Patient-aware EEG-based Feature and Classifier Selection for e-Health Epileptic Seizure Prediction

Mohammad Nassralla[1], Mortada Haidar[1], Hussein Alawieh[1], Ahmad El Hajj[2], and Zaher Dawy[1]
[1] Department of Electrical and Computer Engineering, American University of Beirut, Lebanon.
[2] Department of Electrical and Computer Engineering, Beirut Arab University, Lebanon
Email: zd03@aub.edu.lb

***Abstract*—With the advent of wireless EEG headsets and the improved computational capabilities of mobile devices, the ubiquitous monitoring of patients suffering from epilepsy has recently gained notable interest from research and industry under the umbrella of emerging e-health systems. Among others, the problem of seizure prediction plays a central role in reducing the adverse effects of epileptic seizures by trying to foresee the occurrence of an epileptic seizure before its onset. This paper presents a patient-aware seizure prediction approach which combines an optimized spatio-temporal EEG feature extraction algorithm and classifier selection to maximize the prediction accuracy and reduce the false alarm rate. Experimental results demonstrate that the proposed approach leads to promising prediction results when tested on real clinical data from the Freiburg database, with accuracy exceeding 90% for most subjects. The key differentiating aspect of the proposed approach is its flexibility to be efficiently fine-tuned and optimized per subject in order to enhance sensitivity and minimize the false alarm rate.**

I. Introduction

Around 1% of the world population (50 million) suffer from epilepsy [1]. Epileptic patients may suffer from intractable seizures which are likely to result in many injuries such as damage to the neural tissues, fractures, submersion, burns, accidents and even death. The high percentage of epileptic patients and the number of deaths per one thousand people per year of sudden unexpected death in epilepsy motivate researchers to design algorithms that predict the occurrence of seizure some time before its onset [2]. Real-time prediction of epileptic seizures using e-Health systems helps in preventing unwanted effects if it is done with a certain level of accuracy and short enough delay to allow the subject to take precautionary actions.

Recent studies reveal that monitoring of brain activities at the scalp to discover electrochemical disturbances in the neurons is possible by using the electroencephalogram (EEG) which in turn can help in predicting seizures [3]. EEG is a typical method for acquiring the brainwaves, where electrodes are positioned on the scalp and sense the micro-volt sized signals that end outside the head due to the synchronized neuronal action within the brain [1]. The frequency of EEG brain signals ranges from 0 to 60 Hz; *Delta* (0-4 Hz), *theta* (4-8 Hz), *alpha* (8-12 Hz), *beta* (12-30 Hz), and *gamma* (30-60 Hz) [4]. The amplitudes of EEG signals typically range from $5\mu V$ to $200\mu V$ [5].

Extensive work has been done to predict the epileptic seizures over the years, where prolonged EEG recordings are used to distinguish the three states of epileptic brain activity, namely, the interictal state (period between seizures), preictal state (period before the seizure onset) and ictal state (seizure period). This is done by extracting distinctive features from the EEG traces that allows to delimit the three states and provide clues for accurate seizure prediction. The feature extraction approach is followed by a machine learning stage where feature vectors corresponding to interictal and preictal stages are used to establish a classification rule that will eventually categorize EEG traces as belonging to an interictal stage (no forthcoming seizure) or preictal stage (incoming seizure predicted).

In [6], the short time Fourier transform is evaluated for each EEG segment then the fractional energy of each time-frequency window is measured and used to train an artificial neural network. The spectral power in the *delta*, *theta*, *alpha*, *beta*, and *gamma* bands are adopted as features for EEG signal in [7]. In [8], the first coefficients of the discrete cosine transform are used as the representative features of the EEG signal. The Hilbert-Huang transform that decomposes the signal into basic components (intrinsic mode functions) is adopted in [9]. The auto regressive model is also presented in [10]. Chaos theory is employed in [11] to bring out the non-linear behaviour of each wavelet level of EEG signals. In [12], statistical measures such as the energy, mean energy, coefficient of variation, median absolute deviation, mean curve length, mean cross correlation, entropy, and interquartile range are used to characterize the EEG signals. A combination of image and signal processing techniques is used in [5], where the EEG signals are converted to an image and Canny edge detector is applied to extract distinctive features. These features are combined to the computed EEG zero crossing rates and to the recorded amplitude value in order to subsequently optimize the prediction process. Correlation dimensions, correlation entropy, phase correlation, dynamical similarity index and largest Lypanouv exponents are adopted in literature as nonlinear feature extraction methods in [13]. In [14] the ECG signals are inspected and processed to predict seizures, while the normalized Gabor entropy is presented in [15] and employed to multichannel intracranial EEG (iEEG). The instantaneous amplitude and frequency are extracted out of the different constituting EEG bands [16] followed by machine learning process. In [3] the focus was on the robustness of different machine learning algorithms in predicting seizures.

