drentrup, M. Ihle, M. le van Quyen et al., "Anticipating the unobserved: prediction of subclinical seizures," *Epilepsy and Behavior*, vol. 22, no. 1, pp. S119–S126, 2011.
- [11] A. Aarabi and B. He, "A rule-based seizure prediction method for focal neocortical epilepsy," *Clinical Neurophysiology*, vol. 123, no. 6, pp. 1111–1122, 2012.
- [12] A. S. Zandi, R. Tafreshi, M. Javidan, and G. A. Dumont, "Predicting epileptic seizures in scalp EEG based on a variational bayesian gaussian mixture model of zero-crossing intervals," *IEEE Transactions on Biomedical Engineering*, vol. 60, no. 5, pp. 1401–1413, 2013.
- [13] K. Gadhoumi, J. Lina, and J. Gotman, "Seizure prediction in patients with mesial temporal lobe epilepsy using EEG measures of state similarity," *Clinical Neurophysiology*, vol. 124, no. 9, pp. 1745–1754, 2013.
- [14] J. R. Williamson, D. W. Bliss, D. W. Browne, and J. T. Narayanan, "Seizure prediction using EEG spatiotemporal correlation structure," *Epilepsy and Behavior*, vol. 25, no. 2, pp. 230–238, 2012.
- [15] V. Joshi, R. B. Pachori, and A. Vijesh, "Classification of ictal and seizure-free EEG signals using fractional linear prediction," *Biomedical Signal Processing and Control*, vol. 9, pp. 1–5, 2014.
- [16] W. C. Stacey and B. Litt, "Technology insight: neuroengineering and epilepsy—designing devices for seizure control," *Nature Clinical Practice Neurology*, vol. 4, no. 4, pp. 190–201, 2008.
- [17] M. J. Cook, T. J. O'Brien, S. F. Berkovic et al., "Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study," *The Lancet Neurology*, vol. 12, no. 6, pp. 563–571, 2013.
- [18] A. B. Gardner, A. M. Krieger, G. Vachtsevanos, and B. Litt, "One-class novelty detection for seizure analysis from intracranial EEG," *Journal of Machine Learning Research*, vol. 7, pp. 1025–1044, 2006.
- [19] C.-Y. Chiang, N.-F. Chang, T.-C. Chen, H.-H. Chen, and L.-G. Chen, "Seizure prediction based on classification of EEG synchronization patterns with on-line retraining and post-processing scheme," in *Proceedings of the IEEE Conference of Engineering in Medicine and Biology Society (EMBC '11)*, pp. 7564–7569, Morgan Kaufmann, Boston, Mass, USA, 2011.
- [20] A. T. Tzallas, M. G. Tsipouras, and D. I. Fotiadis, "Epileptic seizure detection in EEGs using time-frequency analysis," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 703–710, 2009.

- [21] F. Duman, N. Özdemir, and E. Yildirim, "Patient specific seizure prediction algorithm using Hilbert-Huang transform," in *Proceedings of the IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI '12)*, pp. 705–708, January 2012.
- [22] N. Özdemir and E. Yildirim, "Epileptic seizure prediction based on hilbert huang transform and artificial neural networks," in *Proceedings of the 20th Signal Processing and Communications Applications Conference (SIU '12)*, pp. 1–4, April 2012.
- [23] A. K. Tafreshi, A. M. Nasrabadi, and A. H. Omidvarnia, "Empirical mode decomposition in epileptic seizure prediction," in *Proceeding of the 8th IEEE International Symposium on Signal Processing and Information Technology (ISSPIT '08)*, pp. 275–280, Sarajevo, Bosnia and Herzegovina, December 2008.
- [24] R. Schuyler, A. White, K. Staley, and K. Cios, "Epileptic seizure detection," *IEEE Engineering in Medicine and Biology Magazine*, vol. 26, no. 2, pp. 74–81, 2007.
- [25] M. I. Gürsoy and A. Subaşı, "A comparison of PCA, ICA and LDA in EEG signal classification using SVM," in *Proceedings of the IEEE 16th Signal Processing, Communication and Applications Conference (SIU '08)*, pp. 1–4, Aydin, Turkey, April 2008.
- [26] S. Santaniello, D. L. Sherman, M. A. Mirski, N. V. Thakor, and S. V. Sarma, "A Bayesian framework for analyzing iEEG data from a rat model of epilepsy," in *Proceedings of the 33rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS '11)*, pp. 1435–1438, September 2011.
