Wavelet Domain Approximate Entropy-Based Epileptic Seizure Detection

A.S. Muthanantha Murugavel#1, S. Ramakrishnan*2

Department of Information Technology

Dr.Mahalingam College of Engineering and Technology

Pollachi - 642003, Tamilnadu, India

[1]as_ananth2k1@yahoo.com

[2]ram_f77@yahoo.com

***Abstract*— The electroencephalogram (EEG) signal plays an important role in the detection of epilepsy. The EEG recordings of the ambulatory recording systems generate very lengthy data and the detection of the epileptic activity requires a time-consuming analysis of the entire length of the EEG data by an expert. The aim of this work is to develop a new method for automatic detection of EEG patterns using wavelet based approximate entropy (ApEn) and probabilistic neural network (PNN). Our method consists of EEG data collection, feature extraction and classification stages. ApEn is a statistical parameter that measures the predictability of the current amplitude values of a physiological signal based on its previous amplitude values. In feature extraction stage we use best basis mother wavelet functions and wavelet thresholding technique. For the feature selection we have used a new methodology, that is minimal variance within class and maximal absolute difference between classes are used for feature selection. In classification stage we implement PNN to detect epileptic seizure detection. It is known that the value of the ApEn drops sharply during an epileptic seizure and this fact is used in the proposed system and overall accuracies as high as 100% can be achieved by using the proposed system..**

***Keywords*— approximate entropy (ApEn), wavelet transform, artificial neural network (ANN), electroencephalogram (EEG), EEG classification, epilepsy, seizure detection, probabilistic neural network (PNN).**

I. INTRODUCTION

Epilepsy is a chronic disorder characterized by recurrent seizures which may vary from muscle jerks to several convolutions. Estimated 1% of world population suffers from epilepsy [1], while 85% of them live in the developing countries. Epileptic detection is done from electroencephalogram (EEG) signal as epilepsy is a condition related to the brain's electrical activity. Electroencephalogram (EEG) is routinely used clinically to diagnose, monitor and localize epileptogenic zone. Occurrence of recurrent seizures in the EEG signal is characteristics of epilepsy. In majority of the cases, the onset of the seizures cannot be predicted in a short period, a continuous recording of the EEG is required to detect epilepsy. The entire length of the EEG recordings is analyzed by expert to detect the traces of epilepsy. The traditional methods of analysis are tedious and time-consuming and so many automated epileptic EEG detection systems have been developed [2]. This paper discusses an automated epileptic EEG detection system using probabilistic neural network (PNN) using a time-frequency domain feature of the EEG signal called approximate entropy (ApEn). EEG data is first digitized. The digital EEG data is fed as an input to an automated seizure detection system in order to detect the seizures present in the EEG data. Approximate Entropy drops abruptly due to the synchronous discharge of large groups of neurons during an epileptic activity. Hence, it is a good feature to make use of in the automated detection of epilepsy.

Entropy is a thermodynamic quantity describing the amount of disorder in the system. From an information theory perspective, the above concept of entropy is generalized as the amount of information stored in a more general probability distribution. First Shannon applied the concept of information or logical entropy to the science of information theory and data communications. Recently, a number of different entropy estimators [2] have been applied to quantify the complexity of the signal. Entropy estimators are broadly classified into two categories spectral entropies and embedding entropies. The spectral entropies use the amplitude components of the power spectrum of the signal as the probabilities in entropy calculations. It quantifies the spectral complexity of the time series. The embedding entropies use the time series directly to estimate the entropy. Kolmogorov—Sinai entropy and the approximate entropy are the embedding entropies discussed here [3].