The robustness and accuracy of a prediction algorithm depends on an optimized performance of its main two constituents, namely, feature extraction and classification. The

978-1-5386-4727-1/18/$31.00 ©2018 IEEE

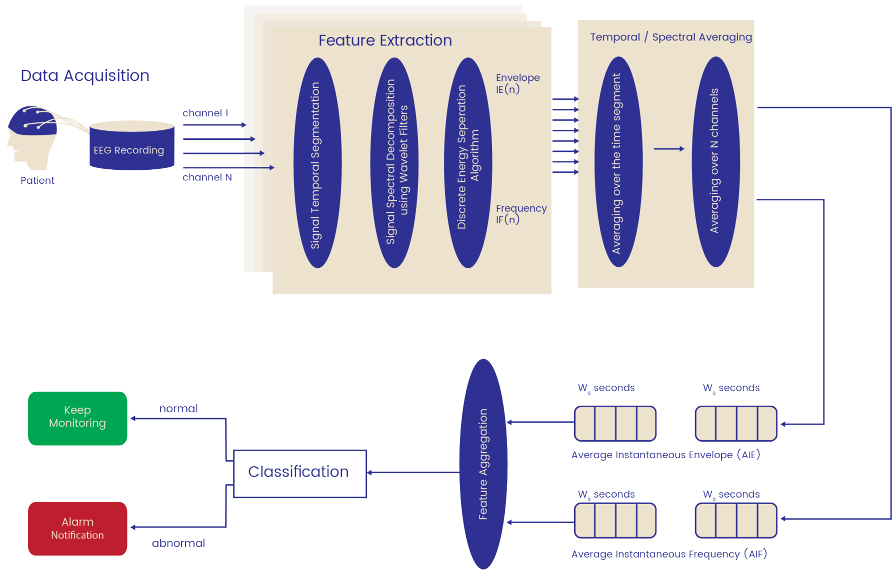

Fig. 1. Proposed epileptic seizure prediction approach.

aforementioned references considered a plethora of feature extraction techniques that are followed by a chosen classification method. While such approached may yield good results, accuracy can be notably improved if both the feature extraction and classification stages are jointly optimized and various parameters are fine-tuned based on the specificities of each patient data. This is the key novelty of the approach proposed in this work, which relies on multiresolution wavelet decomposition with instantaneous envelope and frequency estimation and spatio-temporal averaging, in addition to machine learning classifier selection among multiple known binary classification techniques.

The rest of the paper is organized as follows. Section II explains the details of the proposed approach including both feature selection and classification components. The experimental evaluation methodology is presented in Section III with extensive results on real patient datasets. Conclusions are drawn in Section IV.

II. Proposed Seizure Prediction Approach

The main objective of this paper is to properly foresee the onset of epileptic seizures some time before their occurrence in an automated manner. As mentioned earlier, the first stage in the proposed prediction approach involves extracting distinctive features from recorded EEG signals. In particular, this work performs a multiresolution analysis of the instantaneous envelope and frequency of EEG signals, and employs some special filtering to reduce the temporal and spatial fluctuations of features. Wavelet transform [17] is adopted to localize the varying characteristics of EEG signals both in time and frequency; then a discrete energy separation algorithm (DESA) [18] is applied to extract the instantaneous envelope and frequency at each resolution level. The extracted features are spatially averaged and aggregated before being introduced to a machine learning algorithm for classification. The parameters of the proposed algorithm (the mother wavelet, the number of resolution levels, window size, etc.) are optimized based on exhaustive experimental testing. For each patient, and given the extracted features, a classifier selection is performed to further improve the prediction accuracy. The details of the proposed algorithm are illustrated in Fig. 1 and discussed in further details in the sequel.