- [27] P. Jallon, "A Bayesian approach for epileptic seizures detection with 3D accelerometers sensors," in *Proceedings of the 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC '10)*, pp. 6325–6328, September 2010.
- [28] N. E. Huang, Z. Shen, S. R. Long et al., "The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1971, pp. 903–995, 1998.
- [29] A. Eftekhar, F. Vohra, C. Toumazou, E. M. Drakakis, and K. Parker, "Hilbert-huang transform: preliminary studies in epilepsy and cardiac arrhythmias," in *Proceedings of the IEEE Biomedical Circuits and Systems Conference (BIOCAS '08)*, pp. 373–376, Baltimore, Md, USA, November 2008.
- [30] A. Eftekhar, C. Toumazou, and E. M. Drakakis, "Empirical mode decomposition: real-time implementation and applications," *Journal of Signal Processing Systems*, vol. 73, no. 1, pp. 43–58, 2013.
- [31] R. Aschenbrenner-Scheibe, T. Maiwald, M. Winterhalder, H. U. Voss, J. Timmer, and A. Schulze-Bonhage, "How well can epileptic seizures be predicted? an evaluation of a nonlinear method," *Brain*, vol. 126, no. 12, pp. 2616–2626, 2003.
- [32] N. E. Huang, S. R. Long, and Z. Shen, "The mechanism for frequency downshift in nonlinear wave evolution," *Advances in Applied Mechanics*, vol. 32, pp. 59–117, 1996.
- [33] E. Saff and S. AD, *Complex Analysis for Mathematics, Science and Engineering*, Prentice-Hall, New York, NY, USA, 1976.
- [34] L. Cohen, *Time-frequency Analysis*, Prentice Hall, Englewood Cliffs, NJ, USA, 1995.
- [35] E. Funke and E. Mansard, "On the synthesis of realistic sea states," in *Proceedings of the 17th International Conference on Coastal Engineering*, vol. 80, pp. 2974–2991, 1980.
- [36] J.-B. Saulnier, A. Clément, A. F. de O. Falcão, T. Pontes, M. Prevosto, and P. Ricci, "Wave groupiness and spectral bandwidth as relevant parameters for the performance assessment of wave energy converters," *Ocean Engineering*, vol. 38, no. 1, pp. 130–147, 2011.
- [37] G. H. John, R. Kohavi, and K. Pfleger, "Irrelevant features and the subset selection problem," in *Proceedings of the 11th International Conference on Machine Learning (ICML '94)*, pp. 121–129, Morgan Kaufmann, New Brunswick, NJ, USA, 1994.
- [38] R. Kohavi and G. H. John, "Wrappers for feature subset selection," *Artificial Intelligence*, vol. 97, no. 1-2, pp. 273–324, 1997.
- [39] M. Hall, *Correlation-based feature selection for machine learning [Ph.D. thesis]*, University of Waikato, 1999.
- [40] R. E. Neapolitan, *Learning Bayesian Networks*, Prentice Hall, 2003.
- [41] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Francisco, Calif, USA, 1988.
- [42] N. Friedman, D. Geiger, and M. Goldszmidt, "Bayesian network classifiers," *Machine Learning*, vol. 29, no. 2-3, pp. 131–163, 1997.
- [43] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: an update," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 10–18, 2009.
- [44] R. R. Bouckaert, "Bayesian network classifiers in weka for version 3-5-7," Tech. Rep., Waikato University, Hamilton, New Zealand, 2008.
- [45] T. Maiwald, M. Winterhalder, R. Aschenbrenner-Scheibe, H. U. Voss, A. Schulze-Bonhage, and J. Timmer, "Comparison of three nonlinear seizure prediction methods by means of the seizure prediction characteristic," *Physica D: Nonlinear Phenomena*, vol. 194, no. 3-4, pp. 357–368, 2004.
- [46] Y. Park, K. Parhi, and T. Netoff, "Seizure prediction by spectral power in nine bands using cost-sensitive svm," in *Proceedings of the 4th International Workshop on Seizure Prediction (IWSP4 '09)*, pp. 1761–1770, Kansas City, Mo, USA, 2009.
- [47] P. Mirowski, D. Madhavan, Y. LeCun, and R. Kuzniecky, "Classification of patterns of EEG synchronization for seizure prediction," *Clinical Neurophysiology*, vol. 120, no. 11, pp. 1927–1940, 2009.