The discrete wavelet transform is a versatile signal processing tool that has many engineering and scientific applications. DWT employs two sets of functions called scaling functions and wavelet functions, which are associated with low-pass and high-pass filters, respectively. The decomposition of the signal into the different frequency bands is simply obtained by successive high-pass and low pass filtering of the time domain signal. Subasi [4] deals with a novel method of analysis of EEG signals using discrete wavelet transform, and classification using ANN. The pseudo Wigner-Ville and the smoothed pseudo Wigner-Ville distribution [5] was used for extracting features from the time-frequency plane. PNN is predominantly a classifier since it can map any input pattern to a number of classifications. Among the main advantages that discriminate PNN is: Fast training process, an inherently parallel structure, guaranteed to converge to an optimal classifier as the size of the epresentative training set increases and training samples can be added or removed without extensive retraining. Accordingly, a PNN learns more quickly than many neural networks model, which led to its success on variety of applications. Based on these facts and advantages, PNN can be viewed as a supervised neural network that is capable of using it in system classification and pattern recognition [6].

The detection of epilepsy, which includes visual scanning of EEG recordings for the spikes and seizures, is very time consuming, especially in the case of long recordings. In addition, bio-signals are highly subjective so disagreement on the same record is possible, so the EEG signal parameters extracted and analyzed using computers, are highly useful in diagnostics. Automatic analysis of EEG recordings in the diagnosis of epilepsy started in the early 1970s and lot of seizure detection algorithms have been developed. In this paper, approximate entropy-based epileptic EEG detection proposed by the author Vairavan Srinivasan et al [11], is used with some modified approach such as wavelet domain. Since wavelet has several advantages, it is both time and frequency based and it can simultaneously possess compact support, orthogonality, symmetry, and short support, and high order approximation.

Therefore, main objective of this paper is to propose a novel feature extraction technique for the detection of epilepsy. The wavelet transformation is used for extracting Approximate Entropy and a new methodology is presented for feature selection. The methodology is applied to two different groups of EEG signals for analysis of EEGs and EEG sub bands for detection of epileptic seizure: 1) healthy subjects; 2) epileptic subjects during a seizure (ictal EEG). Each EEG is decomposed into two constituent EEG sub bands: delta, theta, alpha, beta, and gamma using wavelet-based filters. The features such as Approximate Entropy of the wavelet coefficients are used to represent the time frequency distribution of the EEG signals in each sub-band of the wavelet transformation and the probabilistic neural network is used to detect epileptic EEG signals.

II. PROPOSED METHODOLOGY

As in traditional pattern recognition systems, the epileptic seizure detection consists of main modules such as a feature extractor that generates a wavelet based feature from the EEG signals, feature selection that composes composite features, and a feature classifier (PNN) that outputs the class based on the composite features. The data flow of the proposed approach is illustrated in Fig. 1.

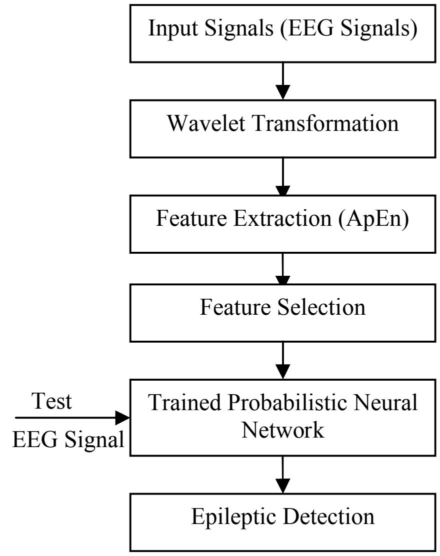

Fig. 1 Data flow diagram of the proposed system

A. *Dataset Description*

The data used in this research are a subset of the EEG data for both healthy and epileptic subjects made available online by Dr. Ralph Andrzejak of the Epilepsy Centre at the University of Bonn, Germany (http://www.meb.unibonn.de/epileptologie/science/physik/eeg data.html) [1]. EEGs from two different groups: group H (healthy subjects) and group S (epileptic subjects during seizure) are analyzed. The type of epilepsy was diagnosed as temporal lobe epilepsy with the epileptogenic focus being the hippocampal formation. Each group contains 100 single channel EEG segments of 23.6 sec duration each sampled at 173.61 Hz. As such, each data segment contains N=4097 data points collected at intervals of 1/173.61th of 1s. Each EEG segment is considered as a separate EEG signal resulting in a total of 200 EEG signals or EEGs. As an example, the first 6s of two EEGs (signal numbers in parentheses) for groups H (H029) and S (S001) are magnified and displayed in Fig. 2.