A. Feature extraction

The feature extraction stage proceeds in two steps. The first includes EEG data preprocessing where the records are organized into vectors of predetermined length. EEG signal in each channel is segmented into non-overlapping W_s epochs, where W_s is the window size in seconds. These vectors are then processed using hybrid wavelet-DESA approach.

1) Wavelet decomposition: Time localizations of spectral components of EEG signals can be achieved by doing a multiresolution analysis using wavelet transform (WT). WT is a multi-resolution technique that localizes features in *time* and *frequency* domains and it finds a wide range of applications including time-scale signal analysis, signal decomposition, and signal compression [17].

The continuous wavelet transform (CWT) can be described using the following integral:

$$X_w(\tau, s) = \frac{1}{\sqrt{|s|}} \int_{-\infty}^{\infty} x(t) * \Phi\left(\frac{t-\tau}{s}\right) dt \tag{1}$$

where $x(t)$ represents the signal to be analyzed, $\Phi(t)$ is the mother wavelet or the basis function, and $X_w(\tau, s)$ is the wavelet coefficient at a certain scale s and time shift τ. All the wavelet functions used in the transform are derived from the mother wavelet through translation (shifting) and scaling (dilation or compression). The translation parameter τ relates to the location of the wavelet function as it is shifted through the signal. Thus, it corresponds to the time information in

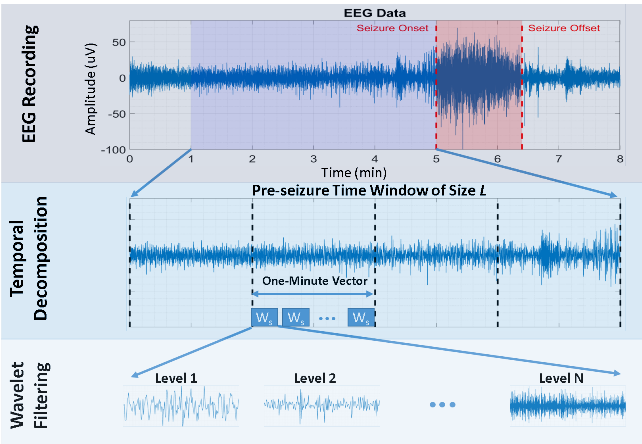

Fig. 2. Decomposition of EEG signals into N different wavelet levels.

the wavelet transform. The scale parameter s is defined as the inverse of the signal frequency $|1/\text{frequency}|$ and corresponds to frequency information. Scaling either expands or compresses the signal. Large scales (low frequencies) expand the signal and provide overall information about it, while small scales (high frequencies) compress the signal and provide detailed information hidden in it [19]. The discrete wavelet transform (DWT) is a fast computation algorithm of CWT that decomposes each segment into N_L levels and generates an output of $X_i[n]$, where i corresponds to the index of each decomposed level. These steps are illustrated in Fig. 2.

2) Discrete energy separation algorithm (DESA): The DESA algorithm proposed in [18] is adopted to obtain instantaneous envelope and instantaneous angular frequency one epoch at a time from each wavelet level. DESA is a signal processing technique that estimates the instantaneous envelope (IE) $a[n]$ and instantaneous frequency (IF) $\Omega[n]$ functions of a multicomponent amplitude and frequency modulated (AM-FM) signal. DESA has a wide range of applications in speech processing and biomedical applications [20]. It is based on a Teager Kaiser energy operator, which is a non-linear operator that estimates the $a[n]$ and $\Omega[n]$ of a signal.

The estimation of $a[n]$ and $\Omega[n]$ can be achieved by three different algorithms as described in [18]. DESA-l is the best algorithm among the three ones in terms of performance [18], and it is defined as follows.

$$Y_i[n] = X_i[n] - X_i[n-1] \tag{2}$$

where $X_i[n]$ are the outputs generated by the DWT when applied on the input EEG signal. $\Omega[n]$ and $a[n]$ can be calculated as

$$\Omega[n] = \arccos\left(1 - \frac{\Psi\left(Y_i[n]\right) + \Psi\left(Y_i[n+1]\right)}{4\Psi\left(X_i[n]\right)}\right) \tag{3}$$

$$|a[n]| = \sqrt{\frac{\Psi\left(X_i[n]\right)}{1 - \left(1 - \frac{\Psi(Y_i[n]) + \Psi(Y_i[n+1])}{4\Psi(X_i[n])}\right)^2}} \tag{4}$$

where Ψ is the energy operator and it is defined as

$$\Psi\left(X_i[n]\right) = X_i^2[n] - X_i[n-1]X_i[n+1] \tag{5}$$

A median filter is then used to smooth the ripples in the estimated $a[n]$ and $\Omega[n]$ [18].

3) Temporal and spatial averaging: Averaging of smoothed $a[n]$ and $\Omega[n]$ is implemented as follows:

i) *Temporal averaging*: At each wavelet level, $a[n]$ and $\Omega[n]$ are generated using DESA, then the temporal IE and temporal IF (TIE and TIF), are calculated to remove the fluctuations over time, where TIE and TIF represent the mean of $a[n]$ and $\Omega[n]$ respectively.

ii) *Spatial averaging*: Channel variability is compensated by averaging TIE and TIF across channels to produce the AIE and AIF vectors, where AIE and AIF represent the average of instantaneous envelope temporally and spatially, and the average of instantaneous frequency temporally and spatially, respectively.

At the end of these two steps and for each EEG epoch, the dimension of AIE/AIF feature vector is equal to the number of wavelet levels N_L.

4) Feature aggregation: Two aggregation procedures are then applied to these N_L dimensional vectors.

i) Incorporation of the AIE and AIF vectors into one vector, denoted as the AIEF vector, and which has a dimension of $2N_L$.

ii) Concatenation of N_e AIEF vectors to form one feature vector of dimension $2N_L \times N_e$, where N_e is the number of segments contained in one feature vector.

B. Classification

The extracted features offer no further insight into seizure unless they are *optimally* utilized in the context of a machine learning algorithm which classifies the incoming EEG data as preictal or interictal. The classification problem can be cast as a function approximation problem which tries to find a function that maps a d-dimensional input to appropriately encoded class information (both inputs and outputs must be encoded, unless they are already of numerical nature). Once the classification is defined as a function approximation problem, a variety of mathematical tools, such as optimization algorithms, can be used. In this work, we have considered a pool of four widely-known machine learning algorithms that are listed below.

1) Random forest (RF) : The RF classifier is a tree-based ensemble machine learning process. It is a group of thousands of decision trees that sequentially partition the feature space into different classes. Each tree is developed and then authenticated using a bootstrap sample of the categorized training data [21].

2) Logistic classifier: The logistic classifier is a classification algorithm with an accompanying supervised learning that is based on a combination of decision tree learning and logistic regression [22].

3) Support vector machines (SVM): SVM attempts to get an optimal decision boundary by maximizing the margin between the boundaries of diverse classes [17]. For this purpose, SVM recognizes those examples of each class that describe the boundary of that class in the feature space. These examples, considered to be the most instructive ones, are designated the support vectors, which are evaluated through a quadratic programming-type constrained optimization algorithm [17]. In this paper the radial basis function is adopted to be the SVM kernel.

4) Gaussian classifier: The Gaussian classifier is a machine learning algorithm that is based on the Gaussian processes models.

These four binary classification schemes are implemented using WEKA software and tested on each of the patients' data separately. The classifier yielding the highest accuracy is selected as the classifier of choice for that particular patient. For a set of N seizure records (including preictal and interictal segments), $N-1$ records are selected for training and the remaining one for testing.

C. Decision making

The generated feature vectors contain information corresponding to the interictal, preictal and ictal states. As the purpose of the seizure prediction algorithm is to distinguish between preictal and interictal activity, segments corresponding to these states should be extracted from the feature vector. Let L denote the selected preictal duration. Then, the corresponding portion of the feature vector is extracted and labeled as preictal. A similar duration of time is extracted from the interictal part as well. Thus for every seizure, we get one preictal and one interictal segments which are fed for binary classification.