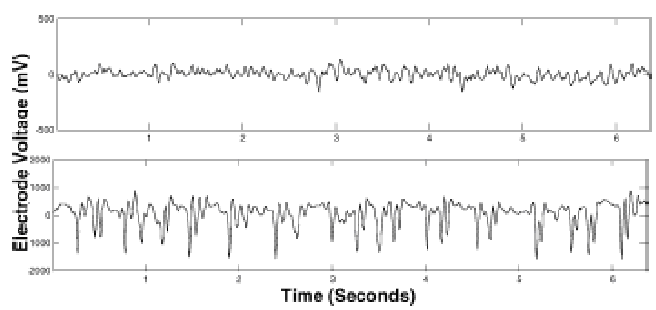

Fig. 2 Sample unfiltered EEGs (0–6 s) for (from top to bottom) Group H (H029) and Group S (S001)

B. Wavelet Transformation

Wavelet transform is a spectral estimation technique in which any general function can be expressed as an infinite series of wavelets. The basic idea underlying wavelet analysis consists of expressing a signal as a linear combination of a particular set of functions (wavelet transform, WT), obtained by shifting and dilating one single function called a mother wavelet. The decomposition of the signal leads to a set of coefficients called wavelet coefficients. Therefore the signal can be reconstructed as a linear combination of the wavelet functions weighted by the wavelet coefficients. The key feature of wavelets is the time-frequency localization. It means that most of the energy of the wavelet is restricted to a finite time interval.

The wavelet technique applied to the EEG signal will reveal features related to the transient nature of the signal, which is not made obvious by the Fourier transform. Adeli et al. [7] gave an overview of the discrete wavelet transform (DWT) developed for recognizing and quantifying spikes, sharp waves and spike-waves. In general, it must be said that no time-frequency regions but rather time-scale regions are defined. All wavelet transforms can be specified in terms of a low-pass filter, which satisfies the standard quadrature mirror filter condition. One area in which the wavelet transformation has been particularly successful is the epileptic seizure detection because it captures transient features and localizes them in both time and frequency content accurately. The wavelet transformation analyses the signal at different frequency bands, with different resolutions by decomposing the signal into a coarse approximation and detail information [8]. The decomposition of the signal into the different frequency bands is merely obtained by consecutive high-pass and low-pass filtering of the time domain signal. The procedure of multi-resolution decomposition of a signal x[n] is schematically shown in Fig. 3. Each stage of this scheme consists of two digital filters and two down-samplers by 2. The first filter, h[n] is the discrete mother wavelet, high pass in nature, and the second, g[n] is its mirror version, low-pass in nature. The down-sampled outputs of first high-pass and low-pass filters provide the detail, D1 and the approximation, A1, respectively. The first approximation, A1 is further decomposed and this process is continued as shown in Fig. 3. The EEG sub bands of a2, d2 and d1are shown in fig. 4.

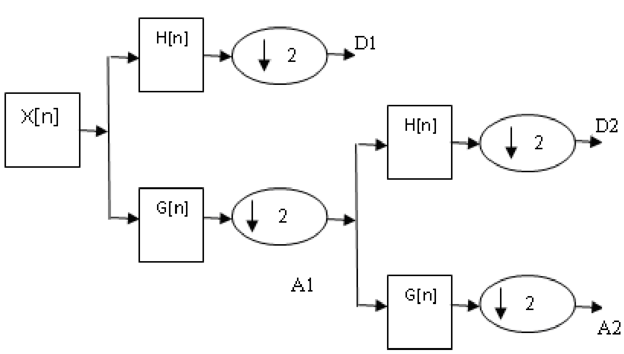

Fig. 3 Two level wavelet decomposition

Selection of suitable wavelet and the number of decomposition levels is very important in analysis of signals using the wavelet transformation. The number of decomposition levels is chosen based on the dominant frequency components of the signal. In the present study, since the EEG signals do not have any useful frequency components above 30 Hz, the number of decomposition levels was chosen to be 2. Thus, the EEG signals were decomposed into details D1–D2 and one final approximation, A2. Usually, tests are performed with different types of wavelets and the one, which gives maximum efficiency, is selected for the particular application. The smoothing feature of the Daubechies wavelet of order 4 (db4) made it more appropriate to detect changes of EEG signals. Hence, the wavelet coefficients were computed using the db4 in the present study. The proposed method was applied on both data set of EEG data (Sets H and S). In the discrete wavelet analysis, a signal can be represented by its approximations and details. The detail at level j is defined as