The decision process considers one-minute equivalent feature vector in order to decide whether a record of unknown label does actually refer to a preictal state. In particular, the prediction metric used for decision making is the number of one-minute feature vectors that were classified as preictal and thus raised a seizure alarm. A time window of size L is considered as preictal (a seizure alarm is raised) if at least K out of the total number of one-minute vectors in the block of size L get a positive outcome. Otherwise it is classified as interictal. Consequently, the fraction of intervals, K, directly affects both *sensitivity* and *false alarm* rate. These measures, along with *accuracy*, are defined as follows [23]:

$$\text{Sensitivity} = \frac{\sum \text{TP}}{\sum \text{TP} + \sum \text{FN}} \tag{6}$$

$$\text{False alarm} = \frac{\sum \text{FP}}{\sum \text{FP} + \sum \text{TN}} \tag{7}$$

$$\text{Accuracy} = \frac{\sum \text{TP} + \sum \text{TN}}{\sum \text{TP} + \sum \text{FP} + \sum \text{FN} + \sum \text{TN}} \tag{8}$$

where $\text{TP}, \text{FN}, \text{FP}$ and TN are explained as follows:

- True positive (TP): Seizure occurs, and the algorithm detects it.
- False negative (FN): Seizure occurs but the algorithm does not detect it.
- False positive (FP): Seizure did not occur, but the algorithm announces a seizure.
- True negative (TN): The patient is in normal state, and the algorithm correctly does not declare any seizure.

Hence, as we increase K (i.e., tighter prediction constraint), sensitivity is expected to decrease as well as false alarm while a lower K (loose constraint) usually increases false alarms as more segments will be classified as preictal. To keep a balance between sensitivity and false alarm, we optimize the parameter K to maximize the overall accuracy.

III. Experimental Results

The EEG datasets used in this work were obtained from the Freiburg EEG database [24] which includes non-invasive recordings for 21 patients having a total of 87 seizures. The data were recorded using the Neurofile NT digital video EEG system with a sampling rate of 256 Hz, and an analog to digital converter of 16 bit per sample. Freiburg datasets can be categorized as ictal, preictal, postictal, and interictal periods, where the ictal signals involve epileptic seizures with at least 50 minutes of preictal signals. Also there are 24-25 hours of interictal signal and 2–5 hours of ictal with preictal and postictal signal. Thus, the data set includes about 509 hours from 21 patients.

A. Feature extraction optimization

The feature extraction process involves segmenting the input EEG signal into different windows. This is followed by the selection of L minutes of the window before the seizure onset which is considered as preictal. A similar time stretch is extracted far from the seizure and labeled as interictal. Then feature extraction using wavelet transform is applied. The purpose of our proposed approach is to optimize the performance of the feature extraction stage to obtain the best prediction performance in terms of high accuracy and low false alarm rate. To this end, the parameters related to the segmentation and feature extraction steps have been studied extensively to obtain the best combination. In particular, the parameters which were tested are the mother wavelets, number of wavelet levels N_L, the window size W_s in seconds, and the preictal time period L in minutes.

Table I illustrates the impact of the different combination of parameters on the accuracy of the prediction algorithm,

TABLE I
IMPACT OF DIFFERENT PARAMETER COMBINATIONS ON THE PREDICTION PERFORMANCE FOR PATIENT 89 OF THE FREIBURG DATABASE.

N_L	L	W_s	Accuracy	Sensitivity	False alarm
5	10	5	80%	60%	0%
5	10	20	90%	80%	0%
5	20	5	100%	100%	0%
5	20	20	90%	100%	20%
5	30	5	100%	0%	0%
5	30	20	100%	100%	0%
7	10	5	80%	60%	0%
7	10	20	70%	80%	20%
7	20	5	90%	100%	20%
7	20	20	90%	100%	20%
7	30	5	100%	100%	0%
7	30	20	100%	100%	0%
9	10	5	70%	40%	0%
9	10	20	70%	40%	0%
9	20	5	90%	100%	20%
9	20	20	80%	80%	20%
9	30	5	100%	100%	0%
9	30	20	100%	100%	0%

TABLE II
TESTING RESULTS OF WAVELET BASED SEIZURE PREDICTION APPROACH.