$$D_j = \sum_{k \in Z} a_{j,k}\, \psi_{j,k}(t) \tag{1}$$

and the approximation at level J is defined as

$$A_j = \sum_{j > J} D_j \tag{2}$$

It becomes obvious that

$$A_{j-1} = A_j + D_j \tag{3}$$

and $$f(t) = A_j + \sum_{j \le J} D_j \tag{4}$$

Wavelet has several advantages, which can simultaneously possess compact support, orthogonality, symmetry, and short support, and high order approximation. We experimentally found that time-frequency domain feature provides superior performance over time domain feature in the detection of epileptic EEG signals.

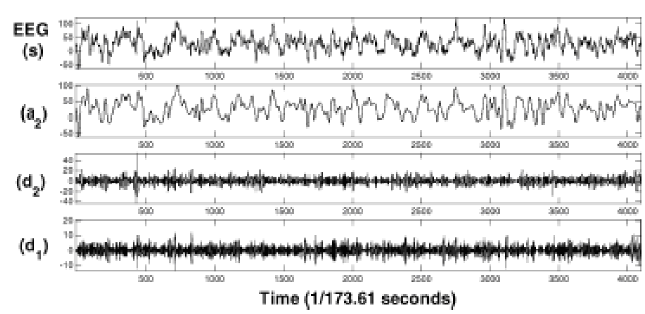

Fig. 4 Level 2 decomposition of the band-limited EEG into three EEG sub bands using fourth-order Daubechies wavelet (s = a2+d2+d1)

C. Feature Extraction

The proposed system makes use of a single feature called ApEn for the epileptic detection. The ApEn is a wavelet-domain feature that is capable of classifying complex systems. The value of the ApEn is determined as shown in the following steps [9], [10].

1) Let the data sequence containing N data points be X = [x(1), x(2), x(3), . . . , x(N)].

2) Let x(i) be a subsequence of X such that x(i) = [x(i), x(i + 1), x(i + 2), . . . , x(i + m − 1)] for $1 \leq i \leq N - m$, where m represents the number of samples used for the prediction.

3) Let r represent the noise filter level that is defined as

$$r = k \times SD \tag{5}$$

for k = 0, 0.1, 0.2, 0.3,…, 0.9 where SDis the standard deviation of the data sequence X.

4) Let {x(j)} represent a set of subsequences obtained from x(j) by varying j from 1 to N. Each sequence x(j) in the set of {x(j)} is compared with x(i) and, in this process, two parameters, namely $C_i^m(r)$ and $C_i^{m+1}(r)$ are defined as follows:

$$C_i^m(r) = \frac{\sum_{j=1}^{N-m} k}{N-m} \tag{6}$$

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

$$\text{and } C_i^{m+1}(r) = \frac{\sum_{j=1}^{N-m} k}{N-m} \tag{7}$$

with conditions depicted by (A) as shown at the bottom of the page.

5) We define $\Phi_m(r)$ and $\Phi_{m+1}(r)$ as follows:

$$\Phi_m(r) = \frac{\sum_{i=1}^{N-m} \ln(C_i^m(r))}{N-m} \tag{8}$$

$$\Phi_{m+1}(r) = \frac{\sum_{i=1}^{N-m} \ln(C_i^{m+1}(r))}{N-m} \tag{9}$$

Small values of ApEn imply strong regularity in a data sequence and large values imply substantial fluctuations [11]. In the proposed approach, ApEn is calculated for one approximation and for detailed information such as a2 and d2.