Patient ID	Nb of Seizures	Accuracy	Sensitivity	False alarm	N_L	$W_s(s)$	Mother Wavelet	L	Classifier	K
1	**11**	91%	100%	18%	7	5	coif1	20	SVM	9 out of 20
73	**7**	93%	100%	14%	4	5	db6	10	SVM	6 out of 10
852	**10**	90%	80%	0%	7	5	db6	30	Logistic	19 out of 30
110	**6**	92%	83%	0%	7	10	coif1	30	Gaussian	15 out of 30
586	**20**	95%	95%	5%	9	20	bior2.6	10	Logistic	6 out of 10
226	**10**	95%	100%	10%	4	10	db6	30	SVM	12 out of 30
1030	**7**	100%	100%	0%	4	5	db6	20	SVM	9 out of 20
300	**10**	90%	90%	10%	9	5	sym3	10	Logistic	5 out of 10
308	**9**	89%	78%	0%	5	5	db6	20	SVM	11 out of 20
454	**4**	100%	100%	0%	9	5	bior2.6	10	Gaussian	6 out of 10
89	**5**	90%	100%	20%	5	5	db6	10	RF	3 out of 10
327	**6**	100%	100%	0%	5	20	db6	30	Logistic	16 out of 30
467	**6**	92%	83%	0%	7	20	db6	20	Gaussian	11 out of 20
1095	**10**	100%	100%	0%	9	10	db6	10	RF	4 out of 10
325	**8**	100%	100%	0%	7	15	bior2.6	30	Logistic	12 out of 30
162	**6**	100%	100%	0%	7	5	db6	10	Logistic	6 out of 10
1119	**6**	100%	100%	0%	9	5	db6	25	Gaussian	13 out of 25
219	**6**	92%	100%	17%	5	5	bior2.6	10	Gaussian	2 out of 10
Average		95%	95%	5%						

the sensitivity, and false alarm rates for patient 89 in the Freiburg database. As shown in Table I, prediction accuracy is highly affected by the variation of N_L, L, and W_s. For instance, the accuracy increases from 80% to 100% by the variation of $N_L = 5$, $L = 10$, $W_s = 5$, to $N_L = 5$, $L = 20$, $W_s = 5$. Moreover, it can be clearly noticed that different combinations yield an accuracy and sensitivity of 100% and a false alarm rate of 0%. This highlights the complexity of the seizure prediction problem as there is no single way to process the EEG signals to achieve the best performance. Among others, delimiting the preictal period remains problematic. From Table I, different combinations with a preictal length of 20 and 30 minutes yielded a 100% accuracy. Hence, knowing the length of the preictal period remains an open question. The optimization performed in this work rather bounds the preictal length to a smaller interval. Based on the combinations tested, the preictal period of patient 89 is longer than 20 minutes and possibly between 20 and 30 minutes.

B. Optimized feature extraction and classifier selection results

Table II demonstrates the results of optimizing the feature extraction process, classifier selection, and decision making (through the parameter K) for 18 patients in the Freiburg database. Two sets of EEG segments are selected for each patient, one set corresponds to preictal state and the other to interictal state. The purpose of the classifier is then to categorize incoming EEG traces into one of these two classes.

A patient-specific binary classification is conducted continuously to classify the input feature vectors into interictal or preictal groups. Once pre-seizure observations are found to last for a certain period, alarms are raised. Simulation results demonstrated efficient prediction performance with average sensitivity exceeding 95% and false alarm rate below 5%.