D. Feature Selection

As discussed in the above section, 30 ApEn features have been obtained from each sub band leading to a total of 60 ApEn features. As it consumes more time in processing these 60 ApEn features, there is a need to select the best thirty features. Table. 1 shows the extracted features (ApEn) for the sub bands for the sample set A. These best features are selected by our novel approach which involves choosing the feature having minimal variance within the class and maximum absolute difference between the classes. Variance has been calculated for each class of sample set to find the minimal variance. And absolute difference between classes of sample set to find the maximal difference.

TABLE I
FEATURE EXTRACTION SAMPLE DATA – SET A

Set	Sub-bands	ApEn
H	D1	-12513000
	D2	295
	A2	-101890
S	D1	-4391700000
	D2	47289
	A2	-16616000

E. Probabilistic Neural Network Classifier

The classification of EEG signals into healthy and epileptic signals is done using the probabilistic neural network. The architecture of the PNN is shown in Fig. 5. In machine learning, a classifier is essentially a mapping from the feature space to the class space. An Artificial Neural Network (ANN) implements such a mapping by using a group of interconnected artificial neurons simulating the human brain. An ANN can be trained to achieve expected classification results against the input and output information stream, so there may not be a need to provide a specified classification algorithm. There is no need to train the network over the entire data set again, so we use PNN to enable quick updates of our network as more patients' data becomes available. Our PNN has three layers: the Input Layer, the Radial Basis Layer which evaluates distances between the input vector and rows in the weight matrix, and the Competitive Layer which determines the classification with maximum probability of correctness. Dimensions of matrices are marked under their names.

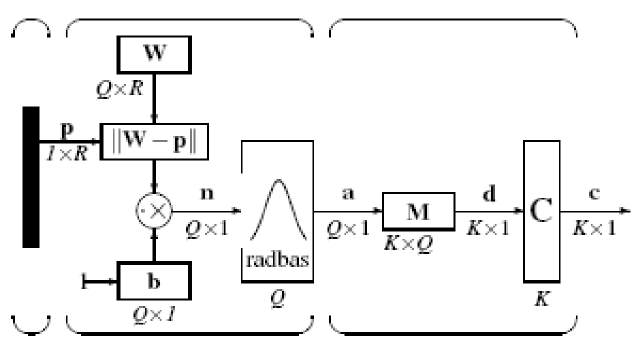

Fig. 5 PNN structure, R: number of features, Q: number of training samples, K: number of classes.

1) *Input Layer:* The input vector, denoted as p, is presented as a black vertical bar in Fig. 5. The input layer unit does not perform any computation and simply distributes the input to neurons in the pattern layer. On receiving a pattern x from input layer, the neuron x_{ij} of the pattern layer computes its output using the below formula.

$$\phi_{ij}(\mathbf{x}) = \frac{1}{(2\pi)^{0.5d}\sigma^d}\exp\left(-\frac{(\mathbf{x}-\mathbf{x}_{ij})^T(\mathbf{x}-\mathbf{x}_{ij})}{2\sigma^2}\right) \quad (10)$$

Where d denotes dimension of the pattern vector x, is the smoothing parameter and xij is the neuron vector.

2) *Radial Basis Layer:* In the Radial Basis Layer, the vector distances between input vector p and the weight vector, made up of each row of the weight matrix W are calculated. Here, the vector distance is defined as the dot product between two vectors. The dot product between p and the i-th row of W produces the i-th element of the distance vector matrix, denoted as ||W− p||. The bias vector b is then combined with ||W−p|| by an element-by-element multiplication, represented as "•×" in Fig. 5. The result is denoted as n = ||W−p||•×b. The transfer function in PNN has built into a distance criterion with respect to a center. In this paper, we define it as radbas(n) = e−n2 . Each element of n is substituted into the transfer function and produces corresponding element of a, the output vector of Radial Basis Layer. We can represent the i-th element of a as ai = radbas(||Wi−p||•×bi), where Wi is the i-th row of W, and bi is the i-th element of bias vector b.

Radial Basis Layer Weights: Each row of W is the feature vector of one training sample. The number of rows is equal to the number of training samples.