Table II extends the optimization described in Table I through the addition of the classifier and decision making parameter. The table shows that for each of the patients, the highest prediction accuracy is achieved by a given set of feature extraction parameters, one classifier and a decision metric K which ranges from 20% in the case of patient 219 (2 minutes out of L =10 minutes) to 63.3% in the case of patient 852 (19 minutes out of L =30 minutes). Practically, this translates into the observation that for some patients, distinctive markers should be depicted for a large amount of time before being able to accurately predict a seizure while for others, only some markers are enough to predict the forthcoming seizure with high confidence. The classifier selection process reflects the separation among the EEG features for different patients. This indicates that each patient has unique brain activity during each state of the epileptic activity. Thus to distinguish a preictal from an interictal state, it is essential to select the most appropriate classifier that would be able to categorize the features in the best way.

C. Comparison with another algorithm

In order to further assess the performance of the proposed approach, the results in Table II are compared to the prediction algorithm described in [16]. In [16], the EEG signals are filtered with bandpass filters at the $delta$, $alpha$, $beta$ and $gamma$ bands, then DESA is adopted to extract the instantaneous envelope and frequency. Feature selection is then applied before classification using SVMs. Feature selection is a method in machine learning used to prevent the model from becoming more complex and reduce the chance of overfitting, by selecting a subset of the features. Then, the accuracy of the prediction approach is presented with and without the feature selection algorithm. Although the results in [16] show that feature selection has an impact on increasing the prediction accuracy, no such technique will be applied for the proposed algorithm to demonstrate that a 100% accuracy can be reached relying solely on parameter optimization as described before. The results of the comparison between the proposed approach and an implementation of the prediction algorithm in [16] are summarized in Table III.

TABLE III
COMPARISON BETWEEN THE PROPOSED APPROACH AND THE EXISTING WORK IN [16].

	Algorithm based on [16]			Proposed algorithm		
Pat_ID	Acc	Sens	FA	Acc	Sens	FA
1	86%	73%	0%	91%	91%	9%
110	75%	83%	33%	92%	83%	0%
300	80%	80%	20%	90%	90%	10%
308	67%	56%	22%	89%	78%	0%
327	58%	83%	67%	100%	100%	0%
1030	50%	86%	86%	100%	100%	0%

The results show that the proposed approach achieves better performance than the filter-based technique. The multiresolution analysis adopted in this paper is able to capture more distinctive features than a technique based on [16]. This translates into an increased sensitivity and lower false alarm rates.

IV. CONCLUSIONS

In this work, a patient-aware seizure prediction approach was proposed. It combines multiresolution analysis, which involves extracting the instantaneous envelope and frequency out of wavelet levels of EEG signals then applying temporal and spatial averaging, to a classifier selection stage. The most appropriate classifier for each of the patients is selected to increase accuracy. The merits of the algorithm were demonstrated through extensive testing on EEG records from the Freiburg dataset. A high sensitivity level of 95% and low false alarm rate of 5% were achieved through an exhaustive search on the different combinations of parameters.

ACKNOWLEDGMENT

The work presented in this paper was fully supported by NeuroPro AG. The statements made herein are solely the responsibility of the authors.

REFERENCES

[1] A. Casson, S. Smith, J. Duncan, and E. Rodriguez-Villegas, "Wearable EEG: what is it, why is it needed and what does it entail?", *IEEE International Conference on Engineering in Medicine and Biology Society (EMBS 2008)*, pp. 5867-–5870, Aug. 2008.

[2] S. Miri and A. Nasrabadi, "A new seizure prediction method based on return map", *18th Iranian Conference of Biomedical Engineering (ICBME)*, pp. 244-–248, Dec. 2011.

[3] S. Kumar, L. Nigam, D. Karnam, S. K. Murthy, P. Fedorovych, and V. Kalidindi, "Machine learning for seizure prediction: A revamped approach", *International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, pp. 1159—1164, Aug. 2015.

[4] Y. Park, T. Netoff, and K. Parhi, "Reducing the number of features for seizure prediction of spectral power in intracranial EEG", *Asilomar Conference on Signals, Systems and Computers*, pp. 770—774, Nov. 2012.

[5] V. Zeljkovic, V. Valev, C. Tameze, and M. Bojic, "Preictal phase detection algorithm based on one dimensional EEG signals and two dimensional formed images analysis", *International Conference on High Performance Computing and Simulation (HPCS)*, pp. 607—614, July 2013.