Radial Basis Layer Biases: All biases in the radial basis layer are set to √ln0.5/s, resulting in radial basis functions that cross 0.5 at weighted inputs of ±s, where s is the spread constant of PNN. According to our experience, s = 0.1 can typically result in the highest accuracy. Summation layer neurons compute the maximum likelihood of a pattern x being classified into Ci, by averaging the output of all neurons that belong to the same class using

$$P_i(\mathbf{x}) = \frac{1}{N_i(2\pi)^{0.5d}\sigma^d}\sum_{j=1}^{N_i}\exp\left(-\frac{(\mathbf{x}-\mathbf{x}_{ij})^T(\mathbf{x}-\mathbf{x}_{ij})}{2\sigma^2}\right)$$

(11)

Where N_i denotes the total number of samples in class C_i.

3) *Competitive Layer:* There is no bias in the Competitive Layer. In this layer, the vector a is first multiplied by the layer weight matrix M, producing an output vector d. The competitive function C produces a 1 corresponding to the largest element of d, and 0's elsewhere. The index of the 1 is the class of the EEG segment. M is set to a K×Q matrix of Q target class vectors. If the i-th sample in the training set is of class j, then we have a 1 on the j-th row of the i-th column of M. The decision layer classifies the pattern x in accordance with Bayes decision rule based on the output of all summation layer neurons using

$$\hat{C}(\mathbf{x}) = \arg\max\{P_i(\mathbf{x})\}, i = 1, 2, .., m \quad (12)$$

Where denotes the estimated class of pattern x, and m is the total number of classes in training samples.

Hence, PNN employed in this work possesses 30 nodes in the input layer and 2 nodes in the output layer (the number of nodes in the output layer is the number of classifications of EEG signals). The performance of the neural model was evaluated in terms of training performance and classification accuracies and the results confirmed that the proposed scheme has potential in classifying the EEG signals.

III. RESULTS AND DISCUSSION

ApEn values are computed for selected combinations of m, r, and N. The values of m, r, and N that are used for the experiments are as follows:) m = 1, 2, 3; r = 0%–90% of SD of the data sequence in increments of 10%; and N = 4097. ApEn values are computed for both normal and epileptic EEG signals and are fed as inputs to the two neural networks. Among the available 100 EEG data sets, 50 data sets are used for training and the remaining data sets are used for testing the performance of the neural networks. The potentiality of the ApEn to discriminate the two signals, namely, normal and epileptic EEG signals depends on the values of m, r, and N. Fig. 6 shows Receiver Operating Characteristics Curve for the overall detection accuracy (%) obtained by the PNN using ApEn as the input feature. The experimental results show that our PNN using wavelet based ApEn can well preserve the most discriminant information of EEG signals and improve the performance over the exiting system in terms of detection rate. PNN gives good overall accuracy values in the range 98% - 100%, only for a few combinations of m, r, and N (e.g., m = 1, r = 0× SD for all the values of N and m = 3, r = 0.2 × SD, N = 4097). Though the use of ANNs increases the computational complexity, the high overall detection accuracies are achieved with this system surpasses its disadvantage as in any automated seizure detection system, the detection of the seizure with high accuracy is of primary importance.

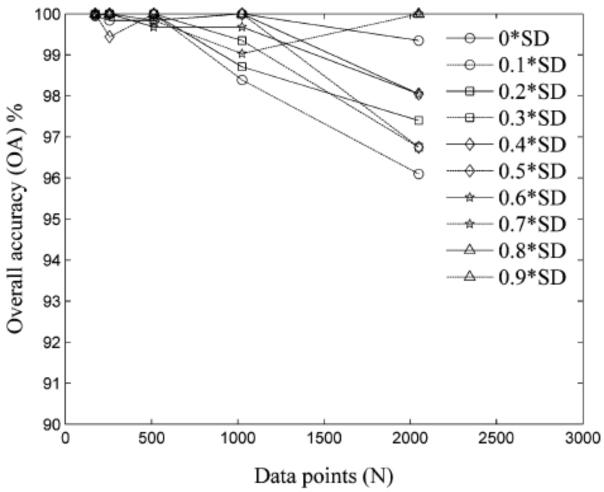

Fig. 6. ROC Curve: Overall Classification Rate

Our experimental results are based on data sets corresponding to five different subjects only. The optimum ApEn parameter values obtained based on this data may not hold good for a general case. Hence, using a linear separator with known ApEn parameter values may not give good results in situations where a large number of different subjects are involved. This problem will not arise in the proposed PNN-based method as it has performed well irrespective of the ApEn parameter values used. It is shown that our wavelet based ApEn possesses good characteristics such as robustness in the characterization of the epileptic patterns and low computational burden. Hence, an automated system using wavelet based ApEn as the input feature is best suited for the real time detection of the epileptic seizures. The proposed system is based on two types of EEG, namely, EEG signals of awake and epileptic subjects. It can be made more robust by acclimatizing it to the other manifestations of EEG like sleep EEG.

IV. Conclusion

In this paper, the neural network namely Probabilistic Neural Network (PNN), has been employed for the automated detection of epilepsy. A robust and computationally low-intensive feature such as wavelet domain based Approximate Entropy (ApEn) has been used for the proposed epileptic detection system and a new approach has been used for feature selection in order to reduce the dimension and increase the computation speed. Experimental results show that overall accuracies as high as 100% can be achieved by this system. As the proposed system is based on a single feature that has a low computational burden, it is best suited for the real-time detection of epileptic seizures from ambulatory recordings.

Acknowledgment

The authors wish to thank Andrzejak et al., 2001 for the benchmark eeg dataset which is publicly available: (http://www.meb.unibonn.de/epileptologie/science/physik/eegdata.html).

References

[1] Guler, I., Kiymik, M. K., Akin, M., & Alkan, A. AR spectral analysis of EEG signals by using maximum likelihood estimation. Computers in Biology and Medicine, 31, 441–450, 2001.

[2] Zoubir, M., & Boashash, B. Seizure detection of newborn EEG using a model approach. IEEE Transactions on Biomedical Engineering, 45, 673–685, 1998.

[3] Adeli H, Zhou Z, Dadmehr N. Analysis of EEG records in an epileptic patient using wavelet transform. J Neurosci Methods; 123(1):69–87, 2003.

[4] Pradhan, N., Sadasivan, P. K., & Arunodaya, G. R. Detection of seizure activity in EEG by an artificial neural network: A preliminary study. Computers and Biomedical Research, 29, 303–313, 1996.

[5] Weng, W., & Khorasani, K. An adaptive structure neural network with application to EEG automatic seizure detection. Neural Networks, 9, 1223–1240, 1996.

[6] Petrosian, A., Prokhorov, D., Homan, R., Dashei, R., & Wunsch, D. Recurrent neural network based prediction of epileptic seizures in intraand extracranial EEG. Neurocomputing, 30, 201–218, 2000.

[7] Folkers, A., Mosch, F., Malina, T., & Hofmann, U. G. Realtime bioelectrical data acquisition and processing from 128 channels utilizing the wavelet-transformation. Neurocomputing, 52–54, 247–254, 2003.

[8] Guler I, Ubeyli ED. Application of adaptive neuro-fuzzy inference system for detection of electrocardiographic changes in patients with partial epilepsy using feature extraction. Expert Syst Appl; 27(3):323–30, 2004.

[9] Subasi, A. Automatic recognition of alertness level from EEG by using neural network and wavelet coefficients. Expert Systems with Applications, 28, 701–711, 2005.

[10] Kandaswamy, A., Kumar, C. S., Ramanathan, R. P., Jayaraman, S., & Malmurugan, N. Neural classification of lung sounds using wavelet coefficients. Computers in Biology and Medicine, 34(6), 523–537, 2004.

[11] M. Boulougoura, E. Wadge, V.S. Kodogiannis, H.S. Chowdrey, "Intelligent systems for computer-assisted clinical endoscopic image analysis", 2nd IASTED Int. Conf. on Biomedical Engineering, Innsbruck, Austria, pp. 405-408, 2004