[6] N. Wang and M. Lyu, "Extracting and selecting distinctive EEG features for efficient epileptic seizure prediction", *IEEE Journal of Biomedical and Health Informatics*, vol. 19, no. 5, pp. 1648—1659, Sep. 2015.

[7] A. Tzallas, M. Tsipouras, and D. Fotiadis, "Epileptic seizure detection in EEGs using time-frequency analysis", *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 5, pp. 703-–710, Sep. 2009.

[8] 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, Oct. 2011.

[9] S. Ghosh-Dastidar, H. Adeli, and N. Dadmehr, "Mixed-band wavelet-chaos-neural network methodology for epilepsy and epileptic seizure detection", *IEEE Transactions on Biomedical Engineering*, vol. 54, no. 9, pp. 1545—1551, Sep. 2007.

[10] E. Kollialil, G. Gopan, A. Harsha, and L. Joseph, "Single feature-based non-convulsive epileptic seizure detection using multi-class SVM", *International Conference on Emerging Trends in Communication, Control, Signal Processing Computing Applications (C2SPCA)*, pp. 1–6, Oct. 2013.

[11] D. Birvinskas, V. Jusas, I. Martisius, and R. Damasevicius, "EEG dataset reduction and feature extraction using discrete cosine transform", *UKSim/AMSS European Symposium on Computer Modeling and Simulation (EMS)*, pp. 199-–204, Nov. 2012.

[12] F. Duman, N. Ozdemir, and E. Yildirim, "Patient specific seizure prediction algorithm using Hilbert-Huang transform", *IEEE-EMBS International Conference on Biomedical and Health Informatics (BHI)*, pp. 705—708, Jan. 2012.

[13] M. Parvez and M. Paul, "Epileptic seizure prediction by exploiting spatiotemporal relationship of EEG signals using phase correlation", *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 24, no. 1, pp. 158—168, Jan. 2016.

[14] S. Iqbal, Y. U. Khan, and O. Farooq, "Nonlinear analysis of EEG for seizure prediction", *IEEE India Conference (INDICON)*, pp. 1—5, Dec. 2015.

[15] B. C. Biswas and S. V. Bhalerao, "A real time based wireless wearable EEG device for epilepsy seizure control", *International Conference on Communications and Signal Processing (ICCSP)*, pp. 149–153, Apr. 2015.

[16] R. Liu, I. Vlachos, B. K. Karumuri, J. A. Adkinson, and L. Iasemidis, "Normalized Gabor entropy analysis of iEEG for prediction of epileptic seizures", *Southern Biomedical Engineering Conference (SBEC)*, pp. 15—16, Mar. 2016.

[17] R. Polikar, "The wavelet tutorial", Rowan University, 2001.

[18] P. Maragos, J. F. Kaiser, and T. F. Quatieri, "Energy separation in signal modulations with application to speech analysis", *IEEE Transactions on Signal Processing*, vol. 41, no. 10, pp. 3024—3051, Oct. 1993.

[19] M. Azim, M. Amin, S. Haque, M. Ambia, and M. Shoeb, "Feature extraction of human sleep EEG signals using wavelet transform and Fourier transform", *International Conference on Signal Processing Systems (ICSPS)*, pp. 701–705, July 2010.

[20] P. S. Rathore and R. B. Pachori, "Instantaneous fundamental frequency estimation of speech signals using DESA in low-frequency region", *International Conference on Signal Processing and Communication (ICSC)*, pp. 470—473, Dec. 2013.

[21] T. K. Ho, "Random decision forests", *International Conference on Document Analysis and Recognition*, pp. 278—282, Aug. 1995.

[22] D. Dancey, Z. A. Bandar, and D. McLean, "Logistic model tree extraction from artificial neural networks", *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 37, no. 4, pp. 794—802, Aug. 2007.

[23] 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", *IEEE International Conference on Engineering in Medicine and Biology Society, EMBC*, pp. 7564-–7569, Aug. 2011.

[24] "EEG dataset: epilepsy center of the university hospital of Freiburg", http://epilepsy.uni-freiburg.de/freiburg-seizure-prediction-project/eeg-database, 2014.